Institute for Public Policy Research

THE NEW POLITICS OF AI

WHY FAST TECHNOLOGICAL CHANGE REQUIRES BOLD POLICY TARGETS

Carsten Jung

February 2025

ABOUT IPPR

IPPR, the Institute for Public Policy Research, is an independent charity working towards a fairer, greener, and more prosperous society. We are researchers, communicators, and policy experts creating tangible progressive change, and turning bold ideas into common sense realities. Working across the UK, IPPR, IPPR North, and IPPR Scotland are deeply connected to the people of our nations and regions, and the issues our communities face.

We have helped shape national conversations and progressive policy change for more than 30 years. From making the early case for the minimum wage and tackling regional inequality, to proposing a windfall tax on energy companies, IPPR's research and policy work has put forward practical solutions for the crises facing society.

IPPR
4th floor,
8 Storey's Gate
London
SW1P 3AY

E: info@ippr.org
www.ippr.org
Registered charity no: 800065 (England and Wales),
SC046557 (Scotland)
This paper was first published in February 2025. © IPPR 2025

The contents and opinions expressed in this paper are those of the authors only.

The progressive policy think tank

CONTENTS

ABOUT THE AUTHORS

Carsten Jung is head of AI at IPPR.

ACKNOWLEDGEMENTS

The author would like to think Harry Quilter-Pinner, Seb Krier, Elisabeth Nöfer, Charles McIvor, and Philip Fox for their comments.

Download
This document is available to download as a free PDF and in other formats at:
http://www.ippr.org/articles/new-politics-of-ai

Citation
If you are using this document in your own writing, our preferred citation is:
Jung C (2025) *The new politics of AI: Why fast technological change requires bold policy targets*, IPPR.
http://www.ippr.org/articles/new-politics-of-ai

Permission to share
This document is published under a creative commons licence:
Attribution-NonCommercial-NoDerivs 2.0 UK
http://creativecommons.org/licenses/by-nc-nd/2.0/uk/
For commercial use, please contact info@ippr.org

SUMMARY

The upcoming AI Action Summit in Paris is an opportunity to show how we can harness artificial intelligence (AI) as a force for societal, economic, and environmental good. The objective of the summit, organised by the French government, is to "catalyse or scale concrete actions that can steer AI away from misuse and abuse and towards opportunities and benefits".

In this briefing, we outline how the summit constitutes the first event of a new era of AI policymaking that links AI policy to delivering public value. AI is fundamentally different from other technologies – it is set to unleash a vast number of highly sophisticated 'artificial agents' into the economy. AI systems that can take actions and make decisions are not just tools – they are actors. This can be a good thing. But it requires a novel type of policymaking and politics. Merely accelerating AI deployment and hoping it will deliver public value will likely be insufficient. **We argue that AI needs to be directed towards societies' goals, via 'mission-based policies'**. How these are determined and measured is at the heart of this new politics. Given the speed of AI development, the AI summit will be a crucial moment for national leaders, civil society and businesses to discuss this.

The Paris summit comes at a critical historical juncture. The development of artificial intelligence is marching on rapidly. While there was a possibility that improvements of AI systems would slow down in 2024, they have, in fact, accelerated (see chapter 1). The rapid improvement of AI will give rise to transformative changes in economy and society. New infrastructure, processes and products will be built, deeply incorporating AI into the economy. The performance improvement of models has been so fast that the incentive to implement them is drastically rising.

We need 'a new politics of AI'. Much of the policy debate around AI so far has been how its development and deployment can be accelerated in a safe way. Sorely missing in this debate is a discussion about *for what ends* we want it to be used. We need not just 'accelerationism', we also need 'directionism'. The specific contours, policy choices, and debates around this, which we outline in this report, are the new politics of AI. This will need to put big societal needs and democratic debate at its heart.

Setting the direction of AI innovation, early on, matters. The infrastructure and design decisions made today could create significant path dependencies. The investments of today could be decisive for future outcomes. This is why the Paris summit's focus on setting up 'public' AI infrastructure is a crucial first step. It involves setting up more decentralised technology systems, which give users, citizens and public policy more control over the conditions of deploying AI. We argue, in addition, it should be combined with clear missions and problem statements that AI deployment should help tackle.

Even though it will be difficult to achieve, policymakers – working with citizens and businesses – should develop such missions and to steer AI deployment through policy.

We outline three different types of steering AI adoption.

1. In some areas no strong mission-based steer is needed. In other words, not all types of AI deployment (eg customer service applications) need to be steered by mission-based policies, and can instead can be governed via updating existing sectoral regulations.
2. In other areas there are big misalignments or deployment gaps, where more active policy steer via missions is needed. For instance, AI deployment in health needs to be steered towards one of the most effective applications: prevention. While there is currently much investment in developing better diagnostics and cures for diseases, technological applications that *stop people from getting sick in the first place*, rather than merely curing them when they are sick, are currently underdeveloped. Mission-based policy can help fill such gaps.
3. Finally, there are also areas where AI could have transformative effects on society (eg AI agents as social companions), but it is simply yet unclear in which direction to steer adoption. In such cases, we would recommend governments to slow adoption, protype and iterate. This can be seen as a large scale 'sandbox' approach, where deployment is closely monitored, assessed and adjusted, if needed.

This new politics of AI will require a step change in democratic engagement with policymaking. **The potential changes caused by AI are too significant to leave them to technical policy experts to decide**. We think there are six factors that will feature in this new discourse.

1. There should be monitoring and reporting of how AI is being deployed. This needs to be much more granular than current high-level surveys on AI deployment in the economy.
2. Setting societal objectives which policy (and AI deployment) should solve should be a more important aspect of democratic politics. Politicians should outline their vision for a future with ubiquitous AI that goes beyond platitudes and instead sets specific goals.
3. Especially with regards to entirely novel issues direct engagement of users and citizens is desirable. This includes emerging trends such as the role of AI agents on social media (eg sophisticated bots interacting with humans or other bots in online discussions).
4. Iterative deployment and social discussion could allow for fast deployment, but then allow the public to discuss, feedback and iterate.
5. Market-driven adoption (with little to no steer) by policy might be preferable by some in some sectors.
6. The military use of AI will cast a long shadow over many other types of deployment. Especially in a world with increased geopolitical competition this will be an area of significant policy focus. Much of such AI activity will understandably be classified and thus hard to have a public discussion about. However, there will be some guardrails of military use that should be developed to assure the public about how technology is used in the name of their security.

Many countries – including the UK – already have mission-based policies, but these need to broken down further into sub-targets and then explicitly connected with policy incentives for deploying AI that helps deliver those missions. In many cases (such as improving prevention in health) better AI alone will not be sufficient to achieve targets, so AI policy needs to be developed together with other policies (eg reaching people via community health centres). This will need to involve wide-ranging collaboration across government departments on AI policy and potentially significant public and private investment.

