JULY 2024
ISABEL ATKINSON
JAMES BROWNE

The Potential Impact of AI on the *Public-Sector Workforce*

A Companion to *The Economic Case for Reimagining the State*

TONY BLAIR INSTITUTE FOR GLOBAL CHANGE

Contents

01

Executive Summary

The public sector in the United Kingdom is a major employer with nearly 6 million employees, or 5 million full-time equivalents,[1] and had an annual wage bill of £240 billion in 2022–23[2] or almost 10 per cent of GDP.

Public-sector productivity has been flat for more than a quarter of a century,[3] but there is now scope for artificial intelligence to radically reshape the way that the public sector operates and to deliver large improvements. This is particularly the case for administrative, back-office and policy functions that largely involve cognitive tasks.

Our analysis in this paper examines the impact AI could have on public-sector productivity by estimating how much time AI could save workers and the associated cost savings. We then examine how quickly AI could be rolled out across the public sector and the associated costs, leading to a comprehensive estimate of the net savings expected from the use of AI in the public sector. Our results indicate that:

- More than 40 per cent of tasks performed by public-sector workers could be partly automated by a combination of AI-based software, for example machine-learning models and large-language models, and AI-enabled hardware, ranging from AI-enabled sensors to advanced robotics.

- These efficiency gains could help save a fifth of public-sector workers' time in aggregate. These potential time savings could lead to significant fiscal savings if the government chose to bank them by reducing the size of the public-sector workforce accordingly.

- However, not all these time savings will lead to a lower wage bill and fewer public-sector workers. Many public-sector professions already face severe staff shortages, with workers doing significant amounts of unpaid overtime to keep the system afloat. We assume that the benefits of AI in these stretched professions (which account for 2.3 million workers and include maths, science and language teachers, doctors, nurses and care workers[4]) do not lead to job cuts, but instead allow frontline workers to work fewer unpaid overtime hours and deliver better outcomes. Excluding these cases from the analysis reduces the overall potential time savings from AI from one-fifth to one-sixth, but still implies savings of £41 billion a year to the public-sector wage bill (1.5 per cent[5] of GDP) if AI were used to its fullest possible extent.

- To achieve these gains, the government will need to invest in AI technology, upgrade its data systems, train its workforce to use the new tools and cover any redundancy costs associated with early exits from the workforce. Under an ambitious rollout scheme, we estimate these costs equate to £4 billion per year on average over this parliamentary term (0.15 per cent of GDP) and £7 billion per year over the next (0.24 per cent of GDP). In the long run, we estimate annual running costs of about £4 billion in today's terms. This implies the net savings from fully utilising AI in the public sector to be nearly 1.3 per cent of GDP each year, equivalent to £37 billion a year in today's terms. This equates to a benefit-cost ratio of 9:1 in aggregate and, since the setup costs of the programme are relatively small, the net benefit is positive almost straight away. Indeed, even after five years we estimate the programme could cumulatively save 0.5 per cent of annual GDP (or £15 billion in today's terms), implying a benefit-cost ratio of 1.8:1 is possible if the technology is rolled out quickly.

The speed with which these gains are realised is within the government's gift to determine. Government IT projects typically take nearly four years to complete, but these projects tend to be more targeted. Deploying AI across the entire public sector would present a bigger delivery challenge. We set out an ambitious scenario whereby the rollout of AI across the public sector is largely completed within two parliamentary terms. TBI's recent paper, *Governing in the Age of AI: A New Model to Transform the State*, sets out a plan for the new government to meet this timetable.

The government will have a choice on how to spend any dividend from AI-enabled efficiency. It could choose to reinvest the savings in the public sector and boost the number of frontline public-sector workers. For example, a saving of 1 per cent of GDP would be enough to boost the size of the NHS workforce by around a third.[6] Alternatively, the government could choose to shrink the UK's public-sector workforce by around a sixth (equivalent to around a million workers) and bank the fiscal savings. If it were to take this approach, it could lower the explosive path for debt shown in the paper *The Economic Case for Reimagining the State* as a share of GDP by 27 percentage points by 2050. Overall, the impact on annual public-sector net borrowing would be 2.4 per cent of GDP by 2050, including savings from lower debt interest.

All the above figures are based on a snapshot of AI's potential capabilities as they are today, but the technology is advancing quickly. We therefore also explore a scenario where AI's capabilities continue to advance. If AI advances such that it becomes possible to save a further 1 per cent of public-sector workers' time each year, then we estimate that the annual net cost savings could reach 1.9 per cent of GDP by 2050 under similar rollout-speed assumptions. The size of the prize could grow further in the years to come.

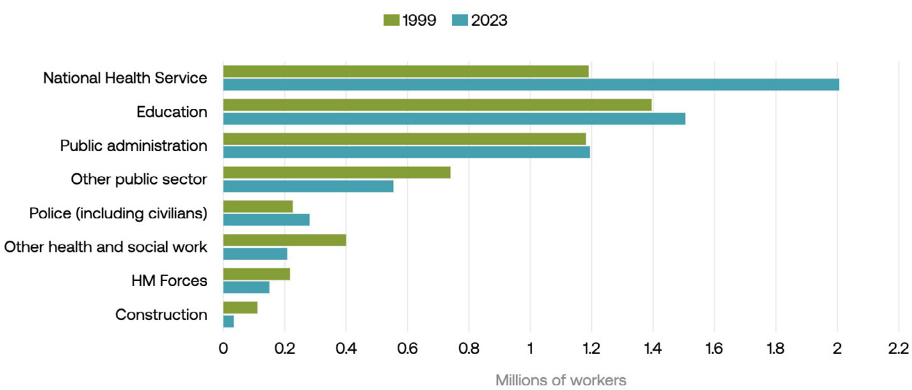
02

The Potential Impact of AI on Public-Sector Productivity

The public workforce has grown by half a million people over the past 25 years, with the majority of workers employed in health, education and public administration (Figure 1). Public-service productivity, however, remains stagnant, with essentially no growth[7] since official statistics began in 1997.[8] All growth in the output of the public sector can be accounted for by greater inputs. To some extent this is understandable; Baumol's cost disease[9] predicts that productivity in service industries will lag behind other sectors since, until now at least, they have shown little capacity for automation.

FIGURE 1

The public-sector workforce has grown by 500,000 workers since 1999



Source: ONS (Figures are taken from the ONS public-sector employment by industry headcount figure provided for December of each year.)

New AI tools have the potential to change this trend. AI is well suited to performing the kind of routine cognitive tasks that take up a large share of workers' time. Indeed, AI is already starting to have an impact in both the public and private sectors. For example, Walmart has implemented Ask Sam, a generative AI tool that reduces the amount of time it takes shop assistants to locate products by two-thirds,[10] indicating a productivity gain of over 200 per cent on this task. Experimental evidence suggests that customer-service agents can resolve 14 per cent[11] more customer issues per hour using generative AI. For less experienced and lower-performing workers, this productivity boost can be up to 34 per cent.[12] Professional writing tasks can be completed 37 per cent faster[13] and at a higher quality using generative AI, indicating a productivity increase of 59 per cent.[14] Similarly, economic researchers can become significantly more productive[15] if they use large-language models (LLMs) to automate microtasks such as formatting references, synthesising text and translating code.[16] In health care, the Hywel Dda University Health Board in Wales took six months to implement AI-powered tools that reduced delayed discharges and freed up 3,000 hospital bed days annually.[17]

These examples show the potential of AI to improve public-sector productivity in particular tasks and occupations. In what follows, we take a broader approach to examine the potential for AI to improve productivity across all the different tasks performed by public-sector workers and how this could reduce the cost of providing public services.

Methodological Approach

To estimate the potential for time savings for public-sector workers using AI tools, we need to estimate how much time can be saved on each task performed by these workers. We begin with information on nearly 20,000 tasks associated with different occupations from the O*NET database.[18] To identify which of these could be performed by AI, there are several approaches we could take. One approach, used in other research in this area,[19] is to identify broad groups of tasks that are well suited to being performed by AI. However, we would ideally want to assess the ability of AI to perform each individual task, drawing on a wide range of sources on AI's current capabilities. But relying on expert judgements to make these individual decisions would make our analysis intractable. Fortunately, this is exactly the sort of task to which LLMs, trained on a vast amount of source material, are well suited. Another approach that has been used by researchers in this area is to use OpenAI's GPT-4 to determine whether each task in the O*NET database can be done using AI.[20] But given the nature of these models – LLMs are not deterministic[21] and it is not possible to trace how they have sourced information – there is a concern that these models are a mysterious "black box" that may or may not give reliable results.

We therefore use a hybrid of the two approaches. We first use GPT-4 to categorise each of the 19,281 tasks in the O*NET database in several different respects that we consider to be important determinants of whether the task can be performed by AI or not. These were chosen following an initial analysis of GPT-4's unguided assessment of the automatability of some sample tasks, in which it struggled with some assessments. This categorisation enables us to generate a prompt to GPT-4 that contains an initial assessment as to whether it is likely that the task can or cannot be performed by AI. In the final stage, we use GPT-4 to categorise the type of AI tool that could be used to perform the task, estimate the time saving and give an assessment of whether it would be cost effective to deploy the tool. We then merge this information with UK Labour Force Survey data to calculate overall time savings given the numbers in each occupation group. A fuller description of the methodology is given in a technical annex.

