Organisation for Economic Co-operation and Development

GOV/SBO(2025)4

Unclassified **English text only**

1 April 2025

PUBLIC GOVERNANCE DIRECTORATE
COMMITTEE OF SENIOR BUDGET OFFICIALS

Using Artificial Intelligence in Public Financial Management

25th Annual Meeting of the Working Party on Financial Management and Reporting

13-14 March 2025

This paper explores how artificial intelligence (AI), including generative AI, can be applied in public financial management.

For further information, please contact:
Delphine MORETTI (delphine.moretti@oecd.org)

JT03563437

Table of contents

FIGURES

INFOGRAPHICS

BOXES

1 Artificial Intelligence and Public Financial Management[1]

1.1. What is Artificial Intelligence?

1. Artificial Intelligence (AI) is an interdisciplinary field of computer science that aims to design and train systems capable of performing tasks that would typically require human intelligence. To do this, AI relies on algorithms and computation models to process usually large sets of data to identify meaningful connections and patterns. It is defined by the OECD as "a machine-based system that, based on explicit or implicit objectives, processes inputs to produce outputs like predictions, content, recommendations, or decisions that can impact the physical or virtual environments." (OECD, 2024[1])

2. AI, a dynamic concept, manifests through a myriad of methodologies, technologies, and applications, evolving rapidly over time, making it challenging to delineate its landscape comprehensively. To gain a nuanced understanding of AI, it's therefore useful to contextualise it within three distinct categories of systems, forming a technological continuum:[2]

- At one end of the spectrum are the "transaction-based" systems that have been in use for a relatively long time. Such systems use specific inputs (generally structured data, known as "transactional") to generate an output according to a pre-determined rule. These systems respond to current situations with limited or no memory of past actions, functioning purely on the inputs they receive at any given moment.
- At the other, theoretical, end of the spectrum lies Artificial General Intelligence (AGI), envisioned as systems capable of understanding, learning, and performing intellectual tasks at a human level and beyond. True AGI does not yet exist and remains a conceptual goal in AI research, with significant debate in the scientific community about its feasibility.
- Between these extremes lie the systems that are at the heart of recent advancements, commonly designated as AI or Generative AI (GenAI). These systems are designed to manage a diverse array of inputs and generate original outputs in one or more delineated domains, such as speech recognition, face detection or image classification. Unlike transaction-based systems, AI can learn from training, interactions and adapt to new situations, thereby improving its performance over time. However, these systems do not match the human mind, which is far more versatile (OECD, 2019[2]).

[1] This paper was drafted by Delphine Moretti, Senior Policy Analyst in Public Management and Budgeting Division. It benefitted of conversations with and inputs from Nicolas Botton, OECD Consultant.

[2] Such categories of system are illustrative only. Concerning AI technologies, the OECD's Framework for the Classification of AI Systems supports a more detailed classification by identifying five different dimensions of AI systems. (OECD, 2022[27])

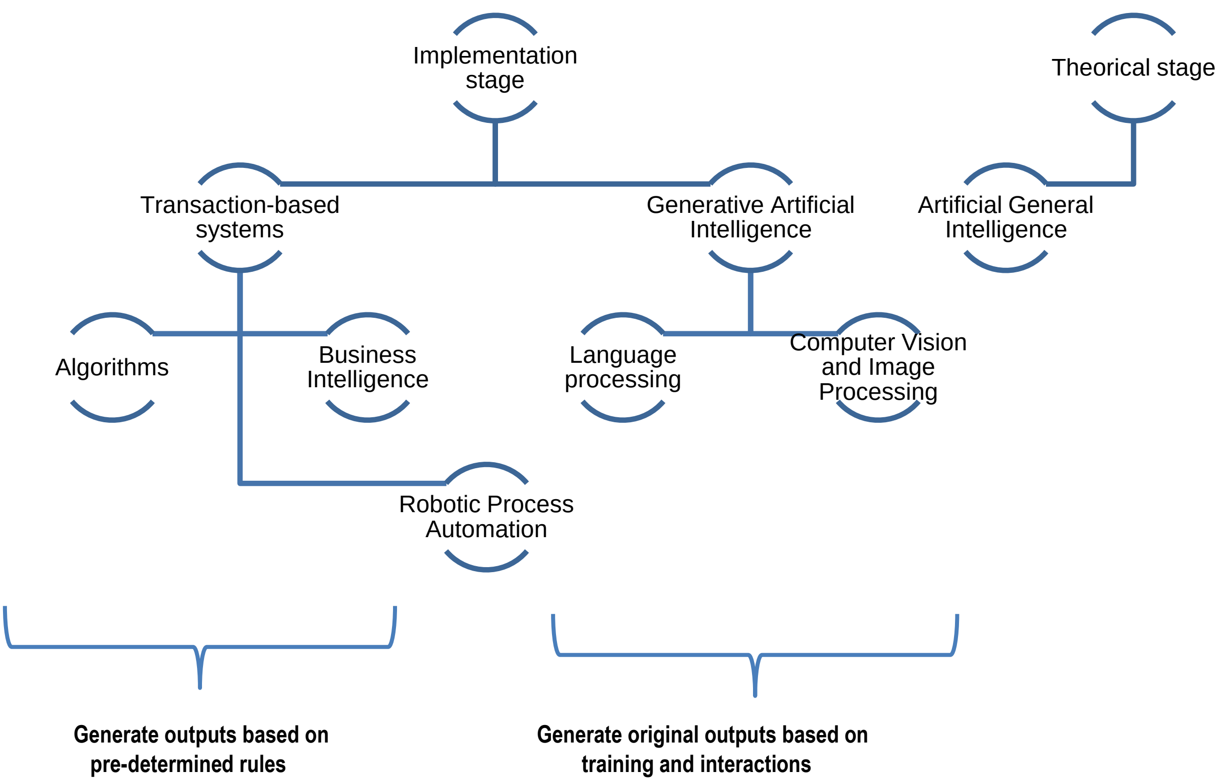
3. Transaction-based systems include:

