THE PRESENTATION OF PUBLIC BUDGET DATA WITH ARTIFICIAL INTELLIGENCE: TECHNOLOGICAL TRANSFORMATION AND CHALLENGES

YAPAY ZEKÂ İLE KAMU BÜTÇESİ VERİLERİNİN SUNUMU: TEKNOLOJİK DÖNÜŞÜM VE SORUNLAR

Berat KARA

45

Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi
Yıl: 2026 Sayfa: 1-19
https://dergipark.org.tr/tr/pub/cusosbil

DOI: 10.35379/cusosbil.1628837

Geliş Received: **28.01.2025**
Kabul Accepted **18.02.2026**

THE PRESENTATION OF PUBLIC BUDGET DATA WITH ARTIFICIAL INTELLIGENCE: TECHNOLOGICAL TRANSFORMATION AND CHALLENGES

YAPAY ZEKÂ İLE KAMU BÜTÇESİ VERİLERİNİN SUNUMU: TEKNOLOJİK DÖNÜŞÜM VE SORUNLAR

Berat KARA [1]

Keywords:
Artificial Intelligence,
Public Budgeting,
Data Reliability,
Decision-Making Systems,
AI Models Evaluation.

Anahtar Kelimeler:
Yapay Zekâ,
Kamu Bütçesi,
Veri Güvenilirliği,
Karar Alma Sistemleri,
YZ Modellerinin Değerlendirilmesi.

ABSTRACT

The use of artificial intelligence (AI) technologies in the public sector and finance has become a significant research topic in recent years. This study aims to evaluate the accuracy of AI models in providing and analyzing public budget data. In this re-search, the accuracy levels of six different AI models (ChatSonic, Claude, ChatGPT, Perplexity, Gemini, Copilot) in relation to the public budget data for the period 2009-2023 have been examined. Gemini and Copilot were excluded from the evaluation due to their inability to provide complete or reliable data. The ChatSonic and Claude models provided data closer to the actual values compared to ChatGPT and Perplexity. However, even the data provided by these models exhibited unacceptable discrepan-cies from the actual data, particularly in the revenue and expenditure categories. Furthermore, the models presented the expenditure data more accurately, while the error rate for revenue data was higher. Additionally, it was observed that accuracy was higher for older data, while error rates increased as the data approached more recent years. It was also found that the AI models did not account for fractions, which could result in significant discrepancies amounting to millions of Liras. The study reveals that while AI models are limited in meeting the high accuracy requirements for sensitive data such as public budgets. However, ChatSonic and Claude can provide more accurate data, holding potential as decision-making tools in future budgeting processes. These findings offer important implications for the future role of AI tech-nologies in public finance.

ÖZ

Yapay zekâ (YZ) teknolojilerinin kamu sektörü ve maliye alanındaki kullanımı son yıllarda önemli bir araştırma konusu olmuştur. Bu çalışma, YZ modellerinin kamu bütçesi verilerini sağlama ve analiz etme süreçlerindeki doğruluklarını değerlendirmeyi amaçlamaktadır. Çalışmada, 2009-2023 dönemi için altı farklı YZ modelinin (ChatSonic, Claude, ChatGPT, Perplexity, Gemini, Copilot) kamu bütçesi verilerinin sunumundaki doğruluk seviyeleri incelenmiştir. Gemini ve Copilot eksik yahut hiç veri sağlayamadığı için değerlendirmeye alınmamıştır. ChatSonic ve Claude, ChatGPT ve Perplexity'ye göre gerçeğe daha yakın veriler sunmuştur. Ancak bu modellerin sunduğu veriler dahi, gerçek verilerden, özellikle gelir ve gider kalemlerinde, kabul edilemeyecek seviyede farklılık göstermiştir. Ayrıca modellerde harcama verilerinin daha doğru sunulduğu, gelir verilerindeki hata oranlarının ise daha yüksek olduğu görülmüştür. Ek olarak, eski yıllara ait verilerde doğruluk oranı daha yüksekken, güncel verilere yaklaşıldıkça hata oranının arttığı gözlemlenmiştir. YZ modellerinin, milyonlarca lira anlamına gelen küsuratları dikkate almadığı da anlaşılmıştır. Çalışma, YZ modellerinin kamu bütçesi gibi hassas verilerde yüksek doğruluk gereksinimlerini karşılama noktasında sınırlı kaldığını, ancak gelecekte, daha doğru veri sunabilen ChatSonic ve Claude modellerinin, bütçeleme süreçlerinde karar destek aracı olarak potansiyel taşıdığını göstermektedir. Bu bulgular, YZ teknolojilerinin kamu maliyesindeki geleceği için önemli çıkarımlar sunmaktadır.

[1] Dr. Öğr. Üyesi., İstanbul Medeniyet Üniversitesi, Siyasal Bilgiler Fakültesi, e-mail: berat.kara@medeniyet.edu.tr , Orcid No: 0000-0002-6948-2197

Alıntılamak için/Cite as: Kara B. (2026) The Presentation Of Public Budget Data With Artificial Intelligence: Technological Transformation And Challenges, Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, s.1-19

INTRODUCTION

Artificial Intelligence (AI) is more than a mere technological change. Instead, it is shaping how societies interact in socioeconomic and governance processes. The interaction of human intuition with the analytical capabilities of AI is turning around the course of decision-making, while its increasing integration into the governance sector is reformulating principles such as transparency, accountability, and efficiency. From decision-making systems to budget process optimization, advanced AI shows promise in assisting the government in financial complexities with the caveat that it must not only ensure the techno-infrastructural norms but also match the societal and ethical expectations along the way.

AI provides support to decision-makers in several ways, such as providing a second opinion, suggesting expert advice, guiding education, and managing the automation of decision-making processes (Bader et al., 1988; Edwards et al., 2000). Some roles undertaken towards improving decision-making systems are trend analysis, consistency checking, uncertainty assessment, data prediction, anticipation of data needs and proposing alternative actions (Duan et al., 2019; Phillips-Wren & Jain, 2006).

Humans interact with AI by combining their intuition and creativity with AI's analytical power, transforming social and technical practices (Antretter et al., 2020). In public services, AI must meet societal expectations and moral values. Thus, it becomes both a technological driver and a catalyst for social change.

Technical infrastructure serves as an additional requirement. The AI projects in public domain achieve success when multiple stakeholders work together to execute their functions throughout testing process (Puron-Cid, 2014). Public acceptance of AI technology depends on three factors which include transparency and accountability and social inclusion, but ethical questions continue to create doubt. The human factors and social dynamics which interact with technology systems will drive administrative changes throughout organizational development.

AI creates new possibilities, but its potential dangers need experts from multiple fields to assess (Eisikovits et al., 2024). In as much as technological advancement may need to reshape systems and enhance human well-being, the needs of human society, and ethical responsibilities need to take center attention. Bi-directional integration of technical and ethical issues is important toward the application of AI in the public sector.

AI is creating significant transformations in areas such as growth, equality, innovation, and employment, with effects that span all aspects of human society. These developments are not limited to increasing productivity and fostering economic growth but also have the potential to balance the negative impact of an aging population on economic growth by supporting innovation (Agrawal et al., 2018; Liyanage et al., 2019).

With the evolution of AI, automated data-driven decision-making has become increasingly common in organizations, as it has facilitated model and forecast development in ways not seen before (Agrawal et al., 2018; Kassania et al., 2021). Such systems provide decision-makers with tools to contemplate their decisions and highlight the relevant criteria to process in a rapid and timely manner.

