Humanity's Greatest Existential Crisis

How We Learn to Love Our New AI Overlords

David L. Shrier

IT'S difficult to escape the headlines about artificial intelligence (AI) these days. The magnitude of chatter about AI is only exceeded by the scale of failure to understand its true impacts and current status.

AI technology has arrived at a speed which is bewildering for the typical government apparatus. Despite having been monitoring and preparing for AI for years, as evidenced by the arrival of the EU AI Act, Europe is (like every other domain on the planet) as yet unprepared for the deeper implications of AI. Globally, more than 80 governments are considering or are implementing some form of AI intervention. How can these disparate efforts be harmonized? Are they focused on the critical issues? Will they create a latter-day Tower of Babel, sowing regulatory confusion in a misguided effort to provide safety from AI harms?

The speed with which AI now is being adopted, after many years of slow progress, is one motivator for a greater sense of urgency from policy bodies. Before recommending a course of action for prudent AI policy, it is important to decode the trends taking shape.

Responding in an Era of Flash Growth

The term 'flash growth' is a new way to conceptualize the pace at which technologies can emerge and disrupt at society scale. The concept emerged out of conversations among Evercore ISI (the research arm of top-ranked investment bank Evercore Partners) and Imperial College Business School, during an effort to understand, dimensionalize, and forecast the impacts of AI on the global economy.

Marc Harris, the Director of Research for Evercore ISI, Julian Emanuel, its head of Equity, Derivatives and Quantitative Strategy, and myself were endeavoring to describe the lightning-fast adoption seen for generative AI. It was challenging to model mathematically, because it broke the conventions of historic technological adoption frameworks.

We saw something quite different with ChatGPT. The adoption "curve" was essentially a straight vertical line.

Normally, new technology is adopted in approximately a sigmoidal pattern. There was a excellent illustration of this in Malcolm Gladwell's breakthrough book *The Tipping Point: How Little Things Can Make a Big Difference* (2002). Every major technology we have seen of the past few hundred years have followed a similar adoption curve. Adoption rates have been getting faster and faster, but generally look like an elongated 'S.'

If we take this theory and map it to practice, we can see similar shapes playing out across everything from usage of television to the smartphone:

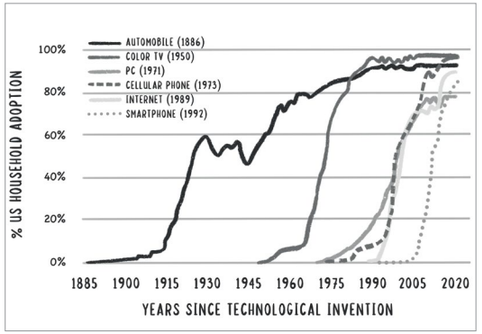
However, we saw something quite different with ChatGPT. The adoption "curve" was essentially a straight vertical line. ChatGPT recruited 100 million users within about six weeks. This kind of widespread and pervasive adoption caused seismic shocks in business, academic, and government circles.

This is not just a one off. This concept of 'flash growth' as a structurally different model was validated a few months later when Meta's Threads app achieved a user count of 100 million in five days.

The emerging theory is that, to quote Sir Isaac Newton, "if I have seen further, it is by standing on the shoulders of giants."

Over the past five decades, trillions of euros of investment have been made into telecommunications infrastructure, computer chip capabilities, and network communications software. What this translates to is that, on the one hand, we have very high-capacity networks which can rapidly transmit data, and on the other hand, we have high-performing computing that can analyze and use this data at speeds and scale never seen before. ChatGPT didn't simply arise *de novo* from Silicon Valley cleverness and magically appear on every desktop. It is part of a clear evolution we have seen of commercially viable AI technology, drawing a straight line from the early machine learning

David L. Shrier *is a Professor of Practice in AI and Innovation at Imperial College Business School in London and co-head of the Trusted AI Institute. He is also Managing Director of venture studio Visionary Future LLC. You may follow him on X @DavidShrier or connect with him on LinkedIn.*



Source: Evercore ISI Research

Figure 1: Adoption Curves of Select Technologies

companies of the 1980s to the generative AI revolution of today.

