

+++

Generative AI (Gen AI) Trends

White Paper on Short, Medium &
Long Term Trends

Morgan Fitzgibbon



Generative AI(GenAI) Short, Medium & Long Term Trends

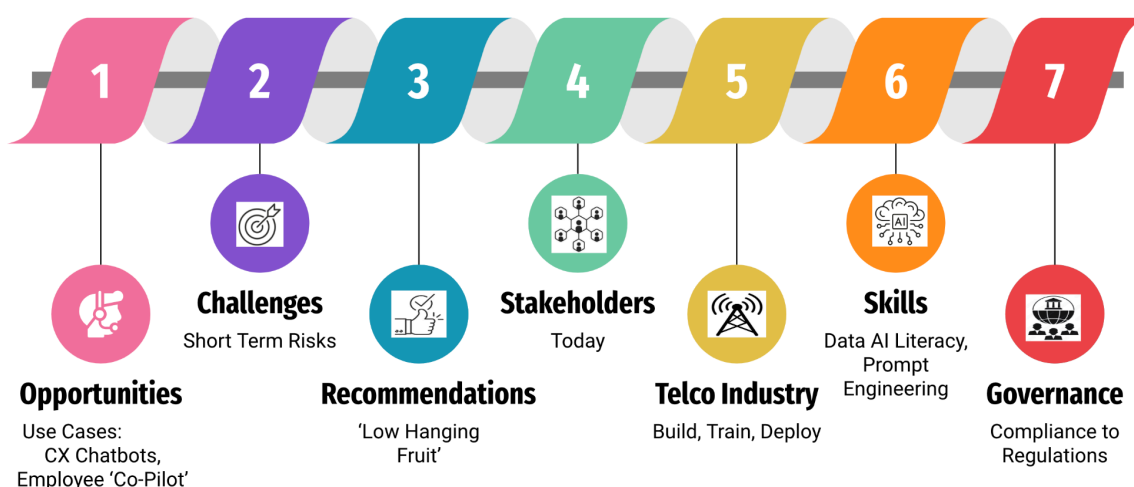


Table of Contents

Introduction to Generative AI (GEN AI)	3
Large Language Models (LLMs) :	4
Parameters:	4
In the Short Term What are the Opportunities?	6
What are the Challenges?	10
What are the Immediate Recommendations?	11
Today, Who are the Main Stakeholders in GenAI/LLM?	12
Medium Term Considerations: Build, Train, How to Deploy?	15
Proprietary versus Open Source?	16
Short to Medium term Public Cloud LLM usage by Telcos	16
Trend towards Smaller Specialised LLM?	17
Vodafone & Microsoft Use Cases	18
Global Telco Alliance between SKT, Deutsche Telekom,e&e, SingTel	20
What are the Skills Required?	21
Long Term Trend as LLMs become more mature ?	23
EU AI ACT :	23
Corporate Governance	26
How to address Data Privacy Concerns in AI Prompt Engineering?	27
Biography	29
References	30

Introduction to Generative AI (GEN AI)

What is Generative AI(Gen AI)?

Gen AI refers to any machine language model capable of dynamically creating output after it has been trained. GenAI works by processing huge volumes of data to find patterns and determine the best possible response to a question or situation, which is then generated as output. GenAI involves the use of foundational models - deep learning networks to drive new applications more quickly and cost effectively(see definitions below for more details).

Large Language Models(LLMs) are a type of GenAI that can understand and generate human-like language. With OpenAI's ChatGPT4 the most well-known example. LLMs are foundational models trained on massive data sets, with billions or trillions of parameters allowing them to learn complex language patterns and perform tasks that would be impossible for smaller models. There are many types of LLMs in existence today designed for a specific purpose, for example LLMs designed specifically for code trained using data from GitHub(an AI-powered developer platform that allows developers to create, store, manage and share their code.)and Bloomberg GPT(LLM for finance). Once a model has been developed, any developer can build applications on top of it to create content. For example, the main application today for ChatGPT4 is the ChatGPT chatbot.

What can GenAI do? GenAI has a deliverable, which makes it different from other types of AI. Gen AI can be applied to any content. In the short term, GenAI is a tool to help individuals and teams to carry out their work more efficiently using Chatbots. It has a user interface to enable users to interact with GenAI models and copilots are conversational interfaces that use LLMs to support uses in various tasks, such as writing documents and summarising emails. GenAI outputs can be created immediately but the challenge then is how much fine tuning needs to be done to improve the output. There are risks with GenAI models such as hallucinations - inaccurate or inappropriate responses to inputs. The real benefits of GenAI is when it is combined with classical AI such as conversational and enterprise search to make it easier for users to navigate websites.

Foundational Models :

General purpose AI models trained on massive data sets. Form the basis of ChatGPT4 application . Rather than develop AI from scratch, data scientists use a foundational model as a starting point to develop machine learning models that power new applications more quickly and cost effectively.

Large Language Models (LLMs) :

Are AI systems designed to process and analyse vast amounts of natural language data and then use that information to generate human-like responses. A LLM is any statistical model of language built on large data volumes and billions of parameters, LLMs are the basis for most of the existing foundation models.

Parameters:

are configuration variables that are learned during a machine-learning process. Such variables are used to control the behaviour of the model. LLMs have tens or hundreds of billions of parameters (GPT-4 is said to have 1.76 trillion). The number of parameters tends to be a measure of the size and the complexity of the model. The more parameters a model has, the more data it can process, learn from and generate. There is now a trend towards smaller, more specialised LLMs with fewer parameters and which are easier and cheaper to train.

Tokens & Tokenization : A token is a basic unit of text or code that an LLM uses to process and generate language. Tokens can be characters, words or other segments of text or code. Tokenization is the process of splitting the input and output texts into smaller units that can be processed by the LLM AI models. Tokenization affects the amount of data and the number of calculations that the model needs to process. The more tokens that the model has to deal with, the more memory and computational resources the model consumes.

Training : is the process of teaching an AI system to perceive, interpret and learn from data. Data scientists can spend years creating a new AI model and training it to perform complex tasks. In the first phase of a training process the model is fed massive amounts of data and is then asked to make decisions based on the information. The training part involves making adjustments to the model until it produces satisfactory results.

Fine tuning : a LLM involves making small adjustments to a pre-trained model to improve its performance in a specific task. This can yield better results than training a model from scratch as the model already works well and can leverage its existing knowledge to learn new tasks more quickly.

Prompt Engineering : This is the process of refining LLMs with specific prompts and recommended outputs, and of refining input to various GenAI services to generate text or images.

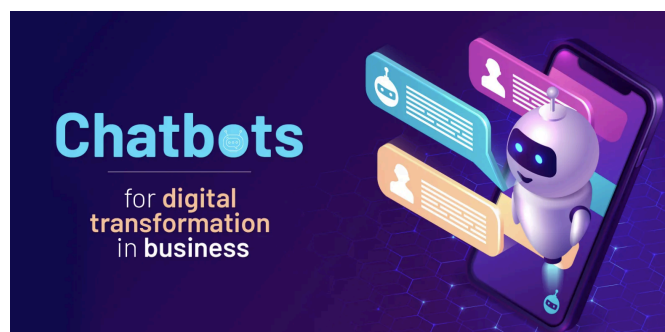
Hallucinations : These are inaccurate or inappropriate responses to inputs presented as facts, such as false or misleading information or race/gender bias. AI hallucinations are often caused by limitations or biases in training data and algorithms, and can result in producing content that is wrong or even harmful.

