Lisbon School of Economics & Management
Universidade de Lisboa

LISBOA
UNIVERSIDADE DE LISBOA

Generative AI

Carlos J. Costa

- Defining GenAI
- Types of generative AI models
- Prompts and Prompt engineering
- Ethics and AI

Generative AI

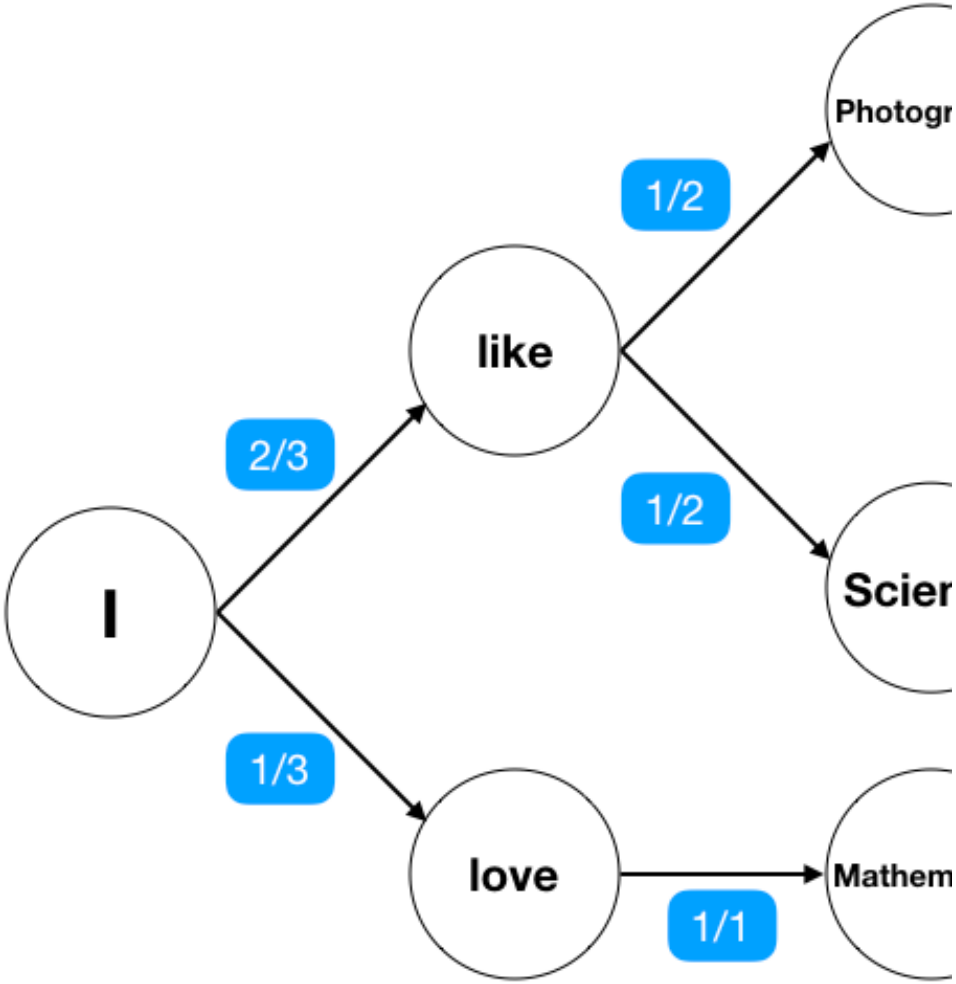
- Class of AI algorithms and models that are designed to generate new, original content.
- Gen AI learn the underlying patterns and structures in the data and can generate novel outputs.
- *Instead of being trained on specific examples and then making predictions or classifications*
- These models are particularly good at creating content that resembles or is similar to the data they were trained on.