AI mitigates personal biases and expedites decision-making (Hogan-Doran, 2017). However, developments in AI have also raised concerns regarding increased unemployment, privacy violations, individual rights and obligations, insecurity, injustice, bias, and the widening digital divide (Gries & Naudé, 2018; Gries & Naudé, 2020; Xu et al., 2018; Autor & Salomons, 2017; Liu et al., 2019; Engin & Treleaven, 2019; Wirtz et al., 2020). Acemoglu and Restrepo (2018c) argue that governments should consider slowing down or suspending AI advancements to address unemployment among low and medium-skilled labor forces. Meanwhile, some researchers contest the notion that AI will fully replace human jobs. Upchurch (2018) suggests that technological singularity is not an imminent issue, as several critical constraints, technical, social, and economic, prevent AI from entirely replacing the workforce. Similarly, Nolan and Slater (2011) assert that the substitution theory lacks practical evidence and empirical analysis.

AI holds promising potential in the domain of public finance and budgeting. This potential is particularly evident in optimizing budgeting decisions, prioritizing expenditures, improving the organizational structure of budgeting, and advancing efficiency, transparency, and accountability (Hollander & Icerman, 1991; Corvalán, 2018; Valle-Cruz, 2019; Anastasopoulos et al., 2020; Valle-Cruz et al., 2022).

In public budgeting, AI identifies expenditure categories, analyzes their impacts on social welfare and economic growth, and improves spending efficiency. It supports scenario building, simulations, and decision-making by helping administrators understand data dynamics and multivariable interactions. Additionally, AI analyzes government actions and economic growth data, evaluates events affecting spending, and processes large datasets to detect patterns and guide budget allocation (Sun & Medaglia, 2019; Valle-Cruz et al., 2020; Valle-Cruz et al., 2022).

The use of technology in governments, facilitated through e-government initiatives, ensures the efficient delivery of public services online, improves data collection and processing, and enhances decision-making processes. AI and smart technologies have introduced opportunities for better-informed citizens and interactions with global economies. Governments have thus turned to these technologies to improve services, foster citizen engagement, and transition toward the concept of smart government by leveraging the power of data and AI-driven analytical insights (Gil-Garcia et al., 2014; Harsh & Ichalkaranje, 2015). Smart government is recognized as one of the fundamental trends with the potential to improve planning and decision-making processes by integrating technologies such as the internet, big data, open data, and AI, supported by the collaboration of public and private sectors, civil society organizations, and citizens (Mellouli et al., 2014; Sun & Medaglia, 2019; Valle-Cruz et al., 2020).

The AI's ability to analyze data makes it possible for citizens, researchers, and students together with public finance specialists to interact with public budget data. The reliance of these groups on the assessments of AI with regard to fiscal transparency and policy analysis opens a Pandora's Box of erroneous or ambiguous claims that may miseducate public discourse and undermine social trust.

Many uneducated users of AI may not exactly know, for what other people have said, how AI processes data and also might take the predictions' reliability for granted. These types of facts may convey a false sense of being informed over government transparency and accountability. Thus, creating an opinion for different representatives. This evidence influences public opinion and leads to major decisions based on speculation, turning AI from a useful tool into a source of distrust.

In this context, various aspects of AI systems are valuable when accuracy is emphasized in the provision of data, not only for enhancing the access and analytical capabilities of users all together, but to also safeguard the institution's public reputations and preserve greater social trust. If reliability malfunctions, which can shake the competence of citizens, the expectation and belief in their governments would suffer, thereby reversing some of the potential that could lead to a greater democratic existence.

Checking how AI presents public budget data precisely is necessary for assessing its potential and the risk that inaccurate information poses. This research thus examines the accuracy of AIs in presenting public budget data in terms of errors engendered by non-expert users, such as students and citizens. The research analyzes budget data from Türkiye between 2009 and 2023 and evaluates the accuracy of six AI models which include ChatGPT, ChatSonic, Perplexity, Claude, Copilot and Gemini. The research comparison will demonstrate how AI functions as a trustworthy data source for investigating budget matters while showing the technology's effects on social systems.

RELATED LITERATURE

AI has developed into an advanced research discipline that now plays a critical role in academic publications. The body of research studies addressing AI through its theoretical frameworks and its real-world uses to demonstrate its effects on economic systems and social structures and organizational activities. AI impacts various industries which include manufacturing and services and education and healthcare. The literature includes a multitude of research examining both possible rewards and threats posed by AI.

Gillham et al. (2018) demonstrated that AI technology increases work efficiency in Chinese healthcare and educational institutions which helps to boost economic growth through higher Gross Domestic Product (GDP) results. According to Chu et al. (2020) AI will enhance production efficiency while decreasing the requirement for human workers. Along similar lines, Jimeno (2019) reviewed the substitution effects case of AI and robotics on the labor market, revealing its contribution to increased macroeconomic uncertainty.

Acemoglu and Restrepo (2018a, 2018b) with Aghion et al. (2017) centered on the role of AI in automation technologies, maintaining that AI was closely related to robotics. Gries and Naudé (2020) similarly assessed the economic impact of AI-based automation. Earlier works such as Hollander and Icherman (1991) proposed AI as a logical complement to mechanical decision-making approaches, in support of decision-making frameworks.

Duan et al. (2019) studied how AI influences different levels of decision-making which include strategic and tactical and operational processes. The research demonstrated that expert systems provide their highest value to organizations at operational and tactical levels while their effectiveness decreases when used for strategic planning. Expert systems function as supportive decision-making tools which help users make better choices, but their success depends on how well users understand the system and how much they use it.

AI has become an important field of research which public sector organizations study. The growing financial pressures along with social needs and operational demands have compelled governments to depend on AI for their public policy development and budget control and public service provision. The research examines how AI affects public sector economics and management and studies its potential to transform government institutions. These studies comprehensively examine the potential benefits of AI in public sector applications, as well as the challenges that may arise.

Zhu et al. (2022) tackle the challenge posed by the competitiveness that arises from government budget restrictions with the effects of AI by constructing a general equilibrium model that incorporates several fiscal policies. Their analysis suggests that purchasing power alleviation would mitigate the adverse economic effect with AI. Government borrowing and direct spending-in contrast to tax reduction-appears to be a more potent solution. They also indicate that regulation in the form of taxes on capital and robots could have positive effects on employment, income distribution, and budget balance.

Santschi et al. (2024) examined the budget surplus and public finance challenges facilitated by AI. The results of this study indicated that an AI model using Swiss data outperformed human experts in forecasting budgets and that this might allow the collaboration between humans and machines to serve public finance by reducing flexibility in the budget.

AI for all transversal processes using multi-layer perceptron and genetic algorithms for the study of public budgeting processes, Valle-Cruz et al. (2022) argue that the choice of AI is the viable means to achieve the cumulative macroeconomic objectives of increasing GDP, combating inflation, and rectifying income inequality. Their study found similar AI-based analyses in public budgeting as the conventional statistical methods used for the public budgets.

The research of AI by Sun and Medaglia (2019), Valle-Cruz et al. (2020), and Wirtz et al. (2019, 2020) deals with public policy processes stressing automation, ethics, explainability, and trust. These studies demonstrate that AI functions as more than a technical instrument because it serves as a fundamental force that drives change in public sector operations. The Dutch government uses AI technology for developing budget forecasts which establish a reference framework to assess government policy under applicable uncertainty according to Kolkman (2020). The research shows that AI analysis results lead to beneficial outcomes for public sector decision-making processes.

The basic government function of public budgeting requires resource allocation decisions which Fernandez-Cortez et al. (2020) identified as essential since they determine all public policies and programs. The researchers developed an AI-based system which distributes public funding through their study. The Mexican federal government distribution of public budgeting through their study developed an AI-based system which proposed recommendations for increasing social development investments while decreasing off-budget spending.

Yetkin Ataer (2025) investigates how AI affects public budget systems by analyzing revenue generation and expenditure control. AI technology improves tax administration effectiveness which leads to increased public revenue collection according to the research findings. The study shows that AI technology will change employment patterns which results in extra obligations for social security systems. The study requires public budget reallocation through adaptable financial resources which operate under permanent strategic frameworks. The author proves that the state needs both regulatory powers and supportive functions to help AI technologies enter Türkiye's public budget system.