Retrieval Augmented Generation(RAG): provides a way to optimise the output of an LLM with targeted information without modifying the underlying model itself. The targeted information can be specific to a particular organisation and industry, so the GenAI system can provide more contextually appropriate answers to prompts.

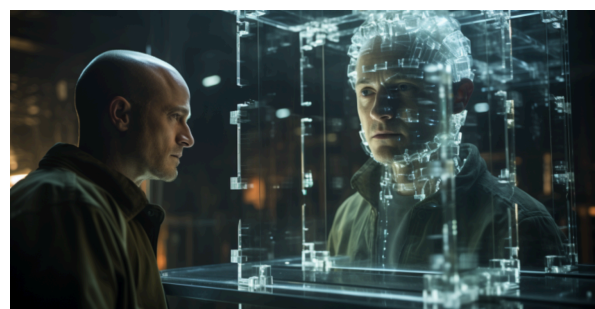
(sources: TMForum GenAI Insight, AWS, Ada Lovelace Institute, IBM)

In the Short Term What are the Opportunities?

Despite Generative AI (Gen AI) technology being less than a year old (ChatGPT launched last November 2022) a *TMForum survey of 104 Telecoms (Telco) of 104 C-Level Executives from 73 Telco Operators identified over 57% of Telcos are experimenting with GenAI/ Large Language Models (LLMs) expecting significant business impacts over the next 1-2 years!* (1) The initiatives are mainly being driven from the Board Level down to the CTOs to identify quick wins across their Business Support Systems (BSS) Customer Interaction Use Cases - Chatbots, Co-Pilot Service supporting Customer Agents and also Operation Support Systems (OSS) to assist Field Maintenance Engineers (FMEs) out in the field interacting with Network Operations Centre (NOC) using Co-Pilot to assist Operations and Maintenance (O&M) activities such as Root Cause Activities (RCA) to debug, fix, and get the system back online as quickly as possible. Already Telco Vendors are starting to use Classical AI/ML (Machine Learning) to automate their networks to improve performance and that activity will continue. But there is a role for Gen AI to compliment this activity, more about that later.



According to *Gartner 55% of organisations are now experimenting with or fully implementing GenAI technologies* (2). Organisations see GenAI reshaping their operations and how their employees will become more productive. *Anaconda report highlights that nearly 50% of data specialists are concerned by job security in face of these advancements* (3). But looking at the



'glass half full' point of view GenAI should enable employees to 'move up the value chain', such that GenAI should help to remove the mundane repetitive tasks and in this way the employee can consider GenAI as either a tool to support their activities, as such as clone or even a subordinate (as of today GenAI is such an immature technology that it is too risky to leave GenAI work directly with end customers, so the employee is considered the 'Pilot'(signs off on the GenAI's work) and the GenAI tool 'Co-Pilot' (does the mundane/repetitive tasks).

So, 2024 will be a defining moment for Organisations and individuals as they navigate this GenAI landscape. What trends and developments can we expect? How will the skills evolve? How will the data industry change? How will it be Governed/be compliant? Our white paper for Short, Medium and Long Term trends aims to explore these questions and more.

The Integration and utilisation of GenAI in Enterprises' internal operations and customer facing products have room for significant growth. *MIT technology review research reveals that only 9% of organisations have successfully productionized Gen AI . More importantly at least 50% of organisations have struggled to deploy generative AI due to high costs, lack of skills and more (4) . Regarding internal usage of AI tools, around 16% of professionals regularly use ChatGPT at work. Organisations are still defining their internal AI usage policies. According to Gartner , 45% of Executives Say ChatGPT has prompted an Increase in AI Investment, while 14% had issued a blanket ban on using tools like ChatGPT (5).*



Today, Co-Pilot assistant solutions are offered out-of-the-box from Hyperscalers(e.g. Microsoft) to Telcos who can fine tune with their own Telco data (e.g. Prompt engineering, more about that later).

GenAI has the capability to democratise AI(available for all) because co pilots give team members extra leverage to become more efficient and productive helping improving Operations Expenditure(OPEX) in Telco O&M departments. But democratising AI can also put the organisation at risk from issues such as hallucinations (incorrect responses to questions asked but stating them as fact).

GenAI Chatbot(Voice services) represent a new revenue opportunity for Telcos, particularly in the Enterprise area. Voice calls could be transcribed and translated(referred to as unstructured data) in real time and then stored and archived. Combined with Classical AI/ML for Predictive customer sentiment could open up new use cases to understand what the customer really wants. Therefore GenAI is ideally suited to working with unstructured data (Telcos have a large base of unstructured data that sits within, or adjacent to their organisations) and complements quite well with Classical AI/ML(which performs well on structured or labelled data).

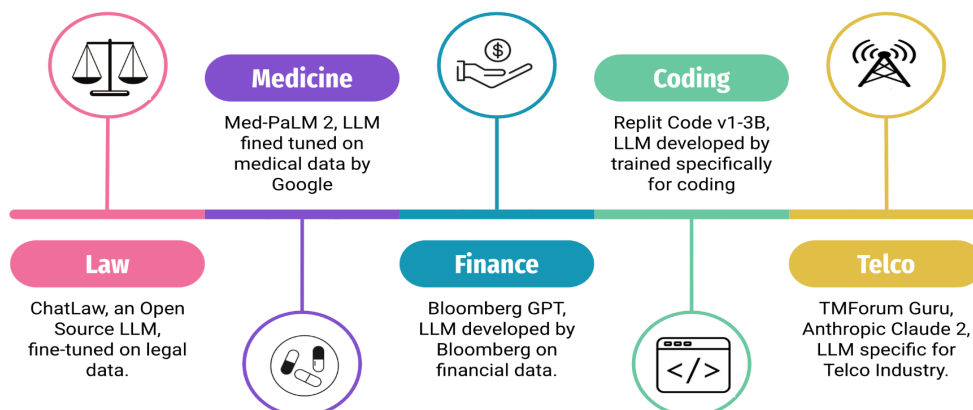
The Telco business model , like the early days with Cloud computing, allows an API based LLM pricing model to experiment and build Proof-of-Concepts(POC) without particular concerns



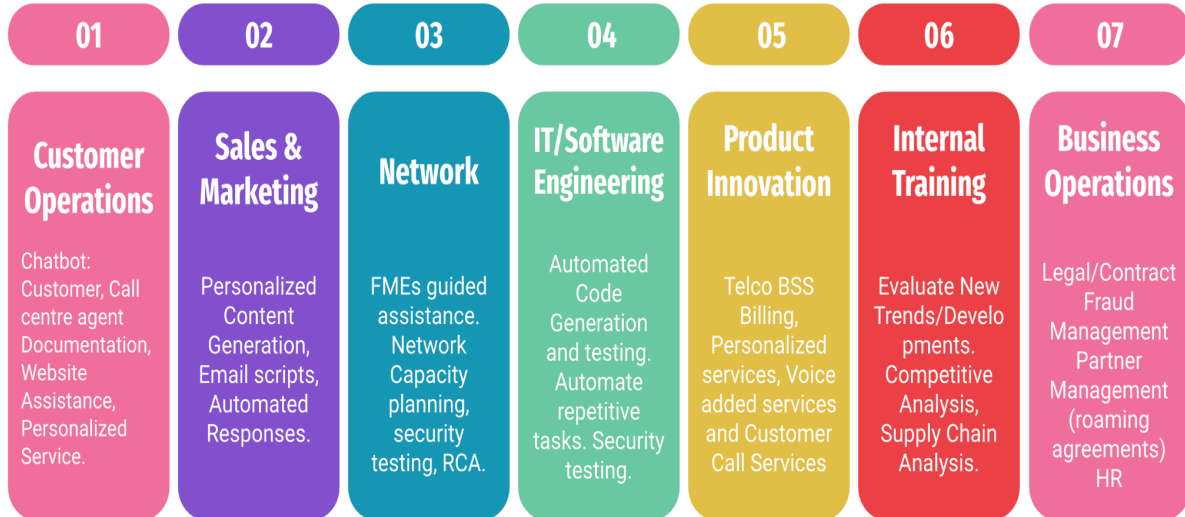
about the costs. But deploying GenAI at SCALE such pricing models may become less attractive. Telcos experience with the early deployments of Public Cloud services working with Hyperscalers(Azure, GCP, AWS) should allow the Telco to negotiate better terms and influence product development . For example, to incentivize the Hyperscaler(LLM owner) on a successful deployment, has an associated KPI(Key Performance Indicator) such as the number of customer queries successfully handled by the chatbot or a reduction in the time taken by the FMEs to address network outages.

