Responsible AI in Action

Bite-sized guidance series

Grow as you learn

This issue covers:

- the right mindset for approaching AI roll-out
- 6 steps for 'iterative unleashing'

Read this alongside other issues in the bite-sized AI guidance series

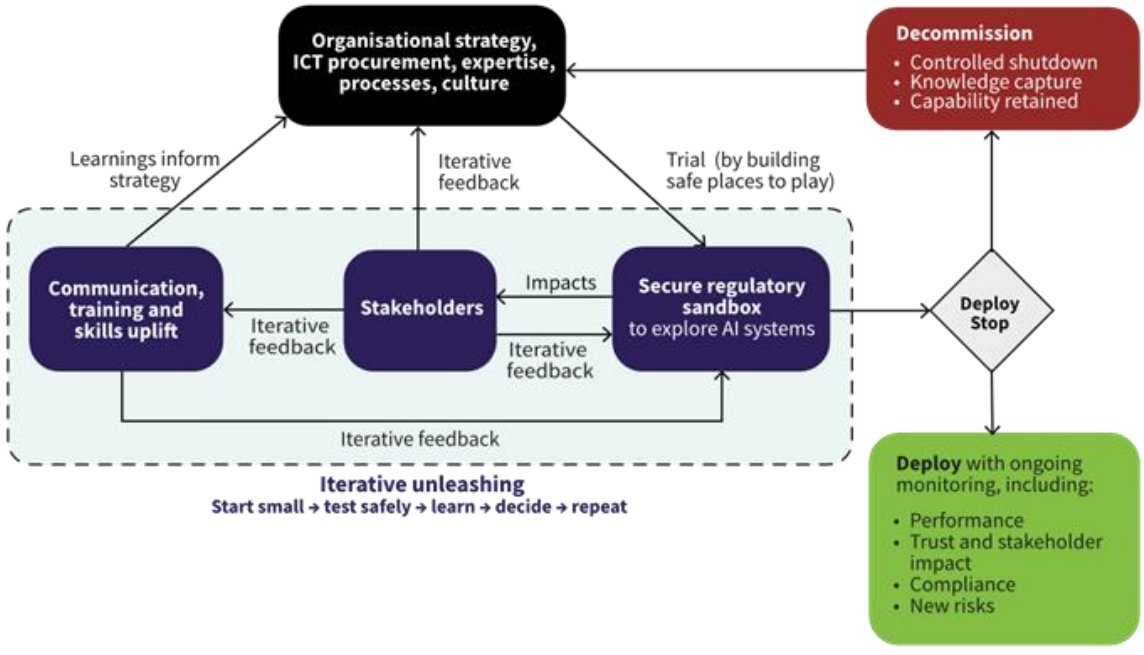
Across conversations with regulators, academics, and system stewards, one idea comes up consistently: capability does not grow from designing the perfect plan. It grows from doing small things well, learning from them, and building from there.

Regulators routinely pilot, test, evaluate, and adapt. Bringing AI into regulatory practice should follow the same pattern. Starting in lower-risk areas, where outputs are easy to check and errors are unlikely to cause harm, allows organisations to build real experience with AI before extending its use to more consequential functions.

This approach is consistent with the AI system lifecycle described in the Public Service AI Framework, which emphasises ongoing testing, monitoring, and improvement rather than one-off deployment.

What iteration is and why it works for AI

The diagram below describes 'iterative unleashing'. This means starting with contained, low-risk experiments, watching closely what happens, and keeping the ability to pause, adjust, or stop if something does not work as expected.



This is a practical method for identifying where AI genuinely adds value, what safeguards are needed, and how to build the organisational confidence to extend use responsibly over time **(see the *Empowering your people* issue)**.

This is closely related to the principles behind regulatory sandboxing. Just as a sandbox creates a bounded space to test whether a regulatory change works before committing to it, an AI pilot creates a bounded space to test whether a tool works in a specific regulatory context before extending its use. The same discipline applies; clear limits, active monitoring, and a genuine commitment to acting on what is learned, including stopping if the evidence points that way.

Iterative unleashing reinforces Human-in-the-loop in practice **(see the *Organisational integration is key* issue)**. Early use cases focus on functions where AI supports people rather than replacing judgement, such as triage, completeness checks, or information synthesis. Statutory decisions about suitability, approvals, or enforcement remain with authorised decision-makers. Starting small and low risk allows regulators to build confidence in how human oversight works before AI is used in decisions that affect rights, obligations, or public trust **(see the *AI in Regulation* issue)**.

