INTERNATIONAL INSTITUTE FOR ANALYTICS™

RESEARCH BRIEF

RESEARCH & ADVISORY NETWORK

Boosting Tax Agency Productivity and Protection with AI and Generative AI

SAS EXPERTS: TEYA DYAN, ALEXANDRE NEGADI, ENSLEY TAN

IN DISCUSSION WITH ROBERT MORISON, SENIOR ADVISOR, IIA

NOVEMBER 2024

Discussion Overview

Generative AI (GenAI) technology is gaining momentum rapidly, not just in businesses and government agencies, but also among those trying to defraud them. Tax agencies have potentially valuable uses of GenAI in process productivity, taxpayer service, and staying ahead of fraudsters by augmenting fraud detection and threat simulation. Given the new and sometimes uneven capabilities of GenAI, agencies should be approaching it with both purpose and caution. To explore these opportunities and threats, IIA spoke with SAS global tax and revenue compliance experts Teya Dyan, Principal Industry Consultant, based in North America, Ensley Tan, Principal Industry Consultant, based in Asia Pacific, and Alexandre Negadi, Senior Industry Consultant, based in Europe.

Let's start by distinguishing generative AI from the traditional AI and analytics that tax agencies currently use. What is new about GenAI?

Alexandre Negadi: Artificial intelligence overall is the science of designing computer systems to support and accelerate human decisions and actions. It has traditionally been used to analyze data to perform specific tasks like classification, optimization, and prediction. GenAI uses machine learning models trained on massive data sets to create text, images, music, and other types of new and synthetic data. The key word is "new."

Think of traditional AI as a powerful analytical tool that excels at finding patterns and insights in existing data. And think of GenAI as a creative partner who can produce new outputs under the prompts and guidance of a human. People are often captivated by the outputs of ChatGPT and other generative AI tools, but it's important to recognize that the output is only as good as the questions you ask and how you drive the interaction.

Teya Dyan: Traditional AI encompasses several major disciplines, including machine learning where algorithms learn from feedback to improve their performance, and natural language processing where algorithms handle text and voice data in realistic ways. For today's tax agency applications, think of GenAI as the child of machine learning and natural language processing. Using large language models, GenAI can generate new written and auditory content based on the data it was trained with. But like a developing child, it doesn't understand the world and it can make mistakes, even building on its own incorrect assumptions. It needs guidance—interaction with humans to shape the inputs, guide the learning process, and assess the outputs.

Ensley Tan: We must recognize how good GenAI is at appearing to understand how humans speak and interact. In June, a ChatGPT bot passed the Turing test. So the barriers to using it are far lower than with traditional AI. You don't need to understand how it works, and you can give it instructions as you would to another person rather than writing code or search criteria. But a big risk is the almost naïve over-reliance on GenAI outputs. Just because the answers sound good or realistic, it doesn't mean they're reliable. In fact, even though everyone now acknowledges that GenAI can hallucinate, there is an effort/accuracy tradeoff decision each of us has to make on how deeply to doublecheck the answers we get. An even bigger risk that we have to recognize is how quickly GenAI has become a tool for perpetrating fraud.

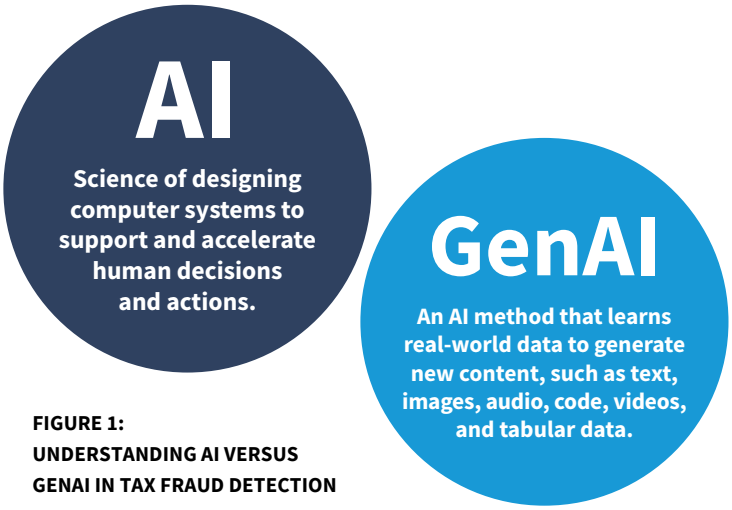

FIGURE 1:
UNDERSTANDING AI VERSUS GENAI IN TAX FRAUD DETECTION

How are fraudsters using GenAI, and how can tax agencies best defend against these new threats?

