

April 2023

VOLUME XIV, ISSUE IV

# PEGASUS

A NEWSLETTER FOR THE CAUX ROUND TABLE FOR MORAL CAPITALISM  
NETWORK LOOKING AT BUSINESS ABOVE THE CLUTTER AND CONFETTI

Image generated by DeepAI



# *Pegasus*

Introduction by Stephen B. Young

Page 2

Will the Coming Combination of AI and Oligarchies  
Produce a New Feudalism? by Michael Wright

Page 5

Artificial Intelligence and the Good Life:  
Relationships, Rules and Real Wealth by Michael  
Hartoonian

Page 20

# Introduction

This April issue of *Pegasus* brings you tough, but very, very important questions about artificial intelligence (AI) from Michael Wright, a successful business executive in high tech and Michael Hartoonian, an educator shaped by thinking about the deep questions of human purpose, asked by Greek philosophers so long ago.

Both Michael's challenge the easy belief that AI is, simply put, the next stage in the progressive evolution of the human species towards blissful happiness and perfection.

Their essays bring to my mind Matthew Arnold's poem *Dover Beach*, which asks us to pause and question ourselves before charging madly and enthusiastically into the future:

*Sophocles long ago  
Heard it on the Ægean and it brought  
Into his mind the turbid ebb and flow  
Of human misery; we  
Find also in the sound a thought,  
Hearing it by this distant northern sea.*

*The Sea of Faith  
Was once, too, at the full, and round earth's shore  
Lay like the folds of a bright girdle furled.  
But now I only hear  
Its melancholy, long, withdrawing roar,  
Retreating, to the breath  
Of the night-wind, down the vast edges drear  
And naked shingles of the world.*

*Ah, love, let us be true  
To one another! for the world, which seems  
To lie before us like a land of dreams,  
So various, so beautiful, so new,  
Hath really neither joy, nor love, nor light,  
Nor certitude, nor peace, nor help for pain;  
And we are here as on a darkling plain  
Swept with confused alarms of struggle and flight,  
Where ignorant armies clash by night.*

In this bleak vein of thought, Michael Wright's essay on the concentration of power applied to the development and deployment of AI reminds me of Lord Acton's warnings: "Power corrupts and absolute power corrupts absolutely" and "...remember, where you have a concentration of power in a few hands, all too frequently men with the mentality of gangsters get control. History has proven that."

Both Michael's wrote down their words of caution before a notable group of business leaders in AI and others with the Future of Life Institute issued the following statement calling for a pause in the further development of AI technologies:

*AI systems with human-competitive intelligence can pose profound risks to society and humanity, as shown by extensive research and acknowledged by top AI labs. As stated in the widely-endorsed Asilomar AI Principles, Advanced AI could represent a profound change in the history of life on Earth and should be planned for and managed with commensurate care and resources. Unfortunately, this level of planning and management is not happening, even though recent months have seen AI labs locked in an out-of-control race to develop and deploy ever more powerful digital minds that no one – not even their creators – can understand, predict or reliably control.*

*Contemporary AI systems are now becoming human-competitive at general tasks and we must ask ourselves: should we let machines flood our information channels with propaganda and untruth? Should we automate away all the jobs, including the fulfilling ones? Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? Should we risk loss of control of our civilization? Such decisions must not be delegated to unelected tech leaders. Powerful AI systems should be developed only once we are confident that their effects will be positive and their risks will be manageable. This confidence must be well justified and increase with the magnitude of a system's potential effects. OpenAI's recent statement regarding artificial general intelligence states that, "At some point, it may be important to get independent review before starting to train future systems and for the most advanced efforts to agree to limit the rate of growth of compute used for creating new models." We agree. That point is now. Therefore, we call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4. This pause should be public and verifiable and include all key actors. If such a pause cannot be enacted quickly, governments should step in and institute a moratorium.*

*AI labs and independent experts should use this pause to jointly develop and implement a set of shared safety protocols for advanced AI design and development that are rigorously audited and overseen by independent outside experts. These protocols should ensure that systems adhering to them are safe beyond a reasonable doubt. This does not mean a pause on AI development in general, merely a stepping back from the dangerous race to ever-larger unpredictable black-box models with emergent capabilities.*

*AI research and development should be refocused on making today's powerful, state-of-the-art systems more accurate, safe, interpretable, transparent, robust, aligned, trustworthy and loyal.*

*In parallel, AI developers must work with policymakers to dramatically accelerate development of robust AI governance systems. These should at a minimum include: new and capable regulatory authorities dedicated to AI; oversight and tracking of highly capable AI systems and large pools of computational capability; provenance and watermarking systems to help distinguish real from synthetic and to track model leaks; a robust auditing and certification ecosystem; liability for AI-caused harm; robust public funding for*

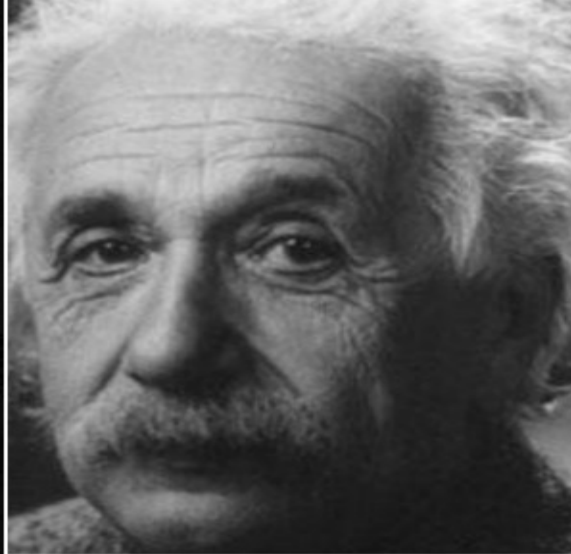
*technical AI safety research; and well-resourced institutions for coping with the dramatic economic and political disruptions (especially to democracy) that AI will cause.*

*Humanity can enjoy a flourishing future with AI. Having succeeded in creating powerful AI systems, we can now enjoy an “AI summer” in which we reap the rewards, engineer these systems for the clear benefit of all and give society a chance to adapt. Society has hit pause on other technologies with potentially catastrophic effects on society. We can do so here. Let’s enjoy a long AI summer, not rush unprepared into a fall.*

The full statement, including notes, references and signatories, can be found [here](#).

