



Data Protection in Digital Tax Administration: A Bibliometric Review of Research Trends, Influential Studies, and Emerging Themes

Regain H. Hutahaean*, Rizky Pratama Putra Karo-Karo

Universitas Dirgantara Marsekal Suryadarma, Indonesia

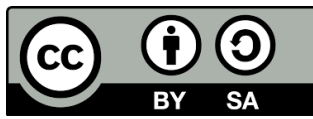
Email: regain.h@gmail.com*, rizkykarokaro@unsurya.ac.id

Keywords:

Data Protection; Digital Tax Administration; Taxpayer Privacy; Cybersecurity; Bibliometric Analysis

Abstract

Digital transformation has reshaped tax administration into an integrated, automated, and data-driven system, making data protection essential to its legitimacy. Electronic tax systems process sensitive taxpayer information; however, knowledge concerning data protection in digital tax administration remains fragmented, creating a need for a systematic mapping of the literature. This study aims to map the development, influential publications, and thematic structure of research on data protection in digital tax administration. A systematic review combined with bibliometric analysis was conducted using 69 Scopus-indexed journal articles published between 1998 and 2026. The dataset was analyzed using Bibliometrix in RStudio through annual scientific production, local citation analysis, and thematic mapping. The findings show that publication activity remained limited during the early period but increased substantially after 2019, reaching its highest levels in 2024 and 2025. Local citation analysis identifies Scarcella (2019) as the most influential publication within the corpus. The thematic map reveals three major motor-theme clusters: GDPR, transparency, and automated decision-making; fraud, cybersecurity, and identity theft; and tax administration, tax compliance, and tax evasion. Data privacy remains a basic theme, while data protection and information exchange occupy a central transitional position. The novelty of this study lies in integrating the technological, legal, and administrative dimensions of taxpayer data protection within a unified bibliometric landscape, thereby demonstrating how privacy, cybersecurity, and automated governance are increasingly shaping research on digital tax administration.



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright © 2026 by Author. Published by International Journal Labs

INTRODUCTION

Digital transformation has reshaped tax administration from a transaction-based process into an integrated, automated, and increasingly data-driven system. Digitalization enables tax authorities to consolidate information, accelerate public services, strengthen oversight, and expand their capacity to process large volumes of data (Luo & Xu, 2024; Darono & Panggabean, 2026). Luo and Xu (2024) show that big data tax administration can improve transparency and governance, although it may also create new implications for organizations subject to tax supervision. In addition, Darono and Panggabean (2026) emphasize that successful technology adoption in tax administration depends not only on formal compliance but also on institutional readiness and effective implementation. Therefore, digital tax administration should be understood not merely as a technological upgrade but as a broader governance transformation that changes the relationship among tax authorities, data, and taxpayers.

This transformation makes data protection an integral component of the quality of digital tax administration. Electronic tax systems process identity information, financial records, transaction histories, and compliance data that may cause substantial harm when accessed, disclosed, or used without authorization. Setyowati et al. (2020) demonstrate that the use of new technologies in taxation must continue to balance system transparency with taxpayer privacy. From a governance perspective, digital maturity and information technology audit frameworks must also be supported by cybersecurity readiness and effective risk controls so that digital transformation can genuinely strengthen fiscal integrity (Umbet et al., 2025). Accordingly, data protection should not be treated as a secondary technical function but as a core condition for building legitimate and trustworthy digital tax administration.

The issue becomes more complex as tax authorities increasingly use artificial intelligence, risk analytics, and large-scale data processing for monitoring and decision-making. Angelov (2026) shows that the adoption of artificial intelligence (AI) by tax administrations across the European Union varies considerably and raises concerns related to transparency, accountability, confidentiality, and data protection. In the Indonesian context, the digitalization of the Core Tax Administration System can improve efficiency, transparency, and tax supervision, but it also creates risks related to data security, privacy, and regulatory compliance (Djinarto et al., 2024). Moreover, Aisy et al. (2025) highlight the importance of an independent data protection authority with sufficient supervisory and enforcement capacity. These developments indicate that the greater the digital capacity of a tax authority, the stronger the need for accountability mechanisms capable of protecting taxpayer rights (Aisy et al., 2025).

This need is closely related to the tension between taxpayer confidentiality and the growing demand for transparency in modern tax administration. Tax authorities require broad access to data to detect noncompliance and improve enforcement, while taxpayers retain a legitimate interest in confidentiality and proportionate data use. In the context of information exchange and reporting obligations, confidentiality protections may directly conflict with disclosure requirements imposed on intermediaries and taxpayers, making clear legal boundaries essential (Tofan & Arseni, 2023). Digitalization also expands the number of actors, systems, and channels involved in the collection and exchange of tax-related information. Therefore, the central challenge is not to choose between transparency and confidentiality but to establish a governance framework in which both principles can operate proportionately and accountably.