- Algorithms that sort, search, or make decisions by processing inputs through a series of "if-then" statements. They are commonly used in environments such as medical diagnosis, legal compliance, or customer service operations.
- Business intelligence (BI) that collects, processes, and visualises data from various sources to uncover insights, trends, and patterns that can drive strategic and operational improvements within an organisation.
- Robotic Process Automation (RPA) systems that automate routine and repetitive tasks by using software robots ("bots"). They are commonly used for streamlining data entry, processing transactions, and managing records.

4. The core technology behind most AI systems is machine learning (ML), especially deep learning (DL) and reinforcement learning using big data and computing resources. AI systems include:

- Natural Language Processing (NLP) systems providing content translation, text classification, speech recognition, writing aids, and chatbots. They are commonly used for conversational interfaces like Google Assistant, Siri, or Alexa.
- Computer Vision and Image Processing systems using deep learning for recognition, classification, conversion, and other tasks. They are used for instance in smartphones security, video surveillance or self-driving cars.

Infographic 1. Technological continuum



1.2. Is Artificial Intelligence a Revolution for Public Financial Management?

5. AI is often presented as a revolutionary force capable of radically transforming economies, industries, and labour markets (OECD, 2023[3]). While the potential for such transformations is undeniable, its actual pace and extent varies significantly from one sector to another, and even among different tasks within the same sector:

- In some sectors, AI is already leading to breakthroughs. This is for instance the case in healthcare, environmental management, and retail services, where they enable real-time insights and responses, as well as productivity and efficiency gains that were previously unattainable.
- In other sectors, AI is used only in a limited way, due to reasons that may include lack of appropriate information technology infrastructure, issues with data, or security/data privacy issues. For instance, cases of GenAI in financial markets remain largely at the development phase, if any (OECD, 2023[4]).
- Finally, AI is sometimes not so much a revolution as it is a continuum of technological evolution from data, big data to data analytics. Where these technologies, which form the building block on which AI can be applied, are already used, the introduction of AI is mainly an enhancement and deepening of existing capabilities – e.g., with new layers of technologies being integrated in existing algorithms.

6. This last pattern is particularly predominant in the field of PFM. Technology has been increasingly integrated in the PFM cycle over the last decades, through the adoption of increasingly sophisticated Financial Management Information Systems (FMIS). Over time, and with quality of data increasing, new layers of technology have been added: data analytics, BI tools, RPA and more recently AI technologies. This has been for instance the case with South Korea's FMIS, called dBrain+, in which new layers of AI technologies are being added (Korea Fiscal Information Service (KFIS), 2023[5]).

2 Main trends of use

2.1. Macroeconomic and macrofiscal forecasting

7. Forecasting is a foundational element for effective PFM, as it provides the basis for predicting future economic conditions and public revenue streams, which in turn allows reliable budget planning. The traditional tools for forecasting within PFM include statistical and econometric models based on time-series data, incorporating variables that influence fiscal outcomes like GDP growth rates, tax revenue collections, and government spending patterns.

8. Establishing the conditions for the production of accurate forecasts has been a focus of many OECD countries, through improvement of statistical and econometric models, establishment of independent fiscal institutions, and greater disclosures of the assumptions underlying the models (Moretti, Keller and Majercak, 2023[6]). Government forecasting has however been put to test in the last decade by a series of economic shocks, from the global financial crisis to the COVID-19 pandemic and energy and cost-of-living crises, which exposed their limits.

9. AI could help address challenges with traditional forecasting methods by enhancing the accuracy and timeliness of predictions. Indeed, AI technologies, particularly ML and DL, can process large volumes of data, including unstructured data, and identify complex patterns that are not easily detectable with existing methods. In fact, a study published in 2018 (Jung, Patnam and Ter-Martirosyan, 2018[7]) found that ML-based models could outperform traditional economic forecasting models.

10. Additionally, ML's predictive capacities are also seen as an opportunity to develop nowcasting, which is the forecasting of the very recent past, the present, and the very near future state of economic indicators that are typically only determined after a delay and are subject to revision, such as gross domestic product (GDP) or inflation. A recent academic study found for instance that ML models were able to provide more accurate and anticipated predictions than traditional time series models, even when using unstructured data (Tenorio and Perez, 2023[8])

11. However, AI-based models that have the best forecasting outputs have been criticised for their "black-box nature" (Jung, Patnam and Ter-Martirosyan, 2018[7]), as their assumptions and calculations cannot be fully understood. They therefore represent both a step forward in accuracy and a step back in fiscal transparency. This leads to a current focus of governments on using analytics and AI to improve human modelling and sensitivity analysis rather than replace it, and to support testing of hypotheses and longstanding assumptions.

12. Governments also aim for developing methods to "unbox" AI models and making their reasoning more transparent - an essential condition for the usability of these forecasts in PFM. For instance, the Swedish National Financial Management Authority (ESV) has worked on developing an application for analysing the impact that each data variable has on the prediction of the "black-box models", as part a wider work programme on as building grounds for integrating AI in the financial management of Swedish government (Boström et al., 2020[9]).

