



Distant Field Labs

The Great Augmentation

Adopting Large Language Models in the Workplace

Public Release - June 2024

Exploring horizons

<https://distantfield.space>

thelab@distantfield.space



Distant Field Labs is pleased to present this report licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license.

If you would like to use this material for commercial purposes, please contact thelab@distantfield.space.

For more information please visit:
<https://creativecommons.org/licenses/by-nc/4.0/>



Introduction

On the 30th of November 2022, a little-known Artificial Intelligence firm called OpenAI presented to the world the latest version of their AI known as a Generative Pre-Trained Transformer, or GPT for short. They did so via a small chatbot called “ChatGPT”.

What has followed since has been nothing short of astounding. Not just for the growth and funding of OpenAI and its competitors, but computing and humanity in general. ChatGPT was fast, accessible, and showed the world what is possible with a large, well-trained Artificial Intelligence which understands our language.

It ushered in the age of Large Language Models and Generative Artificial Intelligence.

With this report, Distant Field Labs undertakes a high level look at Large Language Model (LLM) offerings for workplaces, with the aim of understanding the proposed benefits of LLMs and assisting decision makers in determining if LLMs are useful for their organisation.

We will look beyond the hype, examine vendor promises and the public apprehension, and consider how technologies such as Microsoft Co-pilot, Google Gemini for Workspace and OpenAI Enterprise can impact teams.

It is not the labs intent to do any kind of bake off to determine which product or service is better than any other; and we maintain a vendor-neutral stance. Our research has determined that the leading providers of LLMs are fairly close to a photo finish in capabilities and, despite leapfrogging feature announcements, are all generally following a similar trajectory.

We hope you finish reading this report with a better understanding of LLMs and feel well-armed to determine if adoption is right for your teams in this incoming age of Artificial Intelligence.

“A computer would deserve to be called intelligent if it could deceive a human into believing that it was human.”

Alan Turing, Computing Machinery and Intelligence (1950)¹.



A common language

We have said it before, and will state it again - you shouldn't be expected to be a computer scientist in order to understand the capabilities of Artificial Intelligence.

With that said, it helps if we all have roughly the same understanding of often loaded terminology we encounter while looking at AI. Here are some of the terminology we will use throughout this report:

Large Language Models (LLM)

LLMs are a type of generative AI that process and generate human-like text based on the training from vast datasets of text.

They are capable of understanding context, generating content, and even completing tasks they weren't explicitly programmed to do.

Generative artificial intelligence

Generative AI (GAI) refers to technologies that can create content, such as text, images, videos, and music, that resembles human-like creations.

It learns from a dataset and then generates new, original content that is similar but not identical to what it learned.

Narrow AI

This refers to AI systems that are designed and trained for a particular task (such as voice or image recognition) and cannot perform beyond the set parameters or tasks it was designed for.

Artificial General Intelligence (AGI)

Artificial General Intelligence is a form of AI that can understand, learn, and apply its intelligence to solve any problem, much like a human.

Unlike Narrow AI, which is designed for specific tasks, AGI has the ability to perform any intellectual task that a human being can.

Human In The Loop (HITL)

This is a model of interaction where human judgement is integrated into the AI operational loop, requiring human intervention for training, tuning, and testing the AI system.

It ensures that AI decisions are overseen or guided by humans, combining human intuition with AI efficiency.



A little (AI) learning is a dangerous thing

Artificial Intelligence is not new. Since the late 1940's neuroscientists, psychologists and computer science researchers have worked to understand how humans think, and if it is possible to have computers 'think' on our behalf. Recent advances in hardware and software have simply made AI more accessible to everyone.

In 2017, Google's research paper "Attention is All You Need"² introduced transformer architecture. This innovation, alongside advancements in data availability and hardware, propelled Natural Language Processing (NLP) to the forefront of computer science and laid the foundation for the major AI breakthroughs of the 2020s.

What transformer architecture enabled was continued evolution of NLP capabilities and triggered exponential growth of the complexity that can be handled with modern day compute resources. This has enabled a new era of AI which dominates our headlines today. Systems such as OpenAI's GPT³, Google's Gemini⁴ and Anthropic's Claude⁵ are evolving exponentially and creating new markets.

