

PROACTIVE MANAGEMENT OF CREATIVE DESTRUCTION: TRANSPARENCY AS A FOUNDATIONAL ELEMENT

AI is very much still in its infancy. The path to beneficial and true development lies with transparency and regulation for both governments and companies. There is a wide-spread mistrust of “artificial intelligence” as a technology, and the companies that are developing it, across many stakeholders. There is a similar mistrust of governments’ roles in or ability to regulate tech companies or to govern their own use of these same technologies. Regulation, even though is the key, should be implemented rather carefully so that development is not hindered but is steered in the right direction.

Stacey H. King*



* Stacey H. King, Visiting Policy Fellow – Oxford Internet Institute.

When entering into any kind of multi-party business relationship, transparency among all parties is critical to the success of the outcome. For example, if you want to build a house, making sure the architect, engineer, contractor, city planner, landscape designers, and builders are all onboard with your anticipated needs and future potential use of the house is necessary. Without such transparency, foundations can fail, walls can crack, plumbing can leak, extensions are rejected, and, in some cases, neighbors may complain and prevent plans from moving forward. The same need for transparency and honest dealings is true for the regulation of artificial intelligence (AI).

AI is very much still in its infancy. The AI being built today – narrow AI (ANI), or AI that is taught to manage a particular task – is the foundational building block for the AI of tomorrow – artificial general intelligence (AGI), or AI that is able to replicate broader human thought processes across multiple tasks. Exactly how both narrow and general AI will impact global economies and organizational structures in the short and long-term is still unknown,¹ but recent developments provide enough insight to know that AI will bring not just a gale, but a tsunami of Schumpeter's creative destruction.² Indeed, we may currently be standing at the shore of Schumpeter's gale, watching the water pull away toward sea.

Now is the critical time, while this technology is still in its infancy, for all stakeholders – governments, industry, academia, and civil society – to be more forthcoming and transparent about the knowns and unknowns of AI and its potential impact, to collaborate on ways to support responsible development and use of the same, and to prepare for the inevitable changes to come. But to do so, all stakeholders must embrace a requisite level of transparency, otherwise the walls of this house will inevitably bend under pressure or fail to get off the ground.

Creative Destruction

Just as businesses must organize around and execute on strategies to manage change and protect their relevance and position accordingly,³ so too should governments and society. AI differs from previous disruptive innovations to both market and non-market approaches. Automation itself is nothing new. Humans have automated tasks for centuries and, while initially disruptive, have modified or introduced new ways of organizing our societies accordingly.⁴ More traditional forms of automation sought

¹ Nicholas Crafts, "Artificial Intelligence as a General-purpose Technology: An Historical Perspective," *Oxford Review of Economic Policy*, Vol. 37, No. 3 (2021): p. 521-36.

² J. A. Schumpeter, (1943), *Capitalism, Socialism, and Democracy* (London: Routledge Classics, 2010 ed.).

³ C. O'Reilly and M. Tushman, "Lead and Disrupt: How to Solve the Innovator's Dilemma," (Stanford: Stanford Business Books, 2016).

⁴ C. B. Frey, & M. A. Osborne, "The Future of Employment: How Susceptible Are Jobs to Computerisation?," *Technological Forecasting & Social Change*, (2017): p. 254-280.

to replicate “physical” processes (reproducing works, shipping products, drilling holes in steel, amplifying your voice, or searching through books in a library). And in response society amended or developed structures, regulations, industry standards, and ethical norms for managing the automation of physical processes.

“Artificial intelligence as used today is adding to this mix in ways that are not always perceived as beneficial - from surveillance to deep fakes to harmful bias and discrimination.”

Artificial intelligence, however, is a new type of automation with an important distinction. AI does not seek to replace another human physical process, per se, but rather to replicate part of the human “thought” process or the decision-making behind the doing of something.⁵ This is a fundamentally different form of automation – it challenges the very notion of what humans believe distinguishes them as a species – and our societies do not have as many mechanisms in place for managing the fallout from this type of technology. As a result, transparency around the foundational parts of this technology (ANI) and how it is developing (AGI) by those closest to it – the architects and engineers of this house – is more critical than ever.