2.2. Spending decisions

13. In an environment of high debt levels and competing demands on limited resources, increasing "the bang for the buck" of public spending is a common objective for governments and finance ministries that are under a constant pressure do more (e.g., on climate, security, health or education) with less. Actions commonly explored or already taken by finance ministries include cutting unnecessary costs, eliminating wasteful spending, including through better targeting expenditures to those in need.

14. It is therefore unsurprising that expenditure targeting is one of the areas in which finance ministries had already started exploring the potential of new technologies a few years ago, by leveraging in particular big data, data analytics and data visualisation technologies. Big data allows for the analysis of vast quantities of information from various sources, enabling more comprehensive analysis. Meanwhile, data analytics tools provide the capability to drill down into specific spending categories and beneficiary groups, evaluating the effectiveness of each expenditure based on trends and patterns. Finally, data visualisations allows effective communication of complex information through graphical representations. Specifically, data visualisation can simplify the interpretation of budgetary trends and comparisons, thus facilitating more informed decision-making.

15. For instance, during the mid-2010s, the Australian Government made a strategic move to enhance its data analytics capabilities, aiming to improve the accuracy and efficiency of expenditure targeting and costing (Australian Government Department of Finance, 2018[10]). This initiative introduced a new approach to subsidy targeting and costing, employing techniques that increased precision in policy costing. This improvement was allowed by advancements in computer processing power, sophisticated software, reduced data storage costs, and more recently a new framework for data sharing, and relied on big data, data analytics and data visualisation technologies.

16. ML and DL, which have the capacity to use very large datasets to identify trends and patterns, as well as group data points based on similarity or shared characteristics, offer opportunities to fasten and improve analyses previously done by humans. The capacity of ML and DL to work with unstructured data may also open opportunities for combining datasets previously unused for such exercises.3 For instance, South Korea's dBrain+ includes modules called Korea Fiscal Information System (KFIS) and Korea Risk Assessment and Horizon Scanning, or KORAHS, which collect a large range of economic, fiscal and financial data and use AI to analyse this data in real time to identify risks and support with numerical indicators optimal decision-making on fiscal policies and public spending (Korea Fiscal Information Service (KFIS), 2023[5]).

2.3. Budget planning and monitoring

17. Budget planning is the formulation of realistic budgets based on accurate expenditures projections. Budget monitoring is the continuous assessment of how government resources are used relative to these projections. Budget planning and monitoring involves preparing expenditure baselines, costing new policies, as well as tracking spending against allocations to monitor possible deviations, which are all key and time-consuming functions of central budgetary authorities.

18. AI, and more specifically ML, has the capacity to support these processes by providing outputs that support the formulation of accurate expenditure baselines and costing of new policies. For instance,

---

[3] Importantly, academic research shows that AI could help targeting spending in a way that is compatible multiple, sometimes perceived as conflicting by humas, policy objectives, as well as could support a more dynamic allocation of allocations in case of unexpected economic developments although use cases have not been identified yet. (Valle-Cruz, Fernandez-Cortez and Gil-Garcia, 2022[28])

in Australia, the Department of Veteran's Affairs developed predictive models and tools to help model future financial impacts of the policy decisions. One of the models used is the Priority Investment Approach—Veterans (PIA-V) model, which is an AI-powered actuarial model that simulates the lifetime financial trajectory of beneficiaries. Outputs generated by the models include annual fiscal expenditure for each beneficiary as well as their average years on benefits, which are used for costings, budget estimates, beneficiaries' segmentation and policy evaluation (Australian Government, 2019-2020[11]).

19. Another promising domain of application for ML is the identification, monitoring and mitigation of fiscal risks through the analysis of large data sets. For instance, in the case of subnational governments, fiscal risks can arise due to a variety of causes, including for instance unsustainable spending or investment levels, which need to be identified sufficiently early to take preventive action. AI, and more specifically ML, can support the identification of such risk. Examples of the use AI-powered monitoring of subnational government-related fiscal risks can be found in:

- France, where the Public Finances General Directorate (DGFiP) incorporated AI features into its pre-existing data analytics-based system of monitoring and advice on subnational government finances (Box 1). This system uses the structured financial data collected in the shared and centrally managed FMIS of the French Government, called CHORUS.
- Indonesia, where an AI-powered system called Artificial Intelligence Financial Advisor, or AIFA, uses ML and word-cloud to process unstructured financial and performance data to provide analytics on subnational governments' fiscal performance in real time. It also includes a fraud detection model using Benford's Law (Wisesa, 2023[12]).

Box 1. France: Use of AI in budget monitoring

DGFiP has implemented for several years a "warning system", whose objective is to i) identify municipalities with financial difficulties; ii) provide them with financial advice and iii) proactively support implementation of corrective measures.

This warning system was based initially on an algorithm using historical tax and financial data to score municipalities. More recently, DGFiP developed a predictive model designed for the earlier identification of municipalities' financial difficulties. The model was trained on data spanning four years, to predict outcomes for the fifth year based on ML technology. The predictive model also relies on unsupervised clustering technology to categorise municipalities with similar financial characteristics without predefined outcome examples.

In 2022, an experiment with the model covered 2,500 municipalities, with around 40% of these identified as facing financial difficulties. Of these around 17% had not been detected by the previous algorithm. Additionally, around 35% of municipalities were identified with temporary, non-structural difficulties, highlighting the model's capacity to differentiate between permanent and transient financial problems.

