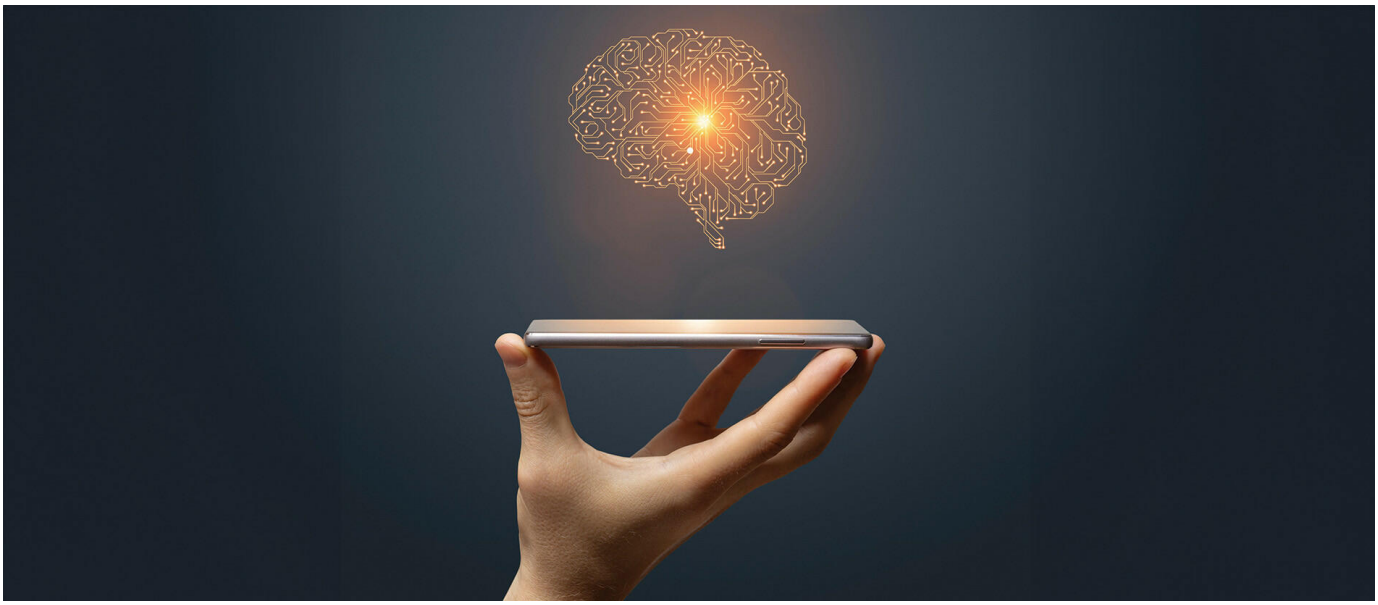
Generative AI: what can it bring to tax?

General Features



21 February 2024

We explore the possibilities that Generative AI can bring to the tax world, along with a consideration of its current limitations.

Key Points

What is the issue?

GenAI is a rapidly evolving field, and there is much debate over its potential impact on tax professionals. This article explores the power of GenAI in balance with its limitations.

What does it mean to me?

The use of GenAI is being considered for a range of tax functions, including document generation and summarisation, knowledge extraction, and data querying,

analysis and visualisation.

What can I take away?

No technique is perfect, however, and there should always be a human to apply a layer of subject matter expertise to review GenAI outputs.

This article aims to provide a simple overview of generative artificial intelligence (GenAI) and its relevance to tax. As artificial intelligence, particularly GenAI, is a rapidly evolving field, there is much debate over its potential impact on tax professionals. Rather than taking a binary viewpoint, this article explores the power of GenAI in balance with its limitations and identifies what tax professionals should focus on in the near term.

What is Generative AI?

GenAI is a form of artificial intelligence that enables users to generate new content across various modalities – such as text, image, code and audio content – based on a broad set of inputs through chat interfaces. There are different types of GenAI models, including text-to-text generators powered by Large Language Models that are trained on massive amounts of text data, amongst others.

These models learn the patterns and structure of natural language and generate text that has similar characteristics. Chat interfaces have transformed GenAI usage by making artificial intelligence more accessible and enabling users to interact with these technologies in a more intuitive way, producing content that can be indistinguishable from that produced by a human.

