

Category

Data & AI

AI Experimentation & Innovation Lab Playbook

A disciplined operating model for running AI pilots that actually reach a build/scale/kill decision instead of dying quietly in "still evaluating" limbo.

- Practitioner
- Intermediate
- Template Included
- Workshop Ready

- Mithun A. Sridharan
- 30 August 2024
- 7 minute(s)

Overview

A practical playbook for running an AI experimentation function, using a three-gate funnel, pilot templates, and a worked retail example that shows how to kill weak ideas fast and scale strong ones.

Isn't killing most pilots a sign the lab is failing?

No — a high kill rate at the Explore and Pilot gates, especially early, is what a working filter looks like. If your lab is scaling nearly everything it pilots, it's likely either testing ideas that were never genuinely uncertain or quietly avoiding hard kill decisions, both of which are worse outcomes than an honest kill rate.

How do we set success criteria for something genuinely novel where we don't have a baseline?

Use a proxy or a manual-process baseline — as in the grocery example, "beat the manual audit process by X%" is a workable criterion even without historical AI performance data. The goal is a number the sponsor agrees in advance would justify scaling, not a perfectly calibrated target.

What happens to the pilot team's work if an idea gets killed?

Document the learnings and the reason for the kill in the Pilot Scorecard and archive the pilot's data and code. Killed pilots are common inputs to future Explore Gate reviews — a kill for weak data availability today can become a viable pilot in a year once that data gap

closes.

Should the lab report to IT, a business unit, or a central innovation function?

This playbook is reporting-structure agnostic, but the gate review committee should always include both technical and business-unit representation regardless of where the lab formally sits, since Scale Gate decisions require both technical feasibility and business ownership commitment.

How many pilots should be running at once?

Size the portfolio to your gate review capacity, not your idea backlog — a committee that can genuinely score 4-6 pilots thoroughly each month should run roughly that many, since the discipline of the framework depends on real attention at each gate, not just paperwork.

```
{ "@context": "https://schema.org", "@type": "FAQPage", "mainEntity": [ { "@type": "Question", "name": "Isn't killing most pilots a sign the lab is failing?", "acceptedAnswer": { "@type": "Answer", "text": "No — a high kill rate at the Explore and Pilot gates, especially early, is what a working filter looks like. If your lab is scaling nearly everything it pilots, it's likely either testing ideas that were never genuinely uncertain or quietly avoiding hard kill decisions, both of which are worse outcomes than an honest kill rate." } }, { "@type": "Question", "name": "How do we set success criteria for something genuinely novel where we don't have a baseline?", "acceptedAnswer": { "@type": "Answer", "text": "Use a proxy or a manual-process baseline — as in the grocery example, \"beat the manual audit process by X%\" is a workable criterion even without historical AI performance data. The goal is a number the sponsor agrees in advance would justify scaling, not a perfectly calibrated target." } }, { "@type": "Question", "name": "What happens to the pilot team's work if an idea gets killed?", "acceptedAnswer": { "@type": "Answer", "text": "Document the learnings and the reason for the kill in the Pilot Scorecard and archive the pilot's data and code. Killed pilots are common inputs to future Explore Gate reviews — a kill for weak data availability
```

today can become a viable pilot in a year once that data gap closes." } }, { "@type": "Question", "name": "Should the lab report to IT, a business unit, or a central innovation function?", "acceptedAnswer": { "@type": "Answer", "text": "This playbook is reporting-structure agnostic, but the gate review committee should always include both technical and business-unit representation regardless of where the lab formally sits, since Scale Gate decisions require both technical feasibility and business ownership commitment." } }, { "@type": "Question", "name": "How many pilots should be running at once?", "acceptedAnswer": { "@type": "Answer", "text": "Size the portfolio to your gate review capacity, not your idea backlog — a committee that can genuinely score 4-6 pilots thoroughly each month should run roughly that many, since the discipline of the framework depends on real attention at each gate, not just paperwork." } }] }

Subscriber access

Unlock this playbook

This playbook — including every framework, template, and step-by-step section — is available free to Think Insights subscribers. Enter your email to unlock it instantly and get our weekly insights newsletter. No account needed, and access is remembered on this device.

First Name

Last Name

Email

Country

I agree to data processing in accordance with GDPR

Leave this field blank

References

Related playbooks

Data & AI

Multimodal AI for Document, Image, Audio & Video Workflows Playbook

A framework for using multimodal AI across document, image, audio, and video workflows — building genuinely integrated processing that handles multiple content

- Practitioner
- Intermediate
- Template Included

[Read playbook](#)

Data & AI

AI for Sustainability, Climate Risk & Transition Planning Playbook

A framework for using AI in sustainability, climate risk, and transition planning — modeling transition pathway scenarios and climate risk exposure with genuine

- Executive
- Advanced
- Template Included

[Read playbook](#)

Data & AI

AI Storytelling, Internal Communication & Adoption Campaigns Playbook

A framework for AI storytelling, internal communication, and adoption campaigns — using specific, concrete narratives of genuine AI-driven outcomes rather than

- Practitioner
- Beginner
- Template Included

[Read playbook](#)

[Talk to our AI innovation team](#)

Author

Mithun A. Sridharan

I'm Mithun A. Sridharan, Founder of this website - Think Insights - on Strategy, Management Consulting, Leadership, Digital Transformation, and Data Literacy. Follow me on social media or connect with me on LinkedIn for updates.