An Old Challenge in the New Era:

How the Public Sector Can Benefit from the Age of AI

Author:
Joel Hoskins
The Productivity Institute

Date:
September 2025

The Productivity Institute
Productivity Insights Paper No.060

Key words

Artificial Intelligence, Public Sector Productivity, Technological Implementation, Political Economy, Service Delivery Chains, Sectoral Capabilities, Policy Shift, Long-Term Investment, Risk Management

Authors' contacts

joel.hoskins@manchester.ac.uk

Acknowledgements

I am grateful to Nina Jörden, Bart van Ark, Diane Coyle, Alan Lewis, Marion Oswald, Muffy Calder, Lucy Veale, Francine Ryan, Mary O'Mahony, and two anonymous reviewers for their valuable comments and support during the preparation of this report. All remaining errors and omissions are my responsibility alone.

This work was supported by the Engineering and Physical Sciences Research Council [grant number EP/Y009800/1], through funding from Responsible Ai UK (IO008)

Suggested citation

J. Hoskins (2025) *An Old Challenge in the New Era: How the Public Sector Can Benefit from the Age of AI*. Productivity Insights Paper No. 060, The Productivity Institute.

The Productivity Institute is an organisation that works across academia, business and policy to better understand, measure and enable productivity across the UK. It is funded by the Economic and Social Research Council (grant number ES/V002740/1).

The Productivity Institute is headquartered at Alliance Manchester Business School, The University of Manchester, Booth Street West, Manchester, M15 6PB. More information can be found on The Productivity Institute's website. Contact us at theproductivityinstitute@manchester.ac.uk

Abstract

While Artificial Intelligence (AI) offers a significant opportunity to boost public sector productivity, the sector's history of struggling with technological implementation suggests that the realisation of this opportunity is far from guaranteed. This report argues that the solution to leveraging AI lies not in the technology itself, but in building the foundational capabilities needed to overcome historical failures and navigate a political economy that often favours short-term wins over long-term investment. This report identifies AI's core value in removing information-intensive bottlenecks from service delivery chains. Further, it introduces a diagnostic framework of twelve capabilities essential for finding and taking advantage of the opportunities presented by AI. Five of these capabilities can be exercised at the organisational level (e.g., risk management, being an ‘intelligent customer’) and seven capabilities can be exercised at the sectoral level to help create a healthy innovation ecosystem (e.g., co-ordinating research, strengthening networks). The conclusion is that to unlock AI's potential, policy must pivot from a focus on technological acquisition to a sustained investment in building these vital organisational and systemic capabilities.

1 Introduction

The rapid proliferation of artificial intelligence (AI) presents the public sector with a transformative opportunity to improve the value and efficiency of its services. Governments have shown a clear ambition to seize this potential, with the UK Treasury recently announcing plans to deliver nearly £14 billion in annual efficiencies, driven in large part by AI and digital transformation (HM Treasury, 2025). However, such ambitions for technological modernisation are not new. From the desire to harness the 'white heat' of the scientific revolution in the 1960s to the e-government initiatives of the early 2000s, successive administrations have staked their credibility on the promise of a more innovative state. Yet, a significant gap has long separated ambition from achievement.

The history of failure has imposed a substantial cost on the public sector. Firstly, a perceived inability to successfully take advantage of new opportunities can lead public sector organisations to not even attempt to adopt new technology, resulting in a reliance on obsolete and unreliable equipment. A clear example of this would be the continued operation of the Police National Computer, a system that first went live in 1974 (The Police Foundation, 2022), making it older than 63% of the UK population. Secondly, a failure to address the challenge can lead to calamitous overspend, such as the NHS National Programme for IT – a technological upgrade programme with a reported cost overrun of more than 440% (Fleming, 2004). Thirdly, a failure to address the challenge can create technology with key weaknesses that go on to have negative consequences for staff or members of the public. One example of this would be the Horizon software, an IT system in the Post Office whose technical weaknesses contributed to the wrongful imprisonment of more than 900 subpostmasters (BBC, 2024). These incidents represent a deep and persistent pattern of failure, with an estimated 28% of government systems being dependent on unacceptably outdated technology, and only 9% of current major IT upgrade projects being assessed as having a high likelihood of success (DSIT & GDS, 2025).

The root cause of these failures is frequently not technical, but institutional[1] - a problem caused by deep-seated issues in the political economy of the public sector. In this realm, the organisational value of long-term investment, learning from failure, and evidence-based prioritisation often collides with a more powerful political imperative to avoid blame (Hood, 2010) and to prioritise presenting an immediate 'victory' through credit-claiming (Mayhew, 1974), particularly in the run-up to an election (Jacobs, 2011). This tension occurs when political actors are not rewarded for diligent effort but are rewarded for initiating exciting new projects that capture the attention of the media. This incentive structure can create a systemic bias against the patient, often invisible work of capability-building, favouring instead the announcement of highly visible 'shiny projects' that serve a symbolic purpose (Edelman, 1985) and offer short-term personal credit for the political sponsor, but limited prospects for increasing productivity. This political logic, while challenging for long-term productivity growth, is a product of the essential democratic requirement for political leaders to be responsive to public concerns.

This report introduces a new diagnostic framework for understanding how the internal and systemic conditions within the public sector can either support or impede the attainment of productivity from AI. It does this by looking at the foundational capabilities that have been identified as relevant for driving an organisation to successfully find and take advantage of the opportunities for improvement that are presented by a frontier technology like AI.

The twelve capabilities in this framework are the result of a deliberate process designed to bridge proven theory with managerial practice. The analysis began with a synthesis of established literature on innovation, economics, and strategic management. These theoretical factors were then critically assessed against the lived reality of government, using case studies of both success and failure to ensure each capability was relevant and robust. To be included, a capability had to meet three clear criteria: it had to be conceptually distinct; it had to be generalisable across different public services; and a deficiency in it had to have a clear and predictable

[1] This conclusion is reinforced by wider economic evidence. While the performance-adjusted cost of computation continues its rapid, historic decline, national productivity growth has slowed (Coyle & Hampton, 2024). This growing gap between technological potential and realised economic performance suggests the primary bottleneck is not the technology itself, but the capacity of organisations to absorb and leverage it.

negative impact on productivity. Finally, this report organises the framework by locus of control, distinguishing between five organisational capabilities that a single leadership team can command, and seven sectoral capabilities that require collective action to build a healthy innovation ecosystem.

Navigating the political economy of the public sector, where political imperatives often clash with the needs of long-term productivity growth, requires public sector managers not to insulate themselves from political realities, but to proactively shape them by providing political leaders with more credible, deliverable, and ultimately more successful opportunities for improvement. The capabilities outlined in this report equip a manager to fulfil this role. Therefore, they are not just tenets of good management; they are the essential institutional armour required to navigate the challenging environment of the public sector. They may even provide the proven competence that earns the trust and autonomy required to undertake more ambitious, long-term reforms.

This report begins by examining the opportunities presented by AI, describing its core technical competencies along with the mechanisms by which these competencies can enhance the productivity of a public service organisation. The report then addresses the central challenge of translating these opportunities into real improvements, by identifying five essential organisational capabilities - such as organisational learning, strategic planning, and acting as an 'intelligent customer' - that enable public bodies to identify, develop, and integrate AI technologies successfully. The analysis then broadens to seven sector-level capabilities that can be exercised by 'strategic institutions', which are the public or quasi-public bodies with a mandate for sector-wide stewardship. Examples of such institutions include the Ministry of Housing, Communities, & Local Government, the National Police Chiefs' Council, or Skills for Care. The capabilities that these institutions can exercise support the overall innovation system by setting rules of interaction, and by creating collective assets. Such capabilities include strengthening innovation networks, co-ordinating research, and ensuring a sustainable flow of human capital. The report concludes by underscoring the central policy implication: that unlocking AI's value will require a shift in policy focus away from technological acquisition and toward the development of organisational capability and the nurturing of the wider innovation system.

2 The Opportunity

Despite ongoing uncertainty about the ultimate scope of its potential, the core value proposition of AI for the public sector appears to concern an improvement in the quality and accuracy of information-intensive tasks while also frequently reducing the associated costs. This is because AI can fulfil data-based tasks that have historically been performed manually by human workers, who can be prone to error and are often significantly slower (Agrawal, Gans, & Goldfarb, 2022). In section 2.1, we outline the competencies that AI has displayed, and in section 2.2 we discuss how the application of these competencies can increase productivity.

2.1 Competencies and Applications of AI

Recent advancements in AI technology have expanded machine capabilities across three key domains: Firstly, they have enabled models that can interpret complex data like imagery (Computer Vision), text (Natural Language Processing), and speech (Speech Analytics) with unprecedented scale and sophistication. Secondly, they have displayed an improved ability to forecast future trends and classify complex data (Predictive Analytics). Thirdly, and most radically, recent models are able to generate novel content and interactions through conversational Chatbots and Large Language Models (LLMs).

Computer vision refers to the ability of some AI models to analyse images or videos to recognise objects, faces, or patterns, enabling the automation of visual tasks at a vast scale and speed. In law enforcement, for instance, AI-driven facial recognition systems can process vast amounts of surveillance footage to identify individuals of interest almost instantaneously, a task previously handled manually (Sarzaeim et al., 2023). Causey et al. (2018) evaluated an AI system for detecting malignant lung cancer in CT scan images and reported a classification accuracy of 95%, compared to 83% for radiologists' manual interpretations.

Natural Language Processing (NLP) models are designed to systematically extract and interpret information from unstructured text (i.e., text that is not organised in a formulaic, pre-defined format), such as emails, reports, and transcripts. This capability can make information retrieval and decision-making more efficient. For example, AI-enabled clinical documentation systems have automated the time-intensive process of capturing patient information, reducing manual effort and errors in healthcare settings (Lin et al., 2018). Similarly, AI models have been deployed in corruption investigation to identify anomalous communication links, deceptive sentiment, or unusual topics of conversation from vast quantities of email and contract data (Bao et al., 2022).

Speech analytics is a competency that combines two technologies: automatic speech recognition (ASR) and natural language processing (NLP). The foundational step, ASR, converts spoken language into text. On its own, this provides significant value; for example, automated speech transcripts can reduce the time and errors associated with manual note-taking (El Hannani et al., 2021), supporting both transparency and efficiency. The analytics component is then enabled by applying NLP to the transcribed text to interpret its meaning and extract actionable data (Alharbi et al., 2021). This combined capability can power more advanced applications, such as real-time language translation to facilitate communication in multilingual public service environments (Pannu, 2015).

Predictive analytics models can uncover complex, non-linear patterns in vast datasets, including unstructured sources like text and images. Its novelty lies in generating accurate forecasts without a human-understandable theory of causation. A striking example is an AI system that was able to predict patient gender from chest X-rays with remarkable accuracy. Even after researchers digitally obscured the anatomical markers that radiologists rely on, the performance was only marginally reduced, and researchers were unable to establish what the model was using to make its predictions (Adleberg et al., 2022). This ability to operate in domains too complex for simple formulas can be used in the public sector. For example, computer-vision models have been used to support faster detection of strokes and cancer, thereby helping to triage waiting lists more effectively (Department of Health and Social Care, 2022). Likewise, the National Grid ESO employs AI to forecast national energy demand, integrating factors like weather and large-scale social events to balance supply and strengthen energy security (National Grid ESO, 2019).

Chatbots leverage NLP and retrieval algorithms to interact with users in conversational formats, serving as virtual assistants for employees and citizens. Internally, organisations have deployed AI chatbots to handle routine HR and IT inquiries, freeing staff from repetitive questions and enabling faster response times (Davenport & Ronanki, 2018; Zheng et al., 2018; Cortés-Cediel et al., 2017). Externally, citizen-facing chatbots provide personalised assistance - such as guiding users through forms or providing policy information - potentially improving service accessibility and user satisfaction (Davenport & Ronanki, 2018). Complementing chatbots, AI-driven knowledge management platforms index and curate documents, allowing employees to search across extensive repositories without prior familiarity (Wirtz et al., 2019).

The development of Large Language Models (LLMs) represents a significant shift from AI focused on analytical tasks like classification and extraction to AI capable of generating novel content. This new generative capacity stems from their ability to generalise knowledge from vast datasets, enabling them to solve novel problems for which they were not explicitly trained. As a result, LLMs can perform tasks that require complex interpretation, reasoning, and synthesis, such as reviewing legal contracts for potential issues (Schwarcz & Choi, 2023), writing functional software code (Chen et al., 2021), and generating novel scientific hypotheses (Boiko et al., 2023).

Despite these capabilities, the efficacy of artificial intelligence is circumscribed by several inherent limitations. A primary concern is AI's potential to function as a 'stochastic parrot', adept at mimicking patterns and statistical correlations within its training data without meaningful comprehension of what real-world concepts the data represent (Bender et al., 2021). This reliance on imitation means the performance of an AI model is intrinsically tethered to the quality, breadth, and representativeness of the available data (Whang et al., 2022); biased, incomplete, or inaccurate datasets can lead to skewed, flawed, or inequitable models, a principle often expressed as "garbage in, garbage out" (O'Neil, 2016). Consequently, AI systems can produce unreliable outputs, including presenting fabricated information as factual, or predictions that lack robustness and which fail unpredictably in novel contexts (Marcus, 2020). Furthermore, as AI systems become more autonomous and capable of pursuing goals with less direct human oversight, new categories of risk may emerge, including the potential for systems to develop unintended sub-goals that conflict with the public interest (Bostrom, 2014).

These limitations are not just abstract technical concerns; they can create serious weaknesses in the technology that can cause it to have an unexpectedly negative (or diminished positive) impact on an individual or society. There are two types of failure regarding the impact that AI applications can create: Firstly, operational failure occurs when an AI application delivers a negligible or even negative impact on productivity relative to its promised benefits. This can happen when the system produces unreliable outputs that are presented with a misleading degree of certainty, causing staff to waste time or make flawed decisions. Secondly, ethical failure occurs when an AI system inflicts harm on members of the public. Such failures include breaches of the Equality Act 2010 (c. 15) arising from algorithmic bias or the violation of procedural fairness when opaque decisions cannot be meaningfully challenged.

