Generative AI and ChatGPT 101

#PositiveImpact

May 2023

What is generative AI?  |  How does it work?  |  How can you use chatbots?  |  What does generative AI mean for the future?
Six months of hype, chat and fears

Next week will mark six months since the launch of ChatGPT, the chatbot based on OpenAI’s GPT-3.5 large language model.

There has been an extraordinary flurry of launches, hype and activity since then. Even in the past week, OpenAI, Google and Meta announced fresh products, OpenAI’s CEO testified in Congress, and the G7 called for “guardrails” on AI, among many other developments.

Whether you agree with the researchers who say the underlying technology has “sparks of Artificial General Intelligence” or you think it is just a “stochastic parrot”, it is vital to understand how it works and what it could mean.

This is the latest in our 101 series for generalists, aiming to give a clear overview of generative AI, addressing what it is, how it works, how to use it, and what it means for jobs, the economy and society. It is also part of an ongoing series starting this week looking more deeply into the topic.

In the meantime, one thing seems certain. Amara’s law, that “we tend to overestimate the effect of a technology in the short run and underestimate the effect in the long run,” could have been written for (or indeed by) ChatGPT.

1950: Alan Turing proposes a learning machine

1966 First chatbot ELIZA developed

1980s Early development of neural networks

2011 Apple releases first iPhone with AI assistant Siri

2012 AlexNet takes leap forward in image recognition

2017 “Attention is all you need” paper announces transformers

2018 Google introduces BERT large language model

2019 OpenAI creates GPT-2 LLM with 1.2bn parameters

Nov 30, 2022: OpenAI launches ChatGPT chatbot
What is special about generative AI?
Almost anyone with internet access can right now operate a chatbot based on a large language model (LLM) at little cost and with little or no training.

1. General (multi-purpose) tech
- Coding
- Search
- Translation
- Sentiment analysis
- Summary
- Human-like Q&A
- Image generation
- Entity recognition

2. Ease of use

Tell me how to access and use ChatGPT in a six-line rhyming poem:

To access ChatGPT, heed this rhyme,
Visit websites where it's prime.
In the chatbox, type your quest,
Hit send, put it to the test.
Read responses as they show,
Keep chatting, watch the knowledge flow.

3. Rapid adoption

<table>
<thead>
<tr>
<th>Key technologies (adoption)</th>
<th>Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Telephone</td>
<td>Two months</td>
</tr>
<tr>
<td>TV</td>
<td></td>
</tr>
<tr>
<td>WWW</td>
<td></td>
</tr>
<tr>
<td>Smartphone</td>
<td></td>
</tr>
<tr>
<td>Uber</td>
<td></td>
</tr>
<tr>
<td>Spotify</td>
<td></td>
</tr>
<tr>
<td>Pinterest</td>
<td></td>
</tr>
<tr>
<td>Twitter</td>
<td></td>
</tr>
<tr>
<td>Facebook</td>
<td></td>
</tr>
<tr>
<td>Instagram</td>
<td></td>
</tr>
<tr>
<td>ChatGPT</td>
<td></td>
</tr>
</tbody>
</table>

Source: Deutsche Bank, company reports, media headlines; ChatGPT content written with ChatGPT.
What is a large language model (LLM) and where does it fit into AI?

GPT-4 is the underlying technology for ChatGPT and is currently the most high-profile LLM, an AI model trained on a vast amount of data, capable of understanding and generating human-like text.

- **Artificial intelligence**: Machines/software capable of intelligent behaviour: three levels
  1. Narrow AI: performs limited tasks
  2. Artificial General Intelligence (AGI)
  3. Superintelligence

- **Machine learning**: Learning and acting on data without explicit instructions

- **Neural networks**: Models inspired by the human brain

- **Deep learning**: Uses multiple hidden layers of artificial neural networks to enhance performance

- **Foundation model**: Understands and generates text and/or images, code etc; need a lot of data for training

- **LLM**: Focused on language tasks

Source: Deutsche Bank.
What is the underlying deep learning technology for LLMs?

An image classification example

<table>
<thead>
<tr>
<th>Input</th>
<th>Feature extraction</th>
<th>Classification</th>
<th>Output</th>
</tr>
</thead>
</table>
| Traditional machine learning: Experts input structured data | ![Feature extraction diagram](image1.png) | ![Classification diagram](image2.png) | CAT
NOT CAT |
| ![Image of cat](image3.png) | | | |

Deep learning with neural networks: Can learn far more complex structures and relationships (e.g., pixels) without needing human help to identify relevant features

<table>
<thead>
<tr>
<th>Input</th>
<th>Feature extraction + Classification</th>
<th>Output</th>
</tr>
</thead>
</table>
| ![Image of cat](image4.png) | ![Hidden layers diagram](image5.png) | CAT
NOT CAT |

"Neuron/node" → "Weight"
What are the key neural networks and what are they made of?

Neural networks have become increasingly sophisticated over the past decade, particularly with the breakthrough of Convolutional Neural Networks in 2012 and Transformers in 2017.

