



FREE EVALUATION GUIDE

AI Chatbox Vendor Evaluation Checklist

20 questions to ask before you choose a chatbox.
Covers hidden costs, AI capability, integrations,
support quality, and scalability.

01 Pricing Traps

02 AI Capability

03 Integrations

04 Support Quality

05 Scale & Flexibility

20,000+

Shopify stores

\$1M+

in attributed sales

4.9 stars

Shopify App Store

How to Use This Checklist

Use this checklist during vendor demos or before signing any contract. For each question, note the vendor's exact answer. Vague or deflecting answers are as informative as direct ones. The goal is not to eliminate every vendor who scores below perfect - it is to surface hidden risks before they cost you money.

Step 1	Step 2	Step 3
BEFORE THE DEMO Send your top 3-4 questions to the vendor in advance. Their willingness to answer in writing tells you a lot about transparency.	DURING THE DEMO Work through each section systematically. Ask follow-up questions. Watch for deflection or repositioning instead of direct answers.	AFTER THE DEMO Score the vendor using the guide on the last page. Compare scores across vendors before making a final decision.

Quick Scoring Guide	
17-20 questions answered clearly	Strong fit. Proceed to a paid trial.
12-16 questions answered	Acceptable. Clarify remaining gaps in writing before signing.
Below 12	High risk of hidden costs or capability gaps. Keep evaluating.

The 5 buyer pain points this checklist covers:

01	Pricing Traps	Hidden fees that make cheap plans expensive.
02	AI Capability	Real intelligence vs. rule-based keyword matching.
03	Integrations	Native Shopify vs. "we have a Zapier workaround".
04	Support Quality	Who you call when the AI starts giving wrong answers.
05	Scale & Flexibility	What happens during your biggest sale of the year.

Pricing Traps

The plan looks affordable. Then the invoice arrives.

WHY THIS MATTERS

Most chatbox vendors advertise a low monthly fee - but bury per-resolution charges, AI conversation overages, and mandatory onboarding fees in the fine print. A \$49/month plan can quickly become \$400+/month if your store has high chat volume. Understand the full cost model before you ever sign a contract.

QUESTIONS TO ASK

☐

1. Is the monthly price fixed, or are there usage-based fees on top?

GOOD: Fixed flat-rate pricing (like Chatty's model) means no surprises.

RED FLAG: Any answer involving 'per-resolution', 'per-AI-conversation', or 'it depends on volume'.

☐

2. Are there per-resolution or per-ticket charges beyond the base plan?

GOOD: A clear 'No' - or a clearly capped overage rate you can budget for.

RED FLAG: Gorgias charges \$0.36 per AI resolution after plan limits. This can multiply costs fast.

☐

3. What happens when I exceed my plan limits?

GOOD: A soft cap with advance notice, or a flat overage rate (\$0.40/conversation).

RED FLAG: Hard cutoffs that disable the chatbot mid-sale, or automatic upgrades to the next tier.

☐

4. Can I cancel month-to-month, or is there a minimum contract?

GOOD: Monthly flexibility - especially important while evaluating.

RED FLAG: Annual contracts with no trial, or cancellation fees.

AI That Actually Learns

Real intelligence vs. fancy keyword matching.

WHY THIS MATTERS

Many chatbots marketed as 'AI-powered' are still running on rule-based decision trees. They match keywords to pre-written answers. They cannot understand nuanced questions, handle product-specific queries, or adapt to new inventory without manual updates. GPT-4-powered AI reads intent, not just words - the difference shows immediately in customer experience and conversion rates.

QUESTIONS TO ASK

☐

1. What AI model powers the chatbot - GPT-4, proprietary, or rule-based?

GOOD: GPT-4 or equivalent LLM. Ask specifically - do not accept 'advanced AI' as an answer.

RED FLAG: Vague answers like 'our proprietary engine'. This often means a decision tree.

☐

2. Can the AI learn from my product catalog, FAQs, and order history automatically?

GOOD: Yes - upload catalog or connect Shopify and the AI trains itself overnight.

RED FLAG: Manual input required for every product. Not scalable beyond 50 SKUs.

☐

3. How does it handle questions it cannot answer?

GOOD: Escalates gracefully to a live agent with full context of the conversation.

RED FLAG: Sends a generic fallback message or, worse, guesses and gives wrong information.

☐

4. Can I review AI responses and correct them before they affect customers?

GOOD: Yes - an admin panel where you can audit, edit, or block specific responses.

RED FLAG: Black-box AI with no visibility into what it's saying on your behalf.

Integrations That Work on Day One

Native Shopify? Or 'we have a Zapier workaround'?