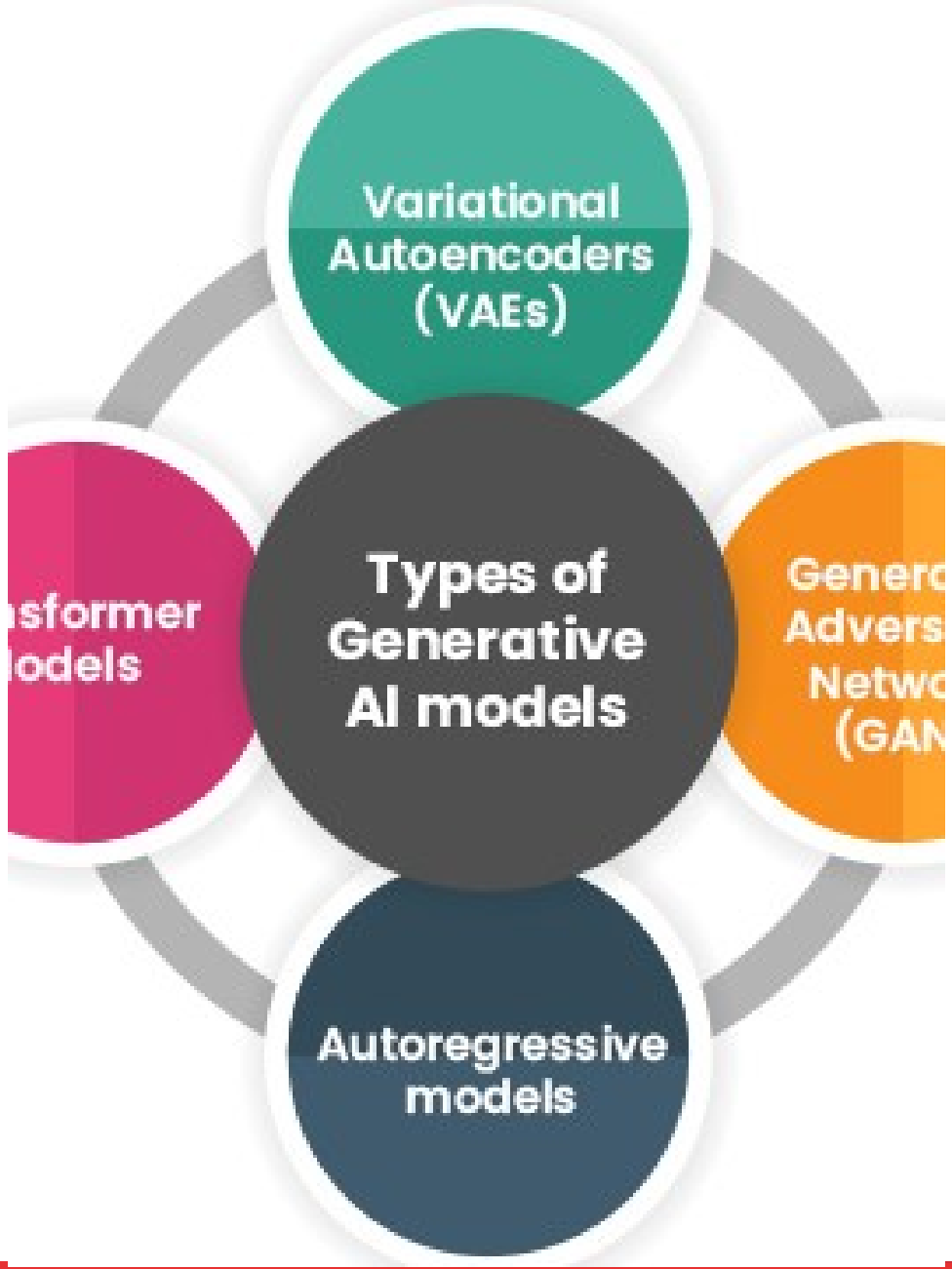
Generative AI

Guessing next word

Markov Chain

Training model

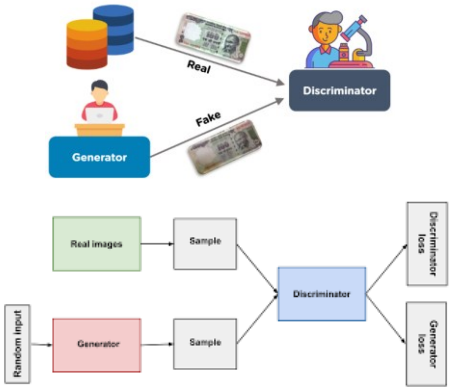


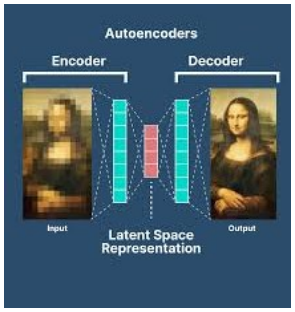


Types of generative AI models

- Generative Adversarial Networks (GANs)
- Variational Autoencoders (VAEs)
- Autoregressive Models
- Recurrent Neural Networks (RNNs)
- Transformer-based Models
- Reinforcement Learning for Generative Tasks

Types of generative AI models



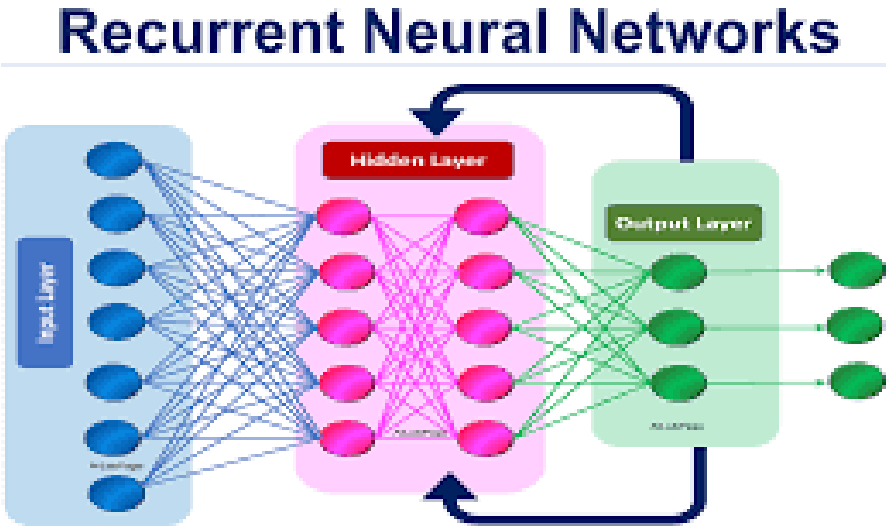


$$y_t = c + \sum_{i=1}^{p} a_{t-i} y_{y-i} + e_t$$

- Generative Adversarial Networks (GANs):
  - a **generator** and a **discriminator** are trained simultaneously through adversarial training.
- Variational Autoencoders (VAEs):
  - learn a probabilistic mapping from the observed data to a latent space.
  - Good to generate new samples from the learned latent space.
- Autoregressive Models:
  - the probability distribution of the next value in a sequence depends on the previous values.

