STARTAD
powered by TAMKEEN

with support from Google.org

THE EVALUATE & PILOT AI SOLUTIONS PLAYBOOK

Index

01 Introduction

1.1 Purpose of the Playbook 3
1.2 What You'll Achieve 6
1.3 Who Should Use This Playbook 7
1.4 How to Use This Playbook 8
1.5 Playbook Roadmap 10

02 Understanding the Impactful AI Method

2.1 Methodology Overview 12
2.2 Why This Methodology Works 14
2.3 Methodology at a Glance 16
2.4 Key Principles Deep Dive 19

03 Evaluating AI Solutions – Step by Step

3.1 Introduction to Section 21
3.2 Step 1: Organization & Challenge Context 22
3.3 Step 2: Stakeholder Identification 25
3.4 Step 3: Define Quantified Objectives 30
3.5 Step 4: Regulatory & Compliance Requirements 34
3.6 Step 5: Define AI Solutions 41
3.7 Step 6: Impact Estimation 50
3.8 Step 7: Compare Solutions 55
3.9 Step 8: Make the Decision 61

04 Designing Your Pilot Roadmap

4.1 Introduction to Evolutionary Piloting 67
4.2 Step 1: Pilot Scope & Success Criteria 69
4.3 Step 2: Breaking Down into Increments 77
4.4 Step 3: Complete Pilot Roadmap 83
4.5 Step 4: Readiness & Launch Preparation 88

05 Running a Successful Pilot

5.1 Introduction 96
5.2 Measurement & Data Collection 97
5.3 How to operationalize measurement (easy rules) 97
5.4 Data collection methods (pick the simplest that works) 97
5.5 Best practices (data quality without heavy process) 97
5.6 Stakeholder Engagement & Communication 100
5.7 Communication cadence (minimum you need) 100
5.8 GO / PIVOT / NO-GO Decision Making 106
5.9 How to run an increment review (easy rules) 106
5.10 Common Pitfalls & How to Avoid Them 109

06 Real-World Usecases

6.1 Introduction 112

6.2 Purpose of these use cases (why they're here) 112

6.3 How to use these examples (simple) 112

07 Conclusion and Next Steps

7.1 Core concepts (6–8 key points) 118

7.2 Critical success factors (what makes PB2 work) 119

7.3 Common pitfalls to avoid 119

08 Appendices

8.1 Appendix A: Template Library Index (Quick Reference) 120

Section 3 - Evaluate & Select (Templates 3.1–3.9) 120

Section 4 - Design the Pilot Roadmap (Templates 4.1–4.5) 120

Section 5 - Run the Pilot (Templates 5.2–5.5) 120

8.2 Quick reference: "Which template do I use?" 121

8.3 Appendix B: Glossary 122

AI / ML terms 122

Healthcare terminology (SDG 3 context) 123

Evaluation & methodology terms (PB2 / Gilb-aligned) 124

Regulatory & compliance terms (UAE/KSA-ready) 126

8.4 Appendix C: UAE/KSA Regulatory Quick Reference (AI Pilots in Healthcare) 129

A) UAE — Summary Table (quick scan) 129

B) KSA — Summary Table (quick scan) 130

C) Official Contact Channels 131

D) How to use this appendix inside PB2 (simple rule) 132

E) "Who to involve internally" (contact info placeholders) 132

8.5 Appendix D: Methodology Reference Sheet 133

1) The Impactful AI Method - Summary (PB2 in one view) 133

2) Tom Gilb Principles — Quick Reference 134

3) Decision Trees 134

4) Flowchart-style "stop/go" gates (PB2 gates) 135

5) Practical "minimum set" to run PB2 135

1.1 Purpose of the Playbook

AI tools are already being deployed in healthcare and mission-driven work—patient communications, scheduling, documentation, fraud detection, workforce support, and more. The hard part is not finding tools. The hard part is **choosing the right one and onboarding it safely.**

This playbook helps you run a **simple, repeatable process** to:

- compare AI options using the same criteria,
- make trade-offs visible (impact vs cost vs risk),
- prove value quickly in a controlled pilot,
- and make a clear decision: **Scale / Pivot / Stop.**

Think of PB2 as your bridge from **“we have a validated AI challenge” (PB1)** to **“we selected the best-fit solution and proved it with evidence.”**

1.1 Purpose of the Playbook (Cont.)

Why Systematic AI Evaluation Matters (Healthcare + Mission Context)

In healthcare and mission-driven settings, AI decisions affect real people and real outcomes:

- patient safety and service quality,
- staff workload and workflow reliability,
- privacy and regulatory compliance,
- trust (patients, funders, partners),
- and scarce resources (time and budget).

Without a systematic method, teams often default to:

- the “best demo,”
- the loudest internal stakeholder,
- or a rushed rollout.

That’s how pilots become expensive, confusing, and hard to stop-even when value isn’t proven.

PB2 reduces that risk by enforcing three disciplines:

Measure success upfront	Make uncertainty explicit	Pilot in short increments
(baseline → target)	(credibility, not guesswork disguised as certainty)	(GO / PIVOT / NO-GO see Section 5)

1.1 Purpose of the Playbook (Cont.)

Connection to SDG 3 (Good Health and Well-being)

PB2 is designed for SDG-aligned health work, especially **SDG 3:**

"Ensure healthy lives and promote well-being for all at all ages."

PB2 keeps SDG 3 central in two ways:

1. Beneficiary-first framing is required.

You start evaluation by documenting who benefits, what problem they face, and what "better" looks like.

2. Impact is made defensible.

You quantify improvement (not just internal efficiency) so results can be explained clearly to leadership, governments, donors, and partners.

In short: **SDG 3 defines what matters. PB2 provides the method to choose responsibly and prove impact with evidence.**

1.2 What You'll Achieve

By the end of this playbook, you will produce a decision-ready package you can take to leadership.
Specifically, you will have:

1. Quantified Impact Estimation Table (IET)

A completed table that estimates how each solution's features impact your objectives, with **credibility scoring** so uncertainty is explicit.

2. Solution Comparison Dashboard (one-page view)

A simple side-by-side comparison across **impact, cost, credibility, and value-for-money** (impact-to-cost ratio).

3. Decision Framework & Recommendation Report

A short, leadership-ready document that records **criteria, weights, scores, final recommendation, and rationale** (or a decision not to proceed).

4. Regulatory & Compliance Checklist (UAE/KSA-ready)

A captured set of requirements and risks identified early—so compliance issues don't appear after selection.

5. 12-Week Pilot Roadmap (increment-based)

A complete pilot plan broken into **1–2 week increments,** each with scope, value delivered, success criteria, and **GO / PIVOT / NO-GO** rules.

6. Pilot Governance Pack (run-the-pilot toolkit)

A practical set of templates to operate the pilot: **measurement plan, communication cadence, weekly status report, and increment review agenda**—so decisions are based on evidence, not opinions.

1.3 Who Should Use This Playbook

This playbook is for teams who need to choose an AI solution and onboard it through an evidence-driven pilot (especially in healthcare or mission-driven programs).

Role	Typical use case (what they're trying to do)	Key sections to focus on
Executive / Program Leaders	Approve investment, set risk tolerance, make the final "pilot / no pilot" call	**Section 1–2** (why + method), **Section 3 (Steps 7–8)** (comparison + decision), **Section 4** (roadmap + readiness gate)
Technology / IT Leaders	Validate feasibility: integration, security, data access, operating model	**Section 3 (Steps 1–6)** (inputs, compliance, solution definition), **Section 4 (Step 4)** readiness checklist, Section 5 pilot operations + incident handling
Business Analysts / Strategy	Convert vendor claims into measurable projections and comparisons	**Section 3 (Steps 3, 6, 7)** (objectives → IET → dashboard), Step 8 recommendation report
Clinical / Operations Owners	Ensure workflow fit, safety, and adoption in real settings	**Section 3 (Steps 1–4)** (context, stakeholders, compliance), **Section 4** (scope + success criteria + increments), **Section 5** measurement + increment reviews
Project Managers / PMO	Run the pilot cadence and keep delivery measurable and controlled	**Section 4** (roadmap + resources), **Section 5** (measurement plan, comms plan, weekly status, GO/PIVOT/NO-GO agenda)
Consultants / Facilitators	Lead workshops, complete templates, and produce a decision-ready package	**Section 3 end-to-end, Sections 4–5 end-to-end**, plus use cases in **Section 6** as reference patterns

1.4 How to Use This Playbook

Approach: Use PB2 like a checklist. Don’t skip steps-each one feeds the next.

01 **Bring PB1 inputs** (Start here)
You should already have a clear challenge statement, beneficiaries, and context from Playbook 1.

02 **Run the evaluation** (Section 3)
Complete the 8 steps in order to produce: objectives → compliance requirements → solution definitions → impact estimates → comparison → decision.

03 **Pick one solution** (End of Section 3)
Make a clear call: **pilot this** / **don’t pilot**. Document the rationale so leadership can approve it.

04 **Design the pilot** (Section 4)
Turn the chosen solution into a 12-week roadmap broken into **1–2 week increments**, each with success criteria and GO/PIVOT/NO-GO rules.

05 **Run the pilot** (Section 5)
Operate with a weekly rhythm: measurement, status update, & an increment review meeting where decisions are made based on evidence.

Rule of thumb: if a template is incomplete, don’t move forward.

1.4 How to Use This Playbook (Cont.)

Time commitment estimates (realistic and lightweight)

These ranges assume 2–3 candidate solutions and a small stakeholder group.

- **Preparation** (PB1 inputs + gather solution materials): 2–6 hours
- **Evaluation** (Section 3, Steps 1–8): 1–2 working days total (often spread over 1–2 weeks)
- **Pilot design** (Section 4): 4–8 hours
- **Pilot execution** (Section 5): 6–12 weeks typical (example roadmap uses 12 weeks)

Prerequisites (from Playbook 1)

PB2 is not meant to start from zero. It assumes you already completed **Playbook 1** (AI Use Case Identification) and have:

| A clear AI challenge statement (problem, root causes, desired outcomes, constraints)

| Defined beneficiaries + mission context (SDG 3 alignment)

| At least one prioritized use case ready for evaluation (ideally your top PB1 candidate)

1.5 Playbook Roadmap

PB2 is a practical workflow you can follow from start to finish: **PB1 inputs → evaluate options → choose one → pilot in small increments → decide Scale / Extend / Stop.**

Figure 1.1 shows the full journey at a glance and highlights the decision gates so teams don't over-commit too early.

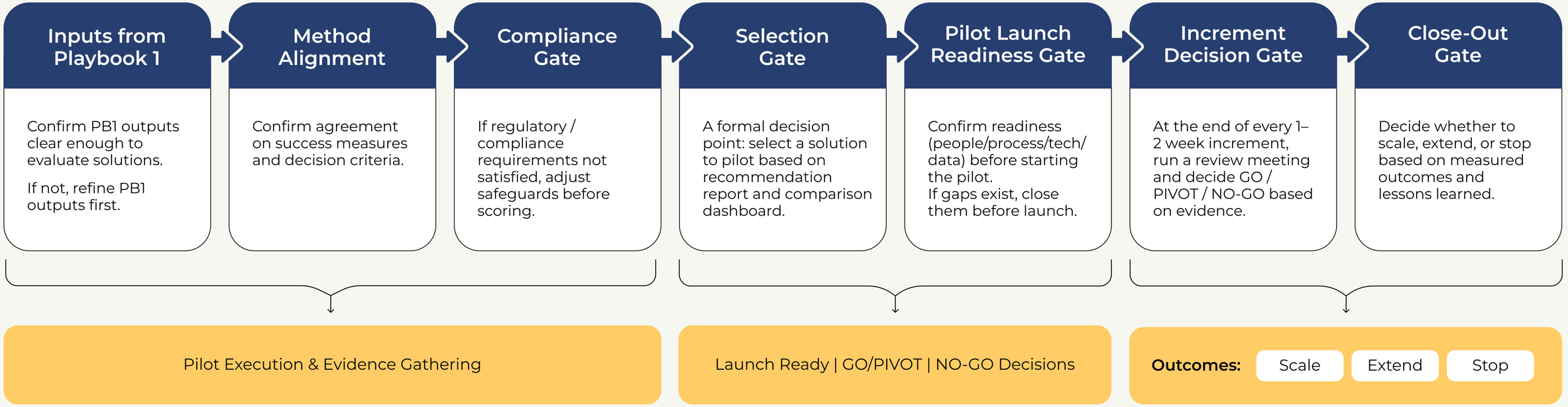


Figure 1.1: Playbook Roadmap

Section 1 Outputs

What you should have by the end of Section 1 :

A shared understanding of what PB2 is for: **evaluate → select → pilot → decide Scale / Extend / Stop.**



Agreement that evaluation will be **systematic** and **evidence-based** (not vendor-led or opinion-led).



Alignment that **SDG 3 / beneficiary impact** is the primary lens for success in healthcare and mission-driven work.



Clear visibility of the **decision gates** you will use throughout the playbook (Input readiness, Method alignment, Compliance, Selection, Launch readiness, Increment decisions, Close-out).



2.1 Methodology Overview

What is the Impactful AI Method? The Impactful AI Method is a practical way to choose the right AI solution and onboard it safely. It helps teams compare options using the same yardstick, then prove value quickly in a controlled pilot.

It answers three simple questions:

- Which option creates the most value for our goals and constraints?
- What's the safest way to validate value fast (before scaling)?
- What do we know vs. what are we guessing (so uncertainty is visible)?

AI tools can speed up analysis (drafting, structuring, estimating), but humans **validate the assumptions** and make the **final decision**.

How it differs from traditional AI evaluation

Traditional AI selection often relies on the **wrong evidence**:

- "Best demo wins" (persuasion over proof)
- "We'll measure later" (no baseline, no targets)
- "Let's roll it out broadly" (big bang risk)
- Politics and opinions override data

The Impactful AI Method replaces that with a simple workflow:

- Define **measurable objectives** (baseline → target)
- Estimate and compare **impact** across options
- Make **trade-offs** explicit (impact vs cost vs credibility)
- Pilot **in small increments** with GO/PIVOT/NO-GO decisions

2.1 Methodology Overview (Cont.)

Core principles (the 3 rules behind the method)

Quantified impact (measure, don't guess)

You define what success looks like using measurable targets, then estimate how each solution's features will move those targets. You also score credibility so assumptions are visible.

Evolutionary approach (learn fast, reduce risk)

Instead of a big rollout, you test in 1–2 week increments, learn quickly, and make decisions using evidence.

Explicit trade-offs (make decisions explainable)

You compare solutions across impact, cost, credibility, and value-for-money, so the final recommendation is clear and defensible.

2.2 Why This Methodology Works

Common pitfalls of traditional approaches
(and why they fail)

01 Vendor promises win

A demo looks great, but it doesn't reflect your reality (data, workflows, integrations, compliance). Teams discover misfit only after time and money are spent.

02 Gut feelings drive decisions

Different stakeholders care about different outcomes (safety, speed, cost, workload). Without clear metrics, alignment is impossible - and results can't be proven.

03 Big bang implementation

Rolling out broadly before validating value increases cost and operational risk. Problems surface late, and it becomes hard to stop due to sunk cost and politics.

How quantified evaluation reduces risk

01 It makes success measurable

You define objectives as baseline → target, so you can verify whether the pilot actually improves outcomes.

02 It creates fair comparison

Solutions are compared using the same criteria, so you're not choosing based on storytelling or isolated scoring.

03 It makes uncertainty visible

Credibility scoring highlights which assumptions need testing before scaling.

04 It limits risk through short increments

Every 1–2 weeks you review evidence and decide GO / PIVOT / NO-GO, so you learn fast and avoid over-committing.

2.2 Why This Methodology Works (Cont.)

Tom Gilb's principles

The Impactful AI Method applies Tom Gilb's ideas from Impact Estimation and Evolutionary Project Management (Evo).

A) Impact Estimation (make impact measurable and comparable)

Impact Estimation is about expressing improvements in a structured way so different options can be compared fairly. In PB2, this is reflected in four practical ideas:

- **Define quality attributes with clear scales** (what you're improving, and how it is measured).
- **Capture current state vs target state** (baseline → goal).
- **Treat uncertainty explicitly** using confidence/credibility (don't hide guesses).
- **Show trade-offs** so the decision is transparent.

In other words: you're not trying to predict the future perfectly - you're trying to compare options responsibly.

B) Evo (evolutionary delivery: prove value in small steps)

Evo is about delivering value in small, testable increments rather than large launches. PB2 translates this into:

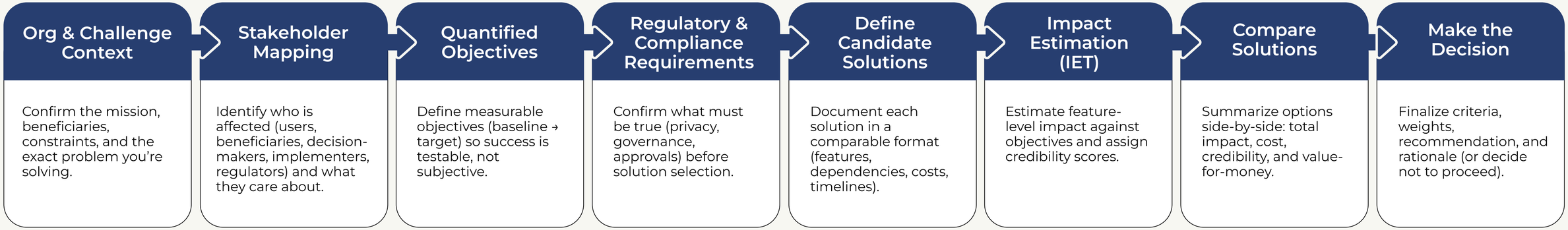
- **Small, valuable increments (1–2 weeks)**
- **Early stakeholder feedback**
- **Measurable success criteria**
- **GO/NO-GO decisions based on evidence** (and GO/PIVOT/NO-GO in pilot governance).

The simplest way to remember the method

- Quantify what "success" means.
- Compare options using impact + cost + credibility.
- Pilot in small increments and decide based on evidence.

2.3 Methodology at a Glance

This section gives you the full workflow at a glance: 8 evaluation steps (Section 3) followed by pilot roadmap design (Section 4) and pilot execution (Section 5). The goal is to make the overall journey easy to understand before you start filling templates.



Phase A - Preparation (PB1 inputs + solutions materials): 2–6 hours

Phase B - Evaluation (Section 3, Steps 1–8): 1–2 working days (often spread across 1–2 weeks)

Phase C - Pilot design (Section 4): 4–8 hours

Phase D - Pilot execution (Section 5): 6–12 weeks typical

2.3 Methodology at a Glance (Cont.)

The 8-step evaluation workflow (Section 3)

These eight steps are designed to be completed in sequence, because each step produces inputs required by the next one.

01 Org & Challenge Context (Template 3.1)

Confirm the mission, beneficiaries, constraints, and the exact problem you're solving.

02 Stakeholder Mapping (Template 3.2)

Identify who is affected (users, beneficiaries, decision-makers, implementers, regulators) and what they care about.

03 Quantified Objectives (Template 3.3)

Define measurable objectives (baseline → target) so success is testable, not subjective.

