Distant Field Labs

Wielding the AI Hammer

Operationalising Artificial Intelligence for Executive Leadership

Public Release - April 2024

Exploring horizons

https://distantfield.space

thelab@distantfield.space

Please note: This report is an early showcase of
Distant Field Labs ongoing research in Artificial Intelligence.

We would appreciate feedback at **thelab@distantfield.space**

Introduction

As we enter 2024, a significant change in how industries leverage Artificial Intelligence (AI) is underway. It has already been called the Age of Artificial Intelligence and we believe it to be true.

With the launch of ChatGPT by OpenAI, humanity obtained a glimpse of what capabilities next-generation AI can deliver and solidified its place in the future of computing.

AI is not new and the current wave of AI advancements are the result of almost a decade of determined work by researchers and enterprises to democratise AI across business, government and education.

Given the speed of growth in Artificial Intelligence, we are confronted with the challenge of how to make sense of rapidly-changing capabilities, and how do we best leverage them?

Many decision makers will find these challenges familiar from times of significant advancements such as the rise of mobile applications or the adoption of public cloud. Both led to navigating significant technological change and having to evolve quickly to stay competitive in market.

As we approach exponential growth in capabilities with AI, we are finding executive leadership and investors having to navigate a tsunami of misinformation, vendor-slanted information or generally incomplete information about AI capabilities and which challenges it can help solve.

It is not the purpose of this report to raise concerns or discourage the adoption of AI. At the Lab, we are early adopters of AI and have leveraged it in our internal processes to aid analytical work.

Our goal is to help provide clarity for executive leaders who are looking to implement AI within their business or services.

Specifically, we are focused on Narrow AI (often referred to as Weak AI) as opposed to implementing workflow augmentation for staff using the likes of Microsoft Copilot or Google Gemini for Workspace.

We walk through key concepts and decisions to help ensure appropriate investments in next-generation capabilities that compel not only individual organisations, but entire industries forward in the coming months.

It is our hope that this report assists decision makers, investment leaders and executive teams in being informed as they work with teams to harness the immense potential of AI.

> "The development of AI is as fundamental as the creation of the microprocessor, the personal computer, the Internet, and the mobile phone. It will change the way people work, learn, travel, get health care, and communicate with each other. Entire industries will reorient around it. Businesses will distinguish themselves by how well they use it. "
>
> **– Bill Gates, The Age of AI has begun.** [1]

A common language

To begin, we must understand the key terminology that will be at play throughout the AI journey. Artificial Intelligence is complex and involves detailed computer and mathematical sciences to achieve desired outcomes.

As we don't have to be an auto-mechanical or electrical engineer to drive our car, we shouldn't be expected to be computer scientists or mathematicians when approaching the operationalisation of AI across environments.

It's important to understand some of the base terminology that will come up during the decision making process and the following are some key terms we will encounter throughout this report:

Model

A model is a mathematical representation of a real world process.

It's how AI predicts (decides) based on the provided information (data). Models are created by training with datasets.

Training

The process of teaching an AI model how to make predictions.

This is achieved by providing datasets to models to learn patterns, relationships and other key characteristics to achieve the desired objectives.

Dataset

Data which you would like to have the AI learn from.

This may be text, images, videos, code or any other type of data. It may also be a combination of different data types.

Foundational models

Foundational models are generally very complex models which have been pre-trained on large datasets.

OpenAI GPT and Google Gemini are two examples.

Parameters

A model's parameters are numerical values learned during training that influence its predictions.

By adjusting these parameters based on the training data, the model learns to identify patterns and relationships that allow it to make accurate predictions on new inputs.

Fine-Tuning

Is a technique to adjust model configurations to improve accuracy or performance.

This includes the model parameters, learning rates, and any other values which improves model accuracy.

Exploring the Modelverse

Proprietary or open source AI models

Like all technological adoption, AI has direct and indirect costs. One of the most significant influences on cost is the **model** being chosen for the given task.

AI models come in all shapes and sizes and with various complexities. Ensuring an appropriate model for the task is critical and it will influence all additional effort during the operational processes.

There are generally two types of models organisations must decide on at the architectural phase of a project: Proprietary (commercial) and Open Source (non-commercial).

Proprietary models

Many vendors provide proprietary models for use via their platforms. They are generally pre-trained off vast datasets and provided via Application Programming Interfaces (API's) for use by your development teams.

State of the art models, such as OpenAI GPT-4[2] or Google's Gemini[3] models have parameter counts in the trillions, meaning they have extreme capabilities when it comes to prediction activities.

These models are provided on a pay-per-use basis and most often via the vendors' cloud platforms and are easy to adopt for use in generalised tasks.

