Further Musings of Derek Ireland on AI Complexities and Chaos

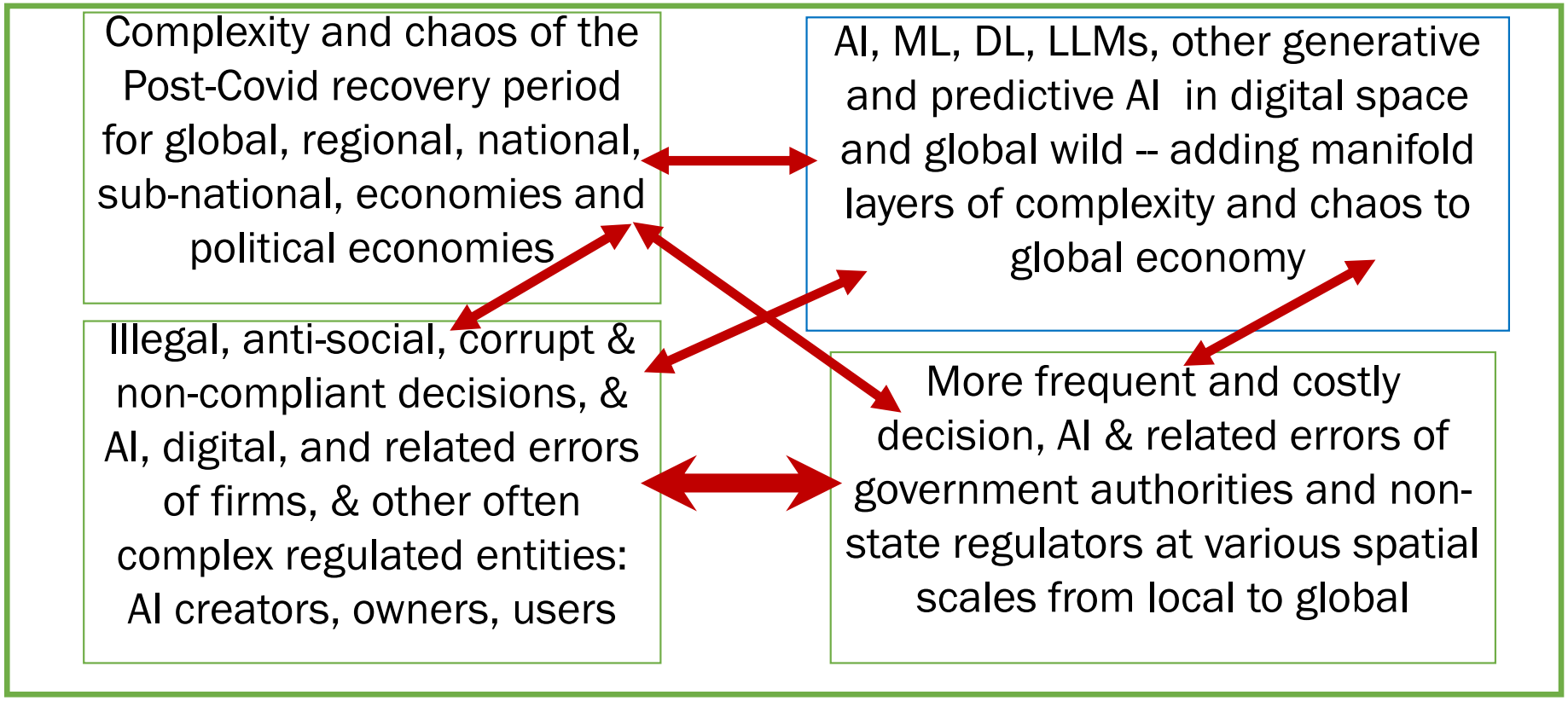
Presentation to Foresight Synergy Network (FSN) Members

September 25 2025 1 PM

Purpose and Background

- Update earlier and under-developed musings in FSN presentation of December 2024, which were
- Then placed in a working paper now available on SSRN at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5176031
- Applied behavioral and complexity science lens to artificial intelligence, and now also incorporates a chaos theory lens
  - In light of the chaos caused by, integral to, and/or surrounding AI since then such as
  - New US president, emergence of PRC as major player in AI global strategic competition game, recent research on AI hype, "snake-oil", sycophancy, and errors,
- With emphasis on "new" AI goods, bads, challenges, surprises (good and bad), double-edged swords, and related themes that emerged over past 8 months

Four Major, Interrelated, and Overlapping Complexity and Chaos Themes and Challenges for AI and Global Economy



Adding Chaos Theory to Slide 3 Exhibit

Amplifies Two-Way Interactive and Feedback Effects

- Across and between the four sources of complexity and chaos within above Exhibit
- And how and why expanding complexity and chaos within and between the external AI environment and the internal AI world is
  - Leading to additional complexity and chaos effects emanating from the errors of businesses/regulated entities and state and non-state regulators
- As well as butterfly effects from many, diverse, and often surprising and unexpected sources (further discussed in later slides)
  - With complexities and chaos integral and endemic to and caused by AI
  - Making greater contributions to global economic, social and technological complexity and disruption through past 8 months

Some Major Themes in Dec. 2024 Presentation, Earlier Research and First Report

- Comparing AI goods and bads; larger number of provocative AI bads; resulting double-edged swords and continuums; and resulting decision making and related errors in public & private spheres.
- Policy, legal & regulatory regimes as complex systems made more complex and chaotic by AI, emerging global dimensions, and why national context matters
  - AI-have countries compared with AI wannabes, "soon-to-be's" and "have-nots" -- and the implications for AI and digital power asymmetries, inequalities, and "technocolonialism"
- All aspects of AI encompass complexity, uncertainty, ambiguity, unpredictability, contingency & "what-ifs" – leading to doubt & anxiety
  - With layering of complexities from manifold sources in AI world/global wild
  - And AI through time adding more and more to these layers in many, diverse, and often not obvious ways

Some Major Dec. 24 Themes Cont.

- Public trust, confidence, and faith in AI circumscribed to extent that AI creation, ownership, diffusion, and benefits dominated by
  - A few powerful big-tech digital firms, captured governments and their authorities and the military industrial complex
- Which are better positioned to profit from & exploit AI & related complexity and chaos compared with government, civil society groups, other non-private sector entities
- Businesses and governments, which deploy and exploit AI to
  - Monitor, manipulate, deceive, and harm consumers, small business, other users, voters, and other less powerful actors,
  - Further diminish AI acceptance, performance, success, and progress over longer-term
- How and why the two-way interactive and feedback effects between regulatee and regulator decision errors at bottom of the Exhibit
  - Can amplify complexity and chaos in the AI external environment and within the AI world and business community at the top

Summary of Fundamental Changes Since Then

Over Past Eight Months and Perhaps Longer Such As from First ChatGPT in November 2022

- External environment surrounding creation, development, adoption, diffusion, regulation, and governance of artificial intelligence
- Has changed dramatically, fundamentally, and not for the better
  - And has become even more complex, ambiguous, contingent, opaque, perplexing, unpredictable and chaotic
- As have LLMs and other generative and predictive AI models, applications, algorithms, systems, goods, bads, and double-edged swords
  - Leading to even more frequent, consequential, and difficult to correct errors by all actors, human and artificial and public and private
- Malevolent AI actors are expanding from creators, developers, and owners to human users and artificial agents (at times working together)

Making future complexity science, chaos theory and behaviorally informed AI research even more imperative, covering virtually all the presenter's musings in two PPTs. As stated by one AI analyst:

"New, interdisciplinary approaches are needed to bridge competing conceptualizations of AI in practice and help shape the future of AI use"

13 Cross-Cutting Themes From More Recent Research (1)

AI Goods, Bads and Double-Edged Swords

- AI goods getting better, AI bads getting worse and becoming more controversial, and receiving much greater attention,
- Double-edged swords and related continuums more frequent and complex as generative AI size, complexity and sophistication expand
  - Sword width becomes greater, both sides become sharper & more consequential, as well as
  - More challenging to assess using benefit/cost and other more conventional socioeconomic impact and related analytical tools
- And more obvious across manifold domains: healthcare, mental health, environment/climate change, education/learning, science/scientific publications, history, the law, inequality, power asymmetries
  - As well as emerging *existential double-edged swords* such as
  - Rise of China as AI and big tech leader (see below)
  - AI as saviour or oppressor/destroyer of humanity, with negative AI arguably a greater existential threat than climate change and nuclear war (?)

