Three Ways to Make AI Work

KPMG. Make the Difference.

Three Ways to Make AI Work

A few months after the introduction of ChatGPT, the maelstrom of interest in the transformative potential of Generative AI (GenAI) reached a fever pitch. Before this, the challenge with AI was always how to make it move beyond a novel, yet clever, technology.

While an AI that could beat the world's most accomplished board game players was an exceptional feat of mathematics and engineering, most of the business world recognised it as an expensive parlour trick.

Even the more commercial hints provided, such as the way in which algorithmic trading fundamentally changed the way all investors participated in capital markets, or the more hidden successes of advanced mathematic techniques in the energy sector, didn't materially turn the heads of corporate leaders.

ChatGPT, driven by an advanced GenAI technology did turn heads. And since then, practically every board room and leadership team in the world has asked a series of critical questions that seem harder than they should be to answer. If we are to realise or even lead on the value that GenAI seems like it will create, where do we start, what do we need, and how do we do this safely?

Paul Henninger
Partner, Head of UK Connected Technology and Global Lighthouse

To answer those questions, it's important to try to unpack some of the confusion that persists around how GenAI and AI in general can produce change. And this is not accomplished by unpacking the details of the algorithms used. Instead, it's useful to understand the **three dominant models** that are driving AI change in different ways:

01 **Agent Augmentation**: the broad deployment of AI as a productivity tool across large user and employee populations (e.g., Microsoft Co-Pilot or ChatGPT itself).

02 **Enterprise Augmentation**: the deep insertion of AI-led multi-technology applications to solve specific operational and functional problems (e.g., bespoke applications for automating risk processes or customer interactions).

03 **Cognitive Agents**: the next generation development of AI native to the data of a discrete real-world process (e.g., an AI trained to 'speak' Procurement or Product Design in the way that the early GPTs were trained to 'speak' a language).

These three models are likely to have very different levels of productivity or value impact on an organisation that invests in and deploys them successfully, as well as having very different risk profiles. They require access to different skills and types of infrastructure and lend themselves to very different types of data programmes.

Although a truly expansive AI strategy can connect all of them, they do not necessarily organically lead from one to another. The answers to where we start, what we need and how to do this safely depends on which model or models of AI change an organisation decides to make a priority.

As ever, the best way to understand this is by example.

01 Agent Augmentation

Agent Augmentation offers the promise of rapid, almost non-technical, broad-based AI change. In some respects, it is the most revolutionary of the three models. The promise is that by deploying a little bit of AI technology to a lot of places, users or employees in an organisation can bring the innovation promised by GenAI to the tens or hundreds of thousands of tasks performed across a department or entire organisation.

Super-users in the agent augmentation model are trained as Prompt Engineers, able to manipulate the GenAI toolset in the way an Excel wizard can model practically anything you can think of in a business.

Many of the initial applications of Agent Augmentation have inserted intelligent automation into parts of our organisations that were very painful for users and that seemed resistant to digital transformation. The deployment of GenAI as a front end for HR Policies and Learning resources, for example, suddenly transformed every employee's ability to manage their employee experience. Self Service HR went from a promise to an advanced capability overnight, where employees saved time in simple tasks. An employee researching parental leave or a travel policy, with a properly deployed toolset, could get the information and context needed for their specific decision or problem in seconds, compared to hours of hunting through policy documents.

When focused on the right outcomes, Agent Augmentation quickly introduced 15%-20% productivity improvements across large tracts of the employee base for certain tasks, and often led to improvements in the quality of outcomes in parallel. Used to compose a digital strategy, for example, an agent assisting a CIO could do the work of a team of 5 in 10% of the time, and when exposed to the right source material, at a good level of quality[1].

The potential for Agent Augmentation is limited in some respects by the power of the underlying models and the data exposed to the platforms. The risks of the Agent Augmentation model are to some extent controlled by the fact that in its current form, while capable of automating knowledge work and similar tasks, the models are not particularly ready to rewire a process inside an organisation.

Even a very skilled member of a finance team can't restructure an end of quarter close process with a well-designed prompt.

So, the limits of this type of AI mean that the risks are limited as well. And the limits of GenAI to, for example, distinguish or prioritise different versions of the same information (e.g., different versions of an HR policy for different countries or over time) mean that data sets required to scale the use of the technology need a high degree of curation, whether by the individual user with knowledge of the data set as they use the tool, or increasingly as part of a well-planned deployment of the technology.

Agent Augmentation is literally productivity tools on steroids. Almost as easy to deploy, and only incrementally riskier than a lot of tools already widely used in organisations, but with possibly unintentionally safe limits on the value and change we can create with this model. The sharp-eyed leader may also observe that one of the most helpful by-products of Agent Augmentation is its incredibly useful way to prepare large numbers of users or employees for the significant change likely to be introduced by the other two models.

1. Martinez & Mezitis, 2023.

02 Enterprise Augmentation

In highly effective deployments, Enterprise Augmentation can introduce a truly profound transformation of a relatively narrow functional slice of the organisation.

Led by GenAI, for example, Enterprise Augmentation applications deployed to perform analysis, search and summarisation of stacks and stacks of evidence in a class action lawsuit, or to check reams of marketing material for green washing can perform work that would otherwise require the deployment of hundreds of highly skilled lawyers or risk professionals. The limits of these applications – often composed from a mix of GenAI, machine learning models, graph analytic tools and other automation technologies – are in the specifics of their design. An Enterprise Augmentation AI designed to do legal work can't recommend a gift for a 5-year-old's birthday and won't be able to process a question about a financial issue. But the degree of productivity and value impact on the function these AI applications are pointed at can be dramatic.

