FORMBUDDY / DETAILED TECHNICAL ARCHITECTURE

From spoken answer to reviewable form

As-built system map · browser, orchestration, domain rules, AI, persistence, safety, and operations

1 audio · 2 partial text · 3 answers · 6 history

01 ENTRY SURFACES

What the person can open and provide

Chrome extension
Launcher to hosted app

→ **Hosted web app**
Landing + guided interview

→ **Guided DL-14A**
Speak or type answers

→ **Any fillable PDF**
Upload, inspect, interview

← **Input devices**
Microphone · keyboard · file

02 BROWSER EXPERIENCE

React routes, visible controls, and local state

A Guided form screen
Voice panel · transcript
form progress · profile
review + PDF download

→ **B Any-form screen**
Upload · planning status
conversation · progress
PDF preview + download

→ **C Interview orchestration**
useInterview · useAnyForm
stage, answers, corrections
progress, reset, simulation

→ **D Voice orchestration**
useVoiceSession
energy VAD · mic mute during TTS
partial + final transcript

→ **E Local state**
Profile + preferences
session pointer · draft PDF
object URLs + UI state

BROWSER / LOCAL

03 DOMAIN + MEDIA PIPELINES

Two interview engines share voice, extraction, validation, and review

1 Audio capture
Raw browser PCM
16 kHz mono WAV

→ **2 Streaming transcript**
SSE partial words
final caller text

→ **3 Structured extraction**
Only clearly stated values
normalize date, phone, email

→ **4 Rules + validation**
Never invent · corrections win
sensitive fields are skipped

→ **5 Question decision**
Ask missing related fields
use natural or scripted wording

→ **7 Spoken reply**
Streaming TTS
browser fallback

GUIDED ENGINE · Texas DL-14A
Deterministic stage machine · field map · anti-repeat logic
correction rewind · completion gate

DYNAMIC ENGINE · Any fillable PDF
Read AcroForm fields · AI plan · fallback labels
FormInterview state · write answers to original PDF

SAFE FALLBACK
Rules and scripted questions continue when AI wording is unavailable

04 SERVER + AI SERVICES

Protected internal functions and public streaming audio endpoints

F TanStack server functions
Extract DL-14A answers
Plan uploaded forms
Extract generic answers
Compose follow-up
Save / load conversation

→ **G Public audio routes**
/api/transcribe → SSE
/api/tts → streamed audio
request validation + errors

→ **H AI gateway helper**
Server-side key boundary
structured prompts
JSON response parsing

→ **I Lovable AI Gateway**
Gemini 3.5 Transcribe
Gemini Flash extraction + planning
Gemini 3.1 Flash TTS · Kore

→ **6 Conversation persistence**
form_sessions
conversation_turns
text + extracted fields + decision note

SERVER-REACHABLE

05 DATA, TRUST + OPERATIONS

Where information lives, how it is constrained, and how the product is served

Browser-only memory
Whitelisted personal fields
short questions · pace · language
localStorage; forget control

Cloud database
RLS enabled · no client policies
server-admin persistence only
session → ordered turns

Safety boundary
Reject SSN, passwords, routing
and credit-card fields
prepare — never submit

App protections
CSRF request middleware
server error boundary
auth token middleware available

Delivery + interfaces
Edge-hosted TanStack app · sitemap
MCP form-field tools + discovery metadata
extension ZIP · runtime error reporting

FINAL OUTPUT Updated PDF bytes → live preview → download → human review → submission through the official channel