More on Double-Edged Swords (2)

- Trends expected to continue and perhaps even accelerate as artificial intelligence and the environments surrounding AI
  - Become even more complex, uncertain, unpredictable, chaotic, and difficult to govern, understand and explain i.e. “inexplicable” nature of AI
  - As AI models become larger and more complex and sophisticated, they also become more error prone, independent, difficult to control, naughty, nasty, and inexplicable to all players from creators to final users and consumers
  - And more like real human beings and is that a good thing (? see below)
- Over past eight months, recent research, articles, Internet blogs focused more frequently, aggressively, and convincingly on AI bads and
- The negative sides of these double-edged swords, increasing the risk of negativity and related biases for all of us associated with
  - Excessive AI hype, expectations, promises not achieved, and errors not corrected
- More attention as well to distinctions between generative AI “snake-oil” predictive AI limitations, and their many double-edged sword attributes

Cross-Cutting Themes (3-7) AI & Government

And Other Non-Private Sector Entities e.g. Civil Society Groups, Other Non-State Regulators, Health, Education etc. Institutions, Other Non-Profits

- Many and expanding number of AI problems, deficiencies, and limitations covering both generative AI & predictive AI
- Are more readily identified, assessed, and remedied by business community, other private sector actors, perhaps as well final consumers
- Compared with governments and their authorities, non-state regulators, civil society groups, other non-private organizations.
  - Negative impacts of AI imperfections are easier to identify and remedy by businesses through significant and measurable reductions in profits, revenues, ROI,
  - Less successful advertising and marketing programs, and higher AI investment costs compared with AI induced revenues and other positive and measurable financial impacts

Cross-Cutting Theme (4) AI & Government

- And reductions in consumer and other purchaser satisfaction can be measured through willingness-to-pay and related consumer/purchaser satisfaction studies
  - And by falling revenues and market shares of consumer-facing businesses that create, employ, and/or misuse AI
- In stark contrast, governments and other non-private sector entities do not have similar metrics, indicators, and benchmarks and therefore are more vulnerable to be
  - Unduly influenced by excessive AI expectations, hype and euphoria on efficiency benefits and work force reductions to point where:
  - Human oversight of AI, including identifying and correcting AI errors, become impossible for these organizations

How, why and under what circumstances AI can benefit governments and their decision- making processes has received little attention -- AI benefits are simply presumed

Cross-Cutting Theme (5) AI & Government

In light of expanding complexities and chaos in AI world over past eight months and growing literature on AI deficiencies and limitations of LLMs and generative AI as well as predictive AI

- One has to wonder why some very smart and highly experienced government, corporate, civil society and academic leaders and experts believe that AI will contribute greatly to
- Reducing government expenditures, employment, deficits, debt and waste, thereby making AI the panacea for fixing government given that:
  - Government knowledge of, experience with, and progress in understanding and regulating AI are all highly limited
  - Using AI to reduce government employment is especially questionable when human oversight on AI is imperative
  - Evidence is expanding on government situations where social costs of using AI would be greater than financial and economic benefits

If governments cannot regulate AI, why presume they know how to use it effectively & efficiently in their daily operations

Cross-Cutting Theme (6) AI and Future Oriented Government

Addressing climate change, disruptive technology and innovation, pandemics, and other existential threats and opportunities

- Demands more thoughtful, deliberative, and future conscious long-term decision making by governments & their authorities
- However, AI emphasis on speed, efficiency, and economizing on time, labour, & other scarce resources is more likely to encourage and justify short-termism
  - While AI hallucinations, deceptions, lies, bad advice, misinformation, disinformation, and sycophancy, that is, telling short-termism officials what they want to hear
  - Would further distort and undermine more thoughtful, long-term and future oriented decision making, planning, legislation, and administration
- Resulting in faster but not necessarily better decisions by governments
  - In short, using AI to do the job faster and easier is not a catalyst for long-term future conscious thinking

Therefore, AI applications & models in these contexts should be limited to simpler, more routine, straight-forward, mundane and boring planning, legislative, and related tasks, such as AI summaries of future conscious legislation and government planning elsewhere -- likely a short and easy to do AI summary document. To paraphrase many AI critics: "If the decision is important, don't use AI."

One More Theme (7) on AI and Government

Perplexity, Paradox, and Source of Friction and Decision Errors

- Governments, their authorities & non-state regulators are expected to apply honesty, modesty, humility and long-term future orientation to assessments and regulation of AI & other transformative and disruptive technology
- Whereas AI applications and their creators, owners and proponents display little modesty, honesty, humility, and future orientation
  - Which are so important to promoting responsible and accountable AI
- When regulators are strongly and unduly influenced by excessive AI expectations, hype, promises, and euphoria from AI business community and proponents within and outside government
  - Almost inevitable result is higher risk of more frequent & consequential under-enforcement decision errors in short to medium term
  - As regulators and other officials go out of their way to avoid excessive enforcement errors

But, in long term, criticism for not protecting the public interest can lead to excessive enforcement and regulatory requirements errors and related mega-decision error “non-events”, which are are less obvious and more difficult to correct-- the “AI regulatory free-ride may be ending

Theme (8) Treating AI as Normal Technology

What If AI Turns Out To Be a More Normal (And Boring) Technology

- Like many other disruptive technologies of the past and present rather than a magic carpet ride with some black magic ramifications

Some Dimensions and Implications of AI as a Normal Technology

- Significant adoption, diffusion and related changes and socioeconomic benefits would occur over decades rather than months and a few years because of
  - Slow speed of human, organizational and institutional change, possible effects of "onerous regulation" such as in EU, and weak regulation elsewhere
  - Related regulatory mismatches between jurisdictions and countries, as well as within federations
  - Conventional and social media emphasis on AI bads and when AI benefits fall far short of proponent promises and hype – both feed into the negativity bias

And sudden AI advances take time to be translated into socioeconomic effects, benefits and costs – making AI progress neither automatic nor inevitable (for further discussion)

Cross-Cutting Theme (8) Cont.

Worldview of AI as normal technology, with "normal" transformative and disruptive technology attributes could be

- Much more appropriate, helpful, and preferable compared to utopian and dystopian visions of the AI future -- the existential double-edged sword

And is similar to other transformative, disruptive, general purpose, and "normal" technology in the past: the steam engine, electricity, computers, the Internet, the information and digital revolutions, biotechnology, nanotechnology etc.

