CUSTOMER SERVICE

WhatsApp Charges for Replies Again: How to Design Your AI Agent

Since 1 October 2026, Meta has been charging again for service messages on the WhatsApp Business Platform. These are the non-template replies a business sends inside the 24-hour customer service window that opens when a customer writes in. They had been free since 1 November 2024. Each business number now gets 1,000 delivered service messages free per month, with no roll-over, and every one from the 1,001st is charged.

The change applies to the platform (the Cloud API that agents and providers use), not to the free WhatsApp Business app. It arrives alongside another Meta rule, in force since January, on which artificial intelligence may run there.

What Meta’s Pricing Says

According to Meta’s pricing documentation, billing has been per delivered message since 1 July 2025, and only what the business sends to the customer is charged. Service messages include those written by a third-party AI agent and those written by a person on your team. Venezuela (+58) has no rate of its own: it falls under Rest of Latin America. Meta may change prices only on the first day of each quarter.

Without a payment method on file, Meta delivers service messages within the free tier and stops delivering them once it is used up. Since July 2026 there has been a fifth category, Meta Business Agent, for Meta’s own agent, billed by tokens at USD 2 per million, which Meta itself estimates at roughly 4 to 5 US cents per message.

The conditions are set out in Meta’s pricing documentation for the WhatsApp Business Platform, updated on 30 September 2026.
https://developers.facebook.com/documentation/business-messaging/whatsapp/pricing

Which AI Meta Allows on the Platform

Since 15 January 2026, the platform terms (now in clause 4.7) have barred AI providers when their technology, a general-purpose assistant for example, is the primary function of what they offer rather than an incidental one. Meta decides which side of that line a provider falls on, at its sole discretion. The same clause lets a business retain an AI provider to use the platform on its behalf, but in that case bars it from letting platform data be used to train or improve AI models, with one exception:

fine-tuning a model for the business's exclusive use, as long as the data does not end up training other models.

The clause can be read in Meta's terms for the WhatsApp Business Platform, last updated on 23 September 2026.
https://www.facebook.com/legal/Meta-Terms-for-WhatsApp-Business-Platform

There are exceptions where the law requires them. Italy's competition authority (AGCM) ordered Meta on 24 December 2025 to suspend those terms over their effects in Italy, and on 9 June 2026 the European Commission imposed interim measures to restore free access for rival assistants until it reaches a final decision. The Commission said at the time that its investigation on the merits was still open. According to Slaughter and May, the law firm advising Meta, the company appealed against those measures to the EU General Court in August 2026. Meta's page for AI providers lists the markets where it allows general-purpose assistants: Italy and the rest of the European Economic Area, Brazil and Türkiye. Venezuela is not on that list.

The European order is explained in the European Commission's press release of 9 June 2026.
https://ec.europa.eu/commission/presscorner/detail/en/ip_26_1276

How to Design an Agent Under These Rules

- Build an agent for one business's catalogue, prices and service, and have it politely decline general questions unrelated to it.
- Write complete answers in a single message, with price, availability and next step, instead of a chain of short bubbles each charged separately.
- Prepare approved templates for what happens outside the 24-hour window, such as order updates, because free-form messages cannot be sent there.
- Count delivered messages per number and per month, and set an alert before the 1,000 free ones run out: your team's replies count too.
- Confirm that a payment method is on file in Billing Hub so replies are not cut off when the free tier is used up.
- Ask any model provider in writing what it does with conversations: whether it trains on them, how long it keeps them and who can see them.

How the WhatsApp AI Agent Handles This

We design the agent around one business: a knowledge base limited to its catalogue, prices, availability and policies, flows for the questions it really receives, and handover to a person when

judgement is needed. We write answers that settle in one message what might otherwise take three, and anything outside the window goes out as an approved template. We log which messages are charged, from the pricing data Meta includes in each message's status, so you can see how many there are per conversation, per month and per number. We check what the model provider's terms say about conversations and prefer terms that rule out training shared models. We document what to adjust if Meta changes its rates or terms. Before going live, we test the agent on authorised real conversations.

> *Rates, terms and the decision on what counts as a general-purpose assistant are Meta's and can change. The European case is not closed. This is not legal advice: if your business also operates in the European Economic Area, Brazil or Türkiye, ask a lawyer what applies there.*

A Check You Can Run

Open Billing Hub in Meta Business Suite and confirm that your WhatsApp messaging account has a valid payment method. Then find how many service messages each business number delivered last month, in your provider's panel or through Meta's pricing analytics, and compare the figure with the 1,000 free per number. If you went past it, read ten recent conversations and count the replies that could have been merged into one complete message. You change nothing while you look.

Supporting guidance: Meta's pricing documentation for the WhatsApp Business Platform.
https://developers.facebook.com/documentation/business-messaging/whatsapp/pricing

Continue with: Your Business Already Lives on WhatsApp. It Just Doesn't Answer the Same at Every Hour.
https://systems.kavaliero.com/en/blog/your-business-already-lives-on-whatsapp

For help with this work, see the scope of WhatsApp AI Agent.
https://systems.kavaliero.com/en/services/whatsapp-ai