Today, Telcos are experimenting with both Proprietary and Open-Source LLMs(more about that later). Trends in LLMs are for smaller more domain specific models in verticals:

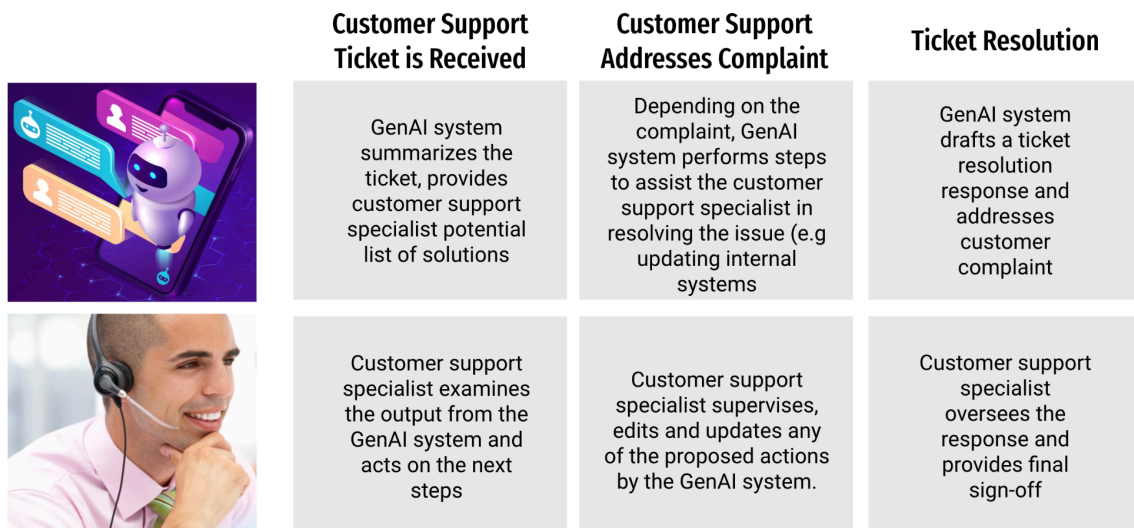
Trends in LLM for smaller more Domain Specific Models



7 Families of Use Cases - Example Telco Industry



Example GenAI Chatbot supporting Customer Agent



What are the Challenges?

In the Telco Industry, for example, there are *many issues to be addressed before deploying GenAI at SCALE* - siloed data, the lack of a common ontology and incomplete information architectures. Some Telcos are working on **Unified Data Models, Data Governance, the awareness of the data that is available across the organisation and the availability of data in real time to address this issue.**

Concerns about LLM accuracy, for example bias and hallucinations represent important challenges today. New techniques are emerging that will deliver significant improvements and as such, they may represent less of a long term concern for Telco businesses than **issues around privacy, security, IP and the cost of deploying GenAI at SCALE.**

The **ability to measure GenAI/LLM model performance** is a significant challenge In the world of predictive AI and machine learning, this was done through techniques such as precision and recall. However, approaches in GenAI are still being developed.

Telcos are desperate to find new sources of revenue and generating efficiencies BUT the AI landscape changes so quickly in terms of technology and players that Telcos need to be flexible and agile to change direction quickly and easily too(comes natural to Hyperscalers but not to the Telco industry).

Telcos and Companies in general will have to hire/retrain staff with NEW AI Skills to experiment with different models(to gauge accuracy, bias and the level of hallucinations), To do Trade-off Analysis Domain Specific Models versus Generic LLMs(which one yields better results?), Different Use Cases will require different levels of accuracy for the use of GenAI to be worthwhile, learn how LLMs work and Fine Tuning the models.

GenAI has a more transformative capability than other branches of AI and as such needs to be prioritised: Telco expenditure for example is forecasted in a *TMForum report to increase by as much as 6x over the next 2 years* and Telcos will be wary of committing to pricing models which SCALE with usage because of the risk of costs spiralling out of control.

What are the Immediate Recommendations?

Adopt aggressive GenAI experimentation near term off-the-shelf LLMs and Application built on top of them. In the medium to long term be cognizant of how to Deploy GenAI at SCALE, what tools are needed to ensure that costs are managed efficiently - ROI, the risk of vendor lock-in, and the integration of structured and unstructured data.

Find the right balance between 1) building families of use cases and applying GenAI to them and 2) democratising access to GenAI to allow different teams to discover their own use cases. GenAI should be seen as a tool for innovation within the wider Telco business.

Develop their own LLMs in the area of Telecoms BUT in more generic functions the speed of innovation in GenAI will likely outpace what CSPs can achieve with their own internal GenAI technologies and platforms. It is better to use off the shelf LLMs from partner Hyperscalers in these cases.

Utilise the lessons learned from Public Cloud Journey to inform their approach to GenAI - initially the business cases around public cloud migration was extremely attractive, BUT when different workloads and applications were migrated to the public cloud the efficiencies did not always emerge.

Digital Transformation Network & IT Systems using Cloud Native principles - adopt architecture concepts such as TMForum's Open Digital Architecture(ODA) , it is important that AI/GenAI become an integral part of the system design rather than a 'bolt on'.

New Data Architecture - specifies where the data is located, how it is accessed by AI and how frequently it is refreshed. TMForum is developing a new data architecture as part of the evolution of the ODA. This will build on SID, TMForum's existing Information Framework(asset common data model and a common data dictionary and offers a foundation for function, application, component and API development.

GenAI offers the prospect of building better, closer, more lucrative relationships - Telcos can fully exploit the potential of GenAI by using new customer chatbots, developing other messaging channels, and

leveraging voice conversations to understand, diagnose and address customer queries and to upsell and cross-sell.

Today, Who are the Main Stakeholders in GenAI/LLM?

In **2024**, the **GenAI/LLM Value Chain is dominated by the Hyperscalers/Cloud Platforms and Infrastructure/semiconductor companies(see graphic below)**. Hyperscaler service providers are extremely active in building LLMs. Stable Diffusion created the stable diffusion Text to Image Model. The trend is towards more specialist LLMs which are cheaper to create/train and geared towards specific market segments(verticals) and geographies. The immediate opportunity for the Hyperscalers is to monetize their investments in Gen AI by using Co-Pilots services to bolster sales of existing productivity tools(Cloud Computing)

Hence, why **Hyperscaler and Infrastructure vendors are very active in investing in startups** (Foundation models, specific LLM models for specific industry verticals as highlighted earlier for example in the **Telco sector, Anthropic Claude2 which is a rival to OpenAI ChatGPT received investment of \$100m from SKTelecom**).