Aydın (2024) provides a complete study of AI applications in public service delivery by reviewing existing literature according to The Organization for Economic Co-operation and Development's (OECD) service classification system. The author demonstrates that AI functions throughout all government operations which include public services and defense and security and economic affairs and environmental protection and housing and health and culture and education and social assistance; these applications provide efficiency and speed and cost advantages which improve service delivery. The research reveals that AI will create major benefits for taxation which directly supports public budgets when it performs auditing and informal sector work and standardization tasks although the study also identifies essential challenges which include data security and ethical issues and privacy breaches and the potential for AI systems to choose personal gain instead of public good.

Lee, Hayes, and Maher (2025) conduct experimental research to evaluate AI capabilities in budgeting and financial analysis for small local governments through their study of Fort Calhoun, Nebraska. The authors show that the free version of ChatGPT-3.5 can only process limited data because it provides essential descriptive analysis tools together with simple policy recommendation functions. ChatGPT-3.5 together with Python provides users with advanced analysis tools but small local governments face challenges because they must acquire coding skills to use these features. The monthly subscription service of ChatGPT-4 provides users with a powerful tool that enables them to upload data directly while offering advanced visualization features and financial condition analysis and forecasting and policy recommendation capabilities. The study shows how AI can help small local governments improve their public budgeting processes through increased efficiency and analytical capabilities while showing that output quality depends on three factors which include data quality and prompt accuracy and human monitoring.

Yavuz and Özgül (2025) use the national AI strategies and actual AI implementations of seven nations from the Government AI Readiness Index to study AI's effect on public budget systems together with the case of Türkiye. The authors show that public budget processes can be managed through AI which functions in all three budget phases of planning and preparation and implementation and auditing. The system functions as a planning system which forecasts future requirements and uncertainties for France and the UAE while it serves as a budgeting tool for Australia and Estonian organizations use it to streamline administrative work and Turkish Court of Accounts audits use it as an advisory solution. The study shows how AI technology boosts efficiency and productivity and accountability through complete public budget processes, yet it needs more resources and trained workers to achieve its full potential.

Aldemir and Ucma Uysal (2025) investigate how AI reshapes public financial management and governmental governance through their qualitative study which analyzes Estonia, Singapore, and Finland. The authors demonstrate that AI-driven predictive analytics, fraud detection systems, and automated reporting tools significantly improve operational efficiency, transparency, and decision-making processes in public financial management. The automated tax filing and budget forecasting systems of Estonia together with Singapore's "Moments of Life" initiative and ACQAR system and Finland's AuroraAI program demonstrate how this transformation has occurred. The study shows how AI technology helps public budget processes of accountability and transparency while showing that algorithmic bias and data privacy problems and the requirement for strong ethical guidelines prevent fair technology use.

Özdemir and Yelboğa conducted their study by examining how AI affects public expenditure auditing through the Turkish Court of Accounts' 2022-2023 audit results. The authors demonstrate that recurring compliance and financial reporting errors can be proactively identified through rule-based systems, anomaly detection, and natural language processing techniques. AI-supported continuous audit systems enable better detection of structural weaknesses that exist in immovable property records and depreciation transactions and procurement processes. The study demonstrates that AI systems can improve public budget expenditure by delivering better accountability and transparency, but these changes require data quality and explainability and human oversight to succeed.

The study conducted by Chung et al. (2025) examined LLMs through their application of ChatGPT to predict local government revenue streams using anonymized monthly data from the Government Finance Officers Association (GFOA). The research evaluated LLM-generated predictions together with conventional forecasting techniques which included Holt's exponential smoothing and ARIMA and machine learning methods which included KNN and GRNN. The study discovered that LLMs produced forecasts with high errors which reached sMAPE values of 85.13%. The integration of human-in-the-loop interventions which included detrending data and hyperparameter configuration led to forecast accuracy enhancement which achieved error reductions down to 9.9% for annual predictions. The authors present LLMs as an economical solution for small and medium governments that operate on limited budgets while they stress the need for human supervision to reduce system biases and hallucination errors. The researchers recommend institutional backing from organizations such as GFOA to create customized LLM solutions for public finance applications which demonstrate the technology's capability to function as an additional forecasting method instead of an exclusive forecasting solution.

RESEARCH QUESTIONS

Researchers highlight widely accepted grounds for viewing AI technologies as decision-making tools in the public sector. However, studies which examine the accuracy and reliability of information provided by these technologies have been relatively few. The public finance sector requires assessment of AI-generated data reliability because it handles sensitive information and needs precise results. Most papers published so far are based on strategic contributions of AI to the budgeting process, while scant evaluation exists on whether AI is actually helping to ensure accurate and reliable budget data provision and analysis. The testing process of AI-based solutions needs to demonstrate their reliability and effectiveness for non-professional users because public finance data analysis and budgetary item evaluation carry critical importance. The testing process will identify risks and opportunities which will help evaluate the advantages and disadvantages of AI as a source of accurate information.

The research establishes its value because it addresses a critical research gap through its examination of AI-generated budget data accuracy which enables non-experts to access information about public finance. The study examines how AI systems function within public finance. The research tests the ability of AI to present public budget information in an accurate and trustworthy manner to users who lack professional expertise. The study advances research developments while it fills a major knowledge void which exists in this area of study. The study investigates two primary research questions.

Research Question 1: *What is the accuracy of the information provided by AI models regarding the public budget data in Türkiye?*

Research Question 2: *Which AI model provides accurate or more accurate public budget data, and which has more potential for future development in this field?*

These research questions will help in showing how AI technologies are likely to give valid and reliable data on public budgets to non-professional users. The study also intends to provide valuable input related to how access to such data can be extended to non-professional individuals in a more efficient manner in assessing the accuracy of AI models presenting the data on public finance. Eventually, these findings will bring forth the accuracy of data provided by such technologies and give a roadmap to further applications. The research paper goes more in-depth into how possible uses of AI can enable assurance that non-professional persons have safe and accurate access to public budget data.

DATA SET AND METHODOLOGY

The study utilizes a comprehensive and systematically compiled public budget dataset for Türkiye covering the period 2009–2023. This temporal span was deliberately selected to capture the post-Global Financial Crisis fiscal environment and to ensure analytical consistency within a structurally comparable policy framework. The endpoint of 2023 represents the most recent year with fully published and finalized budget data available at the time of the study, thereby ensuring both relevance and data completeness. This dataset encompasses various essential components of public finance, categorized under the following headings: Total Budget Revenue (TBR), Income Tax (INC), Corporate Tax (COR), Motor Vehicle Tax (MVT), Domestic Value Added Tax (DVT), Imported Value Added Tax (IVT), Excise Tax (EXC), Bank and Insurance Transactions Tax (BTT), Special Communication Tax (SCT), Total Budget Expenditure (TBE), Personnel Expenditures (PER), State Premium Expenditures to the Social Security Institution (SSE), Goods and Services Procurement Expenditures (GSP), Interest Expenditures (INT), Current Transfers (CUR), Capital Expenditures (CAE), Capital Transfers (CAT), and Non-Interest Budget Expenditures (NBE). These components serve as vital indicators of Türkiye's economic activities and public finance, and they have been thoroughly analyzed to assess the reliability of the budget data. By examining this data, the research seeks to identify the areas where AI models yield more accurate results and where potential accuracy challenges exist.

In the research, six different AI models were used to perform accuracy analyses on Türkiye's public budget data: ChatGPT, Perplexity, ChatSonic, Claude, Copilot, and Gemini. The data provided by these AI models were compared with the actual budget data obtained from the Ministry of Treasury and Finance of the Republic of Türkiye (MTF), and the accuracy of each model was carefully evaluated. These comparisons have more clearly revealed the data presentation capabilities of the models and their accuracy rates concerning public finance data.

RESULTS

The performance of AI models in analyzing Türkiye's 2009-2023 budget items showed significant differences in accuracy and scope. ChatGPT, Perplexity, ChatSonic, and Claude demonstrated strong data processing abilities, providing comprehensive and understandable answers, showing great potential for detailed budget analysis. However, Gemini failed to provide relevant data, indicating its limited capacity. Copilot also showed notable deficiencies, offering incomplete and limited data, making its budget presentations unreliable. Comparisons with actual budget data from the MTF provide a critical reference for evaluating the models' accuracy and reliability in presenting budget information.

