

A Checklist for AI Evaluation

This is a condensed checklist that you can use to help you distinguish between AI tools that deliver sustainable value and those that create expensive problems. Use it during vendor assessments, pilot planning, and post-deployment reviews.

Each item represents a critical decision factor derived from the detailed evaluation framework. For more detail on each checklist item, including strong and weak signals to listen for, tips on common pitfalls, and guides on cross-functional alignment, **download the full Repeatable Framework for Evaluating AI**

Strategic Alignment

- [ ] Clear business problem defined with measurable current and desired states
- [ ] Stakeholder input gathered, including clients, partners, and regulators
- [ ] No stakeholder prohibitions or constraints that would block AI usage
- [ ] Executive sponsorship with committed resources and realistic timeline
- [ ] Organizational readiness assessed for required changes

Use Case Definition and Boundaries

- [ ] Specific use cases identified and documented for this AI evaluation
- [ ] Use case risk levels assessed independently (not all-or-nothing)
- [ ] Vendor capabilities reviewed for granular feature control
- [ ] User populations defined for each approved use case
- [ ] Technical or policy controls identified for enforcing use case boundaries
- [ ] Process established for evaluating additional use cases in the future

Data and Risk

- [ ] Required data identified with confirmed access and quality metrics
- [ ] Data governance structure in place with clear ownership
- [ ] Legal and regulatory compliance requirements documented
- [ ] Vendor provides compliance certifications and transparent data policies
- [ ] Bias and fairness testing conducted with documented results
- [ ] Liability and data ownership clearly defined in contract terms

Technical Integration

- [ ] Integration requirements documented with existing system compatibility confirmed
- [ ] APIs and connectors validated for your technology stack
- [ ] Performance requirements defined and infrastructure capacity verified
- [ ] Skills assessment completed with training or hiring plan if gaps exist
- [ ] Technical debt and long-term maintenance implications understood

Workflow and Adoption

- ☐ User workflows mapped with AI tool integration points identified
- ☐ User experience tested with representative users, not just champions
- ☐ Training requirements assessed with realistic time-to-proficiency estimates
- ☐ Change management plan developed addressing organizational impacts
- ☐ Manual fallback procedures defined for tool unavailability

Pilot Validation

- ☐ Success criteria defined with measurable metrics before pilot begins
- ☐ Representative users and production-like data selected for testing
- ☐ Pilot duration sufficient for statistical validity and edge case discovery
- ☐ Structured feedback collection from users and technical teams
- ☐ Decision-making process established for post-pilot evaluation

Financial Justification

- ☐ Total cost of ownership calculated including all direct and indirect costs
- ☐ Business value quantified with specific, measurable outcomes
- ☐ ROI timeline realistic with risk-adjusted assumptions
- ☐ Alternative approaches considered including lower-cost options
- ☐ Decision criteria defined (i.e. when does this investment stop making sense)?

Vendor Partnership

- ☐ Vendor stability and track record verified with reference customers
- ☐ Product maturity assessed with honest evaluation of limitations
- ☐ Support quality and service level agreements documented
- ☐ Data portability and exit terms negotiated before contract signing
- ☐ Vendor roadmap aligned with your strategic direction

Scaling and Governance

- ☐ Scaling plan developed with phased approach and rollback procedures
- ☐ AI governance structure established with clear authority
- ☐ Performance monitoring plan includes automated checks for drift and bias
- ☐ Model maintenance and retraining processes documented
- ☐ Continuous improvement mechanisms defined with ongoing resource commitment