

Reality Check?

Moving GenAI from Prototype to Production

Nitin Kumar



Nitin Kumar

AI & ML Strategist | Driving Business Value through
AI, GenAI | AI, GenAI Leader



Beyond the Pilot Phase

Making GenAI Actually Work in the Real World

The Pilot Cliff

Great demos.
Little impact.

95%

of GenAI pilots show no measurable P&L impact

MIT, TechRadar, Tom's Hardware

OLD QUESTION:

"How smart is the model?"

NEW QUESTION:

"Can the business absorb it?"

1

Balancing Innovation with Oversight

Especially in Regulated Environments

Guardrails before gadgets



Policy

Data

People

Technology

3 Lines of Defense

DEFENSE 1

At the Prompt

- System prompts that request help vs. guessing
- Role-based restrictions
- Input validation and sanitization

DEFENSE 2

At the Output

- PII and toxicity filters
- Business rule enforcement
- Hallucination detection

DEFENSE 3

At the Environment

- Access controls and RBAC
- Audit trails and logging
- Enterprise risk mapping

2

Embedding GenAI into Workflows

Without Blowing Things Up

Today – Manual Workflow

- 1 Agent opens customer email
- 2 Searches through 3 systems manually
- 3 Copies notes into CRM
- 4 Writes response from scratch

With GenAI – Co-Pilot Workflow

- 1 Email arrives, GenAI auto-reads
- 2 AI pulls context from all systems
- 3 AI drafts response & highlights risks
- 4 Human reviews, edits, and sends

HUMAN REMAINS IN CONTROL

Embedding Playbook

Start at the edges

Begin with cognitive grunt work: summarization, note-taking, draft replies. Not decisions that move money or affect compliance.

Make it default, not mandatory

The co-pilot always offers help, but humans can ignore it. Forcing usage before trust collapses adoption.

Let the frontline tune it

Best prompts come from people on phones, at desks, in operations—not the lab. Give them a sandbox and a path to production.

3

Hallucinations & Measurement

What Really Matters?

Hallucinations are not equal

LOW-RISK

Creativity: brainstorming, ideation, storyboarding. Making things up might be useful.

MEDIUM-RISK

Internal ops: SOPs, meeting summaries, code snippets. Annoying but caught by humans.

HIGH-RISK

Customer, money, law: communication, financial decisions, legal/medical content. Unacceptable.

Mitigation Stack

Grounding (RAG, tools)

Model answers from your verified knowledge and data, not the internet

Policies (system prompts)

Make the model say "I don't know" instead of confidently hallucinating

Filters (toxicity, PII)

Scan outputs for PII, toxicity, bias, regulatory violations before delivery

Humans (review, override)

Human-in-the-loop for high-risk flows, complex cases, escalations

Measure What Matters



Business Value

- Handle time reduced
- First-contact resolution
- Revenue uplift per campaign
- Hours saved per employee



Risk Reduction

- % outputs flagged by filters
- Hallucinations caught pre-customer
- Incidents & complaints
- Regulatory findings



Adoption & Trust

- Suggestion acceptance rate
- Override frequency
- Voluntary usage patterns
- User satisfaction scores

4

Governance for Scalable GenAI

Ship Faster. Sleep Better.

4 Layers of GenAI Governance

1 Strategy

Clear north star: Why are we using GenAI? Portfolio view of use cases, not random experiments.

2 Use-Case Lifecycle

Standard path from idea → pilot → limited rollout → scale. Clear gates at every step.

3 Controls & Guardrails

Catalog of approved models, standard patterns for grounding/filtering, regulatory templates.

4 Monitoring & Audit

Continuous monitoring of performance, drift, incidents. Regular audits for regulators and board.

From Hype to Habit

PICK ONE

Workflow

MANAGE ONE

Risk

MEASURE ONE

Metric

Make GenAI part of how your company works every day

Thank You

Questions?



Nitin Kumar

AI & ML Strategist | Driving Business Value through
AI, GenAI | AI, GenAI Leader



Connect with me on LinkedIn