Table 1. Comparison of Budget Expenditures: Actuals vs. AI-Generated Results (Million TRY)

Period	Source	TBE	PER	SSE	GSP	INT	CUR	CAE	CAT	NBE
2009	**Actual**	268.219	55.946	7.208	29.798	53.200	91.975	20.071	4.319	215.019
	ChatGpt	267.300	61.100	39.500	27.200	37.000	115.300	23.700	6.000	233.600
	Perplexity	283.100	72.900	12.800	32.800	30.000	110.500	30.900	6.700	234.000
	Chatsonic	268.200	55.900	7.200	29.800	53.200	91.900	20.100	4.300	215.000
	Claude	268.200	55.900	11.100	29.800	53.200	91.900	200.100	4.300	215.000
2010	**Actual**	294.359	62.315	11.063	29.185	48.299	101.857	26.010	6.773	246.060
	ChatGpt	293.600	68.900	44.800	32.000	38.700	132.600	26.300	6.900	259.100
	Perplexity	319.000	80.300	13.500	35.000	32.000	115.000	31.500	7.000	258.000
	Chatsonic	294.400	62.300	11.100	29.200	48.300	101.900	26.000	7.400	246.100
	Claude	294.400	62.300	12.900	29.200	48.300	101.900	260.100	7.400	246.100
2011	**Actual**	314.607	72.914	12.850	32.797	42.232	110.499	30.905	6.739	272.375
	ChatGpt	313.300	80.000	52.700	38.500	44.100	153.700	31.500	7.800	278.000
	Perplexity	368.300	86.500	14.700	32.500	36.000	129.300	34.200	6.000	272.400
	Chatsonic	314.600	72.900	12.900	32.800	42.200	110.500	30.900	6.700	272.400
	Claude	314.600	72.900	13.900	32.800	42.200	110.500	30.900	6.700	272.400
2012	**Actual**	361.887	86.463	14.728	32.894	48.416	129.477	34.365	6.006	313.471
	ChatGpt	360.500	88.500	59.900	42.400	50.200	173.500	35.800	9.500	314.600
	Perplexity	422.400	92.100	15.800	32.500	40.000	142.000	34.200	6.500	312.100
	Chatsonic	361.900	86.500	15.300	32.900	48.400	129.500	34.400	6.600	313.500
	Claude	361.900	86.500	15.300	32.900	48.400	129.500	34.400	6.600	312.700
2013	**Actual**	408.225	96.235	16.306	36.386	49.986	148.743	43.767	7.666	358.239
	ChatGpt	407.900	97.300	66.800	47.900	57.100	194.100	40.700	10.200	362.500
	Perplexity	486.000	100.000	17.000	35.000	42.000	150.000	36.000	7.000	357.900
	Chatsonic	408.200	96.200	17.100	36.400	50.000	148.700	43.800	7.700	358.200
	Claude	408.200	96.200	16.300	36.400	50.000	148.700	43.800	7.600	358.200
2014	**Actual**	448.752	110.370	18.929	40.801	49.913	162.282	48.201	7.707	398.839
	ChatGpt	448.800	107.100	73.900	53.200	64.600	212.000	45.600	11.500	402.200
	Perplexity	527.000	110.000	18.500	37.500	45.000	160.000	38.000	7.500	396.000
	Chatsonic	448.800	110.400	19.400	40.800	49.900	162.300	48.200	8.300	398.800
	Claude	448.800	110.400	1.900	40.800	49.900	162.300	48.200	7.600	398.800
2015	**Actual**	506.305	125.051	21.045	45.563	53.004	182.671	57.199	10.438	453.301
	ChatGpt	506.300	119.500	82.100	59.400	70.900	229.400	49.800	12.600	454.400
	Perplexity	590.000	120.000	20.000	40.000	50.000	170.000	40.000	8.000	438.000
	Chatsonic	506.300	125.100	21.800	46.000	53.000	182.700	55.000	9.600	453.300
	Claude	506.300	125.100	20.700	46.400	53.000	182.700	55.000	8.800	453.300
2016	**Actual**	584.071	148.864	24.699	54.100	50.247	224.847	59.677	8.881	533.824
	ChatGpt	584.100	129.100	92.300	63.700	85.600	247.500	54.200	13.800	523.400
	Perplexity	635.000	130.000	22.000	42.500	55.000	180.000	42.500	8.500	485.000
	Chatsonic	584.100	148.900	25.200	54.100	50.200	224.000	59.000	11.800	516.300
	Claude	584.100	147.600	24.800	53.700	50.200	201.600	66.500	9.000	520.400

2017	**Actual**	678.269	162.146	27.272	63.600	56.712	270.924	70.983	13.342	621.557
	ChatGpt	677.700	141.500	102.000	70.100	98.300	266.400	58.700	15.000	620.500
	Perplexity	700.000	140.000	24.000	45.000	60.000	190.000	45.000	9.000	530.000
	Chatsonic	678.300	164.000	28.200	60.400	56.700	265.200	74.800	14.300	598.600
	Claude	678.300	164.000	27.800	62.800	56.700	254.000	85.000	12.800	621.600
2018	**Actual**	830.809	200.903	34.379	71.946	73.961	322.879	88.324	16.746	756.848
	ChatGpt	830.500	151.300	109.200	74.600	118.100	285.400	62.300	16.800	759.600
	Perplexity	750.000	150.000	26.500	47.500	65.000	200.000	47.500	nda	580.000
	Chatsonic	830.500	199.300	34.400	67.600	74.000	323.100	88.300	15.400	737.000
	Claude	830.500	199.300	33.900	73.100	74.000	329.600	88.500	1.300	757.500
2019	**Actual**	1.000.027	249.892	43.045	84.363	99.940	400.316	80.717	16.316	900.087
	ChatGpt	999.500	168.200	120.400	83.900	118.100	305.700	70.100	18.400	877.400
	Perplexity	800.000	160.000	28.000	50.000	70.000	210.000	50.000	10.500	620.000
	Chatsonic	999.500	241.700	41.200	75.000	99.900	391.300	78.800	20.500	895.400
	Claude	999.500	241.700	40.500	83.400	99.900	397.100	78.800	19.500	911.900
2020	**Actual**	1.203.737	287.785	48.294	96.971	133.962	498.063	93.742	15.171	1.069.775
	ChatGpt	1.202.300	186.100	135.100	93.500	167.500	338.600	76.400	19.800	1.044.100
	Perplexity	1.000.000	170.000	30.000	52.000	75.000	220.000	52.000	11.000	700.000
	Chatsonic	1.203.700	282.500	48.000	74.000	134.000	480.000	101.500	24.500	1.068.300
	Claude	1.202.200	282.800	47.700	96.100	134.000	478.200	101.800	23.300	1.068.300
2021	**Actual**	1.603.545	346.279	57.380	133.455	180.852	626.828	131.282	25.492	1.422.693
	ChatGpt	1.599.500	204.200	151.800	104.900	210.200	374.500	81.300	21.200	1.465.100
	Perplexity	1.500.000	180.000	32.000	55.000	80.000	230.000	55.000	11.500	850.000
	Chatsonic	1.599.600	320.100	54.400	93.700	180.900	621.000	123.000	36.500	1.419.400
	Claude	1.599.600	348.400	58.900	122.700	180.900	649.200	144.700	38.600	1.454.500
2022	**Actual**	2.942.748	615.296	96.864	257.660	310.903	1.126.363	276.896	48.822	2.631.845
	ChatGpt	3.133.900	225.900	169.100	119.500	297.900	410.200	89.600	23.700	2.907.400
	Perplexity	7.158.000	200.000	35.000	60.000	90.000	250.000	60.000	12.000	1.200.000
	Chatsonic	3.107.900	536.700	88.700	193.400	276.600	1.330.200	245.900	97.500	2.831.300
	Claude	3.113.800	615.200	99.100	235.600	282.100	1.495.800	238.200	64.500	2.831.700
2023	**Actual**	6.585.456	1.324.530	185.735	452.855	674.615	2.373.641	542.997	858.059	5.910.841
	ChatGpt	6.585.000	250.300	186.400	134.200	331.500	448.300	95.200	25.900	5.810.000
	Perplexity	nda	nda	nda	nda	nda	nda	nda	nda	nda
	Chatsonic	5.901.200	1.091.000	175.900	379.500	620.000	2.571.000	463.000	189.000	5.281.200
	Claude	5.626.000	1.095.000	177.000	421.000	385.000	2.851.000	486.000	114.000	5.241.000

nda: No data available.
Actual: Compiled by the author from the annual budget implementation result reports of the MTF.