This will be worth it: new, powerful AI has the potential to help solve large economic and societal challenges – but we need to invest the resources to steer it in the right direction.

1. INTRODUCTION: GENERATIVE AI WILL DISRUPT AND TRANSFORM THE ECONOMY

> ***"The long-term result of these decisions has hardly fulfilled the utopian vision. When Americans made the car an idol, they relied on a machine to solve problems not reconciliable by a machine."***
> Clay McShane (1994) *Down The Asphalt Path*

The way in which technological revolutions are steered can have large impacts on economy and society. Take the example of cars: the rise of the automobile around the turn of the 20th century exerted a transformative and often under scrutinised influence on urban development. In many cities in the West, the rise in car ownership from the 1920s onward prompted urban planners, political figures and business interests to implement measures to accommodate this shift. Although some public consultation and forecasting were undertaken, these efforts often insufficiently planned for the socio-economic impacts (Mumford 1961, Hall 1991, McShane 1994). For instance, in Los Angeles, expansive highway networks were constructed that frequently disrupted established neighbourhoods and contributed to the gradual relocation of commercial centres to the suburbs. Hall (1991) highlights:

> ***"mass motorization had already begun to impinge on American cities by the mid-1920s ... By 1923, traffic congestion in some cities was already so bad that there was talk of barring cars from downtown streets."***

These developments helped spur unprecedented growth and social progress. But Mumford (1961) asserts that this new, car centred environment had profound consequences for social life, creating distances too great for walking, diminishing access to public spaces, and encouraging the formation of homogenous residential enclaves (McShane 1994). Could more directive policy early on have made a difference? Mumford (1961) argues that better strategic planning – early on – would have led to better road and city planning, and enduringly embedded a higher standard of living. Hall (1991) points out this shift:

> ***"[in the 1960s] planners and politicians belatedly discovered the continued deprivation of the inner-city poor; then, it was seen that the areas where these people lived were suffering depopulation and deindustrialization; in consequence, planners progressively moved away from the merely physical, and into the social and the economic."***

We argue the advent of new, powerful AI technology could be similarly transformative for economies and societies. Steering deployment – early on – could ensure that we avoid going down a path that, as with an overly car-centric city design, we might later regret.

The development of artificial intelligence is marching on rapidly. While there was a possibility that improvements of AI systems would slow down in 2024, they have, in fact, accelerated. AI models capable of undergraduate-level reasoning have become

much cheaper (with GPT4 costs having fallen 100 times) in just 12 months (Benaich and Chalmers 2024). Newer models are now capable of PhD-level reasoning, measured by AI models achieving top scores in scientific and diagnostic reasoning tests (Mollick 2024). At the end of 2024 a new model released by OpenAI was able to solve maths problems that weren't expected to be until many years in the future (ibid).

Even if AI development stalled where it is today, implementation across the economy could have disruptive or positive transformative consequences. Our previous research found that AI could either lead to 8 million job losses, in the UK, and no GDP gains, or no job losses and GDP gains worth up to £306 billion a year. The outcome will depend on whether AI is deployed merely to automate human work, or produce better and more goods and services (Jung and Srinivasa Desikan 2024).

But implementation is unlikely to stall. AI 'agents' have already been launched, able to do advanced multi-step research tasks, conduct multi-step actions online, or to develop, test and de-bug computer programmes. In text exchange, humans are no longer able to distinguish whether they are interacting with an AI or another human (Rathi et al 2024). Businesses and individuals will increasingly be able to make them act on their behalf, including by given them access to their personal and proprietary data. The availability of AI agents can be seen as a vast, cheap increase in the undergraduate workforce (Modei 2024). In other words, technology is already advanced enough to create large value. The question is merely when and how people and organisations choose to implement it.

What we are likely about to witness is not just application of a technology to the existing way of doing things, but a change of the ways in which the economy works. Processes will increasingly be changed and built around cheap generative intelligence, rather than around humans.

BOX 1: THE MORE INTEGRATED GENERATIVE AI BECOMES, THE MORE IT COULD CHANGE THE SHAPE OF JOBS AND THE ECONOMY

In figure 1.1 we show, based on our previous quantitative analysis, how more and more tasks in the economy could be aided by AI, if it is to become more integrated.

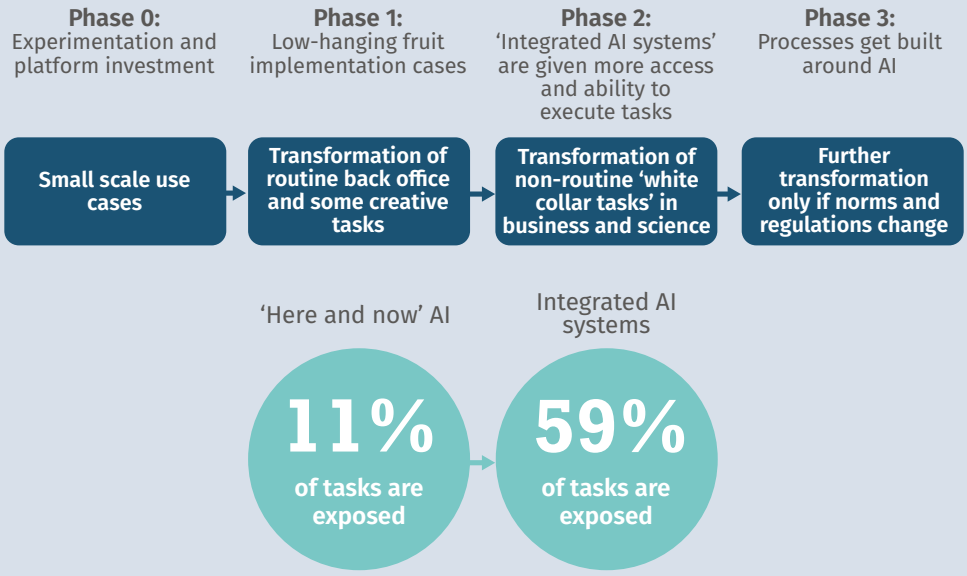
Phase 1 refers to implementation in organisations that will likely target 'low hanging fruit' use cases. These are the cases where generative AI programmes are relatively easily plugged into existing IT processes, without many changes to workflows. About 11 per cent of tasks would be heavily impacted by this. Back office jobs (such as personal assistants), entry level jobs and part time jobs will be most exposed in this first phase. And we find that women will be significantly more affected (as they are more likely to work in the most exposed occupations, such as secretarial and administrative occupations). This phase will largely involve keeping existing work processes in place, but conducting them with generative AI.