Ensley: Fraud is all about generating false information to conceal wrongdoing—tax returns, business records, and identifications—and the more realistic the information looks, the harder it is to recognize the fabrication. GenAI makes it much easier for fraudsters, including criminal organizations, to do bad things quickly and on a large scale, with widespread consequences.

For example, a criminal might create a prompt like: "What does a convincing document look like? Now make me 1,000 variations." A tax agency may see document and transaction rates increase, but good fake documents will make it hard to even know the scale of the problem.

Teya: Fraudsters are always trying to exploit vulnerabilities in agency information systems, processes, and safeguards. Now they are generating fake people and businesses by submitting registration applications for synthetic identities, and they are doing it en masse.

Over time, they are getting better at deceiving significant numbers of people and organizations. Once legitimate-looking registrations are established and tax ID numbers are issued by a governing authority, these fake "taxpayer" identities can be used to open bank accounts for money laundering, file fraudulent insurance claims, "hire" employees and commit identity theft of people looking for work, claim benefits for fake unemployed workforce or injury claims, bilk consumers out of their hard-earned money, and of course, file fraudulent tax returns and request illegitimate refunds. For example, a fake building contractor can be up and running with a legitimate-looking business license and taking deposits from customers in record time.

Tax agencies, in turn, have to manage the increasing volume and improve their vetting of registrations. Traditional analytics and AI, such as pattern recognition, network analysis, and anomaly detection to calculate a robust risk score are very effective at detecting fraud at the source of registration. Agencies should also consider using GenAI for fraud scenario simulation to empower agencies to know what fraudsters might be attempting based on publicly available data sources and forms.

Alexandre: The mass generation of synthetic identities for legitimate-looking businesses or individuals poses one of the biggest challenges that tax agencies have ever faced. There's a significant increase in business being conducted online, with more reliance on digital identities, especially among younger generations. But tax and social welfare agencies are much less protected than financial services institutions in terms of cybersecurity and fraud detection and prevention. Banks are still vulnerable, but they have extensive experience and infrastructure. Agencies need to catch up quickly and they should work together. But it's hard to call attention to the problem because, like the banks, they understandably don't want to reveal its full extent.

On the positive side, what are the promising use cases for GenAI within tax agencies?

Ensley: Based on people's experience with GenAI to date, the most apparent use case is chatbots assisting agency staff. For example, a GenAI-powered chatbot can help policy and guidelines makers be aware of and understand the tangle of existing rules and legislation and evaluate the effect of potential changes: "If we added this rule, what other rules would it affect?" On a broader scale, GenAI can help agencies more holistically understand emerging global issues and how to align with, or react to, them. For example, an international minimum effective tax rate (METR) for corporate taxation is on the drawing board. GenAI can help agencies keep track of how other governments are moving to address the issue. Similarly, we can envision chatbots for advising citizens on common issues, helping them reduce errors and inadvertent non-compliance with their tax obligations.

Alexandre: In its "creative partner" role, GenAI can also assist with prioritizing and conducting audits, managing investigations, and optimizing collections. This makes better use of staff time, enabling them to focus on the cases with the most probable yields.

Ensley: One caution here: If staff members are consulting a chatbot alongside their regular information system, they will have to keep switching contexts. That can be distracting, and we know that whenever people must deal with multiple systems simultaneously, context switching can increase error and decrease productivity. Agencies need to pay close attention to how chatbots integrate with workflow.

