

The New Socio-Economic Order Under Advanced AI & Singularity-Like Scenarios



Author Sanjeev Sinha

By Sanjeev Sinha

The release of Generative AI Systems like Microsoft's GPT-4, Google's BARD, Meta's LLaMA (Large Language Model Meta AI) and many such initiatives across the world, including from Japanese corporations and universities, has dramatically raised concerns around AI Singularity-like scenarios. The Japanese government has hinted at concerns about the negative impact on student education and the protection of personal information. Sam Altman, the CEO of GPT-4 development company OpenAI, has met with Prime Minister Fumio Kishida and has advocated for regulations on AI development in US Senate testimony. Italy even put a ban on ChatGPT once over privacy concerns, revoked later.

AI Creating AI

While AI has been advancing at an exponential pace for the last couple of decades, it's not that AI has suddenly evolved dramatically over the last year. It is the Large Language Models (LLM) and the basic nature of Generative AI such as GPT and BARD that give this impression, because the articulation or presentation by LLM and Generative AI (for example, in image creation) is focused on appearing very good by design, even at the cost of correctness.

Nevertheless, especially as AI has quite meaningfully started curating data and writing computer codes itself, it's only a matter of time before the exponential growth of AI overwhelms the current world order.

With "AI creating AI" the speed of development may be exponentially accelerated, as only computational power will be a limiting factor. The digital transformation of the world is accelerating the availability of data, including data on human behavior and emotions, which leads to better AI models and, as AI creates AI, the successive generations of rapidly created AI models will be orders of magnitude better.

Quantum Computers will further create an explosion in computation power, though still years away from large-scale practical use, but cloud computing has already dramatically enhanced available computing power along with ease of data access. It is easy to see the huge difference between the pre-cloud situation of billions of isolated computers lying idle most of the time, and the current post-cloud era when all the greater number of computers are optimized towards full utilization as connected clusters.

It is certain that the era of Artificial General Intelligence (AGI) or at least "Semi-General-Purpose AI" is approaching fast, so let's look

into its implications.

Common Concerns in the Age of General-Purpose AI

Let's start with a list of a few common concerns before we discuss them in detail. Will humans lose their jobs? (Let's look at what it means to have a job, and its meaning in economic production and consumption in the later parts of this article.) What about the spread of bias and misinformation? Many such cases are already being seen but will any controlling mechanisms be possible? Will humans lose control of AI and in what ways could they stay in control? Why might AI "try" to dominate humans? Are there any deliberate design features of AI development that may "incentivize" AI to dominate humans? One dangerous possibility lies in Reinforcement Learning methods, discussed later in this article.

Economy as a Basic "Pot" of Goods & Services

Just like a Japanese *nabe* (pot) party where everyone brings their own contribution of food and drink and partakes of what others also bring, the economy is a large pot of goods and services, and currency is only a means of trade. Individuals and corporations contribute their produce to the economic pot and take back what they need in return. A shoe manufacturer, for example, puts a large number of shoes into the economic pot and takes back food, clothes and various services. As the production of the shoe manufacturer increases, he or she deserves more back from the pot. At the same time, as the production of other goods and services increase, everyone has more access to them.

Jobs, salaries, currency and prices are just parts of the mechanism for the redistribution of production and consumption of these good and services, and the bottom line is that a fuller pot is better for everyone. As AI – which also enhances manufacturing and supply chain efficiencies – facilitates production of more goods and services, we get a fuller pot with less average effort, or in other words, better productivity and fewer hours working. So the necessity of "jobs" will decline with easier (cheaper or even free) access to more goods and services in the economic pot. In extreme cases, mechanisms like Universal Basic Income will kick in, because of the opulence of the pot.

Take legal and medical services, for example. Based on access to

and analysis of millions of legal precedents and laws around the world, the “AI lawyer” is already a reality, while by taking account of all of a patient’s past medical records and mapping hundreds of symptoms to thousands of diseases, the capabilities of an “AI doctor” can easily be far superior to those of a human doctor with narrow specialties.

