

AI Agent Platform

Evaluation Checklist



1

Alignment with Support Goals

- Can the platform be configured to prioritize your primary support objective (e.g., ticket deflection vs. resolution speed) without custom development?
- Does it allow different goal configurations across channels (e.g., chat vs. email)?
- Can you segment AI behavior by customer type, region, or support use case (billing vs. technical issues)?
- Does the platform track how AI actions impact KPIs like SLA adherence, first response time, or repeat ticket rates?
- Can goals be updated without requiring a major workflow redesign?

2

Workflow Design and Logic

- Can workflows handle multiple inputs—such as CRM data, order history, subscription status, and SLA priority—in a single flow?
- Can the AI take actions, such as updating records or triggering APIs, directly within workflows?
- Can the platform handle uncommon situations without giving vague or generic answers?
- Can workflow updates be deployed without disrupting existing support operations?
- Can multiple steps or agents work together within a single workflow?

3

AI Reasoning and Support Intelligence

- Can the AI remember context across multiple messages without losing track of the conversation?
- How does the system handle vague or incomplete queries without making wrong assumptions?
- Can the AI handle follow-up questions within the same context, or does it treat them as new issues?
- Can the AI agent prioritize and adapt responses based on signals like urgency, issue type, or escalation risk?
- Does the AI handle uncertain situations smoothly without creating a poor experience?

4

Omnichannel Support

- Can a conversation continue across channels without losing context?
- Are responses consistent across entry points (chat, email, social, voice)?
- Does the platform support channels natively, or does it depend heavily on third-party connectors?
- Can the AI agent access unified interaction history across all channels in one view?

5**Human Handoff and Collaboration**

- Can you set rules for when conversations should be escalated based on confidence or failure?
- Does the full conversation, including actions taken, carry over during handoff?
- Are AI-generated summaries clear, accurate, and ready for agents to use?
- Can agents take over or intervene without restarting an interaction?
- Are internal collaboration features embedded within the same workflow environment?

6**Knowledge Management and Training**

- Can the AI use both structured (KB articles) and unstructured data (tickets, documents) as sources?
- How quickly do knowledge base updates show up in AI responses?
- Can knowledge content be updated without disrupting existing AI responses?
- Are approval workflows, versioning, and rollback controls available for knowledge updates?

7**Integration Ecosystem**

- Does the system offer prebuilt integrations with popular business tools (CRM, help desk, ERP, collaboration apps, etc.)?
- Can the platform pull real-time data like account status, order updates, and subscription changes from connected systems?
- Do integrations allow two-way updates or only read customer data?
- Does the system support low-code or no-code integration builders?
- Can the platform be customized for specific business integrations or workflows?

8**Analytics and Reporting**

- Can you clearly track which issues the AI resolved, assisted with, or failed?
- Is deflection measured based on reduced workload, not just the number of sessions?
- Are insights available at the level of workflows, intent, and automation steps?
- Can reporting be customized for operational vs. leadership needs?
- Does the system automatically highlight performance issues or declines?

9

Security, Privacy, and Compliance

- Does the platform meet required regulatory and industry standards (e.g., GDPR, SOC 2), with verifiable certifications or audit reports?
- Is data securely encrypted both in transit and at rest across all system components?
- Are access permissions easy and configurable across users, roles, teams, and environments?
- Are all actions logged clearly for auditing and compliance checks?
- Can you control how data is stored, masked, and deleted to meet compliance needs?

10

Pricing and Total Cost of Ownership

- Is the pricing model clear (per interaction, per agent, or resolution-based)?
- Are set-up, onboarding, and integration costs included or charged separately?
- Will costs rise as usage grows, and how predictable are those increases?
- Are key features like automation, APIs, and integration included or charged extra?
- Are there any hidden or unexpected charges as you scale?

11

Vendor Reliability and Final Validation

- How responsive is the vendor when issues arise, and are escalation paths clearly defined?
- Is documentation consistently updated and detailed enough for troubleshooting?
- Does the vendor provide ongoing support beyond initial onboarding?
- Can the vendor demonstrate consistent performance across real customer deployments?
- Does the vendor have a track record of supporting teams with similar scale or complexity?

12

Implementation Readiness and Operational Scalability

- How complex is the initial setup, and does it require developer involvement?
- How quickly can teams configure and launch workflows using templates or prebuilt components?
- How are AI agents monitored, retrained, and optimized over time?
- Can the platform scale across teams, regions, or growing ticket volumes without significant reconfiguration?
- Is operational ownership clearly defined with appropriate governance controls?