Key neural network models:
(Approximately ranked from simple to complex)

1. **Feedforward**: one-directional, no looping/feedback

2. **Recurrent (includes Long Short-Term Memory, LSTMs)**: good for sequential data like text; have a feedback connection, enabling them to “remember” previous inputs

3. **Convolutional**: layers “convolve” the input, making them particularly good at handling images

4. **Generative Adversarial Networks**: consist of two networks that compete and can generate new data

5. **Autoencoders**: unsupervised, used for data compression

6. **Transformers**: best for natural language processing (NLP):
   - Process all pair-relationships between input tokens
   - Use “self attention” to weigh the relevance of, for example, different words in a sentence and capture long-range dependencies in data
   - Form basis for models like OpenAI’s GPT, Google’s BERT and PaLM, Meta’s LLaMA and DeepMind’s Chinchilla

Common components:

- **Input layer** – takes in the data, e.g. text, and feeds it into the next layer
- **Hidden layers** – in between input and output layers, responsible for data computation and transformation
- **Output layer** – the final result, the form depends on the initial task – e.g. classification, next word prediction etc
- **Neurons** – the nodes within layers where weights, activation functions and biases are applied to inputs
  - **Parameters (weights and biases)** are learned during training
  - **Activation functions** decide whether a neuron should be activated or not, adding non-linearity
- **Tokens** – unit of text, such as a word or subword, that is an input to a language model

Source: Deutsche Bank.
What was the tipping point for generative AI?
Massive computing power, data and the invention of the transformer model are the oxygen, fuel and heat that have enabled generative AI to catch fire.

Transformer architecture

Transformer model: the Rosetta Stone of SOTA (State of the Art) large language models

Source: Deutsche Bank.

From “Attention Is All You Need”, Vaswani et al, 2017
How large was the advance in computing power to get us here?

Recent models use several orders of magnitude more processing power, or “compute”, than past models, with more parameters and more complex model architectures.

Source: Sevilla et al. (2023), Our World in Data, Deutsche Bank. Note: Computation is estimated based on published results in the AI literature and comes with some uncertainty. The Authors of the study expect the estimates to be correct within a factor of 2.
How many parameters do LLMs have now?

The number of parameters in LLMs, a key measure of size, has surged in the past five years. Even if estimates are true, GPT-4 will not be the first 1tn parameter model. China’s multimodal Wu Dao 2.0 reportedly has 1.75tn parameters.

Selected multimodal and language AI systems with more than 50B parameters

- Wu Dao 2.0
- M6-T
- Megatron-Turing
- PALM
- Minerva
- [GPT-4?] *
- LLaMA
- NLLB
- Flamingo
- LaMDA
- FLAN
- Chinchilla
- GPT-3 (davinci)

Source: Sevilla et al. (2023), Our World in Data, Deutsche Bank. *Note GPT-4 parameter count has not been announced; industry estimates that it is around 1 trillion parameters. Parameters are estimated based on published results in the AI literature and come with some uncertainty. The authors of the study expect the estimates to be correct within a factor of 10.
Why do more parameters matter?
LLMs show emergent abilities, where capabilities suddenly (and often unexpectedly) appear when the model reaches a certain size.

Source: Deutsche Bank, “Emergent Abilities of Large Language Models”, Wei et al. (2022); IPA is International Phonetic Alphabet, NLU is Natural Language Understanding
What are the steps in preparing and using an LLM?

LLMs are pretrained then often fine-tuned on data to acquire a skillset before the user prompts them to perform a task.

1) Pretraining (often self-supervised)

2) (Typically) Fine-tuning with specific labelled task-focused data (eg medical, legal)

3) Chatbot overlay sometimes further trained with human feedback to improve how it behaves

4) Prompt (Context input)

5) Inference

6) Output taken into account as context for subsequent prompts in chat mode

Source: Deutsche Bank.
Where do LLMs fit in with other technology?

The “tech stack” is built on hardware and cloud computing. LLMs can often support multiple applications which interact with them via APIs (Application Programming Interfaces). Apps can also be given extra skills with “plug-ins”.

Source: Deutsche Bank.
How do you use a chatbot?
ChatGPT and its competitors look and feel like search engines but it is more effective to interact with them like speaking with a human, giving lots of context, and expecting that answers may be more creative than factual.

Basic approach

1. Ask like you’re having a conversation with a person
2. Give context: more is better, inc goal, tone of voice etc; you can paste in relevant text and data (the amount it can handle depends on the model)
3. Rephrase, correct and clarify if you don’t get the answer you want – you don’t need to restate the whole context

Useful points to include

“Act as a [ROLE]”: doctor, teacher, legal expert, sales manager…
“Addressing an [AUDIENCE]”: university student, 7-year-old…
“In a [FORMAT]”: 500-word essay, sales pitch, analogy, rhyming poem, social media post, table, code, in the style of Mary Shelley…

Digging deeper

1. Specify steps: use chained prompting for complex tasks: ie write an article starting with the outline… [Output]…
   Now write three key messages for each subheading… etc
2. Give sample answers in your question (ie “one-/few-shot prompt”) instead of an open-ended “zero-shot prompt”
3. Ask the chatbot to suggest prompts to achieve your task

Source: Deutsche Bank.

Shortened example prompts for various use cases

**Business**
- Identify 10 ways a finance company can use AI…
- You are a customer service rep. Write an apology for…
- Take this data [eg pasted table] and identify themes…
- Use this job description and cv to write a cover letter…
- Tell me which angel investors have invested in [name], give their email addresses and draft emails asking for…
- Write a business plan and slogan for a business that…

**Education**
- Summarise War and Peace in 500 words for kids…
- Teach me basic calculus with analogies and a quiz…

**Code**
- Write code to convert a video file into a smaller format…
- Explain the security flaw in this code snippet…

**General text and reasoning**
- Write a humorous poem welcoming wedding guests…
- I have this [list] in the fridge… Give me a healthy recipe…
- Parrot crosses a road. Complete this joke… Why funny?
- Compose a song about a fund manager inc chords…
- Why might my dog be sick? It has these symptoms…
- Write a five-day itinerary in a table format for a 30-year-old wheelchair user visiting Berlin for the first time…
What are the immediate limitations and concerns?

