

VOXAPP

AI Receptionist Platform

Volledig Projectdocument

Versie 2.0 - Eigen Orchestratie Architectuur

Datum: 28 februari 2026

Status: VERTROUWELIJK

www.voxapp.tech

INHOUDSOPGAVE

1. Wat is VoxApp?
2. Tech Stack
3. EU-Only Regels
4. Architectuur - Eigen Orkestratie + Vapi Audio
5. Bestaande Code (39 API routes, 13 lib bestanden)
6. Database Schema (15+ tabellen)
7. Wat Gebouwd Moet Worden (4 fases)
8. Vapi Custom LLM Configuratie
9. Multi-Tenant Regels
10. Beveiliging
11. Conversatie Kwaliteitsregels
12. Git & Deploy Regels
13. Pricing Model
14. Key Files Referentie

1. WAT IS VOXAPP?

VoxApp is een AI-receptionist platform voor KMO's in België, Nederland en Duitsland. De AI neemt de telefoon op voor bedrijven (kappers, restaurants, dokters, garages, advocaten, etc.), boekt afspraken, beantwoordt vragen, en stuurt bevestigingen. 24/7. Met de eigen stem van de eigenaar.

Kerngegevens

- Website: www.voxapp.tech
- GitHub: github.com/rudiaerden19-blip/voxapp.git
- Doelgroep: Kappers, tandartsen, dokters, garages, restaurants, advocaten
- Regio: Benelux + DACH
- Talen: Nederlands (Vlaams), Frans, Duits

2. TECH STACK

Frontend

- Next.js (App Router, TypeScript)
- Tailwind CSS v4 - gebruik @import 'tailwindcss', NIET oude @tailwind
- Lucide React icons
- Mobile-first design - altijd EERST telefoon, daarna desktop

Backend / Runtime

- Vercel (serverless, Edge Runtime)
- Supabase (PostgreSQL) - EU regio
- Database URL: <https://bkjqadaamxmjeenzsl.supabase.co>

Voice AI

- Orchestratie: Vapi EU (api.eu.vapi.ai - Frankfurt) als audio-laag
- STT: Deepgram (via Vapi) - nova-2 model
- LLM: OpenAI GPT-4o-mini (EIGEN aanroep, niet via Vapi)
- TTS: ElevenLabs (via Vapi) - eleven_flash_v2_5

Telephonie

- Provider: Telnyx
- Belgisch nummer: +32480210478
- Anchorsite: Amsterdam, Netherlands (NOOIT San Jose)

Deployment

- Hosting: Vercel - automatisch via GitHub push
- CI/CD: GitHub -> Vercel (nooit lokaal)
- Cron: vercel.json - /api/health elke 5 minuten

3. EU-ONLY REGELS (NIET ONDERHANDELBAAR)

Alle data en processing MOET binnen de EU blijven. Dit is niet optioneel.

- Vapi API: ALTIJD api.eu.vapi.ai (Frankfurt), NOOIT api.vapi.ai (US)
- Telnyx anchorsite: Amsterdam, Netherlands - NOOIT San Jose, CA
- Supabase: EU regio
- Vercel: automatische regio (EU edge functions)
- Geen data naar US servers
- GDPR compliant: opt-in recordings, retention policy

Vapi EU Environment Variables

Variable	Waarde
VAPI_API_BASE	https://api.eu.vapi.ai
VAPI_ASSISTANT_ID	7a57ac55-1ca0-4395-a78b-7c78d4093d78
VAPI_PHONE_NUMBER_ID	ad5fb1f0-1db8-4a64-a56b-2f4254101481
VAPI_CREDENTIAL_ID	5b86a7bb-5d64-4518-ae15-f261e7ea7c19

4. ARCHITECTUUR - EIGEN ORCHESTRATIE + VAPI AUDIO

Vapi wordt ALLEEN gebruikt als audio-infrastructuur (STT + TTS + telephony). De conversatie-logica, state machine, LLM aanroepen en booking logic draaien op ONZE backend via een Custom LLM endpoint.

