

From Pilot to Production

A government's field guide to scaling AI and digital pilots - the conditions, the checklist, and the failure modes to watch for

Scaling from pilot to production was one of the most discussed problems at the GovTech 4 Impact World Congress 2026. It came up in sessions on AI in public services, legacy modernisation, smart cities, and municipal innovation and the pattern was consistent everywhere. Governments are good at starting pilots. Very few make it to production, and even fewer make it there before the technology moves on. This guide distils what practitioners across those sessions said actually makes the difference, the conditions that need to be in place, the ways pilots most commonly die, and a short checklist to run before you commit to your next one.

This guide is based on insights from two G4I 2026 sessions: "AI in Public Services: Pilots, Scale, and Digital Sovereignty" and "Smart Cities."

The Conditions for Scaling

A pilot has a realistic chance of reaching production when these conditions are in place from the start, not added retroactively once the pilot has already succeeded.

1. A defined business purpose and a named owner

Every pilot needs someone accountable for the outcome, not a department or a political sponsor in the abstract. Without a responsible owner, pilots stall the moment attention moves elsewhere.

2. A committed budget for what comes after the pilot

Integration, maintenance, training, and change management cost more than the pilot itself. If that budget is not committed before the pilot begins, production is unlikely to follow.

3. Foundations addressed before AI is layered on top

Fragmented data, outdated infrastructure, and legacy processes need to be fixed first. AI accelerates what already exists, including its problems.

4. A repeatable, high-volume starting point

The strongest early candidates are narrow, well-defined, high-frequency tasks — document validation, case management, routine enquiries — where quick wins build trust and momentum for larger deployment.

5. Interoperable, modular, and replicable design

Solutions built to connect with existing systems, and structured so components can be reused elsewhere, scale further than bespoke, isolated builds.

6. Workflow redesign, not just tool deployment

Return on investment disappoints when AI is layered onto an unchanged process. The institutions that see real value are the ones that redesign the workflow around the tool, not the other way round.

Common Failure Modes

These are the patterns practitioners named again and again as the reasons pilots stall - worth checking your own pilot against each one.

- **Too many pilots, not enough filtering**

Organisations run parallel pilots without a clear purpose, sometimes duplicating the same idea in different units. Volume is mistaken for progress.

- **The pilot was built around an easy win**

A project that only succeeds with unusually motivated early volunteers is not proof of scalability. If it only works under ideal conditions, it is not ready to scale.

- **The solution doesn't outperform the process it replaces**

Every new system competes against an existing workflow, however inefficient. If the replacement is not reliable enough to consistently outperform it, adoption will not hold.

- **The economics were never real**

Public organisations cannot fund solutions on theoretical value alone. If the benefit does not translate into an actual budget line, the project will not survive its first review.

- **Freemium entry points quietly became lock-in**

A vendor's low-cost or free tool can become deeply embedded in workflows before its true cost, or its exit difficulty, is understood. Public administrations often lack the flexibility to renegotiate once locked in.

- **The technology was chased because it was fashionable**

Some initiatives begin with the solution rather than the problem - pursued because funding was available or the concept was politically attractive, not because it solved a genuine civic need.

- **Production took too long, and the pilot arrived obsolete**

If production takes one or two years after a successful pilot, the solution may no longer be relevant in a fastmoving AI environment. Speed to production is itself a design requirement.

The Pre-Scale Checklist

Before committing resources to move a pilot toward production, you should be able to check off each of the following:

- ☐ We have a named, accountable owner for this pilot — not a department or a political sponsor in the abstract.
- ☐ A budget for integration, maintenance, training, and change management is committed, not just the pilot's initial cost.
- ☐ We have addressed the underlying data and process problems this pilot depends on, not layered it on top of them.
- ☐ This pilot targets a repeatable, high-volume task, not a one-off or highly bespoke use case.
- ☐ The workflow itself is being redesigned around this solution, not left unchanged with a tool added on top.

- We have tested this beyond ideal conditions and motivated early adopters — with a realistic cross-section of users.
- We understand the total cost of this vendor relationship over time, including what it costs to leave.
- We checked whether a mature solution already exists elsewhere before building our own.
- We have a realistic timeline to production, and we have asked whether the solution will still be relevant when it gets there.
- We can articulate the civic need this solves in one sentence, without reference to the technology itself.

The Underlying Message

Success in public-sector AI and digital transformation is not defined by how many pilots are launched. It is defined by whether governments can translate promising ideas into services that are trustworthy, sustainable, and genuinely used. A pilot without a realistic path to production is not an experiment. It is a delay.