Investment in GenAI/LLM startups globally reach a staggering \$17.8 billion in the first 9 months of 2023 versus \$3.9 Billion for all of 2022! Microsoft invested in **OpenAI(ChatGPT) \$10 Billion**. **Cohere's AI** platform which is enterprise focused on LLM Applications - Chatbots, Search Engines and Copywriting received \$270 million. **GenAI(A121 Labs) received £336 million from NVIDIA, Google, Comcast & Intel.**

Already in early 2024, AI scientists from Google's Deepmind are considering 2 spinoffs: 1. **Holistic AI in the UK** received €200m investment, with Karl Tuyls & Laurent Sifre(from Google Deepmind) joining this outfit which focuses on AI Governance(Responsible AI) for Enterprise , harnessing the power of AI at SCALE(Transparency, Accountability and Safety) and 2. Arthur Mensch from Google Deepmind founded **MistralAI in France** together with ex Meta Open Source LLaMa founders Timothee Lacroix and Guillaume Lample a nine month old company with a **\$2Billion valuation (possibly the first unicorn in GenAI/LLM?)** focusing on building a new GenAI/LLM model on the Azure platform. Notably Mistral AI received a recent investment of €105m from

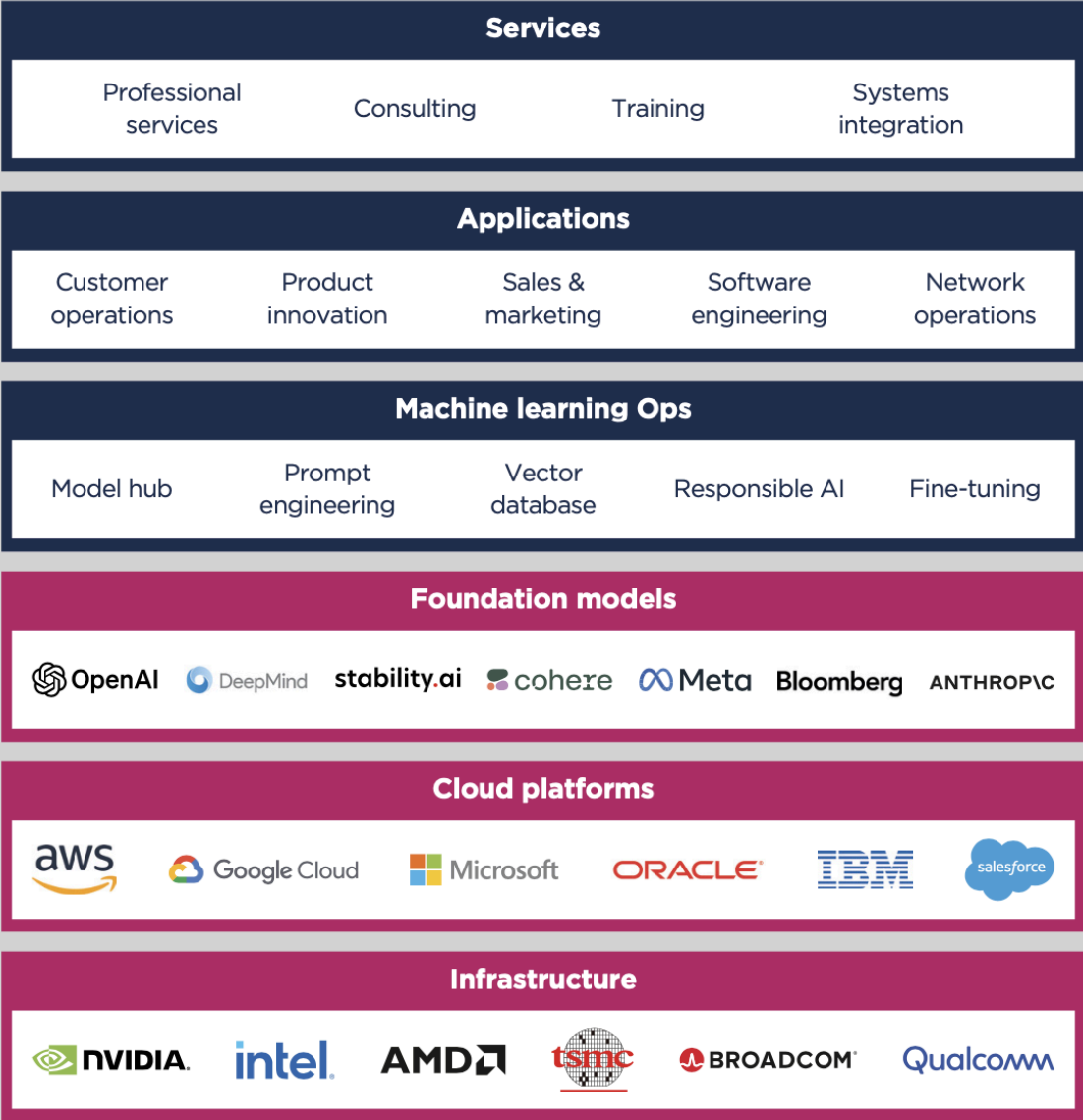
Illiad CEO Group owner Xavier Niel possibly to focus on a new LLM for the Telco Assurance market. Prior investments of €400m were received from a combination of NVIDIA, Andreessen Horowitz and General Catalyst .

In the **Telco space**, Telcos are partnering with Hyperscalers experimenting with off-the-shelf models, training and fine tuning them for specific use cases. BSS Vendors such as **Netcracker** are positioning themselves between the Telco and the LLM(using data adhering to the rule governing security and privacy). **Technotree(CognitiveScale)** BSS capabilities Chatbot to allow customers to ask questions about its products and solutions. **Aria Systems(Billing) & Salesforce(Einstein)** have a new optimised 'concept-to-care' monetization solution that allows Telcos to enhance products in their catalogues, integrate them into their OSS/BSS systems and offer automation and GenAI capabilities.

In a recent report by *McKinsey & Company* , they identified 63 GenAI Use Cases across 16 business functions, VALUE \$2.6 trillion to \$4.4 trillion in economic benefits annually when applied across industries (6).

In **Machine Learning Ops(MLOps)** this addresses the fact that GenAI models are complex and expensive to train & SCALE in production environments . The MLOps solution is to streamline the process to production , maintaining and monitoring the ML model. With regards to SCALE, automating the training and deployment process, the range of tools to curate, host, fine-tune and manage foundation models.

The early-stage GenAI value chain



This graphic is representative of the early-stage GenAI market and is not intended as an exhaustive list of the companies operating in these areas

TM Forum, 2023

Medium Term Considerations: Build, Train, How to Deploy?

In a recent white paper by TMForum they asked the Telcos how they are approaching the GenAI opportunity and why GenAI?

How are you approaching the GenAI opportunity (choose all that apply)?



TM Forum, 2023

For **GenAI the tools can be deployed quickly**, highlighting near term benefits - **Customer Experience Use Cases are quick wins, 'low hanging fruit', 59% of Telcos are addressing these use cases.** GenAI has transformational capabilities than other branches of AI that it needs to be prioritised! **53% of Telcos are deploying their own AI Centres of Excellence** building capabilities to assess the key risk factors - Privacy, Security, Intellectual Property(IP) and to inspire employees to take risks with GenAI agents - what works and what does not work - 'fail quickly'. **53% of Telcos are experimenting with GenAI across different departments and partnering with Hyperscalers (45%) , setting up workstreams, 5 P's : Platform, Partners, People, Process, and Policy and Guidelines.**

Of course **Telcos have challenges** with GenAI and AI in general, and have spent many years trying to improve the Availability, Reliability and Usability of generated data (from Operations and Customers). Commercial viability and availability of AI has made the DATA readiness a bigger priority for CSPs . The **solution being considered by Telcos is a Unified Data Model(Data Readiness)** - data governance, the awareness of the data that is available across the organisation and the availability of the data in real time . Without GenAI Telcos would need to label the unstructured data - process of adding meaningful labels to provide context so that machine learning (ML) models can learn from it. With GenAI/LLM the Telco can feed this unstructured data into the model and immediately start generating insights which over time can be fine

tuned.