The market for AI in relation to widespread use across industries, products, and services is still early.⁶ Nevertheless, wider concerns about the potential impact of AI have stretched from claims that machines will replace all jobs and, eventually, humans,⁷ to claims that machines will replace jobs but lead to an era of creativity and freedom.⁸ In both cases the supposed outcome is based on conjecture and myth, but they are important myths to recognize as they define the fears and hopes that many stakeholders bring to the table and influence how we approach the development and regulation of AI as a society.

⁵“The capacity of computers or other machines to exhibit or simulate intelligent behaviour,” IEEE, *Ethically Aligned Design: First Edition Glossary*, p. 8.

https://standards.ieee.org/wp-content/uploads/import/documents/other/ead1e_glossary.pdf [visited Feb. 29, 2022].

⁶ D. Nathan and N. Ahmed, “Technological Change and Employment: Creative Destruction,” *The Indian Journal of Labour Economics*, (2018): p. 281-298.

⁷ P. Holley, “Elon Musk’s nightmarish warning: AI could become ‘an immortal dictator from which we would never escape,’” 6 April 2018, Retrieved from *The Washington Post*: https://www.washingtonpost.com/news/innovations/wp/2018/04/06/elon-musks-nightmarish-warning-ai-could-become-an-immortal-dictator-from-which-we-would-never-escape/?utm_term=.83898a54f99f

⁸ S. V. Osborn, “The End of Meaningless Jobs Will Unleash the World’s Creativity,” Retrieved from SingularityHub, 23 August 2016: <https://singularityhub.com/2016/08/23/the-end-of-meaningless-jobs-will-unleash-the-worlds-creativity/#sm.0001enmry1fb5dfsxfw17285ra9ll>

AI will have a wide array of impacts that we cannot predict.⁹ We do know from previous general-purpose technologies (steam, railroad, electricity, Internet) that societies and economies are rarely prepared to deal with the wide-ranging impacts such technologies bring.¹⁰ The fallout does not impact one industry or market or regulation.¹¹ It is this wide-scale impact that we, as a society, need to address now. If we know AI is going to have such broad implications, how do we prepare for the negative impacts of a radical shift in the status quo, while simultaneously encouraging innovation and market growth? And why is transparency critical to the same?

State of Play Today

Our global society is already in a period of upheaval, in part arguably a long-tail impact of the introduction of the Internet.¹² The “technology revolution” introduced a myriad of opportunities, economic advantages, and the ability to connect on a global basis, but it was also disruptive to existing societal systems and governance, led to incursions on individuals’ private lives, and has amplified the reach of those who do not have societies’ best interests at heart. The latter has led to a multitude of finger pointing and, in many cases, the refusal of companies to participate transparently in the debate over unintended consequences and how to remedy them.

Artificial intelligence as used today is adding to this mix in ways that are not always perceived as beneficial - from surveillance to deep fakes to harmful bias and discrimination.¹³ In addition, the division of wealth (personal, corporate, and national) is becoming more extreme, market skills are changing, and industrial society norms are being questioned. Finally, communities are becoming more and more polarized, and the environment has become a critical threat to livelihood and economies. Individuals feel less in control of what is happening around them, the infrastructure that is supposed to support them, security for the future, and more and more mistrust of technology itself. And in the midst of all of this, we are asking

⁹ P. Aghion, U. Akcigit and P. Howitt, “What Do We Learn From Schumpeterian Growth Theory?,” Working Paper 18824. Cambridge: National Bureau of Economic Research (2013).

¹⁰ C. Schubert, “How to Evaluate Creative Destruction: Reconstructing Schumpeter’s Approach,” *Cambridge Journal of Economics*, (2013): p. 227-250.

¹¹ P. Aghion, U. Akcigit and P. Howitt, (2013).

¹² Rostam J. Neuwirth, “The ‘letter’ and the ‘spirit’ of Comparative Law in the Time of ‘artificial Intelligence’ and Other Oxymora,” *Canterbury Law Review*, Vol. 26, (November 2020): p. 1-31.

