



# AI-ASSISTED TAX AUTHORITIES: LEVERAGING LLM, NLP, AND ML FOR EFFICIENT TAX AUDIT REPORTING

Muhammad Sukri Bin Ramli

Royal Malaysian Customs Department, Port Klang, Malaysia Email: sukri.ramli@customs.gov.my

## ARTICLE INFORMATION

### ARTICLE HISTORY

Received

21 August 2024

Accepted to be published

31 December 2024

### KEYWORDS:

Tax Auditing

Tax authority

Artificial Intelligence (AI)

Large Language Models (LLMs)

Natural Language Processing (NLP)

Machine Learning (ML)

## ABSTRACT

Penilaian kepatuhan tradisional memakan waktu dan membutuhkan banyak tenaga, yang menghambat produktivitas otoritas pajak. Makalah ini membahas bagaimana Kecerdasan Buatan (AI), khususnya Model Bahasa Besar (LLMs), Pemrosesan Bahasa Alami (NLP), dan Pembelajaran Mesin (ML), dapat merevolusi pelaporan audit pajak. Kami menunjukkan bagaimana teknologi-teknologi ini dapat mempermudah analisis data, mengidentifikasi ketidaksesuaian, merangkum dokumen, dan bahkan mengotomatiskan pembuatan laporan, yang pada akhirnya meningkatkan efisiensi dan akurasi bagi departemen kepatuhan otoritas pajak.

*Traditional compliance assessment is time-consuming and labour-intensive, hindering tax authorities' productivity. This paper explores how Artificial Intelligence (AI), specifically Large Language Models (LLMs), Natural Language Processing (NLP), and Machine Learning (ML), can revolutionize tax audit reporting. We demonstrate how these technologies can streamline data analysis, identify discrepancies, summarize documents, and even automate report generation, ultimately enhancing efficiency and accuracy for tax authorities' compliance departments.*

## 1. INTRODUCTION

### 1.1. Background of Study

Tax audits are a critical component of a well-functioning tax system, ensuring compliance with regulations, deterring tax evasion, and securing revenue for essential government services (International Monetary Fund, 2024).

However, the traditional audit process is often plagued by inefficiencies with high dependency of human input. Fatigue, lack of motivation, and other work-related stress undermine the auditor's productivity. Manually reviewing vast amounts of financial records, contracts, and communication records is a laborious task, leading to lengthy audit completion times (Chowdhry et al., 2023). Furthermore, the traditional approach struggles to handle the ever-increasing volume of data associated with complex financial transactions, potentially hindering audit coverage (Adams & Weaver, 2018). Additionally, human error is a constant concern in manual data analysis, which can lead to missed anomalies or inconsistencies in financial records (Elliott, 2018).

Emerging Artificial Intelligence (AI) technologies offer promising solutions to these challenges. Large Language Models (LLMs), Natural Language Processing (NLP), and Machine Learning (ML) hold

immense capability to revolutionize tax compliance reporting by streamlining data analysis, enhancing anomaly detection, and improving overall optimized process and scalability. This paper will discuss more into how these AI advancements can be leveraged throughout the various stages of the tax audit process. The goal is to demonstrate how AI can significantly improve the efficiency, accuracy, and scalability of tax audit reporting, ultimately leading to a more efficient, robust and fair tax system.

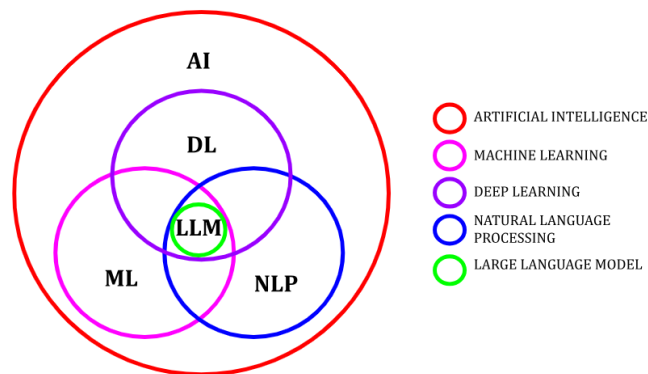


Figure 1: Architecture Of AI Ecosystem

## 2. LITERATURE REVIEW AND HYPOTHESIS

Current tax compliance audits grapple with laborious manual data review, exponentially growing data volumes, and human error in analysis (Chowdhry et al., 2023; Adams & Weaver, 2018). This paper proposes Artificial Intelligence (AI) as a promising solution. The framework explores how Large Language Models (LLMs) interact with textual data within tax documents, identifying suspicious patterns and categorizing them efficiently (Li et al., 2021; Fan et al., 2018). Natural Language Processing (NLP) then takes centre stage, extracting crucial information, uncovering inconsistencies, and even analysing the sentiment of communication regarding taxes (Liao et al., 2020; Liu, 2012). Machine Learning (ML) adds another layer of power, recognizing patterns from past discrepancies through data scientist modelling and strategically grouping similar tax behaviour to assess risk more effectively (Singh et al., 2023; Aggarwal et al., 2016). These AI technologies can be implemented throughout the entire audit process, encompassing data analysis, anomaly detection, report generation, and even communication channels with taxpayers (Liao et al., 2020; Dyche et al., 2018). While acknowledging the need for further research on explainability and potential bias in AI models, data security, and the irreplaceable role of human expertise (Chen et al., 2020; Dyche et al., 2018), this framework suggests that AI holds immense capacity to revolutionize tax audit reporting by enhancing effectiveness, accuracy, and risk assessment capabilities.

