

# AI Validation Checklist

This guide is a companion to the blog [“When the Fourth Firefly Lives in Your Data”](#)



## Map Where AI Is Working

- ☐ List all the steps in your survey process where AI is involved (e.g., sentiment analysis, demographic coding, text categorization, data weighting).
- ☐ Mark which ones are visible (you can easily check the output) and which are invisible (behind-the-scenes, like weighting or classification).

## Establish a Baseline

- ☐ Before updates or changes, record how AI is classifying, coding, or weighting your data.
- ☐ Example: note the percentage of 'neutral' sentiment responses or how many respondents fall into each job category.

## Monitor for Unexpected Shifts

- ☐ Track whether outputs suddenly change from one wave of research to the next.
- ☐ Look out for: words or phrases coded differently, categories merged/split, or demographics shifting unexpectedly.

## Compare Against Human Judgment

- ☐ Every so often, have a human team member manually review a sample of AI outputs.
- ☐ Ask: do the AI's classifications, codes, or weights still make sense?

**Check Across Groups**

- ☐ Look at how AI treats different demographic groups.
- ☐ Are certain groups being coded or classified differently in ways that don't make sense?

**Keep Good Records**

- ☐ Save prompts, outputs, and version notes whenever you use AI in analysis.
- ☐ This creates a paper trail so you can track what changed if something looks off later.

**Build in Oversight**

- ☐ Decide in advance which AI decisions always need a human check.
- ☐ Example: if the distribution of a key demographic shifts by more than a few percentage points, review it before moving forward.

**Decide: Keep It or Correct It**

- ☐ Not every change is bad—sometimes AI picks up on real-world shifts (e.g., new ways people talk about 'value' and 'price').
- ☐ When you spot a change, ask: Is this an error, or an insight worth keeping?

**Train Your People and AI**

- ☐ Make sure your staff know what to look for: sudden jumps in satisfaction, odd category groupings, or demographic inconsistencies.
- ☐ Encourage them to ask, "Could this be a fourth firefly?"
- ☐ Review why the AI makes mistakes, give it clearer context or prompts, and refine how you use it so it improves over time, reducing rework on future projects.

**Share What You Learn**

- ☐ Talk with peers and vendors about your findings—both problems and successes.
- ☐ The more we share, the better our field becomes at handling AI responsibly.

## Quick Start for Any Team

- **This month:** Audit where AI is being used and set a baseline.
- **Each quarter:** Manually review a sample and check demographic consistency.
- **Ongoing:** Document, monitor, and share what you find.