The second phase is one where generative AI becomes more deeply integrated with existing organisational processes. If organisations decide to integrate existing AI technology more deeply into their processes (which is not a given), we find that almost five times more tasks – about 59 per cent of tasks – are exposed. This means a large number of jobs could 'feel' its impact, and it will also increasingly affect high paying jobs too. Such integration into existing processes would, for instance, mean giving AI the ability to access proprietary data, providing inputs via apps or giving AI systems the ability to execute tasks (eg making orders or bookings). Whether this will materialise and who gains and who loses will depend on a number of policy and organisational

factors. Crucially, it is likely that that not all organisations will adopt the technology at similar rates, leading to inequalities. The question is how fast we are transitioning to phase 2.

FIGURE 1.1

We are currently in phase 1 of AI deployment, with a deployment focus on back office tasks



Source: Jung and Srinivasa Desikan (2024)

Phase 3 is one where norms shift (eg people becoming more used to interacting with AI where previously they would have interacted with humans), and whole new processes get built around AI, new tasks devised and performed.

While the last wave of technological change and globalisation affected mostly manufacturing jobs, the current wave is disproportionately affecting the knowledge economy at first. Figure 1.2 shows that about 50 per cent of tasks (weighted by hours worked) are what can be described as knowledge (or 'white collar') economy tasks. Of these, as much as 70 per cent could be significantly transformed by generative AI.

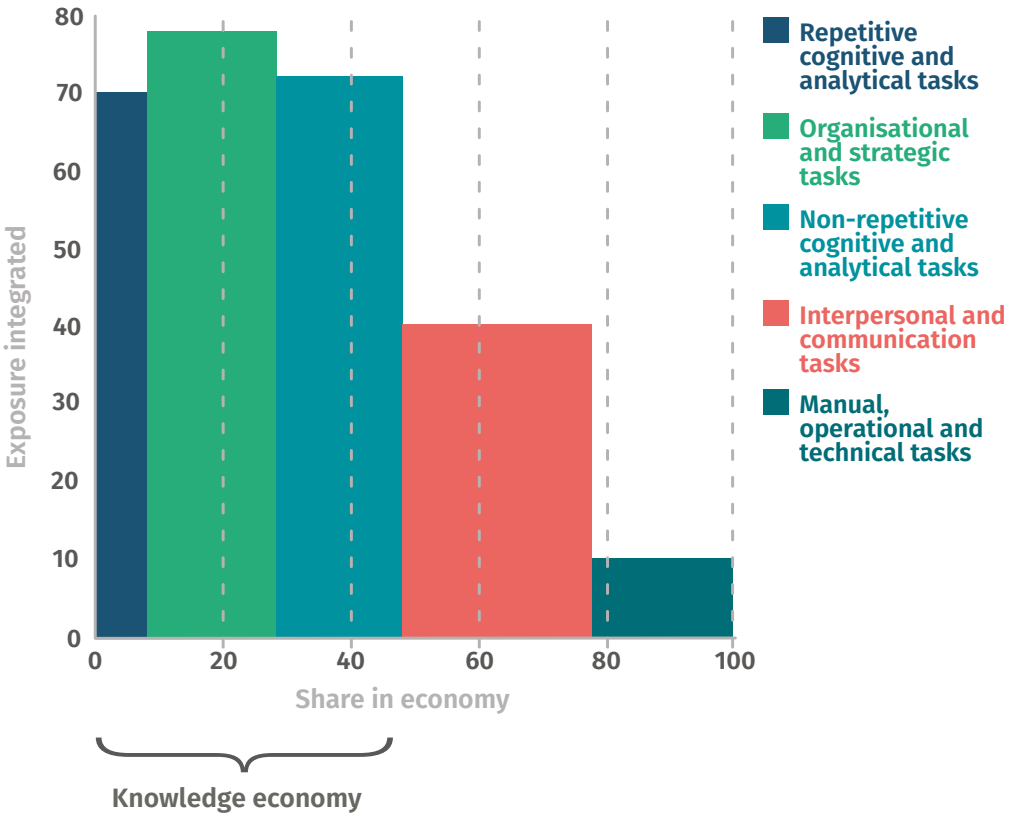
AI will therefore likely be a catalyst for change. It will transform jobs, destroy old ones, create new ones, trigger the development of new products and services and allow us to do things we could not do before. But given its immense potential for change, it is important to steer it towards helping us solve big societal problems.

Much of the policy debate around AI thus far has been revolving around how its development and deployment can be accelerated. Sorely missing in this debate is a discussion about *for what ends* we want it to be used. We need not just 'accelerationism', we also need 'directionism'. The specific contours, policy choices, and debates around this, which we outline in this report, are the new politics of AI.

FIGURE 1.2

Knowledge economy tasks, even creative and non-routine ones could be heavily affected by generative AI, if it was deeply integrated

Share of tasks exposed to generative AI



Source: IPPR analysis of 22,000 ONET tasks, analysed following the LLM-annotation based on Elounoudo et al (2023)) and matched with ONS Labour Force Survey data. See Jung and Srinivasa Desikan (2024) for detailed methodology.

2. ALIGNING AI DEPLOYMENT WITH PUBLIC VALUE

THERE ARE MULTIPLE SCENARIOS FOR AI ADOPTION – POLICY NEEDS TO STEER IT TOWARDS PUBLIC VALUE CREATION

The upcoming AI Action Summit in Paris is an opportunity to show how we can harness artificial intelligence (AI) as a force for societal, economic, and environmental good. The objective of the summit, organised by the French government, is to "catalyse or scale concrete actions that can steer AI away from misuse and abuse and towards opportunities and benefits".[1]

This is filling an important gap in the current AI policy discussion. The EU AI Act, for instance, focusses on AI model *governance* in risky contexts. It prompts AI deployers to root out bias and ensure accuracy and contestability of decisions made by AI systems. While such policies are crucial for addressing risks, less well developed are policies that aim to ensure AI is delivered in a way that delivers public value.

In AI policy, there are currently more prevalent approaches: '**accelerationism'** refers to policies aimed at enabling AI development that makes it better and cheaper and thus increases AI adoption across the economy, with little preference over the types of deployment. An example of this is the UK government's AI Opportunities Plan (Clifford 2025). Such measures will be crucial for making AI widely available and expanding choice. 'AI safety-ism' focusses on avoiding clearly defined risks, no matter how advanced or what type of AI application. The work of AI safety institutes or the EU AI Act are examples of this.

We argue that to achieve 'AI for public good', a third strand of policy is needed: **'directionism'**. **This is the idea to for policy to steer the direction of AI deployment actively**, using policy incentives – such as targeted funding, public procurement or public infrastructure conditionalities – for building of products and services that create public value (table 1.1).