These two categories of failure are not mutually exclusive. An ethical failure can quickly become an operational failure. For example, if a predictive policing tool is perceived as biased, public cooperation may diminish, subsequently requiring greater investigative resources and ultimately undermining overall police productivity (van Ark & Hoskins, 2024). Likewise, an operational failure can have immediate ethical consequences – for example, when an AI algorithm wrongly freezes an innocent person's bank account, and so causes personal hardship.

2.2 How AI Applications Can Increase Productivity

Having surveyed the key competencies of AI technologies, we now consider how their application can contribute to productivity in a public sector setting. We identify three channels by which they can have such a positive effect: firstly, AI applications can directly increase productivity by streamlining the delivery chain; secondly, they can enable the creation of entirely new types of public services; thirdly, AI can indirectly increase productivity by enhancing the very organisational capabilities that are needed to innovate and adapt.

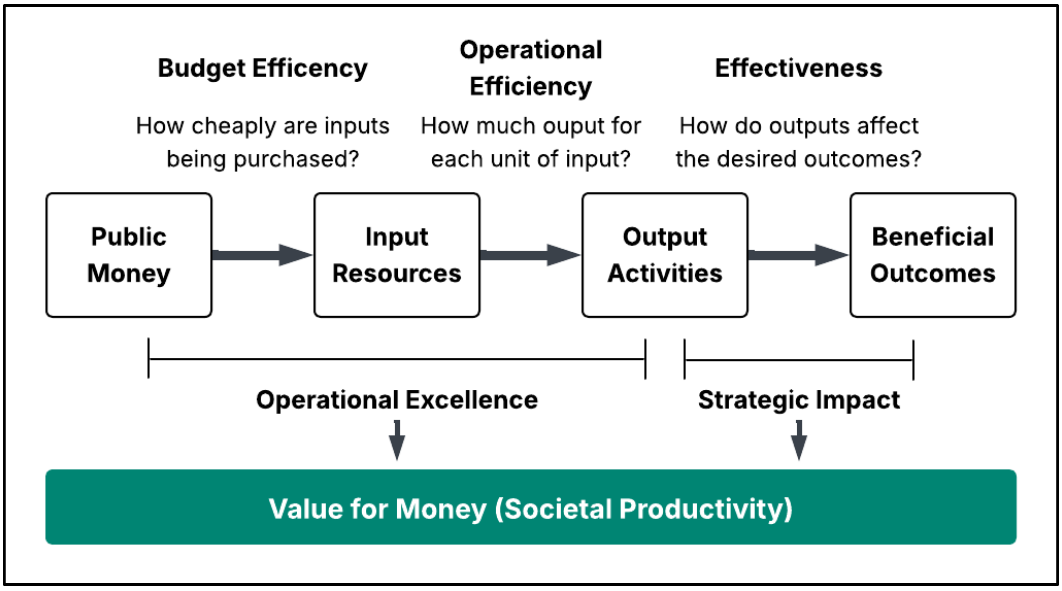
2.2.1 Streamlining Existing Delivery Chains

A bottleneck in a public service is a point where the flow of resources is impeded, creating a backlog and constraining the value that can be delivered downstream. For example, a shortage of detectives in the police service can delay investigations, which in turn reduces the number of cases that reach court and weakens the deterrent effect of policing. Such a bottleneck effectively sets the pace for the entire process. According to the Theory of Constraints, this means that overall productivity can only be significantly improved by addressing this single limiting factor;

investment in any other part of the system will yield little to no overall improvement (Goldratt, 1990; van Ark, 2022).

The delivery chain provides an effective framework for identifying and contextualising bottlenecks in public services. It does this by systematically mapping the flow from budgetary and resource inputs (such as funding, staff, and equipment), through the output activities that constitute the service, to the ultimate beneficial outcomes received by the public. Figure 1 illustrates this chain. On the left is budgetary efficiency (how effectively inputs are purchased), in the middle is operational efficiency (how inputs are converted into outputs), and on the right is effectiveness (how those outputs achieve desired outcomes).

Figure 1: The Public Sector Delivery Chain



Source: van Ark (2022). Adapted from Aldridge, Hawkins and Xuereb (2016).

A failure to consider the entire delivery chain can lead to interventions that reduce overall productivity, even if they seem efficient in isolation. For example, a police force might use AI to automate witness statements, cutting the time officers spend on reports. But if these statements are less accurate or miss important details, they could create more work for prosecutors, who must verify them, and for courts, which must deal with challenges to their validity. In this case, optimising one stage of the chain creates a worse bottleneck downstream, reducing overall productivity. This shows why system-wide planning and cross-sectoral co-ordination are essential for realising the benefits of AI in the public sector.

Productivity gains across the delivery chain can be categorised as either technical or substitution improvements. A technical improvement enhances an existing process or resource, making the current way of working more productive. A substitution improvement, by contrast, replaces a process or resource with a fundamentally different and more productive alternative. The distinction is best understood by its effect on a task: a technical improvement typically augments a worker, making them more effective at their current job, while a substitution improvement automates or replaces their task with a more productive approach.

Figure 2: A Typology of Productivity Improvement

	Budgetary Efficiency	**Operational Efficiency**	**Effectiveness**
Substitution Improvement	(1) An input resource is removed/ exchanged for another that has lower cost	(2) A given task is exchanged for another that requires less time/resource	(3) A given task is exchanged for another that makes a greater contribution to outcomes
Technical Improvement	(4) The same resource is purchased at lower cost	(5) A given task is adapted to take less time/resource	(6) A given task is adapted to make a greater contribution to outcomes

Source: Author's own.

In practice, new technologies often influence multiple types of improvement at once. For example, an AI-driven Natural Language Processing tool for reviewing citizen applications could reduce the time and resources needed for the task (technical improvement, cell 5), free skilled staff for more complex cases (substitution improvement, cell 3), and lower overall staffing costs (budgetary improvement, cell 1). By contrast, predictive policing algorithms may improve the likelihood that police resources are in the right place at the right time (cell 2), but they also risk perpetuating historical biases if trained on skewed data. This can undermine equitable outcomes and public trust, which in turn may reduce both technical efficiency and effectiveness (cells 5 and 6) by lowering public co-operation (van Ark & Hoskins, 2024; Lapuente & van de Walle, 2020; Jackson et al., 2013; Alford & Hughes, 2008).

AI can enhance budgetary efficiency, the first stage of the delivery chain, through both substitution and technical improvements. A substitution improvement (cell 1) occurs where a costly input like human labour is replaced by a less expensive automated system. The "Consult" tool, which analyses thousands of public consultation responses in hours instead of months, is a prime example of this direct reduction in expenditure (Incubator for Artificial Intelligence, 2025). Alternatively, it can be a technical improvement (cell 4), where AI helps acquire the same inputs at a lower cost. Procurement platforms like Simfoni, for instance, use AI to consolidate spending data and identify bulk-purchasing opportunities, thereby improving an organisation's purchasing power (Guida et al., 2023).

Beyond direct cost savings, AI can enhance operational efficiency by reducing the resources needed to deliver a service – which in labour-intensive public services, typically means saving time. Technical improvements (cell 5) make existing tasks faster, as seen with Swindon Council’s "Magic Notes" tool, which significantly reduced the time social workers spent on case notes (Koutsounia, 2024). Substitution improvements (cell 2) replace a process with a more productive one; for example, an AI camera system that proactively detects falls among dementia patients replaces a reactive emergency response model, reducing hospital visits (Xiong et al., 2019). Many AI applications provide both benefits: an AI system analysing medical scans can help radiologists find tumours faster (a technical improvement) while automating the triage of healthy scans (a substitution improvement), fundamentally changing the workflow (McKinney et al., 2020).

The final stage of the delivery chain is effectiveness – corresponding to how much beneficial impact is actually delivered by the activities that the service provides. A substitution improvement in effectiveness (cell 3) occurs when an organisation replaces a lower-impact output with a higher-impact one. For example, an AI scheduling system in eldercare enabled a shift from providing hours of simple supervision to delivering more hours of direct patient care, a more valuable activity that directly improved residents' quality of life (Bajo et al., 2008). In other cases, a technical improvement (cell 6) makes an existing output more potent. For instance, by using AI-powered facial recognition, the investigative leads (the output) produced by Homeland Security Investigations became vastly more accurate and timelier

(Brewster, 2023) - empowering officers to achieve superior public safety outcomes by identifying more perpetrators.

2.2.2 Enabling New Delivery Chains

Beyond optimising existing processes, AI's most consequential potential lies in reconfiguring entire delivery chains, owing to its role as a General-Purpose Technology (GPT) (Bresnahan & Trajtenberg, 1995; Eloundou et al., 2024). Like electricity or the internet, AI is pervasive, improving over time, and functions as a platform for a wide range of novel tools and processes; consequently, it can enable fundamentally new types of service and organisational arrangements. For example, diagnostic imaging could be relocated from hospital radiology departments to local pharmacies, with image interpretation performed centrally by AI systems. Such an innovative approach would not involve improving existing workflows, but rather fundamentally reorganising them. Patient access would be decentralised to community settings, while the scarce resource of expert interpretation is centralised and accelerated by the AI. While a full exploration of such reconfigurations is beyond the scope of this report, organisational capabilities, such as a deep understanding of existing frustrations and bottlenecks, still provide an essential foundation for creating transformative changes that genuinely improve on the status quo.

2.2.3 Supporting Organisational Capability

The third channel through which AI applications can increase productivity is by enhancing an organisation's capacity to identify and respond to opportunities for improvement. For example, a local council might receive thousands of comments from citizen surveys, staff suggestion schemes, and public complaints every year. While a manager might have time to read a few dozen, an AI model using Natural Language Processing can analyse all of them and detect subtle but recurring themes. The model might flag that residents consistently mention "confusion about recycling rules" or "difficulty navigating the online permit application," revealing a systemic bottleneck in communication or service design that was previously invisible in the quantitative data. This granular insight allows for the organisation to target the right bottleneck. In this manner, AI can support the development of the key organisational capabilities required to innovate and adapt. These foundational capabilities are discussed in greater detail in the following section.

3 The First Part of the Challenge: Organisational Capabilities

As established, the realisation of AI's potential is far from guaranteed, particularly when the political and institutional dynamics favour short-term visibility over long-term problem solving. To navigate this environment and deliver sustained improvement in productivity, five organisational-level capabilities are critical: (1) Organisational Learning, (2) Planning & Prioritisation, (3) Developing New Ways of Working, (4) Managing Risk, and (5) Managing Supplier Relations. These capabilities enable public bodies to identify and execute technically sound AI projects and to build the strategic resilience needed to champion them politically. Moreover, these capabilities enable public bodies not only to identify and resolve existing bottlenecks but also to develop the strategic vision and execution capacity required for the more ambitious work of reconfiguring services entirely. While the capabilities are deeply interconnected - for instance, organisational learning is a prerequisite for effective planning - they are separated here for analytical clarity.

Tables 1 and 2 in the appendix provide maturity tables describing how these capabilities manifest at different stages of development. The highest level, (1) Outstanding, describes an organisation where a capability is proactive and embedded in the culture, enabling it to systematically shape and exploit the best opportunities AI presents. Subsequent levels depict a degradation of this ideal. An organisation at the (2) Good level has formalised and consistent processes, allowing it to reliably act on opportunities. At (3) Adequate, capabilities are inconsistent and reactive, meaning the organisation can only capitalise on opportunities when conditions are favourable. At the lower end, (4) Needs Improvement illustrates a passive or siloed organisation where the absence of formal capability means opportunities are rarely identified or captured. The final level, (5) Inadequate, describes an organisation where structures are counter-productive, actively hindering innovation and ensuring that new technologies have a negligible or even negative impact on productivity.

Crucially, this framework should be applied not as a rigid, all-or-nothing checklist, but as a flexible diagnostic tool. This is because 'AI' is not a monolith; the capabilities required for a simple off-the-shelf automation tool are vastly different from those needed for a bespoke, high-stakes predictive system. A leader might therefore use this framework to determine that an 'adequate' level of capability is sufficient for a low-risk pilot, while a 'good' or 'outstanding' level is a necessary prerequisite for a more transformative, system-wide initiative. This allows the framework to function both as a guide for identifying the most impactful areas for intervention and as a sober assessment of the organisation’s readiness to successfully execute innovation.

3.1 Capability 1 – Organisational Learning

Organisational learning refers to the processes by which an organisation examines its delivery chain, not only to correct operational problems but, more importantly, to redesign the core assumptions and workflows that govern it. As described by Senge (1990) and Argyris & Schön (1997), it is this deeper learning that allows an organisation to identify and resolve the root causes of bottlenecks rather than merely treating superficial symptoms. In the context of AI, the greater level of maturity of this capability is what separates organisations that use technology for minor efficiencies from those that leverage it for transformative productivity gains. For example, if the root cause of missed deadlines in an organisation is that roles are poorly defined, then deploying a task-management platform is merely a superficial fix. The tool can only mitigate the symptom; it cannot create the clarity of responsibility that the organisation currently lacks, and which is severely impeding productivity.

The first aspect of organisational learning concerns the structured processes that the organisation can actively design and deploy to identify opportunities for improvement. One such process is the systematic and continuous monitoring of the delivery chain. Supported by instruments like operational dashboards and performance reviews (Kaplan & Norton, 2007), this ongoing activity generates the requisite data to identify areas where observed performance deviates from expected outcomes. This allows the organisation to prioritise areas that warrant a specific diagnostic investigation (Argyris & Schön, 1997). Through the analysis of case files, key performance metrics, and qualitative data from personnel, the investigation aims to identify the underlying determinants of the performance issue. This ensures that any subsequent intervention is precisely targeted at the systemic root cause, rather than only addressing the superficial symptoms.