Results

We find that around 40 per cent of tasks performed by public-sector workers could be done by AI. However, relatively few of these tasks could be fully replaced by AI, meaning some worker involvement would still be necessary in almost all cases. Taking our estimated time savings from using these tools into account, our analysis suggests that 19.8 per cent of public-sector workers' time could be saved using AI tools.[22]

In order to understand how these time savings could be realised in practice, we differentiate between five types of AI technology in our analysis – three consisting of different kinds of AI software to perform cognitive tasks and two which combine AI software with hardware to perform manual tasks. The three software categories are:

- **"Free" AI tools**, including products such as Google's Gemini and OpenAI's ChatGPT and Dall-E, which primarily work by providing users with access to powerful LLMs to answer their prompts. Use cases for these general-purpose tools could include helping teachers produce educational materials, civil servants write reports or police officers transform notes into reports.

- **Low-cost AI tools** which include various kinds of generic professional software such as virtual assistants, like Microsoft Copilot, and payroll-management systems. For example, schools or medical secretaries could use scheduling assistants to make appointments, while procurement professionals could use AI systems to prepare and track invoices.

- **Bespoke or internally trained AI systems** which require training on internal public-sector data to unlock time savings. These could help speed up benefit eligibility checks, monitor bed space in hospitals or answer public queries.

An important caveat regarding the use of these tools within the public sector is that the use of completely “free” models available online may not be suitable or appropriate, given security and privacy concerns. This may therefore require investment in building alternatives within a secure environment. The Generative AI Framework for HMG[23] published in January 2024 sets out many of these concerns in detail and suggests that the use of public generative AI Application Programming Interfaces to build tools, rather than the freely available tools themselves, may be more appropriate. Throughout, we will continue to refer to “free” AI tools, however this refers to a style of tool which, while being low cost, will not actually be totally free for the public sector to implement and use.

The two hardware categories are:

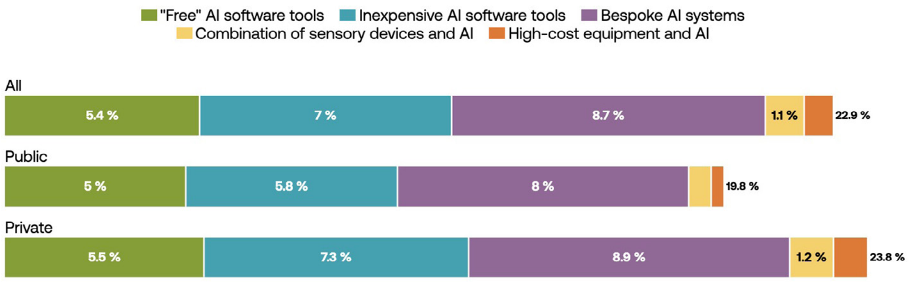
- **AI-enabled low-cost sensory devices** which include elements such as cameras, thermometers, speakers, microphones and radio-frequency identification (RFID) tags. Combining relatively basic AI with these devices could unlock significant time savings, allowing legal depositions to be recorded and directly transcribed, hospital technicians to more easily monitor the status and movement of equipment, and parking enforcement officers to use cameras combined with AI to more rapidly produce assessments and calculate fines.

- **AI-enabled high-cost equipment** which includes a range of more complex and capital-intensive technologies, from medical scanners and surveillance systems to autonomous robots and self-driving vehicles. In clinical settings, AI-enabled medical scanners have the potential to introduce substantial time savings for medical specialists by enabling them to read scans more accurately and efficiently.[24] Specialised drones[25] have the potential to monitor and detect hazards for environmental professionals and firefighters, enabling them to safely cover much larger areas in much less time. At the frontiers of current technology, robotic AI could be used in waste management[26] to sort and break up refuse or in the form of autonomous vehicles to provide transport services. We find that many occupations could benefit at the margin from accessing these technologies, but given the high investment cost, this would not be economically viable unless the tools were being used intensively. In the analysis that follows, we therefore exclude any time savings from these technologies for occupations where the associated time saving is less than 10 per cent.

Our findings suggest that the vast majority of time savings available from AI today could be unlocked at relatively low cost using software-based solutions, with high-cost hardware accounting for only 0.3 per cent of the potential 19.8 per cent time savings offered by AI (Figure 2). These time savings are substantial but tend to be smaller than in the private sector, as capacity for AI to save time for frontline public-service workers remains limited.

FIGURE 2

Around one-fifth of public-sector time could be saved using AI

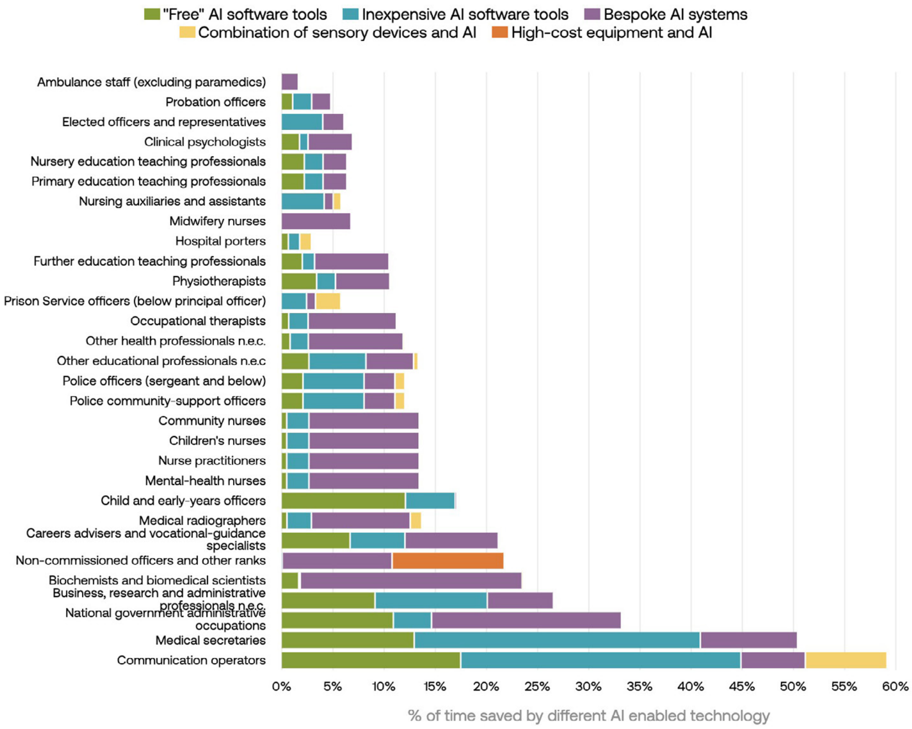


Source: TBI calculations using O*Net and Labour Force Survey data

Within the public sector, the time savings for different workers vary substantially by profession (Figure 3). For example, AI could save less than 5 per cent of ambulance staff time, whereas medical secretaries could do their work in less than half the time if AI were used to the maximum possible extent. “Free” and inexpensive AI tools alone could save more than 40 per cent of time for medical secretaries, for example by helping to schedule tests or procedures for patients and transcribing patients’ medical information to their records.

FIGURE 3

The ability of AI to deliver time savings varies substantially across public-sector occupations

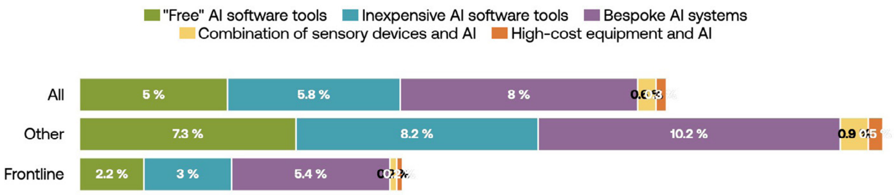


Source: TBI calculations using O*NET and Labour Force Survey data
Note: N.E.C – not elsewhere classified

Looking more broadly across the different functions of government, we find that the biggest time savings are in administrative functions rather than frontline public services (Figure 4). Splitting out frontline roles from other back-office staff,[27] we can see that back-office staff see time savings of around 27 per cent across all public-service functions, whereas frontline staff see time savings of only 11 per cent on average.

FIGURE 4

Back-office functions see larger time savings from AI



Source: TBI calculations using O*NET and Labour Force Survey data

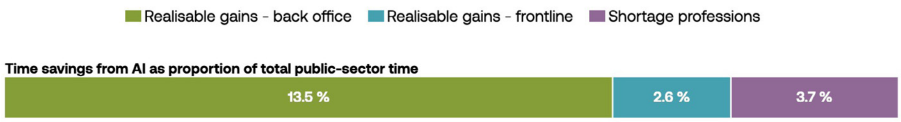
How Could AI Time Savings Affect the Public-Sector Wage Bill?