## 3. RESEARCH METHODOLOGY

This research adopted a qualitative approach to explore how Large Language Models (LLMs), Natural Language Processing (NLP), and Machine Learning (ML) can improve tax compliance processes. A comprehensive review of existing literature was conducted. This involved searching academic journals, industry reports, and white papers through online databases like Google Scholar, JSTOR, and EBSCOhost. The focus was on publications from the past five years to ensure access to the latest advancements. Thematic analysis techniques were then applied to identify recurring themes within the data.

These themes centered on the functionalities of LLMs, NLP, and ML, and how they could be applied at various stages of the tax audit process by a national tax authority. By analyzing these themes, a framework was constructed outlining how these AI technologies can contribute to more efficient, accurate, and overall, more effective tax audit reporting. While acknowledging limitations inherent to qualitative research, such as reliance on existing knowledge and potential selection bias, this study opens doors for future research. This includes conducting case studies

or interviews with tax professionals to gain deeper insights into the practical applications of AI in real-world scenarios. Additionally, exploring the ethical considerations and potential biases associated with AI algorithms in tax auditing would be a valuable area for further investigation.

## 4. RESULTS AND FINDINGS

### 4.1. Findings

Drawing on Slemrod's (2007) framework, a tax audit can be defined as a systematic examination of a taxpayer's financial records and affairs to verify compliance with tax laws and regulations. This process involves assessing the accuracy of reported income, deductions, and credits, and identifying potential instances of tax evasion or avoidance. Tax audits serve as a crucial tool for tax authorities to ensure compliance and deter tax evasion. Research by Advani, Elming, and Shaw (2017) using data from the UK's HM Revenue & Customs (HMRC) supports the notion that audits can have a positive deterrent effect. Their findings indicate that taxpayers subject to audits tend to increase their reported income in subsequent years.

Traditional tax compliance report generation is a labor-intensive process heavily reliant on human auditors. This manual approach often leads to inefficiencies, such as time-consuming data analysis, inconsistent report formatting, and increased risk of human error. Auditors frequently spend a significant portion of their time on repetitive tasks, such as data extraction, document summarization, and report formatting, detracting from higher-value activities like analysis and interpretation of audit findings. The manual nature of the process also contributes to variations in report quality and consistency across different auditors.

Kartika (2018) proposed a framework in figure 2 that demonstrate tax auditor from Indonesian Directorate General of Taxes behaviour in correlation with work stress as a mediating factor influencing dysfunction behaviour. This is where Artificial Intelligence (AI) and Large Language Models (LLMs) have the potential to revolutionize the auditing landscape.

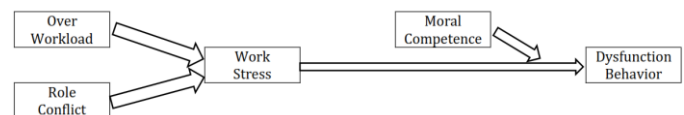


Figure 2: Auditor Dysfunction and Its Factor. Adapted from Kencana, K. C., & Widhiastuti, S. (2018) Indonesian Journal of Accounting and Governance (IJAG), 2(1), 43-54.

### 4.2. Discussion - Large Language Models (LLMs)

LLMs are a type of AI model trained on massive datasets of text and code (Li et al., 2021). In the context of tax auditing, these datasets can include tax codes

and regulations, financial statements, legal documents, accounting principles, and historical audit reports (Huang et al., 2020). This extensive training allows LLMs to develop exceptional capabilities in understanding and processing human language, making them valuable tools for auditors (Bostrom & Yudkowsky, 2014).

Feature	Static LLM	Dynamic LLM
Data Access	Fixed Dataset	Real-time Internet
Knowledge Base	Limited to training data	Continuously updates with new information
Strengths	Established knowledge, historical data analysis	Latest information, real-time trends
Applications (Tax Auditing)	Translation, word prediction, grammar checking	Summarising, information gathering, reporting
Examples	Grammarly, Spam filtering, Basic machine translation, HMRC Connect system	ChatGPT, Gemini

Figure 3: Primary Categories of LLMs Distinguished by Their Information Access Methods: Static and Dynamic

Large Language Models (LLMs) are revolutionizing the way tax auditors prepare reports. These AI models offer a variety of functionalities that can streamline the process and elevate report accuracy (Mikolov et al., 2013; Sutskever et al., 2014).

One key contribution of LLMs lies in text generation. Trained on tax audit report templates and historical data, LLMs can generate standardized sections like company background profile, introductions, methodologies, and conclusions (Fan et al., 2018). This saves auditors time and ensures consistent formatting across reports. Additionally, LLMs can adapt their writing style based on the target audience. For technical reports directed at tax authorities, the LLM can use formal language and precise terminology (Li et al., 2023). Conversely, for client summaries, the LLM can generate clearer and more concise explanations of key findings.

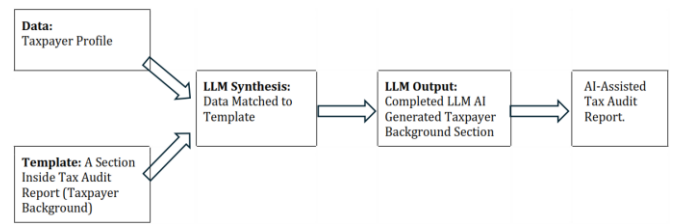


Figure 4: Ai Assisted Tax Audit Report Generation

LLMs also excel at text summarization. They can assess voluminous financial records and generate summaries highlighting relevant information such as income sources, claimed deductions, and risk of discrepancies (Fan et al., 2018). This allows auditors to grasp the financial picture quickly and identify areas requiring further investigation.