Following this diagnostic work, when a significant intervention such as an AI adoption project is delivered, the learning cycle can be concluded with a third organisational learning process, which concerns a post-implementation evaluation (Dunn, 2015). The objective of such a review is to answer two fundamental questions: Firstly, did the project genuinely increase productivity? This requires an assessment of its impact across the delivery chain - including money saved (budgetary efficiency), reductions in the resources, such as staff time, required to deliver the service (operational efficiency), and the extent to which it altered the beneficial impact of the service (effectiveness). Second, what were the determinants of its performance? This requires analysing a wide range of data, such as system-generated error logs and frontline staff perspectives, to understand why the project succeeded or failed, and to uncover unintended consequences (ibid). The preservation and dissemination of these insights are then facilitated by established knowledge codification protocols, such as shared digital repositories (Nonaka & Takeuchi, 1995).

However, the efficacy of such formalised structures is mediated by the prevailing organisational culture. Their value is only fully realised in an environment where staff feel empowered to question established practices (Pärna & Von Tunzelmann, 2007). In a culture where such inquiry is discouraged, staff may see process flaws but remain silent for fear of being labelled unhelpful. This ensures that new technology is either misapplied, layers over deeper causes of inefficiency, or that the best opportunities for adoption are never brought to light. Fostering an innovative culture, in contrast, relies on mechanisms like communities of practice and by delegating authority for experimental pilots to frontline employees, which cultivates a sense of ownership and diffuses tacit knowledge (Fernandez & Moldogaziev, 2013). The literature demonstrates a strong correlation between these attributes of autonomy, ownership, and high levels of employee satisfaction and innovative performance (Hakro et al., 2022; Bartlett & Dibben, 2002). Furthermore, the example set by leadership is instrumental in reinforcing these norms by demonstrating openness to novel ideas and by actively encouraging the candid discussion of organisational challenges (Damanpour & Schneider, 2006; Pil & MacDuffie, 1996). By enabling such clear discussion on the potential for improvement, the organisation is more able to successfully find and take advantage of the right opportunities that AI presents.

3.2 Capability 2 – Planning & Prioritisation

Once organisational learning has revealed the bottlenecks in the delivery chain, the challenge shifts to one of execution. Chiefly, this concerns whether the organisation can commit its resources to the solution that promises the greatest impact on productivity (Goldratt, 1990; van Ark, 2022; van Ark et al., 2023). For example, a local government department responsible for processing planning permit applications may face a bottleneck where a limited number of planners are overwhelmed by a high volume of complex applications. The bottleneck in this process is the manual review stage, and it will dictate the overall speed at which permits are issued, regardless of how quickly applications are submitted or how efficient the final approval step is. Even if the department invests heavily in a user-friendly online submission portal or speeds up the final printing of permits, the fundamental delay caused by the manual review bottleneck will persist, limiting the total number of applications processed per month. Applying the principles of the Theory of Constraints, the most effective way to increase the department's productivity is to direct investment and resources towards alleviating this specific constraint.

However, the decision to invest in a long-term AI project requires navigating a difficult tension between the need for patient investment and the pressure for immediate results. The critical role of resource allocation in this tension is well-documented. The literature consistently shows that organisational innovativeness is strongly predicted by the protection of 'slack resources', which are the discretionary funds or staff time not committed to core operations (Damanpour, 1991; de Vries et al., 2016), and which can be used for exploring, testing, and refining opportunities for improvement. As Flyvbjerg and Gardner (2023) emphasise, realising productivity gains requires maintaining a stable project environment. That requires insulating projects from premature reprioritisation, since the imperative to manage short-term fiscal pressures often results in the diversion of resources from long-term, uncertain projects to meet immediate demands (ibid). Stability is also essential to maintain staff engagement, because employees will not expend discretionary effort on initiatives they reasonably expect to be cancelled (Reichers et al., 1997).

The attainment of productivity from AI requires investing not just in the technology itself, but also in its key complementarities: the additional infrastructure, data, and skills that are essential for the successful usage of the equipment. Damanpour et al. (2009) find that focussing on one innovation is sub-optimal, and this reflects the fact that innovations are often dependent on the infrastructure that precedes them. Planning to take advantage of complementarities is particularly important for artificial intelligence, whose efficacy hinges upon the availability of vast quantities of data for training and optimisation (Whang et al., 2022). As such, AI applications can be complemented by investments in digital infrastructure to collect, store, and process these data. Complementarities can also exist between organisations, as AI performance often improves as more organisations share data (Gregory et al., 2021).

Finally, beyond the technical evaluation of returns on investment, effective planning in the public sector also considers how proposals can be selected and augmented to optimise their political viability. Public managers do not operate in a technocratic vacuum; they must construct a politically viable proposal that can thrive in their 'authorising environment' – which concerns the complex web of political and administrative actors whose support is necessary for the project to proceed (Moore, 1997). This requires a deliberate act of framing (Entman, 1993), where a complex initiative is translated into a coherent policy narrative. By framing a project not as a mere technical upgrade but as a tangible instrument for achieving a stated ministerial objective - be it 'reducing the burden on the taxpayer' or 'cutting waiting times' - the manager transforms the proposal from a budgetary item into a political asset. A compelling narrative provides a politician with the simple, powerful story needed to build the initial coalition required to justify the investment and secure approval (Jones et al., 2014). Moreover, the narrative can act as a durable political asset by providing a compelling and transferable rationale for stakeholder buy-in. This can be used to ensure that the project maintains the necessary stakeholder buy-in through inevitable changes in personnel and leadership in the authorising environment (ibid).

3.3 Capability 3 – Developing New Ways of Working

Even the most well-planned and resourced AI initiative is unlikely to realise its full potential if it is not matched by the capability to develop new ways of working. The necessity of this capability stems from the potential of technology, including AI, to automate certain tasks and augment others, which in turn can significantly alter skill requirements, task distribution, and entire occupational structures (Autor et al., 2003). Furthermore, the effects of AI on the workforce may not follow a single, predictable pattern. Earlier waves of technology, for instance, often primarily benefited the most highly skilled workers, widening the performance gap between them and their less experienced colleagues (ibid). In contrast, emerging evidence on generative AI suggests a different, and potentially levelling, dynamic. In a study of one AI application, Brynjolfsson et al. (2025) found that the greatest performance gains accrued to less experienced workers, as the technology provided them with access to the tacit knowledge that was previously the domain of experts.

The identification and implementation of these reorganisational efforts are critically dependent on the troubleshooting and discretionary efforts of managerial staff (Heyden et al., 2018; Tarakci et al., 2023). As evidenced in South Cambridgeshire District Council's four-day week initiative, active managerial involvement was essential for navigating the operational frictions that invariably arise during such a change. Their efforts included not just identifying training needs, but also proactively modifying established workflows by introducing new 'rules' like making phone calls instead of writing emails, implementing joint management of team diaries, and scheduling mandatory office days to ensure collaboration continued (Jörden & Alayande, 2023). Conversely, weaknesses in the managerial oversight and problem-solving capacity can severely undermine even well-planned projects, as unresolved issues can delay or undermine the work of others on the project, leading to cost overruns and failure to achieve intended benefits (Winch & Leihringer, 2016).

Successful implementation of organisational changes frequently hinges on the quality of the relationship between managers and the workforce. This is because any significant change imposes real costs on staff, including both the short-term disruption of established routines and the long-term threats to professional identity or job security (Allison, 1969; March & Olsen, 1984). Trust serves to mitigate the uncertainty of a proposed change by giving employees confidence that managers will handle the inevitable disruptions and trade-offs fairly (Bouckenooghe, 2012; Oreg et al., 2011). This confidence increases the likelihood that professional staff will engage constructively with the change process (ibid). Conversely, a history of poorly managed change breeds a powerful cynicism that any short-term sacrifices will not be made for an ultimate benefit (Reichers et al., 1997). This cynicism is not passive; it can result in workers forming powerful blocks to change. Pettigrew et al. (1992) provide a stark example, finding that the "historically hostile and indifferent" clinicians in their case study were able to form a "powerful block on change," forcing hospital managers to abandon any attempt at direct implementation of reforms in favour of a drawn-out, multi-year negotiation.

Building trust between management and professional staff is therefore an investment in a long-term asset that raises the likelihood of success for future improvement projects. The investment in this asset does not just require financial expenditure but also involves the deliberate sacrifice of managerial resources that could otherwise be used to meet short-term performance demands (Morrison & Robinson, 1997). These costs include the relinquishment of efficiency to ensure fair processes, compensation for negatively affected workers, the surrender of unilateral control to facilitate employee participation, and the expenditure of executive time on transparent communication (ibid, Oreg et al., 2011). In doing so, management trades immediate operational convenience and control for the long-term returns of trust.

3.4 Capability 4 – Managing Risk

The adoption and implementation of frontier technology like AI is inherently accompanied by uncertainty regarding the necessary investment, or the ultimate performance of the application (Hagendorff & Wezel, 2020). Project teams do not always know in advance the true complexity of the development process, the final system's precise capabilities and limitations, or all the real-world conditions it will encounter (ibid). If these uncertainties are not systematically addressed during development and deployment, then significant latent flaws can be built into the technology without being detected (Kordzadeh & Ghasemaghaei, 2022). These flaws create a high risk of both operational failure, where the system fails to deliver its promised productivity benefits, and ethical failure, where it inflicts adverse consequences upon the public. For instance, an AI-powered recruitment tool developed by Amazon was found to perpetuate gender biases learned from historical data (Dastin, 2018). This case illustrates how a technical weakness, born from an unexamined assumption about the neutrality of past hiring data, created several kinds of failure: an ethical failure through discrimination and a potential organisational failure through legal and reputational damage.

Effective risk management relies on a suite of practices that integrate both preventive and diagnostic techniques. Preventive efforts aim to design out risk from the start, for example by reducing the use of sensitive data, which may trade predictive power for stronger privacy protection (Whang et al., 2022). These are supplemented by diagnostic techniques that seek to uncover latent weaknesses, such as 'pre-mortems' that reason backward from hypothetical failure or the testing of system performance in practice (Klein, 2007; Spring et al., 2020). A key tool that reinforces both is proactive engagement with stakeholders. Such engagement can function as a diagnostic tool by discovering external risks and blind spots, while also acting preventatively by ensuring the project's design is shaped by and respects the public risk appetite (Kallina & Singh, 2024). By demonstrating this level of responsiveness and competence, an organisation earns the social license needed to experiment with more transformative technologies, as public trust often stems more from the perceived competence of the organisation than from the technical safety of the technology itself (Personal Data Protection Commission, 2020).

However, the mere availability of such tools is not sufficient. Their consistent and effective use depends on three deeper organisational factors: the presence of skilled human capital, clear accountability, and a culture that can learn from failure.

Firstly, sufficient human capital is critical for ensuring project managers possess the requisite skills and experience to navigate the complex risks and technical challenges of AI (Zick et al., 2024). Crucially, this includes the experience and education that helps project managers to avoid common cognitive biases like excessive optimism, which might otherwise prevent the identification of key weaknesses (Flyvbjerg, 2021).

The second factor is an oversight structure that clarifies lines of responsibility by linking project outcomes back to managerial decisions (Flyvbjerg et al., 2002). Such systems also serve as a corrective against the common incentive to misrepresent costs or capabilities in order to secure approval (ibid). Mechanisms such as ethics boards can help ensure risks remain within organisational and societal tolerance, while gradually reshaping internal culture. For example, the preliminary review of the West Midlands Police's Data Ethics Committee found that, over time, the formal governance mechanism began to reshape internal culture, and the force's "Data Analytics Lab anticipates the concerns of the Committee in its approach to new projects" (Oswald et al., 2024). This case illustrates how mature oversight structures can encourage managerial staff to identify and manage risks from the earliest stages of the project cycle, rather than mitigate them only once problems emerge.

A third factor is an oversight structure that is designed to learn from failure, which requires a formal capacity to distinguish between 'intelligent failures' at the frontier of experimentation and 'preventable failures' resulting from clear deviations from protocol (Edmondson, 1999; 2018). However, traditional public sector accountability is ill-suited to this, as it tends to emphasise failure more than success (OECD, 2019, p. 46). This emphasis can foster a risk-averse culture where managers focus primarily on minimising their personal liability (PASC, 2011). As a result, weaknesses that undermine the productivity-enhancing potential of new technologies often go unaddressed: managers may avoid taking responsibility for troubleshooting or even suppress information about emerging issues (Edmondson, 2018). This pressure can distort risk management tools into one of two unproductive forms: a rigid, innovation-killing compliance regime that prevents AI adoption, or performative 'ethics washing' that merely gives the impression of managing risk (Bietti, 2020). By contrast, accountability frameworks designed for learning rather than blame can encourage the experimentation and open dialogue necessary to attain productivity from AI.

3.5 Capability 5 – Managing Supplier Relations

The development of AI applications frequently involves collaboration between a client organisation with operational knowledge and a supplier with technological expertise. Successful development requires bridging this information gap: suppliers need clear user requirements while clients need a clear account of the technology's capabilities and limits. However, the information asymmetry can be exploited, with suppliers able to overstate potential benefits, and understate expected costs or timelines (PASC, 2011; Gallivan & Oh, 1999). Moreover, to avoid immediate remediation costs, suppliers may underplay technical vulnerabilities (Holgeid & Thompson, 2013) like model drift or data leakage. This can create significant risks of operational and ethical failure that only manifest once the technology is in use, leaving the user to pay for remediation. Suppliers may also lock clients in through proprietary algorithms, non-portable data formats or bespoke interfaces. These raise the cost of switching to another supplier and so weaken the client’s leverage over price and quality (Farrell & Klemperer, 2007).

The primary remedy is to maintain sufficient in-house technical capability so that the organisation can act as an ‘intelligent customer’ (Holgeid and Thompson, 2013). This capability is characterised by specific skills for monitoring, evaluating, and challenging suppliers, including technical due diligence such as code review, the ability to run independent performance evaluations, and expertise in contract negotiation (PASC, 2011). With these capacities, the client can engage critically with suppliers throughout the project and assess whether a proposed tool is likely to meet its stated performance goals. In addition, an informed client is better placed to spot neglected work or strategic behaviour, for example the downplaying of vulnerabilities or the deliberate design of components to create supplier lock-in. Without sufficient in-house expertise to diagnose problems and assess delivery, an organisation's ability to manage its technology systems and to protect public investment is fundamentally undermined (McManus and Wood Harper, 2007).