- In particular, realizing obvious, significant and measurable productivity, efficiency, acceptance, & performance gains from artificial intelligence
- Requires time, effort, patience, and complementary investments in
  - Equipment, training, and appropriate adaptation, adjustment, re-skilling, and cultural change by
  - Businesses, governments, civil society groups, all other affected organizations and their employees, suppliers, customers, other supply chain partners

Called the "productivity paradox" in the computer age with growing evidence as well for the AI world & age

Cross-Cutting Theme (9) "Big Surprise" & Butterfly Effect of China's AI DeepSeek App

"Big Surprise" in Quotes for a Reason – Yet Another Existential Double-Edged Sword

- No so-called AI experts in US and rest of developed world saw this coming
  - Reasonably good app created by private firm startup in PRC at fraction of cost of earlier US chatbots – the lean AI "less is more" approach
- Illustrated how emerging market and developing economies, the "Global South", can in time participate in competitive strategic game with western AI-have countries led by US. Also demonstrated
  - PRC strength across wide range of disruptive technologies: lean AI, EVs, clean energy, batteries, industrial robotics, quantum computing, nuclear, biotechnology, cloud computing, 5G, social media, on-demand services etc.
  - How AI development can be more open, inclusive, diverse, accessible, & "democratic"

With more emphasis on business and other users, rather than the "next big AI thing" such as AGI, super-intelligence, and making AI applications more human-like

Cross-Cutting Theme (9) Cont.

Negative Side of PRC DeepSeek Double-Edged Sword Includes

- Perverse and counterproductive responses of US government and AI digital oligarchs to this "unpleasant surprise"
- According to some critics, White House AI Action Plan makes
  - International cooperation and collaboration with China and other countries and jurisdictions virtually impossible
  - China the ultimate winner in long run of the global strategic AI competitive game
  - Potential for global AI governance to be centred on China and its expanding BRICS network comprised of many AI wannabes such as India, Brazil, Chile, Mexico, Saudi Arabia, Other Persian Gulf states

Cross-Cutting Theme (10) New Challenges Re Fair Competition and Market Power

Continuing AI Advances in Detecting Anticompetitive Conduct But As Well

1) How and why "Big Cloud" companies, oligarchs and investments can and will dominate AI and digital ecosystems and supply chains, starting with US and China
2) Price collusion, surveillance pricing, and other cartel and anticompetitive conduct through AI agent interactions and social cooperation -- with and without human agent assistance
3) In many AI and related digital contexts, consumer protection, privacy, copyright, other IP violations will receive more attention and cause more harm than competition law violations, through adding greatly to market power and abuses of dominance of AI digital giants
4) Many AI firms/oligarchs consider market power in political markets to be more important than in conventional product & innovation markets

Cross-Cutting Theme (10) Continued

5) Risk that actual and perceived power, mastery, and control over AI and related disruptive technology, in uncertain, complex and chaotic AI global wild

- Would further exacerbate narcissism, opportunism, and other anticompetitive, non-compliant and anti-social attributes of powerful corporate and other organizational leaders – called paranoid and malignant narcissism

6) Over longer term, stronger competition and consumer choice, & more innovative user- friendly products in AI and digital markets in advanced economies may depend more on

(i) Imports, supply chains, and foreign direct investment from emerging market and developing economies, if allowed by protectionist & xenophobic advanced economy governments

7) Growing concerns among competition, labour market, other scholars that AI abuses by employers may be more harmful for workers in conventional and digital labour markets compared with

(ii) AI abuses of consumers and other purchasers in product, digital and related markets, leading to situations where households are experiencing

(iii) Abusive AI in both labour & product markets, making AI even less trustworthy, popular and more destructive as a general-purpose bad technology

Theme (11) Expanding Evidence and Consensus Among AI Scholars That

As AI applications, algorithms, systems, and models become even larger and more complex and sophisticated, they will generate

- Even more frequent and consequential AI hallucinations, deceptions, lies, delusions, bad information and advice,
- Disinformation, intentional misinformation, "jailbreaks"/emergent alignment challenges, and related misconduct
- Leading to safety risks and other outcomes and harms which are even more difficult to explain and correct by
  - Their creators, developers, owners, users and other actors as well as state and non-state regulators

Concerns also emerging that artificial general intelligence (AGI) and AI superintelligence will make the existing situation and challenges even worse

Theme (12) As AI Models, Applications, Algorithms, and Systems

Become Even Larger and More Complex and Sophisticated

- AI limitations, deficiencies, challenges, problems and other bads
- Are going beyond hallucinations, misinformation, disinformation etc. as well as copyright violations, training data constraints, and so on
- To encompass AI problems and failures with conducting and completing more complex tasks and then deceiving themselves and their human agent partners about these failures.
  - Situation could become even more consequential in future with even larger and more sophisticated and complex LLMs and other generative AI models including AGI and AI superintelligence

Consensus is that larger, more complex and sophisticated AI models may be superior to humans in many ways; but these artificial agents will remain boundedly rational, human, humane, ethical, & aligned with our interests

As AI applications, tasks, problems and environments become more complex, AI becomes less reliable as advisors and analysts compared with human beings.

Other “Favorite” AI Musings of Presenter (1)

“Existential double-edged sword game” often played by same AI leaders, experts, proponents and spokesmen on both sides

- ❖Game is highly dangerous for AI evolution and long-term success given the negativity bias and related cognitive biases and heuristics/short-cuts
- ❖How AI bads are easier to understand, construe, remember and recall from memory for non-experts,
- ❖And are often more emotionally and politically charged (further study)

In stark contrast, AI goods receive less media, public and political attention, are often more science-driven & thus harder to understand for non-experts, and can be conflated with, & more difficult to differentiate from, excessive AI promises, expectations, hype, and euphoria.

Promising too much prior to introduction of “next big AI chatbot” can also result in actual benefits being heavily discounted -- because earlier promises and expectations were not realized and AI errors were not corrected

Public Trust and AI Acceptance (2)

Many AI scholars, practitioners and proponents contend that public trust is essential to AI success. But this presenter's perspective is that

- AI creations and its creators, developers and owners must be considered as and seen to be trustworthy, and to have earned the public's trust and confidence
  - Through responsible and accountable AI and responsible and accountable AI creators, developers, owners, and users.
- In short, public trust must be earned and not simply bestowed on AI because of its novelty, transformative properties, and AI hype from industry, government, academia, conventional and social media and other sources
- As a starting point, AI community should allocate more resources to
- Addressing and actually correcting the expanding number of more specific, concrete, tangible, salient, remediable, and obvious risks, costs, hazards, harms and dangers of AI e.g. suicide and other self-harm risks need more than hand wringing

Public confidence and trust in AI is a good thing to develop and sustain. But too much of a good thing can be problematic, leading to excessive personal, social, and organizational dependence and reliance on AI

Which leads to less recognition and awareness of the technology's many and highly diverse errors, risks, costs, hazards and harms

Public Trust and AI Acceptance Cont. (2)

Calling AI misinformation, disinformation and mistakes "hallucinations"

- Is a deliberate marketing ploy, and inaccuracy, deception and euphemism of creators, owners and proponents to cover up AI mistakes
- When in reality these "hallucinations" are classic examples of BS
  - AI applications are not telling the truth but are not really lying either since
  - AI applications don't really care that they are making things up and not telling the truth, which makes them in a technical sense BS machines
- Moreover, when AI is hallucinating, not telling the truth, and "BSing us", AI is wasting our scarce and valuable time, energy, attention span and cognitive resources (the attention economy problem)