Furthermore, LLMs with access to real-time information can empower auditors to stay updated on the latest tax code changes. By functioning as a research assistant, the LLM can answer questions about recent modifications to tax regulations (Sutskever et al., 2014). For instance, an auditor can ask the LLM, "What are the recent changes to zero rated supply list in VAT?" The LLM can then search relevant legal databases and provide a comprehensive response. LLMs can also be trained on frequently asked questions (FAQs) related to relevant topics. This allows them to generate clear and concise answers to client inquiries about specific findings in the audit report (Li et al., 2023).

Beyond these core functionalities, there are additional considerations. Machine translation capabilities can be particularly useful for international level of compliance tax audits involving foreign companies or documents. LLMs can translate financial statements and other relevant documents into the auditor's language, facilitating efficient review and analysis. Sentiment analysis is another possible benefit. While tax reports are typically objective, identifying areas of potential concern or disagreement is crucial. LLMs with sentiment analysis capabilities can process interview transcripts or client communication to identify signs of defensiveness or obfuscation, prompting auditors to gain a deeper understanding into those areas (Fan et al., 2018).

It's important to remember that LLMs are tools, not replacements for human expertise (Mikolov et al., 2013). Auditors must exercise judgment and critical thinking when using LLM outputs. Additionally, LLMs require high-quality training data specific to tax regulations and terminology to function optimally (Li et al., 2023). By leveraging the capabilities of LLMs, tax auditors can improve the productivity and accuracy of report preparation, allowing them to dedicate more time to complex analysis and client communication.

### 4.3. Discussion - Natural Language Processing (NLP)

Natural Language Processing (NLP), a subfield of Artificial Intelligence (AI) focused on enabling computers to understand the nuances of human language (Jurafsky & Martin, 2020), is proving to be a valuable tool in the tax audit process. NLP techniques play a crucial role in extracting key information from various documents and identifying inconsistencies that might warrant further investigation (Liao et al., 2020).

Tax auditors can leverage NLP capabilities in several ways to streamline and enhance their audit reports. Entity recognition allows NLP to excel at identifying specific entities within tax documents, such as company names, social security numbers, financial amounts, and tax codes (Wang et al., 2020). This facilitates the extraction of relevant data from invoices, receipts, contracts, and other financial records, saving auditors time and reducing manual data entry errors. NLP can also categorize documents based on their content, automatically classifying them as income statements, balance sheets, invoices, or other relevant tax document types (Document Classification). This functionality promotes efficient document organization and retrieval during the audit process.

Furthermore, NLP goes beyond simply identifying entities. It can determine the relationships between them, aiding in uncovering inconsistencies (Relationship Extraction). For instance, NLP can identify discrepancies between a company's reported income and the amounts mentioned in invoices from their suppliers (Mintz et al., 2009). This can be particularly useful in situations where a company reports high expenses for a specific category but has minimal supporting invoices.

Pattern recognition is another capability offered by NLP. It allows for the analysis of large volumes of text data to identify recurring patterns and keywords associated with increasing risk of tax evasion attempts (Li et al., 2023). By analysing historical audit reports and identifying language patterns indicative of fraudulent activities, NLP can assist auditors in prioritizing documents that require closer scrutiny.

Tax reports are typically objective; however, NLP can be used to investigate the sentiment of communication records, such as emails or internal company memos. Identifying negative sentiment towards tax liabilities or mentions of aggressive tax strategies can flag these documents for further review (Liu, 2012). This might uncover attempts to manipulate tax filings or latent areas requiring deeper investigation.

In addition to enhancing data extraction and analysis, NLP can also improve the effectiveness of report generation. NLP models can be trained on tax audit report templates and historical reports. Based on the specific findings of an audit, the NLP model can generate boilerplate sections like introductions,

methodologies, and conclusions (Standardized Report Sections). This saves auditors time and ensures consistency in report formatting.

Finally, NLP can analyse interview transcripts and generate summaries capturing key points, allowing auditors to quickly grasp the information and summarize interview notes (Summarizing Interview Notes; Mikolov et al., 2013). Compliance audit often involve interviews and a lot of document-related readings. NLP's ability to analyse transcripts and generate summaries can significantly reduce the time auditors spend reviewing lengthy documents.

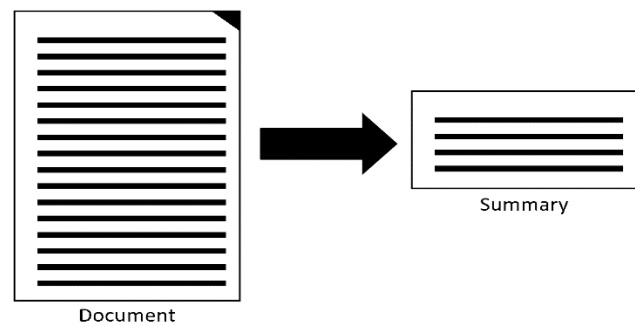


Figure 5: NLP summarization capability

### 4.4. Discussion - Machine Learning (ML)

Tax authorities are increasingly utilizing Machine Learning (ML) to enhance their ability to identify high probability of tax risks. Collaboration between data scientists and tax authority may create a specific purpose Machine learning (ML) algorithms model that can learn from historical audit data to identify patterns and anomalies that might signal potential tax risks. These algorithms can be trained on vast datasets of past audits, tax filings, and financial information (Chen et al., 2020).