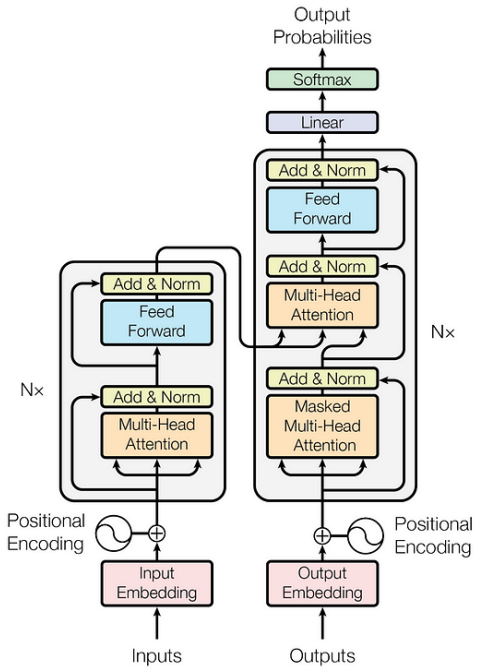
Types of generative AI models

- Recurrent Neural Networks (RNNs):
  - RNNs are commonly used for sequence tasks, including some generative tasks, they are not exclusively generative models.
  - Variants like **LSTM** and GRU are popular choices.
- Transformer-based Models:
  - Transformers, especially large language models.
- Reinforcement Learning for Generative Tasks:
  - can be used in conjunction with generative models, and this combination is powerful in scenarios where the generative model needs to produce sequences or structures guided by a reward signal.



BERT

Encoder

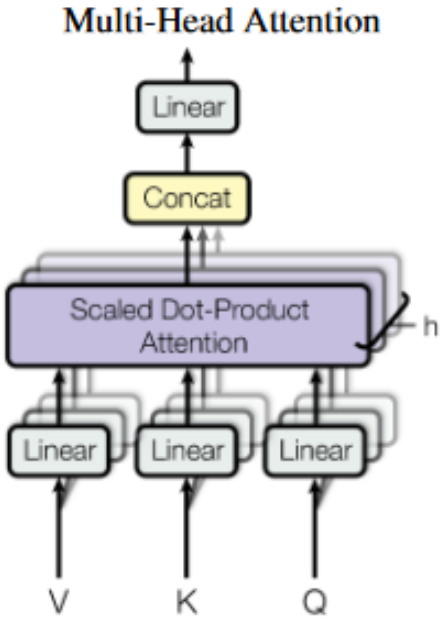


GPT

Decoder

Transformer

- Deep learning architecture based on the multi-head attention mechanism



Attention Is All You Need

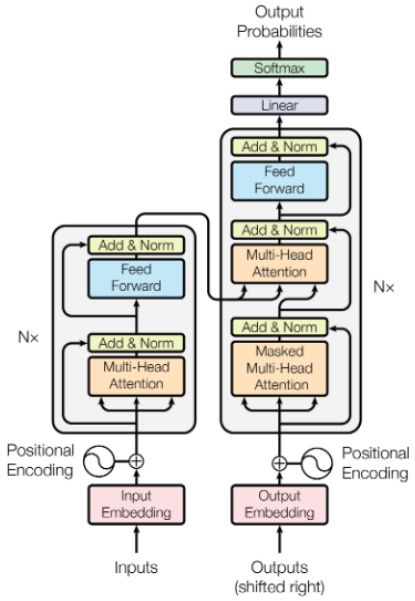
Ashish Vaswani* Google Brain avaswani@google.com
Noam Shazeer* Google Brain noam@google.com
Niki Parmar* Google Research nikip@google.com
Jakob Uszkoreit* Google Research usz@google.com
Llion Jones* Google Research llion@google.com
Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu
Łukasz Kaiser* Google Brain lukaszkaiser@google.com
Illia Polosukhin* ‡ illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.0 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature.

1 Introduction

Recurrent neural networks, long short-term memory [12] and gated recurrent [7] neural networks in particular, have been firmly established as state of the art approaches in sequence modeling and transduction problems such as language modeling and machine translation [29, 2, 5]. Numerous efforts have since continued to push the boundaries of recurrent language models and encoder-decoder architectures [31, 21, 13].



Vaswani, et al. (2017)