This challenge is especially relevant in Indonesia. In September 2024, allegations emerged that approximately six million taxpayer identification number records had been offered for sale online, including national identification numbers, addresses, telephone numbers, and email addresses. However, the Directorate General of Taxes later stated that its examination of access logs did not identify evidence that the alleged breach originated directly from its information systems. The incident nevertheless illustrates the reputational and trust-related risks that arise when taxpayer information is perceived to be insecure. Its relevance increased after Coretax began operating in January 2025 because the system integrates highly sensitive information, including taxpayer identities, income data, business transactions, and tax filing and payment histories. International guidance on the digitalization of revenue authorities

also places data protection and privacy among the key conditions for maintaining taxpayer trust. Thus, the protection of tax data is not merely a technical cybersecurity issue but a broader matter of public governance and institutional legitimacy.

Despite its growing importance, knowledge of data protection in digital tax administration remains fragmented across different research streams. Some studies focus primarily on technology and system security, while others examine information technology audits, artificial intelligence, confidentiality, taxpayer rights, or institutional arrangements for data protection. This diversity can be seen in studies on blockchain-based VAT systems, privacy risks in Coretax, risk governance in tax information technology audits, and AI integration in tax administration (Setyowati et al., 2020; Djinarto et al., 2024; Umbet et al., 2025; Angelov, 2026). As a result, it remains difficult to understand the development of the field, identify its most influential publications, and determine how themes such as cybersecurity, data privacy, confidentiality, and automated decision-making are positioned within a coherent body of knowledge. This fragmentation creates a need for a systematic mapping of the literature on data protection in digital tax administration.

Previous studies have addressed several important aspects of this issue, but they remain focused on specific technical or legal dimensions. Niu et al. (2022) show that taxpayer privacy can be strengthened through a blockchain-based e-tax protocol that balances anonymity with regulatory monitoring. In a different context, Mikiver and Siitam (2024) demonstrate that large-scale data processing by public authorities, including tax authorities, raises concerns related to purpose limitation and the predictability of data-processing rules under the GDPR. Maharani et al. (2025) further show that Indonesia continues to face institutional and compensation-related challenges in personal data protection compared with the European Union. Although these studies provide important insights, they do not map how data protection in digital tax administration has developed as a broader research field. A clear gap therefore remains in understanding the longitudinal development of publications, the most influential studies, and the major and emerging themes within this domain. To address this gap, this study aims to map the development of research, identify the most influential publications, and examine the major and emerging themes in data protection within digital tax administration through a systematic review and Scopus-based bibliometric analysis using Bibliometrix. The findings of this study are expected to provide theoretical benefits by contributing to the systematic mapping of a fragmented research field and identifying the conceptual structure of data protection in digital tax administration, as well as practical benefits for tax authorities and policymakers by supporting the integration of data protection, cybersecurity, transparency, and taxpayer rights into the design of digital tax systems. In addition, this study can serve as a reference for future researchers interested in examining data protection in digital tax administration using bibliometric or comparative approaches across different jurisdictions.

METHOD

Research Design

This study employed a systematic literature review combined with bibliometric analysis to examine the development and thematic structure of research on data protection in digital tax administration. The systematic review was used to identify and delimit the relevant literature through a transparent and reproducible procedure, while bibliometric analysis was used to map

publication trends, citation influence, and the conceptual structure of the field. The review process followed the transparency principles of PRISMA 2020 in documenting the database, search strategy, eligibility criteria, and final corpus included in the analysis (Page et al., 2021).

The study addressed three research questions. RQ1 examined the temporal development of research on data protection in digital tax administration; RQ2 identified publications with the strongest citation relationships within the analyzed corpus; and RQ3 mapped the major and emerging themes characterizing the field. Accordingly, the analytical design used selected bibliometric techniques aligned with each research question to produce focused and interpretable results.

Data Source and Search Strategy

Bibliographic records were retrieved exclusively from Scopus. Scopus was selected because it provides standardized bibliographic metadata, including document titles, abstracts, authors, keywords, publication years, citation information, affiliations, and cited references, which are required for bibliometric analysis. Nevertheless, database selection can influence the composition of a bibliometric corpus because indexing coverage varies across disciplines, countries, and publication languages; consequently, the results of this study represent the Scopus-indexed research landscape rather than the entire universe of publications on the topic (Mongeon & Paul-Hus, 2016).

The search strategy was designed around two conceptual blocks. The first represented data protection and information security, including terms related to personal data protection, privacy, cybersecurity, confidentiality, and data breaches. The second represented tax administration and taxpayer information, covering tax authorities, tax information systems, electronic taxation, and e-filing. The use of multiple synonymous terms was intended to improve retrieval sensitivity without expanding the corpus to taxation studies unrelated to information and data protection. No publication-year restriction was imposed, allowing the historical development of the field to emerge from the retrieved literature itself. Based on the Scopus export, the resulting publications span 1998–2026.

The final Scopus search string was:

(TITLE-ABS-KEY ("data protection" OR "personal data protection" OR "data privacy" OR "information privacy" OR "data security" OR "information security" OR "data breach" OR cybersecurity OR "cyber security" OR confidentiality) AND TITLE-ABS-KEY ("tax administration" OR "tax authority" OR "tax authorities" OR "taxpayer data" OR "taxpayer information" OR "tax information" OR "tax information system" OR "tax administration system" OR "e-tax" OR "electronic tax" OR "tax e-filing")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (SRCTYPE , "j")) AND (LIMIT-TO (PUBSTAGE , "final"))

The database search was conducted on 3 September 2026. The Scopus filters restricted the corpus to journal articles, English-language publications, journal sources, and articles at the final publication stage. These restrictions were applied to improve comparability of bibliographic metadata and to maintain a consistent peer-reviewed publication corpus. The final search yielded 69 records, which constituted the bibliometric dataset used in the analyses presented in this study.