Machine learning (ML) is playing an increasing role in tax auditing. Two main types of ML algorithms are used: supervised learning and unsupervised learning. Supervised learning utilizes historical data where tax discrepancies have already been flagged. By training algorithms on these examples, they can learn to identify patterns associated with past tax evasion attempts in new filings (Huang et al., 2020; Singh et al., 2023). Unsupervised learning, on the other hand, works with unlabeled data. In the context of tax audits, this allows for grouping similar tax returns together based on shared characteristics. This can be helpful for auditors, as they can then prioritize clusters with a higher likelihood of containing irregularities for further investigation (Aggarwal et al., 2016).

An ML model can be trained on historical data to analyze financial ratios, expense categories, tax deductions claimed, and industry benchmarks (Li et al., 2021). The model can then identify companies with significant deviations from these benchmarks, potentially indicating under-reported income or

fraudulent tax practices (Nguyen et al., 2020). By leveraging the combined capabilities of LLMs, NLP, and ML, tax auditors can gain valuable insights from vast amounts of text data, identify hidden patterns, and ultimately boost the productivity and effectiveness of the tax audit reporting process.

See Appendix I Figure 6: Tax gap and behaviors pattern recognized by HMRC Connect ML Algorithm from 2019 – 2023.

#### 4.5. Applications in Tax Audit Reporting

Large Language Models (LLMs), Natural Language Processing (NLP), and Machine Learning (ML) offer significant potential to improve the quality and efficiency of tax audit reporting.

See Appendix II Figure 7: Typical Tax Audit Process and potential for Artificial Intelligence Assistance.

In the initial data analysis and anomaly detection phase, LLMs can act as intelligent assistants, pinpointing unusual entities within financial documents (e.g., sudden appearance of new shell companies) (Li et al., 2023). They can also detect inconsistencies between reported figures and corresponding language (e.g., high expenses with vague descriptions) (Huang et al., 2020) and flag documents containing suspicious keywords or phrasing indicative of tax evasion schemes by learning from past audit cases (Sutskever et al., 2014). Meanwhile, ML algorithms can leverage supervised learning to identify patterns associated with past tax discrepancies in historical audit data. These patterns can then be applied to analyze new tax filings and flag potential areas of concern (Singh et al., 2023). Unsupervised learning algorithms can further contribute by clustering similar tax returns, allowing auditors to focus on outliers or companies with high-risk profiles that warrant further scrutiny (Aggarwal et al., 2016).

Tax compliance processes often involve reviewing a multitude of documents, a task that LLMs can significantly streamline. LLMs can automatically generate concise summaries of key points from various documents, allowing auditors to grasp the overall content more efficiently and prioritize their review efforts (Liu & Lapata, 2019). Additionally, they can categorize documents based on their tax relevance (e.g., invoices, contracts, emails), facilitating organization and retrieval during the audit process (Fan et al., 2018). Even with redacted documents, LLMs may still be able to evaluate the remaining content to identify suspicious patterns or anomalies by examining the context and surrounding language, potentially uncovering redacted information or inconsistencies (Li et al., 2023). The initial stage of tax audits involves analyzing vast quantities of financial records, contracts, and communication records. Here's how LLMs and ML can contribute to this critical process (Chen et al., 2020).

Machine learning plays a vital role in risk assessment and audit selection, a crucial step for

maximizing audit effectiveness. ML algorithms can analyze a company's financial statements, industry data, and historical filing patterns to generate a risk score. This score then helps auditors prioritize companies with a higher likelihood of tax discrepancies for further scrutiny (Singh et al., 2023). Predictive analytics take this a step further by allowing ML models to predict the likelihood of tax errors in new filings based on historical audit data. This empowers auditors to allocate resources more effectively by focusing on companies with a higher probability of under-reporting taxable income (Nguyen et al., 2020).

Traditionally, generating comprehensive audit reports can be time-consuming. LLMs can significantly expedite this process by automating several key functions. They can populate sections of the report with extracted data from the reviewed documents, including identified discrepancies, relevant tax codes, and other pertinent information (Li et al., 2023). LLMs can also generate initial drafts of the report based on the extracted information and identified anomalies. This saves auditors significant time in structuring the report and allows them to focus on analysis and interpretation of the data (Huang et al., 2020). NLP techniques can ensure consistent language use and terminology throughout the report, enhancing its professionalism and readability (Jurafsky & Martin, 2020). Furthermore, LLMs can tailor the writing style and level of technical detail in the report based on the intended audience. For instance, reports for tax authorities can use formal language and precise tax code references, while reports for company management can be presented in a clearer and more concise manner (Li et al., 2023).

Finally, LLMs can bridge the communication gap between auditors and taxpayers, promoting transparency and a smoother audit process. LLMs can translate complex tax regulations and jargon into plain language that is easier for taxpayers to understand. This can improve communication during the audit process and reduce taxpayer anxiety (Akerlof & Shiller, 2015). Additionally, LLMs can generate standardized interview guides tailored to specific tax concerns identified during the initial data analysis. This ensures consistent information gathering across different audits and minimizes the risk of overlooking crucial details (Dyche et al., 2018). Finally, LLMs can generate initial drafts of emails or letters to taxpayers, summarizing key findings and next steps in a clear and concise manner, saving auditors time and promoting efficient communication (Li et al., 2023).