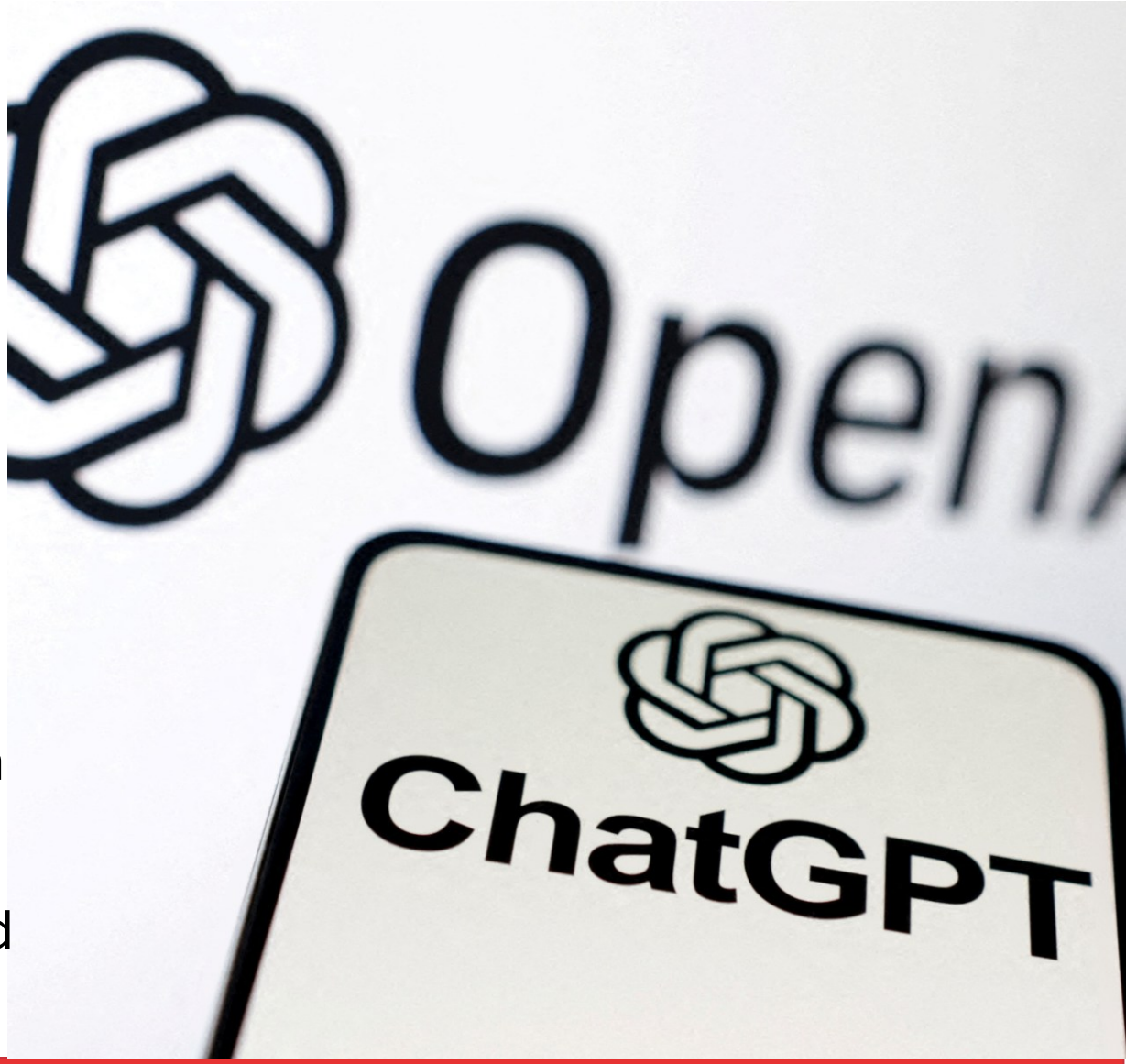
RAG

- retrieval-augmented generation
- is an AI framework for retrieving facts from an external knowledge base to ground large language models (LLMs) on the most accurate, up-to-date information and to give users insight into LLMs' generative process

Source: IBM

GPT

- Generative Pre-trained Transformer
- Is a type of autoregressive language model that uses a transformer architecture.
- Is pre-trained on a large corpus of text data and can then be fine-tuned for specific tasks.



Feature	LaMDA	PaLM	Gemini
Release Date	2021	2022	December 2023
Focus	Conversational AI	General-purpose	Multimodal
Strengths	Realistic dialogue	Large & diverse dataset	Understanding & processing various data formats
Successor	Gemini/ PaLM	Gemini	N/A

Google Gemini

Bard is a conversational AI chatbot powered by a combination of generative AI techniques, including:

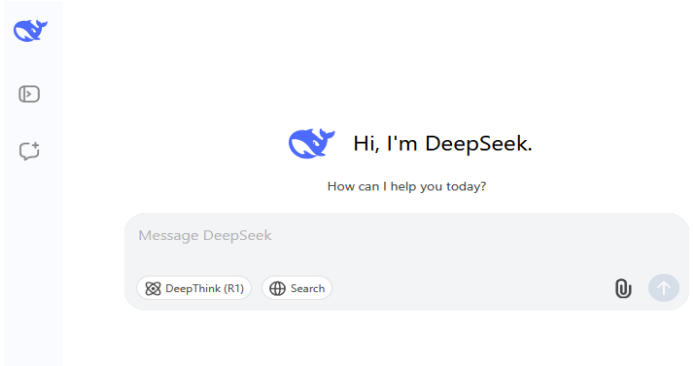
- **Transformer-based models:**
  - Google's Pathways Language Model (PaLM) is used to generate text that is fluent, coherent, and grammatically correct.
- **Autoregressive models**
  - to predict the next word in a sequence, which helps to ensure that its responses are natural and engaging.
- **Reinforcement learning:**
  - it is rewarded for generating responses that are informative, comprehensive, and relevant to the user's query.

