

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)

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)

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.