LLMs are not search engines, though some can be paired with search engines for greater accuracy. Their output reflects the limitations and biases of their training data, encompassing billions of words and images on the internet.

Five limitations and concerns about generative AI and the risks for enterprises

1. Hallucinations → Libel risk
2. Bias → Reputational risk
3. Data privacy → Security, confidentiality risk
4. Unclear sourcing → Legal risk, IP/copyright
5. Deliberate misuse → Operational risk

Additionally: Limited range of training data, eg GPT-4’s “knowledge cutoff” is Sept 2021

Source: Deutsche Bank, media headlines
What job tasks will be most affected?

Roles will be “unbundled” with a net effect of new work tasks, many higher value; occupations requiring science and critical skills are less likely to be impacted by LLMs while programming and writing roles are more likely to be impacted.

Effect of AI on 13 main tasks performed by customer service agents

- **Fully automated tasks, no longer exist**
- **Enhanced**
- **Unchanged**
- **New, high-value tasks**

Regression of occupation-level, human-annotated exposure to GPTs on skill importance

- Programming
- Writing
- Critical thinking
- Learning strategies
- Science

Exposure *:
- High
- Medium
- Low

Source: Customer service example: “Radically Human”, Paul Daugherty, H. James Wilson (2022); “GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models”, Eloundou et al. (2023), Deutsche Bank. Note: Low exposure (alpha in the original study) - lower bound of the share of exposed tasks within a role; medium (beta) – exposure when using the technology via additional tools and applications needs additional investment; high (zeta) – upper bound of exposure to LLMs. See the paper for in-depth description.
How are corporates beginning to use generative AI?

Early adopters have often been individuals doing creative tasks and coding; big tech is now integrating AI into existing products; companies may delay rolling out pure AI solutions until they can ensure quality, reliability and governance.

Six busy months of AI announcements and launches

**Nov 2022**
30: Open AI launches ChatGPT, built on the GPT-3.5 LLM

**Jan 2023**
23: Microsoft makes reported $10bn investment in OpenAI

**Feb 2023**
7: Microsoft launches new GPT-powered Bing search engine
24: Meta announces launch of lightweight LLaMA LLMs for researchers; it leaks a few days later and rapidly becomes the basis for numerous open-source uses

**March 2023**
14: OpenAI launches GPT-4, said in research to show sparks of Artificial General Intelligence; Anthropic launches Claude
21: Google begins widespread rollout of new Bard chatbot
21: Microsoft adds Image Creator to Bing

**May 2023**
2: Amazon said to be bringing AI chat capabilities to Alexa
10: Google presents AI-powered tools for Workspace inc “Help me write” for emails and image generation for Slides

Businesses that say they are using chatbots

**Instacart**: biggest US online delivery company
- Adding ChatGPT to shopping app, so customers can ask for recommendations for recipes, healthy food etc

**WPP**: world’s largest marketing services company
- AI is “fundamental” to future business, used to create ad campaigns and develop personalised advertisements

**Allen & Overy**: one of London’s “magic circle” law firms
- Using Harvey, built using OpenAI’s GPT tech, to help come up with first drafts of legal documents

**Zurich Insurance**: Swiss insurance company
- Investigating using ChatGPT to extract data from claims descriptions and write code for statistical models

**Duolingo**: language learning app
- Launched Duolingo Max offering tailored language lessons and features to roleplay and explain answers

**Salesforce**: software-as-a-service pioneer
- Introducing ChatGPT to Slack platform to summarise conversations, answer questions and draft responses

**Snap**: maker of Snapchat app
- Launched My AI, a customised ChatGPT, for users

**Mattel**: US toymaker, owner of Barbie and Hot Wheels
- Using DALL-E 2 image generator to design new toys

*Source: Media reports, company announcements, Deutsche Bank.*
How fast will things change?
AI text may match the best human writing soonest, followed by code and images, with video games and movies potentially becoming “personalised dreams” by the end of the decade.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Spam detection</td>
<td>Translation</td>
<td>Basic copy writing</td>
<td>Longer form</td>
<td>Vertical fine tuning</td>
<td>Final drafts better than the human average</td>
<td>Final drafts better than professional writers</td>
</tr>
<tr>
<td>Translation</td>
<td>Translation</td>
<td></td>
<td>Second drafts</td>
<td>gets good (scientific papers, etc)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic Q&amp;A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CODE</td>
<td>1-line auto-complete</td>
<td>Multi-line generation</td>
<td>Longer form</td>
<td>More languages</td>
<td>Text to product (draft)</td>
<td>Text to product (final), better than full-time developers</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Better accuracy</td>
<td>More verticals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMAGES</td>
<td>Art</td>
<td>Logos</td>
<td>Mock-ups (product design, architecture, etc.)</td>
<td>Final drafts (product design, architecture, etc.)</td>
<td>Final drafts better than professional artists, designers, photographers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Logos</td>
<td>Photography</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIDEO / 3D / GAMING</td>
<td>First attempts at 3D/video models</td>
<td>Basic / first draft videos and 3D files</td>
<td>Second drafts</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Large model availability:  
- First attempts
- Almost there
- Ready for prime time

Source: Sequoia Capital blog post: “A timeline for how we might expect to see fundamental models progress and the associated applications that become possible. 2025 and beyond is just a guess”
What are potential bottlenecks to AI development right now?

Some experts say that growth will slow as costs spiral and raw materials such as data and hardware become scarce.