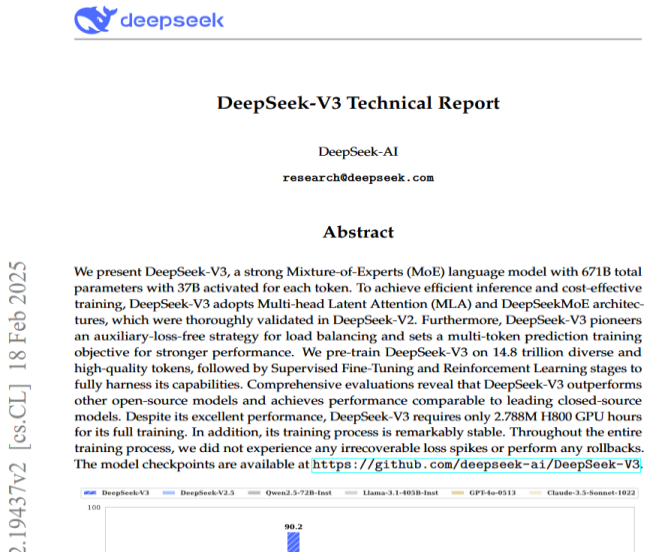
Copilot

- Generative artificial intelligence chatbot
- Developed by Microsoft

DeepSeek

- Founded by Liang Wenfeng
- Headquarters: Hangzhou, Zhejiang, China




deepseek

DeepSeek-V3 Technical Report

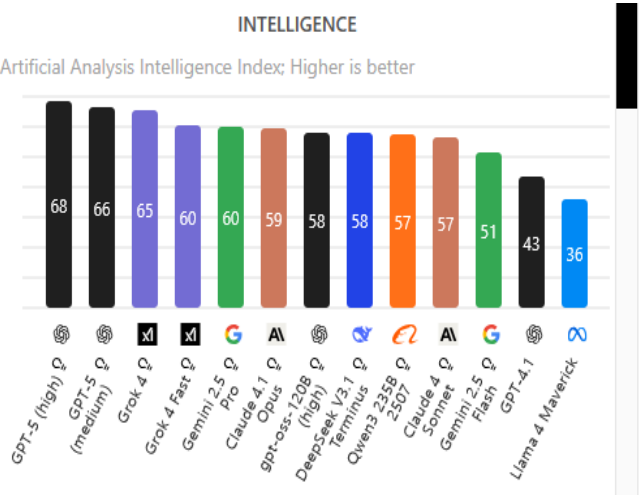
DeepSeek-AI

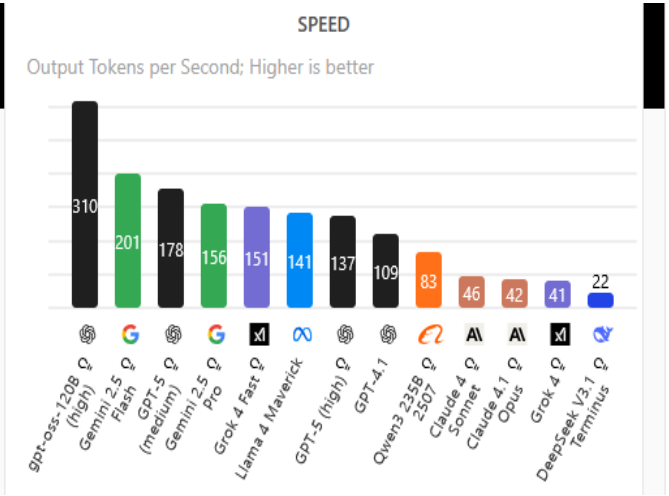
research@deepseek.com

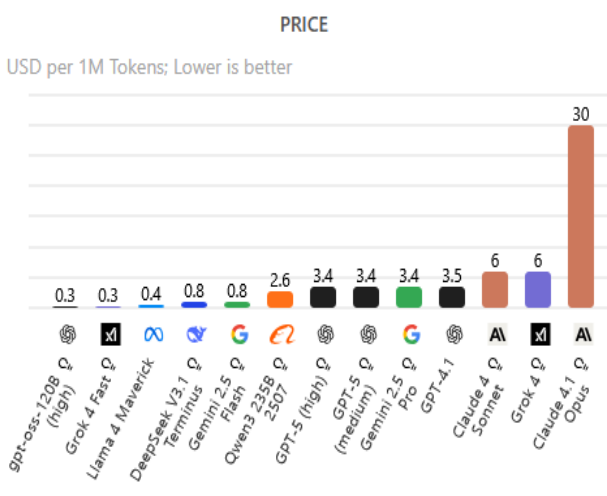
Abstract

We present DeepSeek-V3, a strong Mixture-of-Experts (MoE) language model with 671B total parameters with 37B activated for each token. To achieve efficient inference and cost-effective training, DeepSeek-V3 adopts Multi-head Latent Attention (MLA) and DeepSeekMoE architectures, which were thoroughly validated in DeepSeek-V2. Furthermore, DeepSeek-V3 pioneers an auxiliary-loss-free strategy for load balancing and sets a multi-token prediction training objective for stronger performance. We pre-train DeepSeek-V3 on 14.8 trillion diverse and high-quality tokens, followed by Supervised Fine-Tuning and Reinforcement Learning stages to fully harness its capabilities. Comprehensive evaluations reveal that DeepSeek-V3 outperforms other open-source models and achieves performance comparable to leading closed-source models. Despite its excellent performance, DeepSeek-V3 requires only 2.788M H800 GPU hours for its full training. In addition, its training process is remarkably stable. Throughout the entire training process, we did not experience any irrecoverable loss spikes or perform any rollbacks. The model checkpoints are available at https://github.com/deepseek-ai/DeepSeek-V3.

