

The Intent Toolkit

THREE OBJECTS · ONE HABIT

A 60-second diagnostic, three templates, and a worked example. The job, the truth, the lesson — enough to stop losing the user in one real piece of work, this week.

JEFF NICHOLSON · FROM THE BOOK INTENT · NOTAGILE.AI

You Shipped It. It Works. It Still Missed.

Somewhere in your company, a user said something sharp and specific: what they were trying to do, what kept breaking, what it cost them. Then the machine took over. The quote became a summary. The summary became a persona. The persona became a requirement, the requirement became tickets, and six months later something shipped that technically works — and the user is still exporting to a spreadsheet, because at least the spreadsheet behaves predictably.

Nobody did anything wrong. That is the uncomfortable part. Every handoff was professional, every artifact was polished, every meeting ended aligned. The system worked exactly as designed — and the design loses the user. Every handoff compresses context, and compression removes precisely the things that mattered: the urgency, the stakes, the workaround, the sharp sentence.

The fix is not more research, more process, or better intentions. It is mechanics: a small set of artifacts that make context harder to lose than to keep. This toolkit gives you the three I install first with every team, plus a 60-second test that tells you how much you're currently paying in what I call the handoff tax.

HOW TO USE THIS

- Take the leak test on the next page. Be literal — 'sort of' is a No.
- Then run the three templates against one real, live piece of work. Not a hypothetical. The disagreements that surface while filling them in are the actual work.
- Templates are fillable — use them digitally or print them. Copy one per bet.

These pages are drawn from Intent, the book — where each template gets a full chapter of reasoning, failure modes, and field examples.

60 SECONDS, HONESTLY

Where Are You Losing the User?

Answer for one real, current initiative. Every No is a place context leaks out of your process.

YES

NO

Every active bet has at least one receipt attached — a clip, ticket cluster, trace, or verbatim quote. A summary is not a receipt.

Anyone on the team can name the real moment in the user's day this work is supposed to improve.

You can retrieve a major decision from six months ago and see what was believed at the time, and why.

Your key assumptions are written down somewhere visible, with what would disprove each one.

Decisions carry a check-in date. (A decision without one is just a belief.)

The people building have direct, recurring contact with real users — not just personas and dashboards.

'Done' includes observing what the release did to the user's day — not just shipping it.

The plan is allowed to lose to evidence without anyone paying a social price.

YOUR YES COUNT (OUT OF 8)

7-8 — Your machine preserves context. Rare. The book will sharpen what you already do.

4-6 — You are paying a quiet tax: rework, churn, support volume, products that grow without getting better. The three templates that follow are where to start.

0-3 — Your artifacts exist to hand off, not to build. Every quarter this runs, the gap between what users say and what ships gets wider. Start today, start small.

01 Intent Spec

The backbone artifact. A compact model of a job with enough texture that a reasonable person can judge whether you hit it.

WHO

e.g., Sue, staff accountant, 8 yrs exp, working alone during month-end close

WHEN

e.g., Last 3 days of every month. Hard deadline: books close by the 1st.

JOB

e.g., Reconcile 14 accounts, investigate mismatches, get controller sign-off

CURRENT STATE

e.g., Exports CSVs from 3 systems, pastes into a personal spreadsheet, hunts row by row

BETTER STATE

e.g., "I want to see what doesn't match and why. Not spend 2 hours finding problems."

CONSTRAINTS

e.g., SOX: every adjustment needs audit trail. Controller approves over \$5K.

WHAT DONE MEANS

e.g., Time to identify mismatches drops from 4 hours to 30 min. Zero post-close surprises.

04 Evidence Ladder

Default insert for any decision record. Maps how confident you should be based on evidence type.

RUNG 1: BELIEFS

e.g., "Accountants hate our software." "A better dashboard would fix it."

RUNG 2: STATEMENTS

e.g., Sue: "I spend half my close time finding problems, not solving them."

RUNG 3: BEHAVIOR

e.g., Session traces: Sue exports same CSV 3.2x per cycle. Time to first mismatch: 2.1 hours.

RUNG 4: OUTCOMES

e.g., After shipping ranker: time to first mismatch dropped from 2.1 hrs to 18 min. Zero missed.

CURRENT RUNG

e.g., 3 -- behavioral traces and time data, but haven't shipped and measured outcomes yet.

02 Memory Card

CH 11
INTENT TOOLKIT

The organizational second mind. Records every important bet so the next team doesn't start from zero.

JOB

e.g., Month-end account reconciliation

CURRENT BET

e.g., Auto-surface mismatches with severity ranking. Shaped by SOX audit trail reqs.

APPROACH & ALTERNATIVES

e.g., Reconciliation dashboard. Rejected: auto-reconciliation (no trust), spreadsheet add-in (too fragile).