Source: Authors based on (French Public Finances General Directorate (DGFiP), 2024[13])

2.4. Financial Management

20. Financial management and reporting functions in government are to ensure that spending is executed in compliance with relevant laws and regulations, as well as processed in a timely and efficient manner. These functions involve managing contracts and invoices, disbursing funds after checks and

controls, keeping detailed and systematic records of all financial transactions, and providing clear and accessible information on these to external stakeholders in fiscal reports.

21. Financial management and reporting functions in government involves important but sometimes repetitive tasks that are particularly well-suited for automation. In this area, the functionalities offered by ML, DL and NPL are particularly relevant: these technologies can be used to analyse digital images to extract information from documents (e.g., vendor information); to identify and classify documents (e.g., invoices); perform document comparison (e.g., compare invoice and vendor information); or to identify trends and patterns (e.g., internal controls on payment requests).

22. Uses cases already abound. For instance:

- In France, the DGFiP has developed an AI-based tool as part of the regular internal control processes that "automatises the selection of payment requests to be controlled (and) optimises the workload and the quality of controls performed" (French Public Finances General Directorate (DGFiP), 2024[14]).
- In Brazil, where the National Treasury is using AI to classify subnational governments' expenditures. In the past, data was converted to the relevant standard manually – a time-consuming process which was prone to errors. With the move to an AI model, classification time was reduced from 1 000 to 8 hours, with a model accuracy of 97% (Rocha, Porto and Fabel, 2024[15]).
- In Finland, which has implemented RPA and AI to enhance its financial management and reporting functions within the Finnish Government Shared Services Centre for Finance and HR (Palkeet) in a large range of financial management activities (Box 2).

Box 2. Finland: Use of RPA and AI in financial management

Palkeet' s project on RPA and AI began with a comprehensive feasibility study to identify potential benefits and suitable uses for automation within financial management and HR processes. This study led to a structured project that included hands-on learning in collaboration with selected vendors, which helped build internal expertise.

Subsequently, Palkeet established an in-house Center of Excellence for RPA, focusing on developing and deploying RPA solutions across various financial and HR functions. This initiative was geared towards automating routine tasks such as the management of supplier information, balancing of accounting data, and processing of financial transactions. The deployment of RPA was planned to be scalable and was integrated with AI where complex decision-making or data processing was necessary, such as in the analysis and routing of purchase invoices.

Palkeet identifies three stages in its automation journey, corresponding to increasingly ambitious objectives and new layers of technologies:

- First, automation of single tasks using their FMIS and other simple tools such as Excel.
- Second, from 2016, automation of sequential tasks by adding RPA and Power-BI tools.
- Third, from 2022, end-to-end process automation in some areas with new layers of ML tools ("hyper-automation phase").

Palkeet currently identifies more than 200 detailed uses of automation, including verifying numbers on expense receipts and the content requirements; routing expense receipts; managing supplier information; verifying and balancing expense receipts; balancing accounting data; verifying balance sheet accounts after accounting period; monitoring travel expense bill; or automating the payroll processing.

Source: Authors based on (Finish Agency for financial management and HR (Palkeet), 2024[16])

23. In addition to regular controls, the development of targeted verifications to identify errors (e.g., improper payments in emergency situations) and frauds (e.g., identity theft) have become a major objective for many governments in the wake of the COVID-19 pandemic, which exposed vulnerabilities in several countries' payment systems. ML and DL capacity for identifying trends and patterns can be leveraged to help with this, as illustrated by an increasing number of use cases:

- In Denmark, during COVID-19, the Danish Business Authority developed AI-based controls for aid-applications from companies for the different support schemes (van Noordt and Tangi, 2023[17]). More recently, AI-based control procedures have been tested for subsidy payment schemes to identify abnormal patterns.
- In the United States, the Treasury Department has developed and started implementing processes to detect and prevent fraud in near real-time using AI (United States Department of the Treasury, 2024[18])
- In the United Kingdom, the Department of Work and Pensions is using advanced analytics including ML to identify patterns in claims that could suggest fraud or error, so that these claims can be reviewed by relevant teams within the department (United Kingdom National Audit Office, 2023[19]).

24. Despite capacities of AI to summarise or draft text using NLP, finance ministries seem however to have been cautious in rolling out new technologies in the area of fiscal reporting – e.g., the automatic production of fiscal reports. This cautious approach may stem from the need to ensure compliance of reports with a variety of standards and uncertainties as to whether current technologies would allow such full compliance, as well as what automatisation and use of AI would imply in terms of responsibilities vis-a-vis external auditors and users of the reports, notably the legislature.

2.5. Engagement with external stakeholders

25. NLP-powered chatbots are among the most recognised and frequently utilised AI technologies today. These digital assistants are increasingly employed in the public sector to directly provide services, helping citizens fulfil legal requirements and access desired information. For example, tax office officials, who traditionally engage in regular face-to-face interactions with the public due to the nature of their work, are now supported by chatbots. Several tax administrations are developing these AI tools to field questions on tax legislation, reflecting a shift towards more efficient, automated public interactions.