A modular procurement approach is widely regarded as best practice for reducing project risk, although coordinating multiple suppliers raises the bar for internal capability (Flyvbjerg and Gardner, 2023; Baldwin and Clark, 2000). Breaking a large programme into smaller, modular sub-projects or discrete lifecycle stages counters the well-documented tendency of large, long-term contracts to incur greater average cost overruns and risks of failure (Sauer et al., 2007; Holgeid & Thompson, 2013). Multiple smaller contracts can increase competition, reduce the impact of a single underperforming supplier, and limit lock-in because individual components can be replaced without jeopardising the whole system. This modularity increases the client's control and lowers future transfer costs. It does, however, also require greater technical and managerial capability from the client, since the integrator role requires the client to complete more tasks relating to mediating, monitoring, and co-ordinating multiple suppliers (Kappelman et al., 2006; Lacity & Willcocks, 2000).

Despite the advantages of disaggregation, public sector organisations commonly default to a consolidated model in which a single supplier is given end-to-end responsibility (Holgeid & Thompson, 2013). That choice reflects a tension in incentives: consolidated contracting offers managers a simple, external point of accountability, which can be attractive where the standard organisational response to failure is to search for an individual or group that can be blamed (Hood, 2010). Selecting a single large contract thus becomes a rational strategy for individuals seeking to minimise personal exposure to blame. The result is a transfer of political risk to the supplier that often comes at the cost of a greater likelihood of programme failure.

4 The Second Part of the Challenge: Managing the Public Sector Innovation System

The development of organisational capabilities is essential but not sufficient for realising the full value of AI. This is because the ambitions of a perfect organisation would still be constrained by a dysfunctional innovation system, where systemic challenges like fragmented data standards, skills shortages, or non-competitive supplier markets lie beyond the remit of individual organisations to solve.

The resolution of these sector-level barriers is the defining role of what this report terms 'strategic institutions': public or quasi-public bodies with a formal or widely recognised mandate for sector-wide stewardship. These may be government departments, co-ordinating bodies, professional bodies, or independent charities, but all shape the innovation ecosystem by providing collective assets and setting the rules of interaction. A prominent example of this logic can be seen in the recommendations of the U.S. Commission on Evidence-Based Policymaking. The Commission identified systemic barriers to using government data effectively and proposed the creation of a National Secure Data Service, which would be a central, trusted entity designed to provide the collective data infrastructure and standards that no single agency could create on its own (Commission on Evidence-Based Policymaking, 2017).

This section outlines seven key sector-level capabilities that these institutions can exercise to encourage AI-driven productivity gains. To illustrate how these capabilities manifest in practice, a sectoral maturity framework is presented in the appendix (Tables 3 - 5). This framework shows how the entire innovation system can develop, moving from a state where key issues are unaddressed to one where they are coherently and strategically managed. It is important to note that a sector's maturity level is often the result of the collective, and sometimes uncoordinated, actions of multiple strategic institutions. Conceptually, this means that a system might achieve a 'Good' level of maturity through the fragmented but active efforts of several different bodies. Reaching an 'Outstanding' level, however, would likely necessitate a more deliberately co-ordinated approach.

4.1 Capability 6 – Strengthening Innovation Networks

Innovation in complex domains such as AI depends on combining specialist technical expertise with a grounded understanding of frontline needs and constraints (OECD/BCG/INSEAD, 2025). These forms of expertise are often siloed both within organisations, and across government, academia, and industry, limiting the development of solutions that are both technically robust and practically useful. A key role for strategic institutions is therefore to bridge these silos and connect relevant actors (Provan and Lemaire, 2012). Examples of practical bridging mechanisms include secondments and fellowships, conferences, regular cross-sector workshops, and shared repositories of case studies.

By serving as channels for communication, such networks support organisational learning across the sector by allowing actors to build on existing knowledge, share lessons from both successes and failures, and iterate new approaches (Pärna and Von Tunzelmann, 2007). The innovation literature consistently finds that external organisational networks and active engagement with prior adopters drive greater likelihood of both adopting innovation and attaining productivity from that technology (Burns and Wholey, 1993; Damanpour and Schneider, 2006; Young et al., 2001). This impact is also clear at the individual level; for example, Barrenho et al. (2025) found that a one standard deviation increase in the size of a surgeon's professional network raised their likelihood of adopting a new surgical innovation by 16 per cent. Informal information exchange systems can therefore be a vital asset for turning technological potential into productivity growth.

Beyond informal connectivity, strategic institutions can also play a crucial role in solving the collective action problem of knowledge sharing. Although individual organisations frequently generate useful insights from their AI initiatives, these learnings often remain tacit or unpublished, meaning that their value is not shared across the system. However, codifying and sharing such experience faces two barriers: First, the process itself can be costly and complex. Second, organisations can be deterred by the strategic risks involved, such as reputational scrutiny or the misinterpretation of failures (Nonaka & Takeuchi, 1995). Together, these barriers lead to a chronic underproduction of publicly available knowledge. Strategic institutions can counteract this by establishing clear expectations and incentives for sharing learning. Mechanisms such as standardised reporting templates, shared case study repositories, and peer review programmes can institutionalise collective intelligence collection, and so turn knowledge codification into a standard and expected practice.

4.2 Capability 7 – Providing Clarity on Collaborative Performance

Co-development efforts can stall when public and private actors lack clear information about each other's track records, such as their ability to deliver high-quality AI tools or to act as responsive, capable clients. Without this mutual clarity, the high uncertainty surrounding partner performance can lead to a "lemons problem" (Akerlof, 1970), where the inability to distinguish reliable partners leads to persistent underinvestment, as risk-averse organisations opt out rather than risk failure.

To address this, strategic institutions can maintain mechanisms for reducing uncertainty about the reliability of potential development partners (Ruttan, 2001). On the supplier side, this includes creating and maintaining publicly accessible registers or certification systems for AI vendors, detailing their expertise and past performance, thereby providing a valuable signal of quality (Spence, 1973). On the client side, it can involve establishing accreditation schemes or 'digital readiness' assessments for public bodies themselves, signalling to the market which organisations have the in-house capability to be effective innovation partners. Both sides are reinforced by the establishment of transparent procurement frameworks and standardised contractual terms that clearly delineate responsibilities and reduce the scope for opportunistic behaviour (Williamson, 1985).

4.3 Capability 8 – Funding and Co-ordinating Research

Left to its own devices, the private market is unlikely to produce the specific AI innovations the public sector needs most. This is because public sector-relevant research often suffers from a series of market and co-ordination failures. The benefits of the research can be too diffuse for a single organisation to capture (a public good problem); the upfront investment may be too large for any single agency to bear; and the demand from individual public bodies is often too fragmented to provide a credible incentive for private R&D (Mazzucato, 2011; Bush, 1945). Left unaddressed, these issues lead to a systemic under-provision of the very research that could drive productivity growth in the public sector.

Strategic institutions play a critical role in correcting these failures by cultivating a research ecosystem that supports innovation across the public sector. They can achieve this through four key functions:

Firstly, they can finance foundational research. This type of research cultivates a broad base of scientific understanding whose benefits often spill over widely, meaning they are not fully captured by any single entity (Bush, 1945). A classic example in the field of AI is the early research into the mathematics of neural networks. During the so-called "AI winter" of the 1980s and 1990s, this work was largely curiosity-driven with no immediate commercial application and was advanced by a small number of academic researchers supported by public grants (Toosi et al., 2021). The immense, society-wide benefits of this research, which now form the basis of the entire modern AI industry, only became apparent decades later (ibid). Because individual organisations can only ever capture a fraction of these wide-ranging societal benefits, they lack the incentive to invest at the level that is optimal for the public good. This leads to a systemic underinvestment in the knowledge that the entire sector would benefit from. Strategic institutions can solve this collective action problem by funding this research directly.

Secondly, strategic institutions can "crowd in" further research and development by underwriting the substantial fixed costs associated with establishing specialised research institutes or shared technological platforms. Such investments in the underlying innovation infrastructure can stimulate and enable more targeted R&D contributions from individual public sector organisations (Lundvall, 1992; Freeman, 1987).

Thirdly, strategic institutions can create a credible incentive for private firms to invest in public sector R&D by reducing uncertainty regarding the demand for the end product. They can achieve this by aggregating and co-ordinating demand - for example, by establishing a set of common capability requirements that signals to the market what novel technologies a large number of individual public bodies are likely to procure (Mazzucato, 2011).

Fourthly, they can fund high-risk, high-reward applied research that individual organisations are unable to bear (Foray, et al., 2009). A salient example is explainable AI (XAI), which creates models with transparent and interpretable rationales for their decisions (Gunning et al., 2019; Adadi & Berrada, 2018). XAI is 'high-risk' due to formidable technical challenges, uncertain development timelines, and the need for long-term funding without immediate returns. However, if research in this area can be successfully advanced, then it directly addresses key barriers to AI adoption in the public sector. Firstly, it mitigates legal and ethical concerns over opaque decisions that can undermine principles like due process (Pasquale, 2015; Diakopoulos, 2014), thereby lowering organisational resistance. Secondly, it enhances AI's utility by enabling robust quality assurance, which simplifies integration into existing processes (Minh et al., 2022; Doshi-Velez & Kim, 2017). By funding such high-risk innovation, strategic organisations can facilitate a greater range of available technologies, that are more able to meet the needs of the public sector.

4.4 Capability 9 – Governing Failure

Realising productivity from AI presents a core governance conundrum. When the costs of failure are externalised onto citizens, public bodies lack a powerful incentive to manage the risk of that failure, rendering self-regulation inadequate. Imposing legal liability is the conventional response (Bovens, 2007), but it is a double-edged sword. While it forces organisations to internalise these societal costs and helps to ensure public trust (ibid), it can also foster a punitive culture of blame avoidance, as the pressure of corporate liability is often translated into a search for individual culpability (Hood, 2010). This prompts staff to minimise their personal exposure through strategic actions: suppressing critical information, deliberately not seeking it out, outsourcing key decisions, or burying inconvenient facts within a flood of disclosed data (ibid). This stifles the very open reporting of error and iterative learning that is essential for attaining productivity from innovation, and creates a fundamental tension between accountability and learning.

The National Programme for IT in the NHS offers a clear illustration of how a purely liability-focused governance model, intended to internalise the cost of failure in the contractor, in fact ended up poisoning the collaborative relationships needed for success. Under this system, massive and long-term contracts were awarded to ‘Local Service Providers’ (LSP) who were responsible for delivering all new front-line IT systems in a given region, with payment strictly contingent on meeting rigid, pre-defined milestones (National Audit Office, 2011). The aim was to pass all financial risk onto the contractor, who would then be the single point of responsibility and could easily be “held accountable” when issues emerged. In practice, this liability-focussed governance regime created powerful perverse incentives - the intense fear of triggering penalties led to LSPs consistently avoiding the admission of any kind of error and focussing only on satisfying the literal terms of their contracts (ibid). This was one of the factors that led to the failure of the programme, as necessary adaptions were resisted by the contractors, and minimal effort was made to collaboratively identify and resolve the risks of operational or ethical failure.

The failure of such purely liability-focused models often prompts calls to adopt systems geared towards learning, of which the aviation sector is a preeminent example. Its use of confidential, blameless error reporting has proven highly effective at generating systemic learning from near-misses and mistakes (CHIRP Charitable Trust, 2025; Pellegrino, 2019). However, this model's success is contingent on a direct alignment of organisational and public interests, where a catastrophic ethical failure (loss of life) is inseparable from total operational and financial failure. This alignment can be absent in the public sector. A social benefits agency, for instance, might achieve its cost-cutting targets precisely by creating a severe ethical failure in wrongly denying benefits to legitimate claimants. Therefore, a learning system alone is insufficient when powerful institutional incentives exist to tolerate such failures (O’Neil, 2016).

Successfully navigating this liability-learning paradox depends partly on how strategic institutions design the sector's legal, investigatory, and review processes to change a manager's rational calculus around disclosure. This dilemma is starkly illustrated by the choices facing a manager who discovers a potential weakness in a new technology. In a system where accountability is punitive and unpredictable, the professional and personal risk of disproportionate blame creates a powerful incentive for suppression (Hood, 2010). Even if disclosure is unavoidable, the rational choice is often to minimise liability by providing information that is defensive, legalistic, and which fails to clearly state the tacit knowledge required for genuine organisational learning or for providing stakeholders with a meaningful account of performance (ibid).

A predictable threshold for independent review and a well-trusted review process can fundamentally alter this decision-making calculus (Bovens, 2007; Majone, 1997). When such a system can differentiate between failures - such as justifiable experimentation (which requires disseminating findings), blameless human error (which requires system improvements), at-risk behaviour (which requires coaching), and reckless conduct (which may require sanction) - it removes the concern that disclosure opens the manager to disproportionate blame (Edmondson, 2018; Reason, 1998). This, in turn, allows managerial staff to instead consider the potential positive impacts that disclosure can have. These benefits include the opportunity for organisational learning (Edmondson, 2018), the ability to secure support to fix the weakness, and an opportunity to build stakeholder trust by demonstrating a credible capacity to manage risk (Parker et al., 2025; Gunningham et al., 2004; Bovens, 2007). Such a framework therefore works by reducing the need to avoid blame, allowing the positive incentives for authentic transparency to drive behaviour. As a result, successfully balancing accountability and learning can help to convert transparency from an act of individual courage into a routine aspect of professional competence.

4.5 Capability 10 – Maintaining Pressure for Productivity Growth

Organisational inertia often arises when there is no significant discrepancy between expected performance and what is observed, leaving only a weak impetus for improvement (Cyert & March, 1963). This issue can be particularly acute in growing public service organisations, where consistent financial surpluses can conceal inefficiencies and lessen the willingness to confront difficult trade-offs (Pettigrew et al., 1992). This dynamic in turn suppresses the trigger for "problemistic search" (Cyert & March, 1963), which is the diagnostic inquiry managers undertake when performance falls short of aspirations. Without a recognised performance gap this search for innovation and greater efficiency is unlikely to occur (ibid).

