TRIGLAV MEETING: HUMANITIES & SCIENCES
18 January 2025

Challenging society: Some Thoughts on Artificial Intelligence

Dennis W. MacDonald
Professor Emeritus
Saint Anselm College

Technological development has always been subject to a mixed reception of fear - fear of its impact on a people's way of life or the inability of comprehend the new or threats to livelihood - and promise - promise of an improvement in daily life, expansion capacities to act, a lessening of life's burdens. However, several developments over the last century (or two) have raised serious questions. First, technological developments in transportation and communications ushered in a society that was not only local or national, but also global. In the pre-global era, the impact of technological developments might affect employment opportunities in the local factory or the farm, but now such impact is felt globally, and with little if any opportunity to exert control over it. A second significant factor was the transformation of technology in the period after World War II. A qualitatively different technology, one where human beings create substances not found in nature - various chemical products, artificial fertilizers, plastics, etc. which are incapable of being broken down in nature, produced in enormous quantities with no (or certainly not sufficient) consideration of the impact on the ecosystems into which they must fit.
Technology thus now has a massive impact on a global scale and has brought us to the brink of a climate disaster.

AI or Artificial Intelligence presents an even greater challenge. Its presents capacities significantly beyond the power of social media. Is it - or can it be - a force for societal improvement? Or is it likely to lead to serious problems? Can it be harnessed in a positive way?

I. What is AI?

> Artificial intelligence (AI) is technology that enables computers and machines to simulate human learning, comprehension, problem solving, decision making, creativity and autonomy.
>
> IBM https://www.ibm.com/think/topics/artificial-intelligence

II. Power of AI

There is no question that AI represents tremendous expansion of human capacities in a wide variety of fields. Even though we are still in the early stages of AI development, it already has tremendous accomplishments in many areas.[1] A few examples:

A. Health and Medicine
1. *AlphaFold* by DeepMind has accurately predicted 3D protein structures, solving a 50year-old biological challenge and accelerating drug discovery.
2. AI systems like those from Google Health have outperformed human radiologists in detecting diseases like breast cancer, lung cancer, and diabetic retinopathy in medical scans.
3. AI algorithms have identified potential drug candidates, significantly speeding up the process, such as Exscientia's AI-designed drug entering clinical trials.
4. Robotics and AI, such as the da Vinci Surgical System, have been able to enhance precision in complex surgeries and reduce recovery times in complex surgeries.

B. Language
1. Models like OpenAI's *GPT* series (including ChatGPT) excel at natural language understanding and generation, enabling applications in customer service, creative writing, and education.

2. AI-powered systems like Google Translate provide near-instant translations for multiple languages, facilitating global communication.

C. Autonomous Systems
1. Companies like Tesla, Waymo, and Cruise have developed autonomous vehicles capable of navigating complex urban environments.
2. AI-powered drones and robots are transforming logistics, with companies like Amazon and Zipline leading the charge.

D. Humanitarian
1. AI systems analyze satellite imagery to identify areas affected by natural disasters, assisting in relief efforts.
2. AI aids in tracking endangered species, combating poaching, and managing ecosystems through predictive analytics.

E. Climate
1. AI is used to refine climate predictions, map carbon emissions, and optimize renewable energy usage, aiding global sustainability efforts.

F. Economic
1. AI algorithms identify suspicious transactions and reduce financial fraud in real-time.
2. AI-driven trading systems optimize investment strategies by analyzing vast amounts of market data.

G. Robotics
1. AI-driven robots, such as Boston Dynamics' *Atlas*, exhibit human-like agility and balance, enabling advanced industrial and rescue operations.

H. Education
1. AI platforms like Khan Academy and Duolingo adapt to individual learning styles, making education more effective and accessible.

I. Language Preservation
1. AI helps document and preserve endangered languages through translation and voice synthesis.

J. Art and Creativity - AI tools can create visual images and musical compositions. Whether it creates art is a different question.
1. Tools like *DALL·E* and *Stable Diffusion* create stunning visual art from textual descriptions, democratizing art creation.
2. AI tools like OpenAI's *MuseNet* and *AIVA* compose original music across genres and styles.