26. Conversely, officials involved in government spending usually work behind the scenes, with their public interactions traditionally mediated through official communications and reports rather than direct contact. Thus, the adoption of chatbots in these areas has been notably slower. However, a few countries are beginning to explore AI's potential to foster more direct, seamless and effective communication and exchanges with citizens, as well as other key stakeholders including parliamentarians:

- In Mexico, the government has introduced an AI virtual assistance tool as part of its Intelligent Support Platform, designed to guide users through government programs and supports (Government of Mexico, 2024[20]). It provides information on benefits, eligibility, and application processes for individuals, businesses, and local governments, utilising a simple keyword search function or personalised questionnaires to tailor the information to the user's profile. The data is sourced from units responsible for Budget Programs and is updated in accordance with fiscal regulations.
- In France, the Ministry of Finance is using a large language model for the processing of large number of parliamentary amendments over a very short period of time (French Public Finances General Directorate (DGFiP), 2024[14]). During parliamentary debates on the 2024 budget, the model was able to summarise 94% of 5,400 amendments from the National Assembly's first reading in 10 minutes, with an error rate between 5-10% This both improved speed of engagement with parliamentarians and time management and well-being of staff.

- In the United Arab Emirates, the Ministry of Finance has an ongoing initiative called “Virtual Mona”, which uses AI for financial data analysis and insights, as well as virtual and augmented reality to create interactive visualisations of this financial data. This makes it easier for the Ministry of Finance to share financial information with various stakeholders (both internally and externally) (Ministry of Finance, UAE, 2024[21]).
- In the United Kingdom, an AI tool called Parlex has been developed to provide an intelligent search of the Hansard reports, which document all the parliamentary debates for more than 200 years. It therefore gives government officials insight into policy debates and the views of specific MPs. It is envisaged that this will give officials more time to engage with the relevant parliamentarians (Padfield, 2025[22]).

3 Impact and lessons learned

3.1. An elusive impact

27. Ideally, in the PFM area, the impact of AI should be assessed by using evaluation frameworks to track costs of projects and key performance indicators from completed projects, including cost savings, effectiveness and efficiency gains, error reduction, and compliance enhancement. This would involve, among other things, monitoring full costs of projects, comparing metrics from before and after AI implementation, conducting stakeholder surveys for satisfaction, and analysing data to see how AI outcomes match with fiscal policy goals.

28. However, while many finance ministries or financial management agencies identify a potential for large-scale productivity gains from AI use in PFM, public studies on feasibility and costs are rare and impact assessments are often anecdotal. Evidence on AI projects' costs and impacts are either not collected or not publicly available due to the provisional nature of projects still in pilot or early development phases.

29. Confronted with this lack of data, external oversight bodies have recently called for greater transparency over AI projects across government emphasising that greater scrutiny and evidence collection is required alongside substantial investment. For instance, in the United States, the Government Accountability Office defined a framework of four complementary principles to help ensure responsible AI use by federal agencies and other entities involved in AI systems. The four principles include performance, and monitoring and call for metrics, performance and outputs of AI systems and their components to be better assessed (United States Government Accountability Office, 2023[23]).

30. In measuring impacts of AI, it is also important to discuss risks. Early uses of transaction-based and AI systems show that these risks can be significant, in particular due to potential biases in algorithms. For instance:

- In the Netherlands, data exchange and algorithmic analyses were used to identify potential fraud in childcare benefits to tackle benefit fraud and foster efficiency savings. The system however incorrectly flagged over 20,000 parents as fraudulent (Peeters and Widlak, 2023[24]).
- In Australia, "automated decision-making" was used to identify former and current income support recipients with alleged discrepancies between their earnings as reported to the Department of Human Services and their employment income annually reported by employers. The algorithm incorrectly identified more than half a million Australians, leading to the termination of the project (Royal Commission into the Robodebt Scheme, 2023[25]).

3.2. Early Insights

3.2.1. Matching PFM needs and AI technologies

31. Anecdotal evidence from OECD meetings suggests that the current and envisioned uses of AI in PFM are most commonly related to improving existing processes, rather than redesigning them (e.g., by removing any human intervention) or creating entirely new processes. AI seems to be viewed at the moment as most useful as an "assistant" that:

- Provides high speed and low transaction cost in automating small tasks that civil servants engage in numerous times during the day.
- Analyses historical or real-time data to forecast future events or behaviours to support civil servants' own analyses.

32. Finance ministries that have already engaged in AI projects in the PFM area virtually all emphasise the importance of mapping processes and functions to pinpoint areas of inefficiency and potential efficiency gains as a prerequisite to deploy AI. Once this initial phase is complete, the next step is to assess the suitability of individual AI technology for integration in the process. Broadly, experience shows that:

- Repetitive and time-consuming manual tasks can often be automated with RPA, potentially integrating successive layers of AI technologies.
- Tasks that involve the analysis of large quantities of data, in particular where data analytics technologies are already in use, are likely to be prime candidates for the application of advanced technologies like ML and DL.

33. To scale up their ambitions progressively and safely, finance ministries generally adopt a sequenced approach to the introduction of AI technologies. This is well illustrated by the case of Palkeet, which started with "small" uses of RPA and is now moving towards hyper-automation. According to Palkeet, starting their advanced digitalisation journey with RPA also helped increase the acceptability for civil servants of more sophisticated AI technologies (Finish Agency for financial management and HR (Palkeet), 2024[16])