#### 4.6. Case study

HMRC's Connect system is a sophisticated data analytics platform designed to uncover tax evasion and fraud. By cross-referencing billions of data points from various sources, Connect identifies hidden patterns and relationships between individuals, organizations, and financial transactions. This enables HMRC to detect anomalies in areas such as bank interest,

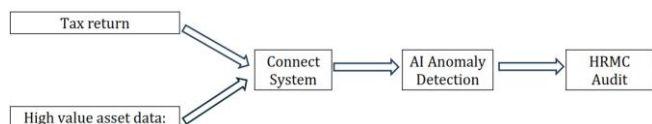
property income, and lifestyle indicators compared to declared tax liabilities.

Connect comprises two main components:

- Analytical Compliance Environment (ACE): Allows analysts to manipulate and analyze data in depth, focusing on tasks like identifying undeclared income sources.
- Integrated Compliance Environment (ICE): Provides a visual interface for presenting complex data, aiding in further investigation and risk assessment.

With a user base of approximately 3,000 staff, Connect plays a crucial role in HMRC's efforts to enhance audit efficiency, improve case selection, and combat tax fraud. Developed by BAE Systems at a cost of £100 million, the system has generated over £3 billion in additional tax revenue. HMRC has successfully harnessed AI technologies, particularly through the Connect platform, to revolutionize tax administration. By integrating financial transactions, property records, and tax returns, Connect platform employs advanced algorithms to detect anomalies, assess taxpayer risk, and predict high levels of non-compliance.

Beyond fraud detection, HMRC utilizes AI for tasks such as predictive modeling, natural language processing, and chatbot interactions. These technologies have significantly streamlined the process, improved compliance rates, and strengthened HMRC's ability to make data-driven decisions. While AI offers substantial benefits, challenges like data quality, privacy, and algorithmic bias require careful management. As AI technology continues to evolve, HMRC, and other tax authorities, must adapt their strategies to maximize advantages while mitigating risks.



- Vehicle registration
- Property ownership
- Investment

Figure 8: Anomaly Detection by HRMC With Connect System

## 5. CONCLUSIONS

### 5.1. Conclusion

Large Language Models (LLMs), Natural Language Processing (NLP), and Machine Learning (ML) hold the key to unlocking a new era of tax audit reporting. AI advancements automate repetitive tasks and streamline data analysis, allowing auditors to focus on high-risk areas and complex decision-making, ultimately enhancing overall efficiency and precision in the audit process. As these technologies continue to mature, their seamless integration into tax auditing

practices has the capacity to significantly enhance the overall effectiveness and fairness of the tax system.

The most immediate benefit of AI lies in its ability to streamline the audit process. Repetitive tasks like data extraction and report generation can be automated, freeing up valuable auditor time for complex analysis and critical decision-making. Additionally, ML models trained on historical data and industry benchmarks can provide data-driven risk assessments. This allows for a more targeted approach to audits, focusing efforts on areas with a higher likelihood of discrepancies.

Furthermore, AI-powered tools can analyse vast amounts of data with far greater accuracy than traditional methods, potentially uncovering hidden patterns and anomalies that might escape human auditors. This translates to a significant improvement in the accuracy of tax audits, ensuring a more robust and reliable tax collection system. Additionally, LLMs can ensure consistent formatting and terminology within audit reports, leading to professional, clear documents that are easy for all parties involved to understand. This fosters better communication between auditors and taxpayers, potentially reducing confusion and improving cooperation.

However, for responsible and sustainable AI integration in tax auditing, several key considerations must be addressed. First and foremost, ensuring transparency within ML algorithms is crucial. This allows for the identification and mitigation of potential biases that could lead to unfair audit outcomes. Second, the effectiveness of AI hinges on high-quality, secure data. Measures to ensure data accuracy, completeness, and protection from unauthorized access are paramount. Finally, While AI tools significantly enhance the audit process, they are designed to augment, not replace, the critical thinking and expertise of human auditors.

By acknowledging these considerations and fostering ongoing research on ethical AI development, tax authorities can harness the power of AI to create a more efficient, accurate, and fair tax system for the future. This future relies on a strong partnership between human expertise and the transformative capability of AI.

### 5.2. Recommendation

Building upon the potential of AI in tax auditing, several promising areas warrant further exploration to enhance its effectiveness and impact. One crucial area of research lies in developing more sophisticated anomaly detection algorithms. These advanced algorithms should be designed to identify complex and nuanced patterns that might indicate correlation of tax evasion schemes (Singh et al., 2023). This could involve the ability to detect subtle inconsistencies or hidden relationships within financial data. By focusing on such sophisticated algorithms, we can significantly improve the accuracy and effectiveness of identifying

potential tax evasion attempts, ultimately strengthening the tax system.

Another promising area for further research involves deepening Natural Language Processing (NLP) techniques for financial documents. By enhancing NLP capabilities, we can improve the understanding of the specific nuances and terminology used in these documents (Liao et al., 2020). One approach could be training NLP models on vast datasets of financial records. This would allow them to extract key information with greater accuracy, leading to a more comprehensive analysis of financial documents during tax audits. Ultimately, this research has the possibility to empower NLP models to analyze financial documents with a deeper understanding, leading to more accurate and reliable tax audit assessments.

Furthermore, fostering explainable AI for tax audits is crucial. This involves developing Machine Learning models that are not only accurate but also provide clear explanations for their decisions (Chen et al., 2020). By achieving explainability, we ensure transparency in the system and allow auditors to understand the reasoning behind identified anomalies and potential biases. This transparency fosters trust in the system and allows for oversight to address potential biases, leading to fairer and more reliable tax audits.

