Punchcard

AI + Jobs to Be Done

A practical framework for delegating real work to AI agents

Daniel Alberson · Founder, Punchcard

Larson Summit · November 5, 2026

From chat messages to delivered work

Agents can work alongside you in real time, or take actions on your behalf.

CHAT

You ask, it answers

Helpful replies. You still do the work, one step at a time.

You own every step

→

AGENT · ALONGSIDE YOU

It works, you steer

Works in real time while you watch, answer questions, and correct course.

You collaborate

→

AGENT · ON YOUR BEHALF

It works, you review

Works on its own, then brings back finished work and asks before acting.

You approve and accept

The more an agent does on its own, the clearer the job has to be.

LEARNING OBJECTIVES

What you'll leave with

Habits that work with any AI tool, starting Monday.

1. Describe a job so an agent can finish it — and you can tell when it's done
2. Decide what an agent may do on its own and what needs your approval
3. Check AI output against its source evidence before you accept it
4. Evaluate AI software on your own work before you buy

TODAY'S PATH

01 The five questions

02 Personal agents

03 Agents for teams

04 Trusting AI output in Excel

05 Buying AI software

Five questions before you delegate

Ask them in order. Weak answers early show up as rework later.

1

JOB

What outcome do I want?

Describe the finished work and what "done" means.

2

CONTEXT

What does it need to know and access?

Documents, tools, history, who owns what. Nothing extra.

3

BOUNDARIES

What can it do on its own?

Prepare freely. Approve consequential actions. Name the never-dos.

4

EVIDENCE

How will we know the work is right?

Sources, completeness, exceptions, and stated uncertainty.

5

OWNERSHIP

Who reviews, approves, and remains accountable?

A named person, with the evidence to decide.

Define done, then scope the access

A prompt asks for a response. A job describes finished work.

1 JOB
What outcome do I want?

A PROMPT

"Summarize these invoices."

A JOB

"Agree each sampled expense to its invoice and list every exception with the document and page."

Done means: the deliverable, the quality bar, the deadline, and the format.

2 CONTEXT
What does it need to know and access?

GIVE IT

- Relevant documents and prior work
- The tools and systems it must use
- History: what was already decided
- Who is responsible for what

Limit it: only the access this job requires — read-only where reading is enough.

Separate preparing work from acting

Decide in advance what needs a human yes — and what is never allowed.

On its own

- Research, read, and compare
- Draft replies and workpapers
- Organize files and flag issues

Easy to undo

With approval

- Send to a client or colleague
- Spend money or submit a form
- Change a system of record
- Commit anyone's time

Consequential or hard to undo

Never

- Share data beyond the job
- Sign, certify, or conclude for you
- Delete records
- Bypass a control

Off-limits, whatever it's told

Ask for help when sources conflict, information is missing, or the next step crosses a boundary.

How will we know the work is right?

Check the work the way a good reviewer would — not just the summary.

Inspect sources
Open the document behind each conclusion. Does it say what the result claims?

Check completeness
Was every item covered, including the hard ones it might have skipped?

Test exceptions
Look hardest at what was flagged — and at what should have been.

Separate support from uncertainty
Which conclusions are backed by evidence, and which are inferred or missing?

A citation is a pointer, not a proof

It shows where to look. You still confirm the source says what the conclusion claims — and that it's the right source.

Who reviews, approves, and owns it?

Three different human roles. Know which one you're playing.

DURING THE WORK

Collaborate

Steer, answer questions, and correct course as the agent works.

BEFORE AN ACTION

Approve an action

Say yes or no before something consequential happens: send, buy, post.

WHEN IT'S DONE

Accept the result

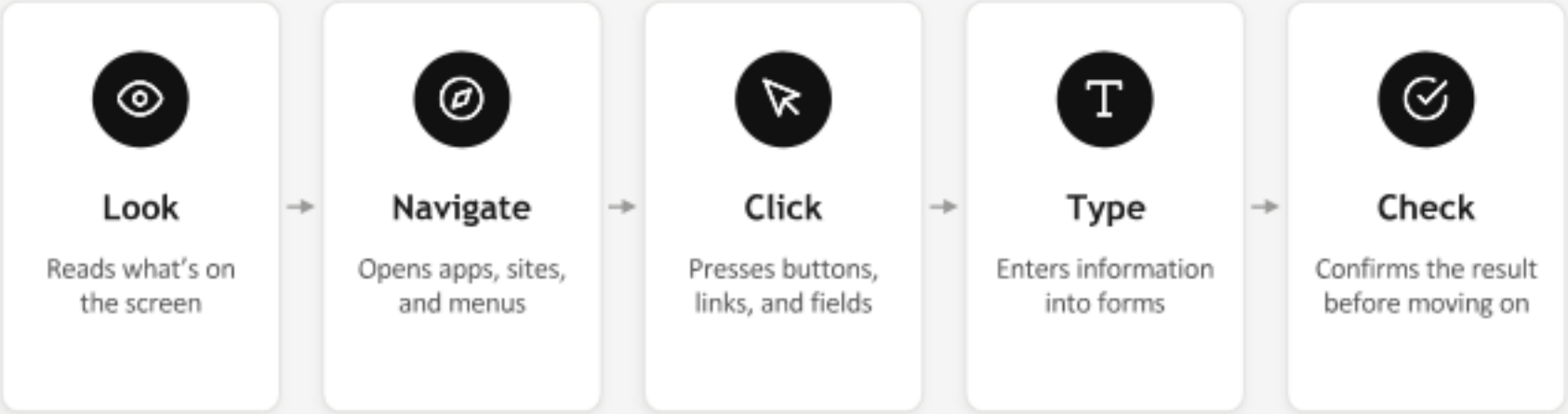
Sign off on the finished work and own the outcome.