Comparing Models



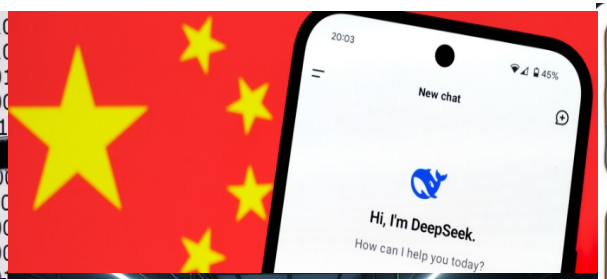




- https://artificialanalysis.ai/

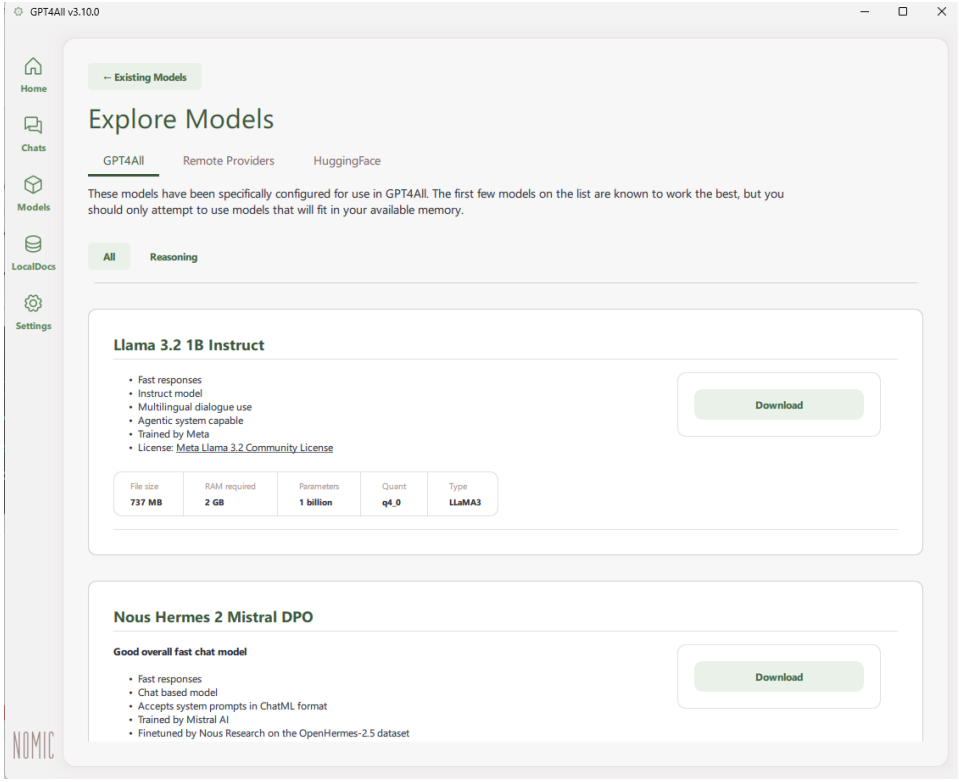
Some issues...

- Hidden costs
- Ethical control vs. Political censorship
- Corporations or government are stealing data?
- Bad practices



Your Own LLM locally

gpt4all




EXAMPLE

Prompts

- “Prompts are your input into the AI system to obtain specific results.
- In other words, prompts are conversation starters: what and how you tell something to the AI for it to respond in a way that generates useful responses for you.
- After that, you can build a continuing prompt, and the AI will produce another response accordingly.
- It’s like having a conversation with another person, only in this case the conversation is text-based, and your interlocutor is AI.”

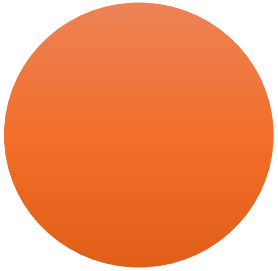
- https://mitsloanedtech.mit.edu/ai/basics/effective-prompts/

Prompt Engineering

- is the practice of designing, formulating, and refining input prompts to guide the behavior of large language models (LLMs) or other generative AI systems toward producing desired outputs.

- Use Copilot (login ISEG)
- Create a text describing ISEG in 2050
- Create also an image illustrating the text

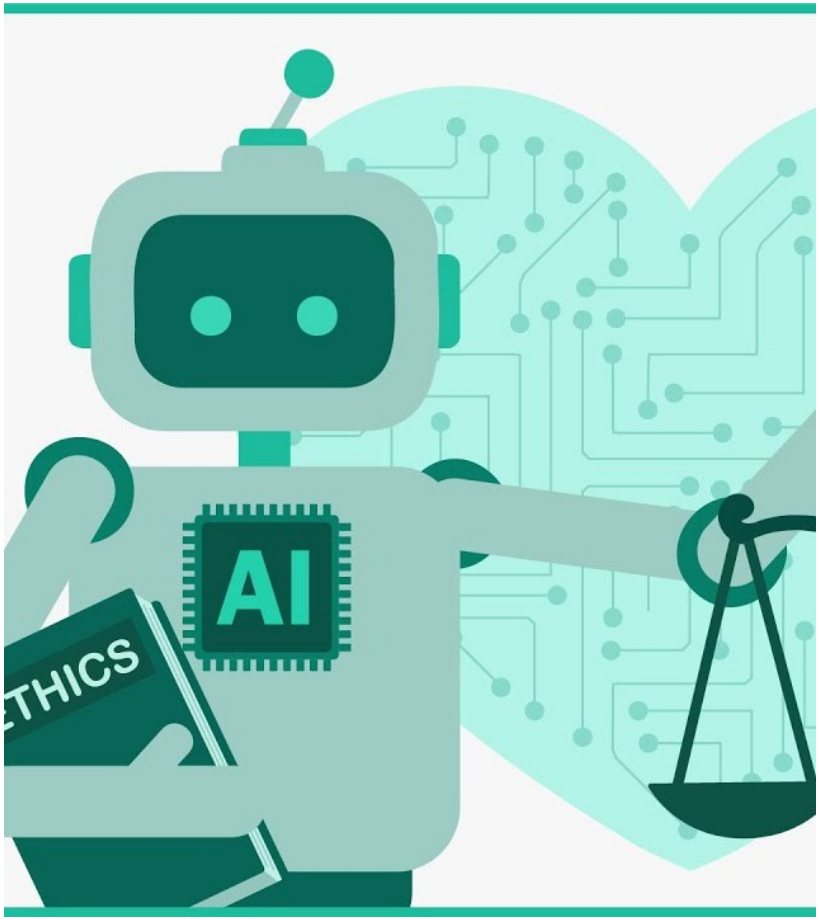
Effective Prompts



- Give **context** (set the role, scenario, or background)
- Be **specific** (clear instructions)
- Show **format**/structure
- Provide examples/**guidance**
- **Iterate**/refine
- Verify/**evaluate**