¹³ E.g., “AI is Sending People to Jail – and Getting It Wrong.” *MIT Technology Review*, 21 January 2019, <https://www.technologyreview.com/2019/01/21/137783/algorithms-criminal-justice-ai/> [visited on 18 February 2022]; “This is What a Deepfake Voice Clone Used in a Failed Fraud Attempt Sounds Like.” *The Verge*, 27 July 2020, <https://www.theverge.com/2020/7/27/21339898/deepfake-audio-voice-clone-scam-attempt-nisos> [visited on 18 February 2022]; “A.I. Bias Caused 80 percent of Black Mortgage Applicants to Be Denied.” *Forbes*, 2 September 2021. <https://www.forbes.com/sites/korihale/2021/09/02/ai-bias-caused-80-of-black-mortgage-applicants-to-be-denied/?sh=2f9b093136fe> [visited on 18 February 2022]. “Exploiting AI: How Cybercriminals Misuse and Abuse AI and ML.” 19 November 2020. <https://www.trendmicro.com/vinfo/us/security/news/cybercrime-and-digital-threats/exploiting-ai-how-cyber-criminals-misuse-abuse-ai-and-ml> [visited 19 February 2022].

our larger societies to trust the companies building systems to eventually replicate human thought process and make decisions, even though little is known about the foundations on which these systems are built and who is training them to make the decision.¹⁴ To develop global trust that these systems are being built for human good, and not just to drive revenue or reinforce a government's power, relevant stakeholders need to be more transparent and provide safety-nets and compliance mechanisms for managing negative impacts. Transparency does not mean that stakeholders need to disclose every business plan or trade secret. Transparency does require stakeholders to come to the table in an open and honest manner, and to negotiate the issues that arise accordingly.

“As algorithms trained by different developers to manage specific tasks become integrated into larger systems, and as these systems become more independent from human intervention in decision-making, questions around jurisdictional oversight will become of greater concern.”

Why are these factors we face today so important to understand in creating a strategy for the promotion of cooperative outcomes in governance for the future? Because it is precisely these issues that lay under the surface of every stakeholder critical to the systems that help us navigate through the changes to come.

Mistrust of AI, Tech Companies, and Government

There is a wide-spread mistrust of “artificial intelligence” as a technology, and the companies that are developing it, across many stakeholders.¹⁵ At the same time, most stakeholders understand that AI is not going to disappear nor is it something we necessarily want to stop. What impacts stakeholder trust is the relatively recent, positive, and hopeful attitude toward technology that was found in the late 20th and early 21st Century, but which has been replaced with a general mistrust or concern around the same today. This leads to a perception that tech companies are concerned more about profit over people and society, are secretive and therefore suspect, and use “self-regulation” as a means for anti-competitive or anti-consumer protection behavior.¹⁶

¹⁴ Jenny Bunn, “Working in Contexts for Which Transparency Is Important: A Recordkeeping View of Explainable Artificial Intelligence (XAI),” *Records Management Journal*, Vol. 30, No. 2 (2020): p. 143-153.

¹⁵ “Trust in Artificial Intelligence: a Five Country Study,” KPMG/University of Queensland, Australia (March 2021). <https://assets.kpmg/content/dam/kpmg/au/pdf/2021/trust-in-ai-multiple-countries.pdf> [visited on 9 February 2022].

¹⁶ KPMG/University of Queensland, Australia (March 2021).

There is a similar mistrust of governments' roles in or ability to regulate tech companies or to govern their own use of these same technologies. On one hand, many are concerned that industry has captured regulators through lobbyists and market power. On the other, there is concern that governments will misuse the technology to exercise control over its citizens or to hold other nations hostage. Governments, like tech companies and industry, need to put covert nationalistic interests to the side and be more transparent about the pressures they face, collaborate on the larger societal needs, and create foundational structures of governance for all.¹⁷

National Strategies Focused on AI Fears

A result of this lack of trust is the development of national regulatory and economic strategies that often reflect fear mitigation as opposed to a holistic and informed approach to how we drive a beneficial development, deployment, and use of AI. In many cases, regulatory approaches are based on short-term thinking around a nascent technology. Overly strict regulations will limit our ability to develop this technology to its full potential or may result in nationalistic or inappropriate use or access by government. Adopting a wait-and-see/let the market settle itself out approach, however, may result in a lack of guardrails needed for the responsible development and use of this technology, and lead to similar misuse by stakeholders.

