AI AGENTS · PDF TEMPLATE

AI Agent Test Plan

Fill this in before an AI agent touches real work. Define success first, test with real cases, and launch in stages.

6 SECTIONS · FILL-IN TEMPLATE · PRINT AND FILL



01 Scope

FILL IN

The job in one sentence

Inputs and tools the agent may use

What "done" looks like

Owner of the agent and its results

02 Limits and approvals

5 CHECKS

- **Permissions:** only the systems and actions the job needs, with the narrowest access available.
- **Escalation:** when the agent is unsure, it stops and asks instead of guessing.
- **Stop condition:** a cap on steps per run.
- **Approvals:** anything involving money, commitments to customers or deleted data waits for a person.
- **Logging:** every run records inputs, tool calls and outputs, so a failure can be explained.

03 Test set built from real work

FILL IN

CASE TYPE	EXAMPLE FROM YOUR WORK	EXPECTED RESULT	PASS?
Ordinary case			
Missing or conflicting information			
Out-of-scope request			
Tool failure			
Hostile content Instructions hidden in input			

04 How each result is graded

4 CHECKS

- **Code checks:** for anything with a fixed right answer (format, fields, totals).
- **Human review:** for the cases that matter most and for calibrating the other two.
- **Model graders:** for judgment calls, against a written rubric, spot-checked by a person.
- **Consistency:** run each case several times; one good run is not proof.

05 Compare with how the work is done today

FILL IN

MEASURE	TODAY	WITH THE AGENT
Share of cases completed correctly		
Time to a usable result		
Correction time per case		
Boundary violations		
Cost per case		

06 Launch gate

SIGN OFF

1 STAGE 1
Shadow mode: the agent drafts and a person does the work.

2 STAGE 2
A person approves every action before it happens.

3 STAGE 3
Approvals only for the risky actions listed above.

- **Retest** after any change to the model, instructions or tools; keep a regression set.
- **Sign-off:** name and date of the person who approved each stage.

STAGE	APPROVED BY	DATE
Stage 1		
Stage 2		
Stage 3		

NEXT STEP

AI agent services

www.doitwithfriday.com/services/agents-skills/

Build with Friday →

BASED ON FRIDAY FIELD NOTES

www.doitwithfriday.com/blog/how-to-test-ai-agents/

www.doitwithfriday.com/blog/ai-agents-for-small-business/

General guidance, not legal or security advice for your specific setup. Questions: hello@doitwithfriday.com