

AI Agent Readiness and Risk Checklist

Decide what deserves to be built. Then build only what passes the test.

Score one real workflow. Surface the hard stop conditions. Define the smallest pilot that can produce evidence without giving the system more authority than it needs.

WHAT YOU WILL LEAVE WITH

- ☐ An evidence-based readiness score across eight operating domains
- ☐ A list of unresolved risks that must be closed before production
- ☐ A bounded pilot canvas with owners, approvals, tests, and fallback
- ☐ A named next action and decision date

HOW TO USE IT

One workflow. The people who run it. Current evidence.

Bring the process owner, one person who performs the work, and the security or IT owner when sensitive data or production access is involved. Do not score a department-wide AI transformation idea.

A confident opinion is not evidence. Any unresolved stop condition overrides the numerical score.

SCORING

0	UNKNOWN OR ABSENT	No evidence, no owner, or not yet addressed.
1	PARTIAL	Informal, narrow, untested, or inconsistently controlled.
2	READY FOR A BOUNDED TEST	Documented, owned, and supported by current evidence.

FIRST QUESTION

Do you need an agent?

Choose the simplest system that can produce the outcome. An agent earns its extra control burden only when the next step depends on what the system finds and it must use tools along the way.

STARTING POINT	USUALLY FITS WHEN	CONTROL POSTURE
Assistant	A person requests analysis or a draft and decides what happens next.	Output stays with a human. No independent action.
Workflow automation	Steps and rules are stable and exceptions are limited.	Deterministic rules, validation, and exception routing.
Agent	The task is multi-step and tool use changes the next step.	Bounded authority, observable actions, failure behavior, and human gates.

- ☐ We can name the business problem without naming a model or vendor.
- ☐ The workflow is recurring or valuable enough to justify ongoing oversight.
- ☐ The trigger, completion point, and accountable owner are identifiable.
- ☐ The outcome can be checked by something stronger than 'looks good.'
- ☐ A mistake can be contained, detected, and corrected.

If most answers are no, simplify or reframe the use case before scoring it.

READINESS REVIEW

Domains 1 and 2

Score each statement 0, 1, or 2. Write the evidence and gaps in the scorecard later.

1. Problem and value

READINESS STATEMENT	EVIDENCE TO INSPECT	SCORE
One specific recurring workflow has a defined trigger and completion point.	Current procedure, ticket path, observation	☐ 0 ☐ 1 ☐ 2
The present cost, delay, risk, or failure rate has a baseline.	Time sample, queue, rework, error, or cost data	☐ 0 ☐ 1 ☐ 2
The desired outcome is measurable in operational terms.	Target time, quality, throughput, or risk measure	☐ 0 ☐ 1 ☐ 2
An agent adds value beyond a form, rule, integration, or human-facing assistant.	Alternatives considered and rejected with reasons	☐ 0 ☐ 1 ☐ 2
Expected value justifies build, review, model, monitoring, and maintenance costs.	Rough total-cost estimate and value range	☐ 0 ☐ 1 ☐ 2

DOMAIN SCORE / 10

2. Workflow clarity

READINESS STATEMENT	EVIDENCE TO INSPECT	SCORE
The actual workflow has been observed, not reconstructed from policy alone.	Walkthrough, screen recording, or work sample	☐ 0 ☐ 1 ☐ 2
Owners, systems, decisions, and handoffs are mapped end to end.	Current-state map with named owners	☐ 0 ☐ 1 ☐ 2
Common exceptions and their frequency are known.	Exception log, tickets, operator interview	☐ 0 ☐ 1 ☐ 2
Decisions that require human judgment are explicitly named.	Approval rules and examples	☐ 0 ☐ 1 ☐ 2
A workable manual fallback exists when the system is unavailable or uncertain.	Runbook, queue, or fallback procedure	☐ 0 ☐ 1 ☐ 2

DOMAIN SCORE / 10

READINESS REVIEW

Domains 3 and 4

Score each statement 0, 1, or 2. Write the evidence and gaps in the scorecard later.