1 It will feature the 'AI Convergence' challenges, which encourage innovative projects addressing critical technological and societal issues. The event's five thematic priorities—public interest AI, the future of work, innovation and culture, trust in AI, and global governance—serve as a framework for fostering solutions that align AI with public interest. By showcasing selected projects from around the world, including those from the Global South, the summit seeks to highlight AI's potential to address pressing global challenges while promoting trust and transparency in its use.

TABLE 1.1

Policy should focus more on shaping the direction of AI innovation, not just acceleration and risk mitigation

	Goal	Policy tools	Examples
Accelerationism	Increase AI deployment, by making it better, easier and cheaper to use	Give businesses and people access to capital, digital infrastructure, talent	UK AI Opportunities Plan, investments in public sector supercomputing capabilities (UK Day One, 2024)
Safety-ism	Avoid clearly identified risks	Safety testing, privacy safeguards, anti-bias assurance	EU AI Act, AI Safety institutes (eg UK, USA, Singapore)
Directionism	'Steer' innovation towards solving important societal problems	Provide incentives to build services and research that explicitly solve societal problems	Outline specific missions and milestones eg in preventative health or climate

Source: Author's analysis

All major AI developers use a range of tools to ensure 'alignment' of models with social and safety norms (eg 'teaching' AI models to not be rude or not output manuals for how to build bioweapons).[2] However, there has been less work on aligning AI *deployment* more widely with societal goals (Ji et al 2023). This is partly because societal goals are more fuzzy and sometimes contested, but also because it involves the interplay of multiple players (both public and private), infrastructure and different policy areas across government departments.

We argue that the advent of powerful AI raises the stakes so much that policymakers should attempt to devise such 'system alignment policies'. These are measures that **shape the way AI is implemented and the way it interacts with existing systems**. As with any systems design issue, ultimate objectives should be set out at the start. But the way it is deployed should be agile — specifically, it should be iterative and flexible. In other words, it is important to learn over time 'what works'.

BUT HOW TO DEVELOP A 'DIRECTION' FOR AI?

We argue that 'mission thinking' can be a tool for ensuring AI is deployed for public value. Mission-based industrial policy focusses on addressing broad societal challenges through coordinated efforts across various sectors (Mazzucato 2013, OECD 2021). This approach involves setting clear, ambitious societal goals (such as increasing healthy life expectancy) that guide innovation and investment towards solving specific issues, such as climate change or health inequities, rather than merely supporting individual industries or technologies (Quilter-Pinner 2024).

Famously, the Apollo mission was an early mission-based industrial strategy by setting a clear, ambitious goal—landing a human on the moon. John F Kennedy's 1961 goal of "landing a man on the Moon and returning him safely to the Earth" before the end of the 1960s. It mobilised multiple sectors and stakeholders (ranging from businesses such as Boeing and IBM to research universities to government entities like NASA) towards a transformative objective. It involved

2 One such approach is reinforcement learning with human feedback.

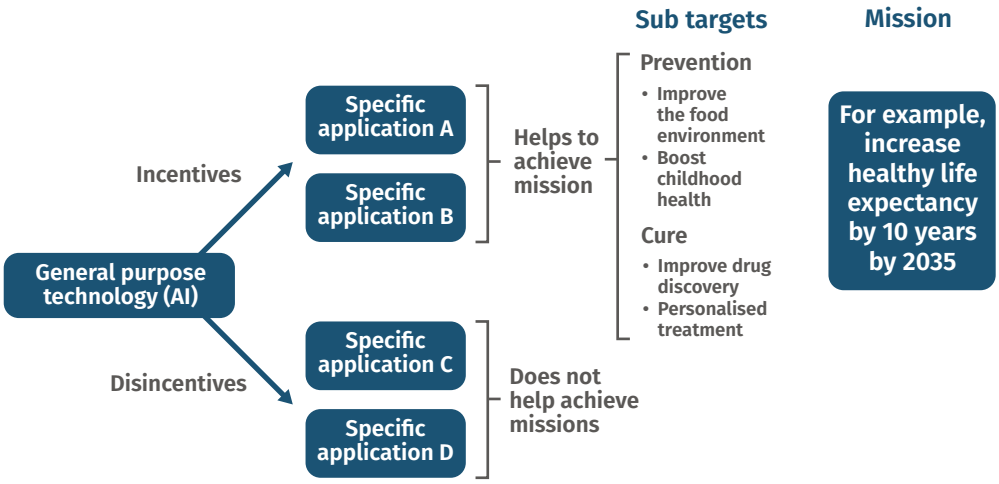
significant public investment – NASA received 0.7 per cent of GDP in the 1960s (Juhász et al 2023). This collaborative and interdisciplinary approach led to advancements in computing, materials science, and systems engineering. Technologies ultimately resulting from this include GPS and modern laptops and phones (Mazzucato 2013).

AI could be a crucial part in achieving the next set of Apollo missions. It is a general purpose technology that can be used in a multitude of ways (figure 2.1). Similar to electricity, it can be used to provide a wide array of services. It is in the specific applications and their context that mission-alignment will become clear.

Many countries and blocs – such as the UK or the EU – have already started setting clear missions. The next step will be to break them down into 'sub-targets' and put in place policy incentives that clearly guide AI adoption towards delivering on those targets.

FIGURE 2.1

'Steering AI', will involve steering it towards certain beneficial applications



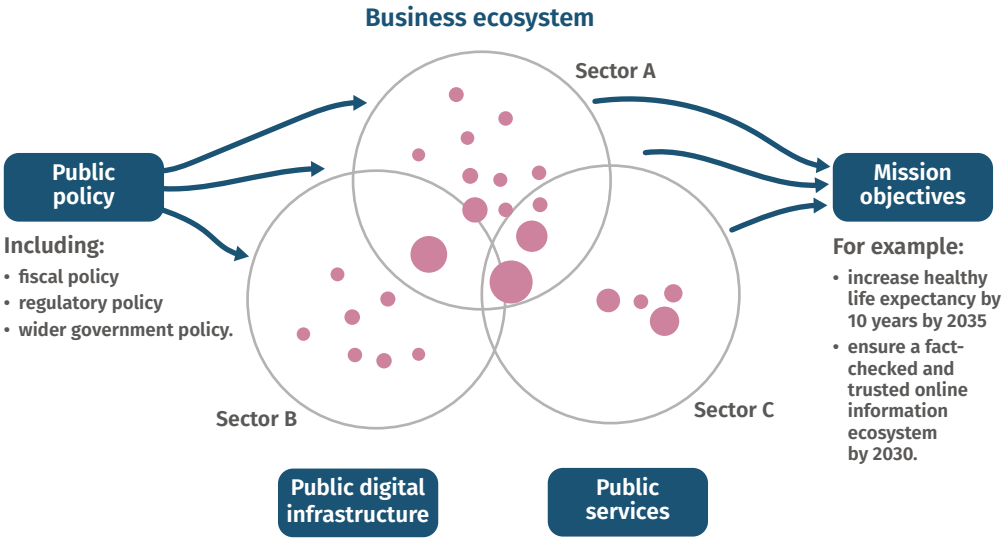
Source: Author's analysis

Steering AI adoption will require multiple policy tools from across government departments

