

□□□□□ · □□□□□□ 2026

प्रदर्शन

ये आँकड़े 14 अक्टूबर 2026 को ऑन-प्रिमाइस प्रोफ़ाइल पर मापे गए, जब चार आउटपुट भाषाएँ एक साथ चल रही थीं (अरबी, अंग्रेज़ी, फ़्रेंच, जर्मन)।

■■■■■■■■ ■■■ ■■■■■■

क्या मापा गया: वाक्य खत्म होने के क्षण से लेकर उसका व्याख्या किया गया ऑडियो उस भाषा को सुन रहे फ़ोनो तक भेजे जाने तक। कमरे के वाई-फ़ाई पर डिलीवरी कुछ मिलीसेकंड जोड़ती है। हर फ़ोन अपना कुल विलंब भी दिखाता है («लाइव से x से. पीछे»)।

■■■	■■■
सामान्य विलंब	वक्ता से दो से चार सेकंड पीछे
मापी गई बैठक की माधिका	3.0 से.
वाक्य	12 बोले गए, 48 डिलीवरी (12 × 4 भाषाएँ)
हानि	0 छोड़े गए · 0 ट्रांसक्राइब नहीं हुए · 0 अनुवाद नहीं हुए · 0 कभी नहीं बोले गए

मापी गई बैठक में लगभग 25 शब्दों तक के लंबे, औपचारिक वाक्य थे। छोटे, बातचीत वाले वाक्य तेज़ होते हैं।

■■■■ ■■■■■■■■■■ ■■(■■■ ■■■■■)

■■■	■■■
वाक् पहचान	~0.6 से.
अनुवाद (एक साथ चार भाषाएँ)	~2.2 से.
आवाज़ संश्लेषण	~0.4 से.

वाक्य खत्म होने से पहले उसका अनुवाद नहीं हो सकता: वाक् पहचान उसे आगे देने से पहले एक छोटी चुप्पी का इंतज़ार करती है। यही इंतज़ार और वाक्य की लंबाई बताते हैं कि लंबे वाक्य थोड़ा ज़्यादा समय क्यों लेते हैं।

○○○○○○○○○○

ऐप्लिकेशन विंडो दिखती है	~1-2 से.
बैठक के लिए तैयार	35-45 से.

○○○○○○

- **श्रोता:** काम हर भाषा के लिए एक बार होता है, इसलिए श्रोता जोड़ने से केवल नेटवर्क ट्रैफिक बढ़ता है। एक भाषा पर तीस श्रोताओं की लागत तीन के बराबर है। व्यावहारिक सीमा कमरे का वाई-फ़ाई एक्सेस पॉइंट है।
- **भाषाएँ:** एक साथ चार आउटपुट भाषाओं तक।
- **लाइव के क़रीब रहना:** अगर कोई फ़ोन 20 सेकंड से ज़्यादा पीछे हो जाए, तो वह कतार में लगे वाक्य छोड़ देता है — कभी भी बोला जा रहा वाक्य नहीं — और श्रोता को बताता है। उससे पहले वह पकड़ने के लिए आवाज़ थोड़ी तेज़ करता है।

○○○○○○○○○○

ग्राफ़िक्स	न्यूनतम NVIDIA RTX 3090 (24 GB) · अनुशंसित RTX 5090 (32 GB)
डिस्क	~60 GB खाली जगह अनुशंसित
नेटवर्क	कमरे के वाई-फ़ाई पर हर श्रोता के लिए एक ऑडियो स्ट्रीम