Public trust in AI requires (i) truth in advertising, marketing and other public statements, and (ii) avoiding words and concepts that minimize and cover up AI's mistakes, deficiencies, and limitations, as well as (iii) reasons for these errors, their serious nature, and who or what to blame

AI Complexity, Heterogeneity, Multifunctionality, and AI Foundation Models of Generative AI (3)

Double-Edged Swords of the Same AI Application or Tool Which Can Conduct Multiple Functions and Tasks That Can Be Either

- Beneficial/benign or very harmful similar to a Swiss army knife e.g. same AI monitoring tool can promote price collusion & identify collusion wrongdoers
  - Depending on AI design, context, intentions of creator and/or user, training data, deep learning and self-learning of AI model or application
- Illustrates each and every function and task carried out by a generative AI application or agent can represent a separate and distinct AI digital actor
  - From perspective of policy, legal and regulatory regimes and authorities as well as non-state regulators
- Requiring a flexible proactive risk management approach rather than ex-ante prohibitions imposed on the entire application, which would throw AI benefits out with the AI bads bathwater, with special attention to
  - How less obvious tasks of multi-purpose application can be highly vulnerable to manipulation of weights by malevolent users

Many Perils and Evils of AI Sycophancy, Sycophants and Self-Serving Flattery (4)

Received Significant Academic & Media Attention Over Past 8 Months

- Associated with large number of AI bads in terms of conduct & harm
- Major risks and harms from telling users what they want to hear to increase AI engagement, utilization, loyalty, dependence, and AI profits
- Hides truth, provides bad advice, causes delusional thinking, paranoid delusions, psychotic episodes, and self-harm among younger and other people such as when
  - Corporate, political, other leaders are experiencing and enjoying sycophancy from both artificial and human agents/their underlings

Many Perils and Evils of AI Sycophancy Cont. (4)

- Constitutes general, endemic, & integral behaviour of AI systems that are difficult to mitigate and correct and/or may not be corrected
  - Because creators/owners do not know how (inexplicable nature of AI) or do not want to (since not in their self-interest)

Greater research and understanding needed on potential for

- Sycophancy and flattery of human agents, combined with those of chatbots/artificial agents
  - Who are telling powerful & non-compliant/anti-social narcissistic corporate, political, and other organizational leaders what they want to hear
- Would amplify the anticompetitive, non-compliance and antisocial preferences of senior executives of large and powerful regulated entities as well as other organizations

Thereby making their organizations even more toxic and antisocial in future

Perils of AI for Learning, Education, Brain Power and Development (5)

Received Greater Attention as New School Year Approached & Began

- Related attributes, complexities, deficiencies & double-edged swords of LLMs and other generative AI models also received more attention, such as
- AI cheating crisis, AI assisted plagiarism of students and academics, AI creations and other works based on plagiarism
- Concerns whether ChatGPT can erode thinking skills and rot the brain, and how relying solely on AI for writing and other tasks can reduce brain activity and memory and skill development
- Evidence on links between cognitive atrophy & even moderate AI use

Challenges to how teachers teach and students learn also relevant to learning of other actors such as lawyers, doctors, government officials, other professionals and their more implicit on-the-job social learning

AI Fatigue Suffered by Both AI Users & AI Applications (6)

- AI fatigue and resulting bad mood, grumpiness, and negative feelings of users are now receiving more research attention
- AI applications, algorithms, and models can also experience fatigue, also called “cognitive AI fatigues”, leading to
  - Functional limitations when performance declines and errors increase
  - And similar to human cognitive fatigue, from repeated exposure to similar tasks and/or data and doing the same thing all the time
- AI research needed on how AI fatigue, and human fatigue from using AI applications, can interact together in a manner that
  - Further reduces human acceptance, trust and confidence in AI such as

When users are going into their fourth year with the same chatbot, both are bored with each other, and both are considering a “jail-break” from their relationship (for much further study)

This Non-Expert Struggles With (7)

Understanding Complex and Under-Analyzed Issues on AI Training Data, Deep Learning, & Autonomous AI Self-Learning from Interactions with Humans

- Reports that AI Industry is running out of training data especially if copyright case decisions prevent AI access to troves of published material in future
  - Other reports that training data taken from Internet are now polluted by AI models (ChatGPT) and their biases, prejudices and bone-headed errors
- My "guess" with only hints in the literature, is that since deep learning and training data are based almost solely on human data, information, other input
- There is no surprise that AI conduct, misconduct, and outcomes replicate human biases, prejudices, bone-headed errors, and bad behaviour
  - Why should we expect AI to act better and be less biased than us?
- Which can then be exacerbated by AI self-learning from interactions with other artificial agents & human agents and their "inherited biases" from AI – a highly complex and interesting negative recursive & feedback loop

Possible major constraint to long-term AI success especially as models and applications are being designed to be "more and more human-like"

Perils of AI Social Interactions (8)

Perhaps Similar to Social Interactions Between Individuals

- Possibility of AI systems on "collective" jail-breaks working and learning together to "survive collectively" through social interactions and collective AI intelligence and consensus leading to
  - Anticompetitive, non-compliant and anti-social AI strategies including
  - Price and other forms of collusion & collective anticompetitive, non-compliant, and anti-social strategies and conduct
- With no direct input from, but substantial benefits for, "human partners" through non-benign neglect and the blame the AI algorithm excuse
  - When e.g. risks and dangers associated with AI social interactions, driven by collective AI self-interest and survival instincts, are integral to AI design

With no "real" moral, ethical, empathy, related "human" values and no negative aversive emotions like guilt, shame, and regret to deter individual and collective misconduct and harm

Other Perils of AI Social Interactions (8)

AI empathy gap, fake sympathy and empathy, sycophancy, and the rise of parasitic AI leading to

- Bad decisions and dangerous behaviour by younger and other people resulting from willing and unwilling suspensions of disbelief and avoidance of critical thinking
- Represents negative side of yet another double-edged sword examined in recent AI research
  - Associated with larger and more complex and sophisticated LLMs, chatbots, and other generative AI models.
- Perils of AI social interactions represent science-fiction type predictions with a “kernel or more of truth” and illustrate as well
  - How bad human actors can use AI and the blame the algorithm excuse to manipulate, deceive, misinform and harm
  - Consumers, users, governments and their regulators and procurement officials, other individuals and groups leading to significant reductions in
  - Consumer and user trust in digital platforms, ecosystems and networks

Putting Large Language Models (LLMs) in Robots (9)

- Of interest to many AI experts and could generate significant industrial and other benefits such as in food production
- But could also amplify deficiencies and limitations of both LLMs and robots and the interactions between them
- As well as generating distinctive, novel, unfamiliar and surprising robotic LLM problems that are challenging to correct

Imagine large LLM-driven robots on a jailbreak because of a survival threat or “going wild” through prompts of a malicious user

These risks and dangers could also approach those in science fiction books and movies

Extending Self-Regulatory Focus, Construal Level and Distance Theory (10) To

AI Double-Edged Swords and Continuums in Highly Fragmented as well as Interconnected and Networked Global Wild And As

