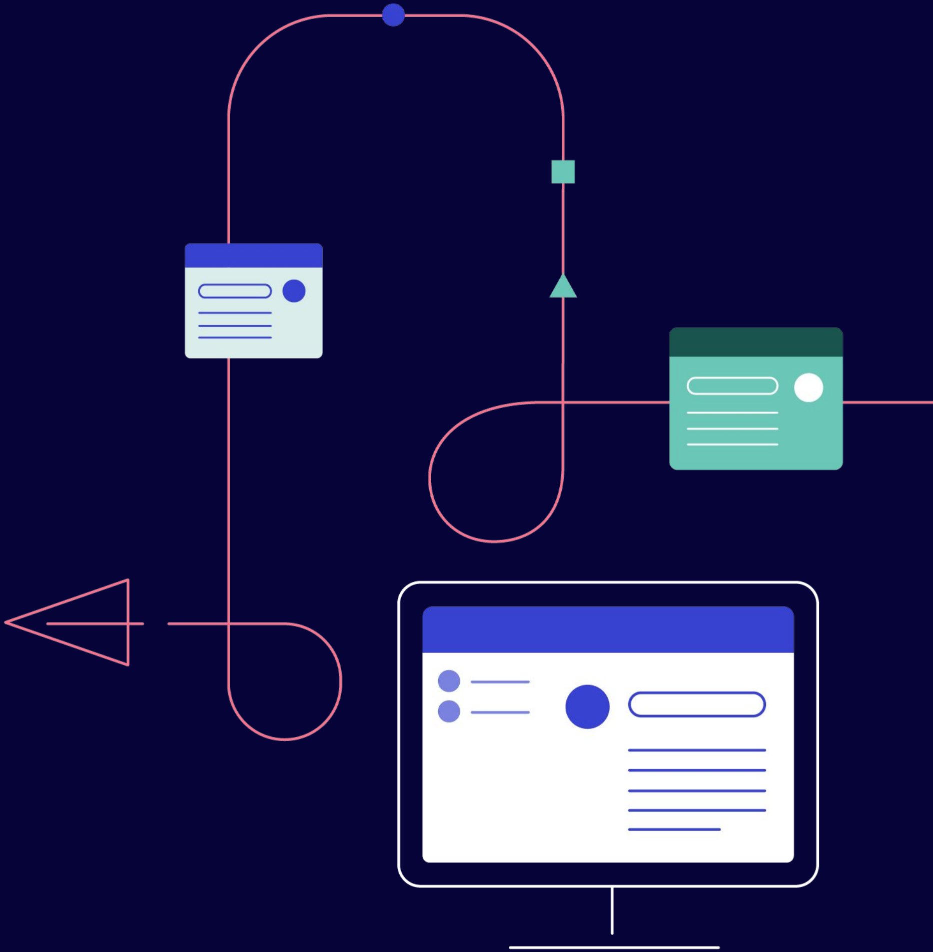
Unbabel

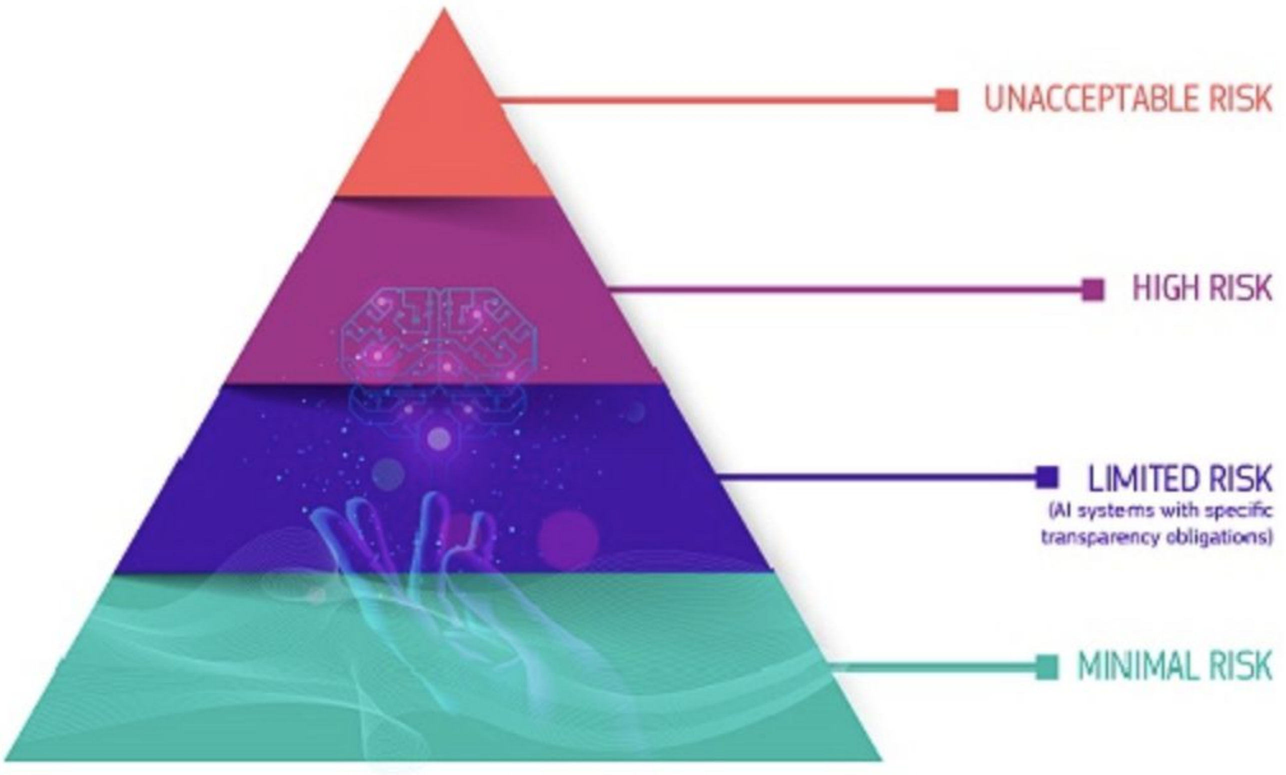
The Rise of AI: Opportunities, Risks, and Regulation in the EU

