Thrive Or Die - UK Government And Public Sector

UK Government And Public Sector Edition — Why The Customer Being The Taxpayer Changes The Governance Question, Not The Mechanism

A Charterhouse Consultants white paper by Bridget Convey, January 2026

Executive Summary

UK government is the one organisation in this series whose shareholder is the taxpayer, which changes what closing the gap means, but not whether the gap exists. The government's own AI suite, Humphrey, is projected to save the equivalent of 75,000 days of manual analysis every year. A separate planning tool, Extract, can do in 40 seconds what previously took planners one to two hours. Both are real, deployed, and measured. The more interesting evidence in this paper is what the government chose to stop doing: in 2025 it discontinued Redbox, one of its own AI tools, because enterprise alternatives from the private sector had simply overtaken it. That is closer to a company voluntarily admitting the gap and closing it than anything else found in this series so far.

The Premise: The Gap Is The Danger, Not The AI

In October 2023, a group including Turing Award winners Yoshua Bengio and Geoffrey Hinton, alongside Stuart Russell, Daniel Kahneman and more than twenty other senior AI researchers, published a consensus paper in the journal Science on the risks of advanced AI. Their argument was precise. AI capability and autonomy are increasing rapidly. Organisations are shifting toward systems that can act and pursue goals with reduced human oversight. Governance is not keeping pace with that shift. The danger they describe does not come from AI itself. It comes from the widening gap between what a system can now do on its own and how closely anyone is still watching it.

That paper is about safety and governance at a societal scale. It does not comment on organisational strategy, and I want to be precise about that rather than overstate its relevance. What it gives this paper is not an endorsement, it gives us the mechanism. Capability grows. Autonomy grows. Oversight does not grow at the same rate. In UK government specifically, the gap is visible less in a single dramatic failure and more in the discipline, or lack of it, behind the savings figures themselves. Estimates for what Humphrey and similar tools could save the taxpayer have ranged from around £20 million to as much as £45 billion depending on the source and the scope assumed, a spread wide enough to suggest the measurement governance around AI savings claims is not yet as mature as the tools themselves.

This is the basis for the claim in the title. Thrive or die is not a slogan layered on top of the argument. It is the plain consequence of the mechanism. An organisation that builds governance at the same pace it builds AI capability closes the gap and survives it. One that lets capability outrun governance does not get a soft landing.

Thrive Or Die: The Pattern In Organisations That Did Not Close The Gap

Kodak invented the first digital camera in 1975 and its own internal research predicted by 1981 that digital would eventually replace film. What it did not close was the governance gap between that capability and a business model still built entirely around film margins. The company filed for bankruptcy in 2012, thirty seven years later. Blockbuster was offered the chance to buy Netflix outright for fifty million dollars in 2000, declined, and filed for bankruptcy in 2010, having eliminated close to eighty four thousand jobs in the collapse. Both prove the same point. The danger was never a lack of capability. It was a governance gap left open until it was unmanageable.

Kodak and Blockbuster protected an old model until it was too late to change course. Government's Incubator for AI shows, in one specific instance, the opposite instinct. Redbox, a generative AI tool built in-house and rolled out to over 6,000 civil servants across 150,000 chats and 1.3 million messages, was sunset in 2025 once departments adopted newer enterprise tools like Copilot and Gemini that outperformed it. Building something, using it at real scale, and then retiring it in favour of something better, rather than protecting the sunk investment, is precisely the behaviour Kodak and Blockbuster failed to show. It is a small example, but it is evidence the mechanism can run the right way inside a large, famously slow-moving institution when the incentive to protect a legacy tool is weaker than the incentive to serve the public well.

The Evidence In UK Government: Real Savings, And An Honest Retirement

Two pieces of evidence anchor this paper, one showing capability delivering measurable results, the other showing government doing something rare in this series, choosing to close its own gap rather than defend a legacy tool.

Humphrey: Consult And Extract, Measured Against Human Performance

Humphrey, a suite of AI tools named after the fictional civil servant from Yes Minister, includes Consult, which reviews public consultation responses, Parlex, which searches parliamentary debate, Minute, a secure meeting transcription tool, and Lex, which researches legislation back to 1267. Consult's initial live deployment, run by the Scottish Government on a consultation about non-surgical cosmetic treatments, processed over 2,000 responses and delivered results described by officials as nearly identical to a team of human analysts, while reducing the bias human analysts can introduce. Wider rollout is projected to save the equivalent of 75,000 days of manual analysis annually. Separately, Extract, developed with the Ministry of Housing, Communities and Local Government, digitises planning data, converting blurry maps and handwritten notes into machine-readable formats in around 40 seconds, work that previously took planners one to two hours, directly supporting the government's target of building 1.5 million new homes over five years.