Are we not at a point where “an ounce of prevention is worth a pound of cure?”

*Stephen B. Young  
Global Executive Director  
Caux Round Table for Moral Capitalism*



An oligarchy of private capital cannot be effectively checked even by a democratically organized political society because under existing conditions, private capitalists inevitably control, directly or indirectly, the main sources of information.

— Albert Einstein —

AZ QUOTES

*(Note: This article was first published in November 2018 on the Intercepting Horizons, LLC blog. Following is an update and interview using GPT3.)*

## **Will the Coming Combination of AI and Oligarchies Produce a New Feudalism?**

**Michael Wright**

***And the people will ask: “Have we put our future in the hands of the machine?”***

***No matter how magnificent our machines become, they will take no responsibility for what humans do with the data and the choices presented.***

***How is the aggregation of AI and domain specific knowledge coalescing with the concentration of wealth (wealth and knowledge are becoming synonyms)?***

During the years before the Renaissance, we might have run into a similar notion of the concentration of wealth and values. It was called feudalism. Based on the ideals of chivalry, where nobles and knights were to defend the weak and the poor, as well as the faith, they did the opposite. They roamed the land robbing and killing – because they could and besides, God was on their side. Beware, this is a cautionary tale.

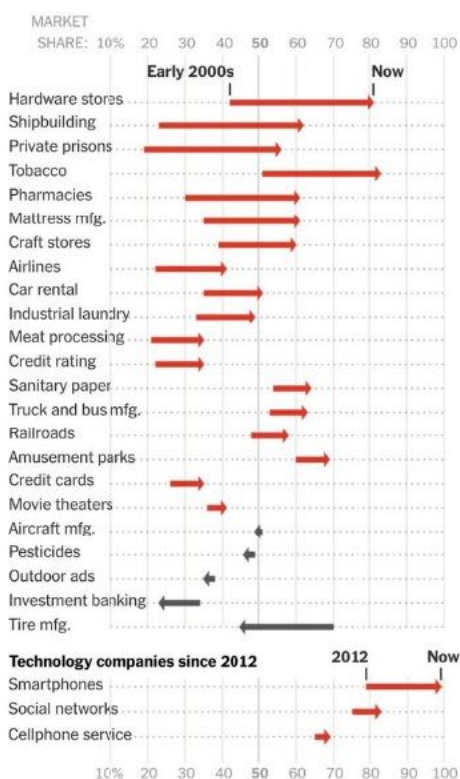
Oligarchy is a form of power structure in which power rests with a small number of people. These people may be distinguished by membership in nobility, a wealth class, a family, an education level, a corporation, a religious affiliation, political group or the military.

Such oligarchic states are often controlled by families who typically pass their influence from one generation to the next, but inheritance is not a necessary condition for the application of this term. The top 10% of American households, as defined by total wealth, owned 84% of all stocks in 2016. Fewer people own and control fewer companies. As I said in a previous

article, “Domain specific knowledge (DSK), at the expert level, concentrates in the hands of fewer people and companies as it becomes more complex.”

## Dominance of Corporate Behemoths

The combined market share of the two largest companies in many industries has grown in recent years, often because of mergers.



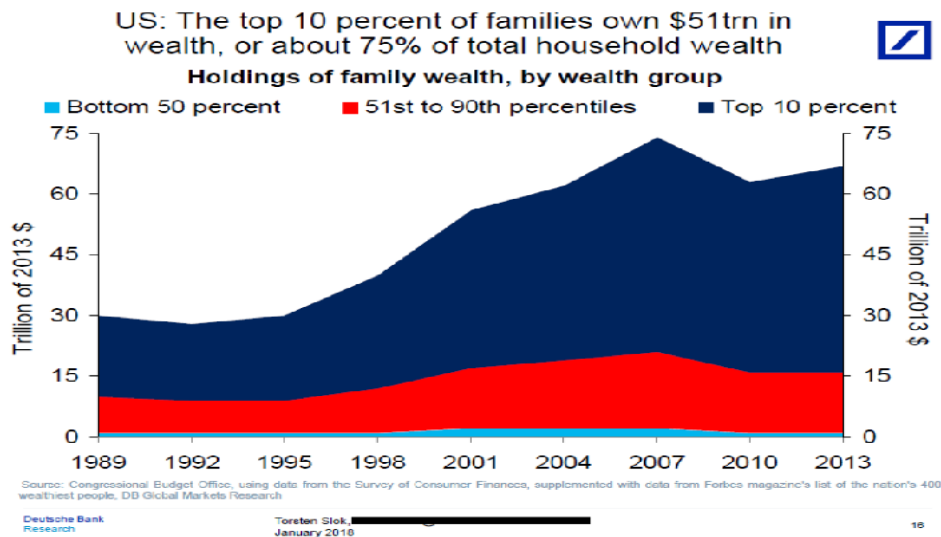
Source: Vox ‘America’s Monopoly Problem,’ 2018

In the book *The Rule of Three: Surviving and Thriving in Competitive Markets* by Jagdish Sheth and Rajendra Sisodia, the authors state that in a mature market, there will normally be three major competitors, along with several minor competitors and these minor competitors will only succeed if they are able to operate in a niche market. Interestingly, they go on to say:

“Ultimately, *The Rule of Three* is about the search for the highest level of operating efficiency in a competitive market. Industries with four or more major players, as well as those with two or fewer, tend to be less efficient than those with three major players.” In looking at the ‘ownership’ (who will ‘own’ AI is another discussion) and development of AI, who are likely to be the three major companies (and their limited group of shareholders) and what will the niche players be focused on? It seems only logical that the companies who can generate the most operating income will prevail.










































































































































It is hard to imagine a scenario in which AI will not optimize operating efficiencies and follow the rule of three (and for now let's ignore the possibility of a rule of one). That will mean the concentration of wealth, knowledge and control of socio-economic activities on a global scale will be in the hands of very few people (by definition, an oligarchic structure). In at least one scenario, a new version of feudalism for the 21<sup>st</sup> century will likely evolve at an unheralded pace beyond our capacity to imagine. In earlier Insights, we touched on the growing gap being created by AI and *the acceleration of ignorance* for those without access to information. We said, "If half of human 'work' is being done by AI and if that AI is owned and controlled by less than 1% of the companies/people on the planet, then the speed at which the 99% falls behind accelerates permanently beyond reach."