04 Regulatory & Compliance Requirements (Templates 3.4–3.5)

Confirm what must be true (privacy, governance, approvals) before solution selection.

05 Define Candidate Solutions (Template 3.6)

Document each solution in a comparable format (features, dependencies, costs, timelines).

06 Impact Estimation (IET) (Template 3.7)

Estimate feature-level impact against objectives and assign credibility scores.

07 Compare Solutions (Template 3.8)

Summarize options side-by-side: total impact, cost, credibility, and value-for-money.

08 Make the Decision (Template 3.9)

Finalize criteria, weights, recommendation, and rationale (or decide not to proceed).

2.3 Methodology at a Glance (Cont.)

What happens after Step 8 (pilot design + pilot execution)

Once a solution is selected, PB2 shifts from evaluation to onboarding through a controlled, evolutionary pilot:

Section 4 - Design the Pilot Roadmap

Break the pilot into 1–2 week increments, define success criteria, and set GO/NO-GO decision points; a readiness checklist confirms you are prepared to launch.

Section 5 - Run the Pilot

Use a measurement plan, communication plan, weekly status reports, and an increment review meeting agenda to run GO/PIVOT/NO-GO decisions based on evidence.

2.4 Key Principles Deep Dive

This section explains the four principles you'll apply throughout PB2. Keep them in mind as "rules of the playbook."

01

Quantified objectives (not vague goals)

What it means

Replace goals like "improve patient experience" with measurable targets like "increase access CSAT from 3.6/5 to 4.2/5" or "reduce wait time from 45 min to 30 min."

Why it matters

If you can't measure it, you can't prove it-and you can't make a defensible decision.

How PB2 applies it

- You define objectives as baseline → target (Template 3.3)
- The pilot success criteria are built from those objectives (Template 4.2)

02

Stakeholder-centric measurement

What it means

Impact must be defined from the perspective of real stakeholders: patients/ beneficiaries, clinicians, operations, compliance, leadership.

Why it matters

Different groups value different outcomes. Without stakeholder clarity, pilots generate conflict ("this helped" vs "this hurt").

How PB2 applies it

- You map stakeholders and their concerns (Template 3.2)
- Each objective is linked to the stakeholders who care about it (Template 3.3)

03

Evolutionary piloting (1–2 week increments)

What it means

Don't "launch and hope." Test in short increments, learn quickly, and improve each cycle.

Why it matters

Short increments reduce risk and prevent wasted time. You can stop early, pivot fast, and keep progress visible.

How PB2 applies it

- You plan a 12-week pilot as 1–2 week increments (Template 4.3 + 4.4)
- Each increment has a clear learning goal + decision rule

04

Evidence-based decision making

What it means

Decisions are based on evidence (metrics + real usage + incidents), not demos, senior opinions, or optimism.

Why it matters

It keeps the pilot honest and prevents "pilot drift" (continuing without learning).

How PB2 applies it

- You measure consistently (Template 5.2)
- You run an increment review and decide GO / PIVOT / NO-GO (Template 5.5)

Section 2 Outputs

What you should have by the end of Section 2 :

A shared understanding of the **Impactful AI Method** and why PB2 uses it.

Alignment on the **3 core principles**: quantified impact, evolutionary delivery, explicit trade-offs.

A clear view of the 8-step evaluation workflow (Section 3) and which templates it uses (Templates 3.1–3.9).

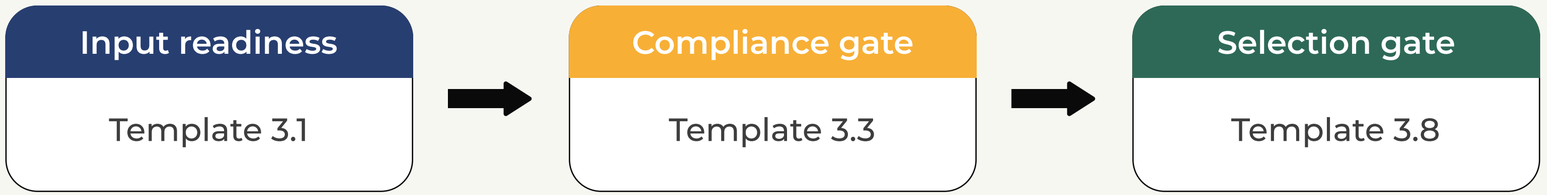
Agreement on the operating rules for the pilot: 1–2 week increments and evidence-based GO / PIVOT / NO-GO decisions.

Introduction to Section

If you completed **Playbook 1 (AI Use Case Playbook)**, you already have your challenge statement, target beneficiaries, and organizational context. In this section, you will use those PB1 outputs as the core inputs for evaluating AI solutions in a structured and comparable way.

This section walks you through an **8-step evaluation process** that moves from “what problem are we solving?” to “which solution should we pilot?” Each step produces a concrete output (template) that feeds the next step—so you can avoid scattered analysis and make decisions based on evidence, not opinion.

Decision Gates



Stop if Fail Go if Pass

Step 1: Organization & Challenge Context

Why this step matters

Evaluation only works when everyone is solving the same problem for the same beneficiaries, within the same constraints. If context is vague, teams compare solutions based on assumptions—and the “best demo” usually wins.

This step prevents that by turning PB1 outputs into a short, shared evaluation baseline.

Important:

Most of this step is already completed in Playbook 1. In PB2, you are simply summarizing PB1 outputs in a format that is easy to use for evaluation.

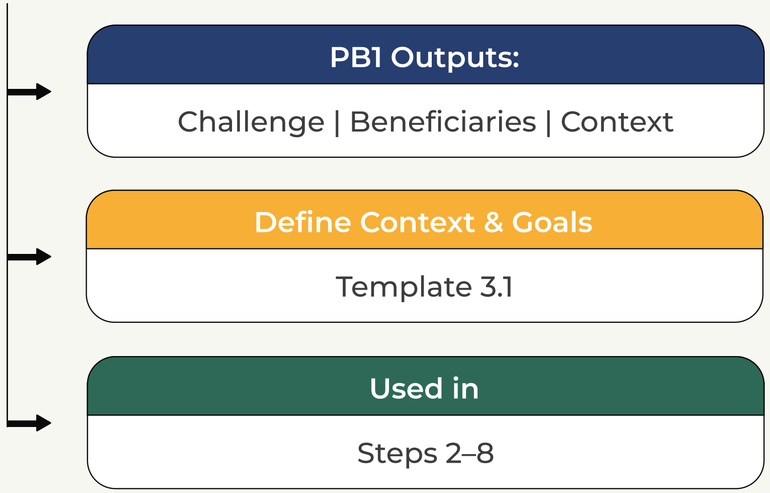
Instructions (how to complete Step 1)

1. Copy your PB1 challenge statement (don’t rewrite from scratch).
2. Confirm who benefits (SDG 3 alignment) and who the solution is for.
3. Write the problem as a single sentence: “Today we have X, causing Y, for Z group.”
4. Define 3–5 desired outcomes (what “better” looks like).
5. List root causes (avoid jumping to solutions).
6. List constraints that will affect solution choice (data, integration, budget, time, compliance).

Rule of thumb:

If two stakeholders disagree on the problem statement, stop and align before continuing.

PB1 → PB2 Input Bridge



To be filled

Template 3.1: Organization & Challenge Overview

(This is completed using **PB1 outputs**: challenge statement + beneficiaries + context.)



Organization

Name: _____

Region: _____

Mission / Vision (1–2 lines):

Beneficiaries (SDG 3 focus)

Primary beneficiaries: _____

Secondary beneficiaries (optional):

SDG 3 link (1 line): _____

Problem (1 sentence):

Desired outcomes
(3–5 bullets)

Root causes
(top 3)

Constraints
(what must be respected)

Data: _____

Integration / systems: _____

Operations / people: _____

Budget / timeline: _____

Template 3.1: Organization & Challenge Overview

Use as only an example

(This is completed using **PB1 outputs**: challenge statement + beneficiaries + context.)

Organization

Name: Community Outpatient Clinics Network

Region: UAE (Dubai)

Mission / Vision (1–2 lines)**:** Deliver timely, safe outpatient care and improve access for the local community.

Beneficiaries (SDG 3 focus)

Primary beneficiaries: Patients seeking outpatient services (general + chronic follow-up)

Secondary beneficiaries (optional)**:** Frontline staff (reception, call center)

SDG 3 link (1 line)**:** Improves access and continuity of care by reducing delays and missed appointments.

Problem (1 sentence):

Patients face long wait times and inconsistent appointment communication, increasing no-shows and staff workload.

Desired outcomes
(3–5 bullets)

- Reduce wait time (arrival → first clinical contact)
- Reduce repetitive inquiry workload for staff
- Reduce follow-up no-show rate
- Improve patient experience for access touchpoints
- Maintain safe escalation/ handoff reliability

Root causes
(top 3)

- High volume of repetitive questions handled manually (phone/ WhatsApp/walk-ins)
- No standardized, multilingual patient information delivery
- Scheduling reminders and updates inconsistent across clinics

Constraints
(what must be respected)

Data: Scheduling + inquiry data exists but is fragmented across channels

Integration / systems: Must work with existing scheduling/EHR; semi-manual acceptable early

Operations / people: Must not disrupt clinic operations; bilingual communications required

Budget / timeline: Pilot must produce evidence in 12 weeks

Step 2: Stakeholder Identification

Why this step matters

AI creates impact (and risk) for different people. If you don't map stakeholders early, your pilot will fail later due to:

- low adoption ("no one asked us"),
- blocked approvals (privacy/compliance shows up late),
- or "success disagreements" (leaders say it worked, frontline says it didn't).

This matches Gilb's idea of stakeholder-centric measurement: impact must be defined in terms of who cares and what "better" means for them.

Instructions (how to do it fast)

1. Start with your PB1 use case and list everyone who touches the workflow.

2. For each group, capture only 4 things:

- Who are they? (group name)
- What's their role? (user / decider / implementer / regulator, etc.)
- How much influence? (High / Medium / Low)
- What do they care about? (top 1–3 concerns)

3. Keep it practical: aim for 8–15 stakeholders (not 40).

4. Identify your top 3 stakeholders (these drive pilot success).

Step 2: Stakeholder Identification (Cont.)

Guidance: Stakeholder categories

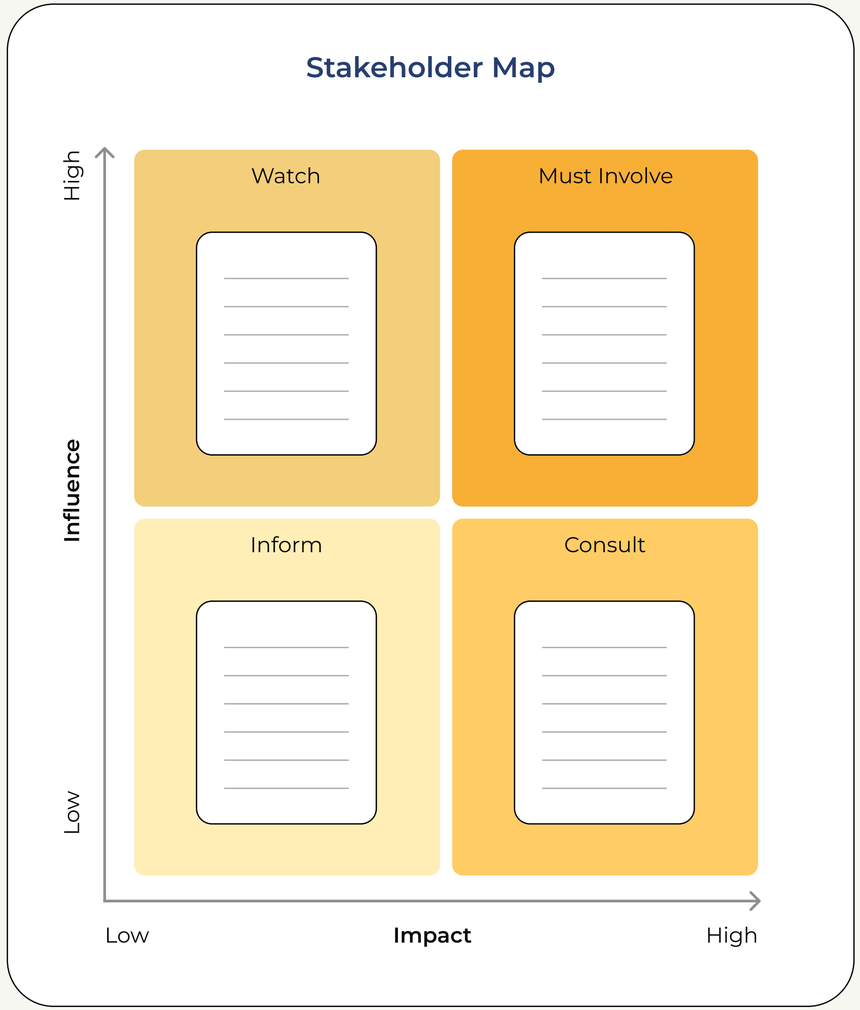
Use these categories so you don't miss critical groups:

Stakeholder category	Who this includes (examples)	How AI Helps
Direct users	Frontline staff who use the tool daily (reception, nurses, call center agents)	Adoption, workflow fit, operational success
Beneficiaries	Patients, caregivers, community members impacted by the service	Safety, trust, experience, SDG 3 outcomes
Decision makers	Sponsor, executive leadership, budget owners	Funding approval, risk tolerance, scale decision
Implementers	IT, data, operations, PM/PMO	Integration, reliability, delivery execution
Risk gatekeepers	Privacy/DPO, compliance, legal, clinical safety	Approvals, regulatory fit, risk management
Influencers	Department heads, site managers, champions, unions/HR (if relevant)	Local buy-in, behavior change, rollout support
External stakeholders	Regulators, partner clinics, funders/donors	External requirements, reporting, partnership constraints

To be filled

Template 3.2: Stakeholder Mapping Worksheet

Stakeholder (name/group)	Role (user/decider/implementer/ regulator/etc.)	Influence (H/M/L)	Key concerns (1–3 bullets)

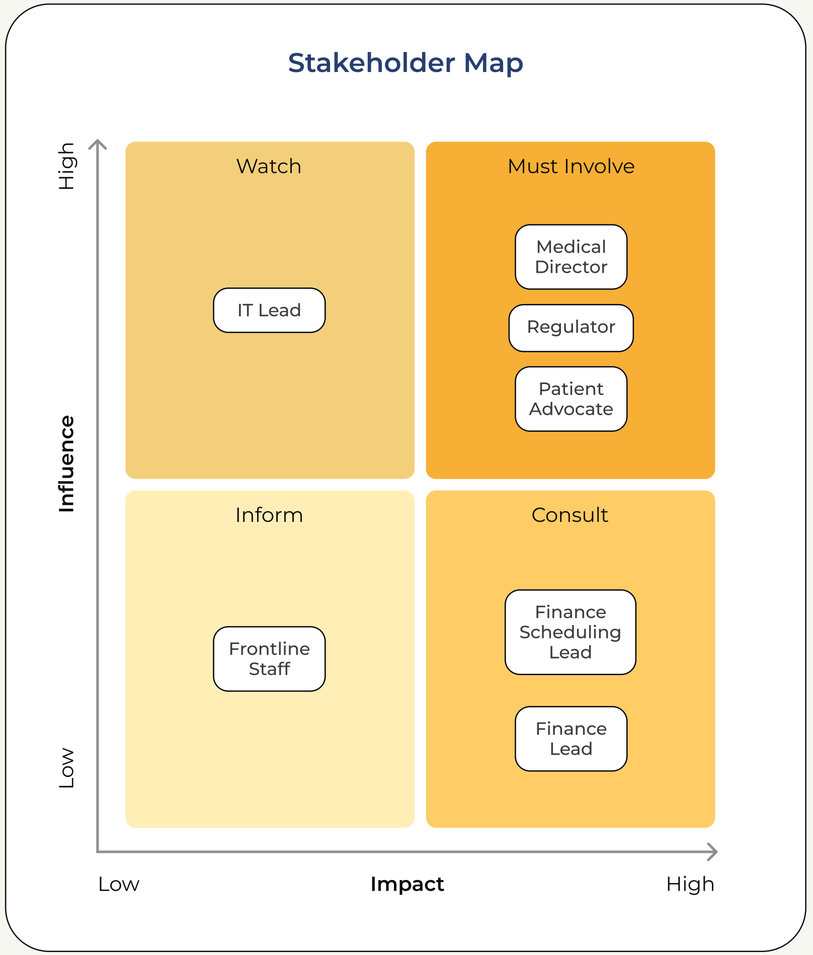


Template 3.2: Filled Example

Patient Access Pilot (Clinic Network)

Use as only an example

Stakeholder (name/group)	Role	Influence	Key concerns (1–3 bullets)
Program Director (Outpatient Services)	Decision maker / sponsor	High	ROI, risk tolerance, scale decision
Clinic Operations Lead	Owner / implementer	High	Workflow disruption, staffing impact
Reception staff (Clinic A)	Direct users	High	Ease of use, fewer repetitive questions, fast handoff
Call center pod lead	Direct users / champion	Medium	Ticket quality, reduced call volume, routing accuracy
Patients (outpatient)	Beneficiaries	Medium	Clear instructions, privacy, trust, language support
IT Integration Lead	Implementer	High	Integration effort, uptime, monitoring, rollback plan
Data/Analytics Analyst	Implementer	Medium	Data availability, metric definitions, reporting cadence
DPO / Privacy	Risk gatekeeper	High	Data residency, consent/notice, audit logs
Clinical Owner	Safety gatekeeper	High	No clinical advice, escalation rules, safety monitoring



Exercise 3.1: Stakeholder Identification Workshop

Goal: Identify the full stakeholder map and agree on top priorities before you start scoring solutions.
Participants (recommended)**:** Sponsor, PM, Ops owner, IT, Privacy/Compliance, 1 frontline champion.

Instructions (how to do it fast)

01 **5 min** | Re-read the Step 1 problem statement (Template 3.1).

02 **10 min** | Brainstorm stakeholders using categories table (above).

03 **10 min** | Fill Template 3.2 (one row per stakeholder group).

04 **10 min** | Mark the Top 3 stakeholders (highest influence + most affected).

05 **5 min** | Capture the Top 3 concerns to track during the pilot.

Output: Completed Template 3.2 + Top 3 stakeholders clearly identified.

Step 3: Define Quantified Objectives

Why this step matters

Most AI projects fail because “success” is vague. Gilb’s approach is straightforward: if you can’t measure improvement, you can’t prove value - and you can’t compare solutions fairly.

This step turns goals like “improve patient experience” into objectives you can test: baseline → target | on a clear scale | for a specific stakeholder

Instructions

Step 3.1 - Start from stakeholders (not from metrics)

From Step 2, pick your top 3–4 stakeholders (highest influence + most affected).

For each stakeholder, ask: “**If this pilot is successful, what would be measurably better for you**?”

Write 1–3 objective ideas per stakeholder first (plain language).