In parallel with the evolving software and hardware that makes AI computation possible, there likewise has been an evolution of the communications networks, data centers, and related technologies that host these AI systems.

Implications for Policy

The implications for policymakers and regulators are significant. Policy and regulation tend to be quite slow, with good reason: regulation made in haste is often not fit for purpose. This begs the question, how do governments respond to a world where the piece of change is several orders of magnitude faster than the typical pace of government intervention?

We have not seen positive results in tech disruption when government moves too quickly. The United States provides cautionary lessons for policymakers and regulators in other domiciles.

For example, the New York State BitLicense, which rather famously hindered fintech in the United States' financial capital for years, allowed centers of excellence in London (and Zurich, Berlin, and Paris) and Singapore to take global leadership.

More recently, we saw two different approaches to blockchain and cryptocurrency in the United States, in the absence of dedicated laws being passed, taken by the Commodities Futures Trading Commission (CFTC) and the Securities and Exchange Commission (the SEC) with respect to another new technology, blockchain.

The CFTC took a measured approach: getting to know the technology and the market, building capacity across an agency, engaging with the ecosystem and market participants. All of this takes time, and it does not typically lead to news headlines. It does lead to a greater fundamental understanding of the market, and an informed perspective for which a regulator can exercise its prudential authority. Regulators acquire insight into potential harms to consumers and how to mitigate them, an understanding of systemic risk, an appreciation of how to promote stability, and a pathway to support innovation. These parallel demands can be met with carefully crafted policy, rules, and interpretation.

In contrast, SEC chairman Gary Gensler has elected to legislate through enforcement. He has been generating numerous media headlines with a series of actions that create new precedent, and in the eyes of a number of individuals deviate from both his regulatory perimeter and the role of the regulator. It is suggested that there is even confusion within the walls of the SEC, because some of Gary's own colleagues disagree rather vigorously with the approach he has chosen to take.

One explanation that has been put forward is that Gensler is openly lobbying for Janet Yellen's job as Secretary of the Treasury. The shift in stance by the SEC under Gensler has been notable based on many conversations with rank-and-file SEC professionals, as well as other SEC commissioners, given their previously measured approach to the technology. Gensler appears bent on carving out a policy domain of one, even against the recommendations, it is understood, of his own experts within his own agency.

Clearly, this is not a scalable or repeatable approach for regulators to take when confronting new technologies.

> *ChatGPT didn't simply arise de novo from Silicon Valley cleverness and magically appear on every desktop. It is part of a clear evolution we have seen of commercially viable AI technology, drawing a straight line from the early machine learning companies of the 1980s to the generative AI revolution of today.*

The U.S. Courts have taken a jaundiced view of Gensler's actions, and he has succeeded in one regard: in a rare show of bipartisanship, the U.S. Senate and House have taken up legislation to clarify the rules around cryptocurrencies, a move that in part has been interpreted as a clear signal to rein in Gensler's overreach.

Europe's Enlightenment

On the other end of the spectrum from the confusion found in the United States, we have the European Parliament (Europarl). Europarl has spent years building sophisticated capabilities, as has the European Commission, with a broad-based consultative process that not only brings in the leading experts from Europe, but also reaches out to global thought leaders from other domains to provide a richly diversified perspective. It has applied this to blockchain and passed blockchain laws.

It is not a fast approach, but it is very thoughtful. And slow as it may be in the abstract, the EU has still proven to be a policy leader in a number of critical areas with regards to AI and related technologies.

Data privacy certainly garnered a great deal of attention and resulted in the EU's General Data Protection Regulation (GDPR), which then prompted California to pass the California Consumer Privacy Act (notable not only because of all the tech companies headquartered in California, but also because if separated from the United States and observed on its own, it would be the fifth largest economy in the world as measured by nominal GDP).