What stands these AIs apart is their ability to go beyond processing complex amounts of information, to the generation of new content.

If we consider previous compute paradigms, our systems are generally built around the concept of information retrieval. Whether a search box, a database query or a list of news articles, we are used to retrieving information and processing it. Often, we need to refine the list of information we are retrieving so we can filter to the context relevant to us. As humans, we curate data to our liking.

Large Language Models are now significantly disrupting this paradigm as AI now seemingly understands the user with more context. It can curate and present information with the context and understanding of why the user requested it in the first place. LLMs can now also create content (text, code, imagery and video) that has never existed before. This ability via LLMs is known as Generative Artificial Intelligence (GAI).

As with all significant technological advancements, it is not without its teething problems. You may have heard how LLMs can confidently present information in the incorrect context, or generate content that is wrong or entirely made up (most often referred to as 'hallucinations').

This is only natural as the new technology is still evolving, however, the message from the market, investors and systems vendors is clear - Large Language Models are here to stay⁶. A new era is upon us and the future of computing appears to be generative.

Artificial General Intelligence

Artificial General Intelligence (AGI) is a theoretical form of AI where a machine would possess the ability to perform 'thinking' on par with a human being. That is, it can understand, learn, apply and modify its knowledge across a wide range of domains.

It has been written and hypothesised⁷ as to whether or not we have reached AGI with current gen AI and it is a complex subject. This is in part, because some foundation models can appear smarter than humans for some tasks.

The evolution of LLMs is likely to play a crucial role in the eventual development of AGI. By pushing the boundaries of language understanding and generation, LLMs are helping to refine the algorithms, data processing techniques, and learning mechanisms that could be the foundations of AGI.

Moreover, the challenges encountered in the advancement of LLMs (ethical use, limitations of model outputs, and managing the societal impacts of automation) mirror those that will need to be addressed as AGI becomes a reality.



Launch not beyond your depth

During our research, we've been surprised by the speed and aggression with which vendors have approached AI, and also at the alarming number of companies now claiming to be "Powered by AI" when they are, in fact, anything but.

Because AI is the buzzword currently synonymous with bleeding-edge technology, efficiency and new ways of working, being 'driven by AI' has become a highly desirable tagline for many companies that traditionally would have monitored and not acted.

Yet for all the hype and swift adoption, we are still learning to understand what is possible with LLM assistance.

And more importantly - from what is possible, what makes sense.

Significant capital and effort has been deployed in advancing LLMs over the last few years⁸. The backers of this new technology understand its value (in both capabilities and financial returns) and are keen to realise the result of these investments with large scale adoption.

The release of ChatGPT created an initial demand for workers who can use these technologies for daily business operations. The market response to this demand has been swift. OpenAI, Microsoft, Google - amongst many others, have all gone to market with enterprise LLM capabilities under the guise of modern workplace technologies.

Common enterprise solutions on offer for organisations include:

- OpenAI ChatGPT Enterprise⁹
- Microsoft Co-Pilot¹⁰
- Google Gemini for Workspace¹¹

Each of these capabilities make promises on how they may augment end user workflows and reduce toil to improve performance and staff happiness.

The general narrative is that users will be enabled to generate content (e.g., new documents, spreadsheets), summarise data (e.g., previous meeting notes, sales data) and augment existing productivity tooling (e.g., email and presentation development).

This is all well and good, however it is key to remember that these are early days. Most workplaces do not move as fast as technology develops. They resist dramatic change in work practices, as these tend to go hand-in-hand with unforeseen complications, and resistance.

Given these initial stages of AI, measuring how many of the sales promises are being met is difficult. To be fair to the vendors, clear use-cases are emerging and the data is interesting, some of which is backed by reliable research.



Hope springs eternal

The application of LLMs can be wide and varying but there are some core trends. Below we summarise some of the common reasons for adoption of LLM augmentation across enterprises:

Increased efficiency and productivity

LLMs may automate routine and time-consuming tasks (such as meeting notes), allowing employees to focus on higher-value work. This not only speeds up workflows but also reduces human error, leading to higher quality outcomes.