1. Access to GPUs

There is an acute shortage of GPUs, almost all of which are made by Nvidia, for supercomputers used to train LLMs. Lead times are at least nine months and more usually 18 months for the new H100 chip. The OpenAI supercomputer used to train GPT-4 reportedly uses 10,000 of the older Nvidia A100 GPUs.

H100s have a retail price of around $10k but can trade on the grey market for as much as $40k.

Nvidia’s stock has almost doubled since ChatGPT launch.

2. Cost of training

Estimated training cost of select large language and multimodal models, $mn *

<table>
<thead>
<tr>
<th>Model</th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
<th>2022</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPT-2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>T5-11B</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Meena</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Wu Dao - Wen Yuan</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>DALL-E</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GPT-Neo</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GPT-J-6B</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>HyperCova</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>ERNIE 3.0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>T5-11B</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PaLM (540B)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Chinchilla</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Stable Diffusion</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>OPT-175B</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Minerva (540B)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>BLOOM</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GLM-130B</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GPT-NeoX-20B</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>AlphaCode</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Gopher</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Codex</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Megatron-Turing</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Wu Dao - Wen Yuan</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>DALL-E</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GPT-3 175B</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GPT-3 (175B)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Gopher (280B)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Gopher (280B)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>OPT (175B)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>OPT (175B)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Car, avg. incl. fuel, 1 lifetime</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>BLOOM (176B)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>American life, avg., 1 year</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Human life, avg., 1 year</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Air travel, 1 passenger, NY-SF</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Source: Bloomberg Finance LP; 2023 AI Index Report; Luccioni et al., 2022; Strubell et al., 2019; Deutsche Bank. * Note: low refers to potential underestimate, mid refers to what is believed to be a mid-level estimate and high refers to a potential overestimate.

3. Environmental costs

CO2 equivalent emissions (tonnes) by selected machine learnings models and real life examples, 2022

- GPT-3 (175B)
- Gopher (280B)
- OPT (175B)
- Car, avg. incl. fuel, 1 lifetime
- BLOOM (176B)
- American life, avg., 1 year
- Human life, avg., 1 year
- Air travel, 1 passenger, NY-SF
How do the general public and experts feel about AI?
Concerns outweighed excitement even before the launch of ChatGPT last November

Source: Pew Research, 2022; Michael et al. 2022; 2023 AI Index Report; Deutsche Bank.
Are we on the brink of Artificial General Intelligence? And is that good?
Experts differ widely between sceptics, who say AI is still rudimentary, despite appearances, and may never match human intelligence, and those who say AGI is imminent, for better or worse; there are also many overlaps

“Intelligence, by humans and machines, is what fuels progress… I think there’s going to be new ideas that are gonna make those systems much more controllable.”
Yann LeCun, Meta

“AI will lead to a low cost and better quality life for millions of people… Right now, I don’t see a clear path for AI to surpass human-level intelligence.”
Andrew Ng, AI entrepreneur and academic

“There are lots of dudes in our industry who struggle to tell the difference between science fiction and real life.”
Andrew Feldman, AI entrepreneur

“Right now, they’re not more intelligent than us, as far as I can tell. But I think they soon may be… So we need to worry about that.”
Geoff Hinton, academic and “godfather of AI”

“A fork is technology. I can use it to eat spaghetti or I can stab you in the hand… Saying “no more forks” means you miss out on all of the benefits.”
Matt Brittin, Google

“AI has much to offer humanity. It could also wreak terrible harm. It must be controlled.”
Stuart Russell, academic computer scientist

“If somebody builds a too-powerful AI, under present conditions, I expect that every single member of the human species and all biological life on Earth dies shortly thereafter.”
Eliezer Yudkowsky, academic

Source: Media reports, Deutsche Bank: NB: Categories are rough and flexible, for indicative purposes only. They reflect emphasis of specific quotes and may not fully reflect views of the speakers.
What are policymakers doing?
Governments and regulators are playing catchup with fast-evolving AI to set “guardrails” to regulate and oversee it without discouraging investment and innovation and losing a competitive advantage.

Challenges

“Like trying to nail Jell-O to the wall.”
Bill Clinton, US President, 2000, on China’s attempts to crack down on the internet

1) Technology still emerging
   - Hard to predict future risk and benefits; new laws may be obsolete by the time are enacted

2) Lack of deep understanding
   - Asymmetry between lawmakers/public and industry/academia

3) Cross-sectoral application
   - AI affects multiple industries with varying risks and impacts

Divergent approaches emerging

US/UK
- Light touch, risk-based regulation
- Applying existing regulations
- No new regulatory bodies
- Sector-specific, focusing on use of AI rather than the technology itself

European Union
- Multifaceted inc specific legislation
- Proposed AI Act in legislative process is detailed and prescriptive, eg classifying AI uses by degrees of risk; minimum requirements for AI systems; new national/EU bodies
- Builds on tech rules such as GDPR

China
- Piecemeal approach
- Focus on security, eg: mandatory security assessment for new products; must meet China’s values

Notable developments

Jan
- US Dept of Commerce releases voluntary AI risk framework

March
- More than 1,000 tech leaders inc Elon Musk publish letter calling for “pause” in AI development
- Italy temporarily bans ChatGPT
- UK publishes AI policy paper calling for “pro-innovation approach”

April
- China’s internet regulator proposes new rules applying to generative AI
- Reports that US Senator Schumer is developing AI regulation plan

May
- OpenAI CEO Sam Altman appears at US Senate hearing
- G7 leaders call for AI “guardrails”

Upcoming...
- Late 2023: EU AI Act set to pass

Source: Government and regulatory announcements, media reports, Deutsche Bank.
What happens next?
The rapid pace of eye-catching launches is likely to slow as the focus turns to refining and integrating the new tech into infrastructure, governance and culture; industry and governments will need to work together to avoid a backlash.