André Martins

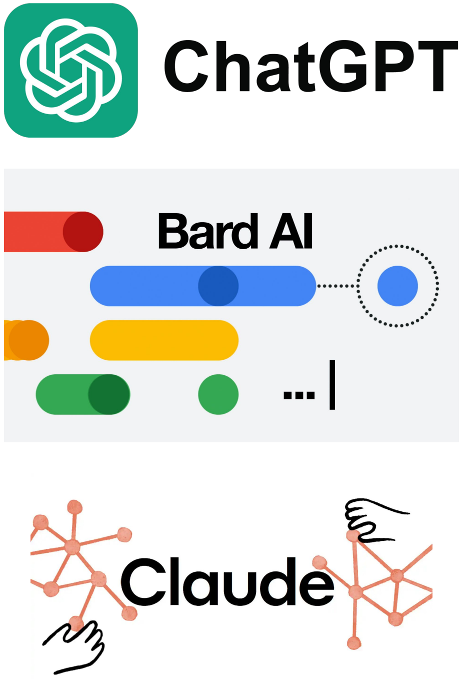
May 24, 2023



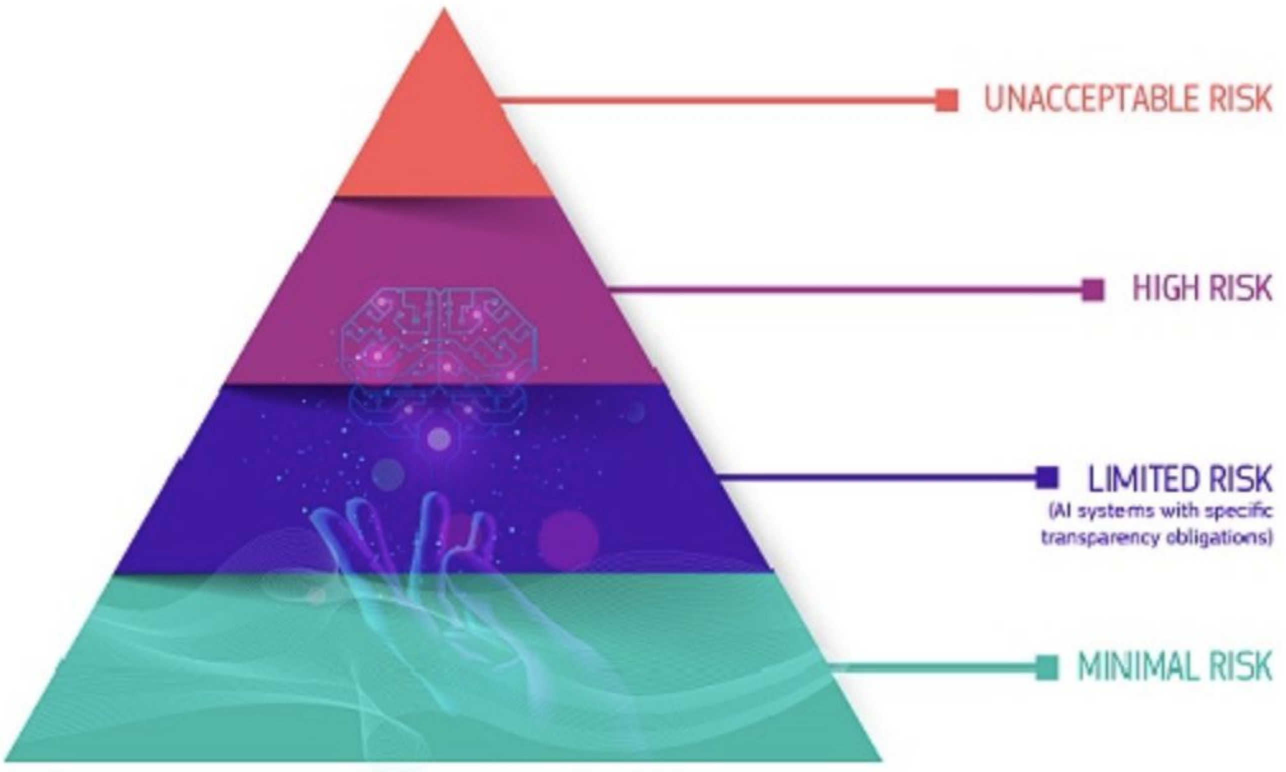
EU AI Act: A Risk-Based Approach



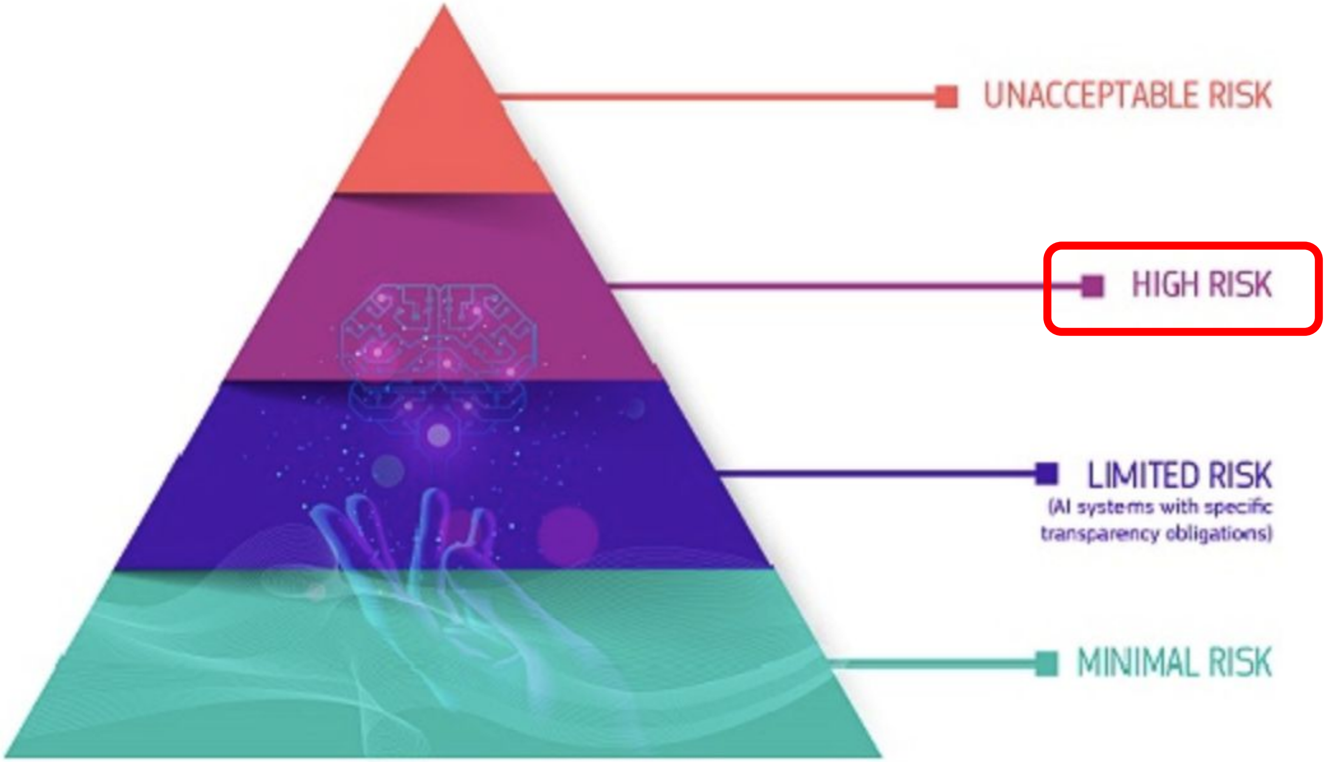
What about general purpose AI (“foundation models”)?