To overcome organisational inertia, strategic institutions can apply constructive tension through two instruments: financial pressure and performance targets. The goal is not to replace the intrinsic motivation of many managers to improve their services (Christensen et al., 2017), but to focus attention on the difficult challenges of productivity and sustainability. Financial pressure operates by setting a predictable resource ceiling, such as a fixed budget or planned annual reductions. This predictable ceiling helps to highlight whether the current operating model is viable in the long run, forcing managerial staff to look beyond short-term spending choices, and consider whether they need to find productivity-enhancing opportunities to ensure their organisation's sustainability (Cyert & March, 1963). Performance targets complement this by converting the general drive for productivity into specific, measurable objectives (Locke & Latham, 1990). Together, these mechanisms make the need for strategic planning both visible and unavoidable.

The effectiveness of such mechanisms is often contingent on their perceived legitimacy. When seen as arbitrary, such pressures may secure reluctant compliance but fail to foster the discretionary effort and creative problem-solving on which genuine, productive change depends (Moorman, 1991; Greenberg, 1990). Benchmarking can be used by strategic institutions to ground financial pressures and performance demands in the demonstrated performance of peer organisations (Camp, 2024). This allows such pressures to be perceived as fair and achievable, which is then crucial for their acceptance by staff as meaningful goals rather than as arbitrary impositions. Goal-setting theory finds that where targets are seen as fair, specific, and appropriately challenging, they engender strong goal commitment (Locke & Latham, 1990). This commitment, in turn, leads to more focused effort, greater persistence, and motivates the learning and problem-solving that is required to leverage new technologies and attain the targets.

Furthermore, the effectiveness of a performance target depends on access to ongoing information that allows workers to track their progress against it (Locke & Latham, 1990; Kaplan & Norton, 2007). Without such information, a goal remains a static endpoint. With it, the goal becomes a guide for day-to-day decision-making, enabling teams to increase effort when behind schedule or critically examine processes that hinder progress. In this way, performance information can turn a distant target into a credible basis for immediate action and adjustment.

However, these instruments carry a significant risk of destabilising the improvement process if miscalibrated. Firstly, excessive or unpredictable financial pressure can lead to short term tactical behaviour, which endangers the long term resource commitments needed for complex innovation untenable (Staw et al., 1981; Pettigrew et al., 1992; Pollitt & Bouckaert, 2017; Flyvbjerg & Gardner, 2023). The resulting uncertainty erodes staff confidence and reduces the discretionary effort from them that is required to develop new ways of working and support organisational learning

(Reichers et al., 1997). Secondly, poorly designed performance targets can produce a narrow focus on measurable indicators and invite strategic gaming (Goodhart, 1984); for example, hospitals may meet waiting time targets by keeping patients in ambulances to avoid officially starting the clock on admission time, and so meet their goals without increasing productivity. An excessive volume of targets further diffuses attention and creates strategic incoherence, which is a well-documented problem in NHS management (Bevan & Hood, 2006). Moreover, when pressure becomes overwhelming, constructive problem search is replaced by reactive crisis management focused on immediate survival, which ultimately undermines the capacity to learn, to commit to long term projects, and to collaborate effectively for implementation (Staw et al., 1981).

4.6 Capability 11 – Ensuring a Flow of Human Capital

The public sector's capacity for innovation is often limited by systemic underinvestment in the specialised skills required for technologies such as AI. Individual organisations can be reluctant to invest heavily in training because newly skilled employees are in strong demand across the economy and can be recruited elsewhere, meaning the sponsoring employer does not capture the full return on its investment. When this rational calculation is repeated across many organisations, it yields a sector-wide shortfall in the competencies needed for innovation. This dynamic is well documented in the human capital literature (Becker, 1964) and is evident in critical public services like adult social care, which suffers from high managerial turnover and significant digital skills gaps (Hoskins & van Ark, 2025). This skills gap is widely perceived as a primary barrier to progress; in a recent survey, 71% of public administration managers (N=415) identified a lack of skills in their workforce as an obstacle to greater adoption of AI (Public First, 2024).

Strategic institutions, with their sector-wide remit, are positioned to address this collective challenge. A key function is to stimulate the development of human capital through the strategic subsidisation of training. This can involve covering the fixed costs of establishing new, specialised programmes or dedicated learning institutions, an approach that overcomes the high initial outlays that individual organisations are often unable or unwilling to bear. Alternatively, these institutions can subsidise the variable costs of upskilling, for example, by providing grants that enable public sector employees to attend external courses or attain higher education qualifications. This leverages existing educational infrastructure to build a more skilled workforce across the sector (Pollitt & Bouckaert, 2017).

Beyond directly supporting training, a complementary function of strategic institutions is to enhance the efficiency of the public sector labour market. They can do this by providing or endorsing robust certification and credentialing mechanisms. By establishing or validating qualifications in areas like AI ethics, data science, or digital project management, these institutions help to reduce information asymmetries between employers and potential employees (Spence, 1973). Such credentials serve as a reliable signal of an individual's skills, enabling public service organisations to more effectively identify and recruit the necessary talent.

Ultimately, ensuring a steady flow of human capital is foundational, as it is the talent and expertise of public servants themselves that underpins the sector's ability to exercise every other capability, from strategic planning to risk management (Ferlie & Ongaro, 2022, p. 83).

4.7 Capability 12 – Standardising and Scaling Best Practice

A key function for strategic institutions is to identify and disseminate best practice, which can take two distinct approaches. The first is a directive stance, where specific solutions are mandated. This can be counter-productive as it risks imposing technology that is misaligned with local context – thereby eroding autonomy and wasting resources (Pollitt & Bouckaert, 2017). For example, rural health systems in England were recently mandated to adopt mobile dental vans - a model successful in Cornwall but which many local leaders, facing different bottlenecks, considered poor value (Timmins, 2024). Such instances illustrate why the innovation literature lacks a clear consensus on the net benefit of such top-down directives (Osborne & Brown, 2013).

An alternative approach addresses this dilemma by providing background infrastructure that enables local innovation within a coherent framework. This entails establishing shared platforms, standards and technical protocols that operate as stable "design rules". The principal objective is to secure interoperability, and so allow disparate systems to interact successfully and exchange data easily (Fishenden & Thompson, 2013; David & Greenstein, 1990). By offering this common foundation, a strategic institution can support modular innovation that meets local needs while preserving the benefits of interoperability (Baldwin & Clark, 2000). The UK Government Digital Service illustrates this model: in response to long-standing concerns around disjointed digital systems (Korteland & Bekkers, 2008), rather than prescribing specific equipment, the GDS established a suite of common rules, data standards and reusable components (National Audit Office, 2017). This institutional arrangement promotes reusability, realises economies of scale and reduces the risk of vendor lock-in.

5 Discussion and Conclusion

The newfound promise of artificial intelligence confronts the public sector not with a new problem, but with a new arena for a struggle that it has faced many times before: the challenge of translating technological potential into public value. This report has argued that while the tools of AI are novel, the productive value they provide is conditional on building foundational capabilities, many of which proved just as essential in previous waves of technological innovation. Therefore, success hinges not on technological acquisition, but on the deliberate development of this suite of organisational and systemic competencies.

To guide this effort, this report has put forward a framework of twelve foundational capabilities - five at the organisational level and seven at the sectoral level - that can support organisations in finding and taking advantage of the opportunities presented by AI. At the organisation level, the report identifies: robust organisational learning to diagnose root-cause problems; disciplined planning and prioritisation to focus resources on true bottlenecks; proactively developing new ways of working to integrate technology with human capital; effectively managing risk to foster intelligent experimentation; and acting as an 'intelligent customer' to co-develop fit-for-purpose solutions. At the broader sectoral level, seven further capabilities are essential for creating a fertile innovation ecosystem. These include strengthening networks, co-ordinating research, maintaining pressure for productivity growth, and establishing governance systems that can both learn from failure and ensure accountability.

The development of these twelve capabilities is not a straightforward exercise - it is profoundly complicated by a political economy that frequently incentivises behaviours that are diametrically opposed to effective capability-building. For example, the value of learning from failure can be overwhelmed by a powerful political imperative to avoid blame. This fosters a risk-averse culture where the maintenance of procedural compliance is regarded as 'safer' than innovation, and so change is consistently resisted. Similarly, the long-term, stable commitment required for strategic planning is persistently challenged by the short-term electoral cycles and shifting agendas that drive political decision-making. This feature of the political economy of the public sector can create a systematic bias against the patient, often invisible work of building the capabilities that are outlined in this report.

The solution to this dilemma, however, lies not in insulating public managers from political realities, but in empowering them to shape those realities. An effective public sector depends on an effective relationship between the managerial and political domain, each contributing a distinct rationale for action – the manager has a mandate to focus on productivity, while the politician has the mandate to decide between the competing demands placed on the public sector. When managerial staff cannot fulfil their side of the relationship by supplying credible, evidence-based improvement options, the policy-making process does not stop. Instead, it defaults to criteria grounded solely in political expediency (Kingdon, 2014). Consequently, policy choices become driven by symbolic appeal and the prospect of immediate credit-claiming, rather than by the potential to enhance long-term productivity. This fuels a proliferation of ‘shiny projects’: highly visible initiatives timed to electoral cycles that siphon resources away from the longer-term reforms needed for sustained productivity growth. However, an agency that has the expertise to generate defensible, evidence-based proposals is able to recalibrate this balance, and so increase the likelihood that policy selections advance both organisational and political objectives in tandem.

While this tilted playing field makes achieving productivity gains from frontier technologies like AI more challenging, it also amplifies the value of the outlined capabilities. The discipline of planning and prioritisation can act as a bulwark against the short-termism of electoral cycles. A robust capability for managing risk and learning from failure provides the institutional resilience needed to counteract a culture of blame avoidance. By crystallising a shared, evidence-based understanding of what truly works, strong innovation networks can build a collective voice that can compete with the siren song of 'shiny projects'. In this light, building these capabilities is the most practical way for public managers to fulfil their mandate for long-term value creation within a system that often incentivises the opposite.

The accumulation of these capabilities can also create a positive feedback loop, by providing the credibility that allows an organisation to influence its statutory constraints. An organisation that can demonstrate its capacity for rigorous planning, intelligent procurement, and effective integration is in a far stronger position to make a credible case to the Treasury for budgetary flexibility. A sector that has widespread risk management and well-used disclosure mechanisms is a more trusted partner for Parliament to grant new regulatory permissions. Both of these features can be leveraged to further strengthen organisational capacity, and hence enable the adoption of more technology that makes a greater contribution to productivity.

AI alone will not fix the public sector. While tactical gains can be found in low-effort projects like deploying simple chatbots and off-the-shelf automation tools, these are merely the foothills of the opportunity. To achieve the greater, transformative prize of AI by reconfiguring service delivery chains and unlocking new forms of public service, the public sector must make sustained investment in building the organisational and sectoral capabilities outlined in this report.

The key policy implication is not to avoid these simpler projects, but to see them as a training ground. We recommend that policymakers and senior public sector leaders treat early AI projects as opportunities to begin building the muscles for finding and taking advantage of opportunities for improvement. The maturity frameworks presented in the appendix offer a roadmap for this journey, allowing an organisation to consciously use these initial projects to advance its capabilities. As they mature, so too can the ambition of the AI applications pursued, moving from workflow improvement to systemic transformation. This is the path to creating a genuinely more capable public sector, one that can translate the promise of new technologies into better public services for citizens and greater value for the taxpayer. In contrast, a public sector leader who looks to AI as a universal panacea will see their own organisation reflected back at them with both strength and weakness laid bare in unsparing detail. Where their organisation lacks capability, so will AI.

Bibliography

Adleberg, J., Wardeh, A., Doo, F. X., Marinelli, B., Cook, T. S., Mendelson, D. S., & Kagen, A. (2022). Predicting patient demographics from chest radiographs with deep learning. *Journal of the American College of Radiology*, *19*(10), 1151-1161.

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). IEEE Access, 6, 52138-52160.

Agrawal, A., Gans, J., & Goldfarb, A. (2022). *Prediction machines, updated and expanded: The simple economics of artificial intelligence*. Harvard Business Press.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. The Quarterly Journal of Economics, 84(3), 488-500.

Aldridge, S., Hawkins, A., & Xuereb, C. (2016). Improving public sector efficiency to deliver a smarter state. Retrieved from: https://quarterly.blog.gov.uk/2016/01/25/improving-public-sector-efficiency-to-deliver-a-smarter-state/

Alford, J., & Hughes, O. (2008). Public value pragmatism as the next phase of public management. The American Review of Public Administration, 38(2), 130-148.

Alharbi, S., Alrazgan, M., Alrashed, A., Alnomasi, T., Almojel, R., Alharbi, R., ... & Almojil, M. (2021). Automatic speech recognition: Systematic literature review. *Ieee Access*, *9*, 131858-131876.

Allison, G. T. (1969). Conceptual Models and the Cuban Missile Crisis. *The American Political Science Review*, 63(3), 689-718.

Argyris, C., & Schön, D. A. (1997). Organizational learning: A theory of action perspective. Reis, (77/78), 345-348.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, *118*(4), 1279-1333.

Bajo, J., Corchado, J. M., & Rodriguez, S. (2008). GR-MAS: multi-agent system for geriatric residences. In ECAI 2008 (pp. 875-876). IOS Press.

Bao, Y., Hilary, G., & Ke, B. (2022). Artificial intelligence and fraud detection. In *Innovative Technology at the Interface of Finance and Operations: Volume I* (pp. 223-247). Cham: Springer International Publishing.

Baldwin, C. Y., & Clark, K. B. (2000). Design rules: The power of modularity. MIT Press.

Barrenho, E., Gautier, E., Miraldo, M., Propper, C., & Rose, C. (2025). Innovation Diffusion Among Coworkers: Evidence from Senior Doctors. *Management Science.* 0(0)

Bartlett, D., & Dibben, P. (2002). Public sector innovation and entrepreneurship: Case studies from local government. Local Government Studies, 28(4), 107-121.