WHY THIS MATTERS

A chatbox that cannot read your order history, product catalog, or customer data in real time is just a generic Q&A widget. The most common integration failure: vendors promise 'Shopify integration' but require a developer to set it up, or the integration only pulls basic order status - not inventory, tags, metafields, or customer history. Ask exactly what data the integration surfaces.

QUESTIONS TO ASK

☐

1. Is the Shopify integration native and maintained by your team, or third-party?

GOOD: Native integration - listed in Shopify App Store, maintained by the vendor.

RED FLAG: Integrations that require Zapier, Make, or a developer to configure.

☐

2. What Shopify data does the AI have access to in real time?

GOOD: Orders, products, inventory, customer history, tags, and metafields.

RED FLAG: Order status only. This limits the AI to basic tracking responses.

☐

3. Does the AI connect to my CRM or helpdesk, and is that included in the plan?

GOOD: Included natively in the plan you're evaluating.

RED FLAG: Integration sold as a separate add-on or requires an Enterprise upgrade.

☐

4. What is the typical implementation timeline from signup to live?

GOOD: Same day or under 48 hours for a standard Shopify store.

RED FLAG: Multi-week implementation with a dedicated 'onboarding specialist' (= hidden labour cost).

Support When It Breaks

What happens when the AI starts confusing your customers?

WHY THIS MATTERS

AI chatbots will occasionally give wrong answers, miss context, or stop responding during high-traffic periods. The difference between a minor issue and a customer service disaster is how fast the vendor responds when you flag a problem. Many vendors offer 'live chat support' on their website but provide email-only support on lower-tier plans - with 24-48 hour response times. Find this out before, not after, your biggest sale.

QUESTIONS TO ASK

☐

1. What support channel do I get on my plan - email, live chat, or dedicated manager?

GOOD: Live chat or dedicated manager for Pro-tier plans and above.

RED FLAG: Email-only support on plans under \$50/month. Unacceptable for a revenue-critical tool.

☐

2. What are the SLA response times if something breaks during a peak sale period?

GOOD: A guaranteed response within 1-4 hours for critical issues.

RED FLAG: No SLA commitment. 'We aim to respond within 2 business days.'

☐

3. Is there a status page where I can monitor uptime and incidents?

GOOD: Yes - a public status page (e.g., status.chatty.net).

RED FLAG: No public status page. You find out about outages when customers complain.

☐

4. Has the platform experienced downtime in the last 6 months? What was the cause?

GOOD: Honest answer with a post-mortem. Every platform has incidents; transparency matters.

RED FLAG: Deflection or inability to answer. This signals poor reliability culture.

Scale and Flexibility

What happens during your biggest sale of the year?

WHY THIS MATTERS

Your chatbox needs to handle a 10x traffic spike on Black Friday without slowing down, hitting limits, or requiring a plan upgrade mid-campaign. Montana West grew to handle a 10x holiday surge with Chatty without adding a single agent. That kind of elasticity must be built into the platform architecture - not something you manually provision. Ask specifically about surge handling and what happens at your plan's conversation limits.

QUESTIONS TO ASK

☐

1. How does the platform handle sudden traffic spikes - e.g., 10x normal volume?

GOOD: Auto-scaling infrastructure with no manual intervention required from your side.

RED FLAG: You need to upgrade your plan in advance or risk service degradation.

☐

2. Can I white-label or customize the chat widget to match my brand?

GOOD: Full widget customization including colors, fonts, position, and welcome message.

RED FLAG: Limited to preset themes, or branding customization requires Enterprise tier.

☐

3. Does the AI support multiple languages, and which ones are included in my plan?

GOOD: Multi-language support included, auto-detected from browser or customer preference.

RED FLAG: English only, or additional languages require a separate add-on fee.

☐

4. What analytics do I get - and can I export data to my own tools?

GOOD: Conversation analytics, resolution rates, revenue attribution, CSV export included.

RED FLAG: Basic metrics dashboard only. No data export or API access on standard plans.

Your Vendor Decision Framework

Score your vendor:

Section	Max Score	Your Score	Notes
01 Pricing Traps	4		
02 AI Capability	4		
03 Integrations	4		
04 Support Quality	4		
05 Scale & Flexibility	4		
TOTAL	20		

How to interpret your score

17-20	Strong Fit	Proceed to a paid trial. Negotiate pricing before signing.
12-16	Acceptable with Gaps	Get written answers on the gaps. Add contractual protections.
8-11	Risky	Hidden costs or capability limitations likely. Request references.
Below 8	Do Not Proceed	Too many unanswered questions. Evaluate other vendors first.

See how Chatty scores on this checklist.

Chatty is flat-rate priced, GPT-4 powered, natively built for Shopify, and scores full marks on every section of this checklist. 20,000+ stores. 4.9 stars. Start free today.

Try Chatty Free

chatty.net | Free plan available | No credit card required