Data-driven insights

LLMs ability to process and analyse data is often faster than human capabilities. Organisations may use LLMs to uncover insights from data that inform decision-making, identify trends, and predict future market movements, enabling more strategic planning.

Starting from zero

LLMs can help staff move past the 'blank page problem' assisting in the development of new documents (such as proposals, project plans and draft emails) while reducing procrastination in the workplace.

Enhanced customer experiences

AI-driven tools like chatbots provide 24/7 customer service, personalised recommendations, and instant responses to inquiries, raising the standard of customer interactions and satisfaction.

Innovation and competitive advantage

By leveraging LLMs for product development, market analysis, and operational efficiency, businesses can innovate faster, responding to changes in the market and consumer demand more swiftly than competitors.

Scalability

LLMs can easily scale with a business's growth, handling increasing volumes of tasks or data without the need to proportionally increase human labour or resources.



If vain our toil, we ought to blame the culture; not the soil

In addition to generally improving worker productivity¹², the adoption of LLMs is in part due to promises of toil reduction for teams. At the lab, we define “toil” as “tasks which may be repetitive and laborious for employees to undertake over a long period of time”.

A look at LLMs in the workplace identifies a number of areas where toil is potentially reduced or eliminated entirely. Each organisation is unique, but common areas of toil reduction the lab has identified includes:

Enhancing creativity and brainstorming

LLMs can assist in the creative process by generating ideas, suggesting improvements, and overcoming writer’s block, enriching the brainstorming process and product development.

Improving research efficiency

LLMs can summarise articles, extract key information from large documents, and conduct literature reviews, reducing the time and effort required for research-intensive tasks.

Automating routine communications

From drafting emails to generating reports, LLMs can automate routine writing tasks.

Optimising workflow and decision making

By integrating with workflow management tools, LLMs can prioritise tasks, suggest optimisations, and even make data-driven recommendations, aiding in more efficient decision-making processes.

Personalised learning and development

LLM-driven platforms can offer personalised learning experiences for employees, identifying skill gaps and recommending tailored training modules, enhancing professional development and performance.

Of all the promises we’ve seen to date, the main points are that LLMs will be used to augment human capabilities, automate mundane tasks, and provide actionable insights, leading to more efficient, innovative, and competitive businesses.

And if we’re not using LLMs in the workplace, we’re being told that our competitors most likely are, or will be very soon.



The point where sense and dullness meet

Decision makers currently face a tsunami of sales material which they must wade through and make informed decisions from.

Due to this vendor barrage, there is a significant challenge in determining the effectiveness of LLMs in the workplace, as the hype drowns the fact that we're still at the dawn of LLM adoption, and instead tries to sell AI as a fully-developed solution.

Many of the vendor-initiated reports highlighting significant organisational benefits and employee happiness are from limited studies and furthermore, often the subjects of those reports are early adopters within private programs, and therefore already supportive of the vendor in question. There are significant biases to work through in available research, and vendor-agnostic reports are more limited in both number and underlying research data.

However, we do have some insights emerging from the early development of the LLM enterprise market.

For example, according to McKinsey, Large Language Model application spans across 16 business functions, potentially delivering \$2.6 trillion to \$4.4 trillion in economic benefits annually when applied across industries. Notably, customer operations, marketing and sales, software engineering, and research and development stand out, accounting for approximately 75% of the total annual value from generative AI use cases¹³.

In customer operations, for instance, LLMs can improve customer and agent experiences through digital self-service and augment agent skills, leading to more efficient issue resolution and potentially reducing human-serviced contact volumes by up to 50% in certain sectors.

One study highlighted by MIT Sloan found that when LLMs are used within its capabilities, it can boost individual contributor performance by up to 40%. This improvement is especially notable in tasks designed to fall within the AI's capabilities, such as creating marketing content or performing data analysis¹⁴.

Similarly, research conducted in collaboration with the Stanford Digital Economy Lab found that access to AI assistance in a call centre increased agent productivity by 14%, with the most significant LLMs observed among less experienced workers.

The AI assistant helped disseminate the tacit knowledge of more experienced workers, aiding new workers in climbing the experience curve more rapidly¹⁵.

Case Study:

Who is gaining most?