Steps for iterative unleashing

1. Treat early adoption as the starting point

Widespread use of tools like Copilot builds familiarity, and that is a reasonable place to begin. But familiarity with a tool is not the same as understanding it well enough to use it safely in a regulatory context. The next step is developing a working understanding of how tools actually produce their outputs, where they are reliable, and where they are not **(see the *Empowering your people* issue)**. This technical grounding does not need to be deep, but it needs to be sufficient for staff to exercise genuine judgement when reviewing AI outputs rather than simply accepting them. From that foundation, regulatory leaders can start to identify where AI fits within specific regulatory functions and where small, safe trials could add value, without moving faster than the organisation's capability to oversee what it is doing.

2. Create an authorising environment for safe experimentation

AI is unlikely to move beyond personal productivity without visible leadership support **(see the *Leadership mindset* issue)**. Staff need to know that experimentation is legitimate, and that trying something that does not work out will not reflect badly on them. But freedom to experiment needs to sit within a clear framework. Leaders need to be confident about what kinds of AI use are appropriate in their regulatory context, and what

safeguards are in place. Encouraging experimentation without that foundation risks AI being used in ways that create problems for accountability and public trust. The goal is an environment where staff feel confident to try AI in low-risk settings and share what they learn, within boundaries that leaders have thought through.

3. Look for regulatory use cases

Start with a narrow, low-risk problem where AI might help. Examples include organising large volumes of public information or identifying themes in submissions or complaints. These are tasks where outputs can be checked easily, errors are unlikely to cause harm, and the work would otherwise consume significant staff time. It is worth considering the social licence implications of using AI for tasks where the public may expect human review, particularly in higher-risk regulatory settings.

Pilots work best when they are designed to generate learning, not just demonstrate success. That means running them in controlled conditions, watching how the tool behaves in practice, and being honest about what the results show. A pilot that reveals problems with underlying data quality, or that staff are not yet confident enough with the tool to use it well, has done exactly what it should.

Ask practical questions like:

- Where are our backlogs, delays, or bottlenecks?
- Where do decision-makers need better information sooner?
- Which high-volume tasks slow down core work?

These questions help identify where AI might add value **(see the *Opportunities and risks* issue)**.

But before settling on an AI solution, it is worth asking whether AI is actually the right tool. Some problems are better solved through process improvement or simpler automation, which carry less risk and are easier to govern. AI should be considered where it offers a genuine advantage over these alternatives.

4. Start small and let progress compound

Given that most of the public sector is still in the early stages of AI use, starting small is the right approach. Small, low-risk pilots build the organisational confidence, capability, and practical understanding needed to use AI well at any scale. This kind of progress requires investment in people alongside investment in tools. Training, communication, and genuine leadership engagement are what turn early experimentation into embedded capability. Each successful small-scale trial creates a foundation for the next one.

5. Build confidence before considering bigger changes

The early goal is building trust that:

- humans remain the decision-makers
- the risks are understood and manageable
- the process remains lawful, fair, and reviewable.

Scaling AI use before this confidence is established creates governance risks that are difficult to unwind.

Build in regular review points to check whether AI systems remain fit for purpose, aligned with delegations and legal obligations, and continue to deliver value without undermining fairness or trust.

Each pilot should be designed so it can be paused or stopped without disrupting core regulatory functions. Ending a pilot that is not delivering is the process working as intended, and it produces information that a large-scale deployment never could. Each cycle finishes with reflection. What worked? What did not? What risks emerged? What governance or oversight needs strengthening? These lessons feed into the next pilot and gradually lift organisational maturity.

6. Learn from AI adoption across the regulatory system

Early AI experiments often stay local to teams or agencies. Progress accelerates when regulators share what they are testing, what worked, and what did not. Peer networks and communities of practice help build collective capability, reduce duplication, and develop a shared understanding of good practice. This learning should feed back into formal guidance and frameworks. Lessons are most valuable when they are captured, shared systematically, and used to update the guidance that shapes how AI is adopted more broadly. The Ministry for Regulation will support this by updating guidance as practice develops.

To get more practical steps for how regulatory leaders can lead AI innovation with confidence, check out the full guidance: **Responsible AI in Action**.