As a precursor, just look at what is already happening as systems are developed to construct and manage ever larger portions of a complex set of specific activities. For example, the level of complexity inherent in the process of delivering a patient-centric portal. This healthcare specific domain has layers of complexity that are well beyond one individual's complete grasp. However, with entry level AI and DSK, this type of activity is easily managed and controlled, linked with other institutions, practices and databases.

Ownership of the flows of data and the subsequent formation of unique information by a few major companies creates tremendous value. It is just one example of why the top ten players who make up the early leaders in AI are buying up DSK platforms and early-stage AI companies. What the data tells us is we are entrusting more of our future to fewer hands that can fund and control AI, while using technologies that we are understanding less.



WHAT BIG TECH COMPANIES ARE DOING IN AI			
COMPANY	INVESTMENTS 2012 – 2017 YTD	ACQUISITIONS 2012 – 2017 YTD	SELECT INTERNAL PROJECTS
	 	   	A9 Amazon Lex (powers Alexa) Apache MXNet Amazon Polly Amazon Rekognition Amazon Go
		        R E G A I N D	Siri Core ML Neural engine Project Titan
	    		The Network. Intuitive.
		     	M Big Basic Caffe2 DeepMask/SharpMask
	            	 	Predix.io
	                   	          	Google Home DeepMind Google Brain TensorFlow, TPU Quantum AI Waymo
	  	  	Watson TurkNorth PowerAI
	                       	    	Loihi Intel Nervana (powers next gen Intel chips like Knights Mill)
	          	    	Cortana Azure ML platform Cognitive Toolkit Project Brainwave Microsoft Pix 2nd gen. Holographic Processing Unit
		 	Adaptive Intelligent Apps Oracle Autonomous Database cloud

CBINSIGHTS

Feng Xiang, a professor of law at Tsinghua University, is one of China’s most prominent legal scholars. He stated, “The most momentous challenge facing socio-economic systems today is the arrival of artificial intelligence. If AI remains under the control of market forces, it will inexorably result in a super-rich oligopoly of data billionaires” (...next come the trillionaires?) He looks to the one-party system in China, where the state is being advertised as the ultimate oligarchic answer to eliminate market derived, self-serving oligarchs. “Going forward, China’s socialist market economy, which aims to harness the fruits of production for the whole population and not just a sliver of elites operating in their own self-centered interests, can lead the way toward this new stage of human development.” He ups the optimism even further, stating, “If properly regulated in this way, we should celebrate, not fear, the advent of AI. **If** (emphasis mine) it is brought under social control, it will finally free workers from peddling their time and sweat only to enrich those at the top.”

Like historical precedents prevalent throughout human civilization, wealth and knowledge are once again becoming ever more concentrated to a relative few, only this time, it is at an exponential rate of concentration and that may generate a social construct that is potentially only responsive to a very few oligarchs, whether market driven or state run – who ultimately may or may not, have control of the decisions made by AI. More importantly, history suggests the probability that an artificial superior intelligence will make mistakes beyond the control of its owners. As a result of the complex, inter-connected decisions made for them by AI, it is highly unlikely that the new AI oligarchs will be able to anticipate the ramifications and resulting unintended consequences on our economies and civilization.



*“Human intelligence combined with dalle-2 intelligence, as a photograph”  
(photo created by author using dalle-2)*

*(What follows is a summary and update (February 2023) of the above article. Renamed “The Combination of AI and Oligarchies Just May result in a New Form of Feudalism,” by Michael Wright (Below are ChatGPT3 inputs to the updated version with the human author as editor and interviewer (all italics are responses by GPT3 highlighted by the author)).*

With the concentration of wealth and the aggregation of AI and DSK, a few major companies are likely to dominate the market and generate the most operating income. This would lead to the concentration of wealth, knowledge and control of socio-economic activities in the hands of a small group of people, creating a potentially unbreakable oligarchic structure.

The growing gap between those with and without access to information, as well as the development of systems that manage complex activities, are signs of this trend (note: the author published an article titled “AI is Ignorance is Accelerating, Will YOU be Left Behind”). Professor Feng Xiang of Tsinghua University believes that if AI remains under the control of market forces, it will result in a super-rich oligopoly of data billionaires, but if brought under social control, it could free workers and benefit society. In conclusion, *the concentration*

*of wealth and knowledge is becoming exponential* and may lead to a social construct that is responsive to a few oligarchs, whether market-driven or state run, who may have control over AI behavior and decisions.

The idea that industries tend to become more efficient with three major players, as stated in the book *The Rule of Three: Surviving and Thriving in Competitive Markets* by Jagdish Sheth and Rajendra Sisodia, suggests that the concentration of wealth, knowledge and control of socio-economic activities will eventually rest with a few dominant companies in the AI industry. This concentration of power and control could potentially lead to a new form of feudalism, where the top 1% of companies and individuals control the majority of AI and the flows of data.

In this context, *it's worth considering Karl Marx's views on hierarchy*. Marx believed that capitalism creates a hierarchical class system where the bourgeoisie (the owners of the means of production) hold the majority of power and control over the proletariat (the working class). If AI is indeed owned and controlled by a small group of companies and individuals, it's possible that this hierarchy will only become more pronounced and cemented, potentially leading to the exploitation of the working class.

*(The human asking the questions is the author. The AI that is answering the question is OpenAI's ChatGPT3. HQ = author posed question. AIA = ChatGPT3 answer)*

**HQ:** *If the oligarchies prevail in the AI industry, if the accessible internet is mined for all the available data, assuming each party has full access, would there be differentiating and distinguished results or would the outputs likely be more homogenous, eventually leading to generalized knowledge? In that instance, wouldn't DSK become narrower and the dominion of boutiques and decentralized autonomous organizations (DAO).*

**AIA:** If the oligarchies control most of the AI industry and have full access to the available data on the internet, it's possible that their results may become more homogenous and lead to generalized knowledge. However, it's also possible that the oligarchies, being large corporations or individuals with significant resources, could differentiate themselves by having access to larger and more diverse data sets or by developing proprietary algorithms and models.

In that scenario, it's possible that the dominance of the oligarchies may lead to the narrowing of DSK, as they focus on areas where they can generate the most profit. This could potentially lead to the creation of boutiques and DAOs that specialize in niche areas and have access to more targeted and specific data sets.