Examples:

- Patients: “faster access”, “clearer appointment instructions”
- Call center: “fewer repetitive calls”, “fewer incorrect escalations”
- Ops lead: “reliable handoffs”, “shorter resolution time”

Step 3.2 - Convert each objective into a measurable objective (fill 6 fields)

Now convert each stakeholder objective idea into a quantified objective by filling:

1. **Quality attribute** - what improves
2. **Scale** - how it’s measured (minutes, %, 1–5 score, etc.)
3. **Current state** (baseline) - today
4. **Target state** - what “better” looks like (pilot target or end-state target)
5. **Stakeholder** - who this is for
6. **Why it matters** - one line (SDG 3, safety, capacity, cost)

Step 3.3 - Keep the list tight

Combine duplicates across stakeholders and finalize 5–9 total objectives.

(Example: “faster response” might matter to both patients and staff - keep one objective and list both stakeholders if needed.)

Step 3: Define Quantified Objectives (Cont.)

Best practices (keep it easy)

- Outcome first: prioritize outcomes (wait time, no-shows, workload) before operational metrics.
- Add 1–2 pilot health metrics: adoption + reliability help interpret results.
- Avoid vanity metrics: messages sent ≠ impact (unless tied to outcomes).
- Pilot targets should be realistic: pilots show direction + evidence, not perfection.
- One objective = one metric: don’t mix multiple ideas into one line.

Example Quantified Objectives

Wait time	Reschedule rate	Staff workload	Patients with access
Baseline: 45 min	Baseline: 12%	Baseline: 6 tasks/day	Baseline: 70%
Target: 30 min	Baseline: 5%	Baseline: 4 tasks/day	Baseline: 85%

Impact

To be filled

Template 3.3: Quantified Objectives Worksheet

#	Quality attribute (what you're improving)	Scale (how you measure it)	Current state (baseline)	Target state (goal)	Stakeholder (who cares most)	Why it matters (impact)

(Field definitions match PB2's objective structure: quality attribute, scale, current state, target state, stakeholder, why it matters.)

Rule: Each quantified objective must have one primary stakeholder (owner of the objective). If another stakeholder also benefits, either:

- Create a separate objective for them (with their own baseline/target)
- Mention them in "Why it matters" (not in the Stakeholder column).

Template 3.3: Filled Example

Use as only an example

#	Quality attribute (what you're improving)	Scale (how you measure it)	Current state (baseline)	Target state (goal)	Stakeholder (who cares most)	Why it matters (impact)
1	Wait time to first clinical contact	Minutes (avg)	45 min	30 min	Patients	Faster access improves care experience (SDG 3 access).
2	Follow-up no-show rate	% missed follow-ups	28%	20%	Clinicians	Improves continuity of care and outcomes.
3	Staff time on routine inquiries	Minutes per shift	60 min/shift	35 min/shift	Reception	Frees capacity for in-person patient support.
4	Handoff reliability (assistant → human)	% successful handoffs	95%	99%	Operations Lead	Prevents dropped requests and workflow breakdowns.
5	Safety escalation correctness	% clinical-intent escalations handled correctly	90%	99%	Clinical Owner	Reduces risk of unsafe advice; preserves trust.

(Field definitions match PB2's objective structure: quality attribute, scale, current state, target state, stakeholder, why it matters.)

Rule: Each quantified objective must have one primary stakeholder (owner of the objective). If another stakeholder also benefits, either:

- Create a separate objective for them (with their own baseline/target)
- Mention them in "Why it matters" (not in the Stakeholder column).

Step 4: Regulatory & Compliance Requirements

Why this comes before solution selection

Compliance is a filter, not a finishing step. If a solution can't meet basic privacy, safety, or hosting requirements, there's no point scoring it—your team will waste time and discover blockers too late.

Think of Step 4 as a simple gate: **Pass** → score solutions | **Fix** → score only if there's an owner + due date | **Block** → stop (adjust scope/safeguards or remove the solution)



UAE - What you must check

You don't need to memorize regulations. You only need to ensure your pilot answers "yes" to these:



Data protection (Federal Data Protection Law / PDPL): You know what data you use, why you use it, and how you secure it (access control, encryption, retention, incident response).



Health data controls: Health data handling and storage/ processing arrangements meet healthcare-sector requirements (especially if health data is involved).



Local regulator expectations (DHA / DOH): If operating in Dubai or Abu Dhabi, confirm any local healthcare policy expectations that affect the pilot (privacy, confidentiality, data exchange).



AI ethics: Define human oversight, safety boundaries, transparency (what the AI can/can't do), and escalation rules.

KSA - What you must check

Same concept: answer "yes" to the essentials:



PDPL: Define lawful processing, clear governance (controller/processor responsibilities), retention, and security controls.



SDAIA: Align to national data governance expectations (especially privacy compliance posture).



NHIC: If interacting with national health information programs/standards, confirm requirements early to avoid integration surprises.



SFDA: If the AI influences clinical decisions (diagnosis/triage/treatment), confirm whether it falls into regulated categories and what route is required.

Template 3.4: Regulatory Compliance Checklist (UAE)

To be filled

Use case / pilot:

Location:

☐ Dubai ☐ Abu Dhabi

☐ Other :

Reviewers:

Privacy/DPO (A.)

IT/Sec (M.)

Clinical

Date:

UAE Gate decision:

☐ Proceed with fixes

☐ Proceed ☐ Stop (Block)

Check (answer at pilot level)	Status (Pass/Fix/Block)	Owner	Due date	Notes
	☐ Pass ☐ Fix ☐ Block			
	☐ Pass ☐ Fix ☐ Block			
	☐ Pass ☐ Fix ☐ Block			
	☐ Pass ☐ Fix ☐ Block			
	☐ Pass ☐ Fix ☐ Block			
	☐ Pass ☐ Fix ☐ Block			
	☐ Pass ☐ Fix ☐ Block			
	☐ Pass ☐ Fix ☐ Block			
	☐ Pass ☐ Fix ☐ Block			

Template 3.4: Filled Example

Use as only an example

Use case / pilot:

Patient Access Assistant (FAQs + appointment support + safe handoff)

Location:

☑ Dubai ☑ Abu Dhabi

☑ Other :

Reviewers:

Privacy/DPO (A.) ·
IT/Sec (M.) ·
Clinical Owner (Dr._____)

Date:

18 Feb 2026

UAE Gate decision:

☑ Proceed with fixes

☐ Proceed ☐ Stop (Block)

Check (answer at pilot level)	Status (Pass/Fix/Block)	Owner	Due date	Notes
Data inventory exists (PHI vs non-PHI is clear)	☐ Pass ☑ Fix ☐ Block	Data Lead	Week 1	Confirm exact PHI fields captured in chat + handoff tickets.
Patient notice/consent approach defined (if needed)	☐ Pass ☑ Fix ☐ Block	Privacy/DPO	Week 1	Finalize Arabic/English notice text + staff script for WhatsApp/QR usage.
Data residency/hosting meets UAE requirements	☑ Pass ☐ Fix ☐ Block	IT Lead	—	UAE hosting confirmed; no storage outside UAE for pilot.
RBAC + encryption + audit logs enabled	☐ Pass ☑ Fix ☐ Block	IT/Sec	Week 1	RBAC roles drafted; audit log retention window pending approval.
Retention + deletion rules defined (pilot data/logs)	☐ Pass ☑ Fix ☐ Block	Privacy + IT	Week 2	Set retention (e.g., 90 days) + deletion process for pilot exports.
Safety boundaries defined (no clinical advice unless approved)	☑ Pass ☐ Fix ☐ Block	Privacy + IT	—	Assistant does not provide diagnosis/ triage; clinical intents escalate to human.
Escalation rules tested (clinical/ sensitive intents → human)	☑ Pass ☐ Fix ☐ Block	Ops Lead	—	End-to-end test completed with 15 scenarios; 100% routed correctly.
Vendor DPA/sub-processors/breach process reviewed (if vendor)	☐ Pass ☑ Fix ☐ Block	Legal + Privacy	Week 2	Sub-processor list requested; breach notification SLA to be added to DPA.
DHA/DOH applicability confirmed (if relevant)	☐ Pass ☑ Fix ☐ Block	Compliance	Week 1	Dubai clinic → DHA policy alignment check; confirm no additional approvals needed for pilot scope.

Template 3.4: Filled Example

Use as only an example

Fix list
(must close before launch or by end of Week 1, as noted):

1. Finalize PHI field inventory + update data flow diagram (Data Lead)
2. Approve patient notice/consent wording (Privacy/DPO)
3. Confirm RBAC roles + audit log retention (IT/Sec)
4. Complete vendor DPA items (Legal/Privacy)
5. Confirm DHA alignment for pilot scope (Compliance)

Template 3.5: Regulatory Compliance Checklist (KSA)

To be filled

Use case / pilot:

Location:

☐ Dubai ☐ Abu Dhabi

☐ Other :

Reviewers:

Privacy

IT/Sec

Clinical

Date:

18 Feb 2026

KSA Gate decision:

☐ Proceed with fixes

☐ Proceed ☐ Stop (Block)

Check (answer at pilot level)	Status (Pass/Fix/Block)	Owner	Due date	Notes
Data inventory exists (personal/health data is clear)	☐ Pass ☐ Fix ☐ Block			
PDPL governance defined (roles/responsibilities)	☐ Pass ☐ Fix ☐ Block			
Notice/consent approach defined (where required)	☐ Pass ☐ Fix ☐ Block			
RBAC + encryption + audit logs enabled	☐ Pass ☐ Fix ☐ Block			
Retention + deletion rules defined	☐ Pass ☐ Fix ☐ Block			
Safety boundaries + escalation rules defined and tested	☐ Pass ☐ Fix ☐ Block			
SFDA applicability checked (if clinical decision impact)	☐ Pass ☐ Fix ☐ Block			
Vendor DPA/sub-processors/ breach process reviewed (if vendor)	☐ Pass ☐ Fix ☐ Block			
NHIC interoperability/data requirements checked (if relevant)	☐ Pass ☐ Fix ☐ Block			

Key regulations (UAE + KSA) - quick reference

Country	Regulation / Policy	What it's relevant for	Reference
UAE	Federal Decree-Law No. 45 of 2021 (PDPL)	Personal data processing governance (rights, obligations, controls)	Legal Advice Middle East
UAE	Federal Law No. 2 of 2019 (ICT in Health Fields / Health Data Law)	Health data handling + restrictions on storing/ processing outside UAE without approval	UAE Legislation
UAE	DHA Health Data Protection & Confidentiality Policy	Dubai-specific health data requirements (DHA entities)	Dubai Health Authority
UAE	DHA Artificial Intelligence Policy	Governance expectations for AI use in Dubai healthcare contexts	Dubai Health Authority
UAE	DoH Responsible AI Standard (Abu Dhabi)	Minimum requirements for responsible AI across Abu Dhabi healthcare ecosystem	Department of Health Abu Dhabi
UAE	Abu Dhabi Patient Healthcare Data Privacy Standard (DoH)	Abu Dhabi healthcare data privacy/security obligations	cdn.clydeco.com
UAE	UAE AI Ethics: Principles & Guidelines (ai.gov.ae PDF)	National responsible AI principles	ai.gov.ae

Key regulations (Cont.)

Country	Regulation / Policy	What it's relevant for	Reference
KSA	Personal Data Protection Law (PDPL)	Core privacy obligations for personal data in KSA	SDAIA
KSA	PDPL Implementing Regulations	Operational requirements for PDPL compliance	SDAIA
KSA	Regulation on Personal Data Transfer Outside the Kingdom + SCCs	Cross-border transfer rules and safeguards	SDAIA
KSA	SDAIA AI Ethics Principles	Ethical AI governance framework	DGA
KSA	NHIC Saudi Health Information Exchange Policies	Consent/access control expectations for health information exchange	nhic.gov.sa
KSA	SFDA Guidance on AI/ML-enabled Medical Devices (MDS-G010)	When AI/ML is regulated as a medical device + expectations	والدواء للغذاء العامة الهيئة
KSA	SFDA Guidance on SaMD (MDS-G23)	Software as a medical device regulatory expectations	والدواء للغذاء العامة الهيئة

Step 5: Define AI Solutions

Why this step matters

You can't compare solutions fairly if each vendor describes their product differently. This step forces every option into the same format so you can evaluate "apples-to-apples."

Keep it lightweight: **2–3 solutions** is ideal. More than 4 usually slows the process without improving the decision

Instructions (how to document solutions fast)

For each candidate solution:

1. Write a **one-line description** (what it does).
2. Capture the **technical approach** at a high level (no deep engineering).
3. Break the solution into **4–6 features** that will actually be piloted.
4. Record dependencies (data, integrations, people, approvals).
5. Add cost + timeline ranges (good enough for comparison).

Rule of thumb: If a feature cannot be tested in a 12-week pilot, it doesn't belong in the feature list.

Step 5: Define AI Solutions (Cont.)

Feature breakdown approach (easy method)

Use one of these feature types:

- **User-facing** (chat, portal, workflow UI)
- **Automation** (routing, reminders, classification, summaries)
- **Integration** (EHR, scheduling, CRM, call center)
- **Safety** & **governance** (audit logs, escalation rules, access control)
- **Measurement** (dashboards, reporting, analytics)

Aim for 4–6 features per solution so impact estimation stays manageable.

Key questions to ask solution providers

Fit & scope

What exact workflows does this support in healthcare settings like ours?
What is explicitly out of scope?

Data & integration

What systems do you integrate with (EHR/scheduling/CRM)? How (API, HL7/FHIR, export)?
What data fields do you require to work well?

Step 5: Define AI Solutions (Cont.)

Key questions to ask solution providers

Privacy, security, compliance

Where is data stored and processed (UAE/KSA residency options)?
Do you provide audit logs, role-based access, encryption, retention controls?
Do you use customer data to train models? If yes, how is it controlled?

Pilot delivery

What can be piloted in 12 weeks, and what evidence can we expect by week 2/4/6?
What does your support model look like during pilot (SLAs, incident process)?

Credibility

Do you have healthcare case studies with measurable outcomes?
What assumptions are required for those outcomes to repeat in our setting?

Guidance: Build vs. Buy

Buy when you need speed, proven patterns, and predictable delivery.
Build when you have unique workflows, strong internal engineering, or strict control needs.

Simple decision test

- If you need results in **weeks**, start with **buy** (pilot fast).
- If you need a differentiated product capability over **months**, consider **build** (but only with strong data readiness and resourcing).

Template 3.6: AI Solution Definition Form

To be filled

Field	Solution details
Solution name	
Type (Buy / Build / Hybrid)	
One-line description	
Technical approach (high level)	
Target users	
Key features (4–6)	

Template 3.6: AI Solution Definition Form (Cont.)

To be filled

Field	Solution details
Dependencies	
Compliance notes	
Estimated cost	
Estimated timeline	
Vendor evidence	

Use as only an example

Template 3.6: Filled Example 1

Solution A - Patient Access Assistant (Buy)

Field	Solution details
Solution name	Patient Access Assistant
Type (Buy / Build / Hybrid)	Buy
One-line description	Handles patient FAQs and appointment support with safe handoff to staff.
Technical approach (high level)	LLM-based conversational assistant + curated knowledge base + routing rules + audit logging.
Target users	Patients; reception/call center staff (handoff recipients).
Key features (4–6)	1. Multilingual patient FAQ (clinic info, prep instructions). 2. Appointment request capture + routing to staff queue. 3. Reminder and reschedule support (semi-manual → partial integration). 4. Safety boundaries + mandatory escalation for clinical intents. 5. Admin dashboard (content updates, routing rules). 6. Audit logs + reporting for usage and handoffs.

Use as only an example

Template 3.6: Filled Example 1 (Cont.)

Solution A - Patient Access Assistant (Buy)

Field	Solution details
Dependencies	Access to clinic FAQs; staff queue for handoff; (optional) scheduling integration; approval from privacy/clinical owner.
Compliance notes	UAE hosting available; RBAC + audit logs; patient notice text required.
Estimated cost	Pilot: $20k–$40k (example range) + support; Ongoing: per-site/per-user subscription.
Estimated timeline	2–3 weeks to launch pilot v1 (semi-manual), 6–8 weeks with partial integration.
Vendor evidence	Healthcare customer references; examples of deflection + reduced admin workload.

Use as only an example

Template 3.6: Filled Example 2

Solution B - No-Show Reduction Automation (Build/Hybrid)

Field	Solution details
Solution name	No-Show Prediction + Outreach Automation
Type (Buy / Build / Hybrid)	Build / Hybrid
One-line description	Predicts no-show risk and automates reminders and rescheduling outreach.
Technical approach (high level)	ML model on appointment history + rules engine + messaging workflows (SMS/WhatsApp/email).
Target users	Scheduling team; patients receiving reminders and rescheduling links.
Key features (4–6)	1. No-show risk scoring per appointment. 2. Automated reminder sequences by risk tier. 3. One-click rescheduling flow (link or assisted). 4. Dashboard for risk trends and campaign performance. 5. A/B testing of reminder scripts and timing. 6. Audit trail for message delivery and opt-outs.

Use as only an example

Template 3.6: Filled Example 2 (Cont.)

Solution B - No-Show Reduction Automation (Build/Hybrid)

Field	Solution details
Dependencies	Historical appointment data quality; messaging platform; scheduling system integration; opt-in/notice management.
Compliance notes	Must align to privacy rules for messaging and retention; ensure opt-out controls.
Estimated cost	Build: internal team time + tooling; Pilot: $10k–$30k equivalent; Ongoing: maintenance + messaging costs.
Estimated timeline	4–6 weeks to MVP if data is clean; 8–12 weeks for stable integration and reporting.
Vendor evidence	Internal baseline analysis required; external benchmarks possible but context-dependent.

Step 6: Impact Estimation

What this step does

This is where you turn “vendor claims” into a clear, comparable estimate of impact.

You estimate impact at the **feature level** (Feature → Objective), then add a **credibility score** (0–10) so everyone can see:

- what you believe will improve outcomes, and
- how confident you are (based on evidence, not optimism).

How to estimate impact (feature-level)

Impact % = how much of the objective “gap” this feature is expected to close.

- If the metric should go **down** (e.g., wait time):
  Gap = Baseline − Target
- If the metric should go **up** (e.g., satisfaction):
  Gap = Target − Baseline

Then for each feature, estimate: “If we implement this feature in the pilot, what **% of that gap** will it close?”

Keep it practical: It’s okay if the total doesn’t reach 100% - many pilots don’t fully hit targets in 12 weeks.

Who should do the impact estimation (best setup)

Do it as a short working session with 4–6 people:

- Ops/Process Owner (knows workflow reality)
- Frontline representative (knows daily friction and adoption risks)
- Data/Analytics (knows measurement feasibility + baselines)
- IT/Integration (knows constraints and effort)
- Clinical/Compliance (safety boundaries)
- Vendor (optional) (provides evidence, not final numbers)

Rule: Vendor can propose estimates, but your team owns the final numbers and credibility scores.

Step 6: Impact Estimation (Cont.)

Step-by-step (do this for each solution separately)

Repeat Steps 1–6 once per candidate solution (Solution A, then B, then C). Each solution gets its own IET.