Open source models

Open Source models are models that have been released for wider distribution or research purposes. They are often for research purposes, but can be adapted for commercial use under the licensing. They are often trained on much less data, however are more flexible for use with custom datasets and fine-tuning.

Given Open Source models are generally trained on less data, this makes them popular amongst use cases where you would like to heavily customise or train the model for specific purposes. Open models are shared via online repositories such as Kaggle[4] or Hugging Face[5] and whilst lesser known, there are hundreds of thousands of different models to choose from to help achieve your objectives including content generation, object recognition and image processing.

Decision approach

As a business, you may end up using proprietary and open source models across projects depending on the requirements. Given the high rate of development, it's also important to ensure teams revisit their decisions around models on a regular basis.

Care should also be given to licensing restrictions of models. For example, many of the commercially developed and released open source models prevent their use in decision making processes for regulated industries such as finance, healthcare and policy development.

When approaching which models to use, we recommend keeping in mind a simple evaluation of the pros and cons.

	PROS	CONS
Proprietary Models	**Ease of Use** Ready to use with minimal setup or customisation	**Cost** Usage-based cost models can be expensive on large volume usage.
	Performance Strong baseline performance, optimised over multiple iterations by the provider.	**Customisation** Often limited ability to fine tune or customise the model for your specific use cases.
	Support Cloud Providers will provide commercial support for using their model.	**Vendor Commitment** Migration across API's may have increased complexity due to vendor lock in.
	Reliability Commercial support means longevity of the solution by the provider.	**Inherited Bias** Large datasets often are often not checked for bias and may contain significant biases which influence the output for consumption.
Open Source Models	**Cost** More accurate cost prediction across large scale use.	**Technical Expertise** Requires inhouse or third-party contracted expertise to appropriately leverage the models. Extended setup time required.
	Customisation Increased control over model customisation.	**Performance Uncertainty** Fine tuning a model for execution within your existing compute environment may increase the complexity of the project.
	Bias Transparency Visibility into the model and training datasets can help reduce biases.	**Limited Support** In-house, third-party or community expertise are relied upon for supporting the model.

We note an increasingly common approach is the concept of hybrid models. This is where organisations may start with Open Source models for prototyping activity and once proven, implement the final product via a more commercially supported model.

Navigating the ethical minefield

Where it can go wrong

AIs are complex and evolving systems which leverage a number of technologically advanced capabilities. Given this complexity, there's plenty that can go wrong within each system.

AI systems epitomise the concept of 'garbage-in, garbage-out' processing. When implementing AI to support both internal and external processes, ensuring strong understanding of the capabilities is critical. In the following section, we explore several example areas where AI introduces potentially new or unique situations where ethics may need due consideration.

Datasets

When an AI model is trained with data, it learns the relationships between data entities. It then attempts to recognise and understand patterns across entities to help predict a great response for the consumer of the system.

If the underlying training dataset is misleading, outdated or wrong, it may directly influence the output of the AI. However, there is also significant risk that the AI learns new patterns relating to the provided dataset which had previously gone unnoticed or unacknowledged.

Often, these relationships can be previously unidentified biases. Early adopters of AI have learnt the hard way that almost all organisations have biases and AI is very good at identifying them. This results in AI systems which can produce obvious biases in their output, which may have never previously been considered.

There are methodologies to mitigate biases within systems, including through the use of synthesised data. Care must be taken to ensure that any system biases are identified and understood during the development process and before production deployment.

In many instances, businesses will look to acquire datasets to train their AI with. This might include market information, near real-time data relating to a particular topic or historical datasets to help strengthen trends identification.

There exists a vast and significantly sized market of both commercial and Open Source data brokers who are willing to provide a variety of datasets to support business needs. Unfortunately, as with many industries that cross geographical borders, many of the data brokers may not be acquiring data via methods inline with what your organisation considers legal or ethical.

At the time of publication, whilst some capabilities exist to help organisations measure AI output for biases, there is little capability to support businesses in understanding the acquired datasets integrity.

It should also be noted that there have been instances where the dataset contained a significant amount of Personal Identifiable Information (PII) resulting in potential privacy breach instances[6]. Bringing privacy teams onboard during dataset development is strongly encouraged.

System output

As AI's are trained, they leverage advanced mathematical algorithms to identify patterns and relationships.

It is possible that the relationships identified may be non-desirable, reputationally-damaging or in some instances (such as Generative AI imagery), illegal.

Safe AI includes leveraging transparency (ensuring explainability of AI decisions) and alignment (ensuring the AI is aligned to its intended goals and only its intended goals) to ensure that the expected system output is not only understandable but expected.

Given the complexities associated with AI, it is strongly recommended that organisations consider adopting expertly developed frameworks (such as the Presidio AI Framework from the World Economic Forum)[7] to ensure safe, expected output from AIs.