Approval needs evidence and context, not a rubber stamp.

If the approver can't see what the agent saw and did, they aren't approving — they're hoping.

How an agent uses a computer

Computer use: it works through the same screens you do, one visible step at a time.



Because every step is on screen, you can watch it work, pause it, or take over.

Email and meeting follow-ups

The useful version depends on one detail most people leave out.

Email triage

JOB

A draft waiting for every email that needs my reply.

THE TRAP

Answered threads keep coming back as “open.”

THE FIX

Check later replies and Sent Items before calling anything open.

BOUNDARY

Drafts only. I review and send.

Meeting follow-ups

JOB

Today’s meeting commitments, captured and moving.

THE TRAP

Everything said in the meeting lands on my to-do list.

THE FIX

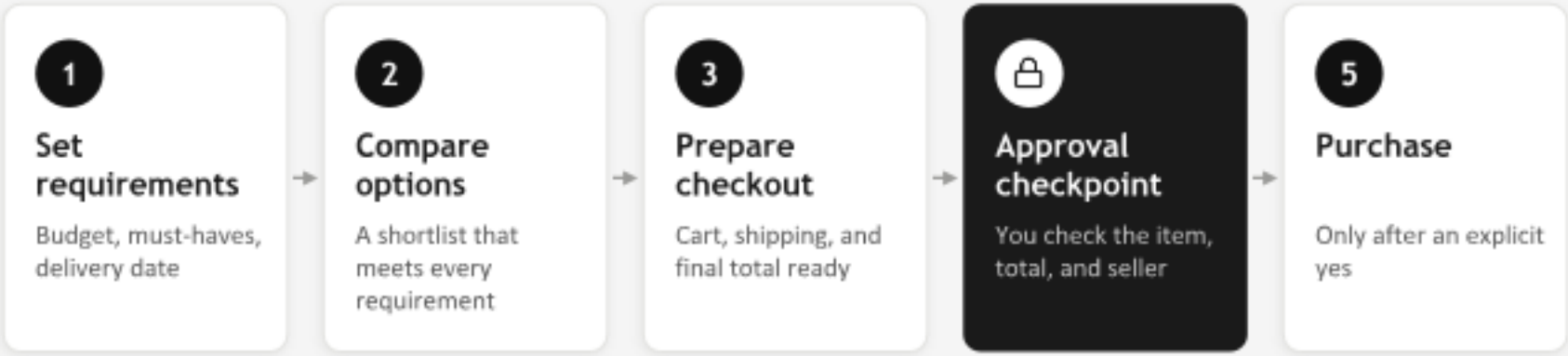
Separate my commitments from other people’s. Do mine; track theirs.

BOUNDARY

Drafts reminders for approval. Never makes new commitments.

Shopping: compare, prepare, then stop

The agent does the legwork. The purchase waits for your explicit yes.



JOB

Right item, within budget, by the date

CONTEXT

Requirements, budget, address, preferred sellers

BOUNDARIES

No purchase without my approval

EVIDENCE

Links, and how each option meets the requirements

OWNERSHIP

I approve the spend and own the choice

Demo: an agent at the keyboard

Watch where it checks its own work — and where it stops for approval.

RUN OF SHOW

1. Email: give the job, drafts only
2. Watch it check later replies and Sent Items
3. Shopping: requirements, budget, compare
4. Stop at checkout for approval



PLACEHOLDER

Live screen: switch to the demo computer

Backup: insert a recording or screenshots from your own rehearsal before the event

Follow one commitment through a team

From a client meeting: “We’ll send the updated reserve rollforward by Friday.”

AGENT 1

Captured

Logs the commitment from the meeting notes

PERSON 2

Owner assigned

Manager names one accountable owner

AGENT 3

Prepared

Gathers files; drafts the rollforward and email

AGENT 4

Handed off

Passes it to the owner with status and open items

PERSON 5

Reviewed

Owner checks the evidence and edits

PERSON 6

Completed

Owner sends; status closes for everyone

Preparing a follow-up is not authority to send it

The agent drafts. The owner sends, sets the date, and makes the promise. The agent never commits another person’s time.

Same five questions, for a team

What changes when the work is shared.