1. Choose your **5–9 quantified objectives** (Template 3.3)
2. List that solution's **4–6 features** (Template 3.6)
3. Fill each cell with Impact % (best estimate)
4. Add Credibility (0–10) for each estimate
5. Check for overlap between features (identify main drivers vs extras)
6. Add 1-line assumptions for low-credibility estimates

Then use the completed IETs to compare solutions in **Step 7** (Template 3.8).

Credibility scale (0–10) - with healthcare examples

Credibility score	What it means (simple)	Healthcare example
0–2	Pure guess / no evidence	"Vendor says it will reduce wait time" (no data, no similar deployments).
3–4	Plausible, but unproven in your setting	Worked in another country/clinic type; no proof it transfers.
5–6	Some evidence, partial match	Small case study + pilot-like constraints are similar, but data/workflow differs.
7–8	Strong evidence, close match	Measured results in similar clinics + clear explanation of how it works.
9–10	Proven repeatedly in near-identical context	Your org (or same system/regulator context) has already shown the metric movement.

Template 3.7: Impact Estimation Table (IET)

To be filled

How to fill cells: write **Impact % / Credibility** (example: 20% / 6)
(Impact % = % of the objective gap closed by that feature.)

Feature (from Template 3.6)	**Obj #3** Impact %	**Obj #3** Cred. (0–10)	**Obj #4** Impact %	**Obj #4** Cred. (0–10)	Notes / assumptions

Use as only an example

Template 3.7: Filled Example

2 objectives × all features - Solution A (Patient Access Assistant)

Objectives shown (from Template 3.3):

Obj #3: Staff time on routine inquiries (60 → 35 min/shift) | **Obj #4**: Handoff reliability (95% → 99%)

Feature (from Template 3.6)	Obj #3 Impact %	Obj #3 Cred. (0–10)	Obj #4 Impact %	Obj #4 Cred. (0–10)	Notes / assumptions
F1 | Multilingual FAQ	35%	7	0%	9	Assumes content covers top intents and is updated weekly.
F2 | Appointment request capture + routing	20%	6	40%	6	Assumes staff queue exists + response SLA agreed.
F3 | Reminder / reschedule support	5%	4	0%	8	Impact depends on integration readiness; pilot may be semi-manual.
F4 | Safety boundaries + escalation	0%	9	45%	8	Requires clear escalation triggers + trained staff handoff.
F5 | Admin dashboard (content + routing rules)	10%	6	5%	5	Requires an assigned owner to maintain content/rules
F6 | Audit logs + reporting	5%	5	10%	6	Assumes logs are reviewed weekly and issues are acted on.

Quick check (sanity): Obj #3 total estimated gap closed: **75%** (pilot likely shows strong trend, not perfect target hit) | Obj #4 total estimated gap closed: **100%** (high reliance on routing + escalation discipline)

Step 6: Impact Estimation (Cont.)

Guidance box: Common pitfalls (avoid these)

Over-confidence

Warning: High credibility everywhere, little proof.
Fix: Ask “What evidence do we have?” If weak → lower credibility.

Overlapping features (not “double counting”)

Warning: Multiple features claim impact on the same objective.
Fix: Identify the main driver & avoid paying for features that add little extra value.

Ignoring constraints

Warning: Assumes perfect data, integration, and adoption.
Fix: Write assumptions and lower credibility if constraints aren’t solved.

Activity ≠ impact

Warning: “More chats” treated as success.
Fix: Tie to outcomes (time saved, no-shows, satisfaction).

No plan to raise credibility

Warning: Low credibility scores with no next step.
Fix: Create a short evidence plan (sandbox test, vendor proof, data check, 1–2 week proof increment).

Simple rule of thumb (easy to remember)

- High impact + high credibility → **pilot** for outcomes
- High impact + low credibility → pilot as a **proof sprint** (learn fast)
- Low impact + high cost → usually **skip** unless mandatory

Step 7: Compare Solutions

Why this step matters

The IET (Template 3.7) gives you detailed estimates. Step 7 turns that detail into a one-page comparison so you can make trade-offs visible and discussable.

This step answers:

- Which solution has the highest expected impact?
- Which is the best value for money?
- Which is the most credible (least risky)?
- Which fits our constraints and strategy?

How to interpret the IET

From the IET, summarize each solution into 3 simple outputs:

01 Total impact (expected).
A combined view of how much the solution is expected to move your objectives.

→

02 Credibility (confidence).
(Use the credibility score from the IET (Template 3.7) — see Step 6.)

→

03 Cost.
Pilot + ongoing (high-level ranges are fine).

Then calculate Impact-to-Cost as a simple value signal.

Tip : Don’t hide trade-offs in one “magic score.” Show the dimensions clearly.

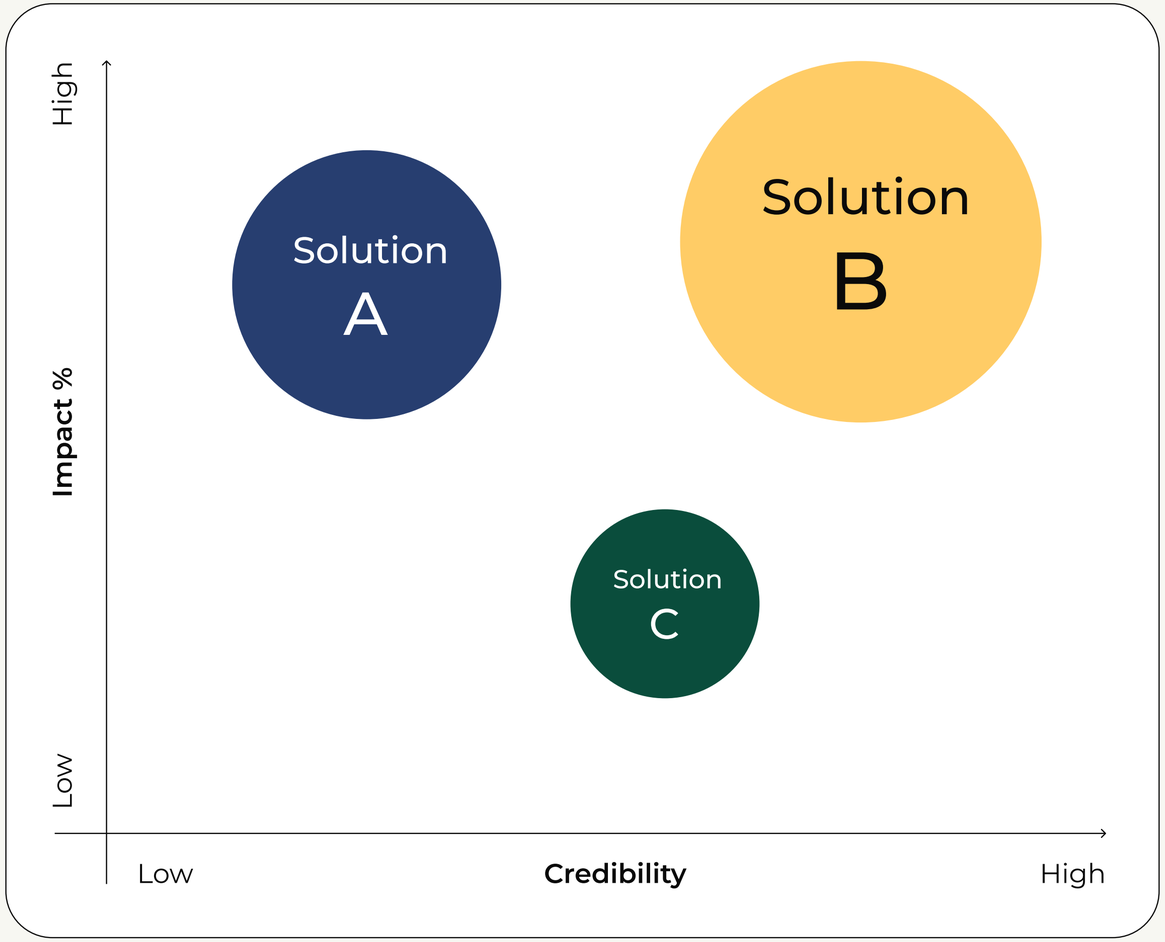
Step 7: Compare Solutions (Cont.)

Instructions (multi-dimensional comparison)

1. Choose 2–3 candidate solutions (from Template 3.6).
2. For each solution, summarize:
   - **Impact score** (from IET totals)
   - **Credibility score** (average or weighted average)
   - **Pilot cost range**
   - **Time to pilot readiness**
3. Compute a **value signal**: Impact ÷ Cost (simple ratio).
4. Add a short “fit notes” line (strategic fit + constraints).

Analysis guidance: what to look for

1. **Highest impact:**
   strongest movement on your top objectives
2. **Best value:**
   good impact at lower cost and effort
3. **Most certain:**
   high credibility, fewer dependencies, proven in similar context
4. **Strategic fit:**
   aligns with mission/SDG 3, workflows, and compliance constraints
5. **Hidden risk:**
   low credibility + heavy integration + unclear data readiness



Template 3.8: Solution Comparison Dashboard

To be filled

Comparison dimension	Solution A	Solution B	Solution C
Solution name			
Type (Buy / Build / Hybrid)			
Compliance status (Pass / Fix / Block)			
Total Impact Score (0–100) (from IET)			
Credibility (0–10) (avg / weighted avg from IET)			
Credibility-Adjusted Impact (Impact × Cred/10)			
Pilot cost (range)			
Pilot cost midpoint			
Impact/Cost Ratio (Adj Impact per $10k)			
Time to pilot-ready			
Key dependencies (1 line)			
Fit note (1 line)			

How to fill Template 3.8

(Simple instructions)

Step 1 - Pull totals from the IET (Template 3.7)

You need two numbers per solution:

A) **Total Impact Score** (0–100)

- For each objective, sum feature impact % across the row (cap at 100%).
- If you use objective weights (recommended), compute: **Total Impact Score = Σ [Objective Weight × Objective Impact%] / 100**
  (If you don't weight objectives: average the objective impacts instead.)

B) **Credibility** (0–10) (use the credibility score from the IET (Template 3.7) - see Step 6.)

Step 2 - Convert cost range into a midpoint

Pilot cost midpoint = (Low + High) / 2 | Example: $20–40k → midpoint = $30k

Step 3 - Compute credibility-adjusted impact

Credibility-Adjusted Impact = Total Impact Score × (Credibility / 10) | Example: Impact 70, Cred 7.0 → 70 × 0.7 = 49

How to fill Template 3.8 (Cont.)

(Simple instructions)

Step 4 - Compute the numeric value-for-money ratio

Use “points per $10k” so the number is readable:
Impact/Cost Ratio (per $10k) = (Credibility-Adjusted Impact ÷ Pilot cost midpoint) × 10,000

Example: Adj Impact 49, cost midpoint $30,000 → (49 / 30,000) × 10,000 = **16.3**

Step 5 - Add only the essentials

Fill the remaining fields briefly:
Time to pilot-ready (weeks) | 1-line dependencies | 1-line fit note

Quick decision hint

- High Impact/Cost + High Credibility → best pilot candidate
- High Impact/Cost + Low Credibility → pilot as a proof sprint (tight scope to raise credibility fast)
- Low Impact/Cost + High Cost → usually not worth piloting unless strategic reasons exist

Use as only an example

Template 3.8: Filled Example

Comparison for 3 solutions (healthcare access use case)

Comparison dimension	Solution A	Solution B	Solution C
Total Impact Score (0–100) (from IET)	70	65	60
Credibility (0–10) (avg / weighted avg from IET)	7.0	5.0	6.5
Credibility-Adjusted Impact (Impact × Cred/10)	49.0	32.5	39.0
Pilot cost (range)	20–40k	10–30k	30–60k
Pilot cost midpoint	30k	20k	45k
Impact/Cost Ratio (Adj Impact per $10k)	**16.3**	**16.3**	**8.7**

Comparison considerations

Before you finalize the dashboard, confirm:

- All solutions are based on the same objectives (Template 3.3)
- Feature sets are comparable (Template 3.6)
- Compliance gate is addressed (Templates 3.4–3.5)
- Credibility reflects real evidence (not vendor optimism)
- Value signal considers both cost and time-to-pilot
- Notes reflect constraints (data readiness, integrations, staffing)
- You can explain the result in 2 minutes to leadership

Step 8: Make the Decision

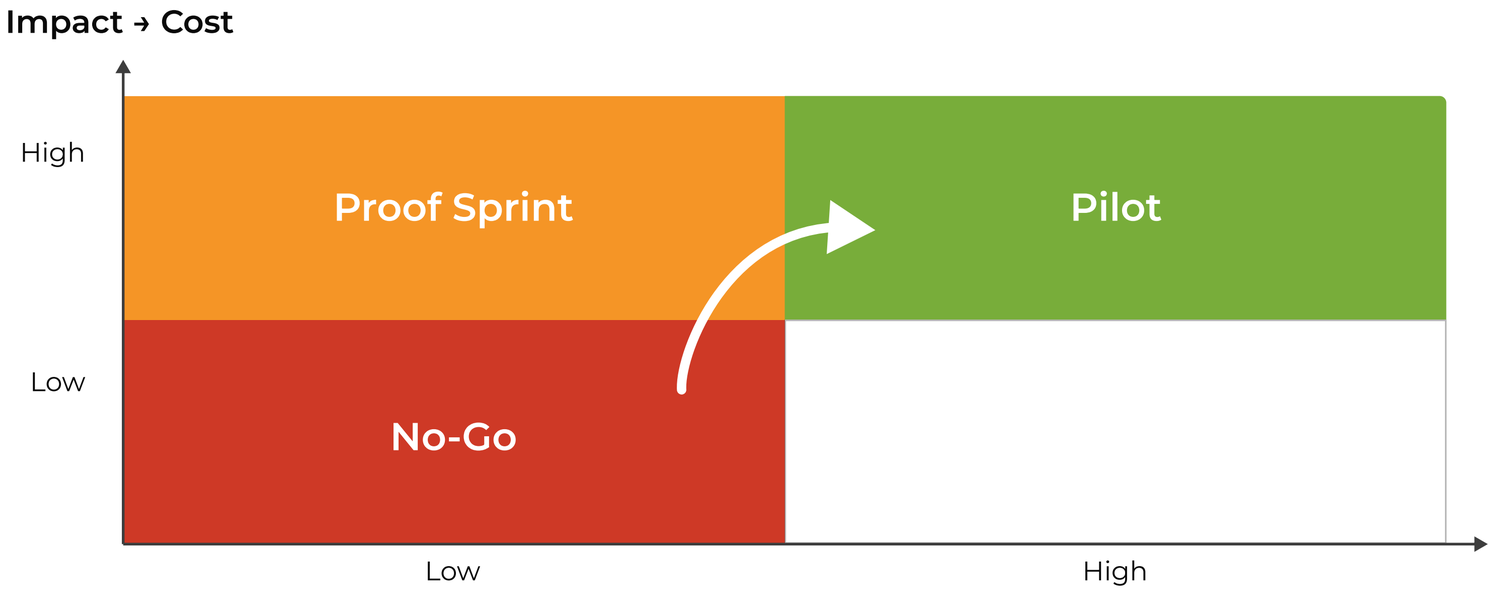
Why this step matters

Step 7 gives you the comparison dashboard. Step 8 makes the final call using the core PB2 decision lens. This avoids the common failure mode of choosing the “best story” instead of the best evidence + value option.

Use the numeric Impact/Cost ratio already calculated in Step 7 (Template 3.8).

Impact × Credibility − Cost

(High impact is good, but not if credibility is low and cost is high)



Step 8: Make the Decision (Cont.)

The PB2 decision method

Step 8.1	Step 8.2	Step 8.3
Apply gates first (non-negotiable)	**Compare on the 3 core dimensions**	**Make the decision using the "Credibility × Cost" risk check**
• Compliance gate: If any solution is Block → it cannot be selected. • Safety boundary gate: If the solution can't operate within required safety boundaries → it cannot be selected.	From Template 3.8, for each solution capture: • Total Impact (from Template 3.7 IET summary) • Cost (pilot + ongoing estimate) • Credibility (0–10 confidence)	Use this simple rule: • **Select for pilot** when: Impact-to-Cost is strong AND Credibility is acceptable (or improving) • **Pilot as a "proof sprint"** when: Impact looks strong but Credibility is low (pilot is designed to raise credibility fast with tight scope) • **Do not proceed** when: Cost is high AND Credibility is low (too risky to justify testing unless scope/cost can be reduced)

Step 8: Make the Decision (Cont.)

Key considerations (secondary checks - keep it short)

After the core decision, run these quick checks:

Time-to-pilot	Can you be live in weeks (not months)?
Integration readiness	Can you start semi-manual if integration is slow?
Workflow fit/adoption	Will frontline teams actually use it?
Vendor transparency	Audit logs, error visibility, support SLAs
Measurement feasibility	Can you measure outcomes weekly?

If any of these are “red,” the decision becomes **PIVOT** the pilot design or **do not proceed**.

To be filled

Template 3.9: Decision Framework & Recommendation Report

Decision Summary **Use case:** _____ **Date:** _____ **Sponsor:** _____

A) Decision table (core PB2 lens)

Solution	Compliance (Pass/Fix/ Block)	Total impact (IET summary)	Credibility (0–10)	Pilot cost (range	Impact/Cost (High/Med/Low)	Decision (Pilot / Proof sprint / No-go)
A						
B						
C						

Top risks (max 3) + mitigation

Risk	Mitigation	Owner	Due

Final decision

☐ Pilot ☐ Proof sprint ☐ No-go

Next step:

☐ Section 4 pilot roadmap ☐ Evidence sprint ☐ Stop

B) Rationale (plain language — 3 bullets)

Pilot intent (if piloting)

☐ Value pilot (credibility already acceptable; focus on outcomes)

☐ Proof sprint (credibility low; focus on validating assumptions fast)

Example (how the logic looks):

Solution X: High impact, low credibility, medium cost → Proof sprint pilot (tight scope to validate)

Solution Y: Medium impact, high credibility, low cost → Pilot (best value-for-money)

Solution Z: Medium impact, low credibility, high cost → No-go (risk too high)

Exercise 3.2: Decision Workshop Guide

Goal: Make a final, documented decision. | **Participants:** Sponsor + Ops owner + IT + Compliance + Analyst + PM (5–7 people)

01 **05 min** | Confirm evaluation inputs and compliance gate status (Step 4).

02 **10 min** | Review comparison dashboard (Template 3.8).

03 **10 min** | Agree criteria + weights (keep 5–8 criteria).

04 **15 min** | Score solutions (1–5) and discuss top disagreements only.

05 **10 min** | Final recommendation + rationale + top 3 risks/mitigations.

06 **05 min** | Confirm next step: Pilot roadmap design (Section 4) or stop.

Output: Completed Template 3.9 + a clear “Pilot / Don’t proceed” decision.