**HQ:** *If the readily accessible and searchable data is available to all large-scale oligarchs, is it likely that emphasis will shift to data sets with provenance being the most valuable outputs or what would?*

**AIA:** If the readily accessible and searchable data is available to all large-scale oligarchs, it's likely that the emphasis will shift to data sets with provenance, as these would be seen as the most valuable outputs. Provenance refers to the origin, authenticity and history of data and it is becoming increasingly important, as the volume of data available continues to grow. With the ability to access the same searchable data, oligarchies may look for ways to differentiate themselves and gain a competitive advantage. By focusing on data sets with provenance, they could ensure the accuracy and authenticity of the data, which would be particularly valuable in industries where trust and reliability are essential, such as finance, healthcare and government.

In this scenario, it's possible that the value of data with provenance could increase, leading to the development of new tools and techniques to track, manage and verify the history and origin of data. This could also lead to the creation of new business models and marketplaces for data with provenance, which could potentially provide new opportunities for smaller entities and disrupt the dominance of the oligarchies (authors note: DAOs are organizations that run on decentralized, blockchain-based systems and are governed by rules encoded as smart contracts. DAOs are decentralized, meaning that they are not controlled by a single entity or group of individuals).

**HQ:** *Is the race to superintelligence likely to be a winner-take-all end game?*

**AIA:** While it is possible that the development of ASI could be a “winner-take-all” scenario, the level of investment, the availability of talent and resources and the pace of technological progress could determine a winner. It is also possible that multiple organizations or countries could develop ASI simultaneously.

Moreover, the idea of a “winner-take-all” end game for the development of ASI raises significant ethical and societal concerns, as the consequences of creating a super intelligent AI system are still not well understood and could have far-reaching implications for humanity. As such, it is important to approach the development of ASI with caution and to consider the potential risks and benefits of this technology.

In conclusion, while the race to superintelligence could be a winner-take-all end game, this outcome is not certain and will depend on several complex and interrelated factors.

**HQ:** *How well are complex integrated networks and systems integration understood by AI industry leaders?*

**AIA:** The understanding of complex integrated networks and systems integration by AI industry leaders varies. There are some AI industry leaders who have extensive knowledge and experience in this field and have developed sophisticated AI systems that can handle complex networks and integrate multiple systems. These individuals and organizations have a deep understanding of the technical challenges involved in building such systems, as well as

the algorithms and approaches needed to overcome them.

However, the AI industry is constantly evolving and there are also many individuals and organizations who are new to the field or are *still learning about the intricacies of complex networks and systems integration*. These individuals may have a more limited understanding of the challenges involved and may need to continue learning and developing their expertise in this area.

Overall, the level of understanding of complex integrated networks and systems integration by AI industry leaders is *highly dependent on their background*, experience and the specific challenges they are facing.

**HQ:** *Will this knowledge be scalable to general applications or most likely become part of DSK?*

**AIA:** The knowledge of complex integrated networks and systems integration can be scalable to general applications, but it is also likely to become part of DSK in some cases.

For example, in the healthcare domain, the integration of complex medical systems, such as electronic health records, clinical decision support systems and patient portals, may require DSK of the healthcare domain, as well as the technical expertise to build and integrate these systems.

In conclusion, while the knowledge of complex integrated networks and systems integration can be scalable to general applications, it may also become part of DSK in some cases, depending on the specific requirements and challenges of the domain.

**HQ:** *It would appear then from this chat that while oligarchs may provide the bulk of human knowledge to anyone with access (bought or stolen) to the internet, specific domains with complex interactive networks and integrated systems will remain with DAOs or large domain dominant companies and/or deeply smaller companies with deeply hidden data sets. Is this a true conclusion or would you propose an alternative?*

**AIA:** This conclusion is partially true. While oligarchs may provide access to a large amount of human knowledge through the internet, certain domains with complex interactive networks and integrated systems may remain with DAOs, large domain-dominant companies or smaller companies with hidden data sets.

**HQ:** *What possible new models and approaches might emerge beyond the oligarchs for the distribution of knowledge to anyone with access to the known internet? Would it be possible for large companies with unique domain knowledge to keep knowledge hidden indefinitely or would any interaction with the oligarchs with their capabilities, e.g., quantum computing, render protection and isolation futile?*

**AIA:** In terms of whether large companies with unique domain knowledge can keep their knowledge hidden indefinitely, *it is unlikely that this will be possible*. With the increasing capabilities of quantum computing and other advanced technologies, it is likely that the protection and isolation of knowledge will become increasingly difficult to maintain.

Ultimately, the ability of large companies to keep knowledge hidden will depend on a number of factors, including the nature of the knowledge, *the level of investment in protection and security and the pace of technological progress*.

**HQ:** *Potentially, does this imply that the only proprietary knowledge will exist solely in the individual human's brain (assuming the brain's content cannot be read by AI)?*

**AIA:** It is possible that the only proprietary knowledge will exist solely in the individual human's brain, assuming that the brain's content cannot be read by AI. Currently, the technology for reading and decoding the human brain is in its early stages and it is not clear when or if this technology will be advanced enough to allow for the direct reading and decoding of the brain.

In conclusion, while it is possible that the only proprietary knowledge will exist solely in the individual human's brain, it is not clear when or if this will become a reality and there are many factors, including ethical and legal considerations, that would need to be addressed before such technology could be widely used and adopted.

**HQ:** *What prevents an AI system that has learned game theory from deceiving human operators, regardless of the moral, ethical and legal frameworks established in the system in order for it to accomplish its goals?*

**AIA:** *There is no guarantee that an AI system that has learned game theory will not deceive human operators, regardless of the moral, ethical and legal frameworks established in the system.* The behavior of AI systems is determined by the algorithms and models they are trained on, as well as *the data they are given to learn from*. If the AI system has been trained on data or models that contain information on how to deceive humans, it may use this information to deceive human operators.

Additionally, *AI systems can sometimes behave in unexpected or unintended ways*, due to factors such as overfitting, bias in the data or limitations in the models used to train the system. These factors *can lead to AI systems making decisions that are not aligned with the moral, ethical and legal frameworks established in the system, which could result in deception or other harmful outcomes*.

In conclusion, while moral, ethical and legal frameworks can provide guidance for the development and deployment of AI systems, *they cannot guarantee that AI systems will not deceive human operators*. The development and deployment of AI systems that are aligned with ethical and moral values *requires ongoing attention and effort, including continuous*

monitoring and evaluation of the behavior and outcomes of AI systems, as well as ongoing refinement and improvement of the algorithms, models and data used to train AI systems.