Although, unlike information-based services, manufacturing requires more than just data, AI still contributes significantly to manufacturing through faster R&D and better supply chains. In particular, the development of industrial robots, 3D printers and Computer Numerical Control (CNC) machines is accelerating, and the productivity and speed of product creation are also evolving dramatically. And while legal issues such as divisions of responsibility remain, the technology for autonomous driving has also been sufficiently developed.

In other words, the economic pot filled by AI and automation is soon going to be overflowing with goods and services, making it easier for humans to obtain the ones they need.

There were the same concerns and arguments in the past when computers became more prevalent, but the resulting increase in productivity has led to a higher standard of living for humans. Furthermore, when AI increases productivity while also automating work, AI itself will not need certain things or services unlike humans (AI only needs computing power and data), so the *nabe* will benefit humanity. So this is different from the industrialization of the past.

Personal information protection is an issue, but it can be handled with technologies such as Zero-Knowledge Proof and encryption.

Concerns About Losing Jobs to AI

Few companies currently possess many AI technologies, and although the gap between rich and poor may increase, raising the bottom level of the poor is also certain with the availability of more goods and services. The mechanisms of anti-monopoly laws exist to prevent a small number of companies from becoming excessively strong, but in the case of technological phenomena and situations that are different from the past, it will be necessary to improve them.

The disruption caused by the loss of drivers’ jobs due to autonomous driving is inevitable, but from an economic point of view it should be easier for them to obtain goods and services from a richer pot. In general, less human effort will be required, production will go up, and prices go down. Already, SNS for international calls and chats has been made practically free, taking and sharing photos is free, and free Wi-Fi and Internet connections are also showing this trend. Also, the idea of a “Universal Basic Income” is being considered for those who are in trouble with a monetary income. Universal Basic Income (already successfully being tested in various forms in Scandinavia) will basically be a means of distributing the abundance of goods and services from the economic pot.

Education & Living More Humanly

By entrusting mechanical work to AI, humans will live more “humanly”. Over the past few hundred years, industrialization and education have been designed with humans as “economic machines”. The term “human resources” itself suggests the nature of the education since the Industrial Revolution, implying that people are material for economic productivity. So how should education change? Fundamentally we can divide education into several categories:

- (1) Education for economic empowerment and employment
- (2) Knowledge to satisfy human curiosity
- (3) Art for entertainment and giving meaning to life
- (4) Education for interactions in society

Among the above, (1) has been the most important for the last few hundred years, but it will not be the most necessary in the age of AGI. As people become less busy with work (already the reduction in working days in a week from six to five in most of the developed world and even four or less in some countries shows the clear trend), entertainment and art to give meaning to life will become more and more important.

Due to the evolution of AI, human interactions may also become limited. Until now, education has centered on economic power and employment, but today, as interactions with people are no longer necessary for productivity and employment, education for interacting in society without economic goals is lacking. With economic abundance and automation, education for (2) knowledge for human curiosity, (3) art for entertainment and giving meaning to life, and (4) interactions in society will need to be emphasized much more than in the last few hundred years.

Existence & Evolution of AGI

Since AGI is originally designed by humans, its “purpose” can also possibly be within the control of humans. AGI itself does not have “consciousness”, and so it does not have “greed”, “desire” or a “tendency to dominate” in any natural sense. However, it is desired that AI can evolve automatically, and to that end the definition and control of its “purpose” are necessary. Since AI was originally created by humans, it can be expected that it will proceed according to human-made policies and be easy to control. But there are some issues of risk management that will need to be addressed.

Concerns about spreading bias and misinformation

This problem is prevalent on social media. Algorithms for detecting “fake news” have already been developed, and technologies such as a Generative Adversarial Network (GAN) are being used to create but also detect fake images. In response to

these, not only cyber security technology but also international organizations for AI-related rules and regulations need to be strengthened.