WHAT HAPPENED

e.g., Prior attempt Q2: built a diff tool. Sue used it once, went back to her spreadsheet. Too noisy.

LEARNINGS

e.g., Showing all mismatches without context is worse than her notebook. She needs severity + likely cause.

INTENT SPEC · COMPLETE, WITH ATTACHMENTS

Month-end budget reconciliation under deadline pressure

Owner: the Product Builder on the budgeting job · Pinned receipt unedited · Last confirmed against a live session: this month

The spec

WHO

Sue, staff accountant at a mid-size company with three cost centers. Ten years of experience. Uses the budgeting system and Excel interchangeably because neither is complete.

WHEN

Last week of every month. Time pressure is real: the controller reviews submissions by 2pm on the first business day of the new month. Sue typically starts two days before close and works evenings if the numbers don't reconcile.

JOB

Verify that posted charges match expected budget amounts across all three cost centers, reconcile discrepancies, and produce a summary the controller can approve without questions.

CURRENT STATE

Sue switches between the system and a spreadsheet. She re-enters data at least three times because the system doesn't carry forward values she already confirmed. She doesn't trust the system's totals, so she calculates them independently. Time to complete: 3-4 hours on a clean month, 6+ hours when discrepancies are complex. She frequently exports to Excel because "at least it does what I tell it."

BETTER STATE

Sue can see all charges in one place, with clear flags for mismatches. Reconciliation takes under 90 minutes. She trusts the numbers enough to submit without a parallel spreadsheet. The controller gets a summary with evidence attached and approves without follow-up questions.

CONSTRAINTS

Finance compliance requires an audit trail for every adjustment. The controller must approve any manual override above \$5,000. GL mapping rules vary by cost center and change quarterly. The system cannot silently auto-correct entries that affect the final submission.

WHAT DONE MEANS

Budgeting closes in under 90 minutes. No spreadsheet detour. No re-entry. Support touches during close drop to near zero. Controller approval happens same day without follow-up.

Pinned receipt — the sharp line, unedited

“I missed my grandson’s game because I was wrestling the budget close again. I don’t have time for this.”

Support call, close week. Linked to the raw recording. This line stays in the header so it cannot be sanded into “reduce confusion.”

Receipts — classified by evidence rung

RUNG 3 · BEHAVIOR

Session trace: 14 back-and-forth navigations and 23 copy-paste events per close cycle; export to Excel mid-task. Average close: 3–4 hours.

RUNG 3 · BEHAVIOR

Ticket cluster #4891: support volume 3x normal during close week, all circling the same complaint — the totals cannot be trusted.

RUNG 2 · STATED

Three other accountants report the same re-entry friction in interviews.

Rule of thumb: the rung sizes the bet. Rung 3 evidence earns a loop-sized release, not a quarter-long project.

Active decision — loop 3

Bet: reconciliation view with mismatch flags and carry-forward values, prioritized over another visual redesign.

Why: behavior receipts show the friction is trust and re-entry, not aesthetics. Bet 1 already disproved the aesthetic theory.

Wrong if: export behavior doesn’t drop by half within 30 days, or Sue stops trusting the flags and re-checks everything manually.

Check-in: two weeks after release. Owner named. A decision without a check-in date is just a belief.

Memory — what past bets taught

Bet 1 (dashboard redesign): adoption was high, time-to-complete didn’t move, accountants kept exporting. Learning: the problem was never aesthetics. The clean design made the wrong thing prettier.

Open question: GL mapping rules change quarterly and still require manual reconfiguration. Candidate for the next bet.

How the pieces connect: the spec defines what Sue needs; the pinned receipt keeps the stakes sharp; the rungs say how confident to be; the decision names the bet, its kill condition, and its check-in; the memory keeps the next team from starting at zero. One job, one living record. From the book Intent by Jeff Nicholson · templates at notagile.ai

If These Pages Hit a Nerve

These three templates are the entry point. The full system — the thirteen polite distortions that crush context, the nine sections of the handoff tax, reality loops, orchestration, metrics that pass what I call the resident test, and the organizational memory that stops your company relearning the same lesson every eighteen months — is in the book.

GO DEEPER

INTENT — the book. The full argument, the frameworks, and the 90-day playbook for running your first intent-first pilot.

OR BRING ME IN

I work hands-on with a small number of teams: a scoped engagement that takes one real, painful job and moves it through the full system in roughly ninety days — measured in the user's terms, not in features shipped. The leak test on page 3 is the conversation starter: send me your score and the job you'd pick.

CONTACT

Jeff Nicholson

jeffnich00@gmail.com

Author of Intent. If your team is losing the user in translation, that's the work I do.