In an interesting study from April 2023 (revised November 2023), researchers Erik Brynjolfsson, Danielle Li, and Lindsey R. Raymond conducted a study to explore the impact of new AI technologies on workplace performance, particularly focusing on a generative AI-based conversational assistant implemented for customer support teams¹⁶.

The study utilised data from 5,179 customer support agents to evaluate the productivity changes following the tool's introduction. The findings revealed that access to the AI tool led to a 14% average increase in productivity, measured by the number of issues resolved per hour.

Most notably, the tool significantly benefited novice and lower-skilled workers, who saw a 34% improvement in their productivity, whereas experienced and highly skilled workers experienced minimal impact. The study suggests that the AI tool effectively shares best practices among workers, accelerating the learning curve for newer employees.

Additionally, it was observed that the AI assistance positively influenced customer sentiment, boosted employee retention, and facilitated worker learning.



Where angels fear to tread

When we began our AI focused research here at the lab, we knew we were stepping into a hype cycle, however, the level of focused and often aggressive marketing for AI workforce solutions has been surprising.

As more companies enter the market with AI capabilities, we have entered a new wild west. Many may find this reminiscent of the wider internet environment prior to the dot-com bubble burst of the early 2000's. The rush for AI superiority and identifying the 'killer' use cases is on, and some days it feels like everyone is in the race.

The problem with this cycle is that marketing largely sells ideas right now. Most, if not all AI solutions, are in early phases and under significant active development. The features are incomplete. The resulting output is unreliable. The legal precedents have not been set. The marketing is so aggressive that it's hard to understand what is functional and available today as opposed to a future feature release.

It's hard to imagine a time where the market may have been so full of incomplete or "minimal viable products" as it is now. This can be taken as yet another indication on how transformer architecture, AI-based computing, is here to stay. The vendors are banking that they will finish the job.

When procuring LLM capabilities for your business, it is important to acknowledge that they are all in roughly the same early stages of development. AI vendors are currently competing for licensing commitments for when the full capabilities of their systems will be realised in the future, not for today.

It is true that there are some benefits to be realised during early adoption but for the most part the current state of LLMs represents a cutting-edge technology that carries a futuristic charm rather than battle-tested and fully-realised enterprise solutions for the workplace. Marketers leverage this futuristic allure to attract attention and investment by promoting AI as a transformative solution for businesses.

The potential conflict we are seeing unfold in real time is that customers of workplace-focused LLMs are seeking to increase efficiency, reduce costs and innovate today - not in the long term when the vision from vendors may be completely fulfilled.

It is for this reason that we strongly urge those looking at adoption of LLMs within their workplace to do so with a mid-long term (2-5 years) planning view.

In taking a holistic view when considering the actual proficiencies of LLMs, the potential benefits for teams along with the aggressive marketing activities and disinformation - decision makers can be forgiven for thinking they have two choices: adoption or obsolescence.

We don't believe the choice is as black and white as presented. Reinforcing the fact we are in the early stages of the age of Artificial Intelligence is best practice at this time.

So far, there have been very few court cases examining the operational and financial impacts of flawed or inappropriate decisions influenced by AI. Future scenarios, including significant pushback from customers against the widespread adoption of AI, are still possible.

The more industries collectively rely on AI backed work, the more cautious they may get and this is likely to trigger a shift in understanding of what is acceptable in their environments.

Such a scenario would inevitably force a market correction, whereby only AI solutions that provide clear and tangible benefits would survive. Vendors would need to adopt more customer-centric approaches, being able to demonstrate the actual value and safety of their offerings, and involving potential users in the development process to ensure that AI technologies meet real-world needs and ethical standards.



And it's not just collective customer demands which may change the landscape. We may also see significant pushback from governments and associated regulation.

It is true that AI development outpaces government policy development but this may not be the case for long. Public sentiment and concern is driving action, and significant effort is being undertaken to understand what appropriate and safe AI looks like, and how frameworks may be implemented to ensure appropriate use in the future.

Regulatory or industry backed challenges may lead to a more balanced and thoughtful approach to AI development, marketing and implementation. It is our belief that the current marketing strategy depends on the following key factors:

1. Companies need to be/remain convinced that they must have LLMs augmenting workflows in order to remain competitive;
2. That the end-game vision is possible today with enough investment;
3. It is going to be difficult to procure non-AI backed services in the near future so therefore best to start procurement and associated activities (such as staff training) today.

