Experiment Type Selector

Chapter 13 . Interrogate Phase

Decision Tree

What type of assumption are you testing?

User behavior	Wizard of Oz or Paper Prototype
Technical feasibility	Integration Spike or Data Quality Audit
Organizational alignment	Stakeholder Alignment Check
Solution accuracy	Shadow Mode

Is the system deployed?

Yes, in production	Shadow Mode (parallel comparison)
No, pre-build	Wizard of Oz, Paper Prototype, or Spike

Selection Steps

1. Identify the assumption category
2. Match to experiment type using the table
3. Check duration against constraints
4. Create Experiment Log entry before starting

Six Experiment Types

Wizard of Oz

Human performs the "AI" function behind the scenes. Tests whether users want the capability at all.

Best for: User assumptions. Duration: 1-2 weeks. Requires: Human performer.

Data Quality Audit

Sample real data and check against requirements. Reveals gaps before training or integration.

Best for: Technical assumptions. Duration: 2-5 days. Min sample: 100 records or 10%.

Integration Spike

Build the smallest possible working connection to external systems. APIs have undocumented behaviors.

Best for: Technical assumptions. Duration: 3-5 days. Use sandbox credentials.

Paper Prototype

Sketch the workflow and walk users through it manually. Tests if users understand the flow.

Best for: User/Solution assumptions. Duration: 1-3 days. Min users: 3-5.

Shadow Mode

AI runs in parallel while humans make actual decisions. Compare outputs to measure accuracy without risk.

Best for: Solution assumptions. Duration: 2-4 weeks. Requires: Deployed system.

Stakeholder Alignment Check

Interview stakeholders separately and compare answers. Surfaces hidden disagreements on objectives.

Best for: Organizational assumptions. Duration: 1-2 days.

COMMON PITFALLS

Using Shadow Mode when Wizard of Oz would be faster. Skipping prerequisites (WoZ and Shadow need human participants). Running a Data Quality Audit with only 10 records.