QUESTION	FOR ONE PERSON	FOR A TEAM
Job	My outcome	A shared outcome with a visible due date
Context	My inbox and files	Shared workspace, prior work, and who owns what
Boundaries	What I'm willing to approve	Permissions by role; can't commit another person
Evidence	I check the work	Status and open items visible to everyone
Ownership	Me	One named accountable person — not "the team"

Expense testing: the prepared result

Fictional sample — four expenses, with invoices given to the agent.

ITEM	VENDOR	GL AMOUNT	DATE	SUPPORT FOUND	PREPARED RESULT
E-01	Tri-County Auto Glass	$1,840.00	03/04/2026	Invoice TC-2207	AGREES
E-02	Lakeside Adjusting Services	$3,275.00	03/11/2026	Invoice LA-0419	AMOUNT MISMATCH
E-03	Meridian Print & Mail	$612.40	03/19/2026	None located	MISSING SUPPORT
E-04	Northgate Records Storage	$945.00	03/28/2026	Invoice NG-1183	AGREES

Prepared is not accepted. Two rows say “Agrees” — would you sign off on them yet?

Inspect evidence before you accept

Open the source behind each result — including the ones marked "Agrees."

FICTIONAL INVOICE

Lakeside Adjusting Services

Invoice LA-0419
Date 03/11/2026
Field adjusting services, March

Amount due

$3,725.00

The cited source for E-02's amount

E-02 · WHAT TO CHECK

Vendor MATCH

Matches the GL vendor

Date MATCH

Inside the test period

Amount MISMATCH

GL $3,275 vs. invoice $3,725

Support PRESENT

Invoice on file

YOUR TURN

What would you verify before accepting this test?

- Right document, not a duplicate?
- Typo, partial payment, or real error?
- Was E-03's invoice missing — or never provided?
- Did it really open the "Agrees" invoices?

Demo: review the test in Excel

Same fictional sample, live in the workbook.

RUN OF SHOW

1. Open the workbook; run the first pass
2. Click a result to open its cited source
3. Investigate E-02 and E-03; flag follow-ups
4. Validate supported rows; sign off when complete



SCREENSHOT PLACEHOLDER 1

Prepared grid, AI-populated cells marked

Replace with your own capture before the event



SCREENSHOT PLACEHOLDER 2

Invoice LA-0419 opened from its cell

Replace with your own capture before the event

Where the human stays in charge

The agent prepares the test. You decide what it means.

1 COLLABORATING

Guide

Set the procedure and the pass/fail rule. Answer its questions when it's unsure.

2 COLLABORATING

Investigate

Follow up on E-02's $450 difference and E-03's missing invoice. It flags; you decide.

3 ACCEPTING

Approve

Validate each row against its source. Sign off once every exception is resolved or documented.

Collaborating is not accepting Sign-off is a separate, deliberate step — with the evidence in front of you.

Buying AI software: test your job

Run the five questions on the product before you sign.

JOB

Test a specific job

Use representative inputs and your hardest cases, not the vendor's demo data.

CONTEXT

Check what it needs

Which documents, systems, and integrations — and the effort to connect them.

BOUNDARIES

Verify permissions

Least access by default, and approval controls you can set and audit.

EVIDENCE

Evaluate the evidence

Citation quality, errors on your cases, review effort, and recovery from mistakes.

OWNERSHIP

Assign an owner

One named owner, and a small pilot with baseline, success, and stop criteria.

VALUE

Measure net value

Time saved after review and corrections, plus implementation effort and total cost.

Run a small pilot before you commit

Decide what success and failure look like before you start.

1 BEFORE

Baseline

- Time per item today
- Error and rework rate
- Reviewer time

→

2 DURING

Pilot

- One job, one owner, real inputs
- Log review and correction time
- Track errors and how they were caught

→

3 AFTER

Decide

- Success criteria met → expand
- Stop criteria hit → stop
- Unclear → adjust and rerun

Net time saved = time today − (human time to set up, review, and correct)
Then weigh implementation effort and total cost. Use your own baseline, not a vendor's ROI figure.

TAKE IT BACK TO WORK

The framework on one page

Answer all five before you hand an agent a job.

1	JOB	What outcome do I want?	Finished work described · "done" defined · deadline and format
2	CONTEXT	What does it need to know and access?	Documents, tools, history, who owns what · nothing extra
3	BOUNDARIES	What can it do on its own?	Prepare vs. act · what needs approval · when to ask · never-dos
4	EVIDENCE	How will we know the work is right?	Sources opened · complete · exceptions tested · uncertainty stated
5	OWNERSHIP	Who reviews, approves, and remains accountable?	Named person · approval with evidence · final acceptance

Punchcard

People can delegate more useful work when the outcome, access, boundaries, evidence, and human responsibilities are clear.

JOB · CONTEXT · BOUNDARIES · EVIDENCE · OWNERSHIP

Questions & answers

Daniel Alberson · Founder, Punchcard · punchcard.com