3.2.2. Addressing IT and data issues

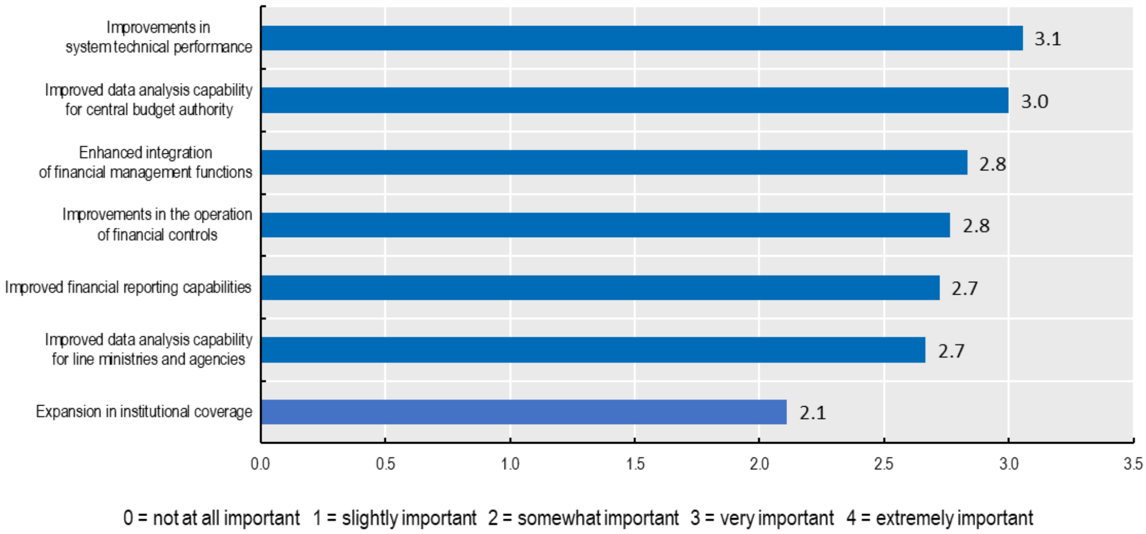
34. Information technology (IT) systems are crucial for finance ministries to be in a position to leverage AI opportunities. However, many OECD countries acknowledge finding themselves currently locked into legacy technologies that are significantly fragmented and often outdated and lack the necessary infrastructure and compatibility to integrate advanced AI functionalities. For example, centrally managed FMIS systems are more than 10 years old in a majority of OECD countries (OECD, 2022[26]), (Question 21).

35. As with any IT system, the quality of output obtained from an AI system depends on the quality of inputs. Anecdotal evidence from bilateral discussions with finance ministries also indicates that the lack of high-quality data, coupled with restrictions on data sharing, frequently impedes the initiation of AI projects. These challenges underscore the need for improved data management practices and policies that facilitate more effective data accessibility and sharing.

36. Accordingly, a large number of OECD countries plan major changes or upgrades to their system, with improvements to technical performance and development of data analytics capabilities as a primary objective for their reforms (Figure 1). They also acknowledge the need to establish appropriate data foundations to unlock the transformative benefits of AI, as large volumes of high-quality data are necessary to train, test, and deploy AI models effectively.

37. In doing so, several finance ministries underline that in-house capacities are being built to manage data analytics and GenAI projects following sometime challenging experiences in relying on external providers for IT systems and services.4 France's DGFiP, for example, has as its strategic objectives to "have our own tools and servers and re-internalize external intelligence to reach our ministry's objective of sovereignty." (French Public Finances General Directorate (DGFiP), 2024[14])

Figure 1. Objectives for FMIS upgrades in OECD countries, 2022



Note: Refers only to countries currently undertaking major development or replacements of their central FMIS (18 countries). Ratings present the average level of importance assigned to each objective on a scale of 0 to 4 by all respondents. Data for Chile, Colombia, Israel, Mexico, Slovenia and the United States are not available.
Source : (OECD, 2022[26]), Question 24

3.2.3. Prioritising time and PFM expertise in AI model training

38. In the case of most sophisticated AI technologies, GenAI including LLMs, the AI models need to be initially trained following three broad steps.5 In the context of PFM, which is a highly technical policy area, implementation of AI models necessarily require substantial training that involves human supervision (and feedback) from PFM experts.

39. Further, it is tempting to overestimate the capabilities of AI. AI systems produce results that are based on probabilities, which means they are by nature uncertain They can produce data that is wrong or text that sounds highly authoritative even when they "hallucinate," i.e., when the content is invented. Therefore, irrespective of how much training models may receive, PFM specialists need to be able to exert critical judgment when using the outputs they generate. Such oversight requires specialists that combine technical PFM skills and understanding of the basics of AI how works.

[4] In a recent OECD survey, commercial off-the-shelves technology capacity to reflect all relevant business processes within government and the performance of the COTS service provider were amongst the top five challenges identified by respondents in relation to their FMIS projects. (OECD, Question 25[22])

[5] The first steps is when the model is fed vast quantities of data and learns the structure of this training data. The second step aims to make the model learn how to respond to the user's instructions. The third step is when human supervisors provide the model with feedback.

40. This underscores the need for continuous recruitment, upskilling and training programmes to accompany both the development and use of AI technologies in PFM. In a number of OECD countries, finance ministries have built-up dedicated teams to manage both analytics and AI projects. These teams often also have a responsibility for developing wider knowledge of new technologies in the ministry. For example, in France, a Digital Transformation Unit (DTNum) has been established within the DGFiP, whose tasks include training staff on IT (20%) and change management (80%, comprising a transformation academy, certifications, communication activities) (French Public Finances General Directorate (DGFiP), 2024[14]).