Table 2. Comparison of Budget Revenues: Actuals vs. AI-Generated Results (Million TRY)

Period	Source	TBR	INC	COR	MVT	DVT	EXC	BTT	SCT	IVT
2009	Actual	215.458	39.668	20.701	4.585	34.034	43.708	4.043	4.274	26.135
	ChatGpt	215.100	47.600	13.500	4.400	2.520	4.410	390	210	3.250
	Perplexity	237.500	28.000	18.000	4.500	12.000	16.000	1.200	500	20.000
	Chatsonic	215.500	38.400	20.700	4.600	34.000	43.600	4.200	4.100	20.100
	Claude	215.500	46.000	20.700	4.900	34.000	43.700	4.200	4.400	28.200
2010	Actual	254.277	41.969	22.854	5.128	39.438	58.038	3.577	4.127	36.210
	ChatGpt	254.000	58.000	22.300	5.000	3.030	5.240	450	240	4.290
	Perplexity	267.500	30.000	20.000	5.000	14.000	18.000	1.500	600	22.000
	Chatsonic	254.300	40.400	22.900	5.000	39.500	58.000	4.200	4.800	30.100
	Claude	254.300	49.400	22.900	5.500	39.500	58.000	4.300	4.500	37.500
2011	Actual	296.824	51.092	29.233	6.060	46.860	64.263	4.309	4.420	48.690
	ChatGpt	296.800	67.900	29.800	5.900	3.740	6.610	560	300	5.610
	Perplexity	307.800	35.000	25.000	6.000	16.000	20.000	1.800	700	25.000
	Chatsonic	296.800	50.600	29.200	6.000	48.700	64.300	4.300	5.000	38.000
	Claude	296.800	59.800	29.200	6.900	48.300	64.300	4.900	500	48.700
2012	Actual	332.475	58.797	32.111	6.773	53.150	71.793	5.491	4.477	50.005
	ChatGpt	331.700	76.300	29.000	6.800	4.040	7.460	610	330	6.420
	Perplexity	347.700	40.000	30.000	7.500	18.000	22.000	2.200	800	27.500
	Chatsonic	332.500	59.800	32.100	6.800	53.200	71.800	5.100	5.800	42.800
	Claude	332.500	69.700	32.100	7.700	53.200	71.800	5.600	5.600	50.000
2013	Actual	389.682	65.914	31.434	7.397	61.144	85.770	6.167	4.567	62.733
	ChatGpt	389.400	87.500	31.000	7.600	4.680	8.370	680	360	7.780
	Perplexity	388.500	45.000	35.000	8.500	20.000	25.000	2.500	1.000	30.000
	Chatsonic	389.700	69.700	31.400	8.100	62.600	85.800	6.200	6.400	48.700
	Claude	389.700	78.700	31.400	8.100	61.700	85.800	6.200	6.100	58.400
2014	Actual	425.383	76.451	35.163	7.834	66.124	91.657	7.494	4.658	64.413
	ChatGpt	425.400	99.300	32.400	8.300	5.150	9.160	740	390	8.630
	Perplexity	427.100	50.000	40.000	9.500	22.500	28.000	2.800	1.200	32.500
	Chatsonic	425.400	77.700	35.200	8.000	68.900	91.700	7.200	7.000	55.800
	Claude	425.400	91.100	35.200	8.500	68.000	91.700	7.300	6.800	65.800
2015	Actual	482.780	88.138	37.009	8.984	79.188	106.646	9.181	4.744	74.655
	ChatGpt	482.800	111.200	39.200	9.200	5.670	9.830	810	420	9.590
	Perplexity	478.600	55.000	45.000	10.500	25.000	30.000	3.100	nda	nda
	Chatsonic	482.800	90.100	37.000	9.100	84.000	106.600	8.500	7.700	65.400
	Claude	482.800	105.900	37.000	9.500	76.900	106.600	900	7.700	76.500
2016	Actual	554.140	99.822	46.898	10.029	91.966	121.221	11.076	4.990	76.842
	ChatGpt	554.400	125.600	41.100	10.100	6.180	10.690	900	460	10.880
	Perplexity	533.000	60.000	50.000	11.500	27.500	32.500	3.300	nda	nda
	Chatsonic	554.100	106.500	46.900	10.100	100.300	121.000	10.800	8.300	72.700
	Claude	554.100	123.700	46.900	10.900	84.700	121.200	11.100	8.500	84.400

2017	**Actual**	630.490	116.043	57.868	10.860	106.573	139.370	13.297	3.865	100.105
	ChatGpt	630.300	141.300	52.900	11.300	6.790	13.070	1.050	510	12.630
	Perplexity	610.000	65.000	55.000	12.500	30.000	35.000	3.600	1.600	40.500
	Chatsonic	630.500	124.300	57.900	11.100	123.500	139.400	12.400	9.300	93.400
	Claude	630.500	143.900	57.900	1.200	106.300	139.400	13.500	9.900	106.500
2018	**Actual**	757.996	142.966	84.132	12.887	128.359	135.133	18.195	3.447	122.301
	ChatGpt	757.800	160.500	74.000	12.500	7.920	13.790	1.240	560	15.780
	Perplexity	650.000	70.000	60.000	13.500	32.500	38.500	3.900	1800	42.500
	Chatsonic	757.800	158.000	84.100	13.100	137.300	133.900	16.400	10.500	120.800
	Claude	757.800	175.400	84.100	14.100	137.100	146.600	17.400	11.100	136.500
2019	**Actual**	875.280	167.754	87.528	14.568	144.860	148.196	22.689	4.075	125.292
	ChatGpt	875.800	182.400	78.600	13.600	8.830	15.970	1.410	620	18.520
	Perplexity	710.000	75.000	65.000	14.500	35.500	40.500	4.100	2.100	45.500
	Chatsonic	875.800	193.600	78.600	14.500	137.400	162.700	20.100	11.400	136.000
	Claude	875.800	236.700	78.500	15.500	146.300	174.400	21.200	12.300	136.100
2020	**Actual**	1.028.446	163.839	112.871	16.044	164.814	209.090	27.110	4.489	159.224
	ChatGpt	1.029.500	204.100	96.300	15.100	9.740	16.740	1.620	680	21.370
	Perplexity	850.000	80.000	70.000	15.500	37.500	42.500	4.400	2.300	50.500
	Chatsonic	1.029.500	203.300	89.000	15.200	145.700	196.300	22.200	12.200	175.300
	Claude	1.029.500	203.800	89.800	16.600	166.000	203.500	24.400	13.200	167.500
2021	**Actual**	1.402.038	224.338	190.148	18.875	271.726	206.996	33.120	6.647	261.944
	ChatGpt	1.407.400	228.700	181.000	18.000	12.360	21.370	2.200	820	31.220
	Perplexity	1.250.000	90.000	80.000	16.500	40.500	45.500	4.800	nda	nda
	Chatsonic	1.407.400	267.400	137.400	19.500	208.300	234.300	30.400	15.100	276.500
	Claude	1.407.400	267.500	189.200	21.700	228.100	261.700	36.400	16.200	235.500
2022	**Actual**	2.800.088	356.464	507.453	23.626	175.143	429.525	58.638	9.298	578.965
	ChatGpt	2.802.400	255.600	361.000	21.000	24.700	38.900	4.000	1.000	62.000
	Perplexity	3.762.000	356.000	507.000	20.000	175.000	167.000	4.000	nda	579.000
	Chatsonic	2.736.800	536.100	270.300	35.000	424.500	349.700	51.600	24.100	576.200
	Claude	2.801.300	536.100	384.700	33.400	465.800	436.600	59.400	26.700	503.100
2023	**Actual**	5.210.488	694.014	786.314	40.175	505.466	928.195	132.742	16.650	949.510
	ChatGpt	5.210.000	281.700	400.000	25.000	30.000	50.000	5.000	1.200	70.000
	Perplexity	nda	nda	nda	nda	nda	nda	nda	nda	nda
	Chatsonic	4.961.700	1.080.000	717.000	60.000	800.000	615.000	100.000	40.000	850.000
	Claude	4.961.000	1.044.000	758.000	58.000	784.000	690.000	98.000	42.000	883.000