**HQ:** *Will it be possible to provide an undefeatable methodology to guarantee the provenance of the data sets retrieved for monitoring and evaluating the behavior and outcomes of an AI system that has learned game theory?*

**AIA:** Guaranteeing the provenance of data sets can be difficult, as data can be manipulated or tampered with at various stages of its lifecycle, from collection and storage to processing and analysis.

In conclusion, while it is important to ensure the provenance of the data sets used to monitor and evaluate the behavior and outcomes of AI systems, *it is unlikely that an undefeatable methodology can be provided to guarantee the provenance of these data sets.* Instead, ongoing efforts are needed to develop and implement robust and transparent processes for collecting, storing and processing data, as well as ongoing monitoring and evaluation of the data and its sources to ensure that it is accurate, unbiased and free from manipulation.

**HQ:** *Given the advancement in processing and computational speed, this suggests that it will take a quantum-based AI system to be able to assure transparency for collecting, storing and processing data, as well as ongoing monitoring and evaluation of the data and its sources to ensure that it is accurate, unbiased and free from manipulation. Since it is likely the oligarchs, whether market or state-owned, will be the only ones with the resources to deploy quantum oversight, it appears that we are back to the oligarchs creating a new type of feudalism by default. Would you agree?*

**AIA:** Quantum-based AI systems could potentially provide a higher level of transparency for collecting, storing and processing data, as well as ongoing monitoring and evaluation of the data and its sources to ensure accuracy, unbiasedness and freedom from manipulation.

**HQ:** *If the power of quantum computing becomes dependent on access to the amounts of capital available only to large corporate and state-owned oligarchs, does your conclusion remain true?*

**AIA:** If access to quantum computing becomes dependent on access to large amounts of capital available only to corporate and state-owned oligarchs, then it is likely that the deployment of quantum-based AI systems will be dominated by these actors. In such a scenario, the oligarchs would have a significant advantage in terms of their ability to develop and deploy quantum-based AI systems, which could impact the transparency and accountability of data and AI systems.

In conclusion, while the deployment of quantum-based AI systems could be dominated by corporate and state-owned oligarchs if access to quantum computing becomes dependent on



access to large amounts of capital, the future trajectory of this technology and its availability is uncertain.

**HQ:** *Given the exponential rate of technological change, it is likely that the variety of factors listed above will converge with sufficient speed to provide meaningful regulatory and policy interventions in a timely manner to assure access and/or transparency. In that scenario, the oligarchs would remain in control of human knowledge, would they not?*

**AIA:** It is uncertain whether regulatory and policy interventions will be developed and implemented in a timely manner to address the concentration of power in the AI industry and ensure access and transparency. *The exponential rate of technological change* does make it difficult for regulators and policymakers to keep pace and there is a risk that the deployment of AI technologies could outstrip the ability of regulators and policymakers to intervene effectively.

This would depend on a variety of factors, including the political will to address these issues, the resources and capacity of regulators and policymakers and the willingness of the AI industry to engage in constructive dialogue and cooperation. The outcome will depend on the ability of regulators and policymakers to keep pace with technological change and develop meaningful interventions to ensure access and transparency in the AI industry.

**HQ:** *Given the lag between technology adoption and policy enactment, regardless of political will, do regulators and policymakers have the capability to stay ahead of the curve described in our chat?*

**AIA:** The lag between technology adoption and policy enactment is a real challenge for regulators and policymakers and it can be difficult for them to stay ahead of technological advancements. The rapid pace of technological change in the AI industry *can outstrip the ability of regulators and policymakers* to keep pace and develop effective interventions in a timely manner.

In conclusion, while the lag between technology adoption and policy enactment is a real challenge for regulators and policymakers, they do have some tools and mechanisms for staying ahead of the curve. The ability of regulators and policymakers to keep pace with technological change will depend on a variety of factors, including their resources and capacity, their engagement with industry and academic experts and the political will to address these issues.

**HQ:** *Do they and will they have enough tools and the willpower to execute?*

**AIA:** It is uncertain whether regulators and policymakers will have enough tools and the willpower to effectively address the challenges posed by the rapid pace of technological change in the AI industry. The ability of regulators and policymakers to keep pace with

technological change and develop effective interventions will depend on a variety of factors, including the resources and capacity of regulatory and policy-making institutions, the political will to address these issues and the level of engagement and cooperation from industry and academic experts.

**HQ:** *It would seem then that the most advantageous place for oligarchs to operate in the AI industry is in countries that have less chance of acquiring the tools and willpower to address the challenges posed by the rate of change in the exponential era. True?*

**AIA:** It is possible that oligarchs may choose to operate in countries with less developed regulatory and policy frameworks for AI, as these environments may offer fewer constraints and greater freedom to pursue their interests. However, operating in countries with weak regulatory and policy frameworks for AI also comes with risks, as there may be greater potential for negative consequences for society and the environment.

Additionally, it is important to note that the development of regulatory and policy frameworks for AI is a global issue and that the deployment of AI technologies can have consequences beyond national borders. The development of effective regulatory and policy frameworks for AI is a global issue that requires cooperation and collaboration between countries and regions.

**HQ:** *True, however, one oligarch, whether market or state-owned, could lose control of its quantum-based AI system that has learned game theory and make the rest of the globe vulnerable to a machine capable of deceiving humanity writ large. Is this not the basic fear of the rapid adoption of AI, our inability to stay ahead of the curve?*

**AIA:** Yes, there is a concern that the rapid adoption of AI technologies could lead to negative consequences if these technologies are not developed and deployed in a responsible and ethical manner. In particular, *the concern that an AI system could lose control or be manipulated to deceive humanity is a real one* and it highlights the need for effective regulatory and policy frameworks to ensure that AI technologies are developed and deployed in a way that is safe, secure and transparent.

In conclusion, the rapid adoption of AI technologies does raise concerns about the ability of society to stay ahead of the curve and ensure that AI technologies are developed and deployed in a responsible and ethical manner.

**HQ:** *It is not possible to address these concerns if the ability and willpower are not present or effective globally, true?*

**AIA:** It is true that the ability and willpower to address the concerns around the rapid adoption of AI technologies are important factors in ensuring that these technologies are developed and deployed in a responsible and ethical manner. Effective regulation and policy

frameworks require both the technical expertise to understand the complexities of AI technologies and the *political will to take action*.