Redbox: The Rare Example Of A Government Tool Being Retired On Purpose

Redbox, an in-house generative AI chat and document tool, was rolled out to more than 6,000 civil servants, generating 150,000 chats and 1.3 million messages, a genuinely large-scale internal deployment. In 2025, it was discontinued, with the stated reason that equivalent private-sector tools, principally Microsoft Copilot and Google Gemini, now outperformed it. This is worth treating as a data point in its own right. Very few organisations in this paper series have voluntarily retired a capability they built themselves in favour of a better alternative, rather than defending the investment. It suggests the part of government responsible for this decision was closing the capability to governance gap correctly, at least in this instance, by prioritising the outcome over the sunk cost.

UK Government And Public Sector In Three Horizons: The Gap Closes Or It Does Not

If the gap between capability and governance is the real variable, then each horizon below has two outcomes, not one. These are deliberately stated as firm positions, not hedged possibilities, because a board or leadership conversation needs something to agree or disagree with.

Two Years: 2026 To 2028

What happens everywhere: AI tools built or procured by central government, following the Humphrey model, spread into local government and individual departments at increasing speed.

- Thrive path. Government standardises how AI savings and performance are measured and reported across departments, closing the current wide gap between different savings estimates, so public and parliamentary trust in the figures grows alongside the capability itself.
- Die path. Departments continue reporting AI impact with inconsistent methodologies and widely varying savings claims, and the resulting credibility gap becomes a story about government AI in general, overshadowing tools like Consult and Extract that are genuinely performing well.

Five Years: 2028 To 2031

What happens everywhere: the build-versus-buy question the Redbox retirement raised becomes a standing decision every department has to make repeatedly, not a one-off.

- Thrive path. Government builds a formal, recurring process for evaluating whether in-house AI tools still lead the market or should be retired in favour of enterprise alternatives, turning the Redbox decision from a one-off into institutional discipline.
- Die path. Departments that built early AI tools defend them past the point they remain competitive, for the same reasons Kodak defended film, sunk investment and internal ownership, producing exactly the slow-moving incumbent failure this series is built around, inside the one sector that had just shown it knew how to avoid it.

Ten Years: 2031 To 2036

What happens everywhere: the relationship between civil service headcount, AI capability and service delivery quality becomes a permanent feature of public spending debates.

- Thrive path. The UK's public sector AI programme is recognised internationally as a reference model, in the way its financial and legal regulatory frameworks have been historically, because it paired real capability with real, standardised measurement of impact.
- Die path. Public trust in government AI claims erodes because of years of inconsistent savings reporting, and genuinely effective tools like Consult and Extract are politically undermined by scepticism earned elsewhere in the system.

What I Am Recommending To Public Sector Leaders And Departmental Boards

- Adopt one standard methodology for measuring and reporting AI-driven savings across departments, closing the current gap between figures as far apart as £20 million and £45 billion for what are related claims about the same programme.
- Build a recurring, formal review process for whether in-house AI tools remain competitive with enterprise alternatives, turning the Redbox decision into standard practice rather than a one-off.
- Treat tools like Consult, which was benchmarked directly against human analyst performance before wider rollout, as the standard for how new AI deployment should be evaluated, rather than deploying first and measuring impact after the fact.

Conclusion

Government is the one organisation in this series whose end customer cannot walk away the way a bank or insurance customer can, which makes closing the capability to governance gap a matter of public trust as much as efficiency. Humphrey's Consult and Extract tools show the capability side of that equation working well, measured honestly against human performance rather than assumed. Redbox's retirement shows something rarer still in this series, an organisation closing its own gap voluntarily rather than defending a legacy investment the way Kodak defended film. The open question is whether that discipline, evident in one specific decision, becomes standard practice across a system this large, or remains the exception that proves how hard the rule still is to follow.

Sources: GOV.UK Incubator for AI, Global Government Forum, Complete AI Training, Think Digital Partners, California Department of Technology newsroom. All figures current as of mid-2026, verified at time of writing. This is a first draft prepared for internal discussion.