

SCALE, PIVOT, OR KILL? Your RetailFlow Project

AI Leadership & Project Management Masterclass

Curtin University

2025-10-31

SCALE, PIVOT, OR KILL?

Your RetailFlow AI Customer Service Pilot

Context: This morning you scoped an AI customer service pilot for RetailFlow. You defined success criteria (Go/No-Go thresholds) in Activity 1.

It's now 6 weeks later. The pilot data is in.

What happened during the pilot:

- **Week 1:** Data quality crisis (Activity 2) - you paused and fixed
- **Week 2:** Team resistance (Activity 2) - you addressed concerns
- **Week 4:** Executive pressure (Activity 2) - you managed expectations
- **Week 6:** Now you have results

Your task: Decide whether to SCALE, PIVOT, or KILL your pilot.

Time available:

- Review data: 5 minutes
- Group discussion: 15 minutes
- Prepare presentation: 5 minutes
- Present decision: 5 minutes per group

Total: 30 minutes

REVIEW YOUR ORIGINAL PLAN (Activity 1)

Pull out your pilot scoping worksheet from this morning.

What were your success criteria?

- 1.
- 2.
- 3.
- 4.
- 5.

What were your Go/No-Go thresholds?

- Accuracy: _____%
 - Customer Satisfaction: _____%
 - Response Time: < _____ hours
 - Cost per Query: \$ _____
 - Team Adoption: _____%
 - [Other criteria you defined]
-

REVIEW THE PILOT DATA

Access the data:

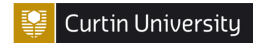
- Online: See companion website → Handouts → Pilot Data Dashboard
- Or printed handout: Pilot Data Summary

Key metrics after 6 weeks:

Metric	Your Target	Actual Result	Met?
Accuracy	____%	88%	/ /
Satisfaction	____%	82%	
Response Time	< ____ hrs	3.5 hours	
Cost per Query	\$ ____	\$14	
Team Adoption	____%	75%	
Escalations	< ____%	3.5%	

Legend: = Met target | = Missed target | = Close (within 5%)

QUICK REFERENCE: Other Pilot Data



Weekly Trends: Accuracy improving (85% → 88%), satisfaction stable (82%)

Query Type Performance:

- Order tracking: 94% accuracy
- Returns/refunds: 91% accuracy
- Product care/sizing: 78% accuracy (most complex)

Team Feedback:

- 3 of 5 agents using AI regularly
- 2 agents still hesitant

Cost Projection: At 60% scale → \$11.50/query (below \$12 target)

Key Insights:

- AI works best on routine queries (order tracking, returns)
 - Struggles with nuanced queries (product care, sizing)
 - Data scientist estimates 4 more weeks → 92% accuracy
-

KEY INSIGHT

Your criteria from Activity 1 are your decision anchor.

Different groups set different thresholds this morning:

- Group A might have set accuracy target at 85% → They MET it (88% actual)
- Group B might have set accuracy target at 92% → They MISSED it (88% actual)

Same data. Different decisions.

That's why defining criteria upfront matters. Your morning decision drives your afternoon decision.

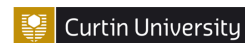
ANALYSE YOUR RESULTS

Step 1: Scorecard

How many of your success criteria did you hit?

_____ out of _____ criteria met

Step 2: Solvable vs. Fundamental Gaps



For each criterion you didn't meet, ask:

- Is this a **solvable gap** (needs more time/training)?
- Or a **fundamental problem** (the approach doesn't work)?

Examples:

- **Solvable:** "Accuracy is 88% vs 90% target - data scientist says 4 more weeks of training gets us to 92%"
- **Fundamental:** "Team adoption is 40% after 6 weeks - the team fundamentally doesn't trust AI"

Unmet Criterion	Solvable or Fundamental?	Why?
-----------------	--------------------------	------

Step 3: What Does the Data Tell You?

Positive signals:

-

Concerning signals:

-

Your key assumptions from Activity 1 - did they hold up?

-

MAKE YOUR DECISION

Our decision: SCALE PIVOT KILL

If SCALE:

What's your rationale?

Rollout plan:

Success criteria for full deployment:

Risks to manage:

If PIVOT:

What would you change?

Why pivot instead of scale or kill?

How long do you need?

What would success look like after the pivot?

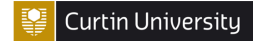
If KILL:

What's your rationale?

What did you learn?

How do you communicate this decision?

DEFEND YOUR DECISION



Prepare to present:

- Your decision (2 min)
- Your reasoning based on YOUR criteria from Activity 1 (3 min)
- How you'll respond to stakeholder pushback (2 min)

Expected questions:

- "You only hit 3 of 6 criteria - why scale?"
 - "Accuracy is 88% vs 90% target - that's close, why not pivot?"
 - "Customers are satisfied (82%) - why kill?"
-

REFLECTION

How did your decision compare to other groups?

What would change your mind?

What surprised you about your own data?

How does this feel different from the practice cases (4A)?

DEBRIEF QUESTIONS (After Presentations)

For the facilitator to ask:

1. **How many groups chose SCALE? PIVOT? KILL?**
 - Notice: Same data, different decisions. Why?
2. **What were the deciding factors for your group?**
 - Which criteria mattered most?
 - Which gaps were deal-breakers?
3. **How did your morning decisions (Activity 1) influence your afternoon decision?**
 - If you'd set different thresholds, would you decide differently?
4. **What does this teach about the importance of Go/No-Go criteria?**

- Why define them upfront?
- What happens if you don't have clear criteria?

5. How does this feel different from the practice cases (4A)?

- Practice cases had clear answers
- Your project is ambiguous (like real projects)
- You owned this decision from the start

The key learning:

Most real projects land in the gray zone. You set criteria, you gather data, you make the best call you can with 70% certainty. That's leadership.