nda: No data available.
Actual: Compiled by the author from the annual budget implementation result reports of the MTF.

As it is possible to see from the tables, the accuracy of data from the four AI models for the 2009–2023 period varies significantly. ChatSonic delivers superior accuracy and reliability, which makes it the top choice followed by Claude whereas ChatGPT and Perplexity deliver inferior results. Perplexity shows data gaps which prevent proper analysis of public budgets while ChatGPT demonstrates high error rates that expose its inability to process detailed datasets.

The study shows that revenue and expenditure total performance differs from the performance of its specific components. The models achieved accurate results for overall figures, yet ChatSonic and Claude produced better results for specific items whereas ChatGPT and Perplexity made critical mistakes. The results show that AI models achieve better performance when working with complete data sets, but they encounter difficulties when trying to process specific details.

Also, it follows from the tables that the accuracy level of data given by AI models considerably differs depending on the period taken into consideration. The older the data, the more correctly it is represented, and vice versa, error rates increase markedly as the timeline approaches the present day. This may be explained by the fact that historical data is more frequently repeated across sources, making it easier for AI models to validate such information. Furthermore, stability and more complete processing of data from older periods may favorably influence the performance of AI, while uncertainties and the dynamic character of recent data contribute to reduced accuracy.

Another striking observation is the tendency of AI models to round figures while presenting budget data. In cases where this involves large-scale public budgets, the differences can range up to millions of Turkish Lira. This is related to a lack of precision and sensitivity of such models and represents further limitations, with serious financial consequences likely or possible. This result emphasizes that, in their current form, AI models are not capable of reaching high accuracy levels that could be expected in such basic documents as budgets.

In addition, the expenditure data has been represented more accurately than the revenue data, which shows a great difference in the processing of different types of data by AI models. This difference shows that AI models may work better with data types that are more structured or have a clearer trend, such as expenditure data. However, the revenue data may contain higher error rates due to its complex structure or sources that are less likely to be repeated. This once again proves that AI models should be adjusted according to the data categories and optimized based on the nature of the data they process.

Though the AI models have been able to attain a certain degree of accuracy in the public budget information, it can be observed that all the information, to a varying extent, is prone to errors. In the context of critical documents like budgets, which are in the order of billions of Turkish Lira, even the slightest mistake can be extremely hazardous. However, the relatively superior accuracy of models such as ChatSonic and Claude indicates the potential of AI technologies to become more accurate in such critical areas. The relative advantages of the two models indicate that AI technology may be effectively used in the future as a decision-making aid in budgeting.

DISCUSSION AND POLICY RECOMMENDATIONS

The implementation of AI technology in public fiscal management demonstrates its ability to revolutionize public financial operations although the technology introduces fundamental operational difficulties which need to be addressed. The existing problems stem from three main factors which restrict access to current detailed public information and create model hallucinations because of incomplete training datasets and create budget document accuracy issues because AI-generated data lacks the required information protection level. The lack of common data standards between fiscal agencies prevents organizations from applying AI technology effectively. The solution to these problems needs two components which include both enhancing existing models and reforming organizations through improved open access data methods and clear data description systems and mixed human-AI work systems.

AI technologies in public finance can enhance public service planning through their ability to automate budgeting operations while generating better results for resource distribution among different groups when implemented successfully. Their capacity to analyze extensive intricate datasets enables them to deliver fresh approaches which help organizations conduct budget assessments and financial forecasting activities. The realization of this potential depends on three key factors which include AI-generated information needing to be both precise and accurate and trustworthy. The use of AI technology in public finance operations creates more problems than it solves because it needs verification of its essential functions.

The study findings demonstrate that AI models fail to deliver accurate and trustworthy public budget information. The error rates identified in the output of these models, particularly when applied with large and complex datasets, suggest that, in current form, AI models may not be suitable for deployment as reliable information sources in sensitive domains. The different accuracy results produced by the models demonstrate that their performance depends on both their core algorithms and their data processing abilities which create a risk of delivering incomplete and inconsistent results. This fact constitutes a major obstacle that prevents effective AI use as a unified decision support system across public sector organizations.

AI models displayed significant differences which affected both their accuracy and reliability during testing. ChatSonic outperformed others, delivering the most accurate and reliable data, establishing its clear superiority in this domain. Claude secured the second position with its moderate reliability, yet it fell short of ChatSonic's performance. In contrast, ChatGPT and Perplexity provided far less accurate and reliable results. The models contained errors, yet ChatSonic and Claude demonstrated their potential to advance AI technology through their relative strengths.

Future research should investigate hybrid modeling methods which combine machine learning with traditional rule-based systems in order to improve AI model performance for public finance applications. The existing fiscal databases should be transformed into open-access databases which contain additional metadata standards that allow machines to read the data. The implementation of automated and manual audit systems will help organizations to identify and eliminate systematic mistakes. The computer science field must work together with public finance professionals to create solutions that reduce algorithmic bias while improving understanding of contextual information.

Model testing has demonstrated that the models produce accurate results which change according to different time periods. AI systems show better accuracy when they present data from previous years, but they experience higher error rates for current date information. The occurrence of this event happens because historical sources maintain a consistent structure which results in their repetition across different instances. On one hand, such data creates a pretty simple learning framework for AI models. The processing performance of AI models becomes harder to handle because newer data requires processing of changing and unpredictable information. The restricted access of AI models to fresh information causes present datasets to show higher accuracy deficits. Dynamic and variable data handling capabilities require further AI technology improvements to enable better operational performance.

The AI systems deliver budget information through rounded numerical values, which results in budgetary discrepancies that reach multi-million levels. The research shows that AI systems cannot deliver correct information because their output generates major financial risks during economic assessments and business choices. AI systems need better accuracy for processing sensitive materials that include budget documentation.

The spending data results in better accurate processing than the receipt data because AI handles structured data more effectively than unstructured data, which leads to revenue errors that occurred because of system complexities and limited available resources. The process requires AI systems to receive specific data type training, which will enable them to achieve accurate results that meet public finance standards.

Non-professionals who include students and amateur researchers face major difficulties because AI-generated public budget data contains multiple inaccuracies. Misinformation may compromise the trustworthiness, thus enforcing distorted views on public finance. The faulty data caused public trust to decline, which resulted in society developing distorted views while academic research and educational standards faced direct threats from the data. The challenges related to AI reliability extend beyond their technical aspects because they create major effects on society.

Confronting these challenges is crucial in light of AI's ubiquity in our daily life, whether academic, professional, or personal. Sensitive topics such as public budgets can mislead people because of inaccurate data, thus resulting in flawed preconceptions and conclusions. As technology democratizes, resolving these issues through technological advancement and a user-centric approach is necessary to provide AI with more powerful leverage in sensitive fields such as public finance.