Eligibility, Data Extraction, and Preparation

Studies were considered eligible when their titles, abstracts, or keywords connected issues of data protection, privacy, confidentiality, information security, cybersecurity, or data breaches with tax administration, tax authorities, taxpayer information, electronic taxation, or related digital tax systems. Publications outside the journal-article format, non-English documents, non-final publications, and non-journal sources were excluded directly through the Scopus database filters. This approach ensured that the corpus remained sufficiently broad to capture different traditions of research, ranging from taxpayer confidentiality and tax-information exchange to contemporary issues involving cybersecurity, artificial intelligence, automated decision-making, and digital tax platforms. The retrieved database therefore encompasses both earlier confidentiality-oriented studies and newer forms of data-protection research associated with increasingly data-intensive tax administration.

All bibliographic records were exported from Scopus in CSV format with the metadata required for Bibliometrix analysis. Data preparation included consistency checks for document titles, publication years, authors, source information, citation fields, and keywords. Keyword standardization was performed conservatively so that obvious spelling or formatting variations could be harmonized without merging conceptually different terms. Such preprocessing is important because inconsistencies in bibliographic metadata can fragment concepts and affect the structure of bibliometric networks and thematic classifications (Aria & Cuccurullo, 2017; Zupic & Čater, 2015).

Bibliometric Analysis

Bibliometric analysis was conducted using the Bibliometrix package in RStudio through the Biblioshiny interface. Bibliometrix is an open-source R package developed specifically for comprehensive science-mapping and bibliometric analysis and supports descriptive performance indicators, citation analysis, and conceptual mapping (Aria & Cuccurullo, 2017).

To ensure that each research question was addressed through a focused and non-overlapping analytical procedure, the bibliometric analysis was structured by assigning one principal technique to each research question. This approach was adopted to maintain analytical clarity and to avoid excessive visual outputs that could dilute the substantive interpretation of the findings. The alignment between the research questions, bibliometric techniques, and their analytical purposes is summarized in Table 1.

Table 1. Alignment of Research Questions, Bibliometric Analyses, and Analytical Purposes

Research question	Bibliometric analysis	Analytical purpose
RQ1. How has research on data protection in digital tax administration evolved over time?	Annual Scientific Production	To identify the historical trajectory and annual development of publications
RQ2. Which publications have been most influential in the development of research on data protection in digital tax administration?	Most Local Cited Documents	To identify documents most frequently cited by other publications within the analyzed corpus
RQ3. What are the major and emerging themes in research on data protection in digital tax administration?	Thematic Map	To identify and classify the conceptual themes characterizing the research field
Source: Developed by the authors based on the research design of this study		

As shown in Table 1, each research question was linked to a specific bibliometric output that directly reflected its analytical objective. Annual Scientific Production was used to capture the temporal growth of the field, Most Local Cited Documents was applied to identify publications with stronger internal citation influence, and the Thematic Map was used to examine the conceptual position of major and less-developed themes. This one-to-one alignment between research questions and analytical techniques enabled the study to maintain a focused interpretation of the bibliometric results while minimizing methodological redundancy.

Interpretation and Reproducibility

The bibliometric outputs were interpreted together with the titles and abstracts of the publications contained in the final corpus. This complementary reading was necessary because bibliometric relationships indicate patterns of scientific association but do not by themselves explain the substantive meaning of a theme or establish causal relationships among research concepts (Zupic & Čater, 2015; Donthu et al., 2021). Consequently, representative articles were examined to support the interpretation of publication trends, locally influential documents, and thematic clusters identified through Bibliometrix.

Methodological reproducibility was supported by explicitly reporting the database, retrieval date, complete search syntax, document restrictions, software environment, bibliometric techniques, and analytical corpus. This procedure enables subsequent researchers to replicate the retrieval strategy or update the analysis as additional studies on data protection and digital tax administration become available.

RESULTS AND DISCUSSION

Evolution of Scientific Production

The temporal distribution of publications indicates that research on data protection in digital tax administration has developed unevenly across the 1998–2026 period. Of the 69 documents analyzed, publication activity remained limited during the earlier years but increased substantially from 2019 onward. The annual publication pattern is presented in Figure 1.

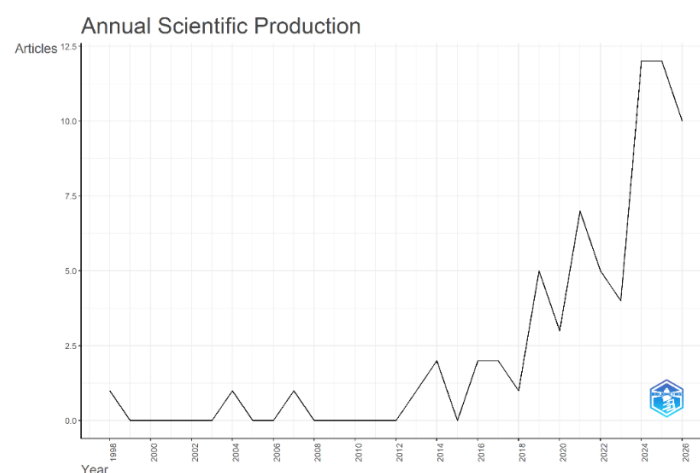


Figure 1. Annual Scientific Production on Data Protection in Digital Tax Administration, 1998–2026