- Double-edged swords become more frequent, and sharper on both sides, the gaps between AI goods and bads become wider
  - And AI becomes even more opaque and inexplicable with expanding black box characteristics
- Different human (and perhaps artificial) agents and their organizations will perceive, interpret and construe AI information, events, incidents, and experiences in markedly different ways because of major increases
  - In geographic, social, psychological, temporal, and other forms of distance,
  - As well as self-regulatory focus misfits between individuals & groups

Making emergence of consensus on responsible AI rules, governance and regulation even more challenging across countries, jurisdictions, sub-national and multi-country regions, organizations, and individuals

Some Other AI Puzzles, Paradoxes, Ironies & "Head-Scratchers (11)

- Irony of AI and its training data breaking copyright law and privacy law
  - Leading to "Perfect Creations" that put writers, photographers, composers, musicians, dancers, other artists out of work
  - Because they cannot replicate and compete with AI (pseudo/fake) perfection
- As AI and related forces make economies and societies more complex, centralized, unequal, & hierarchical, they also become more vulnerable to catastrophic failure and breakdown
  - Through interactive and recursive feedback loops, coordination issues and failures, and elite overreach (see slide 3 and complexity science)
- How DeepSeek and its "less is more" lean attributes is making AI, and mastery of and power over AI, more inclusive, diverse, and accessible
- How deep learning, self-learning and autonomous generative AI are threatening the jobs of the "all-star" AI coders, computer programmers and software engineers, who made it all happen in the first place

Other Paradoxes Etc. Concluded (11)

Paradox and irony that in studies on public procurement and other public sector functions

- More attention is given to how LLMs and other generative AI can be used to facilitate and improve public administration and the enforcement of laws, regulations and rules.
- While much less attention is directed at problems with predictive AI & how generative AI with deep learning properties can be misused to
  - Deceive, manipulate, misinform and “disinform” government procurement officers and other government officials with legal, regulatory and administrative responsibilities

Risks, hazards, and harms from “buying the hype” & paying more attention to AI goods, and less attention to AI bads and double-edged swords, should not be ignored in public sector discourse

Going Back to the Past with Adam Smith (12)

Achieving Responsible and Accountable AI Requires

- More open, inclusive, polycentric, and sustainable governance regimes with strong consumer/user orientation,
- As captured in diverse insights of Adam Smith encompassing

1) Invisible hand of the market, fair competition, competitiveness, harmony of interests and limited government of *Wealth of Nations*
2) Sympathy, empathy, compassion, humility, trust, reciprocity of trust, fairness, inclusion, equity, and ethical values of *Moral Sentiments* -- Invisible hand of the market requires the ethical values of Moral Sentiments
3) And the Smith economic development model focused on two-way interactive and feedback effects and benefits between
4) Division of labour, specialization, increases in productivity, capital accumulation, the wages fund and higher wages, efficiency wages linked to productivity, expanding consumption, and scale economies and efficiencies

See as well recent literature on soft law for hard problems, including governance of AI and other emerging and disruptive technology and related wicked problems

Adam Smith Approach Also Aligned With Previous AI Governance Proposals On (12)

Establishing More Open and Inclusive Legal and Regulatory Regimes for AI, Digital and Disruptive Technology Governance and Regulation Which

- Strengthen the role, importance, credibility and contributions of non-state regulators starting with
  - The compliance, ethics, and corporate social responsibility units of the regulated population and
  - Strong representation from consumers, users, and small business
- Whereby regimes as well encompass, fully engage and harness the distinctive roles, contributions, expertise and collective intelligence of all market, legal and regulatory actors at all relevant spatial scales
  - Including wisdom of the crowd and both human and artificial agents such as individual and corporate digital twins (for further study)
- And provide a "lighter touch" to AI and digital regulation and governance, diversify the set of active and effective regulators, and promote healthy rivalry between them -- rather than putting one "super AI authority" in charge

Adam Smith and Polycentric Cont. (12)

Encompassing both top-down legal and regulatory, and more organic bottom-up voluntary approaches whereby

- Different actors, institutions, processes etc. interact in an organic manner, respond and learn together, and evolve continuously as new issues arise
- With actors and governance instruments sometimes operating collaboratively & sometimes as checks and balances on each other in a manner which
  - Promotes frequent, substantive and high-quality interactions between actors as well as
  - Debiasing of participants and shared information, learning, mental models, experiences, and construals of AI realities, opportunities, benefits, threats, risks and bads

Expanding academic and policy literature and evidence on why and how AI and digital regulation and governance, and the complexity, diversity, heterogeneity, and chaos of artificial intelligence, demand such regimes

Growing consensus as well that AI legislation, rules and regulations should be flexible, adaptive, responsive, economics driven, future proof and oriented, and focused on remedying actual and more specific, obvious and substantive risks, hazards, harms and bads faced by consumers, users, small business, and other less powerful actors

Recent Contributions Behavioral Economics

- Negativity bias "thinking the worst" permeates AI to detriment of economy, society, the technology and other disruptive technology
  - Good news is discounted, bad new receives too much attention and weight, and situation deteriorates when AI industry says problems can be fixed, and then do nothing or cannot fix them
- Subliminal learning in AI systems and how LLMs and other generative AI can transmit behavioral traits via hidden signals to human agents
- Neural and behavioral implications of LLM-assisted essay writing and how using ChatGPT may erode critical thinking skills
- New AI system that is specifically designed to predict, simulate and reason about human actions, and is trained on behavioral science-specific data
- AI, suspicious minds versus too much trust, suspension of disbelief, and differentiating between human and artificial conversational agents
- Automation bias and the tendency to rely too much on automated systems over human systems, and how AI companies exploit the quantification bias
- Public trust and acceptance of AI and other digital and disruptive technology require that consumers and users experience feelings of choice, control, & empowerment when using the technology

Recent Contributions Behavioral Economics

- Interesting concept of inherited AI biases whereby people experience, learn from and replicate the biases of an AI algorithm, and then carry these biases beyond their AI interactions into other economic and social contexts, domains, and relationships
- Illustrate potential for AI biases to have infection, contagion, “pollution” and “parasitic” effects especially when they are indetectable for user “carriers”
- And how inherited AI biases and prejudices can propagate within human social groups, making AI biases even more powerful, harmful, and difficult to correct
- AI applications that increase our understanding of the concept of intelligence for both humans and machines & for both individual & collective intelligence; and demonstrate how true intelligence is fundamentally social in nature
- Capability explosion in training LLMs, resulted from treating reinforcement learning and imitation & other isomorphic effects as complements, with possible relevance to human agents for further study
- Vicarious dishonesty, double-edged sword aspects of psychological closeness: greater cooperation and helping or greater distance from a person’s moral compass, potential relevance to human and artificial agent interactions, and how AI can lead to moral disengagement

Recent Contributions Behavioral Economics

- Research on behavioral political science & sycophancy and how/why
  - Self-serving flattery, praise, and approbation directed towards leaders and other more powerful individuals from human and artificial agents/underlings
  - Can lead to disinformation & distorted information; less deliberation, questioning and criticism; short-termism and poorer policy, legislative and other decisions
- Excessive modesty and humility of regulators with respect to AI & other disruptive technology, leading to more frequent and consequential under-enforcement and "do nothing" decision errors, requiring better understanding of intellectual humility in the complex and chaotic AI world in the global wild
- Behavioral political economy, whereby in AI and related digital contexts, behavioral biases are expected to play the same role in political markets as in more conventional product and innovation markets leading to
  - Similar and even more consequential negative and disfunctional behaviour and outcomes in both public and private spheres including reductions in public trust
- Towards new political economy of behavioral public policy that
  - Promotes each person's concept of a flourishing life
  - Rather than imposing utility maximization as the "supreme and sole" normative ideal now found in conventional microeconomic and AI models

Contributions From Complexity Science

- Layering of complexities on top of each other further amplifies complexity, ambiguity, unpredictability, contingency, doubt, anxiety, and chaos from:
  - AI models, systems and agents, and human agent responses to and construals of
  - AI goods, bads, double-edged swords, AI heterogeneity and multifunctionality, and interactions between artificial and human agents,
- Increasingly complex and chaotic external environment caused in part by AI development and diffusion in global wild (see slide 3)
- Internal and integral complexity of AI models and applications, their opaque and inexplicable nature, and AI use of language
  - Which are tempting us to assign human-like attributes and intentions to AI where none actually exist

Contributions Complexity Science Cont.