The success of this strategy to date indicates that many organisations either optimistically embrace AI solutions, or feel like they have very little choice in the matter.



Unfinish'd things

So far, we have examined the promises of LLMs in the workplace and while we believe there are significant benefits for adoption, we have also advised caution on accepting all of them at face value. The other significant dimension to consider when it comes to the procurement of Large Language Models is risk.

The details of how you implement LLMs within your environment significantly influences the risk profile. For example, adopting Microsoft Co-Pilot as part of your M365 tenancy likely carries less risk than an in-house developed LLM training on customer data (due to pre-existing security controls, training processes etc.).

For the following section, we wanted to touch on what we identify as universal risks for the adoption of LLMs in the workplace. Some of these are inherent in the early adoption of any significantly new and evolving technology stacks, while others are a little more specific to the nature of Artificial Intelligence.

Data privacy and security

Impact: High

In order for LLMs to be effective within a given workplace, the underlying model must be trained on the specific data from the enterprise. Depending on the approach for training, and how data is accessed, LLMs may inadvertently train on sensitive or compartmentalised data and therefore accidentally introduce data breaches in their output.

Either through adversarial user interaction (prompt injection¹⁷) or through the nature of LLM response generation, it is possible sensitive information may be disclosed to users inappropriately.

For industries with strict client confidentiality requirements, such as the legal industry, this can have very real implications for client privacy, however even companies without the same confidentiality requirements need to be considering the cost of data breaches.

Bias and ethical concerns

Impact: Moderate to High.

Artificial Intelligence's achilles heel is bias, you'll see this word mentioned a lot.

When training a LLM with data, it is determining the relationships between data entities and developing an understanding of possible patterns. If the training dataset is misleading, or already has significant but unobserved bias, AI will expose these biases as a matter of course.

This can pose ethical concerns, particularly if discovered in people operations, customer service, or content creation - anywhere where biased responses could lead to reputational damage or legal issues.

Enterprises need to be vigilant to keep humans in the loop and may require additional layers of review for sensitive applications, and if any bias is exposed, take it as a great opportunity to evaluate internal policies.

Dependence on technology

Impact: Moderate.

LLMs can enhance productivity and increase innovation within organisations. This however, is a risk in itself. Given the current AI fanfare, there are very real risks that businesses may adopt LLMs into their day-to-day processes without appropriate checks, balances and preservation of institutional knowledge across teams.

It is critical that enterprises consider the balance of LLM use by keeping humans in the loop to leverage their empathy, judgement and expertise, and ensuring they do not over-rely on AI. You can read more about our thoughts on AI overreliance and cognitive atrophy as we [outlined in the productivity paradox](#).



Regulatory and compliance risks

Impact: High.

Enterprises are required to address many regulatory and compliance checks across both their own individual business governance activities and industry requirements. These include data privacy and management (such as CCPA or GDPR), cybersecurity and sovereign data.

Given the freshness of LLMs, there are many unresolved potential issues with how regulatory compliance is impacted by the adoption of these capabilities to support various activities.

Care must be taken to ensure that teams engage **early and often** with their governance and auditing peers to foresee any potential issues with regulatory compliance after LLM adoption. Care should also be taken to ensure that any liability insurance policies are reviewed and understood prior to adoption.

We're all being told to keep an eye on legal, regulatory and policy frameworks regarding AI and LLMs for valid reasons. Early adopters run the risk of implementing available technology and then having to heavily restrict use depending on future regulations.

For a current view of regulations currently being put in place, a good place to start would be the following links, particularly the incoming EU regulations.

[Blueprint for an AI Bill of Rights \(US\)](#)

[OWASP LLM AI Cybersecurity & Governance Checklist](#)

[Europe becomes the first major world power to enact comprehensive AI regulations](#)

[G7 to agree AI code of conduct for companies \(voluntary\)](#)

This will be a work in progress and already questions are being asked regarding the lag in getting any form of consensus; however technology is moving at an accelerated pace and regulations always move slowly and carefully so this is to be expected.