How GenAI can transform the tax profession

Artificial intelligence has been utilised in tax for over a decade, among other things for automating the analysis of transactional data. This could include, for example, categorising corporate expenses data into allowable and disallowable expenses for tax compliance purposes using supervised machine learning. However, artificial intelligence has not yet automated all the tasks performed by tax professionals.

There are three key aspects that differentiate GenAI from traditional artificial intelligence.

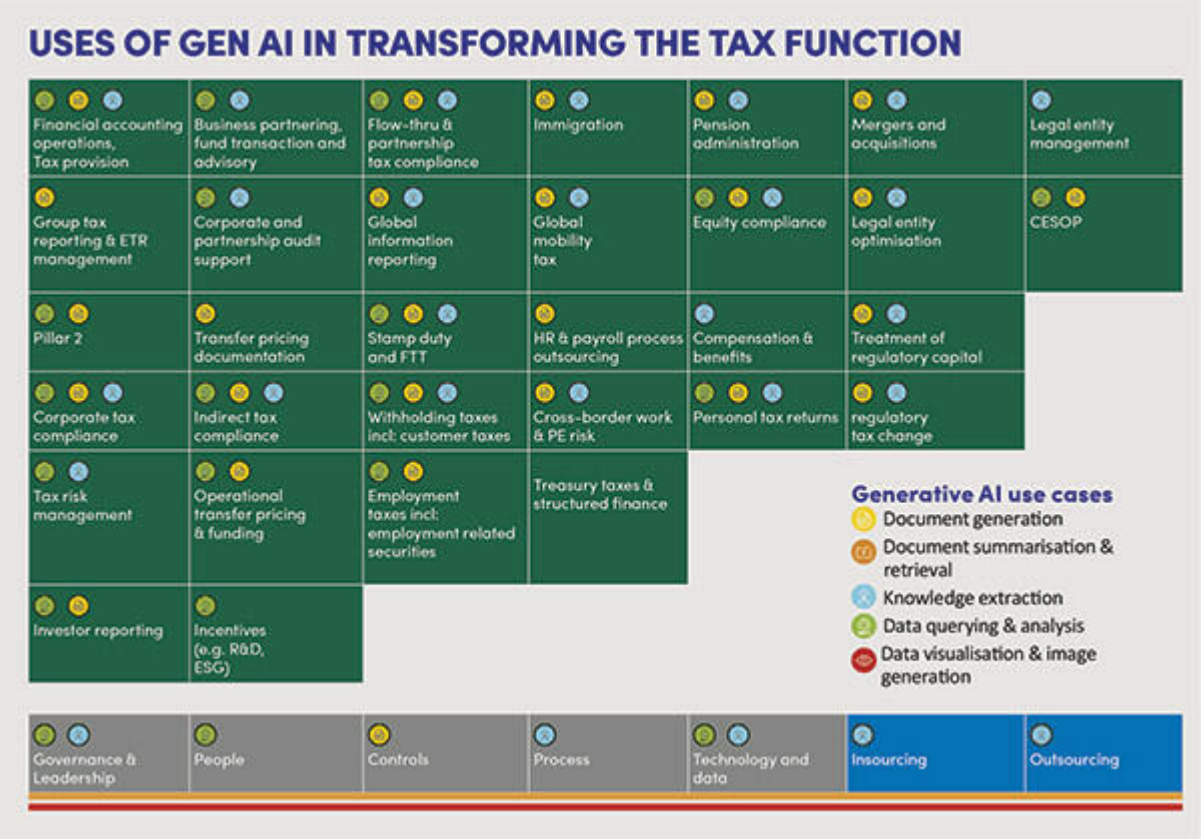
1. Large Language Models use general-purpose text data based on open-source content, such as the web. It is worth noting that the quantum of information available to a Large Language Model is greater than what is available to a single tax professional, and it can distil information at a significantly faster speed.
2. The output of a GenAI application is generally new, novel content that reflects the characteristics of the input data but does not repeat it. In contrast, traditional machine learning aims to produce a desired outcome or target as a prediction, such as the classification of tax data for analysis in a tax return.
3. Finally, GenAI makes artificial intelligence more accessible to a wider audience beyond data scientists. It allows general users to access it in an intuitive way to boost productivity with its ability to process and generate natural language. Additionally, there are several powerful open-source models available, making Large Language Models more accessible for experimentation and application across a variety of use cases.

How does GenAI apply to tax?