- Improve previous example

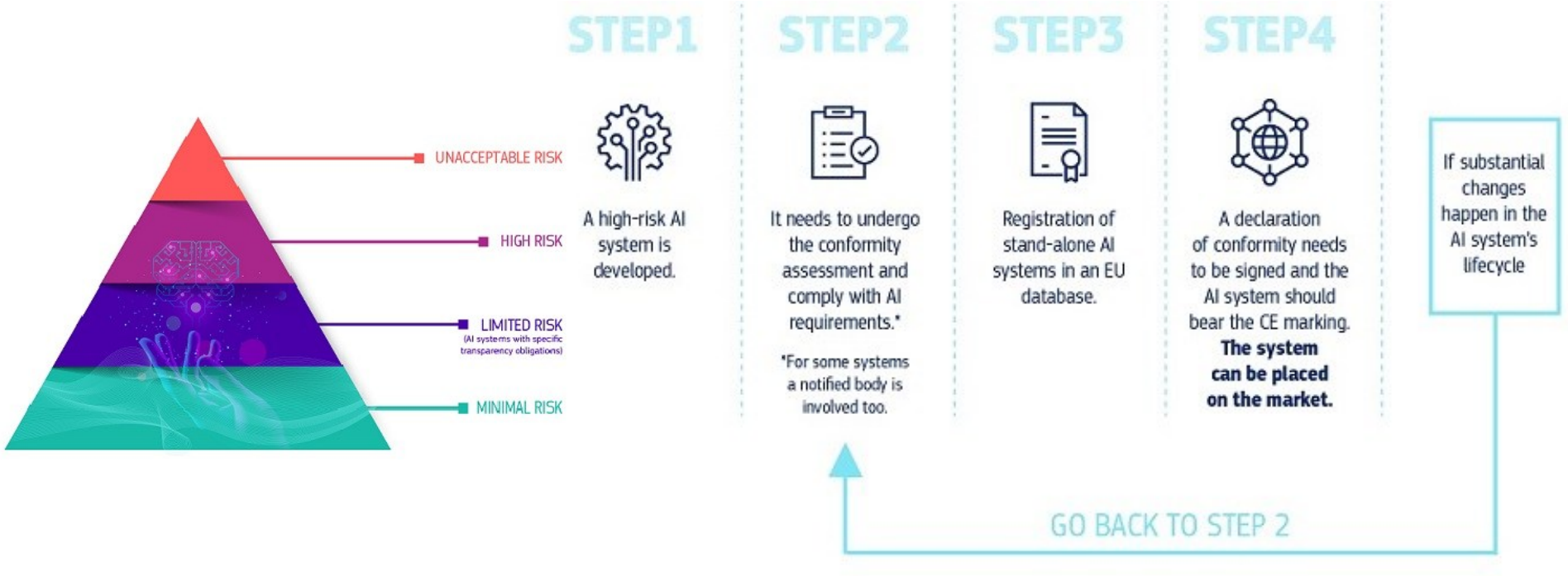
Ethics and AI



- Legal/Ethical
  - Hidden / Indirect Prompt Injection
  - Lawyers Citing Fake Cases Made up by AI
  - Deepfake Imagery Using AI
  - …

Ethics and AI

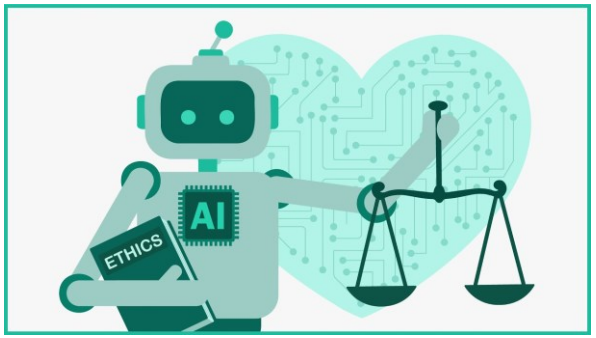
- AI Act
- https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai



Ethics and AI

Recommended usages

Management PhD

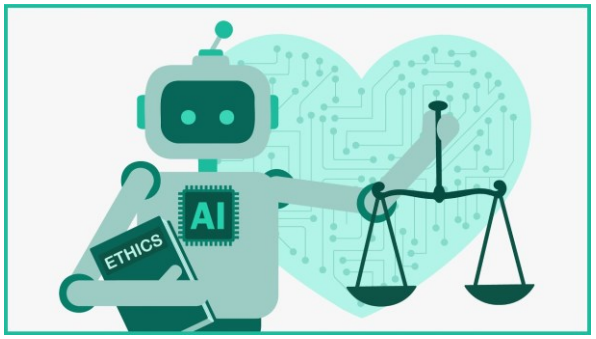


Situation	Recommended Use of GenAI
Academic writing	Improve style, textual coherence, and chapter organization. Suggest chapter structure and reformulate paragraphs.
Literature review	Support in structuring topics — with source validation.
Programming and data analysis	Suggest code, libraries, and tests — always with supervision and validation.
Idea exploration	Brainstorm hypotheses, methodological alternatives, counterarguments. Generate initial research questions.
Personalized tutoring	Explanations of statistical, epistemological, or computational concepts.
Bibliographic summary	Summarize articles, suggest authors or theories (with validation).
Language review	Translation or grammatical correction, especially in English.
Presentation preparation	Generate ideas for slides or communication structures.