Not all potential time savings from AI could be realised in practice. Many frontline public-sector occupations currently face recruitment difficulties and shortages, and involve existing workers doing substantial unpaid overtime.[28] Indeed, a wide range of frontline public-sector professions – including almost all health-care professionals, care workers and secondary-school teachers of maths, science or foreign languages – were on the list of shortage occupations for immigration purposes until recently.[29]

Any AI-related time savings for these professions are more likely to address the existing worker shortfall than lead to frontline job cuts. We can identify those with the appropriate Standard Occupational Classification (SOC) codes in our Labour Force Survey data. If we exclude these shortage professions from the analysis, then the total potential time savings from AI in the public sector fall further from 19.8 per cent to 16.1 per cent. Mapping these time savings onto the public-sector wage bill implies a maximum possible cost saving of £41 billion a year in today's terms.[30]

FIGURE 5

AI could save one-sixth of public-sector workforce time once shortage professions are excluded



Source: TBI calculations using O*NET and Labour Force Survey data

03

The Potential Pace of AI Rollout

The speed with which these gains are realised will be up to the government to determine. The history of large-scale IT projects in the public sector is decidedly mixed. Politicians are still scarred by the memory of big failures such as the NHS computer-system upgrade that took nearly a decade to complete and cost £10 billion.[31] Generally, public-sector IT projects take nearly four years[32] to complete, compared with 2.4 years in the private sector. However, there have also been many success stories in recent years. For example, the EU Settlement Scheme service processed 5.7 million[33] successful applications by the end of 2023. The digitalisation of the passport office, which was able to clear the backlog of applications from the Covid-19 pandemic relatively quickly, is another success.[34] Even the much-maligned universal-credit system eventually proved its worth during the pandemic, giving the government the flexibility to increase benefit levels when additional support was needed.

The fragmented nature of the public sector – there are around 24,500 schools,[35] more than 6,000 GP surgeries[36] and more than 300 local authorities[37] in England alone, all managed separately – adds an additional challenge for rolling out AI tools widely.

For these reasons the recent TBI paper *Governing in the Age of AI: A New Model to Transform the State* advocated for the coordination of AI rollout at the centre of government, including the establishment of an AI Mission Control with regular reporting to the prime minister. This would also require a team of AI specialists in every government department, paid in line with market rates, to identify use cases and drive the development of new AI tools, in partnership with other organisations.[38]

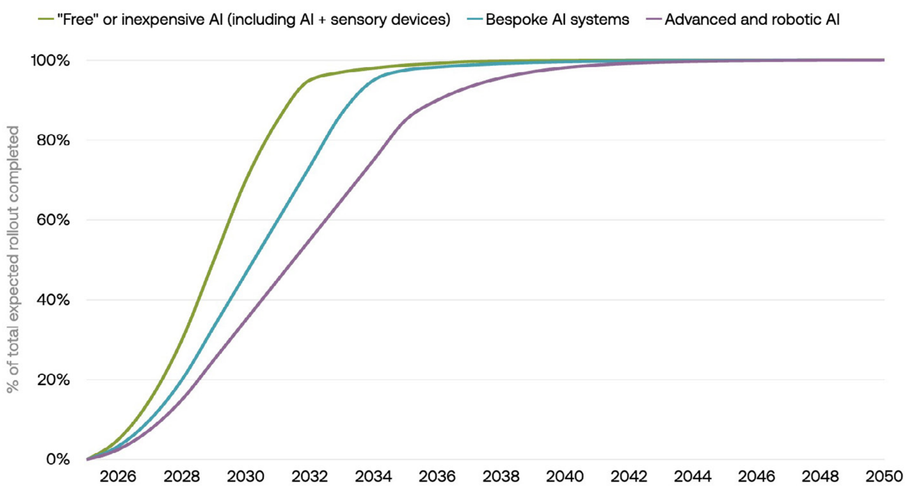
The paper argued that transformative gains from implementing AI in government could be achieved within five years. We take a slightly more conservative, but still very ambitious approach in our analysis. We allow one year for the necessary data-system upgrades and upfront investments in LLMs to be made, and then assume that for off-the-shelf AI software it is possible to move from 5 per cent to 95 per cent coverage in six years.

We assume bespoke AI software takes an additional two years to develop before being implemented and that more complicated AI-enabled hardware takes a further three years to develop and roll out (Figure 6). As a result, the rollout is largely completed within two parliamentary terms.

We also examine a more cautious scenario where rollout takes longer. In this scenario, we allow two years for the preparatory work to take place and then a decade for off-the-shelf AI software tools to reach 90 per cent penetration across the public sector. Again, bespoke AI software and AI-enabled hardware take a further two and five years respectively to reach this level of penetration.[39]

FIGURE 6

Assumed speed of AI adoption across technologies

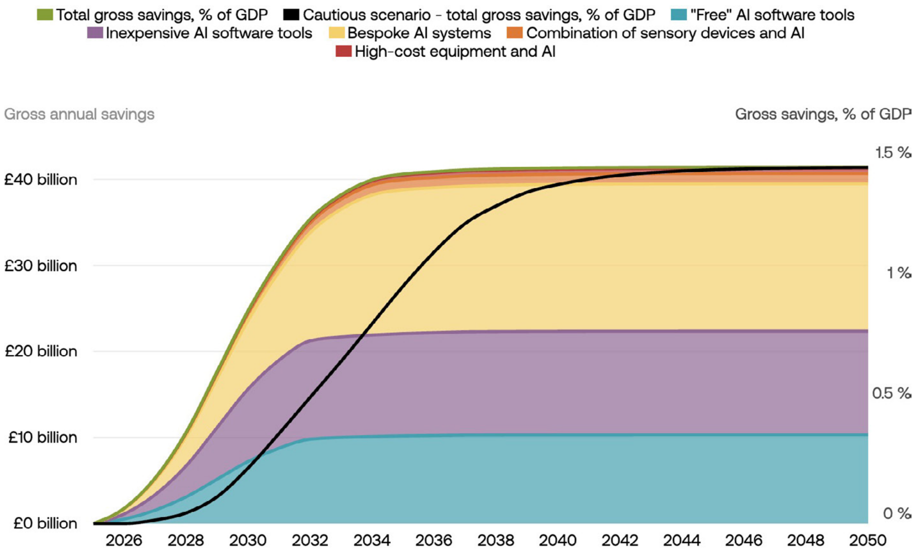


Source: TBI Analysis

Figure 7 maps these rollout assumptions onto potential cost savings. After five years, AI could achieve £18 billion in gross savings per year (0.6 per cent of GDP). These will be mainly unlocked using the cheapest forms of the technology ("free" AI software, low-cost AI software and low-cost AI-enabled sensory devices). By the mid 2030s, we expect the vast majority of gains from AI to have been achieved across all technology categories. Total gross savings are £40 billion a year by 2035 (1.45 per cent of GDP). Under our more cautious scenario, this scale of gross savings is not achieved until 2040. Then by 2050, under both scenarios gross annual savings stabilise at £41 billion a year (1.5 per cent of GDP) as the remaining gains from utilising high-cost AI-enabled hardware are realised.

FIGURE 7

Implications of assumed rollout speed for time taken to achieve public-sector cost savings



Source: TBI calculations using O*NET data and Labour Force Survey data

03 Costs Associated With AI Adoption in the Public Sector

Achieving a one-sixth reduction in the public-sector wage bill will not be cost free. It will require:

1. Upfront investment in public-sector **data systems** to ensure that they are interoperable and can be used to train new AI tools.

2. Ongoing investment in **AI talent** within government via a substantive new team of AI experts responsible for identifying new AI, developing new AI tools and overseeing their rollout.

3. Ongoing investment in **AI technologies**, including the cost of developing and deploying AI tools within government and paying to access private-sector tools.

4. Ongoing investment in **training** to ensure existing public-sector workers are able to fully utilise AI and realise time savings.

5. One-off outlays to cover the **costs of redundancies** if the speed of rollout is sufficiently fast that workforce reductions cannot be achieved by normal staff turnover alone.

In this section we run through each of these costs and link them to the pace of rollout assumed in the previous section to come up with the overall costs of the programme.

Data-System Upgrades

The public sector currently lacks common data standards, with key information often stored in archaic formats. Upfront investment is required to upgrade these data sets to ensure they are usable in the AI era. As noted in previous TBI analysis,[40] schemes to upgrade data in government departments typically cost £75-200 million for large departments with high volumes of citizen transactions and around £15 million for more administrative data sets. Aggregating these figures across departments suggests a **one-off cost of up to £2.5 billion**, payable in the first few years of the rollout programme.

AI Talent

As TBI has previously argued, an effective AI-implementation programme will require a significant investment in AI talent across government. This would include a new team of up to 100 individuals located at the centre of government to set strategy and coordinate AI activity across government functions. In addition, each government department would need specialised teams of AI experts responsible for identifying specific use cases and deploying AI tools to support them. These teams would both develop tools internally and work in partnership with private-sector providers to integrate existing tools into government systems. Leading AI adopters in the private sector such as JPMorganChase have more than 50,000 tech specialists[41] and 2,000 AI, machine-learning or data-science experts.[42] Currently, the civil service employs almost 26,500 people[43] in digital, data and technology functions. If the government were to match JPMorganChase's ratio of AI employees to other tech employees, this would imply an additional 1,000 AI specialists spread across departments (or roughly 50 AI specialists per department). These teams would need to be paid at a level that is at least comparable to the private sector.[44] Tech salaries in the UK tend to be significantly lower than the market-leading rates offered in the United States, so if we benchmarked at 50 per cent of the median market rate[45] in the US, that would equate to an average reward package of £100,000 per person. This would mean the ongoing costs of a 1,100-strong AI workforce (central team plus AI specialists) would be around **£110 million per year** in today's prices.