First, the diversity of data sources utilized by AI models must be enhanced, and the datasets must be kept up to date. One of the primary reasons behind the accuracy issues observed in current models is the limited scope of the datasets and their insufficient inclusion of current data. Integrating broader, more diverse, and reliable data pools into AI systems could improve accuracy rates and enhance their ability to work with dynamic and up-to-date datasets. In this context, incorporating open data resources provided by public institutions into AI development processes should be encouraged.

Second, this could also be achieved by introducing some sort of verification mechanism for budget data provided by AI, as has already become common on most social media platforms. For example, verified information could include a validation badge in AI-generated outputs to allow users easier access to reliable data. A system of this sort would ensure not only reduced the capacity of the AI for misinformation but also quality decision-making in sensitive areas like public finance. In this verification process, the role of human experts and independent auditing mechanisms would be of essence.

Human-centered monitoring of the accuracy of data generated by AI is essential, especially in sensitive areas like public finance. Expert review prevents misinformation and improves the models of AI. The current study emphasizes verification processes like these as the backbone for further mechanisms of oversight.

Third, comprehensive educational programs should be introduced to improve AI literacy. As students, non-experts, and journalists increasingly rely on AI, it is vital they learn to verify information and understand its limitations. In complex areas like public finance, educational efforts across schools and public institutions are needed to ensure accurate interpretation of AI outputs. This would help society use AI responsibly and effectively.

The enhancement in AI's ability to provide more accurate and reliable information will most definitely increase transparency and accountability in public finance. Improved systems in AI could thus enable effective oversight in budgeting, which would allow citizens to better understand and question public spending through AI-supported analyses. AI could provide the exact data that would fast-track budget monitoring for the fair and efficient use of the resources. Therefore, AI's potential should be judged not only as a technological tool but also as a driver of democratic progress.

CONCLUSION

This study provides a comprehensive assessment of the operational boundaries and transformative potential of AI within the specialized domain of public fiscal transparency. The empirical analysis conducted on the 2009–2023 Turkish budget data reveals that while certain AI models demonstrate a promising capacity for data retrieval, significant architectural and reliability limitations persist across the current LLM landscape. A critical finding of this research is the inverse correlation between data recency and model accuracy; as fiscal data approaches the present day, the reliability of AI outputs diminishes markedly. This phenomenon underscores the current inability of AI to serve as an autonomous source for real-time public accounting, as newer data requires a level of precision that existing models, particularly those prone to hallucination, fail to maintain. The frequent occurrence of rounding discrepancies, which may seem minor in a linguistic context but translate into millions of monetary units in a fiscal context, further highlights the precarious nature of relying on unverified AI outputs for budgetary analysis.

The implications of these findings extend beyond technical error margins and intersect with broader concerns regarding democratic accountability and institutional trust. The observed tendency of models to process expenditure data more accurately than revenue streams suggests a structural bias in AI training sets or a higher complexity in the presentation of public income variables. For non-professional users, such as students, journalists, or civil society researchers, these inaccuracies risk creating a "counter-transparency" effect, where the ease of AI accessibility facilitates the spread of granular but false fiscal information. Consequently, the integration of AI into public finance should not be viewed as a mere technological upgrade but as a multifaceted challenge requiring rigorous verification frameworks. The adoption of "human-in-the-loop" oversight and the necessity of cross-referencing AI-generated data with primary official sources remain indispensable to safeguarding the integrity of the social contract between the state and its citizens in the digital age.

Looking toward the future, the role of AI in public finance is poised to transition from a simple data retrieval tool to a sophisticated predictive and analytical engine. As models evolve to better handle dynamic and multi-layered datasets, they could eventually provide real-time budget monitoring and automated fraud detection, significantly reducing the labor-intensive nature of fiscal oversight. However, for this potential to be realized, a paradigm shift in public data infrastructure is required. Future implementations must prioritize the development of specialized, domain-specific AI models that are fine-tuned on verified governmental datasets rather than general-purpose web data. By bridging the gap between advanced algorithmic processing and standardized open-data practices, AI can eventually become a catalyst for a more participatory and efficient fiscal management system. Ultimately, the successful deployment of these technologies will depend on creating a synergy between technological innovation, ethical governance, and expert-led auditing, ensuring that AI serves as a reliable driver of transparency rather than a source of fiscal misinformation.

REFERENCES

Acemoglu, D., and Restrepo, P. (2018a). *Artificial Intelligence, Automation and Work.* NBER Working Papers 24196. doi:10.3386/w24196

Acemoglu, D., and Restrepo, P. (2018b). Modeling Automation. *AEA Papers and Proceedings*, 48–53. doi:10.1257/pandp.20181020

Acemoglu, D., and Restrepo, P. (2018c). The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review, 108*(6), 1488–1542. doi:10.1257/aer.20160696

Acemoglu, D., and Restrepo, P. (2020). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy, 128*(6), 2188-2244. doi:10.1086/705716

Aghion, P., Antonin, C., and Bunel, S. (2019). Artificial intelligence, growth and employment: The role of policy. *Economics and Statistics, 510-512*, 149 - 164. doi:10.24187/ecostat.2019.510t.1994

Aghion, P., Jones, B. F., and Jones, C. I. (2017). *Artificial intelligence and economic growth.* National Bureau of Economic Research Working Paper 23928.

Agrawal, A., Gans, J., and Goldfarb, A. (2019). Economic policy for artificial intelligence. *Innovation Policy and the Economy, 19*, 139-159. doi:10.1086/699935

Agrawal, A., Gans, J., and Goldfarb, A. (2019). *Prediction, Judgment, and Complexity A Theory of Decision- Making and Artificial Intelligence.* National Bureau of Economic Research - The Economics of Artificial Intelligence: An Agenda.

Aldemir, C., and Uçma Uysal, T. (2025). Artificial Intelligence for Financial Accountability and Governance in the Public Sector: Strategic Opportunities and Challenges. *Administrative Sciences, 15*(2), 1-19. doi: 10.3390/admsci15020058

Anastasopoulos, L. J., Moldogaziev, T. T., and Scott, T. (2020). Organizational context and budget orientations: a computational text analysis. *International Public Management Journal, 23*(2), 292-313. doi:10.1080/10967494.2019.1706677

Autor, D., and Salomons, A. (2017). Robocalypse Now–Does Productivity Growth Threaten Employment? *Investment and growth in advanced economies,* 45-118. ECB Forum on Central Banking.

Aydın, M. S. (2024). Kamu Hizmetlerinin Sunumunda Yapay Zeka Kullanımı. *International Journal of Social, Political and Financial Researches, 4*(2), 171-186. doi:10.70101/ussmad.1520892

Bader, J., Edwards, J., Harris-Jones, C., and Hannaford, D. (1988). Practical engineering of knowledge-based systems. *Information and Software Technology, 30*(5), 266-277. doi:10.1016/0950-5849(88)90019-5

Chu, Y.-H., Hsu, S.-M., Chang, C.-C., and Hsu, S.-H. (2020). An Economy-Wide Assessment of Artificial Intelligence Investment on Manufacturing: A Case Study of Taiwan's Semiconductor and ICT Industries. *Modern Economy, 11*(5), 1040-1052. doi:10.4236/me.2020.115078

Corvalán, J. G. (2018). Digital and Intelligent Public Administration: transformations in the era of artificial intelligence. *A&C - Administrative & Constitutional Law Review, 18*(71), 55-87. doi:10.21056/aec.v18i71.857

Duan, Y., Edwards, J. S., and Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *International Journal of Information Management, 48*, 63-71. doi:10.1016/j.ijinfomgt.2019.01.021

Edwards, J. S., Duan, Y., and Robins, P. C. (2000). An analysis of expert systems for business decision making at different levels and in different roles. *European Journal of Information Systems, 9*(1), 36-46. doi:10.1057/palgrave.ejis.3000344