Ethics and AI

Recommended usages

Management PhD



Situation	Incorrect Use of GenAI
Submitting generated text without review	Violates authorship and integrity standards.
Creating non-existent results	Generating data that was not collected constitutes scientific fraud.
Citing fabricated sources	High risk of "hallucinations" — undermines credibility.
Avoiding reading articles	Replacing critical reading with AI reduces scientific rigor and the depth of knowledge developed.
Using AI as the dominant voice	Prevents the development of independent scientific thinking.
Omitting the use of AI	Violates the principle of academic transparency.
Assessments	Using GenAI to answer exams or assessments without authorization.
Omission of GenAI use	Failure to acknowledge the use of GenAI in documents where its contribution was significant.

References

- Belcic, I. (2024) *What is RAG (retrieval augmented generation)?* IBM https://www.ibm.com/think/topics/retrieval-augmented-generation
- Costa, C. J. (2024). Neural Networks: A Comprehensive Overview of Their History, Development, and Future in AI. *OAE – Organizational Architect and Engineer Journal*. https://doi.org/10.21428/b3658bca.13fccc0e
- Costa, C. J. (2025). Introduction to Machine Learning. *OAE – Organizational Architect and Engineer Journal*. https://doi.org/10.21428/b3658bca.35c216ce
- Costa, C. J. (2025). Generative AI Models: A Comprehensive Review. *OAE – Organizational Architect and Engineer Journal*. https://doi.org/10.21428/b3658bca.d5d1872f
- Liu, A., Feng, B., Xue, B., Wang, B., Wu, B., Lu, C., ... & Piao, Y. (2024). Deepseek-v3 technical report. *arXiv preprint https://doi.org/10.48550/arXiv.2412.19437*
- Marcondes, F.S., Gala, A., Magalhães, R., Perez de Britto, F., Durães, D., Novais, P. (2025). Using Ollama. In: Natural Language Analytics with Generative Large-Language Models. SpringerBriefs in Computer Science. Springer, Cham. https://doi.org/10.1007/978-3-031-76631-2_3
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Bibliography

Aparicio, J. T., de Sequeira, J. S., & Costa, C. J. (2021). Emotion analysis of portuguese political parties communication over the covid-19 pandemic. In 2021 16th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Aparicio, J. T., Romao, M., & Costa, C. J. (2022). Predicting Bitcoin prices: The effect of interest rate, search on the internet, and energy prices. In 2022 17th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-5). IEEE.

Aparicio, S., Aparicio, J. T., & Costa, C. J. (2019,). Data Science and AI: trends analysis. In 2019 14th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Arriaga, A., & Costa, C. J. (2023, May). Modeling and Predicting Daily COVID-19 (SARS-CoV-2) Mortality in Portugal: The Impact of the Daily Cases, Vaccination, and Daily Temperatures. In Proceedings of International Conference on Information Technology and Applications: ICITA 2022 (pp. 275-285). Singapore: Springer Nature Singapore.

Costa, C. J. (2023). Artificial Intelligence in Management: Enhancing Planning, Organization, Direction, and Control. *OAE – Organizational Architect and Engineer Journal*. https://doi.org/10.21428/b3658bca.b55ee6ae

Costa, C. J. (2024). Neural Networks: A Comprehensive Overview of Their History, Development, and Future in AI. *OAE – Organizational Architect and Engineer Journal*. https://doi.org/10.21428/b3658bca.13fccc0e

Costa, C. J. (2025). Introduction to Machine Learning. *OAE – Organizational Architect and Engineer Journal*. https://doi.org/10.21428/b3658bca.35c216ce

Costa, C. J., & Aparicio, J. T. (2020). POST-DS: A methodology to boost data science. In 2020 15th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Costa, C. J., & Aparicio, M. (2023). Applications of Data Science and Artificial Intelligence. Appl. Sci, 13, 9015.

Costa, C., Aparicio, M., & Aparicio, J. (2021). Sentiment analysis of portuguese political parties communication. In Proceedings of the 39th ACM International Conference on Design of Communication (pp. 63-69).

Costa, C. J., Aparicio, J. T., & Aparicio, M. (2024). Socio-Economic Consequences of Generative AI: A Review of Methodological Approaches. *arXiv preprint arXiv:2411.09313*

Costa, C. J., Aparicio, M., Aparicio, S., & Aparicio, J. T. (2024). The Democratization of Artificial Intelligence: Theoretical Framework. *Applied Sciences*, *14*(18), 8236. https://doi.org/10.3390/app14188236

Custódio, J. P. G., Costa, C. J., & Carvalho, J. P. (2020). Success prediction of leads–A machine learning approach. In 2020 15th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Hajishirzi, R., & Costa, C. J. (2021). Artificial Intelligence as the core technology for the Digital Transformation process. In 2021 16th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Ortiz, F. C. M., & Costa, C. J. (2020, June). RPA in Finance: supporting portfolio management: Applying a software robot in a portfolio optimization problem. In 2020 15th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Piteira, M., Aparicio, M., & Costa, C. J. (2019, June). Ethics of artificial intelligence: Challenges. In 2019 14th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Samadani, S., & Costa, C. J. (2021). Forecasting real estate prices in Portugal: A data science approach. In 2021 16th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Vaswani, A.; Shazeer, N.; Parmar, N. Uszkoreit, J.; Jones, L.; Gomez, A; Kaiser, Ł; Polosukhin, I (2017). "Attention is All you Need" Advances in Neural Information Processing Systems. Curran Associates, Inc. 30.

Veiga, M., & Costa, C. J. (2024). Ethics and Artificial Intelligence Adoption. *arXiv preprint arXiv:2412.00330*.


Q&A