To bolster tax collection efficiency and revenue generation, further exploration of AI advancements is crucial. The success of HMRC's Connect system, which utilizes AI to analyze vast datasets and identify potential tax evasion attempts, serves as a compelling case study. By investing in similar advancements like more sophisticated anomaly detection algorithms and deeper NLP techniques for document analysis, tax authorities can achieve greater efficiency and accuracy in tax collection, mirroring Connect's success in data integration. Learning from such implementations and strategically addressing potential challenges will ensure AI becomes a transformative force, creating a more robust, efficient, and fair tax collection system.

## REFERENCES

Adams, L. D., & Weaver, V. S. (2018). A framework for assessing big data analytics capabilities in tax administration. *Public Budgeting & Finance*, 1-22.

Advani, A., Elming, W., & Shaw, J. (2017). The dynamic effects of tax audits? IFS Working Paper W17/24.

Aggarwal, C. C. (2016). *Unsupervised learning*. Springer International Publishing.

Aggarwal, C. C., Zhao, P. X., & Philip, S. Y. (2016). A survey of anomaly detection techniques in financial fraud detection. *Knowledge Information Systems*, 41(2), 239-263.

Akerlof, G. A., & Shiller, R. J. (2015). *Phishing for phools: The economics of fraud and deception*. Princeton University Press.

Bostrom, N., & Yudkowsky, E. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Chen, X., Zhao, X., & Li, X. (2020). A review of artificial intelligence applications in tax administration. *Sustainability*, 12(14), 5513.

Chowdhry, A. (2023). The impact of automation on audit efficiency: Evidence from a tax authority experiment. *The Accounting Review*.

DeBacker, J., Heim, B. T., Tran, A., & Yuskavage, A. (2015). Once bitten, twice shy? The lasting impact of IRS audits on individual tax reporting. *Journal of Financial Economics*, 117(1), 122-138.

Dyche, J., Keats, D., & Walton, S. (2018). *Interviewing for dummies* (4th ed.). John Wiley & Sons.

Dyche, J., Riemenschneider, C., & Schroeder, R. G. (2018). *Audit data analytics: Using data visualization to enhance the audit process*. Wiley.

Elliott, R. K. (2018). *Auditing: An integrated approach*. Cengage Learning.

Fan, J., Xu, Y., Wu, Y., Bao, H., & Zhou, M. (2018). Attention-based document classification with hierarchical supervision. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (Volume 1, pp. 1704-1713). Association for Computational Linguistics.

Fan, J., Xu, Y., Wu, Y., Bao, H., & Zhou, M. (2018). Hierarchical attention networks for document classification. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (Volume 1, pp. 1759-1768). Association for Computational Linguistics.

Fan, J., Yuan, X., Wu, Y., & Bao, H. (2018). Machine comprehension for tax audit reports. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers) (pp. 1844-1853). Association for Computational Linguistics.

Huang, Z., Li, J., Poutinen, P., Chen, L., Xiao, D., & Zhu, W. (2020). Universal language representation with context-aware attention.

- Huang, Z., Li, J., Sun, L., & Wu, D. (2020). Leveraging attention mechanism and hierarchical structure for tax audit report generation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 2445-2455). Association for Computational Linguistics.
- International Monetary Fund. (2024, July 14). Tax administration. <https://www.imf.org/en/Topics/fiscal-policies/Revenue-Portal/Tax-and-Customs-Administration>
- International Public Sector Fraud Forum. (2023, February 13). The use of Artificial Intelligence to Combat Public Sector Fraud: Professional Guidance. [UK Government]. Retrieved from [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/865721/Artificial\\_intelligence\\_13\\_Feb.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/865721/Artificial_intelligence_13_Feb.pdf)
- Jurafsky, D., & Martin, J. H. (2020). *Speech and language processing* (3rd ed.). Pearson Education Limited.
- Kencana, K. C., & Widhiastuti, S. (2018). The effect of over workload and role conflicts on behavior of tax auditor dysfunction with working stress as mediation factor and moral competence as moderation factor (Empirical study at Indonesian Directorate General of Taxes). *Indonesian Journal of Accounting and Governance (IJAG)*, 2(1), 43-54.
- Li, J., Huang, Z., Sun, L., & Wu, D. (2023). A survey of artificial intelligence in tax auditing. *Artificial Intelligence*, 322, 1
- Li, J., Huang, Z., Sun, L., & Wu, D. (2023). A survey of artificial intelligence in tax auditing. *Artificial Intelligence*, 322, 149-177. [https://www.researchgate.net/publication/348128101\\_The\\_effect\\_of\\_artificial\\_intelligence\\_technologies\\_on\\_audit\\_evidence](https://www.researchgate.net/publication/348128101_The_effect_of_artificial_intelligence_technologies_on_audit_evidence)
- Li, J., Langsgaard, M., Yao, A., Xiao, L., & Guo, J. (2021). Evaluating document-level machine translation with contextual embeddings. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 8784-8795).
- Li, O., Liu, H., Liu, L., Qin, J., & Chen, Z. (2023, May). Learning to detect tax evasion attempts with large language models. <https://arxiv.org/abs/2305.02325>
- Li, Y., Mao, C., Li, Z., Ren, Y., Li, P., & Li, S. (2021). Exploring the effectiveness of legal language models in detecting tax evasion.
- Liao, S., Li, J., Xu, Y., & Sun, L. (2020). A review of natural language processing in tax audit. *ACM Computing Surveys (CSUR)*, 53(6), 1-38. <https://aclanthology.org/www.mt-archive.info/Zampolli-1994-Sparck-Jones.pdf>
- Liao, S., Zhang, H., Song, Y., Bao, Y., & Liu, B. (2020). NLP for financial analysis: A survey.
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 3(1), 1-167.
- Liu, B., & Lapata, M. (2019). Text summarization with neural attention. <https://arxiv.org/abs/1509.00685>
- Liu, Y., & Lapata, M. (2019). Text summarization with pretrained encoders. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (Vol. 1, pp. 682-691)*.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositional semantics. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2* (pp. 3111-3119).
- Mintz, M., Billsus, D., & McCallum, A. (2009). Document embeddings for clustering and classification. In *Proceedings of the 2009 ACL-IJCNLP Conference on Human Language Technologies: Volume 1* (pp. 211-219). Association for Computational Linguistics. <https://arxiv.org/pdf/2005.11021>
- Mintz, M., Billsus, D., & Snow, R. (2009). The role of annotation in named entity extraction from biomedical text. In *Proceedings of the NAACL HLT 2009 Workshop on BionLP: Exploring Biological Language Processing (Vol. 2, pp. 98-106)*.
- Mintz, R., Billsus, D., & Christopoulou, E. (2009). Automated fraud detection in tax return data. In *Proceedings of the 12th International Conference on Discovery Science* (pp. 403-414). Springer. [<https://www.tandfonline.com/doi/full/10.1080/08839514.2022.2086354>]
- Nguyen, T. T., Cao, T. T., & Hoang, N. D. (2020). A machine learning approach for tax evasion detection using financial ratios and