Proprietary versus Open Source?



So firstly , the Telco and Enterprises in general have to consider whether to go the **Proprietary GenAI/LLM Route** - such as what Bloomberg have done in their field of Finance with BloombergGPT which was extensive to develop, ChatGPT4 for

example was in the range of \$50-\$100m to develop, Bloomberg GPT(Finance) has 50 Billion



Parameters! Gopher Google Deepmind (Science and Humanities) has 280 Billion Parameters . Or **Open Source** which has the advantage to access, use and tweak the model based on user requirements. Open Source models such as LLaMa 2 by Meta &

Microsoft Chatbot or GPT-Neo/GPT-J EleutherAI(academia open source AI toolkit) ? But with Open Source LLMs this might not be a viable option for Telcos(or other Enterprises) who do not have the skills. Having the LLM model on premise might be attractive for dealing with sensitive data but initially for experimentation Public Cloud may be a better option to SCALE compute resources up and down quickly.

Short to Medium term Public Cloud LLM usage by Telcos

more than *65% of them anticipate training off-the-shelf models(AWS/Altman Solom survey) 15% said they would build foundation models in-house. (7)* Currently there are no clear trends using context learning / prompt engineering(i.e. inference) or fine tuning(i.e. training) to train models. Vodafone working with Microsoft is an interesting case to highlight with regards to 'prompt engineering'. Speak more about that later , for now , in the early phase , Telcos are leaning towards context learning which is less resource intensive and provides 'good enough' performance for many use cases.

In any case **Telcos(Enterprises in General) to experiment with different models will have to hire/retrain staff** with new AI/GenAI/LLM skills to gauge the accuracy , the bias and the level of the hallucinations(inaccurate / inappropriate responses). Also Tradeoff Analysis Domain specific models versus Generic LLM which ones yield

better results will have to be considered. Different use cases will require different levels of accuracy for the use of GenAI to be worthwhile. To learn how LLMs work will involve different teams and departments. Fine tuning models process involves understanding how the model works and then framing questions in such a way as to generate the most accurate results.

In general the more parameters the exponential time it takes to train the model, thus OPEX costly.

Trend towards Smaller Specialised LLM?

Hence the **trend to more specific domain LLM with less parameters and the huge investment in startups(highlighted earlier)**. These domain-specific models come with a lot of advantages - firstly they provide higher accuracy, for example PaLM 2 is the first LLM to perform an expert test taker level performance on medical benchmarks. Secondly, it is less costly to train as dataset sizes are much smaller than generic LLMs. These models offer organisations more control over their usage, allowing for more custom applications of AI technology.

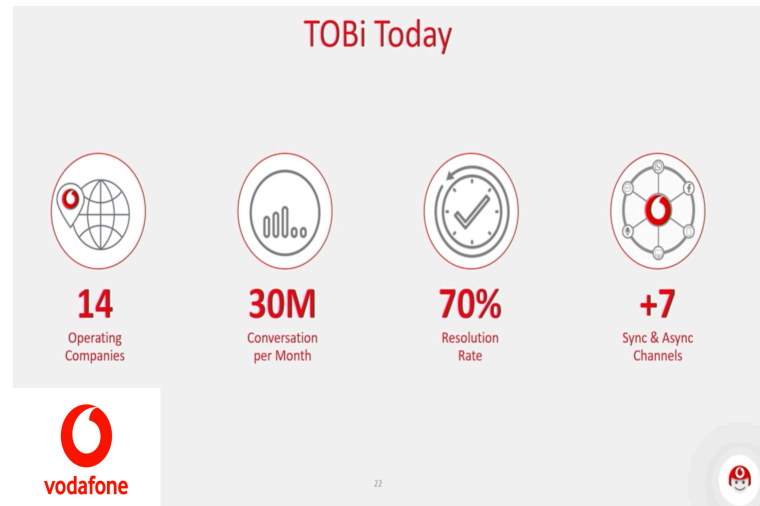
"I think in the next 12 months, there will probably be millions of customised, fine tuned models. Just think about the legal profession. There are differences between US, UK , and European Law. So they will be different models for different countries, regions, domains and more" - Thomas Davenport, Distinguished professor at Babson College.

Arguments For & Against Telcos having their own GenAI/LLM model

Pros and Cons			
	IP Ownership License the model. AI Capabilities. NEW Products & Services	01	What Model? Operations different functions across org. (Finance/Legal etc.)
	Domain Specific More Accurate for Network Operations	02	Fine tuning existing models? Telcos are slow moving org vs. Hyperscalers. Maybe Partner better?
	Control Future AI Direction Not reliant on 3rd Parties. Customize AI model	03	Costs No proper visibility costs could easily spiral out of control
	Pricing Models at SCALE unclear. Better visibility by developing own model	04	Skills Lacks in-house skills to build and train telco specific LLM models.