3. Data and systems

READINESS STATEMENT	EVIDENCE TO INSPECT	SCORE
Every required data source and its owner are known.	Data inventory and ownership list	☐ 0 ☐ 1 ☐ 2
Required data is accessible, timely, and good enough for the intended decision.	Sample records and quality checks	☐ 0 ☐ 1 ☐ 2
Data classes and restrictions are identified before sending data to a model or vendor.	Classification, contract, and policy review	☐ 0 ☐ 1 ☐ 2
Integration methods and a safe test path are available.	API or export docs, sandbox, test account	☐ 0 ☐ 1 ☐ 2
Retention, deletion, location, and export rules are specified.	Data-handling decision and vendor settings	☐ 0 ☐ 1 ☐ 2

DOMAIN SCORE

____ / 10

4. Authority and access

READINESS STATEMENT	EVIDENCE TO INSPECT	SCORE
Allowed and prohibited actions are explicit.	Action inventory and boundary statement	☐ 0 ☐ 1 ☐ 2
Access uses named identities and the minimum required permission.	Role or scope configuration	☐ 0 ☐ 1 ☐ 2
Credentials are separated, revocable, and not exposed in prompts or logs.	Secret-storage and revocation procedure	☐ 0 ☐ 1 ☐ 2
Financial, legal, customer-facing, privileged, and irreversible actions have human gates.	Approval matrix and retained approval evidence	☐ 0 ☐ 1 ☐ 2
Important reads, decisions, writes, and failures are attributable and reviewable.	Audit log with identity, time, input, and outcome	☐ 0 ☐ 1 ☐ 2

DOMAIN SCORE

____ / 10

READINESS REVIEW

Domains 5 and 6

Score each statement 0, 1, or 2. Write the evidence and gaps in the scorecard later.

5. Reliability and evaluation

READINESS STATEMENT	EVIDENCE TO INSPECT	SCORE
A representative test set includes routine, edge, ambiguous, and adversarial cases.	Versioned test cases drawn from real work	☐ 0 ☐ 1 ☐ 2
Acceptance criteria measure the business outcome, not only fluent output.	Quality, time, error, or reconciliation threshold	☐ 0 ☐ 1 ☐ 2
False positives and false negatives are evaluated separately.	Error review by consequence and frequency	☐ 0 ☐ 1 ☐ 2
High-impact outputs have evidence, validation, reconciliation, or structured checks.	Citations, source links, schemas, totals, rules	☐ 0 ☐ 1 ☐ 2
Failures, drift, cost, and unusual behavior have monitoring and a named responder.	Alerts, review cadence, owner, escalation path	☐ 0 ☐ 1 ☐ 2

DOMAIN SCORE ____ / 10

6. Security, privacy, and compliance

READINESS STATEMENT	EVIDENCE TO INSPECT	SCORE
External instructions, documents, and messages are treated as untrusted input.	Injection-resistant design and test cases	☐ 0 ☐ 1 ☐ 2
Secrets and sensitive records are minimized and redacted from prompts and logs.	Data-flow review and sample logs	☐ 0 ☐ 1 ☐ 2
Vendors, regions, subprocessors, and data uses have the required approval.	Contracts, security review, internal approval	☐ 0 ☐ 1 ☐ 2
Data leakage, cross-tenant access, prompt injection, and tool misuse are tested.	Threat scenarios and test evidence	☐ 0 ☐ 1 ☐ 2
An incident can be contained, evidence preserved, and the right people notified.	Stop procedure, contacts, and response path	☐ 0 ☐ 1 ☐ 2

DOMAIN SCORE ____ / 10

READINESS REVIEW

Domains 7 and 8

Score each statement 0, 1, or 2. Write the evidence and gaps in the scorecard later.

7. Human oversight and exceptions

READINESS STATEMENT	EVIDENCE TO INSPECT	SCORE
A specific reviewer and backup are assigned for consequential or uncertain cases.	Named roles and coverage schedule	☐ 0 ☐ 1 ☐ 2
Reviewers receive the source, rationale, confidence signals, and relevant context.	Review interface or approval packet	☐ 0 ☐ 1 ☐ 2
Review volume is manageable during normal and peak periods.	Forecast volume and sampled review time	☐ 0 ☐ 1 ☐ 2
Timeouts, ambiguity, conflicts, and out-of-scope cases route to a defined path.	Exception matrix and queue behavior	☐ 0 ☐ 1 ☐ 2
Authorized people can pause, override, correct, and safely retry the system.	Control test and retained event record	☐ 0 ☐ 1 ☐ 2

DOMAIN SCORE

____ / 10

8. Ownership and operating model

READINESS STATEMENT	EVIDENCE TO INSPECT	SCORE
The workflow has a business owner and the system has a technical owner.	Named ownership and escalation contacts	☐ 0 ☐ 1 ☐ 2
A runbook covers routine operation, exceptions, changes, and recovery.	Current operating documentation	☐ 0 ☐ 1 ☐ 2
Model, vendor, review, and integration costs are capped and monitored.	Budget, limits, and cost dashboard	☐ 0 ☐ 1 ☐ 2
A review date and evidence-based continue, change, or stop criteria are set.	Decision calendar and pilot exit criteria	☐ 0 ☐ 1 ☐ 2
The system can be decommissioned with data exported and access revoked.	Offboarding, retention, and revocation plan	☐ 0 ☐ 1 ☐ 2

DOMAIN SCORE

____ / 10

STOP CONDITIONS

A high score cannot make an unsafe boundary acceptable.