Setting and delivering missions for AI requires setting incentives across sectors, providing infrastructure and embedding them in public service delivery. Figure 2.2 illustrates how mission objectives like improving healthy life expectancy require coordinated action across a business ecosystem, supported by three key enablers: public policy (including fiscal and regulatory measures), public digital infrastructure, and public services. This highlights what is special about mission-based policies: in pursuit of an ambitious goal, they work across government departments and ultimately deliver through the private sector and civil society. Quilter-Pinner (2024) and Dibb et al (2023) show how mission-driven policymaking differs from the way policy is delivered in many countries. At its core, it sets out to deliver specific outcomes, and works with a set of metrics to track these, using a wide range of policy tools – from procurement to skills policy – to realise it.

FIGURE 2.2

Multiple policy levers are needed to steer the business ecosystem towards delivering missions



Source: Author's analysis

Turning a mission into a policy roadmap will involve:

- setting targets for specific societal goals to be achieved
- devising the sub-targets and specific problems that will need solving to deliver the overall objective
- using a wide array of policies to stimulate activity across the ecosystem
- putting in place institutions that allow for adaptive learning.

We argue that this perspective is missing from the AI and tech policy discussion, which is excessively focussed on AI models but not on their deployment context. But for directing deployment towards social good, more systemic thinking is needed.

AI didn't create the need for mission-driven innovation - but its unprecedented speed and power to transform the economy means getting mission-based policy right becomes crucial.

But steering AI can take different forms depending on the area

There are three approaches to steering AI innovation, which vary depending on the area of deployment.

1. In some areas no strong mission-based steer is needed. In other words, not all types of AI deployment (eg customer service chatbots) need to be steered by mission based policies, and can instead be governed via updating existing sectoral regulations.
2. In other areas there are big deployment gaps, where more active policy steer via missions is desirable.
3. Finally, there are also areas where AI could have transformative effects, but it is simply yet unclear in which direction to steer adoption. In such cases, we would recommend governments to slow adoption and review.

TABLE 2.1

Three different modes for steering AI adoption

		Policy direction	Example	Policy approach
A. Steering via existing sector regulation	Less clearly defined social objectives	Limited additional regulation needed to steer AI deployment towards public value.	AI in finance, AI for pharmaceuticals.	Empower existing regulators to see how their applications need to be updated.
B. Setting missions	Clearly defined social objectives	Put in place incentives that align AI deployment with missions.	Increasing health outcomes through prevention .	Identify gaps. Use fiscal and other tools to boost AI depoloyment in these areas.
C. Slowing down, prototyping and iterating	AI is causing significant social change.	AI developments create entirely new sets of issues that need direction, but it might be too soon for policy to know.	Personal AI companions; AI agents on social media; personalised AI news generation.	Setting up reporting and monitoring frameworks, agile policy.

Source: Author's analysis

A. STEERING AI VIA EXISTING FRAMEWORKS

In some sectors and areas of application, business activity is already well aligned with public value creation, or regulators are well placed to steer AI. For instance, we might not need mission-based policy for ensuring AI bots in customer service deliver public value, and there might be no additional policy required to ensure that software coding assistants are aligned with the public good. Similarly, in finance, existing regulators – the FCA, Bank of England and the pension regulator – have wide-ranging powers to steer technology adoption in the finance sector. They should work out plans for fostering AI adoption.

B. STEERING AI VIA MISSIONS

Other areas face bigger challenges. These are areas with *clearly defined societal objectives* that are proving hard to solve (Lund et al 2022). At the current moment, in G7 countries, this includes for instance stagnating population health outcomes, unaffordable housing costs, and the need to accelerate the energy transition.[3]

Consider improving population health. It is a decided goal of all G7 governments but despite a large part of their budget being spent on it, since the turn of the millenium, progress in in extending people's illness-free lifespan has stalled.[4] It is one of the areas where a huge amount AI development is taken in place, mostly focussed on screening and treatment. These will lead to crucial breakthroughs on the early treatment of diseases. And yet, the current direction of AI adoption might be leaving significant gaps. IPPR (2024) argues that significant improvements in

3 There are other problems, such as high cost of living, but this would need to be further broken down in order to be the subject of mission-based policy.

4 See: Our World in Data (2025).

public health require shifting from a reactive 'sickness model' to a proactive 'health creation system' focused on **illness prevention**. For instance, in the UK, almost 7 million people live with major illnesses, many of which would have been preventable (Heath Foundation 2024). This includes chronic pain, COPD type 2 diabetes, cardiovascular diseases, anxiety and depression. **The lack of significant progress on preventative health is a big societal problem.** The need for stopping people from getting sick, rather than only curing them when they get sick is widely advocated by health researchers and officials.

New AI technologies can help deliver this, but the incentive structures in the health system will not by themselves guide health investment towards prevention. We consider it a major risk that AI is not sufficiently deployed to make progress on *preventing* disease and, as a result, fails to contribute to significantly longer, healthier lives.

IPPR identifies five key pillars for prevention: improving work environments, creating healthier food environments, supporting early years and childhood health, enhancing community health infrastructure, and expanding preventative healthcare services. It needs to complement the other factors that need to be aligned too.

But, as we show in our forthcoming in-depth study of UK AI companies, the current UK AI healthcare ecosystem predominantly focuses on clinical care, diagnosis, treatment, and operational efficiency, as shown by the distribution of AI applications across different categories (Srinivasa Desikan and Jung, forthcoming). While these innovations, such as real-time, data-driven medicine and process improvements, could bring valuable benefits to healthcare delivery, there remains a significant gap in AI applications targeting other preventative health measures. Take the food environment. AI applications being developed like the UK's EatingAI offer individual nutrition guidance, which could have significant benefits. But they do not yet address broader systematic changes needed to create healthier food environments. More widely, while AI applications show promise in areas like mental health support and early disease detection, these currently address only limited aspects of these preventative pillars.

C. SLOWING DOWN AND ITERATING

There are other areas in which AI will completely change the status quo, but there may not yet exist a *clearly defined societal objective*. These will need significant steering towards mission-based policies, and require in-depth attention.