BBC. (2024). Fujitsu Japan remains tight-lipped on the Post Office scandal. Retrieved from https://www.bbc.co.uk/news/business-61020075

Becker, G. S. (1964). Human capital: A theoretical and empirical analysis, with special reference to education. National Bureau of Economic Research.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610-623.

Bevan, G., & Hood, C. (2006). What's measured is what matters: targets and gaming in the English public health care system. *Public administration*, *84*(3), 517-538.

Bietti, E. (2020, January). From ethics washing to ethics bashing: a view on tech ethics from within moral philosophy. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 210-219).

Boiko, D. A., MacKnight, R., & Gomes, G. (2023). Emergent autonomous scientific research capabilities of large language models. *arXiv preprint arXiv:2304.05332*.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies.* Oxford University Press.

Bouckenooghe, D. (2012). The role of organizational politics, contextual resources, and formal communication on change recipients' commitment to change: A multilevel study. *European Journal of Work and Organizational Psychology*, *21*(4), 575-602.

Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework 1. European law journal, 13(4), 447-468.

Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies 'Engines of growth'?. *Journal of econometrics*, *65*(1), 83-108.

Brewster, T. (2023). Exclusive: DHS Used Clearview AI Facial Recognition In Thousands Of Child Exploitation Cold Cases. Forbes. Retrieved from https://www.forbes.com/sites/thomasbrewster/2023/08/07/dhs-ai-facial-recognition-solving-child-exploitation-cold-cases/

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. The Quarterly Journal of Economics, qjae044.

Burns, L. R., & Wholey, D. R. (1993). Adoption and abandonment of matrix management programs: Effects of organizational characteristics and interorganizational networks. Academy of Management Journal, 36(1), 106-138.

Bush, V. (1945). Science, the Endless Frontier. United States Government Printing Office.

Camp, R. C. (2024). *Benchmarking: The search for industry best practices that lead to superior performance*. CRC Press

Causey, J. L., Zhang, J., Ma, S., Jiang, B., Qualls, J. A., Politte, D. G., ... & Huang, X. (2018). Highly accurate model for prediction of lung nodule malignancy with CT scans. Scientific Reports, 8(1), 9286.

CHIRP Charitable Trust. (2025). *Improving safety in the air and at sea*. Retrieved from https://chirp.co.uk/about-chirp/who-is-chirp/

Chen, M., Tworek, J., Jun, H., et al. (2021). Evaluating Large Language Models Trained on Code. *arXiv preprint arXiv:2107.03374*

Christensen, R. K., Paarlberg, L., & Perry, J. L. (2017). Public service motivation research: Lessons for practice. *Public Administration Review*, *77*(4), 529-542.

Commission on Evidence-Based Policymaking. (2017). *The promise of evidence-based policymaking: Report of the Commission on Evidence-Based Policymaking*. U.S. Government Publishing Office. https://www2.census.gov/adrm/fesac/2017-12-15/Abraham-CEP-final-report.pdf

Cortés-Cediel, M. E., Cantador, I., & Gil, O. (2017). Recommender systems for e-governance in smart cities. In J. Yang, Z. Sun, A. Bozzon, J. Zhang, & M. Larson (Eds.), Proceedings of the international workshop on citizens for recommender systems - CitRec'17, Como, Italy, August 27, 2017. ACM press.

Coyle, D., & Hampton, L. (2024). 21st century progress in computing. *Telecommunications Policy*, *48*(1), 102649.

Cyert, R. M., & March, J. G. (1963). A Behavioral Theory of the Firm. Prentice-Hall.

Damanpour, F. (1991). 'Organizational Innovation: A Meta-Analysis of Effects of Determinants and Moderators'. Academy of Management Journal, 34(3), 555–90.

Damanpour, F., & Schneider, M. (2006). Phases of the adoption of innovation in organizations: effects of environment, organization and top managers. British Journal of Management, 17(3), 215-236.

Damanpour, F., Walker, R. M., & Avellaneda, C. N. (2009). Combinative effects of innovation types and organizational performance: A longitudinal study of service organizations. Journal of Management Studies, 46(4), 650-675.

Dastin, J. (2018, October 10). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. Retrieved from https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. Harvard Business Review, 96(1), 108-116.

David, P. A., & Greenstein, S. (1990). The economics of compatibility standards: An introduction to recent research. Economics of Innovation and New Technology, 1(1-2), 3-41.

De Vries, H., Bekkers, V., & Tummers, L. (2016). Innovation in the public sector: A systematic review and future research agenda. Public Administration, 94(1), 146-166.

Department of Health & Social Care, NHS England. (2022). *A plan for digital health and social care*. GOV.UK. Retrieved from https://www.gov.uk/government/publications/a-plan-for-digital-health-and-social-care/a-plan-for-digital-health-and-social-care

Diakopoulos, N. (2014). Algorithmic Accountability Reporting: On the Investigation of Black Boxes. Tow Center for Digital Journalism.

Doshi-Velez, F., & Kim, B. (2017). Towards A Rigorous Science of Interpretable Machine Learning. arXiv preprint arXiv:1702.08608.

DSIT, & GDS [Department for Science, Innovation, and Technology, & Government Digital Service]. (2025). *State of digital government review* (Command Paper No. CP 1251). GOV.UK. Retrieved from: https://www.gov.uk/government/publications/state-of-digital-government-review/state-of-digital-government-review

Dunn, W. N. (2015). *Public policy analysis*. *Routledge*.

Edelman, M. J. (1985). The symbolic uses of politics. *University of Illinois Press*.

Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. *Administrative science quarterly*, *44*(2), 350-383.

Edmondson, A. C. (2018). The fearless organization: Creating psychological safety in the workplace for learning, innovation, and growth. John Wiley & Sons.

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*, *384*(6702), 1306-1308.

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of communication*, *43*(4), 51-58.

Equality Act 2010 (c. 15). Retrieved from: https://www.legislation.gov.uk/ukpga/2010/15/contents

Farrell, J., & Klemperer, P. (2007). Coordination and lock-in: Competition with switching costs and network effects. In M. Armstrong & R. Porter (Eds.), Handbook of Industrial Organization (Vol. 3, pp. 1967-2072). Elsevier.

Ferlie, E., & Ongaro, E. (2022). *Strategic management in public services organizations: Concepts, schools and contemporary issues*. *Routledge*.

Fernandez, S., & Moldogaziev, T. (2013). Using Employee Empowerment to Encourage Innovative Behavior in the Public Sector. Journal of Public Administration Research and Theory, 23(1), 155–187.

Fishenden, J., & Thompson, M. (2013). Digital government, open architecture, and innovation: why public sector IT will never be the same again. *Journal of public administration research and theory*, *23*(4), 977-1004.

Fleming, N. (2004, October 12). Bill for hi-tech NHS soars to £20 billion. The Daily Telegraph. London. Archived from the original on 5 May 2013. Retrieved from https://www.telegraph.co.uk/news/uknews/1473927/Bill-for-hi-tech-NHS-soars-to-20-billion.html

Flyvbjerg, B. (2021). Top ten behavioral biases in project management: An overview. Project Management Journal, 52(6), 531-546.

Flyvbjerg, B., & Gardner, D. (2023). How big things get done: The surprising factors that determine the fate of every project, from home renovations to space exploration and everything in between. Signal.

Flyvbjerg, B., Holm, M. S., & Buhl, S. (2002). Underestimating costs in public works projects: Error or lie?. *Journal of the American planning association*, *68*(3), 279-295.

Foray, D., David, P. A., & Hall, B. H. (2009). Smart Specialisation – The Concept. Knowledge Economists Policy Brief No. 9. European Commission, DG Research.

Freeman, C. (1987). Technology Policy and Economic Performance: Lessons from Japan. Pinter Publishers.

Gallivan, M. J., & Oh, W. (1999, January). Analyzing IT outsourcing relationships as alliances among multiple clients and vendors. In *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers*. IEEE.

Goldratt, E. M. (1990). Theory of constraints. North River Press.

Goodhart, C. A. E. (1984). Monetary Theory and Practice: The UK Experience. *Macmillan*.

Greenberg, J. (1990). Organizational justice: Yesterday, today, and tomorrow. *Journal of management*, *16*(2), 399-432.

Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakou, H. (2021). The role of artificial intelligence and data network effects for creating user value. Academy of Management Review, 46(3), 534-551.

Guida, M., Caniato, F., Moretto, A., & Ronchi, S. (2023). The role of artificial intelligence in the procurement process: State of the art and research agenda. Journal of Purchasing and Supply Management, 29(2), 100823.

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI—Explainable artificial intelligence. Science Robotics, 4(37), eaay7120.

Gunningham, N., Kagan, R. A., & Thornton, D. (2004). Social license and environmental protection: why businesses go beyond compliance. *Law & social inquiry*, *29*(2), 307-341.

Hagendorff, T., & Wezel, K. (2020). 15 challenges for AI: or what AI (currently) can't do. AI & Society, 35, 355-365.

Hakro, T. H., Siddiqui, M. B., & Banbhan, A. A. (2022). Investigating the relationship of talent management on employee performance, employee engagement and job satisfaction. Research Journal of Social Sciences and Economics Review, 3(3), 10-20.

El Hannani, A., Errattahi, R., Salmam, F. Z., Hain, T., & Ouahmane, H. (2021). Evaluation of the effectiveness and efficiency of state-of-the-art features and models for automatic speech recognition error detection. *Journal of Big Data*, *8*(1), 5.

Heyden, M. L., Sidhu, J. S., & Volberda, H. W. (2018). The conjoint influence of top and middle management characteristics on management innovation. Journal of Management, 44(4), 1505-1529.

HM Treasury. (2025). Departmental efficiency plans. GOV.UK. Retrieved from: https://www.gov.uk/government/publications/departmental-efficiency-delivery-plans/departmental-efficiency-plans

Holgeid, K., & Thompson, M. (2013, June). *A reflection on why large public projects fail*. In *The Governance of Large-Scale Projects* (pp. 219-244). Nomos Verlagsgesellschaft mbH & Co. KG.

Hood, C. (2010). The blame game: Spin, bureaucracy, and self-preservation in government. Princeton University Press.

Hoskins, J., & van Ark, B. (2025). Productivity and responsible AI in adult social care. *The Productivity Institute*, University of Manchester.

Incubator for Artificial Intelligence. (2025). Developing and evaluating an AI capability. AI.gov.uk. Retrieved from https://ai.gov.uk/blogs/developing-and-evaluating-an-ai-capability/

Jackson, J., Bradford, B., Stanko, E. A., & Hohl, K. (2013). Just authority? Trust in the police in England and Wales. Routledge.

Jacobs, A. M. (2011). *Governing for the long term: Democracy and the politics of investment*. Cambridge University Press.

Jones, M. D., Shanahan, E. A., & McBeth, M. K. (Eds.). (2014). The science of stories: applications of the narrative policy framework in public policy analysis. Palgrave Macmillan.

Jörden, N., & Alayande, A. (2023, April). South Cambridgeshire District Council four-day work week trial: Insights from a focus group study with management and elected members. Bennett Institute for Public Policy, University of Cambridge. Retrieved from: https://bennettschool.cam.ac.uk/wp-content/uploads/2023/07/Four-day-work-week-focus-group-with-managers-and-elected-members.pdf

Kaplan, Robert & Norton, David. (2007). Using the balanced scorecard as a strategic management system. Harvard Business Review. 85.

Kappelman, L. A., McKeeman, R., & Zhang, L. (2006). Early warning signs of IT project failure: The dominant dozen. *Information systems management*, *23*(4).

Kallina, E., & Singh, J. (2024, October). Stakeholder involvement for responsible AI development: A process framework. In Proceedings of the 4th ACM conference on equity and access in algorithms, mechanisms, and optimization (EAAMO '24) (pp. 1-14). ACM.

Kingdon, J. W. (2014). *Agendas, alternatives, and public policies* (2nd ed., Pearson New International Edition). Pearson Education Limited.

Klein, G. (2007). Performing a project premortem. *Harvard business review*, *85*(9), 18-19.

Kordzadeh, N., & Ghasemaghaei, M. (2022). Algorithmic bias: review, synthesis, and future research directions. European Journal of Information Systems, 31(3), 388-409.

Korteland, E. and V.J.J.M. Bekkers. (2008). 'The diffusion of electronic service delivery innovations in Dutch E-policing: The case of digital warning systems', Public Management Review, 10, 1, 71–88.

Koutsounia, A. (2024). AI could be time-saving for social workers but needs regulation, say sector bodies. Community Care. Retrieved from: https://www.communitycare.co.uk/2024/10/04/ai-could-be-time-saving-for-social-workers-but-needs-regulation-say-sector-bodies/

Lacity, M. C., & Willcocks, L. P. (2000). An empirical investigation of information technology sourcing practices: Lessons from experience. MIS Quarterly, 24(3), 363-408.

Lapuente, V., & van de Walle, S. (2020). The effects of new public management on the quality of public services. Governance, 33(3), 461-475.

Lin, S. Y., Shanafelt, T. D., & Asch, S. M. (2018, May). Reimagining clinical documentation with artificial intelligence. In Mayo Clinic Proceedings (Vol. 93, No. 5, pp. 563-565). Elsevier.

Locke, E. A., & Latham, G. P. (1990). Work motivation and satisfaction: Light at the end of the tunnel. *Psychological science*, *1*(4), 240-246.

Lundvall, B. Å. (Ed.). (1992). National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning. Pinter Publishers.

Majone, G. (1997). From the positive to the regulatory state: Causes and consequences of changes in the mode of governance. *Journal of public policy*, *17*(2), 139-167.

March, J. G., & Olsen, J. P. (1984). The new institutionalism: organizational factors in political life. *The American Political Science Review*, 78(3), 734-749.

Marcus, G. (2020). The next decade in AI: four steps towards robust artificial intelligence. arXiv preprint arXiv:2002.06177.

