Captioning & Transcription Compliance Checklist

Run a real-time captioning or transcription feature through the four gates before you ship it. Engineering guidance, not legal advice — confirm specifics with counsel.

GATE 1 · ACCESSIBILITY (when captions are a legal duty)

- [ ] Captions available on demand for deaf/hard-of-hearing patients at no cost (ADA 28 CFR 36.303; Section 1557, 45 CFR Part 92)
- [ ] Patient given 'primary consideration' for the aid they request
- [ ] WCAG 2.1 Level AA as the technical bar; SC 1.2.4 captions identify the speaker
- [ ] Public-entity customers: WCAG 2.1 AA by Apr 26 2027 / 2028 (ADA Title II)
- [ ] Webinars, group sessions, and recorded patient education captioned too

GATE 2 · PHI BOUNDARY (where the audio goes)

- [ ] Signed BAA on file with the ASR vendor before any audio is sent (45 CFR 164.502(e))
- [ ] BAA covers the exact service/tier you call — not just 'the vendor offers one'
- [ ] Browser built-in speech API and free consumer tiers blocked for clinical audio
- [ ] Audio encrypted in transit and at rest; correct region and logging switched on
- [ ] Self-build? A BAA with every vendor in the chain that receives audio

RED FLAGS (stop before you ship)

- [ ] Browser speech recognition wired up for speed — audio leaves with no BAA
- [ ] A transcript pasted into a free public AI assistant to 'clean it up'
- [ ] A misheard drug name in a caption read and trusted as written
- [ ] Treating live captions as the medical record without clinician verification

GATE 3 · ACCURACY + GATE 4 · RETENTION

- [ ] Medical-tuned model chosen — not a general consumer model
- [ ] Accuracy measured on clinical entities (drug names, diagnoses), not just overall WER
- [ ] Speaker labels (diarization) verified — no symptom attributed to the wrong person
- [ ] Unreviewed transcript never stands in for the clinician-verified record
- [ ] Decide deliberately what you keep: ephemeral captions vs a saved transcript
- [ ] Saved transcript treated as PHI at rest and part of the designated record set (164.501)
- [ ] Recording/transcript consent captured; retention period defined and defensible

THE ONE-LINE RULE

Real-time transcription runs the front half of the scribe pipeline in a hurry: streaming speech recognition turns the spoken visit into live captions a patient reads on screen and, optionally, a saved transcript. The captions are frequently a legal duty — deaf and hard-of-hearing patients have a right to effective communication, with WCAG 2.1 AA as the bar — and every word produced is PHI, so the speech-recognition vendor needs a signed BAA before any audio reaches it. Pick a medical-tuned model, measure its accuracy on the words that can hurt someone, decide deliberately whether you keep the transcript at all, and never let an unreviewed transcript stand in for the medical record. The recognition engine is a commodity you can buy; placing it correctly relative to the accessibility duty and the PHI boundary is the work that keeps a launch — and the patients who depend on captions — well served.