Investing in AI Technology

The cost of deploying AI across the public sector will depend on the type of technology involved. Here we consider six different variants of technology and analyse how much each could cost to set up and run. Given AI is still at an early phase of commercialisation, we draw on data from a wide variety of sources to estimate the costs of different elements of the technology, including insights generated by ChatGPT that have been cross checked against other sources for accuracy.

CATEGORY 1: GOVERNMENT VERSIONS OF “FREE” AI SOFTWARE TOOLS

Our workforce analysis earlier in this report indicates that around 87 per cent of the public-sector workforce – around 5.2 million people – could make use of the “free” AI tools that are now widely available online.

- **Upfront costs**: As described in previous TBI work,[46] we estimate the government would need to invest **£50 million upfront** to develop its own multifunctional LLM. This would provide public-sector workers with access to the kind of AI tools freely available online but in a secure environment and trained on official public-sector data.

- **Ongoing costs**: We follow standard industry practice and assume the LLM models would incur annual running costs of just over £5 million per year (10 per cent of the setup costs) to maintain and update. In addition, we assume each public-sector worker that uses the technology incurs a small cost related to the use of computer hardware (including the cost of energy used to power the hardware), as well as data-storage and transfer costs. We conservatively estimate the annual cost is around £30 per user or £155 million per year across all applicable workers.[47] Taken together, this means the annual running cost of this technology would be around **£160 million per year**.

CATEGORY 2: OFF-THE-SHELF PRIVATE-SECTOR AI SOFTWARE

Around 5.6 million public-sector workers could benefit from accessing off-the-shelf AI software tools (such as co-pilots,[48] transcription services[49] or scheduling assistants[50]) that private-sector firms already offer.

- **Upfront costs**: We assume the government will need to invest around **£15 million upfront** in compliance security checks to ensure any off-the-shelf tools are secure, reliable and stable.[51]

- **Ongoing costs**: We assume the government would continue to incur £15 million per year in costs to maintain existing tools and to keep verifying new tools as they emerge. We also assume a per-user cost of accessing these tools. Currently, the market price for existing AI productivity tools ranges from £150–300 per tool per year.[52] We assume a cost of £300 per user per year, on the basis that competition and bulk-purchase discounts will bring down the cost per tool, but that some public-sector workers may need access to multiple tools to achieve the full efficiency gains. This equates to £1,684 million per year in ongoing costs across all applicable users. Adding these user fees to the ongoing cost of verification implies an ongoing cost of **£1,699 million per year**.

CATEGORY 3: GOVERNMENT-TRAINED BESPOKE AI SYSTEMS

Around 5.5 million public-sector workers could benefit from the use of bespoke AI systems to perform specific tasks.

- **Upfront costs**: Our workforce analysis earlier in this report indicates that there are 577 separate tasks with which bespoke trained AI systems could be of assistance. We conservatively assume that the government builds one bespoke AI tool for each of these narrow use cases and that developing each tool costs up to £2.5 million (consistent with the cost of developing a "legal advisor" LLM in our previous analysis[53]). This implies the government would need to invest almost **£1,450 million upfront** to develop this range of tools. We assume these tools are built gradually at the same pace that AI is rolled out across the public sector.

- **Ongoing costs**: We assume that once a tool is developed it costs 10 per cent of its initial investment to maintain and update each year, equivalent to around £145 million in ongoing running costs once fully rolled out. We also assume the same £30-per-user-per-year cost as in Category 1 (to pay for access to computer hardware, data storage and data transfer), which equates to around £168 million a year when scaled up to all eligible workers. This implies a total ongoing cost of **£312 million per year**.

CATEGORY 4: SENSORY HARDWARE AND AI SOFTWARE

Around 2.1 million public-sector workers perform tasks that could benefit from relatively cheap sensory hardware linked to AI software to help perform manual tasks.

- **Upfront costs**: Simple sensory hardware such as headsets,[54] cameras[55] and maintenance sensors[56] typically costs £20–100 per device. On average we assume that each user will require several of these devices to achieve the time savings on offer, equivalent to £250 per user or **£530 million upfront** when fully rolled out across applicable users. We assume these tools are purchased gradually at the same pace that AI is rolled out across the public sector.

- **Ongoing costs**: We assume that each device needs replacing every 10 years, implying an ongoing capital cost of £53 million per year once fully rolled out. We also assume the ongoing cost of linking the sensory devices to AI software would be roughly comparable to the same access costs in Categories 1 and 3 above; £30 per user per year. This equates to around £64 million a year when scaled up to all eligible workers or **£117 million per year** when the ongoing capital cost of replacing devices is included.

CATEGORY 5: AI-ENABLED HIGH-COST EQUIPMENT AND ROBOTICS

Around 150,000 public-sector workers perform manual tasks where advanced AI-enabled electronic equipment could save at least 10 per cent of their time. This equipment, which includes autonomous drones, AI-enhanced medical-imaging scanners, predictive maintenance systems and many other forms of hardware, varies significantly in cost and its potential to replace workers.

Given the wide range of potential tools, we use ChatGPT to explore a selection of use cases to estimate potential cost savings.[57] On average, these use cases imply that AI-enabled hardware would cost around 75 per cent of the equivalent wage rate to perform a given task. Given that eligible workers could save 13 per cent of their time on average using this hardware, this implies an annual equipment cost equivalent to 10 per cent of the wage rate per worker (£30,400), or £3,000 per year. We estimate the lifespan of this equipment is around a decade on average, which implies a total equipment cost of £30,000 over 10 years. We assume that 95 per cent of this (£28,500) is upfront capital expenditure to pay for the machinery itself (consistent with ChatGPT's estimate of the operating costs of running an autonomous drone, including electricity, cloud-computing costs, data storage and transfer). Scaled up across the eligible workforce, this implies the government would need to invest around **£4.3 billion upfront** in the technology and then **£452 million a year** to cover operating costs and annual depreciation.

TABLE 1

Summary of upfront and ongoing technology costs in today's prices

	Upfront cost of technology during rollout	Annual running costs when fully deployed
Government versions of "free" AI software tools	£50m	£160m
Off-the-shelf private-sector AI software	£10m	£1,699m
Government-trained bespoke AI systems	£1,443m	£312m
Sensory devices and AI software	£530m	£117m
AI-enabled high-cost equipment or robotics	£4,293m	£452m
Total	**£6,326m**	**£2,740m**

Source: TBI analysis

Training Costs

To realise the efficiency gains from AI, virtually all public-sector workers will need to receive adequate training in how to use it. In our rollout model, we assume that one in 50 public-sector workers becomes an "AI ambassador". Each ambassador would be given a five-day bootcamp of full-time training to understand the capabilities of AI. These ambassadors are then responsible for disseminating best practice to their colleagues.[58] We also assume that all public-sector workers receive 0.5 days of training in AI technology each year to further enable them to use AI tools in their day-to-day work. In each case, we assume an average training cost of £464 per day, in line with the average cost of training across the health, education and public-administration sectors in the 2022 Employer Skills Survey,[59] uprated to 2024 terms in line with growth in nominal GDP. This implies a one-off **upfront cost of £275m** to train the AI ambassadors, plus an ongoing training cost of around **£1.2 billion per year** to cover the half-day's training for virtually every public-sector worker plus the costs of training new AI ambassadors (we assume a turnover rate of 10 per cent a year).

Redundancy Costs

If the government chooses to reduce staff numbers as a result of a more efficient AI-enabled public-sector workforce, this may require redundancies. Those made redundant will be entitled to redundancy payments. Under our rollout assumptions, we estimate that these will peak at around £4.1 billion a year in today's terms.

In principle, any job losses associated with AI adoption could be managed through natural wastage. Under our rollout assumptions, the highest annual redundancy rate is 2 per cent of the workforce, or approximately 116,000 in a single year. This compares with an estimated turnover rate of 11 per cent of public-sector workers per year, based on longitudinal data from the Labour Force Survey. However, there is significant variation across different parts of the public sector. Only 2.5 per cent of workers in health care move out of the sector from one year to the next, compared with nearly a quarter of workers in social care. Moreover, rather than the rollout happening at the same rate in all parts of the public sector, it is more likely that a larger group of workers in one particular occupation and region will no longer be required in one year, and another the following year. To be conservative in our assumptions about the overall net savings from AI, we therefore assume that all reductions in public-sector headcount are achieved through redundancy rather than natural wastage.

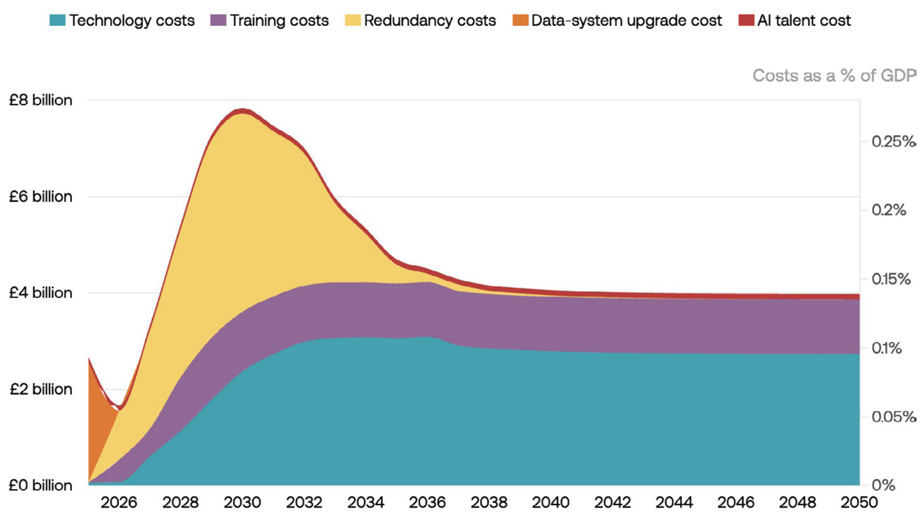
We further assume that all displaced workers will receive redundancy payments on civil-service terms[60] – that is, one month's pay for every year of service up to a maximum of 12. Using data on job tenure of public-sector workers from the Labour Force Survey, we estimate an average entitlement of seven months' redundancy pay. Under these assumptions, there is a total of 1.15 million redundancies. Each person made redundant is entitled to a severance package of just under £21,000, so total redundancy costs are £24 billion in today's terms. These are both spread over a period of 20 years or more. Redundancy costs peak at £4.1 billion in a single year in today's terms.