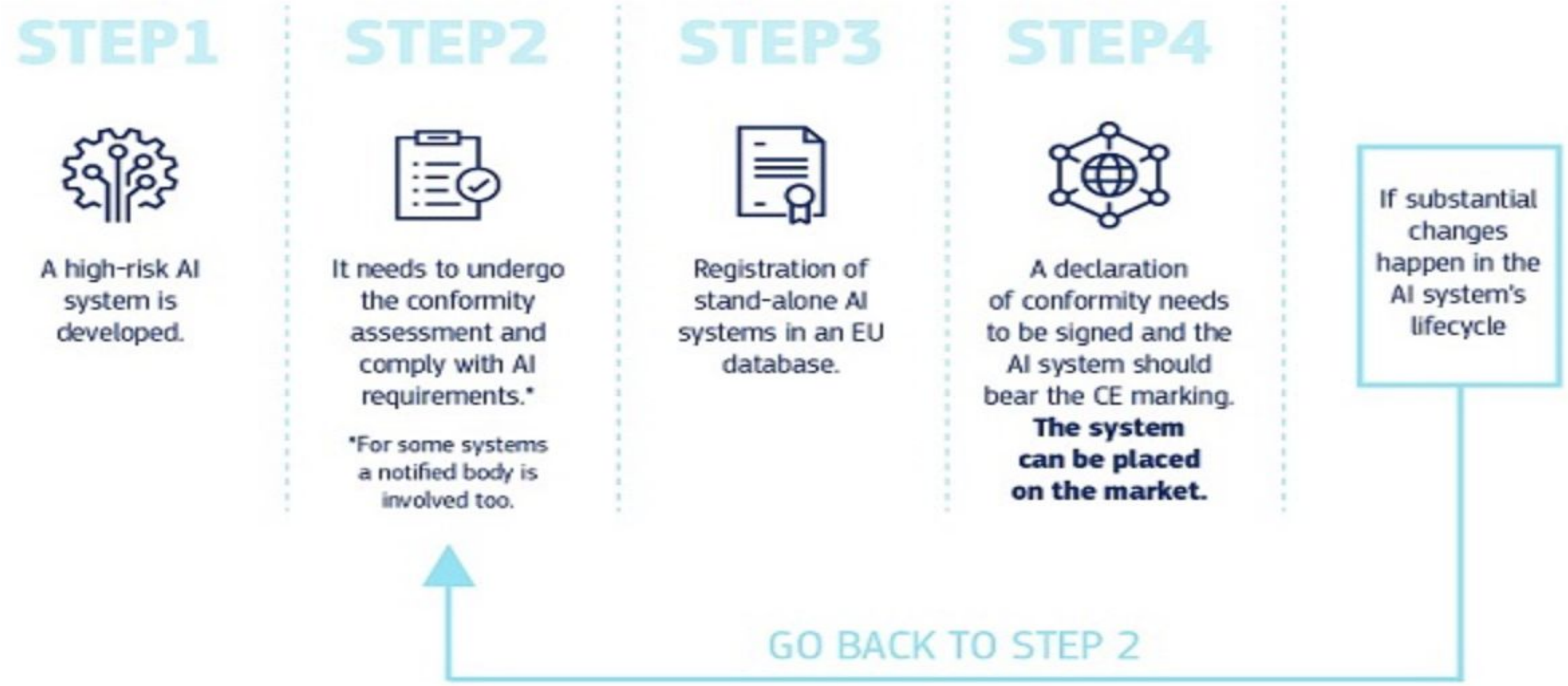
EU AI Act: A Risk-Based Approach



EU AI Act: A Risk-Based Approach



High-Risk AI Systems



STEP1

A high-risk AI system is developed.

STEP2

It needs to undergo the conformity assessment and comply with AI requirements.*

*For some systems a notified body is involved too.

STEP3

Registration of stand-alone AI systems in an EU database.

STEP4

A declaration of conformity needs to be signed and the AI system should bear the CE marking. **The system can be placed on the market.**

If substantial changes happen in the AI system's lifecycle

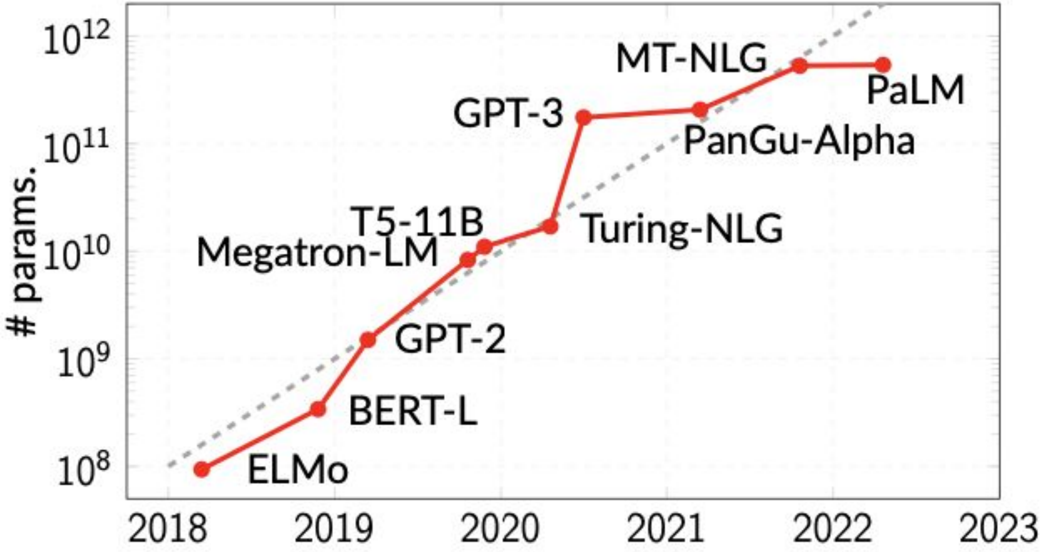
GO BACK TO STEP 2

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

“Foundation” Models

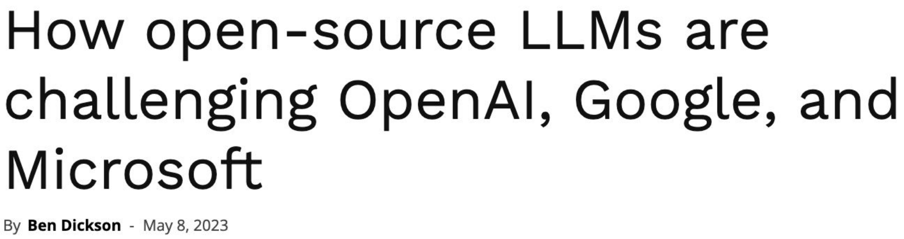
“Foundation” models (LLMs)

- Also known as **large language models** (LLMs) and other names.
- This includes GPT models (OpenAI), PALM (Google), etc.
- Current models have **hundreds of billions of parameters** and are trained on **trillions of words**.