It is also strongly recommended that organisations implementing AI leverage 'Red Team' capabilities[8], to test the AI from an ethical and brand reputation perspective to ensure no concerns exist with the output. For organisations who have previously undertaken Cyber Security Red Teaming, it should be noted that AI Red Team activities are significantly more complex and while sharing the same name, are quite different in nature and cost.

Decision processes

There are significant concerns amongst commercial and Open Source model providers regarding the use of AI in decision making processes. Many licence terms restrict the use of AI for automated decision making activity.

These generally include areas of finance, employment, healthcare, housing, insurance, social welfare or any other area where individual human rights or well-being may be impacted by the resulting decision.

If developing AI to assist with decision making activities within your organisation, it is currently recommended that human review processes are implemented to reduce risk to your organisation.

Adversarial users

We find ourselves reiterating the complexities in the underlying technological implementation of AI capabilities. True to all systems with plenty of moving parts, there are plenty of places where things can go wrong and issues introduced.

In some cases, the introduced issues may come from third-parties trying to influence the behaviour of the model and the output from the system. This can occur in two key ways - influencing the provided data for training (known as data poisoning) or manipulating the query to the AI itself (often referred to as 'prompt injection').

If it is possible to influence the AI in any way, this may lead to unexpected behaviour of the AI, specifically its output. There are known instances where malicious influencing of AI has occurred in production resulting in less than desirable stories on the front page.

Beyond direct influencing of AI for undesirable outputs or behaviour, the complexity and lack of understanding (or emotion) of AI (especially Generative AI) can cause bizarre, blatantly false and wrong output from the system. This is often called Artificial Intelligence Hallucination.

Case Study:

Accidential Bias: AWS's HR Platform

Between 2014 and 2018, Amazon attempted to automate its recruiting process with an AI system. The goal was to improve efficiency by using data to identify top candidates. However, it has been widely reported[10] that the project revealed a major flaw: the system became biassed against women due to the data it was trained on.

Because the majority of Amazon's existing tech employees were male, the algorithm learned to favour male candidates. It would downgrade resumes containing terms associated with women (e.g., "women's chess club"). The AI had unknowingly inherited historical hiring biases where women were under-represented in technical roles.

Amazon scrapped the project after failing to fix the bias. This case highlights the risks of using AI in decision processes without careful consideration, and emphasises the need for human oversight to ensure AI systems make decisions that are fair and unbiased.

The tip of the Priceberg

Considerations of the true cost of AI adoption

Between 2010 and 2020 we witnessed the exponential rise and adoption of public cloud services. Many organisations moved fast to migrate from their legacy on-premises environments to the cloud. While the benefits of cloud computing were obvious, we quickly learnt our cost prediction models and estimated cost savings were often incomplete and, as a result, budgets were routinely blown out and significant adjustments required.

With the adoption of AI, it is important we learn from the lesson of public cloud growth and digital transformation activities to ensure we have a robust understanding of the true cost of AI adoption.

It is too early to create general cost prediction models for the adoption of AI within organisations. However, there is good understanding and evidence of the areas where cost incursion may occur.

Below, we have captured these areas at a high level for consideration.

Economical

There are very real costs associated with the computational requirements of AI.

If using a proprietary model (i.e via cloud APIs), the pay-per-use unit approach may be difficult to predict or measure depending on the system being developed. In the case of Open Source models, the computational costs may also be difficult to predict accurately during early adoption of AI within your organisation.

Good AI, useful AI, is dependent on good data. The business generally either already owns datasets or obtained one for use within AI.

However, ensuring a clean dataset can be an unexpectedly complex problem. Ensuring cleanliness, normalisation and freshness of data for use within training may require significant time, resources and third-party assistance.

Care should also be taken to account for accurate maintenance costs for AI over the lifetime of the system.

If undertaking training and tuning of an AI model for use within the business or service, consideration should be given to the freshness of data. For example, if your business is fast moving, retraining or modifying your models may be required on a monthly or even weekly basis as opposed to yearly or once-off.

Legal and regulatory

AI has already disrupted a number of industries and raised very real legal considerations across the globe. In response to these considerations, there are significant moves underway to ensure safe adoption of AI occurs.

For example, the European Union recently passed the Artificial Intelligence Act[9] which focuses on ensuring 'fundamental rights, democracy and rule of law' is maintained as AIs are adopted across the Union.

If your organisation has compliance requirements in either government or industry regulation, the introduction of AI may significantly influence maintaining compliance costs.

Liability insurance is also an area where unexpected increased costs are occurring (in both premiums and specialised coverage). The insurance industry relies on historical data to help determine their coverage and given the freshness of AI and its use across business, this is proving challenging.