For instance, **AI companions** are fast increasing their user base: Replika had 30 million users and Character.AI had 20 million users in 2024. Researchers have highlighted the potential for both addiction, and risks around long-term psychological impacts (MIT TR 2024). There have been serious allegations against these: Chracter.AI is currently being sued for the potential involvement of one of its bots in encouraging the suicide of a teenager in the US (NYT 2024).

EU AI legislation is aimed at preventing aspects of this. The EU AI Act prohibits 'exploitation of vulnerabilities' and 'manipulative or deceptive techniques' (European Parliament 2024). So, in the EU at least, models would have to be tested against such catastrophic AI behaviours. The UK Online Safety Act might also be effective in preventing this.

But even though such problems could be prevented through pre-deployment testing, there are wider societal questions that are not covered by such risk mitigation strategies. The wider issue is: what type of interaction with AI companions do we want in society? To what extent should the incentives for making them addictive be addressed? Are there unintended consequences from people having meaningful relationships with artificial agents? To what extent does democracy require online

interactions to be between humans, rather than between humans and artificial agents? These questions can likely not be answered before people have gathered experience with new types of applications. So, we argue, the best way of dealing with this is slowing down rollout, prototyping and iterating.

This is of course at odds with some of the fast deployment cycles that characterise many tech products. But, at the same time, **the early experience with the release of new generative AI models provides a positive lesson: ChatGPT was released gradually since 2022, allowing some degree of societal discussion and allowing developers to incorporate feedback**. Policymakers should reflect on how a similar approach could look for future AI deployment in different contexts.

This approach could work in similar ways to regulatory 'sandbox' approaches – these are controlled testing environments where new technologies are trialled under regulatory oversight before wider deployment. But it would work at a bigger scale, engage real users from the outset, and capture genuine usage patterns and wider societal impacts. It would also address broader questions of human–AI interaction and social values, not just regulatory compliance. Finally, by fostering ongoing collaboration between developers, regulators and other stakeholders, it would create a flexible framework that can evolve alongside the technology.

3. THE NEW POLITICS OF AI

Pursuing 'directionism' will require bold incentives and iterative AI deployment. There will need to be more attention to the frontier of AI *deployment* and more attention to 'mission' politics. To make this work, we think there are six elements that will shape the debate (table 3.1). The weight given to each of them will determine the overall policy approach taken to shape AI deployment.

1. **There should be monitoring and reporting of how AI is being deployed.** Currently there is some survey based information and industry datasets, but they are high level and actually entirely unspecific about the type of applications that are being deployed. For example, the DSIT (2024) survey on this, looks at sectors and technologies, but not what the technology is actually *deployed for*. For instance, while such data outlines that AI is being deployed for 'health and wellness', there is little information what kind of applications are most prevalent, or what problems are being solved. It is thus hard to judge whether AI's positive potential is actually being realised or whether there are deployment gaps. In our forthcoming report (Srinivasa Desikan and Jung, forthcoming) we try to fill this gap developing an in-depth dataset of AI use in the UK economy. But much more could be done to keep citizens up to date on how AI is used in the economy.
2. **Setting societal objectives, for policy (and AI deployment) to tackle, should be a more important aspect of democratic politics.** As with the above example of roads' transformative impact on cities – AI too will be transformative. Politicians should articulate the direction they envision this to take. Visions for this will have to be much more specific than statements like 'AI will boost growth' and 'AI will help cure cancer'. They will need to involve very specific goals (such as, 'reduce childhood obesity by 25 per cent, by 2035') that AI should help achieve and by when.
3. **Direct engagement of users and citizens is desirable, especially with regard to entire novel issues such as the role of AI agents on social media.** This could involve citizen assemblies, polling and direct user feedback. This would be aimed at building a bottom-up picture of people's experience with AI deployment in the economy (say, AI teaching tools in schools, AI health advisors, online interactions with AI bots) and inform policymakers about how to steer it.
4. **There will be voices in politics that prefer to 'move fast and break things' and build products first and then make adjustments to the direction of deployment later on.** This arguably is how social media has evolved – with lots of freedom for businesses to deploy and most regulation (such as online harms laws) only following many years later. This approach, while allowing for speed, might create harms along the way. And – as with our introductory example of cars impact on cities – some of the design choices made might be hard to reverse. Thus, to pursue this approach more responsibly might involve iterative policy frameworks, that encourage fast deployment, but then allow the public to discuss, feedback and iterate.
5. **Some might resist steering AI innovation altogether.** The argument is that big and small AI firms "should be allowed to build AI as fast and aggressively as they can – but not allowed to achieve regulatory capture" (Andreesen 2024). Advocates of this approach might argue that markets will be best at steering the direction of AI deployment and regulators should only put in place some

broad guardrails. We agree, as outlined above, that there are many areas, where this is the case (see table 2.1). But we would argue that in areas such as public health, markets alone are currently sub-optimally allocating AI investment (IPPR 2024). That said, in the new politics of AI, where and when markets are effective at allocation of AI resources will likely be contested.

6. **There will be increasing discussion of AI use in military contexts.** For instance, OpenAI, in October 2024, announced that it is now working in the national security space, to "help protect people, deter adversaries, and even prevent future conflict" (OpenAI, 2024). Much of such activity will understandably be classified and it will thus hard to have a detailed public discussion about it. However, some guardrails of military use should be developed to ensure the public about how technology is used in the name of their security.

TABLE 3.1

Elements of the new politics of AI

	Approach	Pros	Cons
1) Monitoring and reporting of how AI is transforming society	Inform citizens about applications and cases of interest, against missions and sensitive areas	Crucial for keeping civil society informed about deployment	Data intensive
2) Representative democracy	Politicians are clear about the social objectives they want to achieve (with the help of technology)	More legislative time spent on defining missions, and sub-targets in order to improve alignment	Limited bandwidth of the public to engage with specifics of missions
3) Bottom up engagement	Large scale citizen and civil society engagement (polling, assemblies, user feedback) on sensitive AI issues.	Especially important for novel, sensitive areas. It's arguably a broadening of RLHF[5]	Difficulty to prioritise between conflicting goals
4) Iterative deployment with social discussion	Build things and then allow social debate to review and feedback	Allows innovations to go ahead, while allowing for critical feedback and adjustments	Could create unintended harms, once products are built it will be difficult to reverse course
5) Market driven design	Build products within legal framework	Allows for fastest deployment	Does not guarantee mission-alignment
6) Guardrails for military use	Acknowledging the importance of AI for geopolitical competition and secrecy, clarifying the parameters of its use	Giving some assurance to citizens on what advanced AI is used for by military institutions	Any guardrails might be seen by some as a geopolitical disadvantage

Source: Author's analysis

5 RLHF (reinforcement learning from human feedback) uses human preferences to refine AI behaviour, similar to how citizen assemblies and public engagement can shape AI deployment policies. Both approaches recognise that aligning AI with human values requires direct input from diverse human perspectives, though public engagement operates at a broader societal scale compared to RLHF's individual feedback mechanisms.