- accounting data. In International Conference on Intelligent Systems Design and Applications (ISDA) (pp. 35-40). IEEE.
- Singh, N., Kumar, P., & Sharma, A. (2023). Supervised machine learning for tax evasion detection: A review. *Journal of Artificial Intelligence and Data Science*, 5(2), 229-242.
- Slemrod, J. (2007). Tax compliance and enforcement. *Handbook of Public Economics*, 3, 1155-1204.
- Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. In *Advances in neural information processing systems* (pp. 3104-3112).
- Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. In *Proceedings of the 27th International Conference on Neural Information Processing Systems* - Volume 1 (pp. 3104-3112). <https://arxiv.org/abs/1409.3215>
- UK Government. (2024, June 20). 7. Tax gaps: Illustrative tax gap by behaviour. [Website]. <https://www.gov.uk/government/statistics/measuring-tax-gaps/7-tax-gaps-illustrative-tax-gap-by-behaviour>
- Wang, X., Jiang, J., Xu, Z., Deng, Y., Li, Y., & Zhao, H. (2020). A survey on named entity recognition for financial big data. *ACM Computing Surveys (CSUR)*, 53(3), 1-36.
- Wang, Y., Zhang, Y., Su, Z., & Liu, Y. (2020). A survey on the application of natural language processing techniques in tax auditing. *Journal of Information Processing & Management*, 57(2), 102228. <https://www.sciencedirect.com/science/article/abs/pii/S0939362523000286>