Cost Summary

Overall, if we aggregate the costs across all five categories highlighted above, we find that the average annual cost of rolling out AI to the public sector is around £4 billion per year in today's prices, or 0.14 per cent of GDP (Figure 8). Redundancy costs cause this figure to edge up slightly in the early 2030s, with costs peaking at nearly £8 billion per year, before falling back.

FIGURE 8

Total costs of rolling out AI across the public sector amount to 0.1 per cent of GDP per year when fully deployed



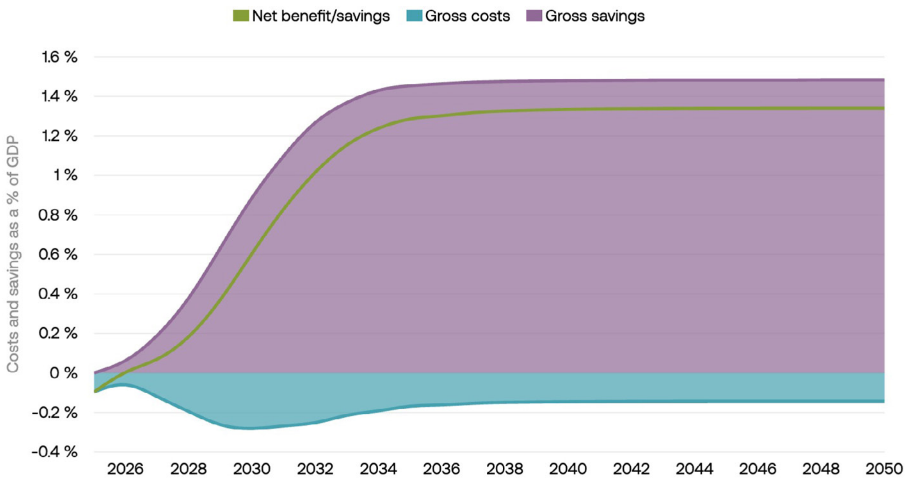
Source: TBI calculations using O*NET and Labour Force Survey data

04 The Net Benefits of Introducing AI in the Public Sector

Drawing the estimated time and cost savings, rollout profile and costs together, we see that the overall net savings from introducing AI in the public sector are substantial. By the middle of this century, the programme could save 1.3 per cent of GDP a year or the equivalent of £37 billion in today's terms (Figure 9). Moreover, because the benefits from AI begin to accrue quickly, notable savings are possible in the short term too. The programme could save 0.5 per cent of annual GDP cumulatively over its first five years – equivalent to £15 billion in today's terms – or a benefit-cost ratio of 1.8:1. This is a lower five-year benefit-cost ratio than that of several of the public-sector productivity programmes announced in the 2024 Budget,[61] demonstrating that these figures, though impressive, are by no means outside the realms of possibility.

FIGURE 9

Direct savings from AI far outweigh the costs of implementation

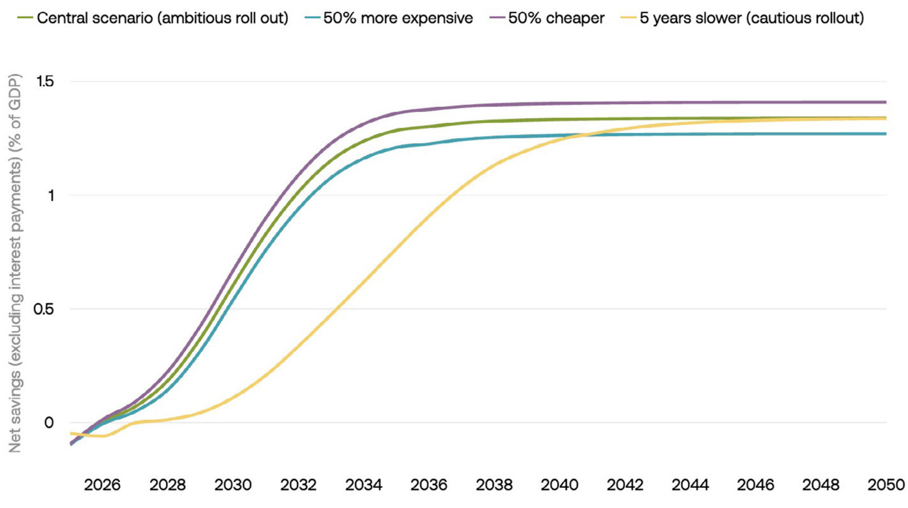


Source: TBI calculations using O*NET and Labour Force Survey data

These estimates are sensitive, though not inordinately so, to our assumptions about the scale of cost savings and the speed of the rollout (Figure 10). Increasing technology and training costs by 50 per cent reduces the long-term net benefit of the programme from 1.3 per cent of GDP each year to 1.27 per cent of GDP (or net savings of £35 billion a year instead of £37 billion in today's terms). And as we have seen, the assumed rollout speed also affects when savings occur. If we assume that it takes 15 years rather than ten to fully roll out the technology in our cautious scenario, then the programme unsurprisingly takes longer to generate material savings. With a slower rollout speed, the cumulative net benefit from the programme for the first ten years would be 1.1 per cent of GDP instead of 1.7 per cent of GDP (or cumulative savings of £31 billion instead of £47 billion).

FIGURE 10

The net benefits from using AI in the public sector are only somewhat sensitive to assumptions around the cost and speed of rollout

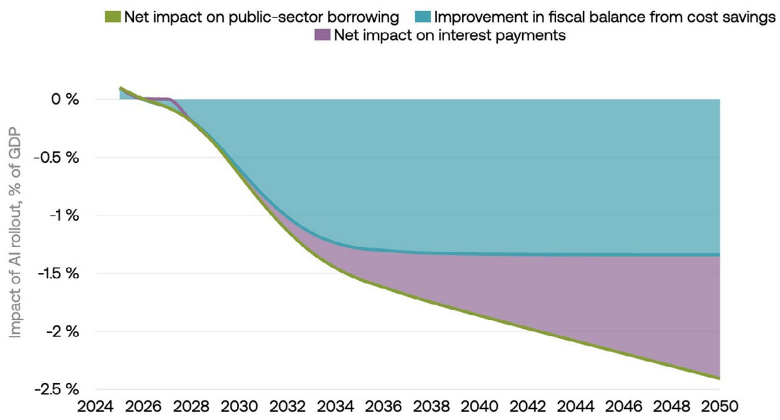


Source: TBI calculations using O*NET and Labour Force Survey data

There are also indirect benefits to the public finances from the rollout of AI. If other spending items and tax receipts are left unchanged, the explosive debt-to-GDP profile shown in *The Economic Case for Reimagining the State* is 27 percentage points lower in our ambitious rollout scenario by 2050. As a result, annual debt interest payments are also lower, to the tune of 1.1 per cent of GDP. The overall impact on public-sector net borrowing is therefore 2.4 per cent of GDP by 2050.

FIGURE 11

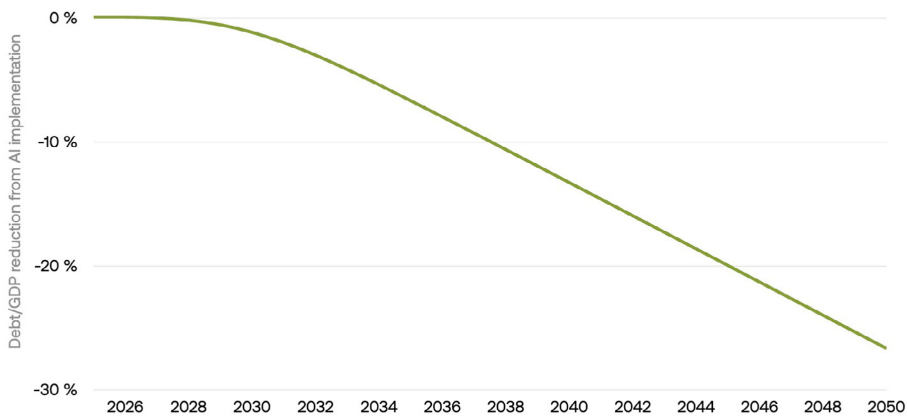
Adopting AI in the public sector could reduce net borrowing by 2.4 per cent per year by mid-century



Source: TBI calculations using O*NET and Labour Force Survey data

FIGURE 12

Implementation could cut public-sector debt as a proportion of GDP by 27 percentage points by 2050



Source: TBI calculations using O*NET and Labour Force Survey data

05

How Might Future Developments in AI Increase Cost Savings Further?

