

Tier Information

Hi-Fi, Studio-Grade

Name	Hi-Fi, Studio-Grade
Tier	Premium
Data modalities	<div style="display: flex; gap: 10px;"> Audio Video Audiovisual </div>
Dataset types	<p>Full-duplex conversational sessions (two-speaker, isolated channels)</p> <p>Monologue ASR (single-speaker read or spontaneous speech)</p> <p>Narrated TTS (single-speaker performance reads for voice cloning)</p> <p>Multilingual read prompts (phonetically balanced sentence sets)</p> <p>Video-only captures (4K UHD, no audio track)</p> <p>Audiovisual sessions (sync'd 4K video + Hi-Fi audio)</p>
File formats	<p>FLAC — lossless audio (default)</p> <p>WAV — uncompressed audio (on request)</p> <p>MP4 (H.264) — video, up to 4K UHD</p> <p>MOV (ProRes) — video master, on request</p>
Sample rate	48 kHz
Bit depth	24-bit
Channel layout	<p>Per-speaker stereo on conversational sessions</p> <p>Mono or stereo on single-speaker sessions</p> <p>Sync'd multi-channel on audiovisual sessions</p>
Video specs	Up to 4K UHD (3840×2160) at 30 / 60 fps on video and audiovisual sessions — color-graded master plus delivery proxy
Microphones	Studio-grade wired — documented per session

Cameras	4K-capable mirrorless / cinema cameras with fixed framing per session; lens, sensor, and lighting documented per recording
Noise floor	Calibrated per session with a pre-recording pass/fail gate — sessions that miss the threshold don't ship
Languages	14 shipping (American English, French, Mandarin, Spanish, +10) · bespoke languages and dialects on request
Speaker metadata	Age, gender, region, dialect, native language, acoustic environment
Annotations	Word-level transcripts Diarization and speaker turns Prosodic markers and disfluency tags Scenario / role labels Custom annotation layers on request

Properties listed here apply to every dataset shipped at the Hi-Fi tier. Per-dataset deviations (additional languages, dialects, or annotation layers) are documented on the relevant marketplace listing.