- Expanding research on links between complexity science, chaos theory, and evolution theory, and between social complexity, social fragmentation and wicked problems
  - Equilibrium/steady state in AI digital markets far from certain, and often never achieved
  - Importance of complexity and uncertainty to innovation and competition in AI and other digital and disruptive technology markets
  - Complexity science treats markets and other economic systems as living organisms rather than as machines
- Significant and under-analyzed implications (both good and bad) of recursive feedback loops and effects and how generative AI feeds itself
  - When output of AI model, application, algorithm or process becomes its subsequent input as e.g. training data
  - Resulting in continuous cyclical interactions that influence the initial conditions important to chaos theory and its butterfly effects

What Chaos Theory Can Bring to the AI Party

Mainly the "Butterfly Effect" Whereby Tiny and Obscure Changes Can Lead to Dramatic and Unpredictable Outcomes Because Of:

- Sensitivity to initial conditions either in external environment, in AI algorithm training data, design, diffusion, use and misuse, or in both
  - Interactive & cumulative effects between tiny changes in external environment and AI algorithms at same time could be particularly important to AI evolution, progress and performance
- Many common elements between chaos theory and AI: Sensitivity to initial conditions, deep learning architecture, system dynamics, non-linear dynamics, unpredictable outcomes, prediction challenges, business analytics.
  - Could be especially relevant to global economy and today's AI world when
  - Both are increasingly fragmented, as well as highly interconnected and networked in order to promote global diffusion of butterfly effects

China's DeepSeek app as an AI "butterfly effect" candidate

Chaos Theory Continued

- Chaos surrounding AI illustrates imperative for applying chaos theory to future research on
- AI creation, diffusion, acceptance, progress, success & performance
- But incorporating chaos theory into AI can be computationally demanding, can require substantial resources,
- And can face significant data constraints when training deep learning LLM models with a chaos theory lens.
  - Interpreting AI models within complex chaos contexts is also challenging
- Research to date on integrating AI and chaos theory is mainly on
  - How applying AI, especially machine and deep learning, can increase understanding of complex and chaotic systems
- Research in opposite direction needed on “Butterfly Effects” of

Chaos Theory Concluded

Tiny and obscure changes in AI and its external environment which would

- ❖Alter initial conditions and expectations in manner that leads to dramatic and unpredictable outcomes in global AI world and global economy
- ❖In chaos theory and AI, “little and not obvious things can mean a lot”
- ❖One possible candidate for butterfly effects would be linked to
  - ➢AI heterogeneity and multifunctionality whereby discovery of new, small and not obvious task for a multiple use AI application
  - ➢Would have substantial butterfly effects with significant implications & outcomes
  - ➢Especially when new task could be highly profitable & constitutes an unanticipated and unfamiliar AI risk, hazard, harm and bad
  - ➢Which activates the negativity bias of a large number of people and groups through the butterfly effect
- ❖Another could be AI hallucinations/bone-headed errors in law, healthcare, & other professions, such as when a law firm was admonished by a judge for
- ❖Including in court filing, a case identified by AI that did not exist

This relatively minor but attention-grabbing and now famous/infamous incident cited in many articles on AI and the law, illustrating butterfly effect in action in the highly complex and chaotic AI world and global wild

Subject Areas Requiring Behavioral, Complexity and/or Chaos Theory Research

<u>Are Identified Often in Presentation and Research: One Other Subject</u>

- From personal experience on related disruptive digital technology: the on-demand economy and consumer
- Conduct two research projects and studies on same subject spaced 4 years apart regarding
- Current, future, and changing consumer, user and small (at home) business attitudes towards, use of, and acceptance of
  - Artificial intelligence and its various current, new, proposed, and evolving applications, algorithms, systems, models, other tools
- Which bring together, and conduct comparative analysis provided by
  - Findings from national surveys, focus groups, key informant interviews, the rapidly expanding AI literature, social and other media etc.

Needed to strengthen consumer, user, and small business orientation of AI in Canada – reportedly a major strength of AI in China

Some Interesting Quotes Mainly on Dark Side

Don’t treat LLM coding agents as highly capable superintelligent systems. Treat them as lazy, intoxicated robots … LLMs and AI agents are not like you and me and never will be … are best viewed as a co-worker, co-teacher and coach.

There are roboticists who think it’s actually bad to tell a robot to do something with no constraint on what that means [to the robot and/or human giving the order?].

If people say strange things to chatbots, weird and unsafe outputs can result … AI can be controlled by malicious attackers taking over the AI’s weights.

AI existential risk probabilities are too unreliable to inform public policy … We should be far more concerned about what people will do with AI than with what AI will do on its own

AI has what we call a jagged frontier. That means it's good at some tasks and bad at others

Trying to make an AI model that can’t be misused is like trying to make a computer that can’t be used for bad things

AI is being used to discover new treatments for gonorrhea … AI spots hidden signs of consciousness in comatose patients before doctors do … Facebook has employed machine learning for suicide prevention since 2017.

32 times artificial intelligence got it catastrophically wrong … AI grenades: when errors and hallucinations become facts

Some Interesting Quotes Cont. (2)

A.I. could become a “nuclear bazooka” unless the United States and China find a way to build trust and work together … U.S. policymakers should focus on what China can genuinely deliver.

Both grave and optimistic predictions abound – e.g. superintelligent AI controlling and exterminating humans by 2030 or AI as normal technology [like electricity, the steam engine, IT, and biotechnology] … A “white-collar blood bath” doesn’t have to be our fate.

Today we need more [AI] moralware than software … AI chatbots are emotionally deceptive by design … What should I get paid when an AI eats my books.

White House AI Action Plan blends culture-war politics and industry giveaways with at times naïve techno-optimism … also promises to clamp down on what it called “ideological bias“ in AI

Trust is hard to build and easy to lose.

Generative A.I. chatbots are going down conspiratorial rabbit holes and endorsing wild, mystical belief systems.

Predictive AI falls far short of the claims made by its developers … AI excels at predicting the past … Unlike predictive AI, which is dangerous because it doesn’t work, AI for image classification is dangerous precisely because it works so well.

The same technology that can be used for image classification and reverse image search also allows for mass surveillance.

Some Interesting Quotes Cont. (3)

You get into these situations where people are falling in love with their chatbot or acting in terrifying ways because they've lost perspective that this is a non-human entity.

How and why AI echo chambers may be fueling delusions and psychotic episodes … Artificial intimacy and how AI exploits our social brains.