Mayhew, D. R. (1974). Congress: the electoral connection. Yale University Press.

Mazzucato, M. (2011). The entrepreneurial state. *Soundings*, *49*(49), 131-142.

McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., ... & Shetty, S. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, 577(7788), 89-94.

McManus, J., & Wood-Harper, T. (2007). Understanding the sources of information systems project failure (a study in IS project failure in Europe). *Management Services* 51(3), 38-43.

Minh, D., Wang, H. X., Li, Y. F., & Nguyen, T. N. (2022). Explainable artificial intelligence: a comprehensive review. Artificial Intelligence Review, 55, 1-66.

Moore, M. H. (1997). Creating public value: Strategic management in government. Harvard university press.

Moorman, R. H. (1991). Relationship between organizational justice and organizational citizenship behaviors: Do fairness perceptions influence employee citizenship?. *Journal of applied psychology*, *76*(6), 845.

Morrison, E. W., & Robinson, S. L. (1997). When employees feel betrayed: A model of how psychological contract violation develops. *Academy of management Review*, *22*(1), 226-256.

National Audit Office. (2011). The National Programme for IT in the NHS: An update on the delivery of detailed care records systems (HC 888, Session 2010-2012). *The Stationery Office.*

National Audit Office. (2017). Digital transformation in government (HC 1059, Session 2016-2017). *The Stationery Office.* Retrieved from: https://www.nao.org.uk/reports/digital-transformation-in-government/

National Grid Electricity System Operator. (2019). *ESO and The Alan Turing Institute use machine learning to help balance the GB electricity grid*. Retrieved from https://www.neso.energy/news/eso-and-alan-turing-institute-use-machine-learning-help-balance-gb-electricity-grid

Nonaka, I., & Takeuchi, H. (1995). The knowledge-creating company: How Japanese companies create the dynamics of innovation. Oxford University Press.

OECD. (2019). The innovation system of the public service of Brazil: an exploration of its past, present and future journey. OECD Public Governance Reviews, OECD Publishing. https://doi.org/10.1787/a1b203de-en

OECD/BCG/INSEAD. (2025). The Adoption of Artificial Intelligence in Firms: New Evidence for Policymaking. OECD Publishing, Paris. https://doi.org/10.1787/f9ef33c3-en.

O'Neil, C. (2016). Weapons of math destruction: how big data increases inequality and threatens democracy. Crown Publishers.

Oreg, S., Vakola, M., & Armenakis, A. (2011). Change recipients' reactions to organizational change: A 60-year review of quantitative studies. *The Journal of applied behavioral science*, *47*(4), 461-524.

Osborne, S. P., & Brown, L. (2013). Handbook of innovation in public services. Edward Elgar Publishing.

Oswald, M., Paterson-Young, C., McBride, P., Maher, M., Calder, M., Gill, G., Tiarks, E., & Noble, W. (2024). Ethical review to support Responsible Artificial Intelligence (AI) in policing: A preliminary study of West Midlands Police's specialist data ethics review committee. *Northumbria University.*

Pannu, A. (2015). Artificial intelligence and its application in different areas. *Artificial Intelligence, 4*(10), 79-84.

Parker, I., Studman, A., & Jones, E. (2025). *Lessons from six years of studying AI in the public sector*. Ada Lovelace Institute.

Pärna, O., & Von Tunzelmann, N. (2007). Innovation in the Public Sector: Key Features Influencing the Development and Implementation of Technologically Innovative Public Sector Services in the UK, Denmark, Finland and Estonia. *Information Polity, 12*(3), 109–25.

PASC [Public Administration Select Committee]. (2011). Government and IT – "a recipe for rip-offs": time for a new approach. Volume I, 28 July 2011. Stationery Office. Retrieved from: https://publications.parliament.uk/pa/cm201012/cmselect/cmpubadm/715/715i.pdf

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.

Pellegrino, F. (2019). The just culture principles in aviation law. *Legal Studies in International, European and Comparative Criminal Law*, *3*.

Personal Data Protection Commission (PDPC). (2020). *Model artificial intelligence governance framework* (2nd ed.). Retrieved from https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf

Pettigrew, A. M., Ferlie, E., & McKee, L. (1992). *Shaping Strategic Change: Making Change in Large Organizations: The Case of the National Health Service*. Sage Publications.

Pil, F. K., & MacDuffie, J. P. (1996). The adoption of high-involvement work practices. *Industrial Relations: A journal of economy and society, 35*(3), 423-455.

Police Foundation. (2022). *A New Mode of Protection: Redesigning Policing and Public Safety for the 21st Century*. The Police Foundation.

Pollitt, C., & Bouckaert, G. (2017). *Public management reform: A comparative analysis-into the age of austerity* (4th ed.). Oxford University Press.

Provan, K. G., & Lemaire, R. H. (2012). Core concepts and key ideas for understanding public sector organizational networks: Using research to inform scholarship and practice. *Public Administration Review, 72*(5), 638-648.

Public First. (2024). AI & the Public Sector. *Public First*. Retrieved from: https://www.publicfirst.co.uk/wp-content/uploads/2024/11/AI-and-the-Public-Sector_final.pdf

Reason, J. (1998). Achieving a safe culture: theory and practice. *Work & stress*, *12*(3), 293-306.

Reichers, A. E., Wanous, J. P., & Austin, J. T. (1997). Understanding and managing cynicism about organizational change. *Academy of management perspectives*, *11*(1), 48-59.

Ruttan, V. W. (2001). *Technology, Growth and Development: An Induced Innovation Perspective*. Oxford University Press.

Sarzaeim, P., Mahmoud, Q. H., Azim, A., Bauer, G., & Bowles, I. (2023). A Systematic Review of Using Machine Learning and Natural Language Processing in Smart Policing. *Computers, 12*(12), 255. https://doi.org/10.3390/computers12120255

Sauer, C., Gemino, A., & Reich, B. H. (2007). The impact of size and volatility on IT project performance. *Communications of the ACM*, *50*(11), 79-84.

Schwarcz, D. & Choi, J. H. (2023). AI Tools for Lawyers: A Practical Guide. *108 Minnesota Law Review Headnotes*. Retrieved from: https://ssrn.com/abstract=4404017

Senge, P. (1990). Peter Senge and the learning organization. *Dimension, 14*.

Spence, M. (1973). Competitive and optimal responses to signals: An analysis of efficiency and distribution. *Journal of Economic theory, 7*(3), 296-332.

Spring, J. M., Galyardt, A., Householder, A. D., & VanHoudnos, N. (2020). On managing vulnerabilities in AI/ML systems. In *Proceedings of the New Security Paradigms Workshop 2020* (pp. 111-126). ACM. https://doi.org/10.1145/3442167.3442177

Staw, B. M., Sandelands, L. E., & Dutton, J. E. (1981). Threat rigidity effects in organizational behavior: A multilevel analysis. *Administrative science quarterly*, 501-524.

Tarakci, M., Heyden, M. L., Rouleau, L., Raes, A., & Floyd, S. W. (2023). Heroes or villains? Recasting middle management roles, processes, and behaviours. *Journal of Management Studies, 60*(7), 1663-1683.

Timmins, N. (2024, July). *Enabling Integrated Care Systems to work better*. Institute for Government. https://www.instituteforgovernment.org.uk/publication/enabling-integrated-care-systems

Toosi, A., Bottino, A. G., Saboury, B., Siegel, E., & Rahmim, A. (2021). A brief history of AI: how to prevent another winter (a critical review). *PET clinics*, *16*(4), 449-469.

van Ark, B. (2022). *Making Public Sector Productivity Practical*. The Productivity Institute and Capita.

van Ark, B., Hoskins, J., & Jörden, N. (2023). Public Sector Productivity – managing the Baumol cost disease. In D. Coyle, B. van Ark, & J. Pendrill (Eds.), *The Productivity Agenda* (Chapter Eight, pp. 86-97). The Productivity Institute.

van Ark, B., & Hoskins, J. (2024). *Investigating Police Productivity: A Literature Review*. The Productivity Institute.

Whang, S. E., Roh, Y., Song, H., & Lee, J.-G. (2022). Data Collection and Quality Challenges in Deep Learning: A Data-Centric AI Perspective. arXiv preprint arXiv:2112.06409.

Williamson, O. E. (1985). The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting. Free Press.

Winch, G., & Leiringer, R. (2016). Owner project capabilities for infrastructure development: A review and development of the "strong owner" concept. *International journal of project management*, *34*(2), 271-281.

Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial Intelligence and the Public Sector—Applications and Challenges. International Journal of Public Administration, 42(7), 596-615.

Xiong, G. L., Bayen, E., Nickels, S., Subramaniam, R., Agrawal, P., Jacquemot, J., Bayen, A., Miller, B., & Netscher, G. (2019). Real-time video detection of falls in dementia care facility and reduced emergency care. American Journal of Managed Care, 25(7), 314-315.

Young, G. J., Charns, M. P., & Shortell, S. M. (2001). Top manager and network effects on the adoption of innovative management practices: A study of TQM in a public hospital system. Strategic Management Journal, 22(10), 935-951.

Zheng, Y., Han, Y., Cui, L., Miao, C., Leung, C., & Yang, Q. (2018). SmartHS: An AI platform for improving government service provision. In The Thirtieth AAAI Conference on Innovative Applications of Artificial Intelligence (IAAI-18) (pp. 7704–7711). Retrieved from https://ojs.aaai.org/index.php/AAAI/article/view/11382

Zick, T., Kortz, M., Eaves, D., & Doshi-Velez, F. (2024). AI Procurement Checklists: Revisiting Implementation in the Age of AI Governance. *arXiv preprint arXiv:2404.14660*

Appendix – Capability Maturity Tables

Table 1 – Organisation-Level Capabilities 1

	Organisational Learning	Planning and Prioritisation
(1) Outstanding (Proactive, Culturally Embedded)	• Learning is a core cultural value, blending qualitative insights with rigorous quantitative analysis. • Continuous improvement cycles are standard practice. Thriving communities of practice diffuse expertise throughout the organisation. • The organisation proactively seeks external knowledge to challenge internal assumptions.	• A formal methodology is used to identify system-wide bottlenecks. The strategic plan maps investment to resolving these constraints and always includes complementarities. • The organisation maintains a dedicated, ring-fenced fund for innovation and experimentation, managed with a tolerance for intelligent failure. • Narrative-crafting and planning are sophisticated enough to resolve the tension between long-term productivity needs and the demand for regular short-term political wins.
(2) Good (Formalised, Consistent)	• Formal post-project reviews are standard and consistently combine qualitative discussion with descriptive data analytics to understand outcomes. • Findings and action items are tracked in a central repository. Emerging communities of practice are encouraged and supported. • Frontline staff have clear and reliable channels for submitting ideas for improvement.	• A formal portfolio management process is used to prioritise projects based on strategic alignment and expected ROI. • Business cases must identify the bottleneck a project addresses. Budgets for strategic projects are protected from in-year reprioritisation, with stable multi-year funding commitments. • Project narratives are consistently crafted from the outset to demonstrate how a project supports key ministerial priorities.
(3) Adequate (Ad-hoc, Inconsistent)	• Post-project reviews are required for major projects, but follow-through is inconsistent. • While some teams use data analytics on an ad-hoc basis, learning remains siloed. Data quality issues are frequently cited as a blocker but are not addressed. • A suggestion box exists but lacks a formal review process.	• Business cases are required for major investments but often lack data-driven analysis of impact. • Long-term infrastructure needs (e.g., data quality) are acknowledged but consistently de-prioritised for short-term deliverables. • Alignment with political priorities is used opportunistically to secure funding, but there is no consistent strategy for narrative-building.

(4) Needs Improvement (Informal, Siloed)	• Learning is isolated within teams and driven by individual initiative. • Data are locked in operational silos, and any 'analysis' is done in individual spreadsheets that are not shared or validated. • Any ‘lessons learned’ exercises are informal, verbal, and rarely documented, meaning insights are lost as soon as key individuals leave. • Feedback from frontline staff is ad-hoc and depends on personal relationships with managers.	• Planning is siloed within departments. There is little or no mechanism for funding cross-cutting initiatives, even when they are identified as strategically important. • The concept of a bottleneck is not part of the formal planning vocabulary. Instead, the organisation has many small uncoordinated initiatives that run in parallel, and which rarely survive or scale-up once the managerial sponsor moves job. • Project justifications are purely technical or operational, failing to connect with the political agenda.
(5) Inadequate (Absent, Counter-Productive)	• Problems are addressed with immediate, short-term fixes with no analysis of root causes. • Data are ignored, mistrusted, or used selectively to assign blame. • The phrase “this is how we’ve always done it” is used to shut down discussion. Questioning established processes is treated as a sign of disloyalty or being ‘unhelpful’.	• Budgeting is purely incremental or driven by short-term political pressures. • The end-of-year budget cycle is characterised by a frantic ‘use it or lose it’ spending spree on un-planned items, ensuring that waste is locked into the following year. • Projects are funded without a written business case or a clear link to strategic goals. • Decisions are often made in response to media headlines or the personal preference of senior leaders, bypassing any formal evaluation process. The loudest voice, not the best case, wins.