Call Flow

```

Klant belt +32480210478
|
Telnyx (anchorsite: Amsterdam)
|
Vapi EU (Frankfurt) - ontvangt audio
|
Deepgram STT - audio -> tekst
|
Vapi stuurt transcript naar ONZE Custom LLM URL
POST https://www.voxapp.tech/api/vapi/chat
|
ONZE BACKEND:
|- Parse transcript
|- Haal session state op (Supabase voice_sessions)
|- State Machine bepaalt volgende stap
|- OpenAI GPT-4o-mini genereert antwoord
|- Check availability inline (Supabase)
|- Slot locking + booking
|- Return antwoordtekst naar Vapi
|
Vapi stuurt tekst naar ElevenLabs TTS
|
Audio terug naar beller

```

Voordelen Eigen Orchestratie

- Deterministische state machine (code beslist, niet AI)
- Eigen LLM controle (model, temperature, prompt per state)
- Inline availability check (VOOR je iets voorstelt aan klant)
- Slot locking (soft lock voordat klant bevestigt)
- Geen tool call overhead (alles in een request)
- Latency winst: minder hops
- Migeerbaar: later Vapi vervangen door eigen audio pipeline

5. BESTAANDE CODE

API Routes (39 endpoints)

Afspraken

Route	Functie	Status
/api/appointments/save	Vapi tool - afspraak opslaan	WERKEND

Beschikbaarheid

Route	Functie	Status
/api/availability	GET slots, POST check tijdstip	WERKEND

Health & Failover

Route	Functie	Status
/api/health	Health check (cron 5min)	WERKEND
/api/telnyx/failover	Telnyx failover bij Vapi down	WERKEND

Admin

Route	Functie	Status
/api/admin/auth	Admin login	WERKEND
/api/admin/tenants	Tenant CRUD	WERKEND
/api/admin/services	Diensten CRUD	WERKEND
/api/admin/staff	Medewerkers CRUD	WERKEND
/api/admin/appointme..	Afspraken CRUD	WERKEND
/api/admin/kennisbank	Kennisbank CRUD	WERKEND
/api/admin/pool-numb..	Pool nummers	WERKEND

Business

Route	Functie	Status
/api/business/[id]	Business data	WERKEND
/api/business/by-email	Zoek op email	WERKEND
/api/business/update	Business updaten	WERKEND
/api/business/modules	Module config	WERKEND

AI Tools

Route	Functie	Status
/api/ai-tools/check..	AI: beschikbaarheid	WERKEND
/api/ai-tools/create..	AI: afspraak	WERKEND
/api/ai-tools/create..	AI: bestelling	WERKEND

ElevenLabs

Route	Functie	Status
/api/elevenlabs/agent	Agent beheer	WERKEND
/api/elevenlabs/voices	Stemmen ophalen	WERKEND
/api/elevenlabs/prev..	Stem preview	WERKEND
/api/webhooks/eleven..	ElevenLabs webhook	WERKEND

Stripe

Route	Functie	Status
/api/stripe/checkout	Checkout sessie	BASIS
/api/stripe/portal	Klantportaal	BASIS
/api/stripe/webhook	Stripe webhook	BASIS

Overig

Route	Functie	Status
/api/tenants/provision	Tenant onboarding	WERKEND
/api/metrics	Metrics per tenant	WERKEND
/api/knowledge	Kennisbank API	WERKEND
/api/usage	Verbruik tracking	WERKEND
/api/forwarding	Doorverbinden	WERKEND

Library Bestanden (src/lib/)

Bestand	Functie
apiAuth.ts	Webhook + admin + ElevenLabs auth
apiLogger.ts	Structured JSON API logging
logger.ts	Call logging (voice engine)
tenant.ts	TenantContext + requireTenant() guard
supabase.ts	Database client + types
adminAuth.ts	Admin authenticatie
planFacts.ts	Support AI feiten + FAQ
stripe.ts	Stripe integratie
tenant-templates.ts	Menu templates per type
translations.ts	Website vertalingen (NIET WIJZIGEN)
modules.ts	Module definities
BusinessContext.tsx	React business context
LanguageContext.tsx	React taal context

6. DATABASE SCHEMA (SUPABASE)

Tabel	Scope	Kolommen
businesses	Hoofd tenant..	id, name, type, phone, email, address, ope..
services	Diensten per..	id, business_id, name, duration_minutes, p..
staff	Medewerkers ..	id, business_id, name, working_hours (JSON..
appointments	Afspraken pe..	id, business_id, service_id, staff_id, cus..
conversations	Gesprekslog	id, business_id, caller_phone, transcript,..
menu_items	Menu items (..	id, business_id, name, price, category, is..
orders	Bestellingen	id, business_id, customer_name, total_amou..
order_items	Bestelregels	id, order_id, menu_item_id, quantity, unit..
voice_sessions	Gesprekssess..	conversation_id, tenant_id, state, collect..
knowledge_base	Kennisbank	id, business_id, category, title, content
phone_numbers	Nummers per ..	telefoonnummer configuratie
pool_numbers	Pool nummers	gedeelde nummers
call_logs	Call tracking	gesprek metadata
usage_monthly	Maandverbruik	minuten per tenant per maand
forwarding_n..	Doorverbindnrs	doorverbind configuratie

7. WAT GEBOUWD MOET WORDEN

FASE 1: Custom LLM Endpoint (week 1-2)

Bouw een OpenAI-compatibel endpoint waar Vapi elk transcript naartoe stuurt. Onze backend beslist wat de AI zegt.