Source: Scopus database, processed using Bibliometrix in RStudio (2026)

As shown in Figure 1, only sporadic publications appeared between 1998 and 2018, generally ranging from zero to two articles per year. Publication activity increased to five articles in 2019, followed by three in 2020, seven in 2021, five in 2022, and four in 2023. A substantial increase occurred in 2024 and 2025, with 12 articles published in each year, representing the highest annual scientific production in the dataset. In 2026, the database records 10 articles, indicating that recent years account for a substantial share of the literature analyzed.

Most Influential Documents Based on Local Citations

The local citation analysis identifies the publications that have been cited by other documents within the bibliometric dataset, thereby indicating their relative influence within the specific research corpus on data protection in digital tax administration. The results show that only a small number of publications received local citations, with Scarcella (2019) recording the highest value, followed by Peeters (2024) and Pellegrino (2024), as presented in Figure 2.

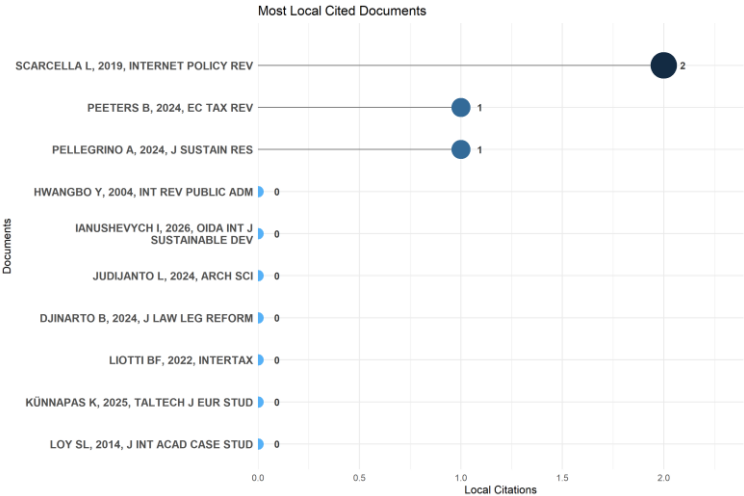


Figure 2. Most Locally Cited Documents in Research on Data Protection in Digital Tax Administration

Source: Scopus database, processed using Bibliometrix in RStudio (2026)

As shown in Figure 2, Scarcella (2019) is the most locally cited document with two local citations, making it the most frequently referenced publication within the analyzed corpus. Peeters (2024) and Pellegrino (2024) each received one local citation, while the remaining documents displayed in the figure recorded no local citations. Overall, the distribution indicates that internal citation linkages within the dataset are concentrated in only a few publications, while most documents have not yet been cited by other studies included in the same corpus.

Major and Emerging Research Themes

The thematic map reveals that research on data protection in digital tax administration is distributed across motor, basic, niche, and emerging or declining themes according to their centrality and density. Several well-developed and highly relevant themes are positioned in the motor-theme quadrant, while other topics remain either foundational, specialized, or less developed within the research field. The overall thematic structure is presented in Figure 3.



Figure 3. Thematic Map of Research on Data Protection in Digital Tax Administration
Source: Scopus database, processed using Bibliometrix in RStudio (2026).

As shown in Figure 3, the motor themes consist of three main clusters: GDPR, transparency, and automated decision-making; fraud, cybersecurity, and identity theft; and tax administration, tax compliance, and tax evasion. Meanwhile, taxation and data privacy are positioned as basic themes, whereas Malaysia and tax secrecy constitute a niche theme. The cluster comprising data protection, exchange of information, and privacy is located close to the intersection of the four quadrants, indicating an intermediate position in terms of centrality and development. Finally, automatic exchange of information and legal professional privilege are located within the emerging or declining themes quadrant.

Development of Research on Data Protection in Digital Tax Administration

The publication trajectory indicates that research on data protection in digital tax administration initially developed as a relatively fragmented field, with early studies concentrating more on the technical security of electronic tax systems than on data protection as an independent legal or governance issue. For example, Julià-Barceló and Vinje (1998) discussed confidentiality and security in electronic communications, including electronic tax declarations, through digital signatures and encryption, while Hwangbo (2004) proposed the Global Electronic Tax Invoice System by emphasizing reliability, transparency, encryption, and information security in electronic taxation. A similar technology-oriented perspective appeared in Lee and Chang (2007), who examined information security management using a tax administration institution as an organizational case. Thus, the low and sporadic scientific production before the mid-2010s appears consistent with a period in which taxpayer data protection was primarily embedded within broader discussions of electronic-system security rather than treated as a distinct research domain.