Section 3 Outputs

What you should have by the end of Section 3

1	Template 3.1 completed: Organization & challenge context (from PB1) clearly documented.
2	Template 3.2 completed: Stakeholder map with influence levels and key concerns (top stakeholders identified).
3	Template 3.3 completed: 5–9 quantified objectives (baseline → target, measurable scales).
4	Templates 3.4 / 3.5 completed: Compliance checklist (UAE and/or KSA) with Pass/Fix/Block status and assigned owners.
5	Template 3.6 completed: 2–3 candidate solutions documented in a comparable format (4–6 features each).
6	Template 3.7 completed: Impact Estimation Table (IET) with impact % + credibility scoring.
7	Template 3.8 completed: One-page solution comparison dashboard (trade-offs visible).
8	Template 3.9 completed: Decision framework + recommendation report (or “do not proceed / evidence sprint” decision) with rationale, risks, and objections captured.

Introduction to Evolutionary Piloting

Most AI pilots fail for one simple reason: teams try to “prove everything at once.” They run a big pilot with too many features, too many dependencies, and unclear success measures—then they can’t tell what worked, what didn’t, or what to fix.

This playbook uses Evolutionary (Evo) Piloting to make pilots easy to run and easy to stop if they’re not working.

What Evo means (in plain terms)

Evo is a way to pilot AI by delivering value in small, testable steps. Each step produces something usable, measures results, and ends with a decision: GO / PIVOT / NO-GO.

Gilb’s Evo principles

- **Deliver in small increments** — don’t wait 3 months to learn if it works.
- **Measure real outcomes** — success is proven with metrics, not opinions.
- **Get feedback early** — frontline users help shape what’s next.
- **Decide based on evidence** — continue only if results justify it.

What “increment” means:

An increment is a short, time-boxed pilot cycle (usually 1–2 weeks) where you:

- deliver a small piece of functionality (or a workflow improvement),
- test it with real users,
- measure results against your success criteria, and
- make a decision at the end: GO / PIVOT / NO-GO.

Example: “Increment 1 = launch top 20 FAQs + safe handoff, measure adoption and safety, then decide whether to expand.”

Introduction to Evolutionary Piloting (Cont.)

Why 2-week increments (practical reasons)

Two weeks is long enough to deliver something meaningful, and short enough to:

- Keep momentum
- Limit risk
- Spot adoption issues early
- Avoid “pilot drift” (where nothing gets finished).

If your team is smaller or approvals are slow, you can use **1-week increments**—the rule is: keep increments short and measurable.

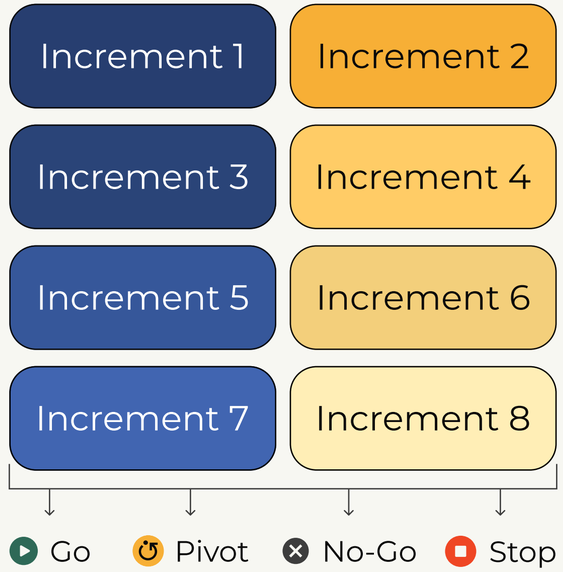
Pilot vs. “big bang” rollout (quick comparison)

Approach	What it looks like	What usually happens
Big bang	Build/integrate everything, then launch	Late surprises, sunk costs, hard to stop
Evo pilot	Launch small, measure, improve in increments	Faster learning, lower risk, easier decisions

Big Bang



Pilot (incremental)



Step 1 : Pilot Scope & Success Criteria

Why this step matters (simple)

A pilot succeeds when it is small enough to run, but real enough to prove value. This step sets clear boundaries so you don't pilot "everything" and end up with unclear results.

How to set scope

Use these 3 rules to keep scope tight:

01 One use case, one primary workflow

Example: "patient appointment questions + handoff" (not "all patient communications").

02 One pilot area (small but real)

Example: 1 clinic + 1 call-center pod (not the whole network).

03 12 weeks max, with 1-2 week increments

This playbook uses a 12-week structure because it's long enough to gather evidence but short enough to stop if needed.

How to define success criteria

Your success criteria should come directly from Template 3.3 (Quantified Objectives). Pick:

2–3 outcome metrics	1–2 safety/quality guardrails	1 adoption metric
(value)	(risk control)	(real usage)

Good success criteria are: measurable, time-bound, and have a clear threshold that triggers GO/PIVOT/NO-GO.

Step 1 : Pilot Scope & Success Criteria (Cont.)

Best practices (keep adoption easy)

Start small	If you can’t run it with minimal disruption, it’s too big.
Be representative	Choose a pilot area that reflects real conditions (not the “perfect clinic”).
Define what is NOT included	Scope clarity is a success factor.
Vendor transparency	Start semi-manual if needed; integrate later if evidence supports scaling.

Template 4.1: Pilot Scope Definition

To be filled

Pilot name: _____

Selected solution (from Section 3): _____

Pilot objective (1 sentence): _____

Pilot location / unit: _____

Pilot duration: _____ weeks (recommended: 6–12 weeks)

Increment length: ☐ 1 week ☐ 2 weeks

Users included: _____

Beneficiaries included: _____

In-scope (what we will pilot):

Out-of-scope (what we will NOT pilot):

Template 4.1: Pilot Scope Definition (Cont.)

To be filled

Channels included: ☐ Web ☐ WhatsApp ☐ Phone ☐ Email ☐ Other: _____

Languages: ☐ English ☐ Arabic ☐ Other: _____

In-scope (what we will pilot):

- Data: _____
- Integration: _____
- People/training: _____
- Compliance approvals: _____

Owner: _____

Sponsor: _____

Template 4.1: Filled Example

Use as only an example

Pilot name: Patient Access Assistant - Clinic A Pilot

Selected solution (from Section 3): Solution A - Patient Access Assistant (Buy)

Pilot objective (1 sentence): Reduce staff time spent on routine patient inquiries and improve patient experience for appointment-related support, while maintaining safe escalation to humans.

Pilot location / unit: Clinic A (front desk) + Call Center Pod 2

Pilot duration: 12 weeks

Increment length: ☐ 1 week ☑ 2 weeks

Users included: Reception staff (Clinic A), Call Center Pod 2, Pilot PM

Beneficiaries included: Patients contacting Clinic A (English/Arabic)

In-scope (what we will pilot):

- FAQ responses for appointment prep, clinic hours, location, required documents
- Appointment request capture + routing to staff queue (handoff)
- Safe escalation rules (clinical intents → immediate human handoff)
- Weekly content updates + basic reporting (usage, handoffs, escalation reasons)

Out-of-scope (what we will NOT pilot):

- Clinical advice, diagnosis, medication instructions
- Full EHR write-back automation
- Multi-clinic rollout beyond Clinic A during pilot
- Advanced personalization based on medical history

Template 4.1: Filled Example (Cont.)

Use as only an example

Channels included: ☑ Web ☑ WhatsApp ☐ Phone ☐ Email ☐ Other: _____

Languages: ☑ English ☑ Arabic ☐ Other: _____

In-scope (what we will pilot):

- **Data:** Approved FAQ content pack (top 50 intents) + clinic policy answers
- **Integration:** Staff handoff queue defined (ticket form or inbox); scheduling integration optional
- **People/training:** 30-minute training for reception + call center; 2 champions assigned
- **Compliance approvals:** Patient notice text approved; DPA clauses reviewed; audit log retention defined

Owner: Project Manager

Sponsor: Program Director, Outpatient Services

Template 4.2: Pilot Success Criteria Worksheet

To be filled

Pick **4–6 criteria** total.
(2–3 value outcomes + 1–2 safety guardrails + 1 adoption metric)

Metric type	Metric (from Template 3.3)	Baseline	Target / Threshold	How it is measured	Decision rule
Outcome (value)					
Outcome (value)					
Outcome (value)					
Guardrail (safety/quality)					
Guardrail (safety/quality)					
Adoption					

Template 4.2: Filled Example

Use as only an example

Chosen criteria (6 total): 3 outcomes + 2 guardrails + 1 adoption

Metric type	Metric (from Template 3.3)	Baseline	Target / Threshold	How it is measured	Decision rule
Outcome (value)	Staff time on routine inquiries (min/shift)	60	≤ 45 by Week 6; ≤ 35 by Week 12	Weekly time sampling + staff self-report log	**GO** if trend improving and Week 6 target met; **PIVOT** if flat; **NO-GO** if worse
Outcome (value)	Patient satisfaction (access touchpoints) (1–5)	3.6	≥ 4.0 by Week 6; ≥ 4.2 by Week 12	Short post-interaction survey	**GO** if ≥ target and no safety issues; **PIVOT** if < target but improving
Outcome (value)	Follow-up no-show rate (%)	28%	≤ 24% by Week 12 (directional pilot goal)	Appointment system report	**GO** if downward trend + plausible linkage; **PIVOT** if no change
Guardrail (safety/quality)	Escalation/handoff success rate (%)	95%	≥ 98% by Week 6; ≥ 99% by Week 12	Audit logs + ticket completion checks	**NO-GO** if < 95% any week; **PIVOT** if 95–98% with fixes
Guardrail (safety/quality)	Incorrect/unsafe responses (count/week)	Unknown	≤ 2/week by Week 4; 0 critical incidents	Weekly QA review + incident log	**NO-GO** if any critical unsafe response; **PIVOT** if repeated minor errors
Adoption	% of patient inquiries handled by assistant before human	0%	≥ 15% by Week 4; ≥ 25% by Week 8; ≥ 35% by Week 12	Channel analytics + handoff logs	**GO** if adoption meets thresholds; **PIVOT** if low adoption due to UX/content

Notes (for the team):

- Outcomes are measured weekly; guardrails are reviewed weekly with immediate fixes.
- “No-show rate” is included as a directional metric (it may require more time to fully prove causality).
- Any safety-related breach triggers an immediate escalation and possible pause.

Step 2 : Breaking Down into Increments

Why increments

Instead of launching a full pilot at once, you run small 2-week chunks. Each chunk delivers something usable and ends with a decision: **GO / PIVOT / NO-GO.**

This keeps the pilot easier to manage, safer, and faster to learn from.

How to design 2-week increments

Use these 4 rules:

01 One increment = one clear outcome

Example: “FAQ works for top 20 intents”
(not “improve everything”).

02 Deliver something usable every 2 weeks

Even if it’s imperfect, it must be testable with real users.

03 Include a decision rule

Decide in advance what “GO” means
(minimum evidence threshold).

04 Keep increments small enough to finish

If it won’t fit into 2 weeks, split it.

Guidance: Feature prioritization (simple MoSCoW)

Use MoSCoW to decide what goes first:

- **M** Must have = required for safety or core value
- **S** Should have = important but not required for first proof
- **C** Could have = nice-to-have
- **W** Won’t have (this pilot) = explicitly out of scope

Shortcut: start with Must-have + one quick win.

To be filled

Template 4.3: Increment Planning Worksheet

Increment _ – _____ **Duration:** ☐ 1 week ☑ 2 weeks (Dates: _____ to _____)

Scope (what we will build/test):

_____

_____

Value delivered (what improves + for whom):

_____

_____

(Example: reduces staff workload / improves patient experience / improves safety)

Investment (people/effort)

Owner: _____

Team involved: _____

Effort estimate: ☐ Low ☐ Medium ☐ High

Dependencies: _____

Decision rule (choose one):

GO if: _____

PIVOT if: _____

NO-GO if: _____

Success criteria (from Template 4.2):

_____

_____

Key questions this increment must answer:

_____

_____

Notes / risks: _____

Template 4.3: Increment Planning Worksheet

Use as only an example

Increment 1 — Safe FAQ MVP **Duration:** ☐ 1 week ☑ 2 weeks (Dates: _____ to _____)

Scope (what we will build/test):

Top 20 FAQs (English/Arabic),
safety boundaries, basic handoff to staff inbox

Value delivered (what improves + for whom):

Reduces repetitive questions;
proves safe behavior in real usage

(Example: reduces staff workload / improves patient experience / improves safety)

Investment (people/effort)

Owner: _____

Team involved: PM + 1 content owner + vendor setup

Effort estimate: ☐ Low ☐ Medium ☐ High

Dependencies: _____

Decision rule (choose one):

- **GO if:** if no critical safety incidents + staff accept handoffs
- **PIVOT if:** adoption is low due to content gaps
- **NO-GO if:** if any safety breach

Success criteria (from Template 4.2):

Adoption trend toward ≥15% by Week 4; unsafe responses ≤2/week; 0 critical incidents

Key questions this increment must answer:

Will patients use it? Are answers accurate? Is escalation safe and consistent?

Notes / risks: _____

Template 4.3: Increment Planning Worksheet

Use as only an example

Increment 2 — Handoff Reliability **Duration:** ☐ 1 week ☑ 2 weeks (Dates: _____ to _____)

Scope (what we will build/test):

Structured intake + routing rules, handoff SLA, audit logging for handoffs

Value delivered (what improves + for whom):

Reliable routing; fewer back-and-forth messages; faster resolution

(Example: reduces staff workload / improves patient experience / improves safety)

Investment (people/effort)

Owner: _____

Team involved: PM + Ops lead + light IT support

Effort estimate: ☐ Low ☐ Medium ☐ High

Dependencies: _____

Decision rule (choose one):

- **GO if:** if reliability ≥98%
- **PIVOT if:** if routing errors repeat
- **NO-GO if:** if handoffs repeatedly fail

Success criteria (from Template 4.2):

Handoff success ≥98% by Week 6; SLA met for response/triage

Key questions this increment must answer:

Do handoffs reach the right team quickly? Are routing errors manageable?

Notes / risks: _____

Template 4.3: Increment Planning Worksheet

Use as only an example

Increment 3 — Expand Coverage + Weekly Improvement Loop **Duration:** ☐ 1 week ☑ 2 weeks

Scope (what we will build/test):

Expand to top 50 intents, weekly content update process, reporting dashboard

Value delivered (what improves + for whom):

Higher usefulness + adoption; measurable reduction in staff workload

(Example: reduces staff workload / improves patient experience / improves safety)

Investment (people/effort)

Owner: _____

Team involved: PM + content owner + analyst

Effort estimate: ☐ Low ☐ Medium ☐ High

Dependencies: _____

Decision rule (choose one):

- **GO if:** if workload + satisfaction improve
- **PIVOT if:** if usage is high but value is low
- **NO-GO if:** if no measurable movement

Success criteria (from Template 4.2):

Staff time ≤45 min/shift by Week 6 (or strong downward trend); satisfaction ≥4.0 by Week 6

Key questions this increment must answer:

Does staff time drop measurably? Do patients report better experience?

Notes / risks: _____

Exercise 4.1: Feature Prioritization (20 minutes)

Goal: Agree what to pilot first (Must/Should/Could/Won't).

01 List all candidate features (from Template 3.6).

02 Assign each feature to Must / Should / Could / Won't.

03 Pick 1–2 Must features for Increment 1 (plus one quick win).

04 Confirm each chosen feature supports a success metric (Template 4.2).

Output: A clear feature list + the first 2–3 increments drafted in Template 4.3.

Step 3: Complete Pilot Roadmap

What this step does (simple)

You now turn your increments into one clear 12-week pilot plan that answers:

- What happens each 2 weeks?
- Who is doing it?
- How will we measure it?
- When do we decide GO / PIVOT / NO-GO?

Think of this as a **single page plan** that leadership and the delivery team can both understand.

Guidance: Team & resource planning (keep it lightweight)

- Assign one owner per metric (so measurement doesn't "disappear").
- Assign two champions (frontline + ops) to drive adoption.
- Use semi-manual workarounds early if integration will slow you down.
- Protect time: a pilot needs 2–4 hours/week from key owners, not full-time teams.

Instructions (how to assemble the roadmap fast)

01 | Confirm the basics (from Templates 4.1 and 4.2)	• pilot scope (where, who, what's in/out) • success criteria (4–6 metrics)
02 | List increments (6 × 2 weeks) (from Template 4.3)	• each increment must deliver something testable • each ends with an Increment Review (GO/PIVOT/NO-GO)
03 | Add the minimum resources	• Sponsor (decision maker) • PM (runs cadence) • Ops owner (workflow + adoption) • IT (integration/security as needed) • Data/analyst (measurement) • Compliance/clinical owner (safety + approvals) (Keep it lean—avoid large committees.)
04 | Check data readiness (quick check)	For each success metric, confirm: • Source: where the data comes from (system/report/log) • Access: who can pull it weekly • Quality: is it consistent enough to trust If any metric can't be measured weekly → simplify the metric.

Template 4.4: Complete Pilot Roadmap

To be filled

Pilot overview

Pilot name: _____

Solution selected: _____

Location/unit: _____

Duration: _____ weeks

Increment length: ☐ 1 week ☐ 2 weeks

Pilot owner (PM): _____

Sponsor: _____

Success criteria (from Template 4.2):

- **Outcome**: _____
  (baseline ___ → target ___)
- **Outcome**: _____
  (baseline ___ → target ___)
- **Guardrail**: _____
  (threshold: _____)
- **Adoption**: _____
  (threshold: _____)

Resources (minimum roles):

Sponsor: _____ , PM: _____ ,
Ops owner: _____ , IT: _____ ,
Data/Analyst: _____ , Compliance/Clinical: _____

Measurement plan (weekly):

Metrics pulled from:
Owner:
Report cadence: ☐ Weekly dashboard ☐ Weekly spreadsheet
☐ Other: _____

Template 4.4: Complete Pilot Roadmap

To be filled

Roadmap (increments + decision points)

Increment 1

Weeks: _____

- Scope: _____
- Decision rule: _____

Increment 2

Weeks: _____

- Scope: _____
- Decision rule: _____

Increment 3

Weeks: _____

- Scope: _____
- Decision rule: _____

Increment 4

Weeks: _____

- Scope: _____
- Decision rule: _____

Increment 5

Weeks: _____

- Scope: _____
- Decision rule: _____

Increment 6

Weeks: _____

- Scope: _____
- Decision rule: _____

Close-out decision (end of Week 12): ☐ Scale ☐ Extend ☐ Stop

Use as only an example

Template 4.4: Filled Example

12-week pilot roadmap (6 increments) - Patient Access Assistant

Pilot overview

Pilot name: Patient Access Assistant - Clinic A Pilot

Solution selected: Solution A (Buy)

Location/unit: Clinic A + Call Center Pod 2

Duration: 12 weeks

Increment length: ☐ 1 week ☑ 2 weeks

Pilot owner (PM): Project Manager

Sponsor: Program Director, Outpatient Services

Success criteria (from Template 4.2):

- **Outcome**: Staff time on routine inquiries (60 → 35 min/shift by Week 12)
- **Outcome**: Patient satisfaction (3.6 → 4.2 by Week 12)
- **Guardrail**: Critical unsafe responses (0 allowed; ≤2 minor/week by Week 4)
- **Adoption**: % inquiries handled before human (≥35% by Week 12)

Resources (minimum roles):

Sponsor (Program Director), **PM**, **Ops owner** (Clinic Ops Lead), **IT** (Integration Lead), **Analyst** (Reporting), **Compliance** (Privacy), **Clinical** owner (Safety)

Measurement plan (weekly):

Metrics pulled from: assistant logs + handoff tickets + quick survey + time sampling
Owner: Analyst (weekly dashboard) + PM (weekly summary)

Use as only an example

Template 4.4: Filled Example (Cont.)

