

APRIL 4, 2026

# WHATSAPP AI CONCIERGE PLAYBOOK: DM TO PAID CALL IN 10 MINUTES (TEMPLATES + COSTS)

A copy-ready, provider-agnostic guide to launch a WhatsApp AI concierge that qualifies in four questions, routes to the right offer, auto-books calls, and can take a deposit — built for traveling founders with café-proof fail-safes.

FROM EPISODE

[THE WHATSAPP AI CONCIERGE FUNNEL: FROM DM TO PAID CALL IN 10 MINUTES \(TEMPLATES + COSTS\)](#)

## CONTENTS

- The 10-Minute Funnel at a Glance
- Stack + Wiring Map (provider-agnostic)
- Copy-Paste Template Pack (12 messages + guardrail)
- Routing Rules (BANT in 4 messages)
- Unit Economics (use these working numbers)
- Compliance, Approvals, and Limits (what to avoid getting throttled)
- Failure Modes + Monitoring (make it café-proof)
- Offer Archetypes (pick one to start this week)

Ship a WhatsApp AI concierge that greets leads, asks four qualifying questions, routes them to the right offer, books a slot, and (optionally) takes a deposit — without you being awake. This playbook gives you the wiring map, copy-paste templates, routing logic, cost math, guardrails, and failure-mode fixes so it runs reliably on café Wi-Fi across time zones.

---

## **THE 10MINUTE FUNNEL AT A GLANCE**

The flow is simple enough to build in an hour and robust enough to run while you travel.

End-to-end path:

1. Entry: Click-to-WhatsApp ad, QR code, or wa.me link →
2. Greeting + permission/opt-in →
3. 4-question qualifier (Budget, Timeline, Use case, Authority) →
4. Offer route: Book call (Calendly/TidyCal) OR pay deposit (Stripe Payment Link) OR workshop invite OR nurture →
5. Confirmations + reminders →
6. CRM/log update →
7. Human handoff on edge cases/errors.

Nomad defaults: keep responses templated, let Calendly own time zones, and set monitoring to catch issues while you're on planes.

---

## **STACK + WIRING MAP (PROVIDERAGNOSTIC)**

Pick any BSP you prefer; the logic below is provider-agnostic.

## Prereqs:

- WhatsApp Business Platform access + dedicated number
- BSP account (e.g., Gupshup/Twilio/WATI) with template submission enabled
- Automation runner (Make.com or n8n)
- Calendar app (Calendly/TidyCal) with API access
- Payment processor (Stripe Payment Links)
- LLM API (e.g., GPT-4o-mini or equivalent) for light intent parsing
- Alert channel (Telegram/Email) for handoff

## Blueprint (Make.com or n8n modules):

1. Trigger: Incoming WhatsApp message webhook → store [lead\_id, wa\_number, source, timestamp, thread\_id].
2. Consent Gate: If no explicit permission recorded, send [OPT\_IN\_REQUEST] template; proceed only on affirmative intent.
3. State Machine: Manage conversation state (Q1→Q4→Route). Enforce allowed replies (A/B/C) to keep the LLM minimal.
4. Questions: Send templated messages Q1–Q4; capture normalized answers.
5. Scoring + Routing: Evaluate Budget/Timeline/Use-case/Authority → set route = [BOOK\_FREE\_CALL | BOOK\_PAID\_CALL | WORKSHOP | NURTURE | HUMAN].
6. Calendar: Fetch available slots via Calendly API (owner = you). Create prefilled scheduling link with [name, email/WA, timezone auto].
7. Payments (optional): Generate Stripe Payment Link for [DEPOSIT\_AMOUNT] when route = BOOK\_PAID\_CALL.
- 8.

Messaging: Deliver booking/payment link template; on success, send confirmation + reminders (24h and 2h before).

9. CRM/Sheet: Upsert lead with answers, route, booking URL, payment status; append transcript pointer.
10. Monitoring: Daily digest of counts/costs; real-time error to Telegram if state stalls >10 min, payment fails, or non-matched intent → HUMAN\_HANDOFF.

---

## **COPYPASTE TEMPLATE PACK (12 MESSAGES + GUARDRAIL)**

Copy these as WhatsApp templates (keep variables in [BRACKETS]). Submit 2–3 variants per critical step to survive rejections.