The field subsequently began to broaden from technical confidentiality toward taxpayer rights, privacy, and the legal implications of increasingly data-intensive tax administration. The 2014 study by Loy, Brown, and Tabibzadeh illustrated the practical consequences of inadequate information-security governance through the South Carolina Department of Revenue data breach, while Alm et al. (2017) examined taxpayer confidentiality in relation to compliance behavior. By 2019, the literature had become more explicitly connected to digital

administration and privacy regulation: Hodžić (2018) identified data-security problems as one of the challenges of tax-administration digitalization, whereas Scarcella (2019) examined the implications of profiling and automated decision-making for taxpayers' privacy rights under the GDPR. This development continued during 2020–2023, when studies increasingly addressed automatic tax-information exchange, mass data processing, cybersecurity, and legal safeguards, as reflected in Ronco (2020), Rojszczak (2021), Liotti (2022), Huiskers-Stoop et al. (2022), Keunen (2023), Tomo (2023), and Hatfield (2023). Collectively, these studies suggest that the rise in publication activity from 2019 onward coincided with a conceptual expansion from information security toward the broader governance of taxpayer data, including privacy, proportionality, legal safeguards, and institutional responsibility.

The sharp increase in publications during 2024–2026 further indicates that data protection has become increasingly intertwined with the latest phase of digital tax transformation, particularly the use of integrated tax platforms, artificial intelligence, automated risk assessment, and large-scale data processing. In 2024, Judijanto (2024) directly examined taxpayer privacy protection in Indonesia's Core Tax Administration System, while Djinarto et al. (2024) highlighted privacy and data-security risks associated with CTAS digitalization. In the same year, Peeters (2024) examined legal restrictions on the use of artificial intelligence by tax authorities, whereas Mohr (2024) focused on the legal basis and safeguards surrounding automated risk-management systems in tax administration. This trajectory became even more visible in 2025 through studies on AI applications by tax authorities, GDPR-based taxpayer safeguards, cybersecurity and IT-audit governance, digital tax transformation, and automatic exchange of information, as reflected in Künnapas et al. (2025), Rizzo and Hassan (2025), Umbet et al. (2025), Pali and Rama (2025), and Obando and Porporatto (2025). Publications in 2026 continued this direction by linking digital tax transformation with AI governance, cybersecurity, confidentiality, and proportional risk management, as demonstrated by Ianushevych and Hubanov (2026), Rizzo and Hassan (2026), Nurbekova et al. (2026), and Tomo (2026). Accordingly, the publication peak in 2024–2025 can be interpreted as evidence of growing scholarly attention to the tension between the increasing data-processing capacity of tax authorities and the need to protect taxpayer information and rights. The lower number recorded for 2026 should, however, be interpreted cautiously because the bibliometric database was compiled before the end of the publication year.

Influential Documents within the Research Corpus

The local citation pattern indicates that only a limited number of publications have become internal reference points within research on data protection in digital tax administration. Scarcella (2019) occupies the strongest position, with two local citations, and its substantive focus helps explain its relevance within the corpus. The study examines the use of profiling and automated decision-making by tax authorities and highlights the need to protect taxpayers' privacy rights under the GDPR, particularly when digital technologies are used to identify taxpayers considered at high risk of non-compliance. Consequently, the article connects three issues that later became increasingly prominent in the literature, namely digital tax administration, data-driven enforcement, and privacy safeguards.

The two documents receiving one local citation each further illustrate the shift of the field toward newer digital technologies and their regulatory implications. Peeters (2024) discusses the growing use of artificial intelligence by EU tax authorities and examines the safeguards

provided by the GDPR and emerging AI regulation, including limitations on automated decision-making and the importance of human intervention. Meanwhile, Pellegrino and Stasi (2024) analyze blockchain adoption in value-added tax systems and identify data protection, cybersecurity, regulatory differences, and institutional adaptation as important challenges in the digital transformation of tax administration. Therefore, although these studies examine different technologies, both extend the earlier privacy-oriented discussion toward broader questions concerning the governance and regulation of increasingly automated tax systems.

At the same time, the predominance of zero local citations among the other documents suggests that the internal citation structure of this research field remains relatively dispersed. Several recent studies published in 2024 and 2025 already address highly relevant issues, such as taxpayer privacy in Indonesia's Core Tax Administration System, data security risks associated with CTAS digitalization, and AI applications within tax authorities, yet these publications have not accumulated local citations within the analyzed corpus. Judijanto (2024), for example, emphasizes the need for clear regulation, strong security protocols, transparency, and accountability in protecting taxpayer information within CTAS, while Djinarto et al. (2024) identify data security and privacy concerns as challenges accompanying digital tax administration. Similarly, Künnapas et al. (2025) demonstrate the increasing application of AI within tax authorities. Thus, the low local citation counts should be read primarily as evidence that the literature is still developing and that many of its recent contributions have not yet become established internal reference points within the field.

Major and Emerging Research Themes

The thematic map shows that the current development of research on data protection in digital tax administration is increasingly driven by the interaction between data protection regulation, transparency, and automated decision-making. This pattern is reflected in the motor-theme cluster containing GDPR, transparency, and automated decision-making, which is consistent with Scarcella's (2019) discussion of profiling and automated decision-making by tax authorities and the need for safeguards under the GDPR. More recent studies reinforce this direction, as Mohr (2024) examines the legal basis for automated risk-management systems in tax administration, while Peeters (2024) and Rizzo and Hassan (2025) focus on the legal constraints surrounding AI use by tax authorities, particularly transparency, human intervention, and taxpayer protection. Therefore, the position of this cluster as a motor theme indicates that the governance of automated tax administration has become one of the most developed and central areas within the field.