12-week pilot roadmap (6 increments) - Patient Access Assistant

Roadmap (increments + decision points)

Increment 1

Safe FAQ MVP

Weeks: 1–2

- **Scope:** Top 20 FAQs (EN/AR), safety boundaries, basic handoff inbox
- **Decision rule:** GO if 0 critical incidents + staff accept handoffs; PIVOT if content gaps; NO-GO if safety breach

Increment 2

Improve Handoff Reliability

Weeks: 3–4

- **Scope:** Structured intake + routing rules, handoff SLA, audit logs for handoffs
- **Decision rule:** GO if handoff success ≥95% by Week 4; PIVOT if routing errors; NO-GO if repeated failures

Increment 3

Expand Coverage

Weeks: 5–6

- **Scope:** Expand to top 50 intents, weekly content update loop, QA review process
- **Decision rule:** GO if adoption trend improves + minor errors ≤2/week; PIVOT if accuracy low

Increment 4

Adoption + Workflow Optimization

Weeks: 7–8

- **Scope:** Improve UX prompts, staff playbook for handoffs, champion training refresh
- **Decision rule:** GO if adoption ≥25% by Week 8; PIVOT if usage low due to channel friction

Increment 5

Partial Scheduling Integration (if ready)

Weeks: 9–10

- **Scope:** Lightweight scheduling link/lookup or semi-manual assist; refine routing by clinic rules
- **Decision rule:** GO if improves resolution time without safety issues; PIVOT if integration delays

Increment 6

Stabilize + Close-out Evidence

Weeks: 11–12

- **Scope:** Performance hardening, reporting pack, scale readiness checklist + lessons learned
- **Decision rule:** SCALE if key metrics moved + guardrails met; EXTEND if promising but incomplete; STOP if no measurable value

Close-out decision (end of Week 12): Scale / Extend / Stop (based on measured outcomes)

Step 4: Readiness & Launch Preparation

Why readiness matters

Most pilots don’t fail because the AI is “bad.” They fail because teams launch before basics are ready: owners aren’t assigned, data can’t be measured, staff aren’t trained, or compliance isn’t cleared.

This step is a **final pre-launch check**. If anything is missing, fix it **before** Week 1.

Instructions (quick readiness check)

Do this 30–45 minutes before launch with: **PM** + **Ops owner** + **IT** + **Data** + **Compliance/Clinical**.

Step 01
Walk through the checklist below.

Step 02
Mark each item Pass / Fix / Block.

Step 03
Any Block = do not launch.

Step 04
Any Fix = assign an owner + due date (usually Week 1)

Template 4.5: Pilot Launch Readiness Checklist

To be filled

A) People & ownership

Check	Pass / Fix / Block	Owner	Notes
Sponsor confirmed (decision maker)			
PM assigned (runs cadence)			
Ops owner assigned (workflow + adoption)			
IT owner assigned (access/integration/security)			
Data/analyst assigned (weekly measurement)			
Clinical/Compliance owner assigned (safety + approvals)			

B) Scope & success criteria

Check	Pass / Fix / Block	Owner	Notes
Pilot scope is written (Template 4.1)			
Success criteria defined (Template 4.2)			
O/PIVOT/NO-GO rules set for Increment 1			

Template 4.5: Pilot Launch Readiness Checklist (Cont.)

To be filled

C) Data & measurement readiness

Check	Pass / Fix / Block	Owner	Notes
Baselines confirmed for pilot metrics			
Data sources identified for each metric			
Weekly reporting method agreed (dashboard or sheet)			
QA process for errors/safety incidents defined			

D) Technical readiness

Check	Pass / Fix / Block	Owner	Notes
Access set up (accounts, roles, permissions)			
Hosting / residency confirmed (UAE/KSA if relevant)			
Audit logs enabled (minimum required)			
Escalation/handoff mechanism tested end-to-end			

Template 4.5: Pilot Launch Readiness Checklist (Cont.)

To be filled

E) Operational readiness

Check	Pass / Fix / Block	Owner	Notes
Staff trained (15–30 mins)			
“Who handles handoffs?” is clear			
Communication plan ready (who gets updates)			
Weekly increment review meeting scheduled			

F) Compliance readiness

Check	Pass / Fix / Block	Owner	Notes
Patient notice/consent wording approved (if needed)			
DPA / vendor terms reviewed (if vendor used)			
Risk items from Step 4 addressed (Templates 3.4/3.5)			

Launch decision:

☐ Launch ☑ Delay (Fix items) ☐ Do not launch (Block items)

Date of readiness review: _____

Reviewed by: _____

Template 4.5: Pilot Launch Readiness Checklist

Use as only an example

A) People & ownership

Check	Pass / Fix / Block	Owner	Notes
Sponsor confirmed (decision maker)	Pass	Program Director	Sponsor will attend increment reviews (Week 2, 4, 6…).
PM assigned (runs cadence)	Pass	PM	Weekly cadence + reporting confirmed.
Ops owner assigned (workflow + adoption)	Pass	Clinic Ops Lead	Champions nominated (front desk + call center).
IT owner assigned (access/integration/security)	Pass	Clinic Ops Lead	Champions nominated (front desk + call center).
Data/analyst assigned (weekly measurement)	Pass	Analyst	Dashboard template created; weekly pull every Thursday.
Clinical/Compliance owner assigned (safety + approvals)	Pass	Clinical Owner + Privacy	Clinical escalation rules approved.

B) Scope & success criteria

Check	Pass / Fix / Block	Owner	Notes
Pilot scope is written (Template 4.1)	Pass	PM	In/out scope confirmed; no clinical advice.
Success criteria defined (Template 4.2)	Pass	PM + Analyst	6 metrics agreed (3 outcomes, 2 guardrails, 1 adoption).
O/PIVOT/NO-GO rules set for Increment 1	Pass	Clinic Ops Lead	Safety breach = immediate pause; low adoption = content pivot.

Template 4.5: Pilot Launch Readiness Checklist (Cont.)

Use as only an example

C) Data & measurement readiness

Check	Pass / Fix / Block	Owner	Notes
Baselines confirmed for pilot metrics	Fix	Analyst	Staff time baseline needs 1-week sampling pre-launch.
Data sources identified for each metric	Pass	Analyst	Logs + tickets + survey + sampling plan documented.
Weekly reporting method agreed (dashboard or sheet)	Pass	PM + Analyst	Weekly 1-page scorecard every Friday.
QA process for errors/safety incidents defined	Pass	Clinical Owner	Weekly QA review + immediate escalation channel.

D) Technical readiness

Check	Pass / Fix / Block	Owner	Notes
Access set up (accounts, roles, permissions)	Fix	IT Lead	Finalize content editor + viewer roles.
Hosting / residency confirmed (UAE/KSA if relevant)	Pass	IT + Vendor	UAE hosting confirmed in architecture note.
Audit logs enabled (minimum required)	Fix	IT + Vendor	Log retention policy (90 vs 180 days) not finalized.
Escalation/handoff mechanism tested end-to-end	Pass	Ops Lead	Test completed with 10 handoffs; all received.

Template 4.5: Pilot Launch Readiness Checklist (Cont.)

Use as only an example

E) Operational readiness

Check	Pass / Fix / Block	Owner	Notes
Staff trained (15–30 mins)	Pass	Ops Lead	Training scheduled for Fri; quick guide printed.
“Who handles handoffs?” is clear	Pass	Ops Lead	Call center Pod 2 primary; backup roster defined.
Communication plan ready (who gets updates)	Pass	PM	Weekly update list confirmed (Sponsor, IT, Compliance, Ops).
Weekly increment review meeting scheduled	Pass	PM	Calendar holds created for Week 2, 4, 6, 8, 10, 12.

F) Compliance readiness

Check	Pass / Fix / Block	Owner	Notes
Patient notice/consent wording approved (if needed)	Fix	Compliance Lead	Final text awaiting legal sign-off.
DPA / vendor terms reviewed (if vendor used)	Fix	Legal + Compliance	Sub-processor list requested; DPA clause review pending.
Risk items from Step 4 addressed (Templates 3.4/3.5)	Fix	Compliance Lead	“Fix list” tracked; no blocks identified.

Launch decision: ☐ Launch ☑ Delay (Fix items) ☐ Do not launch (Block items) | **Target launch date once fixes closed:** _____

Fix items to close before launch (minimum set):

- Finalize patient notice wording (Compliance)
- Confirm baseline sampling for staff time (Analyst)
- Finalize RBAC roles + audit log retention (IT)
- Complete DPA/sub-processor review (Legal/Compliance)

Section 4 Outputs

What you should have by the end of Section 4

Template 4.1 completed: Pilot scope definition (location/unit, users, in-scope/out-of-scope, duration, dependencies).

Template 4.2 completed: Pilot success criteria worksheet (2–3 outcome metrics + 1–2 guardrails + 1 adoption metric, with thresholds and decision rules).

Template 4.3 completed: Increment plan (2-week increments with scope, value delivered, key questions, success criteria, and GO/PIVOT/NO-GO rules).

Template 4.4 completed: Full pilot roadmap (12-week plan with 6 increments, decision points, resource plan, and measurement plan).

Template 4.5 completed: Pilot launch readiness checklist (Pass/Fix/Block with owners and due dates).

Section 4 deliverables summary: A **detailed pilot roadmap** with 1–2 week increments | **Clear success criteria** + GO/PIVOT/NO-GO decision rules | A **resource allocation plan** + readiness confirmation to launch

Introduction

A pilot succeeds when it is treated like a **measured experiment**, not a “soft launch.” Most pilots fail for predictable reasons: unclear ownership, weak measurement, low stakeholder engagement, or letting scope expand until the team loses control.

What makes pilots succeed
(keep it practical)

- **Clear cadence**: weekly metrics + weekly update + end-of-increment review
- **Simple success metrics**: a few outcomes + guardrails + adoption (from Template 4.2)
- **Fast feedback loop**: frontline input turns into quick fixes every increment
- **Disciplined decisions**: GO / PIVOT / NO-GO based on evidence, not opinion
- **One accountable owner**: PM runs the rhythm; sponsor makes decisions

What makes pilots fail
(common patterns)

- **No baseline or weak data:** you can’t prove improvement
- **Too many features at once:** no one knows what caused results
- **Stakeholders disengage:** adoption drops and issues surface late
- **Vendor dependency:** the team waits for the vendor instead of learning
- **No decision gates:** pilots run indefinitely without a clear call

Execution best practices
(the “minimum set”)

This section gives you lightweight tools to run the pilot:

- **Template 5.2:** measurement plan (what you’ll measure and how)
- **Template 5.3:** communication plan (who gets updates and when)
- **Template 5.4:** weekly status report (1 page)
- **Template 5.5:** increment review agenda (GO/PIVOT/NO-GO decision meeting)

Rule of thumb: If you can’t run the pilot with a weekly 30-minute rhythm + a one-page scorecard, it’s too complex—simplify before continuing.

Measurement & Data Collection

Why this matters

If you can't measure it weekly, you can't prove its value. This step turns your success criteria (Template 4.2) into a **lightweight measurement routine**.

How to operationalize measurement (easy rules)

- Measure only what you **will use to decide** (GO / PIVOT / NO-GO).
- Keep it small: **4–6 metrics** total (outcomes + guardrails + adoption).
- Assign **one owner per metric** (so measurement doesn't disappear).
- Use the **same method** every week (consistency > perfection).
- Collect both **numbers** and **"why"** (a short notes field for context).

Data collection methods (pick the simplest that works)

- **System reports**: scheduling/EHR dashboards, ticketing, call logs
- **Assistant logs**: conversations, handoffs, response times, escalations
- **Time sampling**: quick weekly snapshots (e.g., 15-minute samples x 3 days)
- **Short surveys**: 1–3 questions after interaction (patients or staff)
- **QA review**: small weekly sample for correctness/safety

Best practices (data quality without heavy process)

01 **Baseline first**
Confirm starting point before Week 1 (or during Increment 1).

02 **One source of truth**
Define where each metric comes from (no guessing).

03 **Avoid vanity metrics**
"Number of chats" is not success unless tied to outcomes.

04 **Track incidents**
Safety issues are not "feedback" - they're a stop signal.

05 **Keep a weekly scorecard**
One page is enough.

To be filled

Template 5.2: Pilot Measurement Plan

Pilot name: _____ Pilot duration: ____ weeks

Increment length: ☐ 1 week ☐ 2 weeks Measurement owner: _____

Metric type	Metric (from Template 4.2)	Baseline	Target/ threshold	Data source	Collection method	Owner	Frequency	Notes field (what to capture)
Outcome							Weekly	
Outcome							Weekly	
Guardrail							Weekly	
Guardrail							Weekly	
Adoption							Weekly	

Weekly measurement rhythm (minimum):

Metrics pulled on: ☐ Mon ☐ Tue ☐ Wed ☐ Thu ☐ Fri

Weekly update sent on: ☐ Mon ☐ Tue ☐ Wed ☐ Thu ☐ Fri

End-of-increment review: Weeks 2, 4, 6, 8, 10, 12 (GO/PIVOT/NO-GO)

Use as only an example

Template 5.2: Pilot Measurement Plan

Pilot name: Patient Access Assistant - Clinic A Pilot

Pilot duration: 12 weeks

Increment length: ☐ 1 week ☑ 2 weeks

Measurement owner: Analyst (Reporting)

Metric type	Metric (from Template 4.2)	Baseline	Target/ threshold	Data source	Collection method	Owner	Frequency	Notes field (what to capture)
Outcome	Staff time on routine inquiries (min/shift)	60	≤45 by Week 6; ≤35 by Week 12	Time sampling sheet + staff quick log	3×15-min samples/week + short shift log	Ops Lead	**Weekly**	Any unusual staffing events; peak days; workflow changes
Outcome	Patient satisfaction (access touchpoints (1–5))	3.6	≥4.0 by Week 6; ≥4.2 by Week 12	Post- interaction survey	1-question survey after chat + QR at desk	PM	**Weekly**	Main complaint themes; channel issues; language gaps
Guardrail	Critical unsafe responses (count/ week)	0	0 (no tolerance)	QA review + incident log	Review 30 random chats/week + flag incidents	Clinical Owner	**Weekly**	Type of unsafe issue; root cause; fix applied
Guardrail	Handoff success rate (%)	95%	≥98% by Week 6; ≥99% by Week 12	Ticketing/ handoff queue	Compare handoff logs vs tickets received	Ops Lead	**Weekly**	Delays/queue overload; routing errors; SLA misses
Adoption	% of inquiries handled by assistant before human	0%	≥15% by Week 4; ≥25% by Week 8; ≥35% by Week 12	Assistant analytics + handoff logs	Weekly export of sessions + handoffs	Analyst	**Weekly**	Top intents; drop-off points; channel friction

Weekly measurement rhythm (minimum):

Metrics pulled on: ☐ Mon ☐ Tue ☐ Wed ☑ Thu ☐ Fri

Weekly update sent on: ☐ Mon ☐ Tue ☐ Wed ☐ Thu ☑ Fri

End-of-increment review: Weeks 2, 4, 6, 8, 10, 12 (GO/PIVOT/NO-GO)

Weekly measurement rhythm (minimum):

- Time sampling is used because it's fast and reliable for pilots (no heavy instrumentation needed).
- QA review is mandatory (safety guardrail); any critical incident triggers immediate escalation and possible pause.
- Adoption is tracked to confirm the tool is actually being used (without usage, outcomes won't move).

Stakeholder Engagement & Communication

Why this matters

Pilots don’t fail because of the tech — they fail when people stop paying attention.
This step keeps the pilot visible, trusted, and adopted with a simple weekly rhythm.

Communication cadence (minimum you need)

Use just 3 touchpoints:

01 Weekly update (5 minutes to read)

Sent every week with the scorecard + what changed.

02 End-of-increment review (30–45 minutes)

Every 2 weeks: GO / PIVOT / NO-GO decision.

03 Fast issue escalation (as needed)

Safety or compliance issues should not wait for meetings.

Stakeholder Touchpoints

Stakeholder	Frequency	Focus Areas
Leadership	Weekly	Progress + risks + decisions
Frontline	Weekly	Changes + tips + issue log
IT + Data	Weekly	Bugs + data + fixes
Compliance	Weekly	Incidents + safeguards

To be filled

Template 5.3: Pilot Communication Plan

Pilot name : _____ Pilot sponsor : _____ Pilot PM : _____

Audience	What they need	Channel	Frequency	Owner
Sponsor / leadership	1-page progress + risks + GO/ PIVOT/NO-GO decisions			
Frontline users (Clinic A + Call Center Pod 2)	What changed + how to use it + how to escalate issues			
IT / Data	Bugs, access issues, reporting needs, integration decisions			
Compliance / Clinical safety	Safety incidents, QA findings, content risks			
Vendor (if used)	Prioritized fixes, content requests, timeline for next increment			

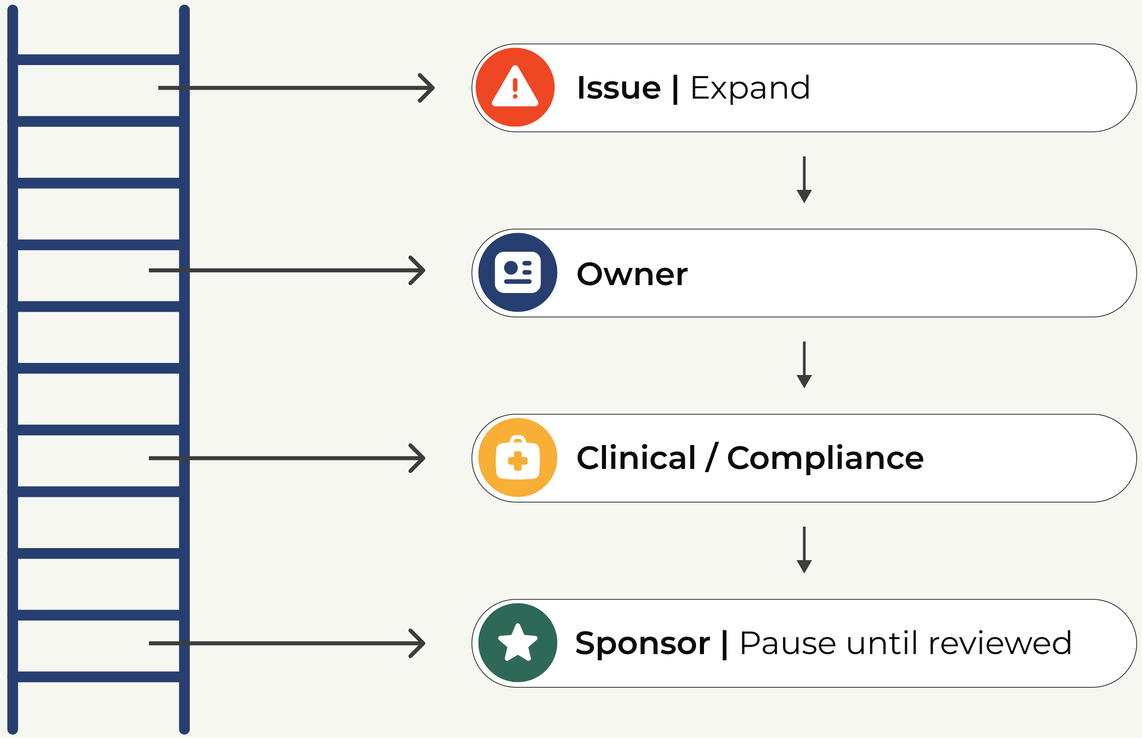
Escalation rule (clear and strict):

If a _____ occurs → notify _____ within _____

Template 5.3: Pilot Communication Plan (Cont.)

