Strategic Approaches to Adopting Generative AI and Large LanguageModels



Co-created by Emerj Artificial Intelligence and Quantiphi

🖬 quantiphi

Solving What Matters



INTRODUCTION

Seemingly overnight, large language models and generative AI went from esoteric forms of artificial intelligence to the most easily recognized and business-changing technological developments of the last century, paving the way for AI capabilities to be embraced by popular culture worldwide.

Following the explosion in awareness after OpenAl <u>released</u> their ChatGPT interface in November 2022, allowing millions of users to interact with their GPT-3 language model, the world stood aghast at the remarkable capabilities of the system.

Instantly, educators decried the ease with which students could fabricate their assignments – often riddled with inaccuracies derived from ChatGPT itself. As the financial and legal industry revealed how the software could replicate sophisticated tasks in their workflows, a certain self-consciousness and fear for job security swept through both industries.

Not all the news coverage was as pessimistic, however. A widely cited report from <u>McKinsey</u> published in June 2023 made huge ripples through the mainstream media, touting a projected **increase of \$2.6-4.4 trillion in value for the economy generated yearly** thanks to generative AI capabilities.

Digging into the report for more nuanced details paints a different, if less searingly optimistic, picture of the seismic changes these technologies will bring to industries. Contrary to much industrial innovation over the last two centuries, the biggest impact of generative AI capabilities will be felt by the economy's highest-educated and highest-paying jobs.

Even advanced software hasn't always had this effect on what gets called "knowledge workers." The National Bureau of Economic Research <u>cities</u> that up to 70% of changes in U.S. wages over the last 40 years can be attributed to wage declines among blue-collar tasks replaced by largely pre-Al automation.

Still, the McKinsey report predicts that 75% of the value that generative AI use cases can deliver is expected to focus on four principal areas: **customer operations, marketing and sales**, **software engineering,** and **R&D**. By sector, it is anticipated that **banking**, **high tech**, and **life sciences** will be the industries most affected – banking itself could see an **increase of \$200-340 billion in value annually** if use cases are fully implemented.

In combination with other AI capabilities and activity, generative AI stands to automate work activities that currently preoccupies 60-70% of employees' time. While that sounds like a recipe for massive downsizing and decreased wages, sources such as the World Economic Forum are <u>optimistic</u> that these changes will be offset by an even greater job demand as early as 2025.

Whether the news is inspiring or daunting, the bottom line is that generative AI has changed – and stands to continue changing – our world drastically and in ways that experts, business leaders, and everyday workers are still trying to understand.



Emerj CEO and Head of Research Daniel Faggella recently spoke with Asif Hasan, Co-founder of Quantiphi, on the AI in Business podcast to discuss the best way for business and financial leaders to think of what these technologies will come to mean for the global economy over the next half-decade.

This article pulls insights from their conversation to shed light on the substance of the changes that generative AI will bring by:

→ Analyzing the workflow impacts of generative AI by dividing them into:

- Task-level disruptions.
- System-level disruptions.
- → Providing a closer look at significant generative AI use cases.
- → Provide holistic criteria business leaders can use to judge their capacity for driving ROI at their organizations.

This article will use the same definition of generative AI from the McKinsey report: applications typically built from foundation models. Our analysis will focus on generative AI more broadly while making key distinctions about the capabilities that large language models alone provide in specific use cases.

WORKFLOW IMPACTS

Task-Level Disruptions and System-Level Disruptions

Asif insists the fundamental change that generative AI will bring to knowledge workflows is bringing down the marginal cost of performing a cognitive task sustainably in the coming years, *"just the way Moore's Law was cutting the cost of compute by a factor of two every 18 months."* He believes that the marginal cost of performing cognitive tasks will come to be seen by business leaders in the next decade as *"the most basic unit of human productivity."*

So why is this happening now? Because, even as machines evolved over the last decade to detect physical objects with telemetric sensory technology, understand language, and recognize patterns, Asif tells Emerj that a model would still need to be custom-built to perform a particular task using supervised learning in the pre-generated AI paradigm – both a cumbersome and expensive process.

Generative AI shows that organizations can pre-train an LLM to perform fundamental tasks, like predicting the next token in a word sequence or pixel in an image. Once those functions are built into the foundational model, *"you can use this as a substrate in an AI system and fine-tune this model inexpensively to perform a huge variety of tasks,"* Asif tells Emerj.

"And this used to be both cumbersome and expensive. Now, what generative AI has shown is that you can pre-train a large language model to perform a fundamental task, like predicting the next token in a sequence of words or predicting the next pixel in an image. And now, when you use this foundation model, you can use this as a substrate in an AI system and fine-tune this model inexpensively to perform a huge variety of tasks."