Users
- Individuals: rapid adoption for personal/semi-professional use
- Companies: FOMO (fear of missing out) balanced by governance concerns; increasing use in customer services, data retrieval, content creation, coding
- General: increasingly sophisticated use of tech (eg plugins) and techniques (eg chained prompts through eg AutoGPT for complex multi-step tasks)

App developers
- Explosion in new applications built on top of LLMs
- Trying out of open-source models, fine-tuned and stacked for specific tasks

LLM developers
- Race to offer integrated enterprise solutions with reliable, ringfenced data
- Further advances in 1-trillion+ parameter multimodal LLMs, notably GPT-5, rumoured to finish training by year-end and to have video capabilities
- Parallel advances in open-source lightweight alternative LLMs that offer similar performance to large models, trained more efficiently and cheaply with pruned datasets, and that users can run on their own hardware

Regulators and society
Widening anxiety threatening a potential backlash with concerns about:
- Jobs: some white-collar jobs will be enhanced while others may be at risk
- Education: eg UK teacher coalition warned in May of “severe” threat
- Privacy: Spain, Canada and others have considered or started probes
- Truth: US presidential election in 2024 will be a litmus test of AI-enabled hyper-personalised campaigning and potential misinformation by “bad actors”
- Ethics: “alignment” to human values

Pressure to legislate and retrain
- Policymakers caught between desire for popular precautions and a “race to the bottom” to encourage home-grown development and use
- Governments will be urged to provide financial and educational safety nets for those displaced by AI to mitigate social upheaval

Users
- Individuals: rapid adoption for personal/semi-professional use
- Companies: FOMO (fear of missing out) balanced by governance concerns; increasing use in customer services, data retrieval, content creation, coding
- General: increasingly sophisticated use of tech (eg plugins) and techniques (eg chained prompts through eg AutoGPT for complex multi-step tasks)

App developers
- Explosion in new applications built on top of LLMs
- Trying out of open-source models, fine-tuned and stacked for specific tasks

LLM developers
- Race to offer integrated enterprise solutions with reliable, ringfenced data
- Further advances in 1-trillion+ parameter multimodal LLMs, notably GPT-5, rumoured to finish training by year-end and to have video capabilities
- Parallel advances in open-source lightweight alternative LLMs that offer similar performance to large models, trained more efficiently and cheaply with pruned datasets, and that users can run on their own hardware

Hardware
- Ongoing pressure on supplies of key GPUs and networking equipment

Source: Deutsche Bank.
Appendix 1
Important Disclosures
*Other information available upon request

*Prices are current as of the end of the previous trading session unless otherwise indicated and are sourced from local exchanges via Reuters, Bloomberg and other vendors. Other information is sourced from Deutsche Bank, subject companies, and other sources. For further information regarding disclosures relevant to Deutsche Bank Research, please visit our global disclosure look-up page on our website at https://research.db.com/Research/Disclosures/FICCDisclosures. Aside from within this report, important risk and conflict disclosures can also be found at https://research.db.com/Research/Disclosures/Disclaimer. Investors are strongly encouraged to review this information before investing.

Analyst Certification
The views expressed in this report accurately reflect the personal views of the undersigned lead analyst about the subject issuers and the securities of those issuers. In addition, the undersigned lead analyst has not and will not receive any compensation for providing a specific recommendation or view in this report. Adrian Cox, Jim Reid, Galina Pozdnyakova
The information and opinions in this report were prepared by Deutsche Bank AG or one of its affiliates (collectively “Deutsche Bank”). Though the information herein is believed to be reliable and has been obtained from public sources believed to be reliable, Deutsche Bank makes no representation as to its accuracy or completeness. Hyperlinks to third-party websites in this report are provided for reader convenience only. Deutsche Bank neither endorses the content nor is responsible for the accuracy or security controls of those websites.

If you use the services of Deutsche Bank in connection with a purchase or sale of a security that is discussed in this report, or is included or discussed in another communication (oral or written) from a Deutsche Bank analyst, Deutsche Bank may act as principal for its own account or as agent for another person.

Deutsche Bank may consider this report in deciding to trade as principal. It may also engage in transactions, for its own account or with customers, in a manner inconsistent with the views taken in this research report. Others within Deutsche Bank, including strategists, sales staff and other analysts, may take views that are inconsistent with those taken in this research report. Deutsche Bank issues a variety of research products, including fundamental analysis, equity-linked analysis, quantitative analysis and trade ideas. Recommendations contained in one type of communication may differ from recommendations contained in others, whether as a result of differing time horizons, methodologies, perspectives or otherwise. Deutsche Bank and/or its affiliates may also be holding debt or equity securities of the issuers it writes on. Analysts are paid in part based on the profitability of Deutsche Bank AG and its affiliates, which includes investment banking, trading and principal trading revenues.