3.2.4. Developing Government-wide standards and community of practice

41. In many OECD countries, the establishment of government-wide frameworks or standards for the responsible and safe adoption of AI is still an ongoing process. Such frameworks are essential to support AI applications effectively in PFM. They should:

- Build a strong legal and regulatory framework that encompasses oversight methods, security measures, privacy rules, validation processes, and adherence to standards to guarantee the ethical and responsible utilisation of AI systems.
- Develop explicit guidelines that incorporate best practices for accountability, transparency, and effective AI deployment.
- Actively tackle challenges related to the management of sensitive information.
- Identify where AI can deliver substantial benefits, thereby preventing the wasteful use of resources on technologies that fail to enhance traditional approaches.

42. Moreover, as many finance ministries engage in pilot projects and lead or sponsor programs to boost AI development and adoption in line ministries, these efforts need a co-ordinated approach and should integrate lessons from other governmental projects, which could greatly minimize risks of failure and enhance the quality of outcomes. For instance, a “community of practice” approach to AI projects can help prevent common and avoidable mistakes and mitigating risks from inherently complex AI projects, stemming for instance from the involvement of multiple stakeholders with diverse interests or lack of understanding of technologies and systems.

4 Conclusion and policy implications

43. Experience with AI in PFM currently demonstrate both undeniable potential but also challenges. While AI offers significant possibilities for bringing PFM into a new digitalisation era, finance ministries generally take an (understandable) cautious approach in developing projects with most focusing on "task automation" and "predictive" rather than "prescriptive" applications of AI:

- Task Automation AI is the use of AI to streamline processes, reduce operational costs, and increase productivity by handling tasks like data entry, appointment scheduling, or form processing efficiently and with minimal error.
- Predictive AI is the use of AI systems to analyse historical and real-time data to forecast future events or behaviours – e.g., to forecast spending or anticipate fiscal risks. These applications help to understand likely future scenarios based on existing data, thereby enhancing finance ministries' ability to prepare and plan.
- Prescriptive AI would go a step further by not only forecasting outcomes but also suggesting courses of action to achieve desired goals or mitigate risks. This could involve AI systems that for instance suggest optimal budget allocations.

44. From the implementation of task automation and predictive AI in PFM, some initial policy implications can be drawn:

- First, addressing the technological infrastructure within finance ministries is vital. Many OECD countries face the challenge of outdated and fragmented IT systems that are ill-equipped to support advanced AI functionalities. To this end, system upgrades and the development of new data analytics capabilities must be prioritised. Moreover, the emphasis on building in-house capacities, as seen in France's strategic objective to "re-internalise", highlights the trend towards greater data sovereignty and the reduction of dependency on external vendors.
- Second, there is an urgent need for continuous professional development and training in AI and data analytics for PFM. Finance ministries must invest in upskilling all levels of their workforce in data literacy; including to handle sophisticated AI applications. This is a challenge that is currently a focus of the data analytics branch of the Australian Department of Finance. This approach not only enhances technical skills but also ensures that personnel are equipped to manage and interpret AI outputs critically.
- Third, establishing robust project selection and evaluation frameworks within finance ministries is critical, and should be aligned with government-wide frameworks where these exist. Such frameworks should track not only cost and performance indicators but also the qualitative impacts of AI, such as error reduction and compliance enhancement. This is needed as multiple projects will likely compete in the future for limited investment resources.
- Fourth, developing government-wide AI standards and guidance for project implementation is essential for the integration of AI technologies in PFM. AI standards should cover key areas for safe implementation of AI, including issues in relation to data exchange, privacy protection, avoidance of intended or unintended bias in models or cybersecurity risks.

45. More broadly, in particular if applications of prescriptive AI in PFM start being more systematically explored, PFM specialists must start discussing the profound implications of AI on the delicate balance of

roles and responsibilities within public finance systems. This includes considering how AI could alter the letter and spirit of PFM institutions and processes and reshape the functions of fiscal stakeholders including external oversight bodies such as supreme audit institutions and independent fiscal institutions. Therefore, as AI technologies can automate fiscal decision-making processes, it is crucial to identify emerging questions regarding the future of transparency and accountability:

- How will accountability be defined if, or when, decisions are made by algorithms? The automation of fiscal decisions could shift the concept of accountability from human judgment to system-based outcomes. This change necessitates a re-evaluation of how responsibilities are assigned and how accountability in fiscal decisions is measured in an automated environment.
- What mechanisms will be needed to ensure transparency in automated decisions? As forecasting and budget planning and monitoring could increasingly be made by AI, there must be robust frameworks to ensure transparency. This involves creating and implementing standards that govern AI processes to make sure that their training (including the datasets used) and processing (including the code) are accessible for audit and review.
- How will the roles of traditional oversight bodies evolve in response to AI adoption? Supreme audit institutions and independent fiscal institutions will need to adapt their methodologies to effectively oversee and audit AI-powered or driven fiscal processes. This might include developing new skills in data science and AI among staff members and updating audit processes to incorporate AI-specific considerations such as data integrity and algorithmic bias.
- What new safeguards will be necessary to protect against misuse of AI in fiscal governance? As AI systems play a more prominent role in fiscal decision-making, identifying possible types of intended or unintended misuses and safeguarding against them becomes paramount.

By addressing these questions, PFM specialists can help shape a new fiscal governance framework that accommodates the full innovative potential of AI, while safeguarding the key principles of PFM.