1. OPT\_IN\_REQUEST  
 “Hey [FIRST\_NAME] — you tapped my WhatsApp link from [SOURCE]. I help [ICP] with [OFFER]. OK to ask 4 quick questions to see if I can help? Reply YES to continue or STOP to opt out.”
2. Q1\_BUDGET  
 “To recommend the right path, where’s monthly revenue? A) < \$5k B) \$5k–\$15k C) \$15k+ (Reply A, B, or C.)”
3. Q2\_TIMELINE  
 “When do you want this live? A) This month B) Next quarter C) Just researching (Reply A, B, or C.)”
4. Q3\_USECASE  
 “What’s the main thing you want to automate? A) Lead capture B) Client onboarding C) Content repurposing D) Customer support (Reply A–D.)”
5. Q4\_AUTHORITY  
 “Are you the decision maker for this? A) Yes B) I share decision C) No (Reply A, B, or C.)”

in 10 Minutes (Templates + Costs)

6. ROUTE\_BOOK\_FREE  
“Looks like we can help. Grab a time here: [CALENDLY\_LINK]. It shows your local timezone. You'll get reminders automatically.”
7. ROUTE\_BOOK\_PAID  
“We can fast-track you with a credited strategy session. Secure your spot with a [DEPOSIT\_AMOUNT] deposit here: [STRIPE\_LINK]. It's fully credited toward any package you choose.”
8. ROUTE\_WORKSHOP  
“Best next step: our [WORKSHOP\_NAME] this week for \$[PRICE]. Seat link: [CHECKOUT\_LINK]. You'll leave with a working mini-system.”
9. CONFIRM\_BOOKING  
“Locked in: [DATE\_LOCAL] at [TIME\_LOCAL]. You'll get a 24h and 2h reminder here. Need to reschedule? Use this link: [RESCHEDULE\_LINK].”
10. REMINDER\_24H  
“Reminder: Tomorrow [DATE\_LOCAL] at [TIME\_LOCAL]. Bring: [PREP\_ITEMS]. Reschedule: [RESCHEDULE\_LINK].”
11. EDGE\_CASE\_HANDOFF  
“That's a great question. Looping in [HUMAN\_NAME]. Expect a reply within 12 hours. Want to book instead? [CALENDLY\_LINK]”
12. OPT\_OUT  
“No problem — you're unsubscribed. If that changes, you can message START any time.”

Prompt guardrail (system prompt for LLM):

“You are a WhatsApp concierge for [BRAND]. You only send pre-approved templates with variables filled. Never invent times, prices, or claims. If a user's message doesn't map to an expected answer (A/B/C/D or YES/STOP), respond with EDGE\_CASE\_HANDOFF and raise an alert.”

---

## ROUTING RULES (BANT IN 4 MESSAGES)

in 10 Minutes (Templates + Costs)

Translate answers to actions with explicit rules so the bot can't wander. Recommended defaults for solo nomads:

Budget (Q1): A (<\$5k) → WORKSHOP; B (\$5k–\$15k) → BOOK\_FREE\_CALL; C (\$15k+) → BOOK\_PAID\_CALL with [DEPOSIT\_AMOUNT=\$97].

Timeline (Q2): If C (Researching) → NURTURE regardless of budget; else continue.

Use case (Q3): Tag intent for agenda + case studies in reminders.

Authority (Q4): If B or C (shared/no), append "Include decision maker?" step; else continue.

Routing pseudo-logic:

```
if TIMELINE == RESEARCH: route = NURTURE
else if BUDGET == C: route = BOOK_PAID_CALL
else if BUDGET == B: route = BOOK_FREE_CALL
else: route = WORKSHOP

# Authority check
if AUTH in {B, C} and route in {BOOK_FREE_CALL, BOOK_PAID_CALL}:
    send CO_DECIDER template, then proceed when confirmed

# Booking mutex
when generating [CALENDLY_LINK]: lock slot_id for 5 minutes; release on success/fail

# Postbooking
REMINDERS = ON (24h and 2h)
if PAID route and payment fails: retry at +1h and +6h; after 3rd fail → HUMAN
```

---

## UNIT ECONOMICS (USE THESE WORKING NUMBERS)

You need two cost lanes: messaging and model. Use conservative planning numbers and adjust to your BSP and region.

Messaging: All-in per outbound message ≈ Meta template fee + BSP/platform markup. A practical US planning number is ~\$0.03/message. If the chat starts from a Click-to-WhatsApp ad or CTA

in 10 Minutes (Templates + Costs)

button, user-initiated messaging is typically free within a limited window — run your qualifier + booking inside that window.