From our own project experience at KPMG UK, when accompanied by the optimal amount of process and people transformation, these types of solutions can take more than 50% of the work out of the tasks they are deployed against and in some cases can result in 80% or even greater than 100% productivity gains making it possible to process as many mortgages in a day as the recently most advanced automated processes could in a week, enabling legal teams to prepare for 'surprise' witnesses or last minute depositions and other legal tactics with a degree of preparedness that wasn't possible before.

Because of the degree of impact, the risks of these systems are narrow in scope, but considerable in potential impact. They must be well designed to produce specific outcomes, and thoroughly tested in the right ways.

But with a comprehensive design, build and testing process that is overlaid with appropriate in-live monitoring, the risks are manageable.

The technology skills and infrastructure involved in deployment, of course, require greater investment and organisational maturity than the Agent Augmentation model. The use of GenAI however, greatly reduces the investment and time required to deploy these systems programmatically within an organisation and the change that this model introduces is profound.

The limits of Enterprise Augmentation, however, are that these ecosystems of intelligent applications seem to fall short of the highly democratic, highly general, almost human promise of those first experiences with ChatGPT and similar Gen AI deployments. The Enterprise Augmentation future looks a lot more like the floor of a modern factory full of robots and robotic machines than the bridge of the Starship Enterprise. Nonetheless, this is how some of the most dramatic changes will be introduced into the enterprise and our lives over the next 24-36 months. In parallel, some enterprises will be investing in what I think many executives understood to be the promise of GenAI in the form of the third model for AI change: Cognitive Engines.

03 Cognitive Engines

The Cognitive Engine delivers on the promise of the almost Sci Fi future of GenAI, or at least it has the potential to.

While OpenAI's GPT models (along with Meta's, Google's and Anthropic's models) gained incredible capability via exposure or training to large tracts of information produced over the last 30 years or so of digitisation, their ability to generalise comes in some respects at the expense of their specificity in a distinct domain. It can have a chat with a user about procurement, for example, but the models aren't quite up to standing in for a procurement employee, team, or department. However, if trained on 10 years of well-formed procurement data and paired with slightly more effective abilities to generate code and therefore interact with other systems, a GenAI model purpose built to 'do procurement' might be able to do just that: stand in for both a procurement platform and many of the highly skilled employees that makeup a procurement department.

These types of AI, sometimes called Industrial AI or even 'small' models are coming and will introduce yet another step change in value creation and risk profile. Imagine the hallucinations capable for a fully deployed network of procurement AIs: half the world's fungible supply of aluminium suddenly showing up in a field on the opposite side of the world would be a more impactful hallucination than an invented legal citation, fictional date, or HR policy detail.

And the level of change required to adapt to this type of machine intelligence both within a function, across functions, and even across supply chains will be significant. While we could tell the AI to hold on to traditional ideas such as Purchase Orders and Invoices, by training the AI to achieve Procurement outcomes, we would likely see a type of invention and transformation of what we consider standard human and corporate processes that achieve significant optimisations but are unrecognisable relative to today's processes and procedures.

The data and technical investment to build these kinds of models is significant and the skills required to train, deploy, and manage these models is still relatively rare, and distributed across firms in sectors who currently consume a lot of technology, sectors that produce the technology, and other sectors that help people use the technology. It may take 2 to 10 years for this type of investment to mature, and there are a series of problems to solve before these models become common, but this is in many respects what many senior leaders had in their heads when they thought about the impact of ChatGPT on their businesses. Managing the risks of this generation of powerful models will require very mature AI trust and safety programmes and will have regulatory hurdles to pass as new AI capabilities are made broadly available. Furthermore, the climate impact of the power consumed in the development of these models needs to be carefully considered.

A portfolio approach

The challenge faced by leadership within all types of organisations is not so much where to start, but more where to invest across these three models.

Each requires different skills and technologies, each could impact the competitive position of an organisation to very different degrees, and each introduces materially different risks into a business. The leaders in the AI revolution will take a portfolio approach that is laser-focused on how AI will help them achieve an organisational goal.

- **Agent Augmentation** is not only a good way to create value through automation, but to help an enterprise understand the cultural and organisational impact of this technology.

- **Enterprise Augmentation** may not deliver the promise of super intelligent AI, but deployed aggressively it will completely transform many functions across the enterprise, from Finance, HR and Risk to Product Development and Customer Interaction.

- What is also important to consider is that an organisation does not have to 'go it alone' when leveraging the potential that **Cognitive Engines** can bring; There are multiple ways, including alliances and partnerships, 'fintech' alignment and investment, government sponsored R&D, and data monetisation to participate in the value created by Cognitive Engines.

Getting started

There are three thing you can do now to accelerate AI change using this model.

First, if you don't have an AI strategy, build one based around the three types of AI Change. Your organisation could be at the cutting edge of agent productivity without taking on significant technical complexity and could drive significant value with a smaller footprint of enterprise transformation with a strong Enterprise Augmentation Strategy and must decide whether to get in on or sit out the investment to create Cognitive Engines. If you have an AI strategy, use this framework to test and establish how it will be implemented within your organisation and market.

Secondly, establish the governance, policies, and risk processes appropriate to the type of AI change you will focus on and determine, based on your strategy, the infrastructure, data, and technology partners required for the profile AI transformation you have decided to focus on for now. And finally prepare and get ready to deploy a broader transformation. No technology produces productivity gains without a parallel restacking of key processes, without changes to a workforce model, and without a change to the economics of the underlying business. With the right strategy, process, and transformation plan, we can move at pace toward the significant promise of value driven by AI Change.

![KPMG logo]

Get in contact

Paul Henninger
Partner, Head of UK Connected
Technology and Global Lighthouse

paul.henninger@kpmg.co.uk
+44 (0)7557 578 902

kpmg.com/uk