From our conversations with major corporates, the types of use cases being considered or implemented are in the following categories:

- **Generation of documents:** assisting with tax compliance, including producing tax memos and supplementary returns;
- **Document summarisation and retrieval:** reviewing new tax legislation and summarising its contents;
- **Knowledge extraction:** highlighting the pertinent points within a tax consultation;
- **Data querying and analysis:** applying tax legislation to a specific company or personal situation in order to produce an optimal outcome; and
- **Data visualisation and image generation:** generating process maps for governance purposes, and producing video training for non-tax professionals on tax matters such as account payable processing.

Image



The diagram above illustrates some of the use cases by area of tax. Whilst this is a rapidly evolving area and it is not an exhaustive list, it gives an overview of the breadth of GenAI. It is striking how many of these uses are emerging in the advisory space – not just in compliance, where traditionally technology has been most impactful to how tax work is done.

What are the limitations of GenAI?

Despite its power, there are significant limitations to the use of GenAI in tax. Below, we set out some of the possible key limitations, why they occur and what can be done to mitigate them.

Hallucination and producing inaccurate results

GenAI Large Language Models can produce content that is wrong, although it looks very plausible. This is a phenomenon called hallucination and is largely due to the

nature of how these models are built and function.

Large Language Models are statistical models and predict the next word in a sentence based on probabilities derived from the vast amounts of text data it has been trained on. They are driven by the patterns they see in the training data, as opposed to understanding the meaning or context in the way humans do. They cannot draw on experiences or common sense. As language is inherently complex and ambiguous, Large Language Models can also struggle to reproduce language nuances.

It should be noted that the impact of outputs being incorrect could be high in a regulated environment, whether in the public, corporate or personal spheres. We have already seen cases come before UK courts where Large Language Models have been used to build the technical arguments and some of these arguments have been hallucinated.

The data which trains the models

Tax work is often based on experience and tailored consultation. Much of this work is not available in the public domain and therefore is not included in the data used to train these Large Language Models.

The other challenge is how up to date the data set is. Tax regulation changes frequently and using the right reference rule set is critical to producing accurate results. At the time of writing, GPT4 was accurate up to April 2023, so would not capture legislative updates from the UK Autumn Statement 2023, for example.

Large Language Models aren’t built for real-time search and retrieval purposes. They are not a search engine as their primary purpose is to understand natural language, not to be up to date with world affairs. For Large Language Models to act like a real-time search engine, they will need to be integrated with search engine technology; i.e. a hybrid model.

Confidentiality and privacy

A person’s tax affairs are a confidential matter. There are cloud platform providers which provide technology solutions to keep data secure from the cloud and Large Language Models. This typically requires the involvement of an organisation’s IT

function. For smaller businesses, however, these solutions may not be practical in the near-term, restricting the use of GenAI to cases where confidentiality may be less challenging.

The economics of using GenAI

Large Language Models split the text into tokens for processing. Tokens are usually individual words, but can be a group of letters. The cost of using Large Language Models varies, based on the tokens processed and the computational power needed.

Much tax legislation is several thousand words long and the context of these words is critical for interpretation. There are token size limits on the interfaces for these Large Language Models which, whilst ever expanding, do restrict processing. For example, the OECD's original Pillar Two guidance and all subsequent versions are too large for current processing so asking for a summarised response of changes between guidance isn't possible in one user query.

Token size is also used as the unit of cost for the use of GenAI models. The balance of cost vs potential productivity saving compared to the cost of manual interpretation needs to be considered. It is worth noting that token size limits and pricing models are ever evolving so the impact of this limitation is likely to reduce over time.

In conclusion

There are a range of techniques that can be used to overcome some of the challenges posed by current state GenAI. No technique is perfect, however, and there should always be a human to apply a layer of subject matter expertise to review GenAI outputs.

GenAI is having, and will continue to have, an impact on the tax profession. The magnitude, timing and types of impact can be debated. As tax professionals, there are four things we can do in 2024 to appropriately consider this new player in our profession.

1. Educate ourselves on the domain area by reading articles, attending webinars, and so on. This is a rapidly changing area, so for most people it is almost

impossible to stay up to date completely, but it is worth engaging with key updates. Use this to develop a view of what GenAI could mean for the tax activities you undertake or lead.

2. Reflect on use cases and where most value can be added within tax.
3. Experiment with GenAI, noting its limitations.
4. Have an open mind to the positive effect that GenAI could have on the productivity of our next generation of professionals.
5. Lastly, review your data and infrastructure readiness to adopt artificial intelligence.

© Getty images/iStockphoto