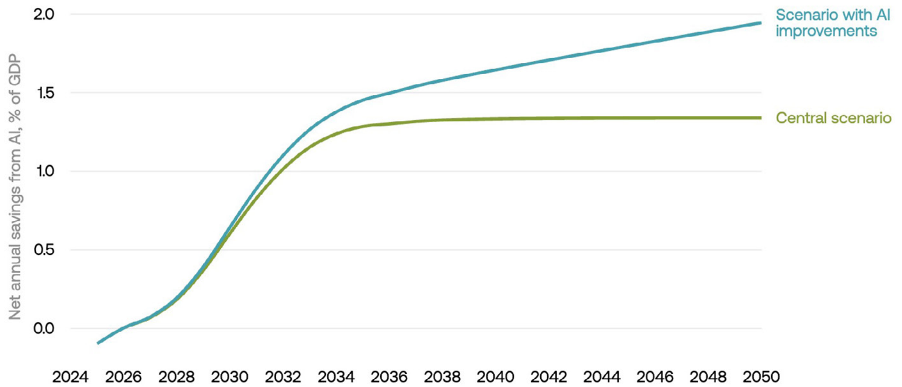
In just a few years, free online generative AI tools have advanced to include high-quality image, video and audio generation as well as new abilities to perform mathematical calculations. The potential savings from adopting AI in the public sector could therefore grow further as AI's capabilities expand in the future.

To model this upside potential, we analyse a scenario whereby the capabilities of AI continue to grow so that AI can save an extra 1 per cent of time (and hence wages) annually for public-sector workers not in shortage professions. We use the same rollout-speed assumptions and per-user costs as in our main analysis. The results show that the net savings from deploying AI in the public sector could rise to 1.9 per cent of GDP by 2050 (Figure 13) or around £54 billion per year in today's terms – £17 billion more than in the static scenario above.

The size of the prize is large and likely to grow further. Obtaining the maximum possible benefit from AI in government should, therefore, be a key element of the next government's agenda.

FIGURE 13

AI's capabilities are likely to expand further in the future, boosting its potential impact



Source: TBI calculations using O*NET and Labour Force Survey data

06

Technical Annex: Detailed Methodology

The nature of AI's integration into different sectors of the economy and different jobs remains highly uncertain. We are still in the foothills of AI adoption, so robust empirical studies assessing AI's impact on specific jobs and tasks remains limited.

Any estimates of AI's potential impact must therefore largely be based on a forward-looking assessment of how AI could affect the future of work rather than how it already has affected work. All studies that provide a macroeconomic-level assessment of AI's impact on the labour market thus rely on an ex-ante assessment of the technology's potential at a point in time and rest on specific assumptions that are open to debate. This study is no different.

Our methodology starts, as is now common in the literature on technology and employment, with the premise that jobs consist of a number of distinct tasks. It is not "jobs" that are replaced by technology but tasks: new technologies might perform some of these tasks, but not others. As with other studies, we obtain a list of tasks associated with each occupation from the O*NET database.[62]

The main aim of our analysis is to identify what type of tasks could be performed by AI and how much worker time could be saved as a result. Here, there are at least three broad approaches that have already been trialled in the literature – each with advantages and drawbacks:

- **Method 1: Broad categorisation of tasks based on individual human judgement:** This method, used by economists Briggs and Kodnani (2023) in a Goldman Sachs research paper, involves identifying broad clusters of "work activity" that could potentially be performed by AI.[63] In the O*NET database there are 39 categories of work activity (for example, "getting information" or "monitoring and controlling resources"). The authors assume 13 of these work activities could be performed by generative AI based on their own judgement and a reading of the existing evidence on AI. Within each work activity there are also seven levels of complexity identified in the O*NET database, and the authors assume that AI can perform up to a difficulty level of four across each of the 13 categories. Then, because the study is only focused on generative AI, they assume that any occupation that involves a significant share of workers' time spent outdoors or performing physical tasks cannot be automated by AI at all. Any task that meets all the above criteria is then assumed to be fully automatable and hence could lead to a time saving of up to 100 per cent. The authors' headline result is that 25 per cent of UK workforce time is exposed to automation from generative AI.

This approach is transparent and defensible, but as with all studies it does have some drawbacks. First, its categorisation of tasks is broad, so it cannot account for how AI's capabilities to perform individual tasks might vary across different professions. For example, it may be easier for AI to perform writing tasks in professions where there are large amounts of machine-readable data already available – such as the legal and financial professions – than in other settings where data is more siloed or expensive to digitise. Another drawback is that the analysis relies on fixed thresholds related to the difficulty of tasks that AI can perform, but this is somewhat arbitrary and may overestimate AI's abilities to perform some basic tasks and underestimate its ability to perform more complex tasks more efficiently. Finally, this method implicitly assumes that just because AI could perform 100 per cent of a task, that all of that task is at risk of automation. However, in some professions this is unlikely to be the case, either because it is more efficient for a human to work in conjunction with an AI tool on a particular task, or where it is socially desirable to keep a human in the loop. For example, for tasks that have a large impact on other human lives – such as judgements in criminal courts – generative AI may be able to produce judgements based on the submissions of both parties in a case, but it is unlikely that society would find this acceptable.

- **Method 2: Broad categorisation of tasks based on the wisdom of the crowd:** This approach, deployed by Felten, Raj and Seamans (2021), creates an AI exposure index by profession.[64] It involves using information from the Electronic Frontier Foundation's (EFF) 2017 AI

Progress Measurement study,[65] which identifies ten broad types of activity where AI has already been shown to be capable of performing particular tasks (for example, generating images, reading comprehension and so on). The authors then map those AI capabilities onto the 52 "occupational abilities" contained in the O*NET database based on a crowd-sourced survey of responses from 2,000 gig workers on the Amazon Mechanical Turk platform. The resulting AI Occupational Exposure (AIOE) index then identifies which jobs are most exposed to disruption from AI.

This method has been deployed by other researchers, including by the UK government[66] and the International Monetary Fund (IMF).[67] The latter has built on Felten et al's study by adding an "AI complementarity" component, which seeks to screen out tasks from the exposure index that would be socially unacceptable for AI to perform (such as decisions by judges). The IMF's headline result is that almost 70 per cent of the UK workforce have some exposure to generative AI.

This approach is novel in that it is using the wisdom of a crowd (rather than experts) to link AI capabilities with particular tasks, and the IMF's extension is innovative in that it accounts for the social acceptability of adopting AI in different settings. But this method also has its drawbacks. First, the assessment of AI's capabilities is static based on the EFF's 2017 study, so if there is a new breakthrough in AI's capabilities (such as its recent ability to generate video content) then these may not be captured by the study. Second, as with Method 1, the broad categorisation approach could mask some differences in AI's ability to perform similar tasks across different professions. Third, it is not obvious that asking a large number of non-experts will definitively give an answer that is closer to the truth than asking a small number of experts or relying on a large language model's (LLM) training data – knowledge about task content is not widely dispersed. Finally, while this approach helpfully identifies exposure, it does not provide a specific assessment of how much time could be saved by adopting AI in each profession. This time-saving component is essential in our analysis to identify the potential efficiency gains of adopting AI.

- **Method 3: Granular categorisation of tasks using LLMs cross-checked against human judgement:** This method, pioneered by Eloundou, Manning, Mishkin and Rock (2023),[68] involves using AI itself to help categorise whether a particular task can be performed by generative AI or not. The authors initially categorise a wide range of detailed work activities from the O*NET database using human annotators familiar with the capabilities of LLMs. They categorise activities into three groups: 1) no exposure to AI; 2) direct exposure to generative AI (where it could help complete the task at least 50 per cent faster); and 3) exposure to generative AI when paired with other software (where the pairing can lead to a time saving of at least 50 per cent). They then provide GPT-4 with a rubric of prompts to perform the same classification exercise and find a high degree of alignment between the human and AI-generated assessments, which suggests GPT-4 is a reasonable proxy for human judgement in this case. Overall, the authors find that LLMs could help complete 47 to 56 per cent of all worker tasks in the United States significantly faster.

  Other studies, including the Institute for Public Policy Research's (IPPR) recent report,[69] have applied Eloundou et al's approach to the UK and found that up to 59 per cent of workforce tasks could be affected by generative AI in the coming years. IPPR's analysis is very similar to Eloundou et al's, but relies more directly on ChatGPT's results, which are then cross-checked against the authors' judgements for a sample of 250 tasks.

  These studies are novel in that they are using ChatGPT as a form of the wisdom of the crowd – on the basis that the LLM was trained on a large volume of data that reflects the accumulated knowledge of the world's population and provides a probabilistic assessment of the most likely outcome based on that data. This approach also has an advantage in that it provides a more granular assessment of AI's capabilities across different tasks and professions, and is more specific about identifying how much time could be saved from adopting AI. However, as with the other studies, these advances also come with drawbacks. Eloundou et al acknowledge that both methods they use to categorise tasks are flawed – the human annotators used in their study are knowledgeable in the capabilities of LLMs but not in how they could be applied to specific professions. Meanwhile, ChatGPT's results, even though they match well with the human annotators, do give some contradictory assessments for some tasks. In addition, this approach does not distinguish between tasks that can be performed by AI in theory and those that should be performed by AI in practice; it includes tasks even if there are strong ethical reasons not to deploy AI in such a setting (as Method 2 adjusts for).

