

Echo Cancellation Selection Checklist

Classical vs AI hybrid, the double-talk weak spot, the G.114 budget, and a pre-flight before you ship

1 · Classical AEC vs AI hybrid at a glance

Stage	What it does	Used in	Adds	Note
Adaptive filter	Predicts + subtracts linear echo	Both	~0-2 ms	Core of every canceller
Classical back end	Hand-tuned residual + freeze on double-talk	Classical	low	Weak: nonlinear + double-talk
AI neural back end	Learns to separate near-end speech from echo	AI hybrid	up to ~20 ms	Strong: nonlinear + double-talk
Browser AEC3	Free Chromium built-in (classical hybrid)	Baseline	low	Try this first

Echo cancellation removes the FAR end's voice leaking back through your speaker — not your own background noise (that is noise suppression). Its superpower is the reference signal: a clean copy of the sound to remove.

2 · The latency budget — fit it under ITU-T G.114 (≤ 150 ms one-way)

Pipeline stage	Adds	Running total vs 150 ms
Network + capture/encode/decode (typical baseline)	≈ 110 ms	—
+ Classical canceller (AEC3)	+ ~2 ms	= 112 ms ✓ room to spare
+ AI canceller (AEC-Challenge budget)	+ up to 20 ms	= 130 ms ✓
Conversation target	—	≤ 150 ms (G.114)

3 · Pre-flight before you ship an echo-cancellation feature

- Tried the browser built-in first: echoCancellation: true on getUserMedia (AEC3) — if echo complaints stop, ship it and stop.
- Identified the echo path: headset/earbuds (weak, easy) vs loud open speaker near the mic (hard — may need AI).
- Checked the hardware: cheap distorting speakers make nonlinear echo classical cancellers miss → points to AI.
- Stress-tested double-talk (both parties speaking) — the case that separates a good canceller from a poor one.
- Ran the G.114 latency arithmetic: classical adds ~0-2 ms; an AI canceller up to ~20 ms — confirmed the headroom fits.
- Decided self-host vs buy: open (DTLN-AEC, SpeexDSP) you operate, vs paid SDK (Krisp, NVIDIA) the vendor operates.
- Avoided stacking two cancellers: did not run an AI canceller on top of the browser's — turn one off (echoCancellation: false).
- Fed the canceller the raw reference + mic signals it was designed for — not pre-processed audio from another stage.
- Measured quality with AECMOS (echo annoyance scored separately), not a vendor before/after demo clip.
- Checked ERLE only as the single-talk easy case — confirmed it does not hide a mangled near-end voice.
- Verified the canceller stays real-time on your target device, with headroom for other audio work.
- Confirmed the privacy path: on-device cancellers (AEC3, Krisp, NVIDIA local) keep audio off the network.