The same stepwise process is now under way for artificial intelligence.

The EU formed an expert advisory after years of study by both Parliament and Commission. Inputs and perspectives were offered to balance the competing priorities of protection, stability, risk, and innovation. Out of this emerged a risk-based approach. Where there is greatest danger, AI is banned entirely. Where there is a high degree of risk, such as jobs screening or law enforcement, a high degree of oversight is required. Where there is minimal risk, such as in video games, there is no real oversight. Then ChatGPT arrived over the winter of 2022-2023.

The EU's AI Act wasn't formulated with the effects and paradigm of generative AI in mind. Algorithmic explainability is all well and good, but probabilistic machine learning systems defy some of the earlier efforts to manage AI. For example, ChatGPT (or Google Bard, or Meta's generative system, or open-source Vicuna) deliberately randomize results. If you put in the same prompt, you get different answers. This is challenging for regulators accustomed to a more deterministic and predictable approach to computers (if I put 1+1 into a calculator, I always get an answer of 2, unless it is broken). ChatGPT isn't a calculator. It is, however, an incredibly powerful and useful tool that can relieve a tremendous amount of drudgery.

Every major technology we have seen of the past few hundred years have followed a similar adoption curve. Adoption rates have been getting faster and faster, but generally look like an elongated 'S.'

Let's ask Chat GPT, "What is the biggest risk Chat GPT represents to Europe?" Here is the response generated in August 2023:

As of my last knowledge update in September 2021, I don't have access to real-time information or events that have occurred after that date. However, based on the information available up to that point, I can provide some insights into potential risks that ChatGPT or similar AI models could represent to Europe or any other region. Keep in mind that these risks are not unique to Europe and could apply globally:

Misinformation and Manipulation: One of the significant risks associated with AI models like ChatGPT is the potential to spread misinformation and manipulate public opinion. Malicious actors could use AI-generated content to spread false information, incite unrest, or even interfere in political processes, which could have destabilizing effects on European societies.

Privacy Concerns: AI models often require large amounts of data to train effectively. The use of personal or sensitive data in training such models can raise concerns about privacy and data protection. If not properly managed, this could lead to violations of GDPR in Europe.

The EU's AI Act wasn't formulated with the effects and paradigm of generative AI in mind.

Job Displacement: The increased use of AI and automation technologies in various sectors could lead to job displacement for certain types of jobs, potentially causing economic and social challenges. This could require proactive measures to re-skill and retrain the workforce to adapt to the changing job landscape.

Bias and Fairness: AI models like ChatGPT can inadvertently inherit biases present in their training data. These biases can perpetuate inequalities and discrimination in automated decisionmaking processes, such as hiring, lending, and law enforcement, which can have significant societal implications.

Security Vulnerabilities: AI systems are not immune to attacks. Adversarial attacks, where specially crafted inputs can fool AI models, could be used to exploit vulnerabilities, or create unexpected behavior. This could potentially be exploited by malicious actors for various purposes.

Ethical and Accountability Challenges: AI models can sometimes generate content that is offensive, inappropriate, or harmful. Determining responsibility and accountability for such content can be complex, raising questions about the legal and ethical frameworks necessary to address these issues.

Dependency on Foreign AI Technology: If Europe becomes overly dependent on AI technologies developed outside the region, it could have implications for technological sovereignty and strategic autonomy. Efforts to support domestic AI research and development might be necessary to mitigate this risk.

Regulatory Challenges: The rapid advancement of AI technology often outpaces regulatory frameworks. Developing effective regulations that balance innovation with potential risks is a challenge that governments and regulatory bodies face.

It's important to note that many of these risks are not unique to AI models like ChatGPT but apply to the broader adoption of AI and automation technologies. Europe, like any other region, would need to address these challenges through a combination of technological development, policymaking, and international collaboration to ensure that the benefits of AI are maximized while the risks are minimized. For the most up-to-date information on this topic, I recommend consulting recent sources and experts in the field of AI ethics, policy, and technology.