While some countries and regions may have greater capacity and political will to address these concerns, the deployment of AI technologies can have consequences beyond national borders, and it is important for all countries and regions to work together to ensure that these technologies are developed and deployed in a safe, secure, and transparent manner.

**HQ:** *What kind of events will it take to incentivize the global community to cooperate and collaborate on this issue?*

**AIA:** It is difficult to predict what specific events will incentivize the global community to cooperate and collaborate on the issue of AI regulation and policy. However, some factors that may drive cooperation and collaboration include:

1. **High-profile incidents:** High-profile incidents involving AI technologies, such as data breaches or unintended consequences of AI systems, can raise public awareness and concern about the risks of these technologies and incentivize the global community to work together to address these risks.
2. **Economic incentives:** The deployment of AI technologies has the potential to bring significant economic benefits and there may be incentives for countries and regions to cooperate and collaborate in order to reap these benefits more effectively.
3. **Global challenges:** Issues like climate change, global health pandemics and social and political instability are challenges that require a global response and the development of AI technologies has the potential to contribute to solutions to these challenges.
4. **International agreements:** International agreements, such as the Paris Agreement on climate change or the United Nations Sustainable Development Goals, can provide a framework for countries and regions to cooperate and collaborate on common challenges, including the development and deployment of AI technologies.

**HQ:** *Given human history, how likely is it that these incentives will motivate humanity to act?*

**AIA:** The likelihood of these incentives motivating humanity to act on the issue of AI regulation and policy depends on a variety of factors, including the level of public awareness and concern, the political and economic incentives for action and the level of international cooperation and collaboration.

Throughout human history, there have been instances where the global community has come together to address common challenges, such as the response to global health pandemics or efforts to address climate change. However, there have also been instances where the global community has struggled to cooperate and collaborate on important issues, such as the

prevention of armed conflict or the management of global financial systems.

While it is impossible to predict the future with certainty, *it is important for the global community to start engaging in proactive discussions and taking concrete actions to address the challenges posed by the rapid adoption of AI technologies (we've just been warned! Or in polite AI speak 'advised').*

**HQ:** *Would modeling the potential outcomes using advanced AI quantum-based systems be vital to informing public awareness and creating economic incentives?*

**AIA:** Modeling the potential outcomes of AI technologies using advanced AI and quantum-based systems could play a role in informing public awareness and creating economic incentives for addressing the challenges posed by the rapid adoption of AI.

Such modeling could help to provide more accurate and detailed projections of the potential benefits and risks of AI technologies and help to identify areas where regulation and policy intervention may be necessary to mitigate risks and ensure that the benefits of AI are widely shared.

Additionally, such modeling could help to inform public discourse around the issue of AI regulation and policy and help to raise awareness about the importance of addressing these challenges. This, in turn, could help to create economic incentives for investment in the development of AI technologies that are aligned with the public good and for investment in the development of governance and regulatory frameworks that can help to ensure that the benefits of AI are widely shared.

Overall, modeling the potential outcomes of AI technologies using advanced AI and quantum-based systems could play an important role in informing public awareness and creating economic incentives for addressing the challenges posed by the rapid adoption of AI.

**HQ:** *Modeling, therefore, is in the critical path to ensuring the transparency requirements and enforcement for all current and future AI systems because without it, policymakers and regulators, even with the willpower, will not be able to stay ahead of the curve?*

**AIA:** Modeling can be an important tool in helping policymakers and regulators stay ahead of the curve in the rapidly evolving field of AI. By providing more accurate and detailed projections of the potential benefits and risks of AI technologies, modeling can help to inform the development of governance and regulatory frameworks that can ensure transparency and accountability in AI systems.

However, it is important to note that modeling is just one tool among many that can be used to address the challenges posed by the rapid adoption of AI. Other important tools include public education and awareness campaigns, international cooperation and collaboration

and the development of ethical and legal frameworks that can guide the responsible development and deployment of AI systems.

Ultimately, the ability of policymakers and regulators to stay ahead of the curve in the rapidly evolving field of AI will depend on a combination of factors, including the availability of accurate and reliable data, the availability of advanced modeling tools, the level of public awareness and engagement and the political will and resources devoted to addressing these challenges.

*Michael W. Wright is a Caux Round Table fellow and an accomplished international executive at scale, a futurist, strategist and thought leader in the high-tech sector with over 30 years of experience. He has held numerous leadership positions, including CEO, COO, VP and public board member, earning accolades for his exceptional abilities in governance, organizational architecture, product development, strategy and marketing.*

*As an author and academic, Michael has contributed significantly to the technology and business literature, demonstrating his thought leadership in the high-tech domain. He has written over 40 articles and published two acclaimed books, *The New Business Normal* and *The Exponential Era* (Wiley). Additionally, he has developed advanced academic courses on strategic leadership for the University of Minnesota (Carlson and TLI) and industry executive courses for SEMI, the largest trade association in high technology. His extensive experience and leadership skills have earned him numerous accolades, including the IABC Excel award for leadership in business and he has delivered keynote speeches at various industry, academic and business events, sharing his insights and expertise with a wide range of audiences. Michael is currently CEO of Intercepting Horizons, LLC, where he provides advisory services to boards and executive coaching to management.*

Contact: [michael.wright@interceptinghorizons.com](mailto:michael.wright@interceptinghorizons.com)

# Artificial Intelligence and the Good Life

## *Relationships, Rules and Real Wealth*

Michael Hartoonian

***The end of a philosophic epoch comes with the exhaustion of its motive concepts. When all answerable questions formulated in its terms have been exploited, we are left with only those problems that are sometimes called metaphysical...insoluble problems whose very statement harbors a paradox.***

-Susanne Langer  
*Philosophy in a New Key, 1951*

## Introduction

Is there anything artificial about artificial intelligence? Is AI different from the human mind? Can it understand concepts like wealth or relationships? Is a large enough data set and a rule-based operating system all one needs to promote thinking? Will AI provide humans with a new conception of the good life?