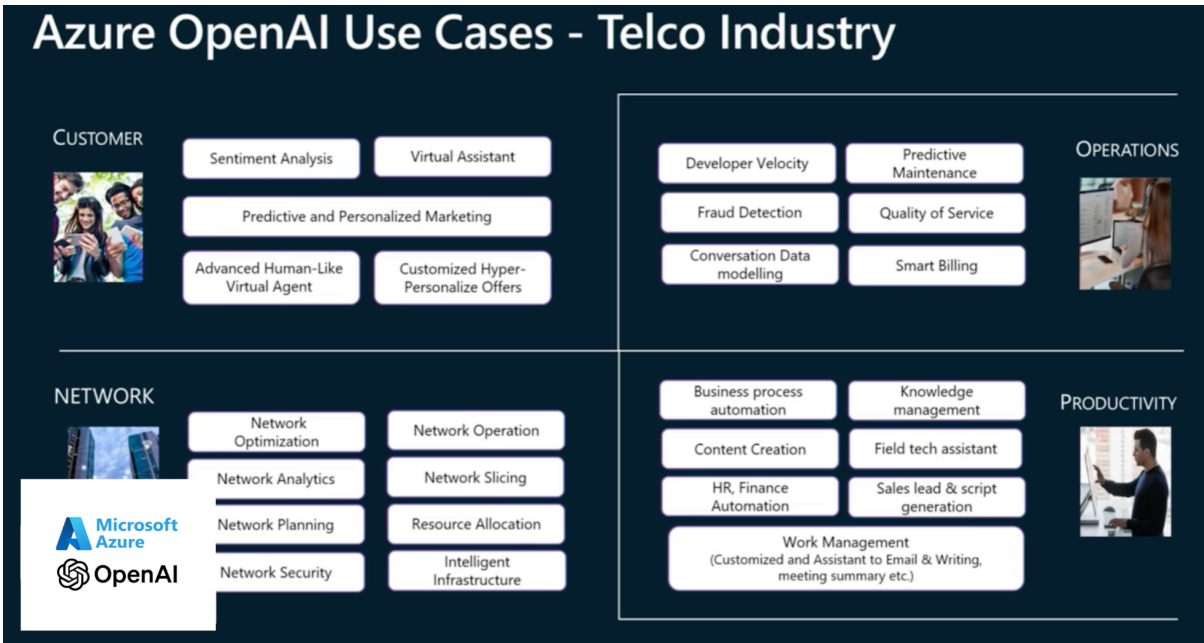
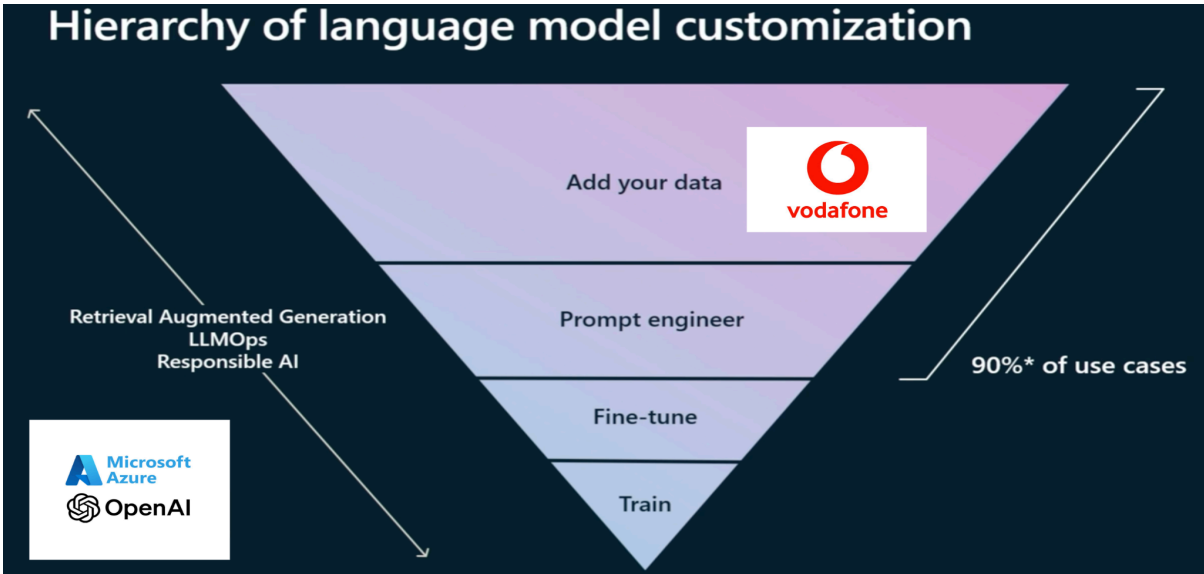
Vodafone & Microsoft Use Cases

Telco Vodafone's chatbot ToBi which is operating in 14 countries, has 30 million conversations per month and has a 70% resolution rate has been working with Microsoft OpenAI Gen AI for some time on experimental use cases which are now in production in "virtually every part of Vodafone's business", where any form of written or spoken human language is used, including supply chain, legal, customer operations and marketing. All of its AI is running in Public Cloud enabling Vodafone to focus on use cases and investment in skills.



"Our Goal is to enable every single employee across Vodafone, in every single function, to derive value, make faster decisions, by using insights and by using AI. Today, Agent Co-Pilot (for minimum customer impact) where GenAI summarises email interactions with customers are sent to the customer agent who deals directly with the customers. In Future GenAI there is a balance between benefit/cost/risk (too risky at the moment to have GenAI deal directly with end customers. But Vodafone is working on a 3rd Generation of ToBi (using prompt engineering) to minimise the risk in this regard (see graphic below)", says Dr. Lester Thomas, Head of New Technologies and Innovation Vodafone.

With prompt engineering enables 1) over 90% of Vodafone Use Cases, 2) Vodafone to input their data without necessarily having to train the underlying LLM model supplied by Azure OpenAI, leaving that to Azure and 3) 'Learn by doing' benefit less cost and risk to address the 'low hanging fruit' use cases. This also allows Vodafone to keep control of the security of its data, enabling privacy and security of its customers. A Win/Win for both the Enterprise and the Hyperscaler.



Use Cases - Customer and Network Operations & Productivity where GenAI/LLM compliment classic AI/ML(that is being used for autonomous networks). GenAI is a great tool for 'generating synthetic data' for testing especially for Security Use Cases. Azure Co-Pilot Tool empowers the Customer Agent and Network Engineer to achieve more.

Global Telco Alliance between SKT, Deutsche Telekom, e&e, SingTel

Journey to build a LLM for Telco Industry, launches in early 2024.

Here is the Journey thus far:



2021-2023:
Developing AI Supercomputer



July 2023: Global Telco Alliance to co develop a platform for operator specific AI services.
DT & SKT : LLM for Digital Assistants in Customer Service



August 2023: invested **\$100m** in GenAI startup Anthropic(also backed by Google)
Companies Agree : jointly develop **LLM that is optimised for Telco Operators**



Launch in early 2024: Variant of Anthropic's generic model called Claude Support Korean, English, German, Japanese, Arabic and Spanish Languages.



Foundation for New Services:
Chatbots, Other Apps

Global Telco AI Alliance Plans:
Add new Operator members
Drive SCALE
Increase richness of the model
Increase Applications built on top

"What became very clear for us is to get really good with this, and use it at massive scale we need to tune the model to make it work for our use cases", says Johnathan Abrahamson, Chief Product & Digital Office Deutsche Telekom. " Focus to deliver a stronger Customer Experience using a chatbot. An LLM's reasoning capabilities make it good at analysing customer intent and searching a telco's back-end system for the answer. We really like the idea of Deutsche Telekom to develop its own IP."

Chung Suk-Geun, SKT Chief AI Global Officer, " GenAI/LLM has the potential to redefine the operator relationship with their customer, to sell new Products and Services , improve the customer experience and to help other enterprise customers outside the Telco Industry. Develop Operator Billing, the sale of third party services, use sms to communicate directly with customers and to initiate transactions, transform internal core processes, marketing, sales and customer service operations. Some examples of marketing cost reductions(automated customer service), networks(automated infrastructure monitoring), and service quality improvements (personalised, targeted offerings)"

What are the Skills Required?

GenAI is proving to be exceptionally adept at data-related applications, especially in the realms of exploratory data analysis and coding tasks. Tools like ChatGPT for Data Analysis and GitHub Co-Pilot showcase GenAI's ability to reduce complex data queries and coding tasks into natural language queries. These advancements are not just about automating tasks; they radically **reduce the barrier to a Data Culture** - to interacting with data. GenAI's ability to understand and execute natural language queries have made data analysis more intuitive and accessible, even for those without extensive technical expertise.

This evolution could be foundational in achieving the 'Holy Grail' of data transformation: building a data culture, integrating GenAI interfaces into coding based and low-code platforms - Excel, Power BI, Tableau and more makes these tools more accessible to wider audiences within any organisation. This is the essence of what Valerie Logan, CEO of the Data Lodge refers to as '**Data Literacy by Design**' where data tools are designed with enabling data literacy in mind.

"More and more data tools are allowing for these augmented capabilities, things like not allowing time series data into pie charts or providing tool tips and hovers that explain data insights. This is what our community defines as data literacy by design", Valeri Logan CEO of Data Lodge.

As we progress into 2024, the widespread adoption of these AI enabled tools is expected to significantly lower the barriers to engaging with data, enabling more people within organisations to leverage data for insights and decision making. However, with great power comes great responsibility. The responsibility of critiquing outputs, methodologies and potential hallucinations. This is why data literacy and conceptual understanding of data science remain crucial, ensuring that data insights are not unquestioningly believed but intentionally curated and understood.