Open-Source LLMs

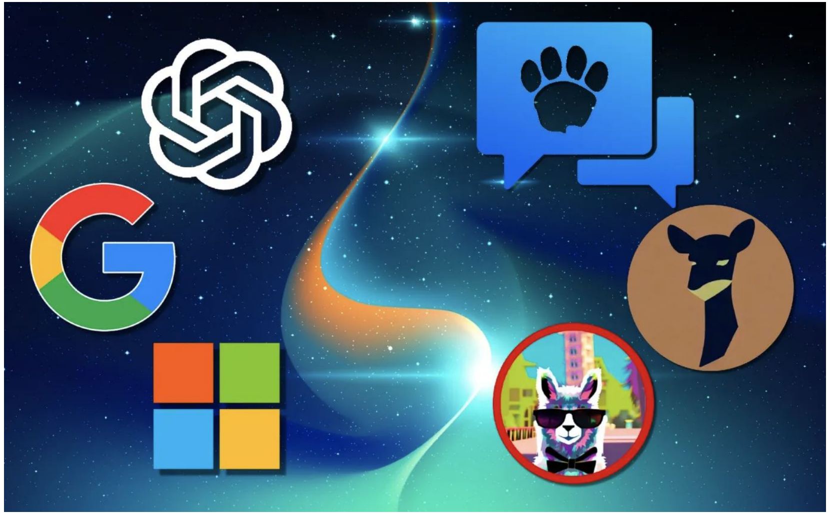
- Traditionally LLMs have been very expensive to train and only big companies could afford it.
- This is being challenged now by several **open source initiatives**.
- Data and model parameters are **made available** for everybody to test and improve upon.
- This contrasts with recent OpenAI and Google models which are closed – **noone knows which data they were trained on, not even the model size**!