Bestand	Functie
src/app/api/vapi/chat/route.ts	Custom LLM endpoint - Vapi stuurt tran..
src/lib/appointment-engine/Sta..	Deterministische state machine (GREETI..
src/lib/appointment-engine/NLU..	Natural Language Understanding - extra..
src/lib/appointment-engine/Ses..	Session state opslag per call (Supabas..
src/lib/appointment-engine/Ava..	Inline beschikbaarheid check (openings..
src/lib/appointment-engine/Slo..	Soft lock mechanisme (lock -> confirm ..
src/lib/appointment-engine/Res..	Natuurlijk antwoord generatie per state
src/lib/appointment-engine/SMS..	SMS bevestiging na boeking via Twilio

State Machine States

GREETING	-> Begroeting, vraag wat klant wil
COLLECT_SERVICE	-> Welke behandeling?
COLLECT_DATE	-> Welke dag?
COLLECT_TIME	-> Welk uur?
COLLECT_NAME	-> Naam van de klant?
CHECK_AVAILABILITY	-> Inline check in Supabase
CONFIRM	-> Bevestiging vragen aan klant
BOOK	-> Afspraak opslaan + SMS
SUCCESS	-> Bevestiging uitspreken
RESCHEDULE	-> Verplaatsen flow
CANCEL	-> Annuleren flow
ESCALATE	-> Doorverbinden naar mens
ERROR	-> Foutafhandeling

FASE 2: Business Logic & Validatie (week 3-4)

- Openingsuren check - businesses.opening_hours JSONB
- Medewerker beschikbaarheid - staff.working_hours + bestaande appointments
- Diensten dynamisch laden - services tabel per tenant
- Annuleer afspraak flow - zoek op naam+datum, bevestig, cancel
- Verplaats afspraak flow - cancel oud + book nieuw
- Nearest alternatives bij conflict - 2 dichtstbijzijnde vrije slots

FASE 3: SMS & Monitoring (week 5-6)

- SMS bevestiging na boeking (Twilio)
- SMS annulering (klant antwoordt ANNULEER)
- Booking metrics per tenant per dag
- Slack alerts bij failures (accuracy < 98%, escalatie > 3%)
- Prompt finetuning per state (50+ test calls)

FASE 4: Multi-Tenant Productie (week 7-8)

- Automatische prompt generatie per tenant uit DB
- Vapi EU assistant aanmaken per tenant via API
- Telnyx nummer provisioning per tenant
- Onboarding wizard in dashboard (7 stappen)
- Stripe betalingen + abonnementenbeheer

8. VAPI CUSTOM LLM CONFIGURATIE

De Vapi EU assistant moet geconfigureerd worden om ONZE Custom LLM URL te gebruiken:

Via Vapi Dashboard

- Login op dashboard.vapi.ai
- Ga naar assistant 7a57ac55-1ca0-4395-a78b-7c78d4093d78
- Onder 'Model': kies 'Custom LLM'
- URL: <https://www voxapp.tech/api/vapi/chat>
- Opslaan

Via API

```
PATCH https://api.eu.vapi.ai/assistant/7a57ac55-...
Authorization: Bearer <VAPI_API_KEY>
{
  "model": {
    "provider": "custom-llm",
    "url": "https://www.voxapp.tech/api/vapi/chat",
    "model": "voxapp-orchestrator"
  }
}
```

Vapi Instellingen

Instelling	Waarde
silenceTimeoutSeconds	20
responseDelaySeconds	0.4
backchannelingEnabled	false
backgroundSound	off
firstMessage	Hallo, waarmee kan ik u helpen?
voice.provider	11labs
voice.model	eleven_flash_v2_5

9. MULTI-TENANT REGELS

- Elke request MOET tenant_id / business_id bevatten
- Request zonder tenant -> HARD FAIL (throw TenantError)
- Elke database query filtert op business_id
- Geen DEFAULT_TENANT_ID fallback in productie
- Geen cross-tenant data leakage
- Alle metrics per tenant
- Alle logs bevatten tenant_id
- voice_sessions zijn per call_id + business_id