To err is human; to mitigate is .. better human

Many of the risks of AI & LLM adoption may be mitigated prior to becoming significant issues, depending on the regulatory landscape of the industry. However as mentioned in our previous section, regulations work slowly and most of us would rather be proactive towards safe AI than wait for it to be regulated.

Below we present some high level guidance on mitigating common risks associated with LLM adoption. The lab recommends consulting with your normal security and risk partners to ensure a complete and comprehensive understanding of appropriate mitigations.

Data management

Work with your privacy functions to determine AI training guidance and transparency. Implement strict guidance for teams on adoption of LLMs and any model fine tuning with enterprise data.

Training and awareness

Educate employees on the responsible use of LLMs, highlighting the importance of manual oversight and ethical considerations. Document use-cases where the enterprise is comfortable with its use, and limit interactions when privacy or security concerns have been identified.

Policy development

Develop clear policies and procedures that align with ethical guidelines, compliance standards, and best practices for AI governance.

Regular audits

Conduct regular audits of the LLMs implementation and staff members interactions with LLM. Monitor outputs to identify potential biases, misuse, or security gaps.

By understanding the risks associated with LLMs, decision makers can adopt a more informed and balanced approach to integrating these technologies into their operations, maximising their benefits while mitigating potential downsides.



Conclusion

It's hard to look at the future of computing and the integration of Large Language Models and not deem it inevitable within enterprise environments. It very much feels like a 'when' not 'if' scenario, driven significantly by vendors, financial incentives, employee and market demand.

While the adoption of LLMs promises substantial benefits, it is accompanied by some very real risks and challenges. The potential benefits to productivity, innovation, and customer engagement are compelling, but must be carefully balanced against the risks and ethical implications these technologies introduce.

Adopting bleeding-edge technology is a mixed blessing; done well it can usher in improvements and provide an array of benefits. Implemented poorly or in haste, and it can introduce significant risks that can dramatically impact the organisation and its reputation. The difference between the two is often the time, consideration and planning prior to implementation; which is in short supply these days particularly when under pressure from stakeholders or vendors.

To harness the full potential of LLMs while safeguarding against their pitfalls, businesses must adopt a proactive stance, with the full acknowledgement that it may be years before either the promised benefits or impending regulations are realised. This involves cultivating a well-informed corporate culture that prioritises continuous learning, ethical considerations, and robust data governance. Equipped with a comprehensive understanding and strategic oversight, decision-makers can effectively navigate this new terrain, making sure that their adoption of LLMs not only enhances competitive edge but also aligns with core business values and responsibilities.

Integrating Large Language Models into your business operations requires careful consideration. We recommend all companies considering LLM integration to take your time to understand and weigh up the potential benefits against inherent risks, resisting market pressures to hastily integrate new technologies at the expense of quality.



References

- 1 <https://academic.oup.com/mind/article/LIX/236/433/986238>
- 2 <https://arxiv.org/abs/1706.03762>
- 3 <https://openai.com/gpt-4>
- 4 <https://deepmind.google/technologies/gemini/#introduction>
- 5 <https://www.anthropic.com/claude>
- 6 <https://www.nvidia.com/gtc/keynote/>
- 7 <https://www.microsoft.com/en-us/research/publication/sparks-of-artificial-general-intelligence-early-experiments-with-gpt-4/>
- 8 <https://www.cbinsights.com/research/large-language-model-llm-developer-market/>
- 9 <https://openai.com/blog/introducing-chatgpt-enterprise>
- 10 <https://www.microsoft.com/en/microsoft-copilot/>
- 11 <https://workspace.google.com/solutions/ai/>
- 12 https://www.nber.org/system/files/working_papers/w31161/w31161.pdf
- 13 <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#introduction>
- 14 <https://mitsloan.mit.edu/ideas-made-to-matter/how-generative-ai-can-boost-highly-skilled-workers-productivity>
- 15 <https://hai.stanford.edu/news/will-generative-ai-make-you-more-productive-work-yes-only-if-youre-not-already-great-your-job>
- 16 https://www.nber.org/system/files/working_papers/w31161/w31161.pdf
- 17 <https://arxiv.org/abs/2306.05499>