Clearly there are trade-offs between the different methods. None is perfect. Greater reliance on human judgement can limit the analysis to a broader categorisation of tasks with less specificity over time savings. On the other hand, pursuing a more detailed categorisation typically involves relying more on AI to support the assessment.

Our approach builds on this existing body of work in a number of ways. First, the scope of our study is broader; while the previous studies focused on generative AI, we attempt to assess the impact of all types of AI, including AI-enabled hardware that can perform physical tasks. Second, we provide a more granular assessment of AI's potential to save time within the workforce. Third, we utilise ChatGPT to help perform our analysis but do so by using it to create a system of filters that we then use to refine its results according to own judgemental assessment.

We began our analysis with a training data set of around 200 work tasks from the O*NET database. We then used this data set to iteratively develop a rubric of prompts that pushed OpenAI's GPT-4 Turbo model to produce results that accorded closely with our own judgement of AI's capabilities. Our own judgements were informed by the AI empirical studies mentioned above, the latest research on AI's existing capabilities, conversations with AI technology experts, and cross-checking the results from GPT-4 for particular tasks against those of each member of our own research team to benefit from the wisdom of crowds.

This process resulted in us using the following sequence of decision-tree prompts to GPT-4 to help categorise each task (which we have simplified here for brevity into short-form questions):

- **Is the task fully cognitive or does it require some sensory input or manual input to perform?** This is to identify whether a task can solely be performed by AI software (for example, generative AI or machine learning) or whether it requires complementary hardware as well (for example, a headset, microphone or more sophisticated AI-enabled hardware such as a drone or autonomous vehicle). This distinction is important as the base responses from ChatGPT tend to overestimate AI's current capabilities to perform physical tasks.

- **a. For cognitive tasks: Does the task require a high degree of human empathy, or does it have a significant impact on peoples' lives if errors are made?** This is to provide an assessment not just of whether AI can perform a task, but whether it is socially optimal to do so – similar to the IMF's "AI complementarity" component outlined in Method 2.

  b. For tasks involving manual input: Is the task repetitive and performed in a stable environment, or is it a complex physical task involving high degrees of autonomy in a changing environment? This is to correct the bias in ChatGPT's assessment of physical tasks, whereby it assumes that AI-enabled hardware can already perform a range of complex physical tasks (such as autonomous driving), whereas in practice most AI-enabled hardware is mainly used in controlled environments to perform repetitive tasks (for example, collecting stock in warehouses and distribution centres).

- **Given these previous characteristics, could AI perform the task?** As part of the prompt for this question, we provide GPT-4 with a strong prior (based on the previous answers) that certain tasks – such as complex physical tasks or cognitive tasks involving high degrees of human empathy – cannot be done by AI. The model should only overwrite this prior if it has a very high degree of confidence in its answer (that is, a high probability of being true).

- **If the task can be done by AI:**

  a. **Is it more efficient to be done solely by AI or in conjunction with a human?**

  b. **Would it be cost effective to use AI?**

  c. **Would it be socially desirable for the task to be performed by AI?**

  d. **What type of AI would be required to perform the task?**

  e. **How much time could AI save relative to a human solely performing the task?**

  The first four prompts are designed to provide filters for the data set enabling us to exclude certain tasks that would not be performed by AI in practice, even if they could theoretically be performed by AI. The final prompt then creates an estimate of the potential time saving from AI for each task.

Once we had refined our prompts on the training data set of 200 tasks, we then deployed the same rubric to GPT-4 Turbo to categorise the other near 20,000 tasks in the O*NET database – effectively giving GPT-4 a steer as to how to efficiently apply our judgements across a much larger data set, but still allowing it to apply the rich information in its training data to the nuances of each individual task.

Once we had this information at the task level, we then aggregated our results up to the occupation level by using information on the importance of each task in a profession to create an importance-weighted average. This gives us a potential time saving for each occupation in the O*NET database at the US Standard Occupational Classification (SOC) code level.

To apply this information to the UK Labour Force Survey (LFS) data, we then use a crosswalk to match US SOC codes to UK SOC codes so that we can calculate the potential saving for each occupation as defined in the UK SOC. We perform our analysis at the economy-wide level first and then use the

LFS data to identify how many public-sector workers are in each occupation and their wages, enabling us to calculate the aggregate time and cost saving across the public sector.

As part of the model-refinement process, we also cross-checked our results at various points to ensure they accord with frontier studies that assess AI's potential impact on work. For example:

- **At the task level:** Dell'Acqua et al (2023)[70] conducted a field study of 758 consultants from the Boston Consulting Group to test the ability of AI to perform a range of consultancy tasks. They found that AI helped perform these tasks 25 per cent faster on average. Our analysis of the "management analyst" and "project management specialist" professions, both of which involve tasks closely related to those in the study, estimated that the use of AI would introduce a time saving of 24.7 per cent and 19.5 per cent respectively – very close to the empirical results from the study. Moreover, both these figures are conservative when compared with other studies on the ability of AI to improve the writing speed of business professionals. For example, Noy and Zhang (2023) show that professionals who use ChatGPT to help with writing tasks can save about 40 per cent of their time.[71]

  Peng et al (2023) conducted a separate study focused on computer programmers and found that access to GitHub Copilot, an AI pair programmer, can save 55.8 per cent of time for some coding tasks.[72] Our estimates, which apply to a wider range of coding and software-based tasks, are more conservative but in a similar ballpark – indicating an average saving of 39 per cent across these tasks, or 29 per cent for all tasks associated with computer programmers.

- **At the occupation level:** Recent evidence from the UK Department for Education (DfE) found that AI can save teachers at least 4 per cent of their time, while a new study by Oak National Academy, a provider of digital teaching resources, suggests a time saving of up to 8 per cent.[73] [74]Again, this range closely matches our own estimate of time saved by primary and secondary school teachers through the use of AI (6.3 per cent and 7.6 per cent respectively). And again, our results are more conservative than some other studies, for example a McKinsey study suggested time savings of 20 to 40 per cent were possible for teachers.[75]

- **At the economy-wide level:** We can also compare our results at an aggregate level with those of the macroeconomic studies mentioned earlier. Our overall potential time saving of 25 per cent across the whole economy closely matches Briggs and Kodnani's estimate[76] highlighted in Method 1. Our results on the share of tasks impacted by AI are also in a similar ballpark to other studies. We estimate 50 per cent of employment-weighted tasks across the whole economy are potentially exposed to AI – in line with Eloundou et al's 47 to 56 per cent range for the United States,[77] but slightly less than IPPR's 59 per cent figure for the UK.[78] Since our study includes a broader range of AI tech than these studies, which focus only on generative AI, this suggests our estimates are generally more conservative than other studies.

These robustness checks provide reassurance that the numbers produced in this paper are consistent with other expert judgements and the emerging real-world evidence. However, as noted earlier, the figures rely on a forward-looking assessment of AI's potential, so both higher and lower numbers are possible. These figures should thus be treated as indicative of the scale of potential gains that could emerge, rather than a precise point forecast of what will happen.

Endnotes

1 https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/publicsectorpersonnel/datasets/publicsectoremploymentreferencetable

2 https://www.gov.uk/government/statistics/public-expenditure-statistical-analyses-2023

3 https://www.ons.gov.uk/economy/economicoutputandproductivity/publicservicesproductivity/bulletins/publicserviceproductivityquarterlyuk/octobertodecember2023#quarter-on-quarter-productivity-estimates

4 Source: TBI analysis of Labour Force Survey data

5 https://obr.uk/download/public-finances-databank-may-2024/?tmstv=1718189148

6 https://digital.nhs.uk/data-and-information/publications/statistical/nhs-workforce-statistics/february-2024

7 https://www.ons.gov.uk/economy/economicoutputandproductivity/publicservicesproductivity/bulletins/publicserviceproductivityquarterlyuk/octobertodecember2023

8 Source: Office for National Statistics, https://www.ons.gov.uk/economy/economicoutputandproductivity/publicservicesproductivity/bulletins/publicserviceproductivityquarterlyuk/octobertodecember2023.

9 https://www.productivity.ac.uk/wp-content/uploads/2023/11/PIP025-Public-Sector-Productivity-FINAL-Nov-2023.pdf

10 https://www.semafor.com/article/10/06/2023/why-walmart-thinks-ai-wont-cut-jobs

11 https://www.nber.org/papers/w31161

12 Brynjolfsson et al. (2022), "Generative AI at Work", NBER Working Paper 31161.

13 https://www.science.org/doi/10.1126/science.adh2586

14 Noy and Zhang (2023), "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence", MIT Working Paper.

15 https://www.nber.org/papers/w30957

16 Korinek (2023), "Language Models and Cognitive Automation for Economic Research", NBER Working Paper 30957.

17 https://www.thetimes.com/uk/healthcare/article/ai-cure-for-bed-blocking-can-predict-hospital-stay-fpvjn5ql6

18 https://www.onetcenter.org/database.html#individual-files

19 See, for instance, https://www.goldmansachs.com/intelligence/pages/ai-may-start-to-boost-us-gdp-in-2027.html

20 Tyna Eloundou and others, "GPTs Are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models"

21 That is, they do not necessarily give an identical response to an identical request.

22 For some tasks, GPT-4 was unable to assign a specific technology type that could reduce the time taken, suggesting that the technology may not yet exist. We therefore exclude any time savings from these tasks in our analysis.