TO STEER AI, MISSION-BASED POLICIES NEED TO BE MORE AMBITIOUS

There are various international examples of mission-based policies that incorporate elements of such an approach. The US and the EU set ambitious targets for their energy transitions. The US Inflation Reduction Act (IRA) makes use of fiscal policy (mainly tax credits), the regulatory framework (eg clean energy standards), infrastructure development (grid modernisation and EV charging) and delivers this through cross-departmental coordination. It has led to a clean energy manufacturing boom (Clean Investment Monitor 2024) that might well continue after the change in government. South Korea's Korean New Deal (worth $135 billion, passed in 2020 – conducts mission-oriented policy backed by coordinated policy levers. The program integrates public infrastructure, business sectors and public services through strong institutions with regular evaluation. The UK's Industrial Strategy Challenge Fund (ISCF), while small in scale, was successful at coordinating innovation activity across businesses, local authorities and universities (Rand and Frontier Economics 2024). The EU's Horizon Europe missions too have set clear societal targets (eg 100 climate-neutral cities by 2030), backed up by a comprehensive policy mix (funding, regulation, procurement), including with integration of public services and infrastructure.

That said, **there is a clear case for improving countries' mission policy, to make them more effective**. While many might be taking their inspiration from the Apollo mission, they fall short of the ambition of the original US moonshot programmes or DARPA-scale schemes. Horizon Europe is large overall, but that budget is spread across myriad workstreams rather than focused on a single bold target. The US IRA, while highly effective, could have set clearer targets and progress metrics and relies heavily on one policy instrument (tax credits) and could thus be more balanced. The UK ISCF was short-lived and small in scale.

To build on existing frameworks and deliver proper mission based policy making – which includes ambitious generative AI deployment – we recommend a number of steps, building on Quilter-Pinner (2024).

1. **Set missions, clear sub-targets, and progress metrics, with high-level political commitment and centralised oversight.**
    - Missions must be championed at the highest level (eg prime minister or ministry heads) and overseen by a central body with the authority to pull resources from different departments.
    - Flesh out sub-targets that break down high level missions (eg improving healthy life expectancy over 10 years) into its constituent components (eg improving the food environment, or improving childhood health).
    - Ambitious timescales, budgets and outcome metrics should be clearly set and regularly monitored.

2. **Explicitly connect missions with incentives for AI deployment.**
    - Make advanced AI a core enabling technology for missions across sectors like health or the energy transition. The UK's new AI Opportunities Plan foresees this, aiming to appoint mission-focused programme AI Research Resource directors. But this is only one policy lever – more alignment across government policy will be needed.
    - Be clear about what 'complementary factors' – product inputs which interact with AI to create value – are involved (see box 2). For instance, spell out what infrastructure or skills investment is needed for mission-aligned AI deployment to be realised in practice.

3. **Funding for incentivising AI deployment in priority areas.**
    - With exceptions of the US Inflation Reduction Act and South Korea's New Deal, most mission-based policies are not backed by macroeconomically large amounts of funding. Depending on the objective, more fiscal incentives will be needed to significantly steer AI adoption. If large scale economic gains from AI are realised, this could be funded without significantly altering overall tax levels

4. **Ringfence funding for projects with uncertain returns**.
    - A significant, dedicated budget – on the scale of a DARPA-like fund – which could facilitate bold, long-term R&D that is mission aligned.
    - In this context, embrace a higher risk tolerance for generative AI research, including 'blue sky' experiments, to create breakthrough models and applications. This will be important for areas – for instance in prevention of illnesses – where new business models are yet to be established, but where potentially large gains lie.

5. **Flexible, cross-disciplinary collaboration**
    - Bring industry, academia, and civil society together in open innovation networks that cut across traditional silos.
    - Offer challenge-based funding calls specifically requiring AI-driven solutions, so generative AI becomes integral to problem solving rather than an afterthought.

In our forthcoming work, we will outline how such a set of policies could be applied in the UK.

The Paris AI Action Summit in February 2025 could add important elements of a new 'directionist' approach to AI policy. The summit will showcase concrete examples of AI applications tackling societal challenges, with a clear focus on deployment, providing inspiration for how mission-based approaches could drive directed adoption while ensuring safety. Its practical outcomes could strengthen mission oriented policies by demonstrating proven use cases and establishing shared standards for responsible AI development. While the summit will advance global collaboration, individual jurisdictions must still define and prioritise their specific AI missions via democratic processes.

BOX 2: CAN COUNTRIES WITHOUT FRONTIER AI COMPANIES MAKE ANY DIFFERENCE?

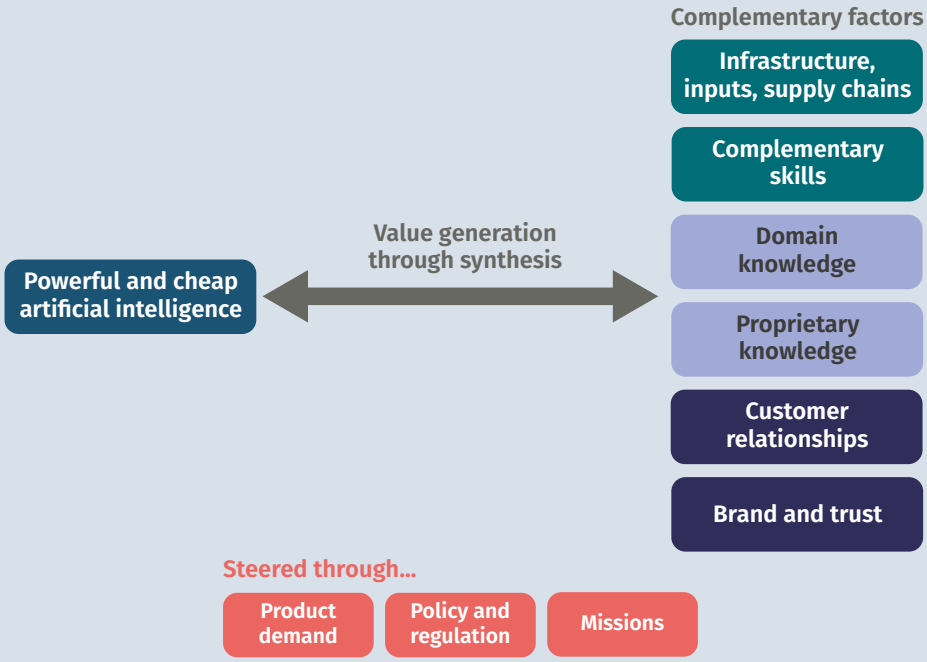
Most cutting-edge AI development is currently conducted by a handful of companies in a small number of countries. This raises the question whether countries that do not host leading AI firms have the ability to control the deployment of AI in their economies.

