What to expect in new Large Language Models?

Rajagopal Tampi

Introduction

Large Language Models (LLMs) have grown to be the preferred resource for search and other purposes like pattern recognition, predictions, classification, coding, music composition and summarisation. Over the last two years countries are developing their own LLMs for fear of technology denial, currently happening between the US and China. China has risen to the challenge and developed DeepSeek, an efficient and frugal competitor reminding us that necessity is the mother of invention. This has taken the US by surprise sending markets and technical stocks crashing. India is developing its own domestic LLM[1] . Malaysia, Sweden, UAE have built or are building their own LLMs.

AI Risk categorisation

The reason for building country specific LLMs is justified purely by internal needs of a country. It is utopian to imagine that a single LLM can serve the human-centric needs of citizens of all countries in the world. As my book "Applied Human-Centric AI[2,3]" explains the building of LLMs is not a technology exercise alone. The book categorises the risk factors which must be considered for LLMs to be "fit for use" and human-centric in their functionality. The highest risk categories (highest to lowest) are Life-threatening like hospital software, autonomous systems including autonomous weapon systems (AWS), driverless cars are examples. Human Rights risk category comprises upholding human rights, jobs, homes, and individual security. The Human Decision Threshold category is where individual's personal decisions and preferences are automated by AI through agents. The dangers here stem not purely from overriding personal decisions and preferences but from the agents capturing data for creating AI models of the individual and profiling them psychologically and biologically. The fourth category is Global and Societal threats category which includes large scale social unrest, societal divisions, and wars between nations. This category includes political misinformation, election interference in other nations, religious division and more. The fifth category is Sustenance which comprises AI processes and AI infrastructure respecting the planet's sustainability laws. The other AI risk categories cover bias, ethics, monopoly, incomplete data sets and end user criterion.

It is unfortunate that these risk categories are not analysed systematically at the design stage by the leading AI companies. The advent of DeepSeek has triggered rising concerns over national security risks and the opportunity for adversaries to interfere through LLMs developed by them.

The evolution of LLMs

Since January 20th, 2025, countries are facing the bleak prospects of an escalating trade war. Technology sanctions and trade protectionism are looming large. The grim situation is aggravated by the Israel and Ukraine wars. Openness of global trade, technology sharing and work being transferred to the place where the cost is minimal has been the practice followed by developed countries for the last 50 years. This has now been reversed. Protectionism, trade barriers, self-sufficiency (manufacturing within the country) and innovative technology sanctions have become the new geo-political tools in a changing world. The proliferation of these new geo-political, security and economic realities will result in the top economies of the world building their own LLMs in the short term. Self-sufficiency at a national level or a block of nations like the European Union, geo-political alignments and agreements will dictate the external facing risk management design philosophies of these LLMs. Local Customs, local laws, cultural and value systems will dictate the internal design of these LLMs. National and State laws and policies will be incorporated into future LLMs. National LLMs will support local laws, languages, cultures, and sensitivities. Information openness will depend on the form of national governance with democracies providing to greatest openness to dictatorships providing the least. Information levels from open information to total secrecy adopted by governments of various nations will determine the amount of information access and the quality (truthfulness, level of obfuscation) of information inferenced from LLMs.

Religions will influence the design of LLMS as in the case of Malaysia. This could be done by diktat as is the case in Malaysia. While the implementation of religious plurality in a national LLM is a complex matter, the introduction of uniform civil codes in a nation can reduce this complexity.

LLMs could be seen as tools for implementing a nation's foreign policies dealing with commerce, defence, security, and international cooperation. To effectively control information access and output, it becomes necessary to identify the LLM user at a national level through information such as IP addresses. Levels of information access to citizens of different nations based on geopolitical realities and emerging AI technology sharing restrictions will supplement trade agreements currently in in place. Recently the US Navy has banned the use of DeepSeek AI based on security and ethical concerns.

The focus on architectural, lean software development and software engineering aspects of AI development and training will increase leading to more cost effective and purpose efficient LLMs. Advanced nationality-based access and role-based use-case authorisation will be incorporated in LLMs.

Conclusion

The weaponisation of AI technology and curated information sourced from AI models will form an integral part of warfare both offensive and defensive. Technology which was viewed as a standalone subject is now linked inextricably with geopolitical, national security, trade, and commerce strategies. Geo-politics will now be a factor which dictates design, development, access, and use of LLMs. The banning of certain AI products based on geopolitical alignments could become a regular form of punitive action employed by nations.

The use of others nation's LLMs could soon become fraught with a new category of risks perpetrated by the adversary's developers which could include intentional misinformation, obfuscation, misleading, malicious intent and propaganda.

The silver lining in this cloud is a change for the better taking place in the level of honesty of speakers from technology companies and academia at AI conferences which the author has attended. This indicates growing acceptance of the black box nature of neural networks and the harms that AI can inflict on us. Yet much more is required to secure the future safety of humans.

One important way to counter negatives such as LLM bias, human rights violations and geopolitical alignments which could lead to wars is by maintaining open uncensored social media contact within and across nations. This presupposes an open and free internet. To enforce an open and free global communication platform for the benefit of the people, the control of the internet must be under UN forces, not with any nation especially not with any of the superpowers.

The younger generations should take the lead in campaigning for AI human-centricity since they are the ones who will endure most of the consequences.

References:

1. https://indianexpress.com/article/business/india-world-class-ai-model-ashwini-vaishnaw-9807337/lite/
2. https://www.amazon.com/dp/B0CYXHYPHC
3. https://www.amazon.in/dp/B0CW1D7B88