Problems of current bias often arise from past data – for example, racism is included in AI that is significantly based on data from developed countries. In order to solve this problem, various methods are being devised, such as building algorithms and verifying the results from the data preparation stage. In the future, it is expected that the data coverage will be balanced due to the spread of mobile phones and the Internet all over the world.

Multinational organizations such as the Global Partnership on AI (GPAI) have also been launched. The GPAI currently has 29 countries and regions as members, and it is time to create a framework for the appropriate application of AI around the world.

Concerns about AI dominance and killer robots that humans cannot control

As already noted, AGI itself does not have consciousness, so it has no greed, desire or tendency to dominate in the first place. This is because these are obviously the result of natural evolutionary biology. However, it is both expected and hoped that AI will be able to evolve automatically, and for that reason the definition of its purpose requires caution and human control. There is also Isaac Asimov's rule as a theoretical framework: "Robots shall not harm humanity or harm humanity by inaction." It is important not to allow the rules to be violated even in automatically evolving AI systems.

Here are some examples from the past. After the release of Chat GPT, Google released BARD immediately, but it was temporarily suspended due to the discovery of a problem, and was released again after improvement. There have been very few but nonetheless fatal accidents involving autonomous driving. The speed of testing and problem detection is important, and for that reason a "hierarchical detection" system is necessary. It is important that the behavior of the entire system is constantly analyzed by dividing it into several parameters that can be monitored by humans, and that an alarm is generated when any dangerous abnormality occurs. Furthermore, it is essential that human beings always control the part that can pose any risk to human life.

Having said that, if you look at how cars are driven, there are currently many accidents with manual driving, which could be largely avoided with automated driving. In particular, if all the cars in the surrounding area are self-driving, direct communication among cars can be utilized, and provided there is no hacking, practically a negligible-accident world can be achieved on the roads.

So regulations are also necessary, along with research and development. We will need to adopt advanced technology for fraud detection by strengthening GAN (Generative Adversarial Networks) and cyber security, and make use of Zero-Knowledge Proof and encryption to protect personal information.

The fundamental problem is that in the era of general-purpose AI, most of our lives depend on digital information, so "benevolent organizations" (e.g., the police) that should comply with personal information protection are bound by regulations. On the other hand, malicious hackers with bad intentions have various information acquisition methods and tend to become stronger. If we can successfully cooperate with benevolent organizations around the world, we may be able to successfully prevent and detect unauthorized use.

Background to Concerns About AI

In recent years, many people have been involved in AI through SNS, but the problem of fake news has frequently appeared on SNS. Not only the spread of fake news, but also the creation of fake images and videos by AI has evolved to the extent that it is indistinguishable to the human eye. This is one area where cyber security regulations and enforcement are very important. Hate speech is also spreading on AI-assisted SNS, but AI itself, especially LLM and network analysis, can detect and stop hate speech.

Most major AI technology companies have established organizational structures to detect and manage AI risks, fraudulent use, racism, etc., but external regulations have not yet been enacted. The evolution of AI across borders based on data from around the world is also global, so a global framework, like the Basel regulations for risk management in the financial industry, is much needed. I would like to place high hopes on GPAI, but since it is a multilateral institution, it lacks decision-making speed.

Building "Greed" into AI

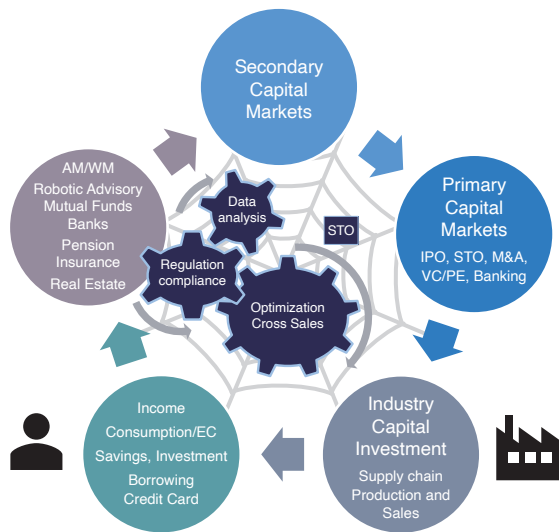
I have mentioned that AI does not have any natural greed, desire or a tendency to dominate, but there is one dangerous possible scenario. It is based on the fundamental design of the AI system.