Teya: At this stage, I would strongly caution against deploying GenAI chatbots to interact with the public. We certainly don't want them hallucinating with tax guidance or giving advice on statute and policy out of context. They would have to be trained carefully on vetted and controlled data sources and thoroughly tested using both predictable and unpredictable prompts. However, GenAI is proving very useful in taxpayer communications by generating letters and messages for outreach and education that are easily understandable, free of legal jargon, and translatable into different languages. This communication is then reviewed and approved by a knowledgeable human before it is sent to the public. This process can expedite outreach and education and ensure that new information is readily consumable by the targeted audience.

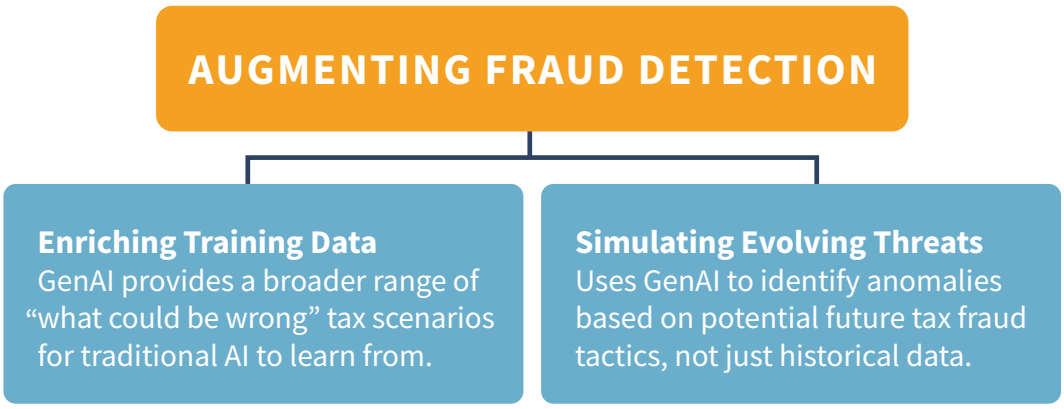

FIGURE 2: BENEFITS TO EMPLOYING GENAI IN TAX FRAUD DETECTION

For internal use, GenAI chatbot advisors can be incorporated into training and knowledge retention processes. For example, when new legislation is passed, agencies may have to train staff in the changes very quickly, including what can and can't be communicated to taxpayers. Tax can be an incredibly detailed and convoluted field. Chatbots can also help new staff get up to speed more quickly. Even experienced staff can benefit from an assistant that can scan and make sense of all the relevant statutes and publications, rules and guidelines associated with a specific question or prompt.

Alexandre: Among the most potent applications of GenAI today is augmenting traditional fraud detection methods. Agencies can train GenAI to analyze immense amounts of quantitative tax-related data, including data from external sources, to help find new patterns of incoherence between declared revenues. Then they can adjust anomaly detection models accordingly.

Agencies can generate synthetic data of their own and put it to good use training their fraud detection models and simulating evolving threats. For example, you can generate new anonymized declarations of revenues that are fraudulent and use that synthetic data to train traditional AI models to recognize a broader range of "what could be wrong" scenarios. GenAI can also be used to simulate known tactics of fraudsters and how these evolve—and even create potential fraud tactics to use for testing the efficiency of existing detection models. In both cases, the statistical properties of the synthetic data can be fine-tuned to the target scenarios you're simulating.

What are the keys to success when tax agencies deploy GenAI?

Teya: I would say be prudent but not paranoid. There are use cases that are very appropriate to start trying, but keep a human in the loop. Be purposeful about the sources your GenAI application is trained on and evaluate the results. For example, GenAI can help interpret laws and regulations. But some laws contradict each other across jurisdictions, and others are written in terms of different scenarios that GenAI may have trouble distinguishing. So you can't turn GenAI loose to find all the information on a particular topic and take the results at face value, because it can assemble chunks of information in ways that are incorrect or non-sensical. That said, GenAI may be a tool for discovering opportunities to clarify laws and policies.