A second motor-theme cluster, comprising fraud, cybersecurity, and identity theft, reflects the growing attention paid to security risks accompanying the digitalization of tax systems. Earlier evidence from Loy, Brown, and Tabibzadeh (2014) demonstrates the institutional consequences of weak information-security governance through the South Carolina Department of Revenue data breach, while Hultgren et al. (2022) document identity-theft tax refund fraud and its implications for cybersecurity and taxpayer trust. Similarly, Ngugi et al. (2022) identify vulnerabilities within electronic tax-filing processes that may increase exposure to tax identity theft, whereas Hatfield (2023) emphasizes the need to strengthen safeguards for taxpayer data in response to persistent cybersecurity weaknesses. In this context, the thematic map suggests that cybersecurity is no longer treated merely as a technical issue,

but has become closely connected with fraud prevention, identity protection, and the broader protection of taxpayer information.

The third motor-theme cluster, consisting of tax administration, tax compliance, and tax evasion, indicates that data protection research remains closely connected to the operational objectives of digital tax systems. Hodžić (2018) identifies data-security problems as one of the challenges accompanying the digitalization of tax administration, while more recent studies examine how digital tools are used to improve compliance, efficiency, and enforcement. For instance, Zakaria et al. (2024) analyze the adoption of tax digitalization, Pali and Rama (2025) discuss the role of digital transformation in improving tax services and compliance, and Weiqun and Xing (2025) examine the broader effects of tax-administration digitalization on organizational transformation. Consequently, the central position of this cluster indicates that research on data protection develops alongside the broader institutional agenda of strengthening compliance and reducing tax evasion through digital technologies.

Meanwhile, the position of taxation and data privacy as basic themes, together with the central cluster of data protection, exchange of information, and privacy, shows that privacy remains a foundational issue connecting different strands of the literature. Schaper (2016) highlights the tension between tax information exchange and data-protection rights in the European Union, while Ronco (2020) similarly examines the compatibility of tax information exchange with data-protection safeguards. This discussion was subsequently extended by Liotti (2022) and Huiskers-Stoop et al. (2022), who examine whether international and regional frameworks provide adequate protection for taxpayers' data in increasingly automated exchange-of-information environments. Accordingly, the map suggests that privacy and data protection function as cross-cutting concepts that connect tax transparency, information exchange, and digital tax administration, even though their thematic development is less concentrated than that of the motor themes.

Finally, the niche and emerging-or-declining quadrants illustrate themes that are more specialized or less integrated into the wider research field. The Malaysia and tax secrecy cluster reflects a relatively specific line of inquiry, as seen in research on confidentiality within Malaysian tax administration and comparative tax-secrecy provisions, including Yussof et al. (2025). By contrast, automatic exchange of information and legal professional privilege occupy the emerging-or-declining quadrant. Studies such as Gennadi & Stanislav (2021) and Tomo (2023) examine legal and data-protection issues arising from automatic exchange and third-party reporting, whereas Tofan and Arseni (2023) and Rao (2026) focus on conflicts between disclosure obligations and professional confidentiality. Because a thematic map alone cannot determine whether these themes are genuinely emerging or declining, their position should be interpreted more cautiously as indicating relatively low centrality and limited thematic development within the current dataset.

CONCLUSION

This study mapped the development, influential publications, and thematic structure of research on data protection in digital tax administration. Scientific production remained limited and sporadic during the early period but increased markedly from 2019 onward, reaching its highest levels in 2024 and 2025. This growth reflected the increasing relevance of data protection as tax administrations adopted more intensive digitalization, automated

processing, and data-driven enforcement. The internal citation structure of the field remained relatively dispersed, with only a few publications serving as key reference points. The thematic map identified three main motor-theme clusters: GDPR, transparency, and automated decision-making; fraud, cybersecurity, and identity theft; and tax administration, tax compliance, and tax evasion. Meanwhile, taxation and data privacy functioned as basic themes, tax secrecy appeared as a more specialized niche theme, and automatic exchange of information and legal professional privilege remained less central and less developed within the current corpus.

From a practical perspective, these findings suggested that digital tax transformation should not be treated solely as an administrative efficiency agenda. Tax authorities and policymakers should integrate data protection, cybersecurity, transparency, and taxpayer rights into the design and operation of digital tax systems, particularly where artificial intelligence, automated risk assessment, large-scale data processing, and cross-border information exchange are involved. The prominence of GDPR, automated decision-making, cybersecurity, and identity theft also highlighted the importance of strengthening governance mechanisms such as privacy by design, clear accountability arrangements, human oversight of automated systems, robust data-security safeguards, and accessible mechanisms for protecting taxpayers affected by digital risks. However, this study was subject to several limitations because the analysis relied exclusively on Scopus-indexed, English-language, final journal articles, potentially excluding relevant studies published in other languages or disseminated through other databases, books, legal commentaries, or policy reports. The relatively small corpus and low number of local citations also indicated that this remained a developing and fragmented research field. In addition, the 2026 publication count should be interpreted cautiously because the year was incomplete when the database was compiled. The thematic map was also influenced by the keywords and metadata contained in the retrieved documents; therefore, the position of a theme did not, by itself, establish whether it was definitively emerging or declining.