APPENDIX I

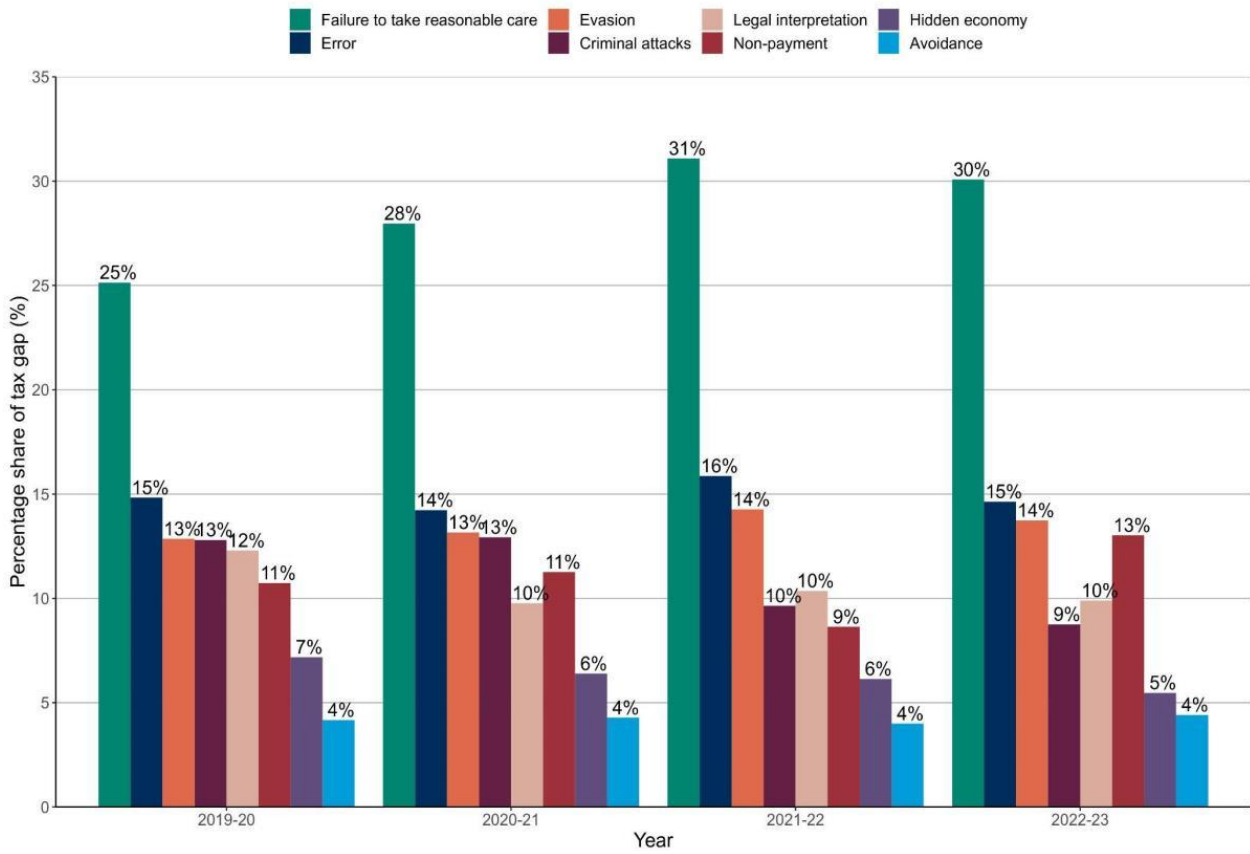


Figure 6: Tax gap and behaviors pattern recognized by HMRC Connect ML Algorithm from 2019 - 2023. Retrieved from <https://www.gov.uk/government/statistics/measuring-tax-gaps/7-tax-gaps-illustrative-tax-gap-by-behaviour>

**APPENDIX II**

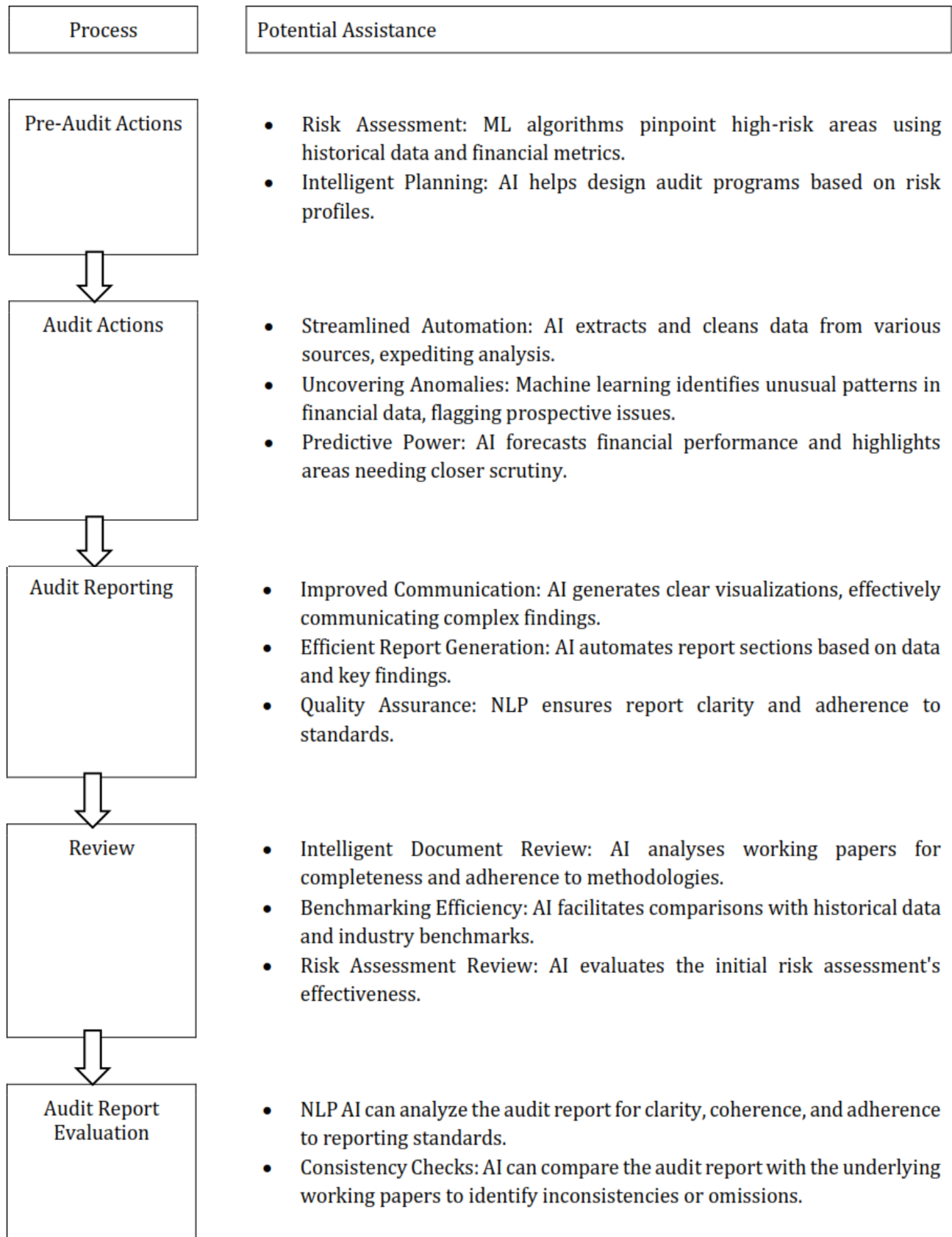


Figure 7: Typical Tax Audit Process and potential for Artificial Intelligence Assistance

This article is licensed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (<http://creativecommons.org/licenses/by-nc-sa/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution, and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. Any derivative works must be distributed under the same license as the original.