III. The Potential Perils of AI

A. Frightening possibility of AI systems taking control over human beings and removing decision-making and judgement from our hands. An early example of this fear is illustrated in a famous scene from Stanley Kubrick's "2001: A Space Odyssey" where the computer HAL 2001 takes control. https://www.google.com/search?client=firefox-b-1d&q=Hal+the+computer+from+2001#fpstate=ive&vld=cid:a2da1dda,vid:5lsExRvJTAI,st:0

B. AI can be harnessed for destructive and exploitative purposes. A couple examples appeared in this week's news. One involved a lawsuit against landlords using AI to raise tenants' rents to the maximum possible level. https://www.washingtonpost.com/business/interactive/2025/realpage-lawsuit-rent-map/?itid=hp-top-table-main_p001_f009 The other involved people's susceptibility to AI generated disinformation and propaganda. https://hai.stanford.edu/news/disinformationmachine-how-susceptible-are-we-ai-propaganda

C. Main categories of concern with AI are as follows:
1. Deepfakes. AI-generated content can mimic real people, spreading misinformation or creating fake scandals. AI tools can be used to create images of your face and body and voice and create a video images people in compromising positions, committing crimes, etc. This takes identity theft to a whole new level.

2. Disinformation. AI can be used to generate and amplify fake news or manipulate public opinion, especially during elections. We have already seen examples of that in U.S. elections. One example combining deepfakes and disinformation occurred during the last New Hampshire primary election where Life Corp, a Texas company, used Biden's voice to call as many as 25,000 New Hampshire voters urging them to skip the primary. The potential, of course, is for far greater harm (and probably was in the US 2016 Presidential election).
3. Cybersecurity Threats. AI tools can be used to find vulnerabilities in software and computer controlled systems which these days is most systems - utilities, government services, weapons systems, industry, etc. Exploiting the vulnerabilities could wreck havoc. On the individual level, AI-generated phishing emails or voice scams can exploit individuals or organizations causing tremendous financial losses and identity theft.
4. Invasion of Privacy and Mass surveillance: Governments or organizations could (and do) use AI to monitor individuals, track movements, or analyze behavior without the consent of the individuals and groups involved. The use of facial recognition in public spaces is very susceptible to misuse and abuse and can lead to discrimination and wrongful surveillance. It does not require much imagination to understand how it might be (and likely is) used by secret police and others in authoritarian societies.
5. Autonomous Weapons. Current international conflicts including Russia's war against Ukraine and the horribly destructive fighting in the Middle East have seen increasing use of autonomous drones, robots, or other weapons that can operate with minimal human oversight, capable of magnifying intended death and destruction and increasing risks of unintended harm. Non-state actors can exploit AI technologies for violent purposes, but use by state actors can be equally horrendous.
6. Economic Exploitation. For many decades, businesses have developed ever more sophisticated approaches to manipulating the public with advertising. AI-driven algorithms have the capability of raising the bar significantly by using personal data to target individuals with manipulative ads or offers. Companies can also use AI to micromanage workers, reduce wage and limit opportunities for collective action.
7. Environmental Impact. AI requires tremendous computational power resulting in very high energy use leading to further environmental degradation.
8. Bias/Discrimination. Bias may be designed into algorithms, intentionally or unintentionally, reinforcing societal biases and leading to unfair treatment in hiring, lending, law enforcement, and other areas.
9. Behavioral Manipulation. The use of AI in social media can exploit psychological vulnerabilities to increase online engagement, potentially leading to addiction, brainwashing, etc. The use of AI in various apps or systems can be used to manipulate behavior subtly, often to extract more data or revenue from users.
10. Exploitation of the vulnerable. AI might be used to facilitate exploitation of labor in supply chains, such as through surveillance of workers or forcing competitive behaviors. Predatory AI systems might be used exploit the elderly or financially struggling individuals with scams.
11. Erosion of Trust. The widespread use of AI-generated media content can make it harder to distinguish truth from fiction leading to the erosion of trust in media and information generally. AI can also lead to distrust of other institutions by creating fake narratives. The widespread skepticism regarding vaccines is an example. Not only can such narratives undermine trust in science and other institutions, it can lead to the vilification of and threats to dedicated public servants (e.g. Dr. Anthony Fauci).