Future research should therefore expand both the empirical and conceptual boundaries of the field. Comparative studies are needed to examine how different tax administrations operationalize data-protection obligations, particularly in relation to AI-based profiling, automated decision-making, cybersecurity incidents, taxpayer data breaches, and cross-border information exchange. Further research could also examine institutional accountability, independent data-protection supervision, breach-notification mechanisms, remedies and compensation for affected taxpayers, and the balance between tax transparency and confidentiality. Empirical research in developing and emerging economies would be especially valuable because existing discussions remain strongly influenced by European data-protection frameworks and may not fully capture the institutional, technological, and regulatory conditions of other jurisdictions.

FUNDING

This study did not receive any specific funding from governmental institutions, private organizations, or other funding agencies.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest related to the conduct of the research, preparation, or publication of this article.

AUTHOR CONTRIBUTIONS

All authors contributed to the preparation of this article. Regain H. Hutahaeen and Rizky Pratama Putra Karo-Karo, was responsible for the conceptualization of the study, data collection and analysis, interpretation of the results, and manuscript preparation. The other authors contributed to the review, editing, and final approval of the manuscript.

GENERATIVE AI STATEMENT

The authors declare that no generative artificial intelligence was used in the preparation of this article.

REFERENCES

- Aisy, F. R., Maskur, M. A., & Amiruddin, A. M. (2025). Establishing Indonesia's Personal Data Protection Agency: Comparative administration sanctions enforcement from Ireland, Australia, and Singapore. *JILS*, 10, 431.
- Alm, J., Bernasconi, M., Laury, S., Lee, D. J., & Wallace, S. (2017). Culture, compliance, and confidentiality: Taxpayer behavior in the United States and Italy. *Journal of Economic Behavior & Organization*, 140, 176–196. <https://doi.org/10.1016/j.jebo.2017.05.018>
- Angelov, A. (2026). Artificial intelligence in European Union tax administrations: A comparative assessment. *Journal of Risk and Financial Management*, 19(4), 295.
- Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959–975.
- Darono, A., & Panggabean, T. (2026). From compliance to adoption: A theory-building study of technology implementation gaps in tax administration. *Journal of Risk and Financial Management*, 19(4), 237.
- Djinarto, B., Suhartono, S., Hadi, S., Setyadji, S., & Nickalus, J. (2024). Reforming tax law enforcement: The role of Core Tax Administration System digitalization and the *ultimum remedium* principle. *Journal of Law and Legal Reform*, 5(4).
- 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.
- Gennadi, T., & Stanislav, A. (2021). The roots of legal problems arising in the course of automatic exchange of information in tax matters between the EU and Russia. *Russian Law Journal*, 9(4), 99–127.
- Hatfield, M. (2023). Safeguarding taxpayer data. *Florida Tax Review*, 27(1), 212–268.
- Hodžić, S. (2018). Tax administrative challenges of the digital economy: The Croatian experience. *eJournal of Tax Research*, 16(3), 762–779.
- Huiskers-Stoop, E., Breuer, A., & Nieuweboer, M. (2022). Exchange of information, tax confidentiality, privacy and data protection from an EU perspective. *Erasmus Law Review*, 15, 86.
- Hultgren, A., Hasseldine, J., & Nash, J. (2022). Identity theft tax refund fraud in the United States. *eJournal of Tax Research*, 19(2), 213–231.
- Hwangbo, Y. (2004). Establishing a trusted third party for taxing global electronic commerce: System architecture of global electronic tax invoice (GETI). *International Review of Public Administration*, 9(1), 33–40.
- Ianushevych, I., & Hubanov, O. (2026). Digital transformation of Ukraine's tax system in line with European Union implementation standards. *OIDA International Journal of*

Sustainable Development, 19(06), 79–90.