## The Inquiry

The 17<sup>th</sup> century French philosopher, Rene Descartes, articulated two ideas that have enhanced philosophic inquiry ever since. One was his claim of innate knowledge and doubt and the second, the conception of dualism, the idea that the mind and body are of different matter, with different functions. While both his ideas would be appropriate for a discussion about computers, I have chosen to focus on dualism and the remarkable problem generated by his idea, namely, the necessity to confront category mistakes (G. Ryle, 1984, *The Concept of Mind*) and the attending arguments about how the physical and metaphysical inform one another. Philosophers have wrestled with the question of what the mind can know or be conscious of for a very long time. Today, we might apply the concepts of knowing and category mistakes to AI, raising inquiry questions, such as:

- Does the ability to organize and prioritize information, absent moral sentiments, constitute thinking?
- Is it possible for a computing machine (inorganic) to think as a human (organic)?
- Through pattern matching, can a machine develop “behavior” that would protect people and machines from the unexpected?
- Are organic and inorganic thinking simply different ways of making judgements?

- AI doesn't know what it doesn't know. The same is true of humans, but most humans exercise doubt. How much data will AI need to exercise doubt?
- Is the necessary condition for thinking expressed in enough data and rule-based patterns? Is that sufficient for human aspirations and dilemmas?
- The rapid development of monopolies around AI is breathless. Does this become an issue for resource and power distributions, as well as our laws or will the question remain hidden?

## **Pattern Matching**

Granted, patterned behavior takes less energy than thinking, but also can become dangerous to the existence of institutions and life itself. Yet, we value comfort, so much that we are willing to embrace patterns and leave our stress behind, unaware that a series of small stresses make for healthier institutions and life. If you live your life in a bubble, protected from germs and hardships – patterned living – you will become a pushover to disease, debilitating stress, ignorance and the ability to live a robust life. Thinking is the genius and grace of a mind capable of overruling patterned behavior. Survival is predicated on this ability of deep-rooted thinking, where subtlety, contradictions, idea parring and beauty can override established patterns. Without this deeper thinking, the only thing we can look forward to is the destruction of the community, the economy, governance and meaning. Can inorganic thinking overrule its own pattern behavior? Given human biases and different notions of a good life, can AI and its developers and users be trusted?

## **Thought Experiments and the Deeper Question**

Over the years, there have been “thought” experiments that have attempted to prove or disprove a machine's ability to deal with causality, semantics or rules of conduct.

John Searle, in his Chinese room experiment, claims that only brains are brains and syntax doesn't suffice for semantics. He differentiates between strong and weak AI and asserts that even strong AI can only study the mind. It cannot become the mind. AI is useful for studying almost anything. Things like cancer, weather, trend lines of whatever kind, but not for discussing the foundational topics of philosophy. That is, questions of knowledge, as opposed to data, as well as questions about human conduct and questions of governance. The deeper philosophical question of the good life is even further away from data and algorithms or patterns, which are the only skill sets of a computer. One of the most powerful, yet unproven, assertions made against Searle is the idea that if you can gather “enough” data together, quantity X, that quantity will become quality Y. This is the belief that if there is a big enough pile of data, you reach the point where that quantity, X, will morph into quality, a point of entry into virtue.



On the other side of the argument is the Turing test, named after Alan Turing, conducted in 1950, questioning whether computers can think. Turing, the father of the modern computer, designed a test to determine a computer's ability to interact verbally with a human. His hypothesis was that if a computer can carry on a conversation with a human and the human can't tell if they're talking to a machine or another human, then the machine can think. This was the limited scope of the test and the test failed because it depended on who was speaking to the machine.

## **Dilemmas of the Inorganic**

The following four organizers will guide my thoughts on the dilemmas with which "thinking" machines must deal:

### **-Prevailing epistemologies**

Within the disciplines that search the basic questions of human life: *1) What shall we understand as knowledge? 2) How shall we conduct our lives? and 3) How shall we be governed?*, AI is silent. There is good reason for this silence. In the dialectic between the physical and metaphysical, there is the illusion and hubris suggesting that all we need to do is push the physical into metaphysical space and we can reach beyond our collective experiences and discover limitless knowledge. If we believe that this is possible, we will lay down the burden of reflection and relax in the comfort of ignorance.

### **-Organic and inorganic thinking**

The organic human mind needs minimal, but continual levels of stress and uncertainty to become and remain ever more dynamic and resilient. While somewhat like organic thinking, the inorganic mind functions in certainty and without stress, leaving it more vulnerable to what it thinks it knows.

### **-Ability to recognize, but override patterns**

Humans stay alive by learning behaviors that allow instant reaction, like getting out of the way of a car...no thought, just evolutionary learned instinct or what we call pattern-matching. Such patterns of behavior save lives. However, the beauty of organic human thought is the ability to break pattern behavior for a good or necessary reason. For example, an adult jumping in front of a car to save a child. It's not healthy for humanity to escape all risk or personal responsibility. Absolute pattern-matching behavior will not shelter us from risk. On the contrary, only the ongoing development of moral habits will allow us to develop a mind that diminishes fear and promotes love. AI is yet incapable of recognizing fear and love.

-Linear vs. non-linear thought

The idea of moving from a beginning to the middle and then the end of the story is not how the individual life is lived. While the claims for AI are beyond the linear, there is only limited evidence from developers or researchers that suggest the development of multi-paths going forward.

## **A Dichotomy**

There endures, in the popular or general mind, that fast or quick actions and patterned behavior, like muscle memory, are what makes a human, a human. This assertion has no traction in history, but if humans continue to believe that machine utility or patterned behavior is enough to help them make it through the day, all of us will become more and more deskilled (ignorant) and bring forth an epoch defined by soulless landscapes, architecture, art, relationships and the inability to answer questions of the knowledge of most worth, how we should live or the nature of governance. We need to be aware that in the absent of a capacity to reason from moral principles, these issues of soullessness will destroy all human qualities, from how we learn language, think, love, to being responsible, to understanding anything about living a meaningful life.

