The Move-On Check

GenAI makes the fast part easy. I'd write down what good looks like for one workflow, then score its output against that before it moves on.

Early in my time at a very large retailer, I worked on test data for its path to production. Teams needed millions of records in minutes, and the real work was getting them to look enough like production that they could confidently move on to the next step. This sheet is that check for one AI workflow.

Pick one workflow

The workflow _____

What the next step does with its output _____

Model it runs on today _____

Write down what good looks like

A person reviewing it is a fine way to score. Write down what they look for, so someone else could do it the same way.

CHECK	WHAT GOOD LOOKS LIKE	HOW WE SCORE IT	WHO OWNS IT
Example: Right shape	*Every summary names the ticket and the next step*	*A script checks both, on every one*	*(a name)*
Hold: Right facts	*Looks accurate*	*Someone skims a few*	
Right shape: the format and fields the next step expects			
Right facts: it matches a source we trust			
Honors its contract: the next step takes it without anyone fixing it by hand			
Holds up at volume: the same pass rate on a full day of real inputs			
Safe to pass on: no customer data or secrets where they don't belong			
Cost: dollars per run (input and output tokens times their prices) against a written budget			

PASS OR HOLD

This grades the sheet itself. Pass if someone new to the team could score tomorrow's output with only this sheet and get the same answer you would. Hold if any row says "looks fine," has no owner, has nothing a stranger or a script could check, or is a Cost row with no budget number, like the Hold example above.

Score one real batch

Take the last batch the workflow produced and run every row against it.

Outputs scored _____

Passed every row _____

Failed, and on which row _____

Minutes per output to score _____

Dollars per run, this batch _____

Decide this week

The row we couldn't fill in _____

Who will fill it in (a name) _____

Date we score the next batch _____

What this sheet is and is not. A written check can still be wrong, so I'd look at it again whenever the model, the prompt or the inputs change. It scores the output. If you're worried about the model itself going away, I'd pair it with The Model Continuity Checklist.

I sell AI software. You don't need anything from me to fill this in.

Ben Pickett

© 2026 Enterprise Field Notes