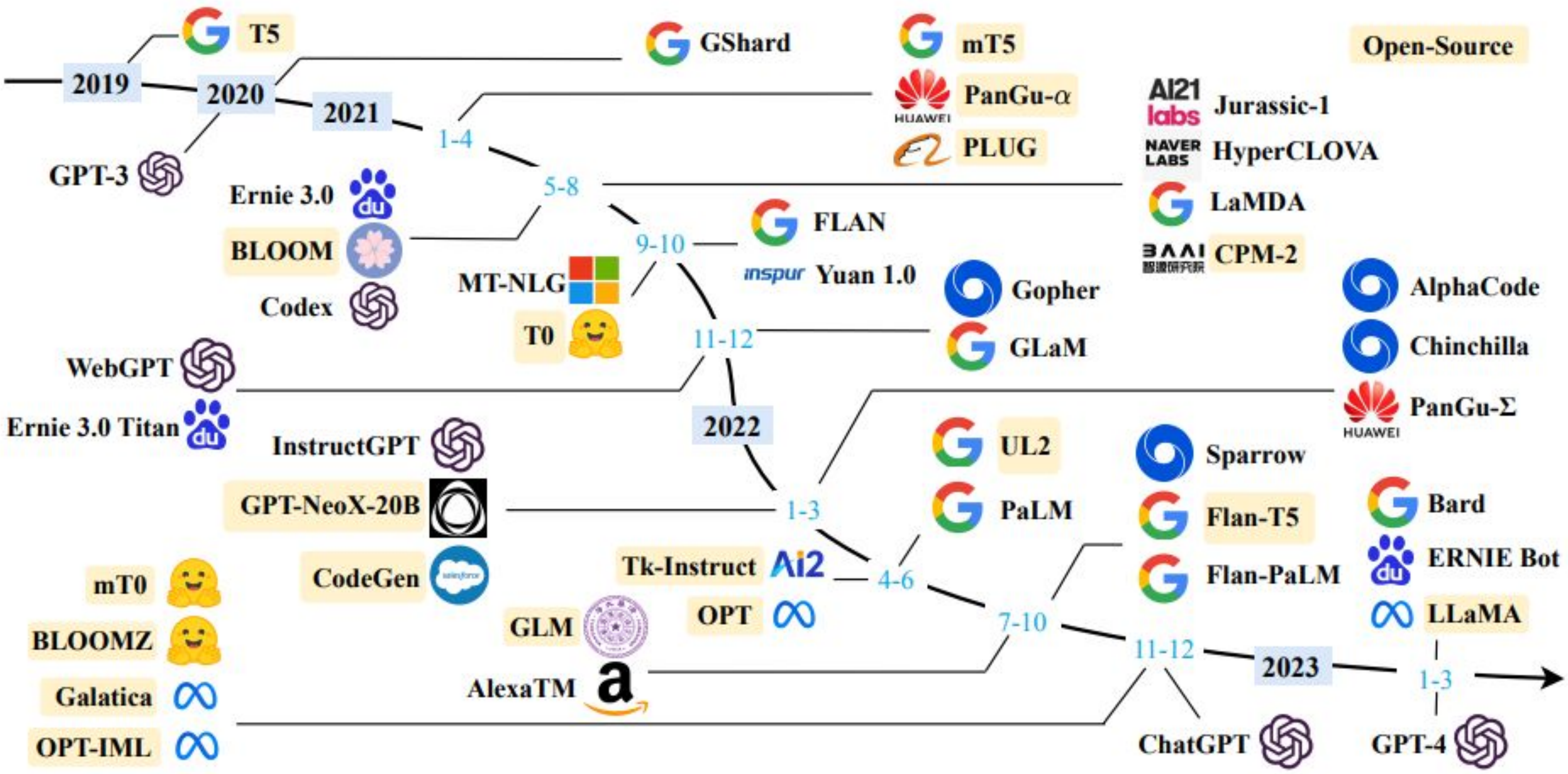


How open-source LLMs are challenging OpenAI, Google, and Microsoft

By **Ben Dickson** - May 8, 2023

https://bdtechtalks.com/2023/05/08/open-source-llms-moats/

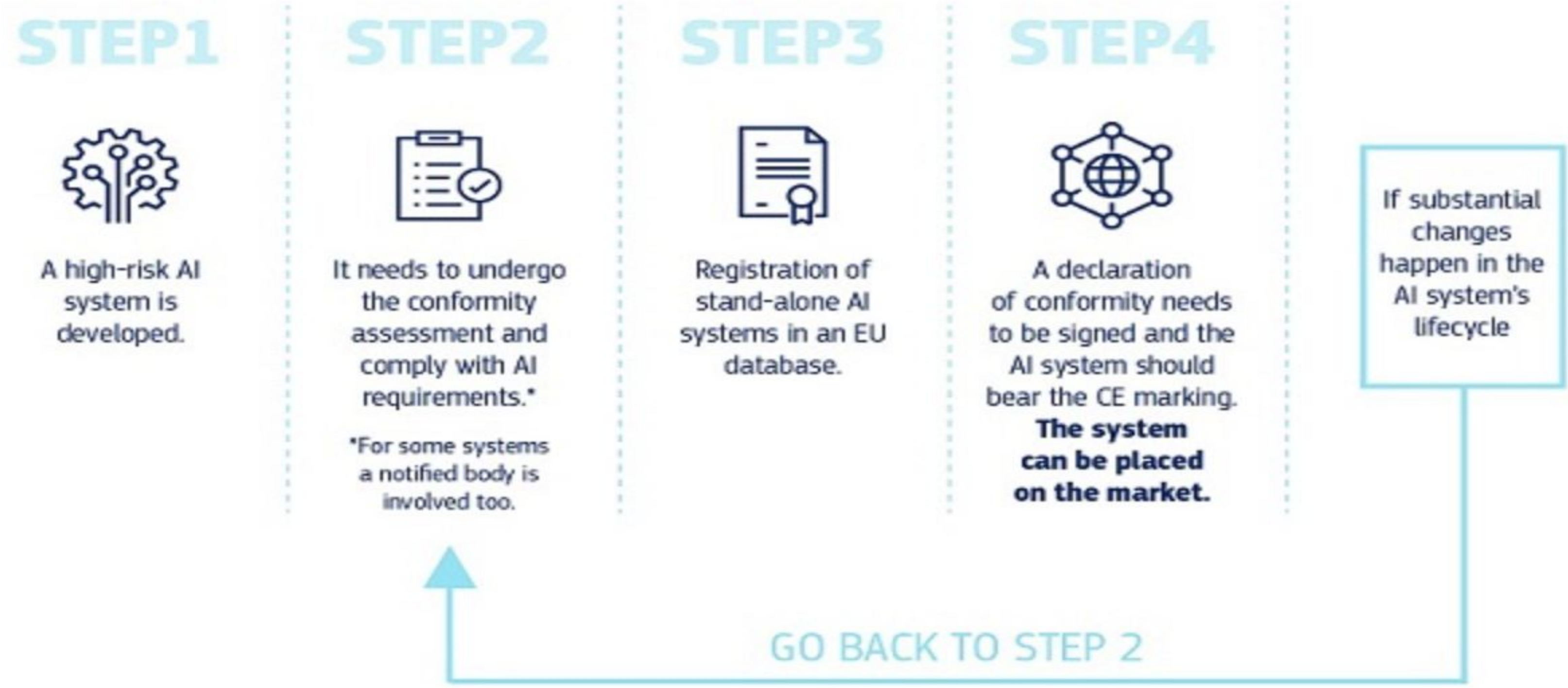




https://github.com/eugeneyan/open-llms .

AI Act and Foundation Models

High-Risk AI Systems



STEP1

A high-risk AI system is developed.

STEP2

It needs to undergo the conformity assessment and comply with AI requirements.*

*For some systems a notified body is involved too.

STEP3

Registration of stand-alone AI systems in an EU database.