Do not connect the system to production while any applicable condition remains unresolved.

STOP CONDITION	MINIMUM RESOLUTION BEFORE A PILOT
No accountable process owner	Name an owner who can decide scope, approve exceptions, and stop the system.
Success and acceptable failure cannot be defined	Establish a baseline, test cases, and explicit acceptance criteria.
The value depends on near-perfect model behavior	Redesign so outputs can be checked, limited, or kept advisory.
Consequential actions lack an approval rule	Require approval each time or prohibit the action.
Broad admin or production write access is the default	Create a named, least-privilege identity with narrow scopes and limits.
Sensitive-data handling or vendor eligibility is unknown	Resolve policy, contractual, privacy, and residency requirements.
Representative testing outside live production is impossible	Build a controlled sample, sandbox, shadow mode, or replay path.
Important actions cannot be audited	Add attributable logs before enabling external or write actions.
There is no safe fallback, pause, or containment path	Add a manual queue, kill control, rollback, and incident owner.
The business case has no current baseline	Measure the manual workflow before claiming improvement.

UNRESOLVED STOP CONDITIONS: _____

DECISION FRAME

Start at the lowest useful authority tier.

The pilot should demonstrate value without giving the system more authority than it needs.

TIER	AGENT AUTHORITY	STARTING CONTROL POSTURE
0 - Draft	Produces analysis, classifications, or drafts for a person.	Human owns every downstream action. Keep source evidence visible.
1 - Read	Searches or reads scoped company systems.	Read-only access, narrow data scope, logging, sensitive-data limits.
2 - Reversible write	Updates internal records or creates reversible artifacts.	Field and volume limits, validation, logs, rollback, sampled review.
3 - Consequential	Affects customers, money, rights, access, production, or irreversible state.	Explicit human approval each time or prohibit the action. Retain evidence.

Minimum useful tier: Highest permitted pilot tier:

READINESS SCORECARD

DOMAIN	SCORE	PREREQUISITE OR GAP	OWNER
1. Problem and value	___ / 10		
2. Workflow clarity	___ / 10		
3. Data and systems	___ / 10		
4. Authority and access	___ / 10		
5. Reliability and evaluation	___ / 10		
6. Security, privacy, and compliance	___ / 10		
7. Human oversight and exceptions	___ / 10		
8. Ownership and operating model	___ / 10		
TOTAL	___ / 80		

64-80 and no stop condition: bounded pilot candidate. 40-63: complete prerequisites and rescore. 0-39: simplify, reframe, or keep human-led. Any stop condition: no production connection.

These bands are operational triage, not an industry benchmark. Inspect the shape of the score: a high total does not compensate for weak authority, security, reliability, or ownership.

PILOT CANVAS

Bound the test before the build expands.

Complete this page with the people who own the workflow, access, and exceptions.

WORKFLOW AND OWNER

Workflow _____	Business owner _____
Technical owner _____	Trigger _____
Completion point _____	Current baseline _____

SCOPE AND AUTHORITY

In scope _____	Out of scope _____
Allowed reads _____	Allowed writes or external actions _____
Explicitly prohibited actions _____	Human approval required for _____

EVIDENCE AND FAILURE BEHAVIOR

Representative test set _____	Acceptance criteria _____
Failure or uncertainty routes to _____	Monitoring and alert owner _____
Manual fallback _____	Pause, rollback, and access revocation _____

DECISION BOUNDARY

Pilot duration or sample size _____	Decision date and owner _____
Continue when _____	Change when _____
Stop when _____	Evidence retained _____

DECISION RECORD

Leave with a decision, not a vague intention.

DECISION: ☐ Bounded pilot ☐ Complete prerequisites ☐ Use a simpler system ☐ Stop

WHY THIS IS THE RIGHT DECISION NOW

TOP THREE ACTIONS

ACTION	OWNER	DUE	EVIDENCE REQUIRED TO CLOSE
1.			
2.			
3.			

NEXT REVIEW DATE:

START WITH THE PROBLEM, NOT THE MODEL

A bounded path from AI idea to working system.

013 Labs helps small and mid-market organizations identify the right AI opportunity, test it against real operating conditions, and build a controlled system their team can actually run.

If this checklist exposes unresolved scope, access, evaluation, or operating questions, an **AI Systems Diagnostic** turns those unknowns into a documented build, buy, or do-not-build decision.

013labs.com/consulting.html | darren@013labs.com

013 Labs grounds its work in systems we have built and operated. That is evidence of how we work, not a promise that your outcome will look the same.