Research from OpenAI, Open Research, and Wharton suggests 80 percent or more of jobs will feel some measure of impact from AI. Other work by EvercoreISI indicates nearly 100 percent of jobs in some way touched by AI, although in many cases perhaps not profoundly.

All of this is well and good. But what happens if we ask the exact same question again of ChatGPT? (omitting the disclaimer it auto-generates)

One of the potential risks associated with AI models like ChatGPT is the spread of misinformation or the amplification of existing biases. These risks could certainly extend to Europe, as they could to any other region.

The European Union has been proactive in addressing AI-related risks and has been working on regulations to ensure that AI systems are developed and deployed in a way that aligns with European values, ethics, and human rights. This includes efforts to address issues like bias, transparency, accountability, and potential societal impacts.

If ChatGPT or similar AI models were to spread misinformation or inadvertently reinforce biases, it could lead to public distrust in information sources, affect decisionmaking processes, and contribute to the polarization of society. This risk underscores the importance of responsible development, deployment, and oversight of AI systems to ensure they have a positive impact on society.

Fairly considerable dispersal in the answer, even if each provide a certain measure of insight. It's this exact kind of variability that fuels apoplexy in supervisory bodies. Regulators and policymakers will need to acquire new knowledge leading to new approaches and new tools for managing systems that operate in this manner.

AI and Global Competitiveness

Regulation, of course, is only part of the story. Artificial intelligence is a subject vital to national interest. If you make laws that are too restrictive around AI, you not only will discourage innovative businesses from coming to your country and creating new jobs and economic expansion, you also may inhibit the ability of your other industries to benefit from various kinds of AI adoption. In fact, according to the forecast developed by Visionary Future, large enterprises will be the chief beneficiaries of artificial intelligence, productivity, and revenue enhancements—much more so than tech startups. By 2032, Visionary Future LLC anticipates approximately €31 trillion of additional potential capitalization for large global enterprise, versus about €4.2 trillion for AI startups, based on an expected market spend of €3.5 trillion globally.

Tremendous value creation means tremendous potential for global prosperity if GDP growth can be harnessed appropriately. Endemic issues including migration, homelessness, longevity, and the environment can be ameliorated by the direct application of the AI dividend to humanity's biggest problems.

The Greater Good and Workforce of the Future

Will AI be one of the greater goods of society? How will ordinary citizens benefit? What are the risks to the average resident of a particular country?

The AI societal transformations described herein are predicated on a massive restructuring of the workforce. Research from OpenAI, Open Research, and Wharton suggests 80 percent or more of jobs will feel some measure of impact from AI. Other work by EvercoreISI indicates nearly 100 percent of jobs in some way touched by AI, although in many cases perhaps not profoundly.

When we look at how AI might be turned into operational reality, there is materially the potential for impact of 20 to 40 percent on any given business at a substantive versus a superficial level. To think of the difference: a person can compose an email better with the type-ahead features that Microsoft and Google have begun implementing in their email systems where it predicts and suggests the end of the sentence. Their job may be more profoundly impacted if they work in a contact center and 90 percent of their colleagues are replaced by AI chatbots.

Evercore examined 800+ occupations, looking at 52 skills and 41 activities, and mapped these to AI capabilities. They determined the exposure of each industry to AI disruption:

What's particularly notable about the Evercore analysis is that high salary jobs are particularly vulnerable to AI disruption in light of the latest technology developments. Substantial re-skilling is required. Although noted economist Erik Brynjolfsson stated that 'For every robot, there is a robot repair

person,' not everyone is capable of learning how to become a robot repair person or an AI model evaluator.

How can we guide AI research and implementation to serve humanity? How can we harness the AI dividend?