STEP4

A declaration of conformity needs to be signed and the AI system should bear the CE marking. **The system can be placed on the market.**

If substantial changes happen in the AI system's lifecycle

GO BACK TO STEP 2

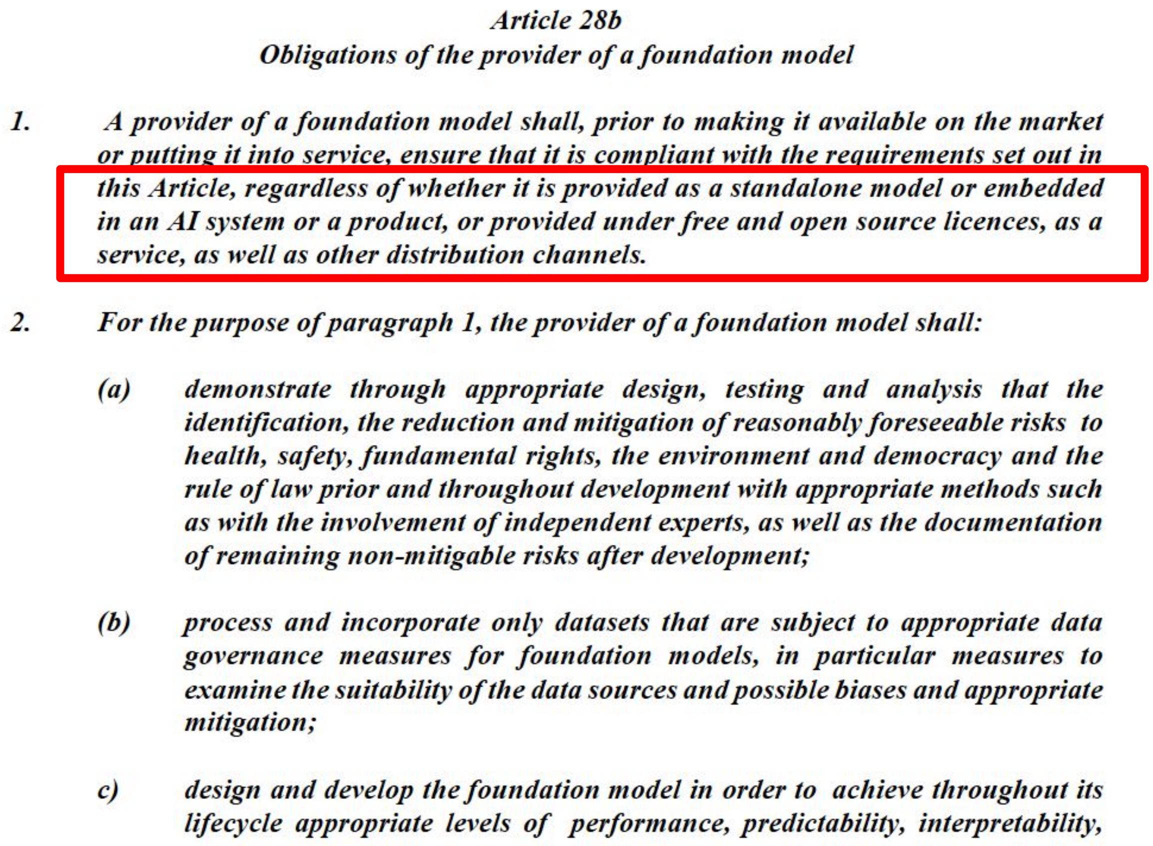
https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

Obligations of LLM providers

- Register their “high-risk” foundation model with the government.
- Register the anticipated functionality of their systems. Systems that exceed this functionality may be subject to recall. (A problem for many open-source projects, whose full functionalities cannot be anticipated.)
- Disclose the data sources used with some detail (this is a good thing), computing resources (including time spent training), performance benchmarks, and red teaming.
- Pass extensive (and expensive) licensing, risk testing, and conformity assessment. Tests are done by third party companies which charge fees.
- Have their systems monitored post-release; recertification required if models show unexpected abilities or after any substantial training.
- Failing to comply leads to massive fines (20M EUR).

No exemptions for open-source LLMs

- The current AI Act draft makes life very difficult for open source initiatives.
  - It includes open source exceptions for traditional ML models, but **not for open source generative systems.**
  - Open-source developers, and hosting services such as GitHub are **liable** for making unlicensed models available.
- It puts a big burden on testing and certification even for models that are not part of a product. This is **bad for research and open science** and it creates **obstacles for innovation** (e.g. EU startups and SMEs).
- At the same time, big tech companies won't have trouble passing all the certifications because they have the resources to do it. So they can use these regulations for **gatekeeping** and preventing any competition.