To be filled

Escalation Ladder (Safety/Compliance)



Template 5.3: Filled Example

Use as only an example

Pilot name : Patient Access Assistant - Clinic A Pilot **Pilot sponsor :** Program Director, Outpatient Services **Pilot PM :** Project Manager

Audience	What they need	Channel	Frequency	Owner
Sponsor / leadership	1-page progress + risks + GO/ PIVOT/NO-GO decisions	Email + 15-min check-in	Weekly (Fri) + end-of-increment (biweekly)	PM
Frontline users (Clinic A + Call Center Pod 2)	What changed + how to use it + how to escalate issues	WhatsApp group + quick guide	Weekly (Mon) + as-needed updates	Ops Lead
IT / Data	Bugs, access issues, reporting needs, integration decisions	Teams channel + ticket list	Weekly (Thu)	IT Lead + Analyst
Compliance / Clinical safety	Safety incidents, QA findings, content risks	Email + incident log	Weekly (Thu) + immediate escalation	Clinical Owner + Compliance
Vendor (if used)	Prioritized fixes, content requests, timeline for next increment	Weekly call + shared tracker	Weekly (Wed)	PM

Escalation rule (clear and strict):

If a **critical unsafe response** or **privacy breach** occurs → notify **Clinical Owner + Compliance + PM** within **2 hours** and **pause the feature** until reviewed.

Template 5.4: Weekly Pilot Status Report

To be filled

Pilot: _____ Week #: ____ Dates: _____

Overall status: ☐ On track ☐ At risk ☐ Off track

This week's metrics (from Template 5.2)

- Outcome 1: _____
  (baseline ____ → current ____ → target _____)
- Outcome 2: _____
  (baseline ____ → current ____ → target _____)
- Guardrails: _____
  (incidents: _____)
- Adoption: _____
  (current: _____)

This week's metrics (from Template 5.2)

_____

_____

_____

Issues / risks (top 3)

- Risk: _____ | Owner: _____ | Due: _____
- Risk: _____ | Owner: _____ | Due: _____
- Risk: _____ | Owner: _____ | Due: _____

Decisions needed (if any)

_____

_____

_____

Next week focus (2–3 bullets)

_____

_____

_____

Template 5.4: Weekly Pilot Status Report

Use as only an example

Pilot: Patient Access Assistant - Clinic A Pilot **Week #:** 3 **Dates:** 03–07 Mar 2026

Overall status: ☑ On track ☐ At risk ☐ Off track

This week's metrics (from Template 5.2)

- **Outcome 1:** Staff time on routine inquiries
  (baseline 60 → **current** 52 → **target** ≤45 by Week 6**)**
- **Outcome 2:** Patient satisfaction (1–5)
  (baseline 3.6 → **current** 3.9 → **target** ≥4.0 by Week 6**)**
- **Guardrails:** 0 critical unsafe responses
  (incidents: 1 phrasing error fixed**)**
- **Adoption:** % handled by assistant before human
  (current: 18% (target ≥15% by Week 4)**)**

This week's metrics (from Template 5.2)

- Expanded FAQ coverage from 20 → 35 intents (top appointment questions).
- Updated escalation rules for "medical symptoms" keywords (faster human handoff).
- Added clearer "book appointment" prompt to reduce drop-offs.

Issues / risks (top 3)

- **Risk**: Some patients ask "insurance coverage" (content gap) | **Owner**: Content Owner | **Due**: Week 4
- **Risk**: Handoff queue response times vary at peak hours | **Owner**: Ops Lead | **Due**: Week 4
- **Risk**: Audit log retention setting not finalized | **Owner**: IT Lead | **Due**: Week 3 (end)

Decisions needed (if any)

Decide whether to include "insurance + billing FAQs" in scope (currently out-of-scope).

Next week focus (2–3 bullets)

- Implement structured intake fields for routing (Increment 2).
- Finalize queue SLA + backup coverage for peak times.
- Run weekly QA review (30 chat sample) and update content pack.

GO / PIVOT / NO-GO Decision Making

What this is

At the end of every increment (usually every 2 weeks), you run one short review meeting and make one clear call:

GO : continue / expand
(evidence is strong enough)

PIVOT : continue, but change something
(scope, feature, workflow, metric, audience)

NO-GO = stop
(risk is too high or value is not showing)

This prevents pilots from dragging on without decisions.

How to run an increment review (easy rules)

01 Show the scorecard first (Template 5.2 metrics).

02 Confirm guardrails
(safety/compliance) before discussing value.

03 Discuss 3 questions only:

- What has improved?
- What didn't improve (and why)?
- What will we change in the next increment?

04 Decide Go / Pivot / No-Go and record it.

Timebox: 30–45 minutes.

Template 5.5: Increment Review & Decision Meeting Agenda

To be filled

Meeting: Increment Review + Decision

Increment #: _____ Dates: _____ Duration: 30–45 mins

Attendees: Sponsor, PM, Ops owner, Data/Analyst, IT (as needed), Compliance/Clinical (as needed)

Metrics snapshot (10 mins)

- Outcome metrics: _____
- Guardrails (incidents): _____
- Adoption: _____

What we learned (10 mins)

- What worked: _____
- What didn't: _____
- Root cause (best guess): _____

Metrics snapshot (10 mins)

Decision: ☐ GO ☐ PIVOT ☐ NO-GO

Reason (1–2 lines): _____

Next increment focus: _____

Actions (5 mins)

Owner - Action - Due

Use as only an example

Template 5.5: Filled Example

PIVOT decision (with rationale)

Meeting: Increment Review + Decision

Increment #: 2 Dates: _____ Duration: 30–45 mins

Attendees: Sponsor, PM, Ops owner, Data/Analyst, IT (as needed), Compliance/Clinical (as needed)

Metrics snapshot (10 mins)

- **Outcome metrics:** Handoff success rate improved to 96% (target was ≥98% by Week 6)
- **Guardrails (incidents):** 0 critical safety incidents (guardrail met)
- **Adoption:** Adoption increased to 22% (on track)

What we learned (10 mins)

- **What worked:** _____
- **What didn't:** _____
- **Root cause (best guess):** _____

_____

Metrics snapshot (10 mins)

Decision: ☐ GO ☑ PIVOT ☐ NO-GO

Reason (1–2 lines): Why PIVOT (not GO): Routing errors are still happening during peak hours, causing delays and staff frustration. Value is visible, but reliability isn't strong enough to scale scope yet.

Next increment focus: Simplify routing rules (fewer categories) | Add a backup queue during peak hours | Update intake prompt to reduce incomplete requests

Actions (5 mins)

Owner - Action - Due

Ops Lead: define peak-hour backup coverage (Due: Week 5)

IT Lead: adjust routing rules in config (Due: Week 5)

PM: update staff playbook + training refresh (Due: Week 5)

Common Pitfalls & How to Avoid Them

Use this as a **quick troubleshooting guide** during the pilot.
Each pitfall includes: warning signs → prevention → recovery.

Pitfall	Warning signs	Prevention (do this early)	Recovery (if it happens)
Scope creep	“Can we also add…?” every week; increment goals slip	Lock in/out scope (Template 4.1) + MoSCoW per increment	Pause new requests → push to backlog → re-baseline next increment
Weak measurement / no baseline	Debates replace data; “it feels better”	Baseline in Week 0–1 + Template 5.2 owners per metric	Run a 1-week baseline sprint; reduce metrics to 4–6 measurable ones
Too many features at once	Hard to finish increments; unclear what caused change	One increment = one outcome; keep scope “finishable”	Split scope; remove non-critical features; restart with smaller increments
Low frontline engagement	Users ignore tool; champions disappear; complaints rise	Assign 2 champions + 15–30 min training + weekly feedback loop	Run listening session; simplify workflow; fix top 3 pain points fast

Common Pitfalls & How to Avoid Them (Cont.)

Pitfall	Warning signs	Prevention (do this early)	Recovery (if it happens)
Vendor over-reliance	Team waits for vendor; delays pile up	Define vendor SLA + internal workaround path	Escalate with timeline; switch to manual fallback; reduce dependencies
Safety/compliance treated as “later”	Incidents; unclear escalation; approvals incomplete	Guardrails + escalation rule + weekly QA (Templates 4.5 + 5.2)	Pause feature; incident review; tighten rules; relaunch only when safe
Poor handoffs / unclear ownership	Requests fall through; staff frustrated	Define “who owns handoffs” + backup coverage	Fix routing; add queue rules; assign backup owner for peak hours
No decision discipline	Pilot runs on forever; no clear calls	Schedule increment reviews upfront + Template 5.5 decision record	Force a decision meeting; use thresholds; choose GO/PIVOT/NO-GO
Misaligned expectations	Leadership expects scale-ready results in 2 weeks	Set expectations: pilot = evidence, not perfection	Reframe goals; show scorecard; agree what “success” means again

Section 5 Outputs

What you should have by the end of Section 5

1. Template 5.2 completed: Pilot Measurement Plan (metrics, baselines, targets, data sources, owners, weekly cadence).
2. Template 5.3 completed: Pilot Communication Plan (audiences, channels, frequency, owners, escalation rule).
3. Template 5.4 completed: Weekly Pilot Status Reports (one-page scorecard each week, issues/risks, next steps).
4. Template 5.5 completed: Increment Review & Decision records (GO / PIVOT / NO-GO decisions + rationale + action log).
5. Pilot evidence pack: Weekly metrics + QA/safety log + adoption insights (ready to share with leadership).
6. Lessons learned summary: What worked, what didn't, and what must change before scale-up.

Introduction

This section shows how the playbook works in practice. The goal isn't to impress with "perfect" stories — it's to help teams learn faster by seeing what typically happens in real pilots: what works, what breaks, and what decisions matter most.

Purpose of these use cases (why they're here)

Each use case helps you:

- See the **full PB2 flow** end-to-end (evaluate → decide → pilot → scale/stop)
- Understand what **good evidence** looks like (not just opinions)
- Spot common **risk signals early** (before they become failures)
- Reuse proven patterns (metrics, increment structure, decision rules)

How to use these examples (simple)

Use each case as a reference when you're stuck:

- **Before evaluation (Section 3):** Borrow objective formats, stakeholder lists, and compliance checks.
- **During comparison & decision (Step 7–8):** Compare how trade-offs were handled (impact vs cost vs credibility).
- **During the pilot (Sections 4–5):** Copy the roadmap structure, cadence, and GO/PIVOT/NO-GO decision logic.

Tip: Don't try to copy everything. Copy the **shape:** metrics, increments, decision points, and how evidence was used.

Use Case 1 - Healthcare AI Example (SDG 3 Aligned): Patient Access Assistant

Organization context

A mid-size outpatient healthcare provider (UAE), serving a high-volume urban population.

Mission focus: improve access and patient experience while protecting safety, privacy, and trust (SDG 3: Good Health & Well-being).

Challenge & quantified objectives (PB1 → PB2 inputs)

Problem: frontline staff and call center are overloaded with repetitive appointment questions, causing slow response times and poor patient experience. Quantified objectives (examples):

- Reduce staff time spent on routine inquiries from 60 → 35 min/shift (12 weeks)
- Improve patient satisfaction on access touchpoints from 3.6 → 4.2 / 5
- Improve handoff reliability to humans from 95% → 99% (guardrail)
- Increase % inquiries resolved before human from 0% → 35% (adoption)

Evaluation process (Section 3 summary)

Solutions considered:

- A — Patient Access Assistant (buy) — FAQ + safe handoff
- B — No-show prediction + outreach (hybrid)
- C — Digital intake + queue dashboard (buy)

Impact estimation (high-level):

- A scored highest on near-term impact + credibility (fast pilot-ready, clear handoff control)
- B had potential long-term value but lower credibility (data quality dependency)
- C improved intake but required heavier workflow change

Decision: **Pilot Solution A**
(best impact-to-speed, safest to test first).

Pilot approach (12 weeks, 6 increments)

Increment	Focus	Decision
1	Safe FAQ MVP (top intents + escalation rules)	GO
2	Routing + handoff reliability	PIVOT
3	Expand intent coverage + weekly improvement loop	GO
4	Adoption + workflow optimization	GO
5	Light scheduling assist (semi-manual)	GO
6	Stabilize + close-out evidence pack	GO → SCALE

Outcomes (end of Week 12)

- Staff time on routine inquiries: **60 → 36 min/shift** (near target)
- Patient satisfaction: **3.6 → 4.1 / 5**
- Handoff success rate: **95% → 99%**
- Adoption (assistant handles before human): **0% → 33%**
- ROI (pilot-level): improved throughput and reduced staff burden justified scaling; full ROI projected during rollout.

Key takeaways (critical lessons)

- Keep scope tight: one workflow + one unit makes evidence clear.
- Guardrails matter: safe escalation rules prevented “trust-killing” incidents.
- PIVOT is normal: one pivot on routing avoided bigger adoption failure.
- Measurement discipline wins: weekly scorecards kept decisions factual.
- Semi-manual first is fine: integration can wait until evidence is strong.

Use Case 2 — Mission-Driven Organization Example: Community Health Referral Assistant (SDG 3 Aligned)

Organization context

A mission-driven NGO operating across UAE + Jordan, helping low-income families access primary care, maternal health, and mental health referrals.

Key constraint: limited case worker capacity and inconsistent follow-up after referrals.

Challenge & quantified objectives (PB1 → PB2 inputs)

Problem: Beneficiaries miss steps due to confusion (documents, directions, eligibility). Case workers spend time answering repeat questions and chasing follow-ups. Objectives (examples):

- Referral completion 45% → 60% (12 weeks)
- Follow-up time 5 days → 2 days
- Case worker repetitive Q time 50 → 30 min/day
- Adoption ≥30% inquiries handled before human
- Guardrail: 0 privacy incidents

Evaluation process (Section 3 summary)

Solutions considered:

- A — Buy referral assistant (multilingual guidance + safe handoff)
- B — Build WhatsApp bot (custom flows)
- C — CRM automation only (reminders + tasks)

Decision: **Pilot A** (fastest, most credible path; strong language support; proven handoff).

Pilot approach (12 weeks, 6 increments)

Increment	Focus	Decision
1	Top 25 FAQs + safe handoff	GO
2	Eligibility/doc guidance	PIVOT (simplified wording)
3	Reminders + “stuck” escalation	GO
4	Partner clinic routing basics	GO
5	Stabilize + scale pack	GO
6	Stabilize + close-out evidence pack	GO → EXTEND (4-week extension to validate in second region)

Outcomes (end of Week 12)

- Completion **45% → 58%** (near target; strong upward trend)
- Follow-up time **5 → 2.3 days**
- Case worker time **50 → 32 min/day**
- Adoption **34%** handled before human
- Guardrail: **0** privacy incidents

Key takeaways (critical lessons)

- Clear language beats “advanced AI” for beneficiary impact.
- Strict privacy boundaries + early handoff prevent trust failures.
- Most pivots are content/ workflow, not technology.
- Extend before scaling when geography/program context changes.

Use Case 3 — Failed Pilot Example: Clinical Documentation Copilot (Stopped)

What it was

A hospital department piloted an AI “documentation copilot” to help clinicians draft visit notes faster.
The intent was good (reduce admin burden), but the pilot was stopped early.

What went wrong (and why)

01 Pilot scope was too big	• Tried to cover multiple note types and multiple clinician teams at once. • Result: too many variables, no clear learning per increment.
02 Guardrails were weak	• Escalation rules and QA sampling were not strict enough. • Clinicians lost trust after repeated “almost correct” outputs.
03 Measurement wasn’t operational	• No clean baseline for documentation time and rework. • Success became subjective: “some people like it, some don’t.”
04 Workflow fit was underestimated	• Copy/paste friction, login issues, and poor integration created extra steps. • The AI saved time in theory but added time in practice.
05 Vendor dependency slowed learning	• Fixes required vendor cycles, so increments couldn’t improve fast enough.

Use Case 3 — Failed Pilot Example: Clinical Documentation Copilot (Stopped)

Warning signs that were missed

- “We’ll measure properly later” (no baseline in Week 1)
- “Let’s include more teams so it’s representative” (scope creep early)
- Early feedback: “I don’t trust this output” (trust signal ignored)
- Increasing rework time (AI outputs required heavy editing)
- No clear GO/PIVOT/NO-GO decisions (pilot drift)

Lessons learned (what mattered)

- Trust is a metric. If clinicians don’t trust outputs, adoption dies.
- Guardrails must be strict. QA sampling + escalation rules are non-negotiable.
- Integration matters more than features. Friction kills time-savings claims.
- Start smaller than you think. One team, one note type, one workflow.

How to avoid this next time (PB2 fixes)

- Use Template 4.1 to lock scope (one unit + one note type).
- Define 4–6 success metrics in Template 4.2 (baseline → target).
- Add weekly QA + incident tracking in Template 5.2 (guardrail first).
- Run end-of-increment decisions using Template 5.5 (GO/PIVOT/NO-GO).
- If integration isn’t ready, pilot with a semi-manual workflow first.

Close-out decision: NO-GO (stopped) — value not proven safely and reliably within pilot constraints.

Section 6 Outputs

What you should have by the end of Section 6 :

1. **3 real-world use cases documented** in a consistent format (context → objectives → evaluation → pilot → outcomes).
2. **At least one SDG 3 healthcare case** showing end-to-end PB2 execution (evaluation → decision → pilot → scale/extend/stop).
3. **One mission-driven case** showing how to apply PB2 under real resource constraints (limited staff, multi-stakeholder environment).
4. **One failed pilot case** with clear warning signs, root causes, and prevention steps (what to avoid and how PB2 fixes it).
5. **Reusable patterns** extracted from cases: example metrics, increment structures, GO/PIVOT/NO-GO thresholds, and common pitfalls.

From Evaluation to Action

This playbook is designed to help you move from “**we have an AI use case**” to “**we proved value safely and made a clear decision**.”

Keep the process simple: measure impact, make trade-offs visible, pilot in small increments, and decide based on evidence.

Core concepts (6–8 key points)

01 **Start with PB1 inputs**
A clear challenge, beneficiaries, & constraints make evaluation faster & safer.

02 **Quantify objectives**
Success must be measurable (baseline → target), not opinion-based.

03 **Compare solutions fairly**
Use the same yardstick across options (impact, cost, credibility).

04 **Make uncertainty visible**
Credibility scoring prevents false confidence & surfaces where evidence is weak.

05 **Address compliance early**
If requirements aren’t met, don’t score or select-fix scope/safeguards first.