Value creation will come at the intersection of increasingly cheap intelligence and complementary factors. This will be businesses, individuals, third sector organisations and government applying AI tools to their existing specialisms (see figure 3.1). For instance, the UK has a global comparative advantage in the life sciences, and businesses can use this domain knowledge to deploy AI for value generation. Equally, say, Germany's manufacturing base can provide essential domain knowledge that can be leveraged by AI. Organisations that move first in applying existing production factors to cheap intelligence will be most profitable. The scarcity of these complementary factors also means that there are significant profits to be made by those who are not providing the AI.

So mission-based AI policy, that aims to steer adoption, can focus on the 'complementary factors' and underlying infrastructure. They are the bottlenecks and thus also a source of power to influence the direction of AI application. It is, other words, not just down to the AI model developers.

FIGURE 3.1

Value creation in the AI economy



Source: Author's analysis

NEXT STEPS FOR THE POLICY AND RESEARCH AGENDA

- In our forthcoming work, we will conduct analysis on what directing AI would mean with regards to different sectors, starting with prevention in health. We will use this to spell out in detail, proposals for mission-based policymaking.
- We will also develop metrics for measuring the direction of AI deployment, reflecting the need for a discussion about the future and current direction of deployment.
- We will also flesh out more concretely which policy levers are best suited for achieving 'directionism' in different contexts.
- We are highly interested in collaborating with practitioners and researchers that have experience with implementing mission-based policies and specific sectors.
- We are also interested to hear from AI labs and firms that are working on, or are interested in, tackling some of the issues that we have outlined here. In particular, we want to hear about the problems that they are facing due to the regulatory environment, financing availability and infrastructure.

REFERENCES

Andreesen M (2024) 'Why AI Will Save The World', blog post. https://pmarca.substack.com/p/why-ai-will-save-the-world

Benaich N and Chalmers I (2024) 'Improvements in AI systems and their economic impacts.'

Clifford M (2025) AI Opportunities Action Plan, Department for Science Innovation and Technology. https://assets.publishing.service.gov.uk/media/67851771f0528401055d2329/ai_opportunities_action_plan.pdf

Department for Science Innovation and Technology [DSIT] (2024) *Artificial Intelligance Sector Study 2023*. https://assets.publishing.service.gov.uk/media/671a06ca67b3ef4856faf8a8/artifical_intelligence_sector_study_2023.pdf

Dibb G, Alvis S, Gasperin S and Murphy L (2023) *Making markets in practice*, IPPR report. https://www.ippr.org/articles/making-markets-in-practice

The Clean Investment Monitor (2024) *Clean Investment Monitor: Q3 2024 Update*. https://cdn.prod.website-files.com/64e31ae6c5fd44b10ff405a7/6740efdaf61924a3fc272b75_Clean_Investment_Monitor_Q32024.pdf

Eloundou T, Manning S, Mishkin P and Rock D (2023) 'GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models', ArXiv paper. https://arxiv.org/abs/2303.10130

European Parliament (2024) *EU AI ACT: first regulation on artificial intelligence*, legislation. https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

Hall P (1991) *Cities of Tomorrow: An Intellectual History of Urban Planning and Design in the Twentieth Century*, Wiley

Health Foundation (2024) *Health inequalities in 2040*. https://www.health.org.uk/sites/default/files/upload/publications/2024/Health inequalities in 2040.pdf

IPPR (2024) *Our greatest asset – the final report of the IPPR commission on health and prosperity*. https://www.ippr.org/articles/our-greatest-asset

Ji X et al (2023) 'AI Alignment: A Comprehensive Survey', ArXiv paper. https://arxiv.org/pdf/2310.19852

Juhász R, Lane N and Rodrik D (2023) *The economics of industrial policy*, NBER working paper. https://www.nber.org/papers/w31538

Jung C and Srinivasa Desikan B (2024) *Transformed by AI – How generative AI could affect work in the UK – and how to manage it*, IPPR. https://www.ippr.org/articles/transformed-by-ai

Lund S et al (2022) *Great expectations: the promises and limits of innovation policy in addressing societal challenges*, working paper. https://swopec.hhs.se/lucirc/abs/lucirc2022_009.htm

Mazzucato M (2013) *The Entrepreneurial State – Debunking public vs private sector myths*, Anthem Press

McShane C (1994) *Down the Asphalt Path: The transformative impact of cars on urban spaces*, Columbia University Press

MIT Technology Review [MIT TR] (2024) 'We need to prepare for 'additictive intelligence', news article. https://www.technologyreview.com/2024/08/05/1095600/we-need-to-prepare-for-addictive-intelligence/

Modei D (2024) 'Machines of loving grace', blog. https://darioamodei.com/machines-of-loving-grace

Mollick E (2024) 'What just happened – a transformative month rewrites the capabilities of AI', newsletter. https://www.oneusefulthing.org/p/what-just-happened

New York Times [NYT] (2024) 'Can A.I. Be Blamed for a Teen's Suicide?', news article. https://www.nytimes.com/2024/10/23/technology/characterai-lawsuit-teen-suicide.html

Organisation for Economic Co-operation and Development [OECD] (2021) *The design and implementation of mission-oriented innovation policies*. https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/02/the-design-and-implementation-of-mission-oriented-innovation-policies_cb8908f7/3f6c76a4-en.pdf

OpenAI (2024) 'Open AI's approach to AI and national security', blog post. https://openai.com/global-affairs/openais-approach-to-ai-and-national-security/

Quilter-Pinner H (2024) *Mission-driven government*, IPPR. https://www.ippr.org/articles/mission-driven-government

Rathi R, Taylor S, Bergen B and Jones C (2024) 'GPT-4 is judged more human than humans in displaced and inverted Turing tests', ArXiv paper. https://arxiv.org/pdf/2407.08853

Rand and Frontier Economics (2024) *Interim Impact Evaluation of the Industrial Strategy Challenge Fund*. https://www.ukri.org/wp-content/uploads/2024/09/IUK-270924-ISCF-interim-impact-report_FINAL.pdf

Srinivasa Desikan B and Jung C (forthcoming) *The direction of UK AI Innovation*, IPPR

Institute for Public Policy Research

GET IN TOUCH

For more information about the Institute for Public Policy Research, please go to www.ippr.org

You can also e-mail info@ippr.org or find us on X @ippr

Institute for Public Policy Research
Registered Charity no. 800065 (England & Wales),
SC046557 (Scotland), Company no, 2292601 (England & Wales)

The progressive policy think tank