Conflict of Laws and Standards

Today over 700 initiatives across 60 nations are in some stage of development.¹⁸ These include national and international regulatory strategies, industry standards, public consultations, and development of oversight boards. Almost 250 of these initiatives are national strategies and plans, each of which will reflect the foundational theories of governance reflected across the disparate jurisdictions: in other words, laws developed around theories of individual rights, market-based drivers, or security of the state over the individual. Just as we have seen the difficulties companies and governments have faced in regulating the Internet and world-wide-web given its global reach, so too will we face similar issues with AI.

As algorithms trained by different developers to manage specific tasks become integrated into larger systems, and as these systems become more independent from human intervention in decision-making, questions around jurisdictional oversight will become of greater concern. In addition, how companies will implement the volume of regulations – many targeting disparate pieces of a system (datasets vs.

¹⁷ Steve Mills, Matthew Mendelsohn, et al., "Responsible AI Builds Trust in Government," *BCG*, January 2021, <https://web-assets.bcg.com/98/15/18c0ec044bb7b9e2a372ea036371/bcg-responsible-ai-builds-trust-in-government-jan-2021-r.pdf> [visited 2 February 2022].

¹⁸ National AI Policies and Strategies, <https://oecd.ai/en/dashboards> [visited on 19 February 2022].

algorithms vs. use vs. data collection vs. access vs. documentation) – is another question. Without coordination across stakeholders, we risk either (a) regulatory interventions with compliance difficulties or (b) a race to develop an initial standard that requires global compliance, like the GDPR, but which may impact future development and use of the technology.

Academia and Industry: A Love-Hate Relationship

Academia and industry are also facing bottlenecks in driving collaboration. Industry is often hesitant to open itself to researchers or academic institutions as it may lead to a perceived lack of competitive advantage, criticism that could impact revenue or market price, or slow a process already in flight. Industry often requires researchers to work with strict guidelines when collaborating, which calls into question the objectivity and accuracy of resulting research and reports. For academic institutions and research that do not take corporate funding and cooperation for research, the results may be well intentioned recommendations for a path forward, but often are theoretical and disconnected from technical or organizational practices and implementation.¹⁹ As a result, outcomes continue to result in silos that do not meet the needs for moving us forward in a meaningful way.

Self-Regulation

In many ways the rise of the Internet and the technologies that it drove biased us toward self-regulation as the most practical approach to early-stage innovation. But we are now experiencing the dark side of that self-regulation, which has forced us to question if self-regulation is still an appropriate path. Indeed, many of the companies,²⁰ governments, and individuals that benefited most from self-regulation have acknowledged that with AI, some form of regulation is necessary. The question over what types of regulatory guidelines around responsible development and use of technology the tech industry must implement is one that no private company will voluntarily put in place if it perceives it will result in a disadvantage in favor of a competitor. We must therefore carefully balance the important rights of a private company to act as a private company if we are to expect transparency from the same.

Private organizations are not non-profit organizations, nor are the tools they develop public utilities. In general, companies build products to meet a customer need (real or anticipated) and to earn revenue from the same. They do not generally develop

¹⁹ Jessica Morley, Luciano Floridi, Libby Kinsey and Anat Elhalal, “From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices,” *Science and Engineering Ethics*, Vol. 26, No. 4 (2019): p. 2141-168.

²⁰ See, e.g., “Microsoft President Brad Smith Calls for AI Regulation at Davos,” *GeekWire*, 21 January 2020, <https://www.geekwire.com/2020/microsoft-president-brad-smith-calls-ai-regulation-davos/> [visited on 18 February 2022].

products for nefarious purposes. But, as we have seen, there are unforeseen consequences that arise when we do not anticipate potential misuses or issues that arise in development processes or product use. Companies often approach development, particularly around technology, as a puzzle to solve. The question of whether it *should* be solved is one that, particularly in anticipation of AGI, becomes potentially more existential than with other technologies and certainly requires a different approach from “move fast, break things, and fix the bugs later.” It serves as a fundamental shift in strategic thinking and one that will take time, resources, and organizational change to make a natural part of the development or business practice. As a result, there is a hesitancy to change, to be more transparent, or to close any door to a potential business opportunity in the future – known or unknown – which has surfaced as a lack of meaningful collaboration with other stakeholders. And this, perhaps, is one of the biggest potential roadblocks we must overcome.