Opinions, estimates and projections constitute the current judgment of the author as of the date of this report. They do not necessarily reflect the views of Deutsche Bank or its affiliates. Deutsche Bank provides liquidity for buyers and sellers of securities issued by the companies it covers. Deutsche Bank research analysts sometimes have shorter-term trade ideas that may be inconsistent with Deutsche Bank’s existing longer-term ratings. Some trade ideas for equities are listed as Catalyst Calls ideas on the Research Website (https://research.db.com/Researchv), and can be found on the general coverage list and also on the covered company’s page. A Catalyst Call idea represents a high-conviction belief by an analyst that a stock will outperform or underperform the market and/or a specified sector over a time frame of no less than two weeks and no more than three months. In addition to Catalyst Calls ideas, analysts may occasionally discuss with our clients, and with Deutsche Bank salespersons and traders, trading strategies or ideas that reference catalysts or events that may have a near-term or medium-term impact on the market price of the securities discussed in this report, which may be directionally counter to the analysts’ current 12-month view of total return or investment return as described herein. Deutsche Bank has no obligation to update, modify or amend this report or to otherwise notify a recipient thereof if an opinion, forecast or estimate changes or becomes inaccurate. Coverage and the frequency of changes in market conditions and in both general and company-specific economic prospects make it difficult to update research at defined intervals. Updates are at the sole discretion of the coverage analyst or of the Research Department Management, and the majority of reports are published at irregular intervals. This report is provided for informational purposes only and does not take into account the particular investment objectives, financial situations, or needs of individual clients. It is not an offer or a solicitation of an offer to buy or sell any financial instruments or to participate in any particular trading strategy. Target prices are inherently imprecise and a product of the analyst’s judgment. The financial instruments discussed in this report may not be suitable for all investors, and investors must make their own informed investment decisions. Prices and availability of financial instruments are subject to change without notice, and investment transactions can lead to losses as a result of price fluctuations and other factors. If a financial instrument is denominated in a currency other than an investor’s currency, a change in exchange rates may adversely affect the investment. Past performance is not necessarily indicative of future results. Performance calculations exclude transaction costs, unless otherwise indicated. Unless otherwise indicated, prices are current as of the end of the previous trading session and are sourced from local exchanges via Reuters, Bloomberg and other vendors. Data is also sourced from Deutsche Bank, subject companies, and other parties.

The Deutsche Bank Research Department is independent of other business divisions of the Bank. Details regarding our organizational arrangements and information barriers we have to prevent and avoid conflicts of interest with respect to our research are available on our website (https://research.db.com/Researchv) under Disclaimer.

Macroeconomic fluctuations often account for most of the risks associated with exposures to instruments that promise to pay fixed or variable interest rates. For an investor who is long fixed-rate instruments (thus receiving these cash flows), increases in interest rates naturally lift the discount factors applied to the expected cash flows and thus cause a loss. The longer the maturity of a certain cash flow and the higher the move in the discount factor, the higher will be the loss. Upside surprises in inflation, fiscal funding needs, and FX depreciation rates are among the most common adverse macroeconomic shocks to receivers. But counterparty exposure, issuer creditworthiness, client segmentation, regulation (including changes in assets holding limits for different types of investors), changes in tax policies, currency convertibility (which may constrain currency conversion, repatriation of profits and/or liquidation of positions), and settlement issues related to local clearing houses are also important risk factors. The sensitivity of fixed-income instruments to macroeconomic shocks may be mitigated by indexing the contracted cash flows to inflation, to FX depreciation, or to specified interest rates – these are common in emerging markets. The index fixings may – by construction – lag or mis-measure the actual move in the underlying variables they are intended to track. The choice of the proper fixing (or metric) is particularly important in swaps markets, where floating coupon rates (i.e., coupons indexed to a typically short-dated interest rate reference index) are exchanged for fixed coupons. Funding in a currency that differs from the currency in which coupons are denominated carries FX risk. Options on swaps (swaptions) the risks typical to options in addition to the risks related to rates movements.
Derivative transactions involve numerous risks including market, counterparty default and illiquidity risk. The appropriateness of these products for use by investors depends on the investors’ own circumstances, including their tax position, their regulatory environment and the nature of their other assets and liabilities; as such, investors should take expert legal and financial advice before entering into any transaction similar to or inspired by the contents of this publication. The risk of loss in futures trading and options, foreign or domestic, can be substantial. As a result of the high degree of leverage obtainable in futures and options trading, losses may be incurred that are greater than the amount of funds initially deposited – up to theoretically unlimited losses. Trading in options involves risk and is not suitable for all investors. Prior to buying or selling an option, investors must review the “Characteristics and Risks of Standardized Options”, at http://www.optionsclearing.com/about/publications/character-risks.jsp. If you are unable to access the website, please contact your Deutsche Bank representative for a copy of this important document.

Participants in foreign exchange transactions may incur risks arising from several factors, including the following: (i) exchange rates can be volatile and are subject to large fluctuations; (ii) the value of currencies may be affected by numerous market factors, including world and national economic, political and regulatory events, events in equity and debt markets and changes in interest rates; and (iii) currencies may be subject to devaluation or government-imposed exchange controls, which could affect the value of the currency. Investors in securities such as ADRs, whose values are affected by the currency of an underlying security, effectively assume currency risk.

Unless governing law provides otherwise, all transactions should be executed through the Deutsche Bank entity in the investor’s home jurisdiction. Aside from within this report, important conflict disclosures can also be found at https://research.db.com/Research/ on each company’s research page. Investors are strongly encouraged to review this information before investing.

Deutsche Bank (which includes Deutsche Bank AG, its branches and affiliated companies) is not acting as a financial adviser, consultant or fiduciary to you or any of your agents (collectively, “You” or “Your”) with respect to any information provided in this report. Deutsche Bank does not provide investment, legal, tax or accounting advice, Deutsche Bank is not acting as your impartial adviser, and does not express any opinion or recommendation whatsoever as to any strategies, products or any other information presented in the materials. Information contained herein is being provided solely on the basis that the recipient will make an independent assessment of the merits of any investment decision, and it does not constitute a recommendation of, or express an opinion on, any product or service or any trading strategy.