- Judijanto, L. (2024). Concerns over the protection of taxpayers' privacy data on the Core Tax Administration System. *Archives Des Sciences*, 74(6), 80–85.
- Julià-Barceló, R., & Vinje, T. (1998). Towards a European framework for digital signatures and encryption: The European Commission takes a step forward for confidential and secure electronic communications. *Computer Law & Security Report*, 14(2), 79–86.
- Keunen, L. (2023). Big data tax audits: The conceptualisation of fishing expeditions. *International Review of Law, Computers & Technology*, 37(2), 166–197.
- Künnapas, K., Maslionkina, P., Hinno, R., & Liberts, R. (2025). Current AI applications by Estonian tax authorities and use case scenarios. *TalTech Journal of European Studies*, 15(1), 179–208.
- Lee, M. C., & Chang, T. (2007). Applying ISO 17799: 2005 in information security management. *International Journal of Services and Standards*, 3(3), 352–373.
- Liotti, B. F. (2022). Taxpayers' data protection: Do international, regional, and domestic instruments guarantee adequate rights in tax (A) EoI? *Intertax*, 50(2).
- Loy, S. L., Brown, S., & Tabibzadeh, K. (2014). South Carolina Department of Revenue: Mother of government dysfunction. *Journal of the International Academy for Case Studies*, 20(1), 83.
- Luo, J., & Xu, J. (2024). The impact of big data tax administration on corporate ESG—A quasi-natural experiment based on Golden Tax Project III. *China Journal of Accounting Research*, 17(3), 100378.
- Maharani, D. P., Kusumadara, A., Widhiyanti, H. N., & Dewantara, R. (2025). Administrative legal liability for data breaches: Securing the right to compensation in Indonesia and the EU. *Law. Human. Environment*, 3(16), 173–190.
- Mikiver, M., & Siitam, N. (2024). Data-driven public administration and the GDPR: Seeing the Court of Justice's judgement in Case C-175/20 in a broader context. *Juridica International*, 33, 52.
- Mohr, A. (2024). A statutory basis for the use of automated risk management systems in tax administration. *British Tax Review*, 2024(2), 283–303.
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213–228.
- Ngugi, B. K., Hung, K. T., & Li, Y. J. (2022). Reducing tax identity theft by identifying vulnerability points in the electronic tax filing process. *Information & Computer Security*, 30(2), 173–189.
- Niu, H., Li, T., & Gong, X. (2022). A blockchain-based certifiable anonymous E-taxing protocol. *PLOS ONE*, 17(7), e0270454.
- Nurbekova, G. T., Greggi, M., & Tussupova, L. K. (2026). Maintaining confidentiality in the exchange of information on tax matters in the Republic of Kazakhstan. *Laws*, 15(3), 41.
- Obando, R. R., & Porporatto, P. (2025). Taxpayer rights in the context of the automatic exchange of information (AEOI): Insights from recent Swiss case law. *European Taxation*, 65(11), 471–478.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372.
- Pali, A., & Rama, S. (2025). Digital transformation and its impact on public services: The case of the tax administration services. *Interdisciplinary Journal of Research and Development*, 12(1 S1), 40–50.
- Peeters, B. (2024). European law restrictions on tax authorities' use of artificial intelligence systems: Reflections on some recent developments. *EC Tax Review*, 33(2), 54–57.

- Pellegrino, A., & Stasi, A. (2024). Critical factors in adopting blockchain technology in value-added tax systems. *Journal of Sustainability Research*, 6(2).
- Rao, S. (2026). Client-attorney privilege: The last barrier to tax transparency? *World Tax Journal*, 18(1).
- Rizzo, A., & Hassan, G. (2025). Addressing the use of AI by EU tax authorities: Towards a common framework of taxpayer protection. *European Taxation*, 65, 15–29.
- Rizzo, A., & Hassan, G. (2026). Proportionate risk management for tax-related AI systems: Analysis and implications under EU and US law. *World Tax Journal*, 18(1), 109–142.
- Rojaszczak, M. (2021). Compliance of automatic tax fraud detection systems with the right to privacy standards based on the Polish experience of the STIR system. *Intertax*, 49(1), 39–52.
- Ronco, S. M. (2020). Data protection in direct tax matters and developments from the EU standpoint: The case of automatic exchange of information. *International Tax Studies*, 2020(4), 2–21.
- Scarcella, L. (2019). Tax compliance and privacy rights in profiling and automated decision making. *Internet Policy Review*, 8(4), 1–19. <https://doi.org/10.14763/2019.4.1422>
- Schaper, M. (2016). Data protection rights and tax information exchange in the European Union: An uneasy combination. *Maastricht Journal of European and Comparative Law*, 23(3), 514–530.
- Setyowati, M. S., Utami, N. D., Saragih, A. H., & Hendrawan, A. (2020). Blockchain technology application for value-added tax systems. *Journal of Open Innovation: Technology, Market, and Complexity*, 6(4), 156.
- Tofan, M., & Arseni, A. A. (2023). The professional conflict pertaining to confidentiality—The obligation of disclosure for intermediaries of financial transactions. *Laws*, 13(1), 2.
- Tomo, A. (2023). Tax information, third parties and GDPR: Legal challenges and hints from the Court of Justice. *EC Tax Review*, 32(4).
- Tomo, A. (2026). Shifting responsibilities and eroding autonomy in tax compliance: Impact of the Directive on Administrative Cooperation on the evolving roles of tax authorities, taxpayers and third parties. *Erasmus Law Review*, 19(2), 62–69.
- Umbet, M., Askarov, D., Rudžionienė, K., Christauskas, Č., & Alikulova, L. (2025). Evaluating the implementation of information technology audit systems within tax administration: A risk governance perspective for enhancing digital fiscal integrity. *Journal of Risk and Financial Management*, 18(8), 422.
- Weiqun, X., & Xing, L. (2025). Can the digitalization of tax administration promote corporate digital transformation? Empirical evidence from China. *International Review of Economics & Finance*, 104634.
- Yussof, S. H., Isa, K., Abdullah, N., & Khalid, M. (2025). Tax secrecy provisions: A comparative analysis between Malaysia and other Commonwealth countries. *Intertax*, 53(12).
- Zakaria, M., Ahmad, W. N. W., Hussin, N. C., Hassan, R. A., Marzuki, M. M., Syukur, M., & Sari, E. N. (2024). Adoption of tax digitalisation among Malaysian tax practitioners. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 22(3), 567–575.
- Zupic, I., & Čater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods*, 18(3), 429–472.