A SYSTEMS VIEW ON AI POLICY

What this leads us to is the need for an integrated systems view of engaging around AI. It doesn't matter if progressive legislators pass a particular new set of laws, if they have failed to fund building capacity in the civil service organizations, then responsible for deploying and enforcing those laws.

Artificial intelligence represents the greatest existential threat that humanity has faced, yet also holds within it the potential for humanity's salvation in the form of addressing fundamental issues from health to climate.

We only have to look at the confusion in the United States to see the dangers of waiting too long to legislate, and allowing opportunistic individuals to seek to use technology disruption issues as a campaign platform for their next job.

One comprehensive approach to technology policy and impact was modeled by the Commonwealth Fintech Toolkit, that was created in collaboration between Visionary Future and the Commonwealth Secretariat's (COMSEC) 53 nations, representing 2.4 billion lives, and spanning the globe.

A similar model could be useful for thinking about artificial intelligence.

A staged consultative approach is driven by the government and is married to capacity building, that spans elected officials, appointed officials in civil service and private sector, with engagement from academia to provide neutrality and perspective.

Horizon scanning can help identify both best and worst practice. The University of Cambridge's spinout Regulatory Genome helps automate this, using AI to help monitor AI developments.

Open-source repositories of information are created and published. These can include not only reference policy frameworks and case studies as with the Fintech Toolkit, but also open-source code repositories highlighting best practices and providing greater utility to promote widespread adoption.

Convenings serve to identify problem statements, align policy initiatives with areas of greatest need and impact, initiate solutions, scale impact, and disseminate findings. They also provide a foundation for community building that can provide momentum and support at a global level.

In numerous conversations with private sector experts, government officials, executives, entrepreneurs, and civil advocates, the need for coherency and resolution of landscape confusion became paramount.

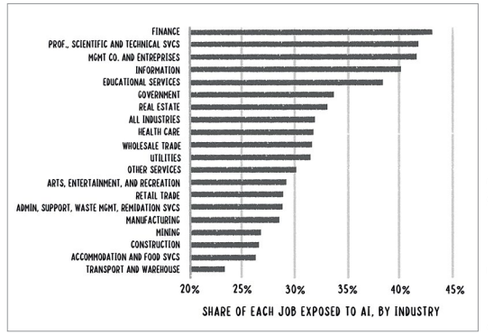
Later in 2023, the UK is hosting a global summit on artificial intelligence, following on the heels of the G7 announcement in May about the need for standards in Trustworthy AI and formulating the Hiroshima AI process to bring this about. Areas such as intellectual property, responsible use of AI, and governance will be addressed.

THE BIRTH OF TRUSTED AI

Imperial College London is leading a large-scale initiative to help convene and guide progress on responsible and trustworthy applications of AI in service of humanity with the formation of the Trusted AI Institute.

Initial university partners, led by Imperial College London, include the University of Oxford, the Massachusetts Institute of Technology, and the University of Edinburgh and represent more than 3,000 of the top AI researchers in the world. The Trusted AI Institute is also engaging with government agencies, multigovernmental bodies like the OECD and the World Economic Forum, and private citizen advocacy groups.

The Trusted AI Institute have just received its first pilot commitment for a country-scale initiative that bridges private sector, public sector, and academia in providing better assurance, inclusion, and cyber security around artificial intelligence implementations.



Source: Evercore ISI Research

Figure 2: Exposure to AI by Industry

BOTH PROBLEM AND SOLUTION

Artificial intelligence represents the greatest existential threat that humanity has faced, yet also holds within it the potential for humanity's salvation in the form of addressing fundamental issues from health to climate. The Trusted AI Institute, and similar initiatives such as those championed by the OECD and the G7, as well as legislative interventions such as the EU AI Act, will help bring forward an enlightened path to AI adoption.

The question is not, should we adopt AI, or should we stop using it. The question is, how can we shape AI so that this innovative technology, created by humanity, helps serve the greatest good of humanity.