The information presented is general in nature and is not directed to retirement accounts or any specific person or account type, and is therefore provided to You on the express basis that it is not advice, and You may not rely upon it in making Your decision. The information we are providing is directed only to persons we believe to be financially sophisticated, who are capable of evaluating investment risks independently, both in general and with regard to particular transactions and investment strategies, and who understand that Deutsche Bank has financial interests in the offering of its products and services. If this is not the case, or if You are an IRA or other retail investor receiving this directly from us, we ask that you inform us immediately.

In July 2018, Deutsche Bank revised its rating system for short term ideas whereby the branding has been changed to Catalyst Calls (“CC”) from SOLAR ideas; the rating categories for Catalyst Calls originated in the Americas region have been made consistent with the categories used by Analysts globally; and the effective time period for CCs has been reduced from a maximum of 180 days to 90 days.

United States: Approved and/or distributed by Deutsche Bank Securities Incorporated, a member of FINRA, NFA and SIPC. Analysts located outside the United States are employed by non-US affiliates that are not subject to FINRA regulations.

European Economic Area (exc. United Kingdom): Approved and/or distributed by Deutsche Bank AG, a joint stock company with limited liability incorporated in the Federal Republic of Germany with its principal office in Frankfurt am Main. Deutsche Bank AG is authorized under German Banking Law and is subject to supervision by the European Central Bank and by BaFin, Germany’s Federal Financial Supervisory Authority.

United Kingdom: Approved and/or distributed by Deutsche Bank AG acting through its London Branch at Winchester House, 1 Great Winchester Street, London EC2N 2DB. Deutsche Bank AG in the United Kingdom is authorised by the Prudential Regulation Authority and is subject to limited regulation by the Prudential Regulation Authority and Financial Conduct Authority. Details about the extent of our authorisation and regulation are available on request.

Hong Kong SAR: Distributed by Deutsche Bank AG, Hong Kong Branch except for any research content relating to futures contracts within the meaning of the Hong Kong Securities and Futures Ordinance Cap. 571. Research reports on such futures contracts are not intended for access by persons who are located, incorporated, constituted or resident in Hong Kong. The author(s) of a research report may not be licensed to carry on regulated activities in Hong Kong and, if not licensed, do not hold themselves out as being able to do so. The provisions set out above in the “Additional Information” section shall apply to the fullest extent permissible by local laws and regulations, including without limitation the Code of Conduct for Persons Licensed or Registered with the Securities and Futures Commission. This report is intended for distribution only to “professional investors” as defined in Part 1 of Schedule 1 of the SFO. This document must not be acted or relied on by persons who are not professional investors. Any investment or investment activity to which this document relates is only available to professional investors and will be engaged only with professional investors.
India: Prepared by Deutsche Equities India Private Limited (DEIPL) having CIN: U65990MH2002PTC137431 and registered office at 14th Floor, The Capital, C-70, G Block, Bandra Kurla Complex Mumbai (India) 400051. Tel: + 91 22 7180 4444. It is registered by the Securities and Exchange Board of India (SEBI) as a Stock broker bearing registration no.: INZ0002S2437, Merchant Banker bearing SEBI Registration no.: INM000010833 and Research Analyst bearing SEBI Registration no.: INH000001741. DEIPL may have received administrative warnings from the SEBI for breaches of Indian regulations. Deutsche Bank and/or its affiliate(s) may have debt holdings or positions in the subject company. With regard to information on associates, please refer to the “Shareholdings” section in the Annual Report at: https://www.db.com/in/en/annual-reports.htm

Japan: Approved and/or distributed by Deutsche Securities Inc. (DSI). Registration number - Registered as a financial instruments dealer by the Head of the Kanto Local Finance Bureau (Kinsho) No. 117. Member of associations: JISDA, Type II Financial Instruments Firms Association and The Financial Futures Association of Japan. Commissions and risks involved in stock transactions - for stock transactions, we charge stock commissions and consumption tax by multiplying the transaction amount by the commission rate agreed with each customer. Stock transactions can lead to losses as a result of share price fluctuations and other factors. Transactions in foreign stocks can lead to additional losses stemming from foreign exchange fluctuations. We may also charge commissions and fees for certain categories of investment advice, products and services. Recommended investment strategies, products and services carry the risk of losses to principal and other losses as a result of changes in market and/or economic trends, and/or fluctuations in market value. Before deciding on the purchase of financial products and/or services, customers should carefully read the relevant disclosures, prospectuses and other documentation. “Moody’s”, “Standard & Poor’s”, and “Fitch” mentioned in this report are not registered credit rating agencies in Japan unless Japan or “Nippon” is specifically designated in the name of the entity. Reports on Japanese listed companies not written by analysts of DSI are written by Deutsche Bank Group’s analysts with the coverage companies specified by DSI. Some of the foreign securities stated on this report are not disclosed according to the Financial Instruments and Exchange Law of Japan. Target prices set by Deutsche Bank’s equity analysts are based on a 12-month forecast period.

Korea: Distributed by Deutsche Securities Korea Co.


Singapore: This report is issued by Deutsche Bank AG, Singapore Branch (One Raffles Quay #18-00 South Tower Singapore 048583, +65 6423 8001), which may be contacted in respect of any matters arising from, or in connection with, this report. Where this report is issued or promulgated by Deutsche Bank in Singapore to a person who is not an accredited investor, expert investor or institutional investor (as defined in the applicable Singapore laws and regulations), they accept legal responsibility to such person for its contents.