Article 28b
Obligations of the provider of a foundation model

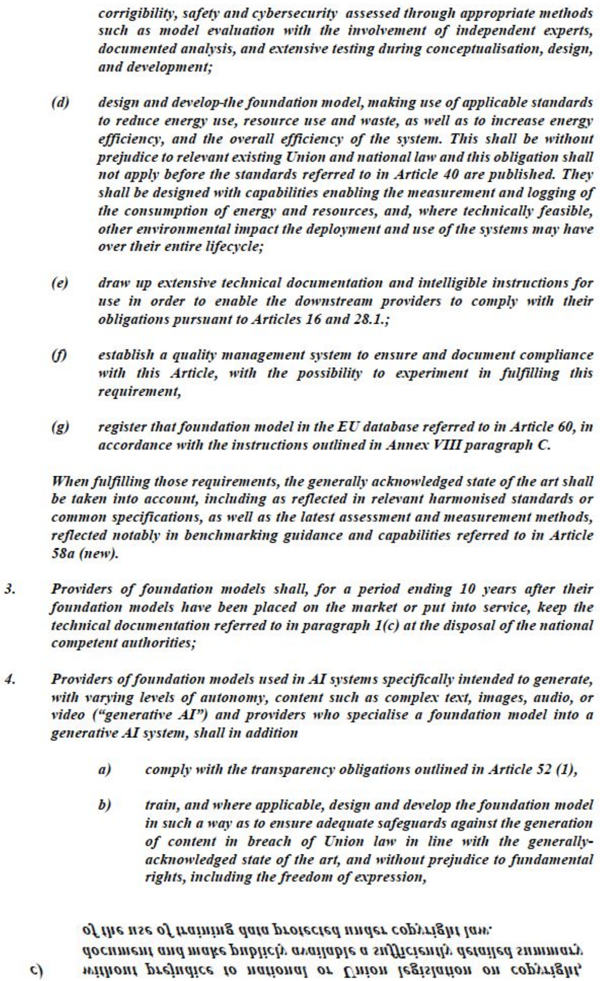
1. *A provider of a foundation model shall, prior to making it available on the market or putting it into service, ensure that it is compliant with the requirements set out in this Article, regardless of whether it is provided as a standalone model or embedded in an AI system or a product, or provided under free and open source licences, as a service, as well as other distribution channels.*

2. *For the purpose of paragraph 1, the provider of a foundation model shall:*

   (a) *demonstrate through appropriate design, testing and analysis that the identification, the reduction and mitigation of reasonably foreseeable risks to health, safety, fundamental rights, the environment and democracy and the rule of law prior and throughout development with appropriate methods such as with the involvement of independent experts, as well as the documentation of remaining non-mitigable risks after development;*

   (b) *process and incorporate only datasets that are subject to appropriate data governance measures for foundation models, in particular measures to examine the suitability of the data sources and possible biases and appropriate mitigation;*

   c) *design and develop the foundation model in order to achieve throughout its lifecycle appropriate levels of performance, predictability, interpretability,*



corrigibility, safety and cybersecurity assessed through appropriate methods such as model evaluation with the involvement of independent experts, documented analysis, and extensive testing during conceptualisation, design, and development;

(d) *design and develop-the foundation model, making use of applicable standards to reduce energy use, resource use and waste, as well as to increase energy efficiency, and the overall efficiency of the system. This shall be without prejudice to relevant existing Union and national law and this obligation shall not apply before the standards referred to in Article 40 are published. They shall be designed with capabilities enabling the measurement and logging of the consumption of energy and resources, and, where technically feasible, other environmental impact the deployment and use of the systems may have over their entire lifecycle;*

(e) *draw up extensive technical documentation and intelligible instructions for use in order to enable the downstream providers to comply with their obligations pursuant to Articles 16 and 28.1.;*

(f) *establish a quality management system to ensure and document compliance with this Article, with the possibility to experiment in fulfilling this requirement,*

(g) *register that foundation model in the EU database referred to in Article 60, in accordance with the instructions outlined in Annex VIII paragraph C.*

When fulfilling those requirements, the generally acknowledged state of the art shall be taken into account, including as reflected in relevant harmonised standards or common specifications, as well as the latest assessment and measurement methods, reflected notably in benchmarking guidance and capabilities referred to in Article 58a (new).

3. *Providers of foundation models shall, for a period ending 10 years after their foundation models have been placed on the market or put into service, keep the technical documentation referred to in paragraph 1(c) at the disposal of the national competent authorities;*

4. *Providers of foundation models used in AI systems specifically intended to generate, with varying levels of autonomy, content such as complex text, images, audio, or video ("generative AI") and providers who specialise a foundation model into a generative AI system, shall in addition*

   a) *comply with the transparency obligations outlined in Article 52 (1),*

   b) *train, and where applicable, design and develop the foundation model in such a way as to ensure adequate safeguards against the generation of content in breach of Union law in line with the generally-acknowledged state of the art, and without prejudice to fundamental rights, including the freedom of expression,*

   c) *without prejudice to national or Union legislation on copyright, document and make publicly available a sufficiently detailed summary of the use of training data protected under copyright law;*

To conclude...

Big AI Tech Companies 💙 AI Regulation?

- We usually think of regulation as something to protect citizens from the hands of big corporations
- But something unprecedented is happening: big tech companies asking for more regulation in the US Congress (!!!)
- We should reflect on this 🤔
- Why would OpenAI want more regulation?

OpenAI's Sam Altman calls for regulation amid fears AI could cause 'significant harm to the world'

https://www.youtube.com/watch?v=9MQbP25u1qE

Final Thoughts

- Regulating AI is very important – I am a strong believer in Responsible AI.
- But I have concerns about some of the added requirements on the AI act about “foundation” models – they’re ill-defined, hard to implement, and might be counterproductive:
    - They discourage AI startups and SMEs to compete in the real game.
    - They compromise open source initiatives and open research.
    - They protect the monopoly of large tech companies who already created and deployed their models *without* regulation.
    - They will create an ecosystem dependent on these large tech companies.
- I believe **openness** is the best way to achieve safety and transparency.

Thank you!