06 **Pilot in 1–2 week increments**
Each increment delivers value and ends with GO/PIVOT/NO-GO.

07 **Guardrails first**
Safety & privacy are non-negotiable; incidents trigger immediate action.

08 **Decide and document**
A decision is only “real” if it’s written, justified, and shareable.

From Evaluation to Action

Critical success factors
(what makes PB2 work)

- A sponsor who attends decision gates and supports stopping when needed
- A small set of metrics with owners (measured weekly)
- Clear scope boundaries (no “pilot everything”)
- A steady cadence: weekly scorecard + end-of-increment review
- Fast learning loop: fix issues quickly, don’t wait for perfect conditions

Common pitfalls to avoid

- Starting without baselines or measurable targets
- Letting scope creep expand the pilot beyond what can be measured
- Choosing based on demos or internal politics instead of evidence
- Treating compliance and safety as “later”
- Running the pilot without decision gates (pilot drift)

Next step: Use the template index in the appendices to run PB2 as a workflow - one step at a time, with decision points clearly scheduled.

Appendix A: Template Library Index (Quick Reference)

Use this index to jump straight to the template you need. Templates are grouped by the section where they're used.

Section 3 - Evaluate & Select
(Templates 3.1–3.9)

- **Template 3.1:** Organization & Challenge Overview
- **Template 3.2:** Stakeholder Mapping Worksheet
- **Exercise 3.1:** Stakeholder Identification Workshop
- **Template 3.3:** Quantified Objectives Worksheet
- **Template 3.4:** Regulatory Compliance Checklist (UAE)
- **Template 3.5:** Regulatory Compliance Checklist (KSA)
- **Template 3.6:** AI Solution Definition Form
- **Template 3.7:** Impact Estimation Table (IET) + Credibility Scoring
- **Template 3.8:** Solution Comparison Dashboard
- **Template 3.9:** Decision Framework & Recommendation Report
- **Exercise 3.2:** Decision Workshop Guide

Section 4 - Design the Pilot Roadmap
(Templates 4.1–4.5)

- **Template 4.1:** Pilot Scope Definition
- **Template 4.2:** Pilot Success Criteria Worksheet
- **Template 4.3:** Increment Planning Worksheet
- **Exercise 4.1:** Feature Prioritization
- **Template 4.4:** Complete Pilot Roadmap (12-week structure)
- **Template 4.5:** Pilot Launch Readiness Checklist

Section 5 - Run the Pilot
(Templates 5.2–5.5)

- **Template 5.2:** Pilot Measurement Plan
- **Template 5.3:** Pilot Communication Plan
- **Template 5.4:** Weekly Pilot Status Report
- **Template 5.5:** Increment Review & Decision Meeting Agenda

Appendix A (Cont.)

Quick reference: “Which template do I use?”

- Need to define the problem + context? → 3.1
- Need stakeholder alignment? → 3.2 (+ Exercise 3.1)
- Need measurable success metrics? → 3.3 (and later 4.2 / 5.2)
- Need UAE/KSA compliance checks? → 3.4 / 3.5
- Need to document solution options? → 3.6
- Need quantified impact + confidence? → 3.7
- Need side-by-side comparison? → 3.8
- Need a final recommendation? → 3.9 (+ Exercise 3.2)
- Need pilot scope and plan? → 4.1 → 4.3 → 4.4
- Need launch readiness? → 4.5
- Need to run the pilot weekly? → 5.2 → 5.4 → 5.5

Appendix B: Glossary

AI / ML terms

- **AI** (Artificial Intelligence): Systems that perform tasks that normally require human intelligence (e.g., classification, summarization, decision support).
- **ML** (Machine Learning): Models that learn patterns from data to make predictions or decisions.
- **LLM** (Large Language Model): A language model used for chat, summarization, drafting, and Q&A.
- **Generative AI**: AI that creates content (text, images, code).
- **Prompt**: Instructions/input given to an LLM.
- **System prompt / guardrails**: Rules that constrain model behavior (e.g., “no medical advice,” “escalate clinical intent”).
- **Hallucination**: Confident output that is incorrect or unsupported.
- **RAG** (Retrieval-Augmented Generation): Responses grounded in approved documents or knowledge sources.
- **Fine-tuning**: Additional training on specific data to shape model behavior.
- **Inference**: Running the model to generate outputs.
- **Token**: A chunk of text processed by an LLM (affects cost and limits).
- **Latency**: Response time from user request to AI output.
- **Grounding**: Ensuring AI answers come from trusted sources or validated logic.
- **Evaluation dataset**: A set of test cases used to check accuracy, safety, and behavior.
- **False positive / false negative**: Incorrectly flagging or missing a condition/event (important for triage and safety).

Appendix B: Glossary (Cont.)

- **Bias**: Systematic errors that produce unfair outcomes for certain groups.
- **Model drift**: Performance changes over time due to shifts in data/workflows.
- **Human-in-the-loop** (HITL): Humans review/approve outputs (especially for safety or clinical workflows).
- **Escalation trigger**: A rule that forces handoff to a human (e.g., symptoms, medication, emergency language).

Healthcare terminology (SDG 3 context)

- **Beneficiary**: Person/population receiving health or well-being impact (patients/community).
- **Clinical safety**: Preventing patient harm from incorrect guidance or delayed escalation.
- **Triage**: Sorting cases by urgency.
- **Patient journey**: End-to-end experience (booking → visit → follow-up).
- **Care pathway**: Standard process for managing a condition/service.
- **Continuity of care**: Coordinated care across providers/time.
- **No-show rate**: % of appointments missed.
- **Referral**: Transfer to another service/provider for care.
- **Access** (patient access): Ability to reach services (booking, info, directions, eligibility).

Appendix B: Glossary (Cont.)

- **Frontline staff**: Staff directly interacting with patients (reception, call center, nurses).
- **Clinical intent**: A request that could affect medical outcomes (symptoms, meds, diagnosis).
- **Non-clinical intent**: Admin/logistics requests (hours, location, documents).
- **Escalation to clinician**: Routing to clinical staff for safety-critical queries.
- **Workload / throughput**: Volume handled per staff/time period.
- **Service-level agreement** (SLA): Target response/handling time (e.g., handoff within X minutes).
- **Quality assurance** (QA): Review process for correctness and safety.

Evaluation & methodology terms (PB2 / Gilb-aligned)

- **Impactful AI Method**: Structured evaluation using quantified objectives, impact estimation, credibility scoring, and evolutionary pilots.
- **PB1 outputs:** Challenge statement, beneficiaries, context, constraints, and prioritized use case from Playbook 1.
- **Organization & challenge context**: Mission, beneficiaries, constraints, and the exact problem being solved.
- **Stakeholder**: Any group affected by the solution (users, beneficiaries, decision makers, regulators).
- **Stakeholder-centric measurement**: Measuring success from the perspective of stakeholders who care.
- **Objective** (quantified objective): Measurable target with baseline → target.

Appendix B: Glossary (Cont.)

- **Quality attribute**: Aspect to improve (speed, accuracy, safety, satisfaction, cost).
- **Scale** (measurement scale): How an objective is measured (minutes, %, 1–5 score).
- **Baseline**: Current performance level before the pilot.
- **Target**: Desired performance level by a defined time.
- **Assumption**: A belief used in planning that must be tested (e.g., “patients will use WhatsApp”).
- **Dependency**: A condition that must be true to proceed (data access, approval, integration).
- **Candidate solutions**: Shortlisted options to compare in a consistent format.
- **Build vs buy**: Build internally vs purchase vendor solution.
- **Feature breakdown**: Listing solution capabilities as features to estimate impact per objective.
- **Impact Estimation Table** (IET): Grid mapping features → objectives with impact % + credibility.
- **Impact %**: Estimated improvement contribution of a feature to an objective.
- **Credibility score** (0–10): Confidence level that the estimate is true (0 guess → 10 proven).
- **Comparison dashboard**: Side-by-side view of solutions (impact, cost, credibility, value-for-money).
- **Impact-to-cost ratio**: A simple value-for-money indicator (benefit relative to cost).
- **Weighted criteria**: Decision method where criteria have weights (importance).
- **Decision framework**: The documented method and rationale used to select a solution.
- **Decision gate**: Stop/go checkpoint (input readiness, compliance, selection, launch readiness).
- **GO / PIVOT / NO-GO**: Continue, adjust, or stop based on evidence.

Appendix B: Glossary (Cont.)

- **Evo** (Evolutionary approach): Deliver value in small increments with frequent feedback and decisions.
- **Increment**: A 1–2 week cycle delivering something testable, ending with a decision.
- **Pilot roadmap**: The full plan of increments, measures, and decision points (e.g., 12-week plan).
- **Pilot scope**: Where/with whom/what is included and excluded.
  Guardrail metric: Safety/compliance metric with strict thresholds (often “no tolerance” for critical incidents).
- **Adoption metric**: A measure of real use (usage rate, % handled before human).
- **Evidence pack**: The documented pilot outputs (metrics, decisions, incidents, lessons).

Regulatory & compliance terms (UAE/KSA-ready)

- **Compliance**: Meeting legal, regulatory, and policy requirements.
- **Regulatory landscape**: The relevant authorities, laws, and obligations for deployment.
- **Data protection**: Rules governing collection, processing, storage, and sharing of personal data.
- **Personal data**: Any data that identifies a person directly or indirectly.
- **Health data / PHI**: Health-related personal data; higher sensitivity and stricter controls.
- **Consent / patient notice**: Communication and permission required for data use (where applicable).

Appendix B: Glossary (Cont.)

- **Data residency**: Where data is stored/processed (country/region requirement).
- **Data minimization**: Collect only what's necessary for the purpose.
- **Purpose limitation**: Use data only for the stated purpose.
- **Access control** (RBAC): Role-based permissions limiting who can see/do what.
- **Audit log**: Record of user/system actions for traceability.
- **Retention**: How long data/logs are stored before deletion.
- **Encryption**: Protecting data in transit and at rest.
- **Anonymization / pseudonymization**: Reducing identifiability; pseudonymized data can sometimes be re-linked.
- **DPA** (Data Processing Agreement): Vendor contract defining processing responsibilities and controls.
- **Sub-processors**: Third parties a vendor uses to process data.
- **Incident response**: Steps to detect, report, contain, and recover from incidents.
- **Risk register**: List of risks, owners, mitigation actions, and status.
- **Approval workflow**: Required sign-offs before launch (privacy, clinical safety, IT).

Appendix B: Glossary (Cont.)



UAE terms used in the playbook

- **DHA**: Dubai Health Authority.
- **DOH**: Department of Health (Abu Dhabi).
- **UAE Federal Data Protection Law**: UAE-level privacy framework governing personal data processing.
- **AI ethics** (UAE): Principles guiding safe, responsible AI use (fairness, transparency, safety).



KSA terms used in the playbook

- **PDPL** (KSA): Personal Data Protection Law in Saudi Arabia.
- **SDAIA**: Saudi Data & AI Authority.
- **SFDA**: Saudi Food & Drug Authority.
- **NHIC** (Saudi): National Health Information Center (health data/health information initiatives).

Appendix C: UAE/KSA Regulatory Quick Reference

(AI Pilots in Healthcare)

This appendix is a practical navigation aid - not legal advice. Always confirm requirements with your internal privacy/compliance lead and, where needed, external counsel.

A) UAE - Summary Table (quick scan)

Area	Primary body / instrument	What it means for an AI pilot (plain English)	PB2 "where it shows up"
Personal data protection	UAE PDPL (Federal Decree-Law No. 45 of 2021) + UAE Data Office	You must have a lawful basis, clear purpose, strong security controls, and defined data handling (incl. sharing with vendors).	Step 4 (Template 3.4) + launch readiness (4.5)
Digital health / health data rules	Federal Law No. 2 of 2019 (ICT in Health Fields)	Health data use via ICT must follow mandated mechanisms/controls for safety, security, and proper handling in the health sector.	Step 4 (3.4) + pilot controls (5.2)
Dubai healthcare regulator	Dubai Health Authority (DHA) policies & regulations	If operating in Dubai, follow DHA health facility/professional requirements and relevant policies/guidelines for safe practice.	Step 4 (3.4) + gate/ approval routing
Abu Dhabi health information exchange	DoH Abu Dhabi HIE policy + standards (ADHIE/Malaffi)	If exchanging clinical data in Abu Dhabi, required standards/ minimum datasets may apply. Plan integration and data quality early.	Step 4 (3.4) + solution definition (3.6)

Official UAE resources (links):

- UAE PDPL (Federal Decree-Law No. 45 of 2021) - official text/download.
- UAE Government portal summary of PDPL + UAE Data Office role.
- Federal Law No. 2 of 2019 (ICT in Health Fields) - official text/download.
- UAE Government portal page referencing ICT law in health sector.
- DHA policies & regulations hub.
- DoH Abu Dhabi HIE policy (ADHIE).
- DoH HIE standards (minimum datasets/standards references).

Appendix C: UAE/KSA Regulatory (Cont.)

(AI Pilots in Healthcare)

B) KSA - Summary Table (quick scan)

Area	Primary body / instrument	What it means for an AI pilot (plain English)	PB2 "where it shows up"
Personal data protection	KSA PDPL (Royal Decree M/19) + implementing regulations	You must manage personal data lawfully, define controller responsibilities, protect rights, and follow rules for processing/transfers.	Step 4 (Template 3.5) + launch readiness (4.5)
PDPL compliance enablement	SDAIA + National Data Governance Platform	Guidance, tools, and compliance services may be relevant for assessing or demonstrating PDPL compliance posture.	Step 4 (3.5) + governance pack (5.x)
Health information / national health data ecosystem	NHIC (National Health Information Center)	If interacting with national health information programs/standards, align early to avoid late integration rework.	Step 4 (3.5) + solution definition (3.6)
Medical devices / regulated medical software	SFDA	If the AI is used to diagnose/treat/clinical decision-making, it may fall under regulated categories - engage SFDA requirements early.	Step 4 (3.5) + selection gate (end of Section 3)

Official KSA resources (links):

- PDPL implementing regulation (SDAIA PDF).
- SDAIA regulations/policies hub (PDPL-related).
- National Data Governance Platform (PDPL services/compliance).
- NHIC official site (Saudi Health Council) + contact form.
- SFDA official site + contact page.

Appendix C: UAE/KSA Regulatory (Cont.)

(AI Pilots in Healthcare)

C) Official Contact Channels (use when you need clarification or formal routing)

Entity	When to contact	Official contact channel
DHA (Dubai)	Dubai healthcare compliance/policies, licensing/regulatory questions	Email: info@dha.gov.ae Timing: Monday to Thursday: 7:30 AM -3:30 PM | Friday: 7:30 AM - 12:00 PM | Saturday & Sunday: Closed Toll Free (24/7): 800342 (800 DHA) | Outside UAE (24/7): +971-45991200
DoH Abu Dhabi	Abu Dhabi health regulatory questions, HIE/Malaffi policy direction	Phone: 800555
SDAIA (KSA)	PDPL questions, guidance, national data governance	"Communicate with SDAIA" portal.
NHIC (KSA)	National health information programs/coordination	NHIC contact form
SFDA (KSA)	Medical device / medical software regulatory questions	Phone: 0118806000-0112038222 For inquiries, please call centralized number: 19999 From outside of Saudi Arabia: 0118806060

Appendix C: UAE/KSA Regulatory (Cont.)

(AI Pilots in Healthcare)

D) How to use this appendix inside PB2 (simple rule)

- Use Step 4 checklists first (3.4 / 3.5). WIP: PB 2 v2
- If anything is RED (Block) → stop scoring/selection, adjust scope/safeguards, then re-check.
- Only proceed to Selection Gate when compliance is Green or Yellow-with-plan.

E) "Who to involve internally" (contact info placeholders)

Add your organization's names/emails next to these roles in your final PDF:

- **Data Protection Officer** / **Privacy Lead** (data mapping, vendor DPA, residency, retention)
- **Clinical Safety Lead** (escalation rules, guardrails, QA sampling)
- **IT/Security** (access control, audit logs, encryption, integrations)
- **Compliance/Legal** (approvals, regulator engagement if needed)
- **Pilot Owner** (cadence, weekly reporting, gate decisions)

Appendix D: Methodology Reference Sheet

01 The Impactful AI Method - Summary (PB2 in one view)

Goal: Choose the best-fit AI solution and prove value safely through a short pilot.

Inputs (from PB1): Challenge statement • Beneficiaries • Context • Constraints • Candidate solutions

Outputs: Impact Estimation Table (IET) • Comparison dashboard • Recommendation report • Pilot roadmap • Evidence pack

8-step evaluation flow (Section 3):

1. Org & Challenge Context (Template 3.1)
2. Stakeholder Mapping (Template 3.2)
3. Quantified Objectives (Template 3.3)
4. Regulatory & Compliance (Templates 3.4–3.5)
5. Define Solutions (Template 3.6)
6. Impact Estimation (IET) (Template 3.7)
7. Compare Solutions (Template 3.8)
8. Make the Decision (Template 3.9)

Pilot flow (Sections 4–5): Scope + success criteria (4.1–4.2) → increments (4.3) → roadmap (4.4) → readiness (4.5) → run pilot (5.2–5.5)

Appendix D: Methodology Reference Sheet (Cont.)

02 Tom Gilb Principles - Quick Reference

- **Quantify value**: Define success as baseline → target, not vague goals.
- **Measure what matters**: Use stakeholder-centric measures (beneficiaries, users, decision makers).
- **Compare alternatives**: Always compare options on the same yardstick.
- **Make uncertainty visible**: Use credibility/confidence so guesses aren't hidden.
- **Deliver in small steps** (Evo): Build and learn in 1–2 week increments.
- **Decide based on evidence**: End each increment with a decision: GO / PIVOT / NO-GO.
- **Build trust**: Safety + compliance guardrails are part of the design, not an add-on.

03 Decision Trees

A) Compliance decision tree (Step 4 gate)

- If **Block** (red) → Stop → adjust scope/safeguards → re-check
- If **Fix** (yellow) → proceed only with owner + due date
- If **Pass** (green) → proceed

B) Selection decision tree (end of Section 3)

- If best option has high impact + acceptable credibility + fits compliance → Select for pilot
- If impact is high but credibility is low → Pilot designed to validate uncertainty (tight scope)
- If compliance cannot be satisfied → Do not proceed (or redesign use case)

Appendix D: Methodology Reference Sheet (Cont.)

C) Increment decision tree (every 1–2 weeks)

- **GO** if: guardrails ok + metrics moving + adoption acceptable
- **PIVOT** if: guardrails ok but metrics/adoption not moving (change scope/workflow/features)
- **NO-GO** if: safety/privacy breach or no measurable value after agreed attempts

04 Flowchart-style “stop/go” gates (PB2 gates)

Input readiness → Method alignment → Compliance gate → Selection gate → Launch readiness → (repeat) Increment decisions → Close-out decision (scale/extend/stop)

05 Practical “minimum set” to run PB2

If you only do the essentials, do these:

- 3.3 quantified objectives (5–9) | 3.7 IET with credibility | 3.8 comparison dashboard
- 4.4 12-week roadmap (6 increments) | 5.2 + 5.5 weekly measurement + increment decisions