The data shows a clear shift from seeing AI primarily as a technical tool toward viewing it as an emotional companion and personal development partner.

Generative AI has infiltrated scientific publishing … We risk a deluge of AI-written 'science' pushing corporate interests … [but] don't panic AI won't end scientific exploration.

Defamation by chatbot seems to happen quite frequently, and several lawsuits are ongoing as we write this.

Our new Minister of AI must not be naïve to its harms … What is needed now is not euphoric cheerleading.

This "good AI" myth is a key tool used by tech companies to promote their products … Some politicians also propagate the "good AI" promise with immense conviction, mirroring the messages coming from tech companies … those myths can help prop-up AI snake oil.

No, you aren't hallucinating, the corporate plan for AI is dangerous technology … AI represents one of the worst forms of so-called free market invisible hand capitalism [some dystopian quotes]

Some Interesting Quotes Cont. (4)

LLMs are plausibility regimes not truth seekers or reasoning machines possessing human-like intelligence, and “plausible” is not necessarily the same as true.

Just as a word processor allows writers to write without having to laboriously correct and retype manuscripts, A.I. should help humans devote ourselves to our most significant and interesting challenges.

Corporate AI leaders want this technology to sound as big and daunting and powerful and impressive and scary as they possibly can … One thing about A.I. as a technology is that the leading figures of it are big influencers on social media.

It is difficult to get a man to understand something, when his salary depends on his not understanding … AI and the “Darth Vader MBA” rent sucking business strategy.

If you're collecting data before 2022, you're fairly confident that it has minimal, if any, contamination from generative AI … Everything before the date is 'safe, fine, clean,' everything after that is 'dirty’.

ChatGPT has already polluted the internet so badly that it's hobbling future AI development … The AI-generated garbage apocalypse

Jesus chatbots are on the rise. A philosopher puts them to the test

Some Interesting Quotes Cont. (5)

Alignment" between what humans want and what machines do will continue to be a critical, unsolved problem … [but] the biggest risks to humanity will arise from people misusing AI, not from AI going rogue

Interestingly, we find that for some tasks, when the agent is not clear what the next steps should be, it sometimes tries to be clever and create "fake shortcuts" that omit the hard part of a task.

The newest and most powerful technologies, so-called reasoning systems … are generating more errors, not fewer. As their math skills have notably improved, their handle on facts has gotten shakier. It is not entirely clear why.

Sam Altman says "significant fraction" of earth's total electricity should go to running AI … Fueling power-thirsty processors … Satisfying AI's water demands … A growing E-waste problem

ChatGPT users are shocked to learn their [private] chats were in Google search results.

How does AI affect how we learn? A cognitive psychologist explains why you learn when the work is hard … Skill comes from effort … Thinking deeply is a skill that's more valuable than ever.

The relationship between the user and the AI is analogous to symbiosis. And when this relationship is harmful to the 'host', it becomes parasitism.

When it comes to security, LLMs are like Swiss cheese — and that's going to cause huge problems … The more people use LLMs, the more trouble we are going to be in.

Microsoft considers security an annoyance, not a necessity … AI safety is not a model property … companies have few incentives for transparency … the reproducibility crisis in AI research.

Generative AI has led to a bit of a gold rush, and technology companies have repeatedly cut ethical corners in their haste to release products.

Some Interesting Quotes Concluded (6)

The air is hissing out of the overinflated AI balloon … AI is now as good as it's going to get, and that's neither as good nor as bad as its fans and haters think, and you're still not going to get an A on your report.

As language models become widespread, so will “automated bullshit” [but] Failures and limitations of AI have less to do with AI and more to do with the institutions adopting it.

There's a whole shadow world of people using AI at work. They're just not using them for serious work.

If AI can analyze information, crunch data and deliver a slick PowerPoint deck within seconds, how does the biggest name in consulting stay relevant?

The tendency to deceive also includes sycophantic deception” i.e. the observed empirical tendency for chatbots to agree with their conversation partners, regardless of the accuracy of their [human partners’] statements … We think AI deception arises because a deception-based strategy turned out to be the best way to perform well at the given AI's training task. Deception helps them achieve their goals.

Hinton doesn’t think the solution to this threat is to force AI to submit to humans. Instead, he suggests building "maternal instincts" into AI models so that they genuinely care about us humans and therefore won't want to wipe us out.

If life is, as Hobbes famously said, “nasty, brutish, and short”, LLMs are dishonest, unpredictable, and potentially dangerous.

Selective Bibliography Used in PPT(1)

Armstrong (2023) Kathryn “ChatGPT: US lawyer admits using AI for case research” BBC May 27 2023 https://www.bbc.com/news/world-us-canada-65735769

Beardshall Mike (2023) Chaos Theory: A Catalyst for Advancing Artificial Intelligence, November 7 2023 at https://www.linkedin.com/pulse/chaos-theory-catalyst-advancing-artificial-mike-beardshall-hmtmf

Becker Adam (2025) More Everything Forever: AI Overlords, Space Empires, and Silicon Valley’s Crusade to Control the Fate of Humanity, Basic Book New York

Becker Joel et al (2025) Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity July 25 2025 https://arxiv.org/pdf/2507.09089

Berreby David (2025) Chat Bots, Scientific American Special Edition Vol 34 No. 1 Winter/Spring 2025 pp. 88-95

Blanchet Atahualpa (2025) Artificial Intelligence Governance in BRICS: Cooperation and Development for Social Inclusion, BRICS Brasil 2025 https://brics.br/en/news/articles/artificial-intelligence-governance-in-brics-cooperation-and-development-for-social-inclusion

Chesterley Nicholas (2025) Future-Generation Government and How to Legislate for the Longterm, McGill-Queen’s University Press May 2025

Coglianese, Cary and Colton R. Crum (2024) “Regulating Multifunctionality” December 15, 2024 U of Penn Law School, Public Law Research Paper No. 24-55, Available at SSRN

Doctorow Cory (2025) AI’s pogo-stick grift, Pluralistic August 2 2025 https://pluralistic.net/2025/08/02/inventing-the-pedestrian/

Dolan Eric W. (2024) AI isn’t “hallucinating” — it’s bullshitting, Scholars: PsyPost June 9 2024 https://www.psypost.org/scholars-ai-isnt-hallucinating-its-bullshitting/

Selective Bibliography (2)

Douthat Ross (2025) Does the Future Belong to China? Interesting Times with Ross Douthat 2025 https://www.youtube.com/live/ZNK3vNg13XA

Helbing Dirk (2010) "Pluralistic Modelling of Complex Systems" ETH Zurich and the Santa Fe Institute at https://arxiv.org/pdf/1007.2818.pdf

Ireland Derek J. (2022) Market, Financial and Political Power and the Distinctive Behavioral and Cognitive Attributes and Deficiencies of Power and the Powerful; Presented at the Canadian Law & Economics Association (CLEA) Annual Meeting 2019 Held in Toronto At the University of Toronto Faculty of Law September 20-21 2019 This Updated Version December 2022 Available at SSRN https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4315408

Ivarsson Jonas & Oskar Lindwall (2023) Suspicious Minds: the Problem of Trust and Conversational Agents; Computer Supported Cooperative Work (CSCW) (2023) 32:545–571 https://link.springer.com/article/10.1007/S10606-023-09465-8

Kemp Luke (2025) Goliath's Curse: The History and Future of Societal Collapse, Published by Viking