References

Australian Government (2019-2020), *Department of Veterans' Affairs Annual Report 2019-20*, https://www.transparency.gov.au/publications/veterans-s-affairs/department-of-veterans-affairs/department-of-veterans-affairs-annual-report-2019-20. [11]

Australian Government Department of Finance (2018), "Using data analytics to target government expenditures", *Paper presented at the 39th Annual Meeting of Senior Budget Officials*. [10]

Boström, H. et al. (2020), *Explaining Multivariate Time Series Forecasts: An Application to Predicting the Swedish GDP*, https://ceur-ws.org/Vol-2796/xi-ml-2020_bostrom.pdf. [9]

Finish Agency for financial management and HR (Palkeet) (2024), *RPA/AI in Financial Management*. [16]

French Public Finances General Directorate (DGFiP) (2024), *Digital Transformation of the Public Finances General Directorate (DGFiP)*. [14]

French Public Finances General Directorate (DGFiP) (2024), *The predictive model*. [13]

Government of Mexico (2024), *Plataforma inteligente de apoyos del Gobierno Federal*, https://www.transparenciapresupuestaria.gob.mx/es/apoyosdelgobierno (accessed on 20 April 2024). [20]

Jung, J., M. Patnam and A. Ter-Martirosyan (2018), *An Algorithmic Crystal Ball: Forecasts-based on Machine Learning*, International Monetary Fund, https://www.imf.org/en/Publications/WP/Issues/2018/11/01/An-Algorithmic-Crystal-Ball-Forecasts-based-on-Machine-Learning-46288. [7]

Korea Fiscal Information Service (KFIS) (2023), *Digital Transformation in Public Finance: Korean FMIS (dBrain+)*. [5]

Ministry of Finance, UAE (2024), *Ministry of Finance Showcases "Virtual Mona": An AI-Driven Financial Data Advisor at GITEX Global*, https://mof.gov.ae/ministry-of-finance-showcases-virtual-mona-an-ai-driven-financial-data-advisor-at-gitex-global/?utm_source=chatgpt.com. [21]

Moretti, D., A. Keller and M. Majercak (2023), "Medium-term and top-down budgeting in OECD countries", *OECD Journal on Budgeting*, https://doi.org/10.1787/39425570-en. [6]

OECD (2024), *OECD AI Principles overview*, https://oecd.ai/en/ai-principles (accessed on 31 January 2024). [1]

OECD (2023), "Generative artificial intelligence in finance", *OECD Artificial Intelligence Papers*, No. 9, OECD Publishing, Paris, https://doi.org/10.1787/ac7149cc-en. [4]

OECD (2023), *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*, OECD Publishing, Paris, https://doi.org/10.1787/08785bba-en. [3]

OECD (2022), “2022 OECD Financial Management and Reporting Survey”, https://data-explorer.oecd.org (accessed on 22 April 2024). [26]

OECD (2022), “OECD Framework for the Classification of AI systems”, *OECD Digital Economy Papers*, No. 323, OECD Publishing, Paris, https://doi.org/10.1787/cb6d9eca-en. [27]

OECD (2019), *Artificial Intelligence in Society*, OECD Publishing, Paris, https://doi.org/10.1787/eedfee77-en. [2]

Padfield, M. (2025), “Why the policy profession must embrace AI”, *Heywood Quaterly*, Vol. 3, https://heywoodquarterly.com/why-the-policy-profession-must-embrace-ai/#. [22]

Peeters, R. and A. Widlak (2023), “Administrative Exclusion in the Infrastructure-level Bureaucracy: The Case of the Dutch Daycare Benefit Scandal”, *Public Administration Review*, Vol. 38/4, pp. 863-877. [24]

Rocha, K., K. Porto and V. Fabel (2024), *AI is Enhancing Fiscal Transparency in Brazil*, https://blog-pfm.imf.org/en/pfmblog/2024/12/ai-is-enhancing-fiscal-transparency-in-brazil#:~:text=Brazil%27s%. [15]

Royal Commission into the Robodebt Scheme (2023), , Vol. 1, https://robodebt.royalcommission.gov.au/system/files/2023-09/rrc-accessible-full-report.PDF. [25]

Tenorio, J. and W. Perez (2023), *GDP nowcasting with Machine Learning and Unstructured Data to Peru*, https://perueconomics.org/wp-content/uploads/2023/11/WP-197.pdf. [8]

United Kingdom National Audit Office (2023), *Report on Accounts - Department for Work & Pensions*, https://www.nao.org.uk/wp-content/uploads/2023/07/dwp-report-on-accounts-2022-23.pdf. [19]

United States Department of the Treasury (2024), *Paying the Right Person, in the Right Amount, at the Right Time*. [18]

United States Government Accountability Office (2023), *Key Practices to Help Ensure Accountability in Federal Use*, https://www.gao.gov/assets/gao-23-106811.pdf. [23]

Valle-Cruz, D., V. Fernandez-Cortez and J. Gil-Garcia (2022), “From E-budgeting to smart budgeting: Exploring the potential of artificial intelligence in government decision-making for resource allocation”, *Government Information Quarterly*, Vol. 39/2, p. 101644, https://doi.org/10.1016/j.giq.2021.101644. [28]

van Noordt, C. and L. Tangi (2023), “The dynamics of AI capability and its influence on public value creation of AI within public administration”, *Government Information Quarterly*, Vol. 40/4, p. 101860, https://doi.org/10.1016/j.giq.2023.101860. [17]

Wisesa, A. (2023), “Inovasi artificial intelligence sebagai financial advisor kementerian keuangan”, *Swatantra*, Vol. XX, pp. 1-8, https://jurnal.umj.ac.id/index.php/SWATANTRA/article/viewFile/15611/8230. [12]