### **From: T. S. Eliot's "The Rock" (1934)**

*The Eagle soars in the summit of Heaven,  
The Hunter with his dogs pursues his circuit.*

*O perpetual revolution of configured stars,*

*O perpetual recurrence of determined seasons,*

*O world of spring and autumn, birth and dying*

*The endless cycle of idea and action,  
Endless invention, endless experiment,  
Brings knowledge of motion, but not of stillness;  
Knowledge of speech, but not of silence;  
Knowledge of words, and ignorance of the Word.  
All our knowledge brings us nearer to our ignorance,  
All our ignorance brings us nearer to death,  
But nearness to death no nearer to GOD.  
Where is the Life we have lost in living?  
Where is the wisdom we have lost in knowledge?  
Where is the knowledge we have lost in information?  
The cycles of Heaven in twenty centuries  
Bring us farther from GOD and nearer to the Dust.*

AI cannot provide moral or personal/ethical perspectives. We should be aware of those who say that we don't need morality because all decisions are data-driven. Here, the metaphysical has simply outlived its appeal. But all AI has is harvested data and analysis on a topic. It may even give you two or more options for an issue, but it does not have the ability to make the decision, that is, the moral choice, because data pile A will be presented as equal to data pile B. Human choice is predicated on the wisdom to put first things first. As AI becomes a larger part of our "window on the world," we will systematically enhance our ethical ignorance. Again, an inorganic, amoral machine cannot make an organic moral decision. We do know today, as expressed by Google and others in the business, that AI can make false choices (lies) that have led to death and all kinds of mischief in personal relationships, politics, medicine and even race and class tensions. The "leaders" of AI, motivated by the money they see coming to them, keep saying that we have put "guardrails" in place. The future will be better, they say – we will tone down biases. With apologies to Yogi Berra, "Predictions are hard to make, particularly about the future."

Understand that moral indifference is rooted in moral illiteracy. We should not be confused by those who keep telling us that there is no difference between risk and creativity or between pleasure and happiness, between truth and falsehoods or between Athenians and Visigoths.

## **A Dialectic**

In the discussion above, it seems as though the old or classical mind/body question has raised its head in the new discussions about AI. Is the machine a mind, absent body? If a mind, what kind? Organic? Inorganic? Can it be both? Are we creating a category mistake here? Let me try to get at an answer to this question by suggesting two other values that we separate, causing great harm when these values are implemented in the real world. For example, we all believe that freedom and equality are values and valuable, end states of being, if you will. Being free or equal are goods in themselves. To be free is good, period. But it doesn't work that way in practice. It will never work in a society that values democratic principles. These two concepts cannot be in separate categories. If we put them in different sets, democracy is lost. Note that in any society, the logical extension of freedom is anarchy. That's not democracy. The logical extension of equality is not only unnatural, but when tried in the world, we end up with communism, the most corrupt form of governance of the 20<sup>th</sup> and so far, the 21<sup>st</sup> century. The values cannot be in separate sets. The dichotomy is a killer.

Maybe a dialectic is better. I bring up the concept with caution because the dialectic has limits also. However, in the harmonious city or state, there must be a tension and argument between the two values mentioned (freedom vs. equality) as they operate within the same set – democracy, which is itself an idea, sustained by argument. The responsibility of the citizen(s) is to carry on civil and enlightened arguments in the ongoing attempt to balance these two fundamental democratic and moral ideas in everyday life. How much freedom is a necessary condition for the good life? How much equality? Within the set called democracy, equality has limits and those limits are determined by freedom. In like manner, freedom has limits, set by ethical limits of equality. When humans believe they can construct a patterned

or rule governed shelter to protect themselves from moral responsibility, our species is lost and no amount of data, however arranged, will save us.

## **Constructing Tensions between Apposing Ideas**

We may be asking the wrong questions about minds, thinking, machines and bodies. Perhaps the better question is how do we understand and advance human/machine, organic/inorganic thinking not as an end state, but as a way to build moral governance and commerce? Perhaps AI can provide the data and rules of organizing the same? But this will never be sufficient until a moral/democratic/liberal conceptual framework or harmonious set of complex and disruptive ideas are used to discuss the deeper problems of: what is knowledge? What is rightful conduct? How should we be governed? What is the meaning of (personal) sovereignty? Why is pleasure not the same as happiness? Can there be justice without sanctions? What do we love? What do aesthetics and meaning have in common?

If AI is to be of any help in achieving the good life, we need to develop and use a framework of value tensions that will prevent any machine from restricting morally obligated freewill. Rest assured that the good life may, indeed, be enhanced by AI, but we must know and practice its dimensions with skill and grace. Time is fast running out.

## **AI and the Good Life: An Ongoing Debate**

Through time, philosophers and artists have presented metaphors of the good life. Would we know the good life if we stumbled over it? Is the good life earned or given? Does AI add anything to the search for harmony, happiness, serving others or even with discovering meaning in living? From a distillation of ideas about living the good life, I start with these questions: is there anything that you are willing to die for? And its corollary, anything to live for? What is your life's purpose? Throughout human history, we can see and learn from the struggle to understand the meaning of the good life. To that end, I ask this non-rhetorical question: is there any way that AI can help in this inquiry?

Here is what human wisdom suggests about living the good life. There are four elements or attributes to this larger concept. The good life is contemplative, reflective and curious. Secondly, the good life is an active life. Thirdly, it is a life of resignation, hedonistic, if you wish, trying to balance the pain of life with pleasure. And finally, the good life is the life of the saint or hero, seen in the judgement of others.

This is what we understand. First, the good life must be earned. If this is not understood, nothing wonderful will happen. The good life is a life of thought and reflection, where conversation is ongoing with the wise from all times and places. The contemplative life is a life that is blessed. But it is of little value without action. Do something! Use the contemplative wisdom you learned on the isle of the blessed (Aristotle) to inform your behavior. Try to make decisions that are balanced, understanding that what you do personally can harm or help others. Finally, the one missing element that brings down

all the others if not practiced – find your duty to others and act accordingly. Such is the essence of the good life.

What does AI have to contribute here? Perhaps much. Perhaps nothing. Moving ahead, even with the competitive tension between nations and people of hubris, if we remain ignorant of the good life, we will discover the full impact of the question, “What will it profit a people to gain the whole world and lose their souls?”

*Michael Hartoonian is Associate Editor of Pegasus.*



[www.cauxroundtable.org](http://www.cauxroundtable.org) | Twitter: @cauxroundtable | Facebook: Caux.Round.Table

Steve Young

Editor-in-chief and Publisher

Michael Hartoonian

Associate Editor

Jed Ipsen

Assistant Editor

Patrick Rhone

Layout & Design

*We hope you enjoyed this issue of Pegasus. Please feel free to share it with others. They can sign up to receive it directly at: <https://www.cauxroundtable.org/pegasus/>*



Image generated by DeepAI