Eisikovits, N., Johnson, W. C., and Markelevich, A. J. (2024). Should Accountants be Afraid of AI? Risks and Opportunities of Incorporating Artificial Intelligence into Accounting and Auditing. *Accounting Horizons*, 1-7. doi:10.2139/ssrn.4748690

Engin, Z., and Treleaven, P. (2019). Algorithmic Government: Automating Public Services and Supporting Civil Servants in using Data Science Technologies. *The Computer Journal, 62*(3), 448–460. doi:10.1093/comjnl/bxy082

Fernandez-Cortez, V., Valle-Cruz, D., and Gil-Garcia, J. (2020). Can Artificial Intelligence Help Optimize the Public Budgeting Process? Lessons about Smartness and Public Value from the Mexican Federal Government. *Seventh International Conference on eDemocracy & eGovernment (ICEDEG)*. 312-315. IEEE. doi:10.1109/ICEDEG48599.2020.9096745

Gil-Garcia, J. R., Helbig, N., and Ojo, A. (2014). Being smart: Emerging technologies and innovation in the public sector. *Government Information Quarterly, 31*(1), 11-18. doi:https://doi.org/10.1016/j.giq.2014.09.001

Gillham, J., Rimmington, L., Dance, H., Verweij, G., Rao, A., Roberts, K. B., and Paich, M. (2017). *The macroeconomic impact of artificial intelligence.* PWC.

Gries, T., and Naudé, W. (2018). *Artificial Intelligence, Jobs, Inequality and Productivity: Does Aggregate Demand Matter?* IZA Discussion Papers No: 12005.

Gries, T., and Naudé, W. (2020). *Artificial Intelligence, Income Distribution and Economic Growth.* IZA Discussion Paper No. 13606. doi:10.2139/ssrn.3679012

Harsh, A., and Ichalkaranje, N. (2014). Transforming e-Government to Smart Government: A South Australian Perspective. *Intelligent Computing, Communication and Devices*, 9-16.

Hogan-Doran, D. (2017). Computer says "no": Automation, algorithms and artificial intelligence in Government decision-making. *Judicial Review: Selected Conference Papers: Journal of the Judicial Commission of New South Wales, 13*(3), 345-382.

Hollander, A. S., and Icerman, R. (1991). Capital budgeting in governments: The feasibility of Artificial Intelligence technology. *Expert Systems with Applications, 3*(1), 109-116. doi:10.1016/0957-4174(91)90091-R

Jimeno, J. F. (2019). Fewer babies and more robots: economic growth in a new era of demographic and technological changes. *SERIEs, 10*, 93–114. doi:10.1007/s13209-019-0190-z

Kassani, S. H., Kassani, P. H., Wesolowski, M. j., and Schneider, K. A. (2021). Automatic Detection of Coronavirus Disease (COVID-19) in X-ray and CT Images: A Machine Learning Based Approach. *Biocybernetics and Biomedical Engineering, 41*(3), 867–879. doi:10.1016/j.bbe.2021.05.013

Kolkman, D. (2020). The usefulness of algorithmic models in policy making. *Government Information Quarterly, 37*(3), 101488. doi:10.1016/j.giq.2020.101488

Lee, M., Hayes, D., and Maher, C. S. (2024). AI as a Budgeting Tool: Panacea or Pandora's Box? *Public Finance Journal, 1*(1), 49–65. doi:10.59469/pfj.2024.6

Liu, H.-W., Lin, C.-F., and Chen, Y.-J. (2019). Beyond State v Loomis: artificial intelligence, government algorithmization and accountability. *International Journal of Law and Information Technology, 27*(2), 122–141. doi:10.1093/ijlit/eaz001

Liyanage, H., Liaw, S.-T., Jonnagaddala, J., Schreiber, R., Kuziemsky, C., Terry, A. L., and Lusignan, S. (2019). Artificial Intelligence in Primary Health Care: Perceptions, Issues, and Challenges. *Primary Health Care Informatics Working Group Contribution to the Yearbook of Medical Informatics*, 41 - 46. doi:10.1055/s-0039-1677901

Mellouli, S., Luna-Reyes, L. F., and Zhang, J. (2014). Smart government, citizen participation and open data. *Information Polity, 19*(1-2), 1-4. doi:10.3233/IP-140334

Nolan, P., and Slater, G. (2010). Visions of the future, the legacy of the past: Demystifying the weightless economy. *Labor History, 51*(1), 7-27. doi:10.1080/00236561003654677

Özdemir, H., and Yelboğa, A. (2025). Yapay Zekâ Destekli Denetimin Kamu Harcamalarındaki Rolü: Sayıştay Örneği. *Sayıştay Dergisi, 36*(139), 771-802. doi:https://doi.org/10.52836/sayistay.1793520

Phillips-Wren, G., and Jain, L. (2006). Artificial Intelligence for Decision Making. *Knowledge-Based Intelligent Information and Engineering Systems*, 531–536.

Puron-Cid, G. (2014). Factors for a successful adoption of budgetary transparency innovations: A questionnaire report of an open government initiative in Mexico. *Government Information Quarterly, 31*(1), 49-62. doi:10.1016/j.giq.2014.01.007

Santschi, D., Grau, M. C., Fehrenbacher, D. D., and Blohm, I. (2024). Artificial Intelligence to Improve Public Budgeting. *Forty-Fifth International Conference on Information Systems*. 1-9. ICIS.

Sun, T. Q., and Medaglia, R. (2019). Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly, 36*(2), 368-383. doi:10.1016/j.giq.2018.09.008

Torben Antretter, Blohm, I., Sirén, C., Grichnik, D., Malmström, M., and Wincent, J. (2020). Do Algorithms Make Better - and Fairer - Investments Than Angel Investors? *Harvard Business Review*.

Upchurch, M. (2018). Robots and AI at work: the prospects for singularity. *New Technology Work and Employment, 33*(1), 205-218. doi:10.1111/ntwe.12124

Valle-Cruz, D. (2019). Public value of e-government services through emerging technologies. *International Journal of Public Sector Management, 32*(5), 530-545. doi:10.1108/IJPSM-03-2018-0072

Valle-Cruz, D., Criado, J., Sandoval-Almazán, R., and Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly, 37*(4), 101509. doi:10.1016/j.giq.2020.101509

Valle-Cruz, D., Fernandez-Cortez, V., and Gil-Garcia, J. (2022). From E-budgeting to smart budgeting: Exploring the potential of artificial intelligence in government decision-making for resource allocation. *Government Information Quarterly, 39*(2), 101644. doi:10.1016/j.giq.2021.101644

Wirtz, B. W., Weyerer, J. C., and Geyer, C. (2018). Artificial Intelligence and the Public Sector—Applications and Challenges. *International Journal of Public Administration, 42*(7), 596-615. doi:10.1080/01900692.2018.1498103

Wirtz, B. W., Weyerer, J. C., and Sturm, B. J. (2020). The Dark Sides of Artificial Intelligence: An Integrated AI Governance Framework for Public Administration. *International Journal of Public Administration, 43*(9), 818-829. doi:10.1080/01900692.2020.1749851

Xu, M., David, J. M., and Kim, S. H. (2018). The fourth industrial revolution: Opportunities and challenges. *International Journal of Financial Research, 9*(2), 13194-13194. doi:10.5430/ijfr.v9n2p90

Yavuz, E., and Özgül, H. B. (2025). The Role of AI in Public Budget Processes: A Comparative Evaluation on National AI Strategies and Practical Examples. *Sayıştay Dergisi, 36*(138), 575-602. doi:10.52836/sayistay.1717763

Yetkin Ataer, M. (2025). Yapay Zeka ve Türkiye’de Kamu Maliyesinin Değişen Rolü. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 34*(Özel Sayı: Uygarlığın Dönüşümü: Yapay Zekâ), 341-354. doi:10.35379/cusosbil.1696963

Zhu, J., Zhang, J., and Feng, Y. (2022). Hard budget constraints and artificial intelligence technology. *Technological Forecasting and Social Change, 183*, 1-12. doi:10.1016/j.techfore.2022.121889