

## 1 - The five stages every meeting bot runs

- 1. Join** Headless browser (or native API) enters the call as a participant.
- 2. Capture** Per-speaker audio arrives over WebRTC; stream IDs keep voices apart.
- 3. Recognise** ASR turns audio into raw text - a commodity engine; none win here.
- 4. Understand** Diarization labels who spoke; an LLM writes the summary + actions.
- 5. Deliver** Push to CRM / docs / chat and fire a webhook to your software.

## 2 - The four products at a glance (2026 - verify before you commit)

Otter	in-house ASR, live transcript	-> Pro \$8.33/user/mo; API on Enterprise
Fireflies	AskFred + AI-credit metering	-> Pro \$10, Business \$19/user/mo
Fathom	free unlimited rec, paid AI	-> Premium ~\$15/user/mo (5 free AI/mo)
Supernormal bot OR bot-free OR extension		-> Starter \$16/mo; SOC2 + HIPAA

Differences live at stage 1 (capture) and stage 5 (packaging + price), not in features.

## 3 - Build it yourself: three paths + the cost math

- A. Rent a capture API Recall.ai -> bot into any platform, audio+transcript via webhook.
- B. Native platform API Zoom RTMS / Google Meet Media API -> no extra participant, lowest latency, but one platform only + host-gated.
- C. Extend your pipeline LiveKit agent / SFU-side ASR -> no per-meeting fee, you own data; only when the meeting is already your product.

Recall.ai math: 8,000 hrs/mo x \$0.50 capture = \$4,000 + x \$0.15 transcribe = \$1,200  
= \$5,200/mo, before your own summary-model + storage cost.

## 4 - Consent checkpoints (engineering context, not legal advice)

- Disclose the AI before it engages - EU AI Act Article 50 makes this a hard requirement.
- A visible bot named 'Notetaker' is NOT, by itself, the consent the law requires.
- All-party-consent states (CA, IL, PA + 8 more) need everyone to agree before recording.
- Diarization builds a voiceprint - biometric data; Illinois BIPA may require written opt-in.
- Keep a retention + deletion schedule and give participants a real way to decline.