Where Do We Go from Here?

We need to regenerate the willingness of stakeholders to work together towards a common ground and to be more transparent with one another to form a workable solution. By pro-actively organizing and working across stakeholders, we can help drive the development of the governance systems we will need both today and tomorrow. Industry and government need to be more transparent and forthright in where we are and where we are going and willing to address the question of *should* we proceed with something (even if it will meet a consumer or shareholder’s needs), and academia and civil society need to be willing to develop realistic regulations that are implementable, do not slow innovation or create more problems in development, or redefine how a private company operates and governs.

So how do we do this when stakeholders have in many respects gone to their corners? First, education is a critical part of collaboration. Where the concentration of technical expertise lies increasingly within smaller and smaller groups, collaboration becomes an uneven playing field. Education and dialogue need to be at a level that is accessible to a wide-variety of stakeholders and not (a) purely academic, (b) understood only by engineers, or (c) developed by a corporate PR-department. This requires a certain amount of transparency by those who hold the concentration of expertise - industry. Understanding the fundamentals of *what* we are looking to regulate or not regulate is critical, as is understanding whether the desired outcome is technically feasible. The narrow AI developed today is in many ways the foundation of AGI in the future. Those building and deploying that technology today, need to document the process and the limitations, and be honest about the same, so we better understand them for tomorrow and can build appropriate guardrails around the same.

Second, dialogue on an international level that allows all stakeholders to be more forthright in their positions, concerns, and competing pressures is necessary. This may involve multi-stakeholder collaborations at the outset employing tools such as Chatham House Rules, giving participants some reassurance that their collaboration will not have negative individual consequences. Development of voluntary ethical guidelines for AI, such as those recently adopted by UNESCO, are excellent resources to use to guide the discussion. And international organizations with the expertise to bring together many stakeholders across the world to negotiate critical agreements on standards and rules are perfect for driving the initial collaboration (UNESCO, WIPO, WTO, ISO, IEEE, and others).

Third, where we take that dialogue – or prioritization – comes next. Part of the collaborative process should be agreement on what aspects of regulatory oversight need to be (a) international, (b) national, (c) standards-driven, and (d) self-regulated. We do not need and will not want regulations for all aspects of AI development and use. Over-regulation, just like lack of regulation, will result in poor development, but also difficulty in managing compliance by both companies and regulators (not to mention limiting experimentation and innovation). Similarly, many regulations that currently exist may well apply to the problems we are concerned about with AI. Identifying what truly needs a novel approach will help narrow down the task at hand. To that end, prioritization of issues arising in the development and end-use of AI will help guide this process. Of critical importance at the outset are issues around: Safety and Human Good; Government/Military Use of AI; Human Rights; Competition and Markets; Intellectual Property; Data Rights and Privacy; Education and Labor; and Fairness and Bias.

Finally, we need to address how governments, standards bodies, and companies will address compliance. What mechanisms and bodies need to be in place to ensure regulations and industry standards are met? Who has oversight and what are the remedies? What requires strict compliance and penalties, and what rights do individuals have to seek remedies for harm? How do we encourage ongoing transparency?

Conclusion

Whether or not we need to regulate AI is no longer a question. The biggest issue is what requires regulation, what industry standards are needed, and what should be left to self-regulation based on those priorities that are most important to (a) protect our larger society from negative unintended consequences and misuse, including incursions on the rights of the individual, (b) prevent a race to dominance at the expense of safety or the principal of use of AI for human good, (c) encourage a

competitive marketplace, and (d) allow for open experimentation and innovation. To accomplish this requires all stakeholders be willing to collaborate in an open and transparent manner, to listen to different approaches to the same problem, and to develop solutions that have the flexibility to move with the technology and society as it matures.

Collaboration is difficult and will require years to get right. Many of these processes have already begun and it is critical that every stakeholder come to the table willing to participate in an open and transparent way, particularly those that are primarily responsible for introducing AI and its positive impacts to the world.