IV. Should AI be promoted and encouraged or restricted or prohibited?

Clearly, artificial intelligence has both great promise and great perils for society. It is a powerful; tool unlike any we have seen before. How should society handle it. Should it be highly

regulated? Should it be given free reign to develop whatever it is capable of? Should we put the brakes on it?

It is important to keep in mind that like other technology developments (products of applied science) before it, it as a tool[2]. As such, it expands the capacity of human beings to act. The thing about tools is that they are ethically neutral in themselves.[3] A hammer can be used to smash someone's head in or to build them a house. The issue is the control of such tools control by whom, in whose interest, toward what ends?

Laissez faire approach. Society - especially in the US - has generally allowed technological development to proceed largely uninhibited by regulation. The results has been tremendous progress in many areas furthering human capacity to act through new forms of energy, new drugs, chemicals of various kinds, transportation systems, etc. At the same time, the impact on the environment, human health and other areas have been destructive, and we are seeing the effects in climate change, agricultural land poisoned by forever chemicals, etc.[4]

Will we make - are we making - the same mistake with AI with potentially disastrous consequences?

Control. Even without HAL of "2001" or AI wresting control from humans, there are serious potential problems of control. Various AI run systems may simply get "out of control," or control could be exerted for strictly private ends with little or no consideration of the public good. Several questions need to be addressed in this connection.

1. Who controls? Is it appropriate to leave control of such powerful instruments in of private hands serving particular interests (individual, corporate, or even national)? Given that the consequences of activities using these powerful tools are largely public, it would seem that there must be a public element of control as well? (Similarly, with respect to powerful social media that have a huge public and global impact, is it appropriate for a Musk or Zuckerburg be able to control these however they see fit?)
2. To what end(s) should these instruments be controlled? What are the principles that ought to guide regulation of them?
3. On what ground should the appropriateness of activities be judged? Subjective - based on interests - what's in it for me/us? Objective - what are its implications for the larger whole - the public, community, environment, world?
4. What would be adequate institutional structures to exert societal control - not to stifle progress or creativity, but to avoid negative, even disastrous, consequences, and to ensure that the public good is being served? (In the US, at least, the already weak regulatory structures are under attack by the oligarchs. It may get much worse before it gets better.)
5. Finally, like so much activity in the modern world, the implications of AI are global, not confined within local or national boundaries. Do we need global standards for governing the use of AI and global institutions, organizations, develop and apply these standards?

Final comment: Artificial Intelligence may share many of the elements of human intelligence logic, perception, reasoning, etc. However, it seems to lack an essential element; a central aspect of genuine human intelligence is the capacity to see relations in the world, to see how something fits, whether it makes sense. Sound judgement is judgement in relation to the whole, to the world. We are, of course, not omniscient. But we do have the capacity to explore the broader implications of actions for the world, the ability to look at the larger context and attempt to judge objectively. The broader the view, the broader our knowledge, the more likely to judge soundly, to create things that improve the world rather than destroy it.

[1]. One of these is the capacity to survey huge databases of information very quickly. Several of the list included here were generated with ChatGPT in a matter of seconds. It would have required many hours of research develop such lists.
[2]. I mean tools in the widest sense of everything that enhances human capacities for action: hammers, chemicals, agriculture, transportation systems, aspirin, global relations, industry. [3]. This is not to say that any technological development is ethically neutral. One can hardly consider the atomic bomb to be ethically neutral.
[4]. In *The Closing Circle*, Barry Commoner wrote this regarding the man-made organic chemicals which were the products of the technological transformation following World War II: "[I]t would be prudent, I believe, to regard every man-made organic chemical not found in nature which has a strong action on any one organism as potentially dangerous to other forms of life. Operationally, this view means that all man-made organic compounds that are at all active biologically ought to be treated as we do drugs, or rather as we should treat them-prudently, cautiously. Such caution or prudence is, of course, impossible when billions of pounds of the substance are produced and broadly disseminated into the ecosystem where it can reach and affect numerous organisms not under our observa tion. Yet this is precisely what we have done with detergents, insecticides, and herbicides."