Taiwan: Information on securities/investments that trade in Taiwan is for your reference only. Readers should independently evaluate investment risks and are solely responsible for their investment decisions. Deutsche Bank research may not be distributed to the Taiwan public media or quoted or used by the Taiwan public media without written consent. Information on securities/instruments that do not trade in Taiwan is for informational purposes only and is not to be construed as a recommendation to trade in such securities/instruments. Deutsche Securities Asia Limited, Taipei Branch may not execute transactions for clients in these securities/instruments.

Qatar: Deutsche Bank AG in the Qatar Financial Centre (registered no. 00032) is regulated by the Qatar Financial Centre Regulatory Authority. Deutsche Bank AG - QFC Branch may undertake only the financial services activities that fall within the scope of its existing QFCRA license. Its principal place of business in the QFC: Qatar Financial Centre, Tower, West Bay, Level 5, PO Box 14928, Doha, Qatar. This information has been distributed by Deutsche Bank AG. Related financial products or services are only available only to Business Customers, as defined by the Qatar Financial Centre Regulatory Authority.

Russia: The information, interpretation and opinions submitted herein are not in the context of, and do not constitute, any appraisal or evaluation activity requiring a license in the Russian Federation.

Kingdom of Saudi Arabia: Deutsche Securities Saudi Arabia (DSSA) is a closed joint stock company authorized by the Capital Market Authority of the Kingdom of Saudi Arabia with a license number (No. 37-07073) to conduct the following business activities: Dealing, Arranging, Advising, and Custody activities. DSSA registered office is at Faisalah Tower, 17th floor, King Fahad Road - Al Olaya District Riyadh, Kingdom of Saudi Arabia P.O. Box 301806.

United Arab Emirates: Deutsche Bank AG in the Dubai International Financial Centre (registered no. 00045) is regulated by the Dubai Financial Services Authority. Deutsche Bank AG - DIFC Branch may only undertake the financial services activities that fall within the scope of its existing DFSA license. Principal place of business in the DIFC: Dubai International Financial Centre, The Gate Village, Building 5, PO Box 504902, Dubai, U.A.E. This information has been distributed by Deutsche Bank AG. Related financial products or services are only available only to Professional Clients, as defined by the Dubai Financial Services Authority.
Australia and New Zealand: This research is intended only for "wholesale clients" within the meaning of the Australian Corporations Act and New Zealand Financial Advisers Act, respectively. Please refer to Australian specific research disclosures and related information at https://www.dbresearch.com/PROD/RPS_EN-PROD/PROD00000000000521304.xhtml. Where research refers to any particular financial product recipients of the research should consider any product disclosure statement, prospectus or other applicable disclosure document before making any decision about whether to acquire the product. In preparing this report, the primary analyst or an individual who assisted in the preparation of this report has likely been in contact with the company that is the subject of this research for confirmation/clarification of data, facts, statements, permission to use company-sourced material in the report, and/or site-visit attendance. Without prior approval from Research Management, analysts may not accept from current or potential Banking clients the costs of travel, accommodations, or other expenses incurred by analysts attending site visits, conferences, social events, and the like. Similarly, without prior approval from Research Management and Anti-Bribery and Corruption ("ABC") team, analysts may not accept perks or other items of value for their personal use from issuers they cover.

Additional information relative to securities, other financial products or issuers discussed in this report is available upon request. This report may not be reproduced, distributed or published without Deutsche Bank’s prior written consent.

Backtested, hypothetical or simulated performance results have inherent limitations. Unlike an actual performance record based on trading actual client portfolios, simulated results are achieved by means of the retroactive application of a backtested model itself designed with the benefit of hindsight. Taking into account historical data the backtesting of performance also differs from actual account performance because an actual investment strategy may be adjusted any time, for any reason, including a response to material, economic or market factors. The backtested performance includes hypothetical results that do not reflect the reinvestment of dividends and other earnings or the deduction of advisory fees, brokerage or other commissions, and any other expenses that a client would have paid or actually paid. No representation is made that any trading strategy or account will or is likely to achieve profits or losses similar to those shown. Alternative modeling techniques or assumptions might produce significantly different results and prove to be more appropriate. Past hypothetical backtest results are neither an indicator nor guarantee of future returns. Actual results will vary, perhaps materially, from the analysis.

The method for computing individual ESG and composite ESG scores set forth herein is a novel method developed by the Research department within Deutsche Bank AG, computed using a systematic approach without human intervention. Different data providers, market sectors and geographies approach ESG analysis and incorporate the findings in a variety of ways. As such, the ESG scores referred to herein may differ from equivalent ratings developed and implemented by other ESG data providers in the market and may also differ from equivalent ratings developed and implemented by other divisions within the Deutsche Bank Group. Such ESG scores also differ from other ratings and rankings that have historically been applied in research reports published by Deutsche Bank AG. Further, such ESG scores do not represent a formal or official view of Deutsche Bank AG.

It should be noted that the decision to incorporate ESG factors into any investment strategy may inhibit the ability to participate in certain investment opportunities that otherwise would be consistent with your investment objective and other principal investment strategies. The returns on a portfolio consisting primarily of sustainable investments may be lower or higher than portfolios where ESG factors, exclusions, or other sustainability issues are not considered, and the investment opportunities available to such portfolios may differ. Companies may not necessarily meet high performance standards on all aspects of ESG or sustainable investing issues; there is also no guarantee that any company will meet expectations in connection with corporate responsibility, sustainability, and/or impact performance.

Copyright © 2023 Deutsche Bank AG