Despite the increased focus on data and AI literacy, there is still a significant gap in education and training. *According to Randstad, while 50% of employees believe AI will fundamentally change their workflows, only 13% have received AI training (8) .*

Moreover, according to a *New Vantage Partners survey, only 1.6% of CDOs cited data literacy as their top investment priority (9).*

In 2024, building robust data and AI literacy capabilities will be essential for organisations. It will serve as an insurance policy against the risks and inaccuracies associated with GenAI and foster better and broader adoption of these technologies across various sectors.

"AI Literacy is understanding what the training data is, and then how that is being used to generate the output of an AI generated insight. So even if you think we know generated AI has a problem with hallucinations, understanding what a hallucination is and why that may happen is foundational for AI literacy" , says Cindi Howson, Chief Data Strategy Officer at Thoughtspot

Throughout 2023, there was a considerable debate about the future of prompt engineering(as discussed earlier with the Vodafone and Microsoft Use Cases) as a distinct career path. While it's still uncertain it will emerge as a standalone profession, it's becoming increasingly clear that prompt engineering skills are in high demand across various roles. *According to Forrester, an estimated 60% of employees are expected to undergo training in prompt engineering (10) ,* indicating its rising importance in the workforce.

As we move into 2024, prompt engineering is set to become a highly sought-after skill set. However, it will likely be sought after within existing job roles rather than alone as a separate role. For instance, engineers working with LLMs must be proficient in prompt engineering to deploy and manage these AI systems effectively. Similarly , marketers will find prompt engineering invaluable for creating more efficient and impactful marketing workflows. The skillset will be essential for various professionals interacting with AI daily, making it a versatile and crucial component of various job functions.

ChatGPT Prompt engineering

- ▶ ● 🔑 Keywords - with examples
- ▶ ● 📄 Basic prompt structure
- ▼ ● 🛠️ Prompt engineering techniques

- ▼ ● Role prompting

In this technique you provide the model with context about how it's supposed to answer. This attempts to constrain the model into replying given the expected knowledge domain of the role.

Example

Act as if you are a kindergarten teacher working with six year-olds.

How does electricity work?

- ▼ ● Zero-shot prompting

This is where you provide the model with a single set of instructions, a statement, or a question.

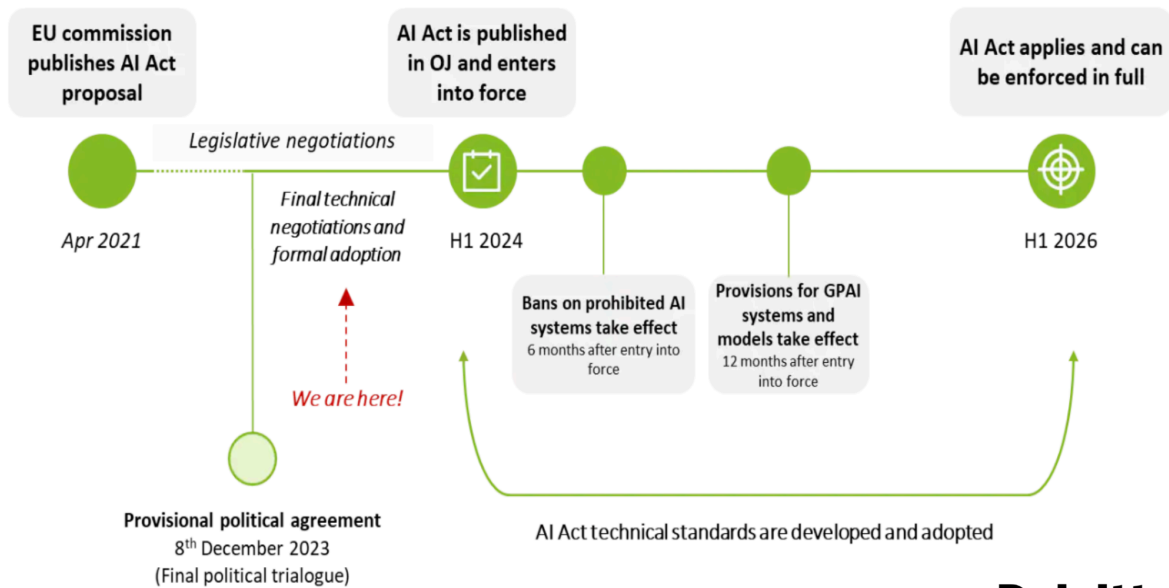
Long Term Trend as LLMs become more mature ?

As GenAI/LLM models become more mature, risks have to be dealt with through **Governance/ Regulation/Compliance**.

EU AI ACT :



The EU reached a provisional political agreement on the **AI Act** on December 8th, 2023. This Act is the first comprehensive and legally binding cross-sector framework for a major global economy. Following the finalisation of further detail, the AI Act is expected to become law in the 1st half of 2024. A 2 year phase implementation period for organisations will then commence.



Deloitte.

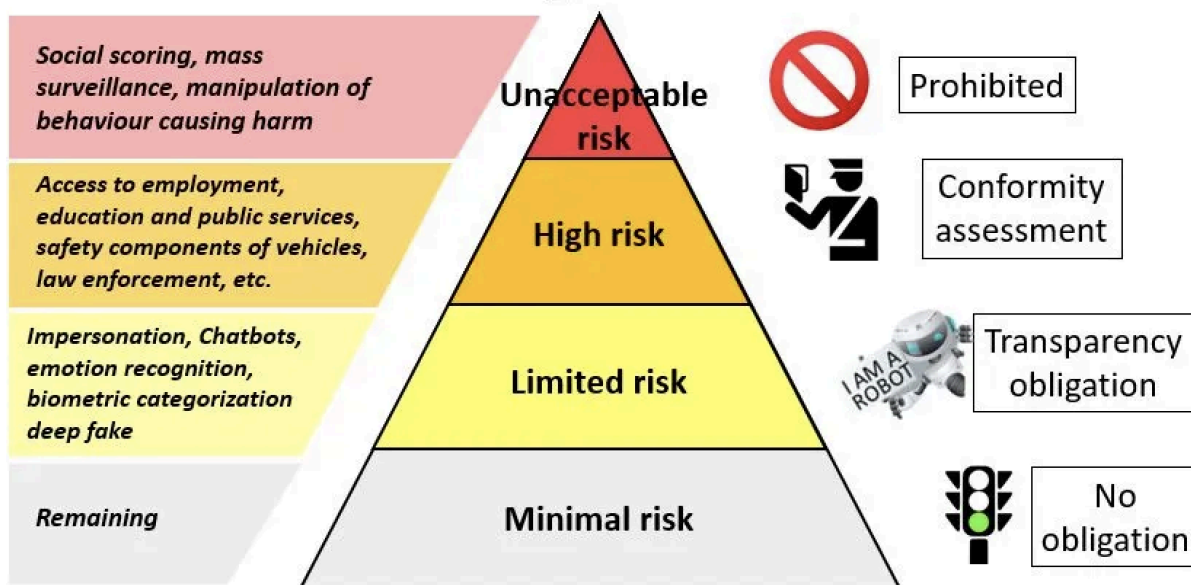
(11) - source Deloitte "The EU AI Act: the finish line is in sight"

Four New Main Elements of " **Provisional Political Agreement**":

- 1) Rules of future high impact General Purpose AI(GPAI) models and high risk 'AI Systems'
- 2) Revised system of governance with some enforcement powers at EU Level.
- 3) Extension of prohibitions list but possibly to use remote biometric identification by law enforcement authorities in public spaces, subject to safeguards(AI Act does not apply)
- 4) Better protection of rights, obligation for deployers of high risk AI systems to conduct a fundamental rights impact assessment prior to putting an AI system into use.