Table 2 – Organisation-level Capabilities 2

	Managing Risk	Developing New Ways of Working	Managing Supplier Relations
(1) Outstanding (Proactive, Culturally Embedded)	• Risk management is proactive and portfolio-based, balancing the risk of inaction against the risk of innovation. • Formal governance, including ethics review and proactive stakeholder engagement helps to define the 'risk appetite' that empowers teams to take calculated risks. • Learning from failure is a formal input into strategic planning.	• Workforce planning is continuous, anticipating future skill needs. • Managers are adept at navigating professional norms and internal politics to drive change. • The organisation systematically pilots and evaluates new organisational models (e.g., agile teams) to complement new technology. • There is a culture of co-design, where both staff and external stakeholders are treated as integral partners in shaping new workflows and services.	• Acts as a sophisticated systems integrator, deliberately disaggregating large initiatives into manageable projects to foster a competitive supplier ecosystem and mitigate risk. • The organisation shapes the market by articulating future needs, setting technical standards for multi-vendor solutions, and actively avoiding proprietary lock-in.
(2) Good (Formalised, Consistent)	• There is a formal risk management framework with clear roles and responsibilities. • There is clear 'line of sight' from project-level risks up to senior management, with established escalation paths for risks that exceed a project's tolerance. • Failures are documented in a 'blameless' post-mortem process to extract lessons for future projects.	• HR teams identify skills gaps and redesign job roles before a technology is deployed. • Change management, including communication and staff engagement, is a well-resourced part of all major project plans. • There is a formal and proactive process for consulting with external stakeholders early in the change process. Their needs and potential impacts are systematically mapped and addressed.	• A dedicated 'intelligent customer' function exists with the technical expertise to co-design solutions and validate supplier claims. • Contracts include clear performance-based incentives and KPIs. The organisation can successfully manage projects with multiple, integrated vendors. • Supplier relationships are managed as partnerships, with regular, structured dialogue to resolve issues collaboratively.

(3) Adequate (Ad-hoc, Inconsistent)	• Risk is managed at the project level, but there is no aggregate view of organisational risk. • 'Pre-mortem' exercises are used sporadically on high-profile projects but are not standard practice. • Responsibility for risk is assigned to individuals, leading to cautious decision-making.	• Major technology projects include a training budget, but it is first to be cut. • The impact on roles is considered, but changes are often reactive and made after implementation. • Some teams experiment with process change, but learning is not shared. • External stakeholder consultation happens, but often too late to meaningfully influence design.	• The organisation has enough expertise to participate in technical discussions but struggles to challenge supplier recommendations. • Contract management focuses on tracking deliverables, not strategic outcomes. • Most projects are single-sourced to avoid multi-vendor integration complexity.
(4) Needs Improvement (Informal, Siloed)	• Risk management is a 'tick-box' compliance exercise. • Risk registers are created at the start of a project but not used as active management tools. • A culture of blame-avoidance discourages staff from raising potential issues.	• Training is considered only after implementation and focuses solely on basic functions ("button-pushing"). • Job roles aren't reviewed or redesigned - leading to skill mismatches and inefficient workarounds. • External stakeholders are only engaged reactively when a problem arises, or a complaint is made.	• The organisation can articulate basic needs but relies entirely on the supplier to define the solution and technical requirements. • Issues are managed reactively through formal contract disputes rather than collaborative problem-solving.
(5) Inadequate (Absent, Counter-Productive)	• A zero-tolerance policy for failure exists, leading to concealment of problems until they become crises. • The concept of risk management is either absent or used purely as a tool for assigning blame. • Near misses are treated as a success ('we got away with it!') rather than as free lessons. There is no incentive to analyse why a potential failure was narrowly avoided.	• Job roles are static and narrowly defined. No budget exists for staff training on new systems. • The implementation of technology is viewed as a purely technical task, completely disconnected from the people that use it. • Staff who attempt to improve workflows are informally sanctioned for 'not following procedure'. • External stakeholders are ignored or viewed as an obstacle.	• Procurement is based on supplier's marketing claims, with no independent verification. • Contracts are awarded without clear deliverables or performance metrics, leading to inevitable vendor lock-in and chronic cost overruns. • Contracts are seen more as a tool for transferring blame than for defining a partnership.

Table 3 – Sector-Level Capabilities 1

	Strengthening Innovation Networks	Providing Clarity on Collaborative Performance
(1) Outstanding (Co-ordinated Ecosystem)	• A dedicated, funded entity (a "network weaver") is responsible for strategic convening and brokering relationships across the sector to solve shared problems. • Participation and knowledge codification are formally incentivised. • There is a central budget for network infrastructure (online platforms, events, staff), ensuring its sustainability beyond individual projects.	• The sector functions as a trusted and predictable partner. • A central body maintains a public "client capability register," rating public bodies on their ability to manage complex projects. This body, combined with standardised contracts and a single, audited supplier register, reduces risk for all parties and fosters a high-trust, high-performance market.
(2) Good (Active but Fragmented)	• Formal, sustainable networks exist within specific sub-sectors (e.g., a national policing network, an NHS-wide forum), each with its own budget and charter. However, these networks are not co-ordinated with each other, leading to sector-wide strategic gaps. • Knowledge codification is an expected output of collaborative projects, but incentives to share are inconsistent.	• Several bodies offer "client-side accreditation" or training programs to improve public sector buyer capability. While not mandatory, this signals quality. • Organisations can access multiple, high-quality supplier registers, but must navigate different systems. • The risk of encountering an unreliable partner (client or supplier) is reduced but not eliminated.
(3) Adequate (Pockets of Activity)	• Convening is ad-hoc, driven by individual champions for specific, short-term projects. • Knowledge sharing relies entirely on personal goodwill; there are no formal processes or rewards for documenting work for others. • Any collaboration is funded out of existing, siloed project budgets, making it fragile.	• Pioneering public bodies publish their internal client-side standards and procurement track records. • A common complaint from innovative suppliers is the 'postcode lottery' of client capability; they actively seek out the few known good partners and avoid bidding for work with others, further concentrating expertise.

(4) Needs Improvement (Passive & Aspirational)	• The ecosystem is passive, with collaboration stated as a goal but no actors resourced or tasked with network-building. • Communication is one-way (e.g., newsletters, websites) and the "network" is limited to events convened by external vendors, not by the sector for the sector. • Central government bodies may issue guidance or policy documents but take no active role in convening practitioners to discuss implementation, leaving individual organisations to interpret the documents in isolation.	• Suppliers face a highly opaque market where they cannot differentiate between competent and incompetent public sector partners, increasing risk and deterring high-quality firms from bidding. • Suppliers rely on informal networks to assess the competence of a potential public sector client, making the market highly relationship-dependent, and closed to new entrants. • There are no specific frameworks, templates, or audited supplier registers. Every procurement team is left to figure it out on their own.
(5) Inadequate (Systemic Vacuum)	• The prevailing culture and incentives across the sector actively work against networking. • There is a complete absence of inter-organisational convening and knowledge-sharing mechanisms. • Each organisation operates as an isolated silo, leading to widespread duplication of failures and an inability to scale success. • There is no perceived collective responsibility for building connections or sharing learnings.	• Partnerships consistently carry high hidden costs in time, risk, effort, and legal fees for both suppliers and clients. • The sector has a reputation for being an unreliable and costly client, characterised by scope creep, delayed payments, and poor project management. This creates a classic "lemons market" where only suppliers who are specialised in navigating bureaucracy (rather than delivering innovation) can succeed. • The sector's poor reputation is a topic of conversation in the tech industry, deterring talented individuals from seeking roles in public-sector focussed firms.

Table 4 – Sector-level Capabilities 2

	Funding and Co-ordinating Research	Governing Failure
(1) Outstanding (Co-ordinated Ecosystem)	• A co-ordinated national research agenda strategically funds foundational, high-risk 'public good' research through complementary funding streams. • Formal public-private partnerships align private R&D with public need, while market-shaping mechanisms like Advanced Market Commitments signal credible demand and de-risk private investment.	• An integrated governance framework, sustained through active leadership, explicitly and transparently balances the imperatives of systemic learning and legitimate accountability. • There are publicly understood criteria for when an incident triggers a non-punitive learning review versus a formal accountability investigation. • Leadership is held accountable for acting on systemic lessons, which builds the trust required for open reporting. • Learning and accountability are seen as mutually reinforcing.
(2) Good (Active but Fragmented)	• Several well-funded research programmes exist, but their strategic priorities are not explicitly aligned. • Strong research is delivered but unaddressed gaps remain in key areas. • Different public bodies may compete for the same limited pool of research talent.	• A formal "just culture" framework is adopted, with a clear triage process to distinguish between system-induced errors, honest mistakes, and blameworthy actions. • Separate, well-defined channels exist for non-punitive incident reporting (for learning) and formal investigations (for accountability). The goal is to separate system failures from individual culpability, fostering greater trust.
(3) Adequate (Pockets of Activity)	• Research is funded on a reactive, ad-hoc basis. Individual organisations may commission small projects to solve immediate problems. • Effort is frequently duplicated as organisations commission similar research projects to solve the same problems. • There is no mechanism for identifying or funding longer-term, sector-wide research challenges.	• Separate and often conflicting processes exist for accountability (e.g., formal audit, compliance checks) and learning (e.g., ad-hoc post-mortems). • The tension between the two is unmanaged and often counter-productive, as staff may withhold information from auditors for fear of blame, undermining both processes. The systems are not designed to complement each other.

(4) Needs Improvement (Passive & Aspirational)	• A government white paper may identify AI as a "priority," but no new funding or co-ordination mechanisms are established, often creating a patchwork of incompatible solutions. • The sector's specific research needs are not articulated or represented in national research funding decisions. • Research funding remains siloed within traditional departmental budgets, with no mechanism for pooling resources to tackle cross-cutting, system level challenges that are identified in policy papers.	• Governance is dominated by a crude and inconsistent blame culture. • Accountability is a highly predictable but crude process: after a failure, a formal investigation is launched with the unstated but clear goal of identifying an individual or group to scapegoat. • Any "lessons learned" exercises are performative and disconnected from meaningful change, resulting in the same systemic failures recurring in different projects over time, as the root causes are never addressed.
(5) Inadequate (Systemic Vacuum)	• There is no concept of a public sector research agenda. • There is a profound lack of awareness of the current state of technological art. Procurement decisions are often based on outdated knowledge and supplier marketing from years prior. • The sector lacks any collective body or voice to influence national research priorities, meaning its unique needs are consistently overlooked by major public and private funding councils.	• There is no predictable process for dealing with failure. Blame is assigned based on expediency or personal animosity, making the environment feel dangerously arbitrary. • Failures are either concealed to avoid repercussions or result in arbitrary, unpredictable blame. • A hostile environment prevents any possibility of systemic learning and erodes the foundations of legitimate accountability, fostering deep risk aversion, organisational paralysis, and public cynicism.

Table 5 – Sector-Level Capabilities 3

	Maintaining Pressure for Productivity Improvement	Ensuring a Flow of Human Capital	Standardising and Scaling Best Practice
(1) Outstanding (Co-ordinated Ecosystem)	• A single, authoritative source publishes trusted productivity benchmarks. • These benchmarks guide both day-to-day operational improvement and are directly linked to predictable, multi-year funding settlements, which include formal "gain-sharing" mechanisms, allowing organisations to retain and reinvest efficiency savings.	• The sector has a co-ordinated Human Capital Strategy for digital skills, co-owned by government, professional bodies, and education providers. • Jointly funded, high-quality training institutions and apprenticeship programmes create a robust talent pipeline. • Sector-wide skills frameworks and certifications ensure talent is mobile and credentials are trusted.	• A central "what works" centre proactively identifies, rigorously evaluates, and disseminates best practice. It maintains a library of interoperability standards and reusable components (e.g., code libraries, APIs) that are mandated for new projects to ensure economies of scale and prevent vendor lock-in. • Standards are co-designed with local entities.
(2) Good (Active but Fragmented)	• Multiple bodies collect performance data, which is used to inform annual budget negotiations. This creates some pressure, but the lack of a single benchmark allows for debate over data validity. • Financial allocations are generally stable year-to-year, but "spend-to-save" or gain-sharing schemes are ad-hoc and inconsistently applied.	• Multiple good training programs and certifications exist, run by different actors (e.g., professional bodies, tech vendors, universities). • There is a flow of talent, but the lack of a unified skills framework leads to duplication, gaps in provision, and difficulty for employers in comparing different qualifications.	• Multiple organisations actively promote best practice within their specific domains (e.g., a police consortium shares body-camera data standards, an NHS body promotes an EPR data schema). • This leads to islands of high performance and interoperability, but system-wide integration remains a challenge due to competing standards.

(3) Adequate (Pockets of Activity)	• Performance data are collected but are treated as a compliance exercise, disconnected from funding decisions. • Financial allocations are determined primarily by historical precedent and inflationary pressures. • Pressure for improvement is applied through reactive, across-the-board budget cuts in response to external fiscal events, not performance.	• Individual organisations invest in their own internal training programs or fund staff to attend external courses. • Talent is developed but often remains "locked-in" to that organisation. • There is no sector-wide view of skills gaps or a co-ordinated effort to address them.	• Best practice is shared informally through personal networks or at conferences. However, attempts at replication are often disappointing due to a lack of information on context. • There are no formal mechanisms for validating these practices or supporting their adoption with standardised tools or templates.
(4) Needs Improvement (Passive & Aspirational)	• The National Audit Office or a regulator may occasionally publish a critical report on inefficiency, but this is an isolated event with no consistent systemic pressure for improvement. • A large volume of centrally imposed targets creates strategic incoherence and diffuses focus. Budgeting is not linked to performance metrics.	• Skills are only discussed in reaction to a crisis (e.g., a major project failure due to lack of expertise). • Responsibility for skills development is devolved entirely to individuals. • Management views specialist skills as a resource to 'buy-in' rather than a capability to 'build up', leading to many costly short-term contracts.	• "Best practice" is what is promoted by large IT vendors in their marketing materials. • There is no independent or public-sector-led effort to identify what actually works. • The default is to procure proprietary, end-to-end solutions, leading to fragmentation and vendor lock-in.
(5) Inadequate (Systemic Vacuum)	• There is no systematic performance measurement or benchmarking. • Financial allocations are arbitrary and unpredictable, operating in a "feast or famine" cycle. This actively punishes good financial management (as surpluses are clawed back) and makes investment in productivity-enhancing technology irrational for managers.	• There is no recognition of a skills gap. • The sector assumes that existing staff can operate complex new technologies with minimal training. This leads to chronic project failure, low user adoption, and significant operational risk. • Suppliers consistently report the lack of client-side technical skills is the single biggest barrier to successful collaboration	• There is no concept of best practice or interoperability. • Every organisation procures or builds its own bespoke systems in complete isolation. This results in a highly fragmented, inefficient, and expensive technology landscape that is impossible to connect or modernise.