

Meeting Copilot Agent - Scoping & Guardrails

Fora Soft Learn · AI for Video Engineering

1 - The copilot loop (the agent loop, pointed at a live meeting)

Listen (STT) -> Decide -> Plan -> Call a tool -> Observe -> Memory (many times a minute)

Then: Draft a result (task / card / recap) -> Human approves (the agent never sends alone). Planning + tool use + memory: the three primitives fire on every turn of the loop.

2 - The six-tool belt (one listens nonstop; a few help live; one writes the recap)

Tool	Job	When it runs
Live transcriber	speech -> tagged text	continuous
Retriever	pull a fact (CRM / docs)	on demand
Action-item writer	spot a commitment	on demand
Task / cal connector	write to tracker / CRM / cal	AFTER human approval
Responder	voice (TTS) or sidebar card	when helping
Recap writer	summarize decisions+actions	after the call

3 - The one-second latency budget (a spoken copilot lives under a clock)

Rule	People talk over the agent past ~1,000 ms per turn.
Cascade:	STT ~300 + LLM ~500 + TTS ~250 + network ~150 = ~1,200 ms (already over).
Fix 1:	one speech-to-speech model ~320-800 ms (fits the budget).
Fix 2:	make the copilot silent - a card, not a voice (no clock at all).
Split:	live = transcribe / retrieve / prompt; after-call = recap / email / CRM writes.

4 - Consent & EU AI Act lanes (design the copilot to clear the lines by default)

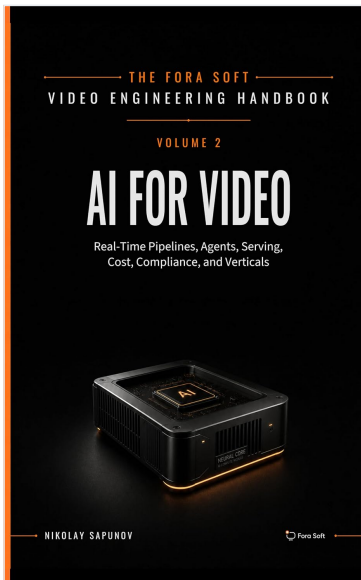
CONSENT	11 US all-party states (CA, IL, FL, PA, WA...). A bot in the list is NOT consent.
PLATFORMS	Google Meet / Teams flag external bots (2026). Use official media APIs or your app.
TRANSPARENCY	EU AI Act Art. 50: tell people they are interacting with an AI (from 2 Aug 2026).
PROHIBITED	EU AI Act Art. 5: inferring emotions in the workplace is banned. Do not build it.
OVERSIGHT	Human approves every send / assign / update. Agent proposes; person decides.

5 - Scoping checklist (ask these before you build, buy, or wrap)

- ☐ Does a spoken copilot fit the 1-second turn budget, or is it silent (a card)?
- ☐ Is live work (transcribe / retrieve) split from after-call work (recap / email / CRM)?
- ☐ Is the tool set small and sharp (~6 tools), not a drawer of thirty connectors?
- ☐ Is there episodic memory, so the copilot walks in knowing the relationship, not just the meeting?
- ☐ Do all participants give clear consent before recording starts, and is it logged?
- ☐ Is scope kept to what was said and decided - never scoring how people felt?
- ☐ Does a human approve every send, assign, and update? The agent never acts alone.

Companion to: 'Meeting Copilot Agent - Scoping & Guardrails Checklist'

www.forasoft.com/learn/ai-for-video-engineering/articles-ai/meeting-copilot-agent-conferencing



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 checklist scopes a copilot against the one-second turn budget, the tool belt and the consent lanes. Volume 7 treats agents as engineering: loop primitives, frameworks, memory and tools, plus call latency, serving infrastructure, cost per stream-hour and EU AI Act compliance for a copilot that runs during the meeting.

WHAT YOU WILL FIND INSIDE

- Agent loop primitives, memory and tool design
- Call latency budget and speech-to-speech options
- Serving the copilot: vLLM, Triton, TensorRT
- EU AI Act Article 50 and Article 5 in practice

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

Scoping a meeting copilot that has to answer inside one second?

We will check your latency cascade and tool belt, then send an architecture with a human approval step and a consent flow. 250+ video products shipped since 2005.

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