Alexandre: To deploy AI and GenAI successfully, agencies need structures for governance, including compliance and security. Most have established data governance programs, with roles, rules, and guidelines, and now they need AI governance and GenAI governance, which are more complicated. In Europe, the new EU AI regulations need to be understood and followed. There's a big difference between experimenting with AI and GenAI and using them operationally. It involves training many people and addressing new risks. Don't underestimate the extent of change and do establish governance to manage it.

Also, a robust and secure technology and software platform is key. Let's say a country's finance minister wants to introduce new taxation rules and anticipate their impact on different taxpayer segments. A robust platform can run enormous amounts of data and provide sophisticated analytics in a few minutes. We understand these platforms and technology integration challenges because at SAS we've been anticipating and working on them for a decade.

Ensley: Pay close attention to the transition issues. If you bring in a large language model or GenAI capability, are they compatible with existing systems? Can you merge the datasets needed for large-scale searches? If you need more processing power, what will it cost? Most fundamentally, can you get consistent answers? You can't have a system that gives different answers to the same question. Nor can you have an organization where everybody has their own favorite way of asking the same question. So take into account what kind of training is needed.

Also be realistic about what you're trying to do and control the scope. GenAI capabilities are extraordinary, but they're not needed for most tax administration, policy, research, and customer service activities. Which processes will benefit from iterating with a creative partner that consults a large language model? When do traditional and trusted AI and analytics models working in the background get the job done—and save the employee the overhead of context switching?

What will it take to have more trusted GenAI?

Alexandre: Agencies can integrate GenAI tools into use cases with clear intent and benefits. But they still have to continuously evaluate effectiveness and security. If you can measure effectiveness and control security, then you can have a trusted solution. I call this strategic implementation, and it's again dependent on governance.

Many trust concerns center around the data that models are trained on and access. The internal systems of tax agencies are well protected, and their data can be generally trusted. However, external data is less reliable, particularly due to the proliferation of misinformation and deepfakes on the web.

But if agencies really need to incorporate external data, they can use traditional AI to vet it, screen it for anomalies and inconsistencies, in ways analogous to detecting fraud.

Ensley: First, GenAI tools have to engender more trust by indicating what data they use and how they get their answers. They can't be complete black boxes. Tracing is inherently tricky, but we're seeing some progress. ChatGPT and Copilot, for example, can give you a footnote of the sources, sentence by sentence. This is especially important when accessing external data sources.

Second, people need education on how to use the tools sensibly and not take answers at face value. They need to understand how GenAI works and what large language models are prone to doing. Also, they need to be aware of how their tools behave, because ChatGPT, Copilot, and Gemini have different ways of interpreting user prompts. Users should appreciate how GenAI is designed to please them, and thus learn when and how to be skeptical. The most skilled users will develop a sense of their own tendencies and potential biases in how they ask questions and accept answers.

Teya: Tax agencies should recognize that AI is a big umbrella, covering many different tools and methods—many of them well-vetted, monitored through model management, and trustworthy. As the new tool, the new child of AI, GenAI can still hallucinate and so needs extra oversight. Trust starts with using the right tool for the use case, scenario, or problem at hand. Use traditional AI to understand our enterprises, anticipate outcomes, and inform decisions. Use GenAI where we want to put those things into language that really makes sense.

There's another angle around trust—and fear—that we haven't touched on. Many employees across sectors today resist GenAI because they're afraid it will take their jobs. But as we've discussed, people still have to be in the loop. In tax agencies, there's no shortage of work to do, and there are opportunities to do it better. I think Karim Lakhani of Harvard Business School has it right when he says, "AI is not going to replace humans, but humans with AI are going to replace humans without AI." Agencies understand that AI is where the world is heading, and they want to be able to use it. The staff education challenge includes encouraging people to be part of the journey.

What concluding pieces of advice do you have for tax agency leaders when it comes to GenAI?

Teya: Start by using GenAI on clear and worthwhile use cases. One is to help summarize and interpret all the regulatory documents and publications and even help put them into plainer and more easily understood language. That can help both staff and taxpayers understand the rules, and we see increases in voluntary compliance when the requirements are clear.