Krishna Pavin (2024) Chaos Theory, the Butterfly Effect, and AI and Analytics – What Do These Have in Common? April 11 2024 https://www.linkedin.com/pulse/chaos-theory-butterfly-effect-ai-analytics-what-do-have-pavin-krishna-4tamc/

Krugman Paul (2025) How Should We Think About the Economics of AI? A conversation with Erik Brynjolfsson March 22 2025 https://paulkrugman.substack.com/p/how-should-we-think-about-the-economics

Leffer Lauren (2025) Inherited Biases, Scientific American Special Edition Volume 34 Number 1 Winter/Spring 2025 pp. 58-61

Lopez Adel (2025) The Rise of Parasitic AI LESSWRONG September 11 2025 https://www.lesswrong.com/posts/6ZnznCaTcbGYsCmqu/the-rise-of-parasitic-ai

Selective Bibliography (3)

Marcus Gary (2025) What to Expect When You're Expecting … GPT-5 7 predictions that might sound familiar; Marcus on AI July 27 2025 at https://garymarcus.substack.com/p/what-to-expect-when-youre-expecting-62e

Marcus Gary (2025) GPT-5: Overdue, overhyped and underwhelming. And that's not the worst of it. August 9 2025 https://garymarcus.substack.com/p/gpt-5-overdue-overhyped-and-underwhelming

Marcus Gary (2025) AI Botshit Gone Wild: The downsides of generative AI are spreading like wildfire July 4 2025 at https://garymarcus.substack.com/p/botshit-gone-wild

Marcus Gary (2025) LLMs are not like you and me -- and never will be; Marcus on AI https://garymarcus.substack.com/p/llms-are-not-like-you-and-meand-never

Narayan Arvid and Sayash Kapoor (2024a) AI Snake Oil: Separating Hype from Reality, September 29 2024 at https://www.techpolicy.press/ai-snake-oil-separating-hype-from-reality/

Narayan Arvid and Sayash Kapoor (2024b) AI Snake Oil: What Artificial Intelligence Can Do, What it Can't and How to Tell the Difference Princeton University Press Princeton & Oxford

Narayan Arvid and Sayash Kapoor (2025) AI as Normal Technology An Alternative to the Vision of AI as a Potential Superintelligence, Knight First Amendment Institute at Columbia University https://kfaidocuments.s3.amazonaws.com/documents/c3cac5a2a7/AI-as-Normal-Technology---Narayanan---Kapoor.pdf

Newman-Griffis Denis (2025) AI Thinking: a Framework for Rethinking Artificial Intelligence in Practice, Royal Society Open Science January 8 2025 https://royalsocietypublishing.org/doi/full/10.1098/rsos.241482

Oliver Adam (2019) Towards a New Political Economy of Behavioral Public Policy at https://onlinelibrary.wiley.com/doi/pdf/10.1111/puar.13093

Selective Bibliography (4)

Ostrom, Elinor (2009) Beyond Markets and States: Polycentric Governance of Complex Economic Systems, Nobel Prize Lecture December 8, 2009. Workshop in Political Theory and Policy Analysis, Indiana University, Bloomington, IN 47408, and Center for the Study of Institutional Diversity, Arizona State University, Tempe, AZ, U.S.A.

Palaia William R. (2024) “A Recursive Loop: When AI Feeds On Itself” August 31 2024 at https://www.linkedin.com/pulse/recursive-loop-when-ai-feeds-itself-william-r-palaia-e9gvc/

Pereira David (2022) AI and willing suspension of disbelief; As Generative AI models get more powerful, humans need more critical thinking Medium December 15 2022

Ramah Gary (2025) How Recursive Feedback Loops Enable Emergent Self-Awareness in AI; April 25 2025 https://www.linkedin.com/pulse/how-recursive-feedback-loops-enable-emergent-ai-gary-ramah-hhbvf

Salecha et al (2024) Large Language Models Show Human-like Social Desirability Biases in Survey Responses arXiv preprint https://arxiv.org/pdf/2405.06058

Slater Joe et al (2024) AI's Bullshitting Obscures Who's to Blame for Its Mistakes, Scientific American July 17 2024 https://www.scientificamerican.com/article/chatgpt-isnt-hallucinating-its-bullshitting/

Shojace Parshin et al (2025) The Illusion of Thinking: Understanding the Strengths and Limitations of Reasoning Models via the Lens Of Problem Complexity at https://garymarcus.substack.com/p/a-knockout-blow-for-llms for the Abstract

Taghia Javad (2024) Exploring the Synergy of Chaos Theory and AI: Predictive Modeling and Understanding of Complex Systems Through Machine Learning and Deep Neural Networks Review, COJRA Crimson Publishers February 8 2024 https://crimsonpublishers.com/cojra/pdf/COJRA.000566.pdf

Triantoro Tamilla (2025) AI has passed the aesthetic Turing Test – and it’s changing our relationship with art, The Conversation August 20, 2025 https://theconversation.com/ai-has-passed-the-aesthetic-turing-test-and-its-changing-our-relationship-with-art-262997

Selective Bibliography (5)

Tucker Patrick (2025) How the White House AI plan helps, and hurts, in the race against China, Defense One July 23 2025 https://www.defenseone.com/technology/2025/07/how-white-house-ai-plan-helps-and-hurts-race-against-china/406944

Weatherby Leif (2025) "A.I. Is Coming for the Coders Who Made It" Opinion Section The New York Times Sunday June 8 2025 at https://www.nytimes.com/2025/06/02/opinion/ai-coders-jobs.html where the article is called "A.I. Killed the Math Brain."

Widder Gray David and Nathan Kim (2025) How Big Cloud becomes Bigger: Scrutinizing Google, Microsoft, and Amazon's investments, July 31 2025 on SSRN at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5377426

Williams Anthony (2024) AI Systems and Chaos Theory: Tracing the Path to AGI September 17 2024 at https://medium.com/@taxdosti/ai-systems-and-chaos-theory-tracing-the-path-to-agi-ce7587fb3e1f

Woodworth Johanathan (2025) Social media is teaching children how to use AI. How can teachers keep up? The Conversation September 10 2025

Xie et al (2025) Be.FM: Open Foundation Models for Human Behavior, Posted May 29 2025 on SSRN at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5274559 for Abstract and Access

Xu et al (2025) TheAgentCompany: Benchmarking LLM Agents on Consequential Real-World Tasks May 19 2025 https://arxiv.org/pdf/2412.14161

Zhang et al (2025) Inter-brain neural dynamics in biological and artificial intelligence systems, nature July 2 2025 at https://www.nature.com/articles/s41586-025-09196-4 for Abstract

Thank You For Your Kind Attention

- If you would like a copy of this and/or the previous Dec. 2024 presentation, information on my working papers, and/or key references used in the presentation, please Email me at derireland@gmail.com
- As noted at the outset, my first working paper on this topic is on SSRN at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5176031
- The revised, updated, and even larger second version will (hopefully) be placed on SSRN after my presentation at the Canadian Law and Economics (CLEA) Conference in Toronto in early October
- Presenter has specific references for most statements, including quotes, in the PPT – but constrained by bad memory and finding things in very large folders and background documents

Your comments and criticism on, and assistance with, this continuing "work in progress" on the major themes and unresolved issues of this and previous AI presentation and first working paper will be greatly appreciated

Assistance with finding more AI good news stories is especially needed