Asif Hasan Co-Founder of Quantiphi

Looking at the next two years of development, Asif tells Emerj business leaders will begin to see many examples of task-level disruptions - or, as he defines the term, "to use the human plus Al combination to perform the exact same task that a human agent is currently performing but just do it better, faster, cheaper." Examples of this work include customer service agents following call scripts and responding to customer requests through a knowledge base that must be verified by paralegal review of commercial contracts and developers writing code to fulfill specifications and user stories.

However, Asif emphasizes that a much more significant impact on ROI will be seen - not at the task level - but in system-level disruptions. As opposed to using AI to complete tasks, organizations routinely conduct "better, faster, cheaper" disruptions at the system level and combine "AI capabilities in novel ways to fulfill the same customer needs, but in entirely new ways."

A potent example of a system-level disruption is going from needing studio distribution support, millions of dollars, and a team of hundreds to make a Hollywood-caliber feature-length film to someday only needing ten to 12 highly creative people using text-to-video models to create quality movies and TV shows.

Navigating the Transition from Task- to System-Level Disruption



Pulling from the work of Ajay Agrawal, Joshua Gans and Avi Goldfarb, Asif notes that the single most significant catalyst technology in enabling system-level disruptions before the emergence of generative AI was the smartphone. Through the expansion of mobile technology following the debut of the iPhone in 2007, companies like Uber were able to subvert the infrastructure of entire industries (in the case of Uber, transportation mobility) by funneling them through the power of devices in everyone's pocket.

It's difficult to anticipate the next iPhone and what that might look like, but Asif points to the dynamic between "attackers and incumbents" in the business technology ecosystem. The attackers, backed by VC institutions, will try to overcome legacy institutions' many strategic advantages. As financial services leaders are particularly well aware, often these advantages lie in primarily already possessing massive amounts of data, however unstructured.

USE CASES DRIVING ROI IMPACT

Criteria for Assessing Use Cases

When asked to cite specific use cases in system-level disruptions, Asif unveils criteria that enterprise leaders must consider regarding adopting Large Language Models (LLMs) or generative AI. The following is a breakdown of each criterion:



→ Where AI matches or exceeds human performance

For example, replying to emails or creating presentations.

→ Repetitive and costly tasks like customer service

Compared to outsourcing, the idea is to identify functions that can be performed better, faster, and cheaper with generative AI.

→ Prioritize tasks that can be transferred into clear written instructions for automation Such as reading and summarizing the financial results of a public company from their regulatory filings.

Key Considerations for Adopting Generative AI

When asked to cite specific use cases in system-level disruptions, Asif unveils criteria that enterprise leaders must consider regarding adopting Large Language Models (LLMs) or generative AI. The following is a breakdown of each factor:

→ Understanding the new generative AI paradigm:

Leaders should understand the difference between generative AI and the older paradigm of task-specific models. Generative AI allows for pre-trained foundation models that can be fine-tuned to perform multiple tasks, making it more accessible and cost-effective.

→ Identifying business implications:

Leaders need a clear perspective on how generative AI can impact their value chain and business model. They should explore how it can improve customer interactions, reduce operating costs, enhance productivity, uncover hidden risks, and enable the development of new products and services.

→ Building, deploying, and managing AI systems:

Leaders should consider the necessary investments, balance external expertise with inhouse talent, choose technology partners, determine the mix of proprietary and opensource technology, and foster a culture of curiosity and experimentation within their teams.

→ Data requirements:

Generative AI differs from task-specific models regarding data needs. Leaders should ensure their organization's data is accessible for AI systems to use as context. They should also curate and manage fine-tuning data, which is crucial in improving the quality of AI responses.

→ Role of champions:

Identifying individuals at the departmental or enterprise level who believe in the power of AI and are willing to champion its adoption. These champions should support experimentation, learning, and evolution while understanding the business context without relying solely on ROI models.

→ Ethical considerations:

Leaders must prioritize transparency, fairness, and accountability in deploying generative AI. Doing so includes preventing misuse, establishing governance frameworks, and ensuring AI's responsible and appropriate use. It's essential to view generative AI as a tool to augment human productivity rather than replace it.

Typically Outsourced Task-Level Use Cases

He cites common use cases with broad application across knowledge-based industries that meet these criteria, including:

- → Replying to emails
- → Drafting documents
- → Creating presentations
- → Answering questions from scanning data kept in spreadsheets
- → Customer onboarding in coordination with KYC compliance requirements
- → Analyzing 10k forms and other business analytics in the capital markets domain

Asif describes these use cases in a basic category of "enabling knowledge workers to navigate their professional lives more efficiently" at the task level. He cites that we're already seeing these functions being built into programs like Office 365 Copilot and Duet AI for Google Workspace, but largely feels that "these applications will start permeating our lives almost by default without any major programmatic changes happening within organizations." That said, organizations will need to train and enable their workforce to learn how to prompt these systems to harness productivity gains.