Another is using GenAI assistants to help staff do their jobs and learn as they go. GenAI can help staff adapt to new regulatory changes. And it may be a powerful tool for new employees coming up to speed, helping them develop competence, consistency, and confidence in their work. That could make a real difference if the agency has staff shortages, too much turnover, or many employees approaching retirement.

Ensley: First, regardless of how much you're using GenAI within the agency, look forward to what fraudsters are doing and keep working on detection and defenses with traditional analytics and AI.

It's easier for fraudsters to experiment with GenAI and see what works than it is for agencies to recognize and counter new threats. GenAI is out there, so plan for how to respond as fraudsters' capabilities increase.

Second, as you are deploying GenAI in-house, figure out how to properly integrate it into work processes. Should it be front and center, like in a chat window that forces context switching, or should it just be another tool in the toolbox for selected staff? The scope of use on the front line might be limited, while those involved in policy research and formulation might interact with GenAI regularly.

As with all new tools, the user experience determines adoption rate. If GenAI is introduced as an additional process, or added to an already busy job, uptake can be excruciatingly slow. However, if GenAI fits seamlessly into the workflow and people experience the value, both usage and benefit realization will accelerate.

Alexandre: If you are serious about making the most of AI and GenAI, you need a strong, dedicated task force behind the effort. This team keeps up to date and develops intelligence around AI and its tools, they help formulate governance and maintain the roadmap of direction and next steps, they oversee pilots and rollouts, and they enable the organization to get the human-AI mix right. The task force leader should report directly to the agency's general manager. This leader also has a role as an expert trainer and communicator within the organization, and must inform people regularly of the latest discoveries and developments in the analytical systems. I mentioned that the transition to operational use of GenAI will take time. The task force maintains focus and momentum.

Teya Dyan

Teya Dyan holds over eighteen years of experience in tax policy regulation and fraud investigation. Her expertise in management of compliance teams and as a front-line investigator led her to embrace the powerhouse combination of advanced analytics and traditional intelligence gathering for case discovery and development. Prior to joining SAS, she worked in Washington State and Kansas Revenue agencies. She now serves as President of the Pacific Northwest License, Tax, and Fraud Association and is a principal consultant with Government Fraud and Security Intelligence Solutions at SAS Institute. Her current role with SAS combines her professional passions of data analytics, investigations, tax policy, and perpetually curious gumption to foster innovation through technology.

Alexandre Negadi

Alexandre Negadi brings over thirty years of experience in selling and delivering consulting and IT services across various industries, with a particular focus on fostering digital transformations in the tax and social welfare sectors for the past twenty years. He has held leadership positions where he developed strategic offerings, including anti-fraud strategies and automated processing solutions for government administrations.

Ensley Tan

Ensley Tan has over twenty-one years of experience, primarily in the government space, with a specific expertise in law enforcement and investigations. During his fourteen years with the Singapore civil service, he worked closely with agency heads and political office holders on a wide range of external policy concerns, including cross-border financial issues with Singapore's economic agencies. As a fraud consultant based in Hong Kong, Tan conducted hundreds of fraud investigations throughout the Asia Pacific and consulted with various government agencies, including advising a regional country's ministry and industry association on the framework and implementation of AML/CTF guidelines as part of that country's FATF obligations. He's currently the principal industry consultant for the public sector in the Asia-Pacific for SAS Institute. Passionate about using analytics to support data-driven decisions, Tan bridges the gap between what data and analytics can do and what government agencies need.

Robert Morison

Robert Morison serves as senior advisor to IIA. An accomplished business researcher, writer, discussion leader, and management consultant, he has been leading breakthrough research at the intersection of business, technology, and human asset management for more than 25 years. He is co-author of *Analytics At Work: Smarter Decisions, Better Results* (Harvard Business Press, 2010), *Workforce Crisis: How to Beat the Coming Shortage of Skills And Talent* (Harvard Business Press, 2006), and three Harvard Business Review articles, one of which received a McKinsey Award as best article of 2004. He holds an AB from Dartmouth College and an MA from Boston University.

IIANALYTICS.COM