For organisations with Technology Errors & Omissions and similar liability insurance policies, it is strongly recommended that you speak with your insurance provider to determine if your coverage includes the use of AI.

Environmental

The incredible complexity of AI and its capabilities are enabled via significant compute resources and modern scaling capabilities. In 2023 alone, NVIDIA is understood to have shipped well over 500,000 units of their H100 GPU for use in AI to meet demand[11].

Entities with a focus on understanding their environmental footprint in pursuit of carbon neutrality may find themselves with significant compute or e-waste cost increases due to the significant change of profile in compute resources.

It should be noted that while many cloud providers (such as Microsoft and Google) provide sustainability reporting for compute resource usage in their environments, they may not currently include an accurate representation of AI consumption. Analysts believe that many AI capabilities (for both AI itself and AI enabled services) are being sold from loss-leader price positions[12] to increase adoption of services. There is a very real risk of inaccurate reporting and carbon sticker shock later if undertaking large-scale AI adoption across services.

Research and development

AI as a discipline is not new, however its recent broad adoption with generative AI has caught many by surprise. As a result there is a very real battle ongoing for talent with specialisation in AI skills[13]. This includes data warehousing and preparation, data scientists, developers and engineers with experience in scaling AI in production.

If undertaking an AI project, there may be significant costs involved during the talent acquisition process or similarly, premium costs charged by system integrators and partners operating across the AI ecosystems.

Another emerging trend across industry is the realisation that initial time and compute estimates are wrong at the start of projects. While there have been significant advancements in AI tooling and capabilities, there are still significant complexities when constructing AI capabilities which may increase project budget requirements over time as they are understood.

Social and ethical

Much has been written (and continues to be written about) the ethical implications of AI implementation across industry. These are very real concerns and there are already very real implications due to how poorly-developed AI can behave.

If your business or organisation participates in local industry or government programs which incentivise job creation and industry growth, these targets may be impacted by the adoption of AI technologies.

Ensuring ethical and clean datasets which prevent inaccurate (or embarrassing) predictions out of the AI, can prove a costly exercise. Paying for expertise and time to ensure a resulting AI trained on a specific dataset does not leak private information nor does it display inherent biases or deliberate inequality in its predictions may be critical for regulatory and public integrity activities.

The AI hammer

When witnessing the power of AI for the first time, it can seem almost magical and the natural response to magic is to wield it.

If AI is a hammer - every business challenge becomes a nail.

Care should be taken to ensure that existing technological processes and capabilities are not ignored when they may be sufficient to solve a given problem. Prior investments in technological capabilities may be sufficient to meet the particular needs of the organisation.

Ensuring clear project goals. Understanding exactly how AI will be leveraged and its benefits over existing or alternative methods is critical to avoid over-engineering.

Conclusion

Like many others around the world awakening to the vast possibilities of Artificial Intelligence, Distant Field Labs believes AI has the power to transform not only individual organisations, but humanity as a whole.

There are however, significant ethical and financial considerations when adopting Artificial Intelligence within your business services or processes.

Artificial Intelligence is becoming a vast and complex discipline across industry and it is difficult to draw robust and complete conclusions on activities that organisations should undertake during operationalisation, however we believe the following key points should be considered as you progress with AI journey:

01

There are significant pros and cons for proprietary and open source models.

Care must be taken to understand how a model is selected and if it is subject to restrictions under licensing agreements.

02

Ensuring clean, ethical datasets for your Artificial Intelligence use is critical to avoid undesirable situations resulting from your AI capabilities.

Ensuring datasets are free from biases and have been ethically sourced is important to ensure safe development and use of AI within your organisation.

03

Beyond obvious economic risks, there exists potential for significant hidden costs during AI adoption, development and deployment.

Care must be taken to ensure a full understanding of project costs including across legal, R&D and environmental areas.

04

The AI regulatory landscape is rapidly changing as exponential adoption occurs.

Ensuring you and your teams are aware of regulatory changes is critical to ensure safe longevity of AI adoption.

References

[1] Bill Gates Notes https://dfgo.io/AIH01

[2] OpenAI GPT-4 https://dfgo.io/AIH02

[3] Google Gemini https://dfgo.io/AIH03

[4] Kaggle https://dfgo.io/AIH04

[5] Hugging Face https://dfgo.io/AIH05

[6] Australian Government https://dfgo.io/AIH06

[7] World Economic Forum https://dfgo.io/AIH07

[8] Google https://dfgo.io/AIH08

[9] European Parliament https://dfgo.io/AIH09

[10] Reuters https://dfgo.io/AIH10

[11] Statista https://dfgo.io/AIH11

[12] The Wall Street Journal https://dfgo.io/AIH12

[13] Financial Times https://dfgo.io/AIH13