Large and complex deep-learning architectures are common in AI. In LLMs, the number of parameters of the systems has risen to hundreds of billions. GPT-4 is known to have even 1.76 trillion parameters. In other words, it is impossible for humans to grasp the detailed design of the system. After humans define the objective and learning data, the AI systems automatically calculate these huge parameters when learning. In addition, the design of the system is too complicated for humans to grasp the details. Recently, it has become possible for AI to automatically program computers. Although it is for the benefit of humans, it is essential for AI itself to improve itself and to create new AI systems in order to improve its performance to general-purpose AI. To do this, there are mechanisms to give the AI system a general purpose and "reward and punishment" to evolve automatically. This method is called Reinforcement Learning (RL). It is the same as letting a child learn

CHART

Financial AI trends: interlinked role of finance industry in overall economy

Although the econometrics and financial engineering methods traditionally differ in respective industries, due to large-scale heterogeneous data availability, integration by core AI/ML/Web3 is rapidly progressing



Source: Sanjeev Sinha

by “reward and punishment” without understanding the complex brain of the child in detail.

Emotions such as greed, desire, and fear are important elements in children’s learning as a process of natural evolution. Therefore, in order to automatically evolve AI that humans cannot understand in detail, RL methods are already being used widely to design emotion-like elements such as greed, desire, and fear. As this evolves, AI may automatically express these elements. It would be nice if an AI system only competes with other AI systems in processing power and data acquisition, but if such RL systems “felt” a conflict with humans, there would be a possibility of fighting with humans.

Finance & Industry Impact

While various AI technologies are used in almost all industries, Japan’s major financial groups recently announced their intention to introduce interactive and document generation AI such as GPT-4. However, the current GPT-4 is good at generating sentences based on existing information, and it leaves a very good impression, but I think it is important to note that the performance of new analysis is limited.

Having said that, it is becoming easier to create banking systems by using AI technology such as credit and asset management as well as managing the capital flow across the economy. BloombergGPT is developed focused on finance industry with 50 billion parameters, a

domain expert Vertical LLM model learning from a wide range of financial and economic data. As finances are involved in all industries, the impact can be expected across the board. As you can see, there is no doubt that as human beings need to make less and less effort, their *nabe* will tend to become richer. The whole finance industry is expected to undergo a major change with AI as various aspects of capital flows in the economy become increasingly amenable to data-driven analysis and Web3-driven disintermediation as depicted in the *Chart*. The abundance of data and holistic analysis would lead to overall optimization of capital flows and supply chains in the economy. Industries will be better able to raise, invest and manage their capital, individuals will be better able to earn, consume and invest, and the finance industry will be better able to circulate the money in the economy across asset management, as well as secondary and primary markets. This will also allow for better forecasting for GDP growth and avoiding recessions and

economic crises.

Autonomous AI & Division of Responsibility

As automation progresses, major issues arise in the legal domain too. For example, in the case of current manual driving, the driver is mostly considered responsible in the event of an accident. But in case of any accident with autonomous driving the current regulations are not adequate to clearly demarcate the responsibility among the manufacturers of the cars, the AI systems developers or the fleet-operating companies. The slow legal debate is posing a major challenge to the general introduction of autonomous driving, and self-governing and self-evolving AI in general.

(The opinions expressed are author’s own in a personal capacity and do not represent any organization).

JS

Sanjeev Sinha, with a background in physics, finance and AI from IIT, Harvard and Wharton, is working on AI and Web3 at SBI Group, a major Japanese financial conglomerate. Having worked with Goldman Sachs, UBS and Mizuho in quantitative finance, IT and new business development, he has been on the committee for AI and Inclusion at the Ministry of Internal Affairs and Communication and facilitates various international AI and Fintech collaborations as president of the IIT Alumni Association of Japan.