Model: Assume ~300 tokens/turn over ~6–8 turns → ~1,800–2,400 tokens/convo. On “mini” models, that’s low cents per convo; even full-fat models are typically <\$0.15 per convo when prompts are small.

Quick estimator (paste into a sheet):

```
Inputs: leads_per_month [100], engage_rate [0.30], turns_per_convo [8], msg_cost [0.03], tokens_per_turn [300], token_price_per_1k [0.0025], ad_initiated_share [0.50]
```

Derived:

```
convos = leads_per_month * engage_rate
messages = convos * turns_per_convo
paid_messages = messages * (1 - ad_initiated_share)
wa_cost = paid_messages * msg_cost
llm_tokens = convos * turns_per_convo * tokens_per_turn
llm_cost = (llm_tokens/1000) * token_price_per_1k
total_ops_cost = wa_cost + llm_cost
```

Example with defaults: convos=30; messages=240; paid\_messages=120; wa\_cost=\$3.60; llm\_tokens=72,000; llm\_cost=\$0.18; total\_ops\_cost ≈ \$3.78. Twelve booked calls at 40% of engaged, three sales at \$1.5k → \$4,500 revenue on <\$10 ops cost. Adjust to your geography/model.

---

## COMPLIANCE, APPROVALS, AND LIMITS (WHAT TO AVOID GETTING THROTTLED)

Stay well inside policy so your number doesn’t get throttled while you’re in the air.

Consent: Always capture explicit opt-in (surface, timestamp, exact language). Honor STOP/unsubscribe immediately.

Templates: Keep copy plain, specific, and non-hype. Submit 2–3 variants per key step (greeting, booking, reminders). Avoid overusing “free,” vague CTAs, and promises.

in 10 Minutes (Templates + Costs)

Messaging limits: Ramp slowly. Week 1: ≤10 new convos/day; Week 2: ≤100/day if quality high; Week 3+: scale. Watch quality rating; reduce volume if it dips.

Data minimization: Store only what you need (consent, answers, booking/payment refs). Auto-purge transcripts at 6 months; retain opt-in proofs.

Task-specific bot: Keep the assistant narrowly focused on qualification/booking — not a general chat AI.

Required footer cues: Periodically include “Reply STOP to opt out.” in non-transactional messages.

---

## **FAILURE MODES + MONITORING (MAKE IT CAFÉ - PROOF)**

Assume things will break. Design graceful failure.

Template rejections: Have alternates ready; strip promo phrasing; resubmit.

Hallucinations: Bot can only send pre-approved templates; never compose free text. If message doesn't match A/B/C/D or YES/STOP → trigger EDGECASE\_HANDOFF and alert you.

Timezone misses: Let Calendly convert time zones; include local time echo in CONFIRM/RE-MINDER. Nightly manual “tomorrow's calls” check.

Double bookings: Implement a 5-minute booking lock (mutex) per slot while user is on the link.

Payment fails: Retry at +1h and +6h; offer alternative method (PayPal/Wise). After 3 fails → HUMAN.

Rate-limit dips: Auto-reduce daily start volume by 50% and swap to lower-friction templates for 48h.

Monitoring: Real-time Telegram alert on errors; daily digest: leads→convos→qualified→booked→show rate→costs. SLA: human reply within 12 hours on handoffs.

---

# OFFER ARCHETYPES (PICK ONE TO START THIS WEEK)

Use one of these positions; the tech stays the same.

1. \$97 Audit Tripwire → \$1.5k+ Implementation  
Flow: PAID route with \$97 deposit to book → 30-min screen share audit → present 3 packages (\$1.5k/\$3k/\$5k). Credit deposit toward purchase. Typical conversion: ~60% to at least the base package (your mileage will vary).
2. \$49 Mini-Workshop → \$1k Cohort/Done-For-You  
Flow: WORKSHOP route for Budget A or “just researching.” Live or pre-record + WhatsApp Q&A. Follow-up sequence pushes cohort/DFY.
3. Free Lead Magnet → \$300 Strategy Call  
Flow: NURTURE route delivers a 3-day mini-course in WhatsApp; Day 3 offers a \$300 strategy session to filter tire-kickers. Expect lower booking rate (~8%) but higher close quality.

Tip: Start with one archetype for 2 weeks, then A/B the deposit/workshop price — not the entire flow.