The AI Act proposes a prescriptive but risk based approach to regulate **Single Purpose AI systems and General Purpose AI (GPAI) systems** and models. Single Purpose AI systems will be permitted but their operator will face strict requirements before placement in the market or use. All GPAI models and systems will be subject to transparency requirements to ensure fair allocation of responsibilities along the AI value chain.

EU Artificial Intelligence Act: Risk levels



High impact GPAI models posing systemic risk will face additional stricter obligations. The AI Act will have significant extraterritorial implications, as it will apply to organisations in other jurisdictions if they market or deploy their AI systems in the EU. This raises the questions for multinational firms, who must choose between adopting AI Act standards globally or scaling back AI in the EU? Firms should assess which of their current and planned AI systems and models fall in the scope of the AI Act and conduct a gap analysis against key requirements. This will provide insight into the scale and challenge of compliance efforts and help identify the impact of the AI Act on strategic choices and product governance.

While the agreement provides clarity on many key elements of the AI Act, the final legal text will **not be available until the 1st half of 2024**, meaning organisations will need to wait to fully assess its operational and strategic impact. However, recently the *Dutch Government in January 2024 announce a vision for Generative AI (GenAI) action plan to support the EU AI Act and to pledge €200m to foster local investment, public-private partnerships (12)* ensuring the use of responsible GenAI application for specific government services and establishing a dedicated team to assess current usage.

" We need to start dealing with what we know instead of waiting for final answers, adding that businesses would benefit from regulatory clarity", says Thomas van Oortmerssen, ministry of the Interior & Kingdom Relations of the Netherlands.

"It is essential that the Netherlands does not remain on the sidelines when it comes to AI, noting that Asia and the US have already taken the lead and Europe will have to catch up!" (13)

Corporate Governance

At the corporate level for example **Microsoft Azure AI Studio (14)** tries to address Governance/Rules , Checks & Balances to compliance.



Acknowledges that there is a big responsibility to offer these services responsibly , reliably and safely to customers. Similarly, **Vodafone AI Framework studio** requires that Use Cases are pitched to the Governance board level and complies with the EU ethical AI Act for lawful and robust AI developments. These initiatives are a good start to ensure a proper governance over responsible AI developments but many challenges still remain to be addressed...

For example, cross sector regulations like **GDPR** and sector specific regulations e.g. governing financial services or digital markets, will be applicable depending on the specific AI use case. Interaction between regulations and AI Act raise some questions:

- 1) Interactions between **FRIA(Fundamental Rights Impact Assessment)** and **GDPR** are a potential challenge.

- 2) Responsibilities of different **actors in the AI value chain** may not always align with those of the organisation that is primarily responsible for protecting personal data under GDPR i.e. the data controller.
- 3) AI Act compliance and **Digital Services Act(DSA)** compliance of **very large online platforms(VLOPS)** together with overall compliance risk management will need to be confirmed in the final text.
- 4) These interactions raise important questions around cooperation and alignment between NCAs responsible for the AI Act and those responsible for other horizontal and sector regulations.

How to address Data Privacy Concerns in AI Prompt Engineering?

(15) Data privacy is a growing concern in AI, especially in the context of prompt engineering(GenAI) , where users' prompts and the AI's outputs can raise questions about data collection, storage, and usage. Here are some ways to address the concerns:

Data Minimization and Anonymization the data before it is inputted so reduce risk of Personal Privacy, such as **1) Pre-defined prompts** that don't require user input, minimising the need for personal data **2) Anonymize the input data and stripped of personally identifiable information** to protect individual privacy **3) Aggregated or de-identified information**. OpenAI may aggregate or de-identify Personal Data so that it is no longer used to identify users and **4) Using a technique called Differential Privacy:** Adding controlled noise to outputs helps ensure statistical accuracy while reducing the risk of re-identification from individual outputs.

Security by Design of OpenAI generator: employs several measures to ensure data privacy and security. These include **encryption of data in transit and at rest, Secure Data Storage, Access Control and Authentication, Regular Security Audits, Compliance with Regulations (EU GDPR, California Consumer Privacy Act (CCPA)), Data Minimization, User Consent and Data Breach Detection and Response.**

Transparency and Control:

- **Prompt Attribution:** Identifying the source of prompts (user or pre-defined) helps trace outputs back to their origin and promotes user awareness.
- **Output Analysis and Filtering:** Providing tools for users to analyse and filter generated outputs based on potential privacy concerns empowers them to control the use of their data.
- **Explainable AI (XAI):** Making the AI's reasoning behind its outputs transparent to users builds trust and allows them to assess potential privacy risks.

Education and Awareness:

- **Clear Data Practices:** Communicating data collection, storage, and usage practices in user agreements and privacy policies builds trust and transparency.
- **Educational Resources:** Providing educational resources on prompt engineering and data privacy empowers users to make informed decisions about their data.
- **Responsible AI Development:** Promoting ethical considerations and responsible AI development principles within companies ensures data privacy remains a priority.

Biography



Over 25 years experience across different industries : Telco, Semiconductor & FinTech. Currently working with C-Levels in Telco strategizing how AI(Artificial Intelligence) & ML(Machine Learning) can be used towards Automating their Networks. How Generative AI(GenAI) can help improve their Customer Experience, and how Gen AI can compliment their Automated Networks towards helping Network Engineers to become more productive and improving Telcos Operations.

References

- (1) "TM_Forum_Generative_AI_Report.pdf"
- (2) <https://www.gartner.com/en/newsroom/press-releases/2023-10-03-gartner-poll-finds-55-percent-of-organizations-are-in-piloting-or-production-mode-with-generative-ai>
- (3) <https://www.anaconda.com/press/anaconda-state-of-data-science-2023>
- (4) <https://www.technologyreview.com/2023/10/10/1081117/generative-ai-deployment-strategies-for-smooth-scaling/>
- (5) <https://www.gartner.com/en/newsroom/press-releases/2023-05-03-gartner-poll-finds-45-percent-of-executives-say-chatgpt-has-prompted-an-increase-in-ai-investment>
- (6) <https://www.mckinsey.com/industries/financial-services/our-insights/capturing-the-full-value-of-generative-ai-in-banking>
- (7) https://pages.awscloud.com/rs/112-TZM-766/images/Altman%20Solon_AWS_Telecoms%20Generative%20AI%20Study.pdf
- (8) <https://www.randstad.com/press/2023/over-50-believe-ai-will-future-proof-their-careers-only-13-have-been-offered-ai-training/>
- (9) <https://www.businesswire.com/news/home/20220103005036/en/NewVantage-Partners-Releases-2022-Data-And-AI-Executive-Survey>
- (10) <https://www.forrester.com/blogs/predictions-2024-data-and-analytics/>
- (11) <https://www2.deloitte.com/uk/en/blog/emea-centre-for-regulatory-strategy/2023/the-eu-ai-act-the-finish-line-is-in-sight.html>
- (12) <https://www.government.nl/latest/news/2024/01/18/dutch-government-dutch-government-presents-vision-on-generative-ai>
- (13) <https://www.mobileworldlive.com/ai-cloud/netherlands-backs-eu-ai-act-in-e200m-action-plan/#:~:text=Ministry%20of%20the%20Interior%20and,would%20benefit%20from%20regulatory%20clarity.>
- (14) <https://azure.microsoft.com/en-us/products/ai-studio>

(15) <https://www.quora.com/How-do-you-address-data-privacy-concerns-in-AI-prompt-engineering>