

Meeting Bot Engineering - Cheat Sheet

Fora Soft Learn · AI for Video Engineering

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)

Product	What it is	Price
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:

$8,000 \text{ hrs/mo} \times \$0.50 \text{ capture} = \$4,000 + x \$0.15 \text{ transcribe} = \$1,200 = \$5,200/\text{mo}$, before your own summary-model + storage cost.

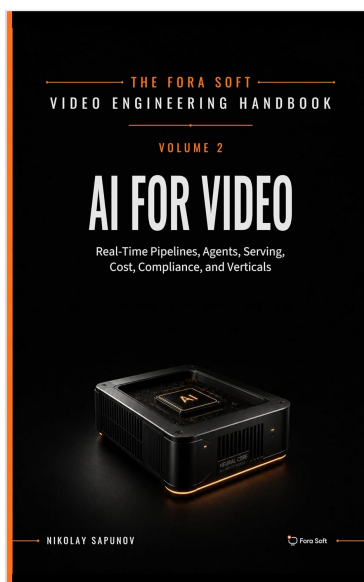
4 - Consent checkpoints (engineering context, not legal advice)

- ☐ 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.

Companion to: 'Otter, Fireflies, Fathom, Supernormal - AI Meeting Bot Engineering Deep-Dive'

www.forasoft.com/learn/ai-for-video-engineering/articles-ai/otter-fireflies-fathom-supernormal-ai-meeting-bot-engineering-deep-dive

Fora Soft · www.forasoft.com



GO DEEPER · VOLUME 7 OF 7

AI for Video: A Complete Guide to Intelligence in Video Products

From Real-Time Pipelines and Agents to Serving, Cost, Compliance, and Verticals

By Nikolay Sapunov, CEO at Fora Soft · 1,221 pages · Kindle & paperback on Amazon

[Read it on Amazon →](#)
amazon.com/dp/B0GX3DGBH4

This sheet maps the five bot stages, the three build paths and the Recall.ai cost math. Volume 7 covers what happens when that bot runs live and at scale: where AI enters a WebRTC pipeline, live captions, serving with vLLM or Triton, cost per stream-hour, and EU AI Act and biometric rules.

WHAT YOU WILL FIND INSIDE

- Where AI plugs into a WebRTC call pipeline
- Live captions and translation under a latency budget
- Serving ASR and LLMs: vLLM, Triton, TensorRT
- Cost per stream-hour for bot fleets
- EU AI Act and biometric rules for voiceprints

THE FULL HANDBOOK · 7 VOLUMES · 9,000+ PAGES · ALL ON AMAZON

Vol. 1 · Video Encoding — pixels and color to H.264, HEVC, AV1, quality metrics and AI

Vol. 2 · Video Streaming I — network transport, contribution and ingest, HLS, MPEG-DASH, CMAF

Vol. 3 · Video Streaming II — ABR, CDNs, players, WebRTC, DRM and the economics of delivery

Vol. 4 · Audio for Video I — PCM, codecs, loudness, accessibility, real-time voice processing

Vol. 5 · Audio for Video II — jitter buffers, lip sync, Atmos, quality metrics, neural codecs

Vol. 6 · AI for Video I — models and inference, computer vision, speech, search, generative video

Vol. 7 · AI for Video II — real-time pipelines, agents, serving, cost, compliance, verticals

Deciding between Recall.ai, native APIs and your own meeting bot?

We will model your monthly hours on each path and give you a build plan with the cost per meeting-hour and the consent flow. 250+ video products shipped since 2005.

[Book a 30-min call →](#)
[WhatsApp →](#)
[Email us →](#)