23 https://www.gov.uk/government/publications/generative-ai-framework-for-hmg/generative-ai-framework-for-hmg-html

24 https://www.sciencedirect.com/science/article/pii/S2666990024000132

25 https://aimlprogramming.com/services/drone-ai-surveillance-for-environmental-monitoring/

26 https://rep-tec.co.uk/products/robotics

27 We define frontline staff as health professionals, teaching professionals, associate professionals and support workers, care workers, armed forces, police officers and other protective-service occupations.

28 Recent research from the TUC suggests that one in six public-sector workers did unpaid overtime in 2023 compared with one in nine in the private sector. For example, teachers did an average of 4.4 hours unpaid overtime per week. See UK workers put in £26 billion worth of unpaid overtime during the last year - TUC analysis | TUC.

29 More recent reforms have replaced the shortage occupation list with an immigration salary list of jobs that have a lower salary requirement than that for a standard Skilled Worker visa. Health care and education workers are still eligible for visas if they are paid in line with a national pay scale. The shortage occupation list can be found here.

30 We use earnings information from the UK Labour Force Survey to account for different wage levels of public-sector professions and their individual susceptibility to AI to map from time savings to an overall fall in the public-sector wage bill.

31 https://publications.parliament.uk/pa/cm201314/cmselect/cmpubacc/294/294.pdf

32 https://www.mckinsey.com/industries/public-sector/our-insights/unlocking-the-potential-of-public-sector-it-projects

33 https://www.gov.uk/government/news/home-office-confirms-changes-to-the-eu-settlement-scheme

34 https://get.kainos.com/rs/935-JPR-705/images/HMPO%20Case%20Study%20Updated%20Jan%2024.pdf

35 https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics

36 https://digital.nhs.uk/data-and-information/publications/statistical/patients-registered-at-a-gp-practice

37 https://www.gov.uk/guidance/local-government-structure-and-elections

38 Alexander Iosad, David Railton and Tom Westgarth, Governing in the Age of AI: A New Model to Transform the State, 2024; https://www.institute.global/insights/politics-and-governance/governing-in-the-age-of-ai-a-new-model-to-transform-the-state

39 In each case, we assume 5 per cent coverage in 2024 and use a logistic S-curve to model the rollout. The time specified is the amount of time it takes to get from 10 per cent to 90 per cent coverage.

40 https://www.institute.global/insights/politics-and-governance/governing-in-the-age-of-ai-a-new-model-to-transform-the-state#footnote_list_item_39

41 https://www.jpmorgan.com/technology/artificial-intelligence/research-awards

42 https://www.jpmorganchase.com/ir/annual-report/2023/ar-ceo-letters

43 https://www.gov.uk/government/statistics/civil-service-statistics-2023/statistical-bulletin-civil-service-statistics-2023#civil-service-headcount

44 https://www.wired.com/story/regulators-need-ai-expertise-cant-afford-it/

45 https://www.levels.fyi/t/software-engineer/focus/ml-ai?countryId=254&country=254

46 https://www.institute.global/insights/politics-and-governance/governing-in-the-age-of-ai-a-new-model-to-transform-the-state#the-return-on-investment-far-exceeds-the-costs

47 We assume that: each user submits 10,000 queries over the course of a year; each query takes 1.65 seconds of compute time to complete (based on average ChatGPT responses) and hence each user requires 4.6 hours of compute time per year. Based on current market pricing, we assume that an advanced GPU costs around £4 per hour, implying an annual compute cost per worker of £18.6 per user. Each 10,000 queries also uses around 1.375 kWh of energy based on recent studies and ChatGPT's own energy-consumption estimates, and that each kWh of energy costs the current market rate of just over 20p - implying annual energy costs of around 30p per user. We then assume each user generates and stores up to 10GB of data permanently on the cloud (which costs 20p per month to store, or £2.40 per year) and transfers up to 10GB of data (which costs just under 8p per month, or 70p per year). Summing up across all these cost categories implies a total annual cost per user of £22. We then conservatively round this figure up to £30 per user to account for other potential costs and uncertainties in the calculations (such as if the AI tools are used more intensively than assumed).

48 https://www.microsoft.com/en-us/microsoft-365/business/copilot-for-microsoft-365#footnote1

49 https://otter.ai/pricing

50 https://calendly.com/pricing

51 Source: ChatGPT to generate initial range of costs, which are then checked and refined against available third-party sources. We then conservatively adopt an estimate towards the upper end of the estimated range.

52 For example, the cost of Microsoft Copilot for business is $30 per user per month, Otter.ai's transcription service is $20 per month, and Calendly's scheduling assistant is $16 per month. This equates to $192-360 per user per year, or roughly £150–300.

53 https://www.institute.global/insights/politics-and-governance/governing-in-the-age-of-ai-a-new-model-to-transform-the-state#the-return-on-investment-far-exceeds-the-costs

54 https://www.amazon.co.uk/s?k=headsets+with+microphone

55 https://www.amazon.co.uk/Surveillance-Cameras/b?ie=UTF8&node=332219031

56 https://www.amazon.co.uk/Amazon-Monitron-Sensors-pack-monitoring/dp/B0851JVLTZ

57 In the limited instances where we use ChatGPT to help plug data gaps in this paper we use it in a similar way to how it could be applied in the public sector – as a co-pilot to help explore analysis in frontier areas but where any AI-generated insights are refined, checked against third-party sources, and subject to sensitivity analysis to ensure accuracy and robustness. In this example, there is a lack of comprehensive financial data or survey data on the potential costs of deploying AI-enabled hardware in different public-sector applications, partly because there is such a diverse range of hardware that could be deployed, each of which has different costs. To address this data gap, we develop our own cost estimates using a case-study approach. We use ChatGPT as a co-pilot to help identify potential real-world applications where AI hardware could save time in the public sector and use it to suggest initial approximate costings for each case study. We then cross-check those figures and applications against available third-party sources to refine them and ensure they are plausible and credible. Then, as a further robustness check, we run sensitivity analysis on the cost assumptions (including in the scenario analysis later in the paper) to explore how different assumptions affect the results. Overall, the final estimates used in this section are an average of a diverse range of cost estimates – so higher and lower costs are possible. However, because such a limited number of jobs are affected by this technology our headline results are insensitive to the assumptions used in this section.

58 This is in line with the 35 hours of Continuing Professional Development that teachers are given each year. We further assume a 10 per cent turnover rate of these ambassadors each year requiring this training, roughly in line with the turnover of workers leaving the public sector.

59 https://explore-education-statistics.service.gov.uk/find-statistics/employer-skills-survey

60 https://www.civilservicepensionscheme.org.uk/your-pension/work-life/redundancy/

61 https://www.gov.uk/government/publications/spring-budget-2024/spring-budget-2024-html

62 https://www.onetcenter.org/database.html

63 https://www.gspublishing.com/content/research/en/reports/2023/03/27/d64e052b-0f6e-45d7-967b-d7be35fabd16.html

64 https://oar.princeton.edu/handle/88435/pr11551

65 https://www.eff.org/ai/metrics

66 https://www.gov.uk/government/publications/the-impact-of-ai-on-uk-jobs-and-training

67 https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379

68 https://arxiv.org/pdf/2303.10130

69 https://ippr-org.files.svdcdn.com/production/Downloads/Transformed_by_AI_March24_2024-03-27-121003_kxis.pdf

70 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4573321

71 Experimental evidence on the productivity effects of generative artificial intelligence | Science

72 https://arxiv.org/abs/2302.06590

73 Teachers work 50 hours a week based on DfE statistics. A recent DfE (2024) report indicated that using generative AI tools can save multiple hours of teachers' time per week. Given that teachers work 50 hours a week on average and "multiple" implies at least two hours, this equates to at least a 4 per cent time saving. Meanwhile, Oak National Academy has shown that teachers can save at least four hours a week by deploying AI tools, which equates to an 8 per cent time saving.

74 b3816c4fd6b7e92d301bf034753f465be334bb7c.pdf (thenational.academy)

75 How artificial intelligence will impact K-12 teachers (mckinsey.com)

76 https://www.gspublishing.com/content/research/en/reports/2023/03/27/d64e052b-0f6e-45d7-967b-d7be35fabd16.html

77 https://arxiv.org/pdf/2303.10130

78 https://ippr-org.files.svdcdn.com/production/Downloads/Transformed_by_AI_March24_2024-03-27-121003_kxis.pdf

TONY BLAIR
INSTITUTE FOR
GLOBAL CHANGE

Follow us

facebook.com/instituteglobal
twitter.com/instituteGC
instagram.com/institutegc

General enquiries

info@institute.global

Copyright © July 2024 by the Tony Blair Institute for Global Change

All rights reserved. Citation, reproduction and or translation of this publication, in whole or in part, for educational or other non-commercial purposes, is authorised provided the source is fully acknowledged. Tony Blair Institute, trading as Tony Blair Institute for Global Change, is a company limited by guarantee registered in England and Wales (registered company number: 10505963) whose registered office is One Bartholomew Close, London, EC1A 7BL.