10. BEVEILIGING

Endpoint Authenticatie

Type	Methode	Functie
Webhook endpoints	x-webhook-secret header	verifyWebhookSecret()
Admin endpoints	Admin cookie of x-api-key	verifyAdminApiKey()
ElevenLabs	x-webhook-secret header	verifyElevenLabsSecret()

Beveiligingsregels

- Geen hardcoded credentials in code
- Alle secrets in .env.local (lokaal) en Vercel env vars (productie)
- .env.local NOoit committen naar git
- Geen PII in logs (namen, telefoonnummers redacten)
- Rate limiting op publieke endpoints

11. CONVERSATIE KWALITEITSREGELS

Toon

Warm, rustig, Vlaams-Nederlands. Niet formeel. Niet overdreven.

Regels

- Maximaal 2 zinnen per antwoord
- Een vraag tegelijk
- Herhaal altijd kritieke info (datum, tijd, naam)
- Nooit Engels spreken
- Geen commerciële taal
- Geen uitroeptekens
- Geen robotzinnen

Verboden Zinnen

- Dit duurt maar een seconde
- Geen probleem
- Absoluut
- Zeker weten
- Super / Geweldig / Perfect

Filler bij Latency > 800ms

- Momentje...
- Even kijken...
- Ik controleer dat even.

Dag+Uur Combo Herkenning

Als de klant dag EN uur tegelijk zegt (bijv. 'dinsdag om 10 uur'), herken beide en sla de aparte tijdvraag over.

12. GIT & DEPLOY REGELS

- Na ELKE taak: git add . && git commit && git push origin main
- Commit messages in het Nederlands
- Gebruik required_permissions: ['all'] voor git push
- Deploy gaat automatisch via Vercel na push
- NOOIT lokaal bouwen of testen
- NOOIT force push naar main
- NOOIT git config wijzigen
- NOOIT --no-verify gebruiken
- Test altijd via productie URL: <https://www.voxapp.tech>

Vercel Environment Variables (vereist)

- NEXT_PUBLIC_SUPABASE_URL
- NEXT_PUBLIC_SUPABASE_ANON_KEY
- SUPABASE_SERVICE_ROLE_KEY
- ELEVENLABS_API_KEY
- ELEVENLABS_WEBHOOK_SECRET
- TELNYX_API_KEY
- VAPI_API_KEY
- VAPI_API_BASE=https://api.eu.vapi.ai
- VAPI_ASSISTANT_ID
- VAPI_WEBHOOK_SECRET
- OPENAI_API_KEY (NIEUW - voor eigen LLM calls)
- TWILIO_ACCOUNT_SID (voor SMS)
- TWILIO_AUTH_TOKEN (voor SMS)
- TWILIO_PHONE_NUMBER (voor SMS)

13. PRICING MODEL

Plan	Prijs	Minuten	Extra/min	Marge
Starter	E99/mnd	300	E0,40	~75%
Pro	E149/mnd	750	E0,35	~80%
Business	E249/mnd	1500	E0,30	~83%

Kostprijs per klant: ~E25/maand

14. KEY FILES REFERENTIE

Bestand	Functie
PROJECT_STATUS.md	SINGLE SOURCE OF TRUTH - altijd updaten
ENTERPRISE_PROJECT_CONSTITUTIO..	Bindende enterprise regels
.cursorrules	Agent instructies
src/lib/apiAuth.ts	Alle auth functies
src/lib/apiLogger.ts	Structured logging
src/lib/logger.ts	Call logging
src/lib/tenant.ts	Tenant enforcement
src/lib/supabase.ts	Database client + types
src/app/api/appointments/save/..	Huidig booking endpoint
src/app/api/health/route.ts	Health check (cron)
src/app/api/telnyx/failover/ro..	Failover
vercel.json	Cron configuratie

SAMENVATTING

- VoxApp is een AI-receptionist platform voor Belgische KMO's.
 - 39 API routes en 15+ database tabellen staan klaar.
 - Vapi EU wordt gebruikt als audio-laag (STT + TTS).
 - De conversatie-logica wordt EIGEN gebouwd (Custom LLM endpoint).
 - Deterministische state machine - code beslist, niet AI.
 - ALLES via EU endpoints - geen US routing.
 - Multi-tenant architectuur met harde tenant isolation.
 - 4 bouwfases: Custom LLM -> Validatie -> SMS -> Onboarding.
 - Na elke taak: git commit + push naar GitHub -> Vercel deploy.
 - Marketingwebsite is AFGESLOTEN - niet wijzigen.
-

VoxApp - www.voxapp.tech - Vertrouwelijk Document