He continues to note that not all of these tasks will be fully automated, but rather only the document processing aspects of these tasks can be streamlined by AI, providing changes to the overall process that will be felt in the system. Examples of these tasks include verifying customer information from mortgage applications or business analysts highlighting important sections of a company's 10k forms in the capital markets domain.

Use Cases Automating Horizontal Business Processes

Asif describes a second category of use cases from generative AI that will derive from automating *"horizontal business processes, such as the customer service function."*

As many firms across financial services have come to realize over much of the last decade, Al capabilities are a driving force in helping turn call centers from cost centers to profit centers. Where decades ago, the role of call centers was to help organizations avoid likely negative customer feedback and costly requests – today, data tools are helping drive satisfied customers into more profitable lines of new business.



Asif describes at length how, up until this point, AI capabilities in the call center have helped to dissolve silo walls and help satisfy customer requests by putting the organization's knowledge base at the fingertips of call agents. He notes that generative AI, in particular, stands to enhance these capabilities, helping call agents to voice more articulate upsell appeals and better organize customer data on agent dashboards.

He feels that largely the call center and accounts payable work traditionally outsourced to foreign locations can be done by generative AI faster and with greater cost efficiency.

He feels that largely the call center work traditionally outsourced to far off locations can be largely done by generative AI faster and with greater cost efficiency.

System-Level Use Cases

The third and most important category of use cases Asif defines as those coming from "critically examining long and expensive steps within the value chain of an industry." Explaining how he applies the criteria as mentioned earlier in a drug discovery context, Asif highlights an example of such a use case from the life sciences sector:

"When you get through the process of first discovering a target, then screening the molecule, then getting to a hit, and then getting to the most promising compound followed by tests to check the safety of the compound – all of this happening in a matter of months, not years. So across many industries, there are similar examples of long, expensive steps in the value chain that, I feel, are ripe for some acceleration with generative AI."



Asif Hasan Co-Founder of Quantiphi Asif suggests that generative pre-trained architectures have the potential to impact any sequence that can be organized. He proposes training a language model to predict the next action based on the user's journey. He believes that exploring such areas and applying generative pre-trained models could enhance the ability to predict future actions or timestamps.

As outsourced tasks are usually a sign they can be easily automated, Asif tells Emerj areas of value chains in industries like financial services, healthcare, and life sciences that heavily focus on research and development are ripe for system-level use cases in generative AI.

While the language of life sciences isn't exactly conversational English, Asif notes it will be large language models – trained specifically in the nuances of coding in genomes, compounds and proteins – leading the way in streamlining these tasks.

With the pre-AI drug discovery process costing \$1-2.5 billion and taking over five years to complete, one can only imagine the effects that decreasing the time spent on incremental tasks like ensuring compound safety down to months will have on the overall timeline and costs.

Works Cited

- 1. Agrawal, Ajay, et. al. Power and Prediction: The Disruptive Role of Artificial Intelligence. Amazon, 2023.
- Autor, David, et al. "Tasks, Automation, and the Rise in US Wage Inequality." National Bureau of Economic Research, NBER Working Paper No. 28920, 2021, <u>https://www.nber.org/papers/</u> <u>w28920</u>.
- 3. "Don't fear AI. It will lead to long-term job growth." World Economic Forum, 26 Oct. 2020, <u>https://www.weforum.org/agenda/2020/10/dont-fear-ai-it-will-lead-to-long-term-job-growth/</u>.
- 4. McKinsey & Company. "The Economic Potential of Generative AI: The Next Productivity Frontier." McKinsey Digital, 14 Jun. 2023, <u>https://www.mckinsey.com/capabilities/mckinseydigital/our-insights/the-economic-potential-of-generative-ai-the-next-productivityfrontier#introduction</u>.
- 5. "The inside story of how ChatGPT was built from the people who made it" MIT Technology Review, 3 Mar. 2023, <u>https://www.technologyreview.com/2023/03/03/1069311/inside-story-oral-history-how-chatgpt-built-openai/</u>.



ABOUT QUANTIPHI

Founded in 2013, Quantiphi is an award-winning Al-first digital engineering company driven by the desire to reimagine and realize transformational opportunities at the heart of business. We are passionate about our customers and obsessed with problem-solving to make products smarter, customer experiences frictionless, processes autonomous and businesses safer.

quantiphi

^{Visit} quantiphi.com

Contact *quantiphi.com/contact*

ABOUT EMERJ AI RESEARCH

Emerj Artificial Intelligence Research is a market research and advisory company focused exclusively on the business impact of AI.

Companies that thrive in Al disruption run on more than just ideas. They leverage data and research on the Al applications delivering return in their industry today and the Al capabilities that unlock true competitive advantage into the future - and that's the focus of Emerj's research services.

Leaders in finance, government, and global industries trust Emerj to cut through the artificial intelligence hype, leverage proven best-practices, and make databacked decisions about mission-critical priorities.



