

Reconocimiento de voz en el dispositivo con NVIDIA Nemotron

Nemotron es el nuevo motor de reconocimiento de voz en el dispositivo de Hedy: totalmente local, más preciso que antes y etiqueta quién dijo qué.

Publicado por Julian Pscheid · 23 de junio de 2026

[Leer este artículo en línea: https://www.hedy.ai/es/post/nemotron-on-device-speech-engine/](https://www.hedy.ai/es/post/nemotron-on-device-speech-engine/)



Seis colegas de diferentes etnias en una conversación activa alrededor de una mesa de conferencias, con una persona hablando y gesticulando mientras las demás escuchan

Respuesta rápida Nemotron es un nuevo motor de reconocimiento de voz en el dispositivo en Hedy, basado en el modelo de voz Nemotron de NVIDIA. Se ejecuta por completo en su dispositivo, transcribe con más precisión que los motores locales anteriores y etiqueta quién dijo qué para que una conversación con varias personas sea fácil de leer después. Lo encontrará en Settings, dentro de la sección Speech & AI, en versiones English Only y Multilingual. Se suma a Whisper y Parakeet y reemplaza a Parakeet de ahora en adelante.

Dos problemas han acompañado a la transcripción local durante mucho tiempo. La precisión quedaba por detrás de la nube, y la transcripción de una conversación entre varias personas volvía como un bloque de texto sin dividir que luego era difícil de leer. Nemotron ayuda con ambos, y se ejecuta completamente en su dispositivo.

Nemotron es el modelo de reconocimiento de voz en el dispositivo de NVIDIA (<https://huggingface.co/nvidia/nemotron-3.5-asr-streaming-0.6b>). NVIDIA lo lanzó en junio de 2026, y Hedy está entre las primeras aplicaciones que lo llevan a un producto publicado en lugar de una demo. Ahora está junto a Whisper y Parakeet como uno de los motores de voz que puede elegir dentro de Hedy.

Un salto claro en precisión local

La transcripción local siempre ha sido la opción adecuada para quienes prefieren mantener sus conversaciones en su propio hardware. El problema era la precisión: los modelos en el dispositivo iban por detrás de la nube, a veces lo suficiente como para notarlo. Nemotron reduce esa brecha. Para reuniones y llamadas cotidianas, la transcripción que produce supone un salto claro frente a lo que los modelos locales lograban antes, sin que nada salga de su dispositivo para conseguirlo.

Si quiere la visión más amplia de cómo encaja la transcripción en el dispositivo dentro de Hedy, vea nuestra visión general de IA local para reuniones (</es/post/local-ai-meetings-hedy-3-2/>) y el análisis técnico profundo de Hedy 3.2 (</es/post/local-ai-engineering-deep-dive-hedy-3-2/>).

Distingue a sus hablantes

El cambio más grande es cómo se ve después una transcripción de varias personas. La transcripción local anterior devolvía un muro de texto sin saber quién estaba hablando. Nemotron separa a los hablantes y los etiqueta: Speaker 1, Speaker 2, Speaker 3, y así sucesivamente. Una reunión de ida y vuelta se lee como una reunión de ida y vuelta.

Dónde aparecen esas etiquetas depende de su plataforma:

- En iPhone y Mac , las etiquetas aparecen en vivo mientras las personas hablan y la conversación pasa de una a otra.
- En Windows y Android , se agregan al final de la sesión durante el procesamiento, así que las verá cuando su transcripción esté lista en lugar de en tiempo real.

Dos versiones: English Only y Multilingual

Nemotron viene en dos variantes, y usted elige la que coincide con su forma de trabajar. La versión English Only está ajustada para conversaciones solo en inglés. La versión Multilingual maneja un conjunto amplio de idiomas desde un solo modelo, que es la opción adecuada si sus reuniones cambian de idioma o no son en inglés.

Encontrará ambas en Settings , dentro de la sección Speech & AI , junto a Whisper. Cambiar de motor requiere solo un par de toques. Para una comparación lado a lado de cada opción, vea cómo se comparan los motores de reconocimiento de voz de Hedy (</es/help/speech-recognition-providers-in-hedy/>).

Dejando atrás Parakeet

Nemotron reemplaza a Parakeet. Hace mejor el mismo trabajo, con una transcripción más precisa. Si hoy usa Parakeet, cambie a Nemotron: Parakeet se está retirando, y Nemotron es donde continúa ese trabajo. Whisper permanece donde está como el motor con la compatibilidad más amplia, y como una buena opción predeterminada si usa un dispositivo más antiguo o simplemente lo prefiere.

Qué necesita para ejecutarlo

En iPhone y iPad, Nemotron necesita un iPhone 12 o más reciente , o un iPad de esa generación en adelante. Los dispositivos Apple más antiguos no verán la opción. También se ejecuta en Mac, Windows y Android.

Abra Settings !' Speech & AI , elija Nemotron en modo English Only o Multilingual, y su próxima sesión se transcribirá en su dispositivo con los hablantes etiquetados. Más fácil de leer después, y nada salió de su dispositivo para conseguirlo.

Hedy AI · Coaching de IA en vivo para conversaciones importantes

Prueba Hedy gratis: <https://www.hedy.ai/es/downloads/>

<https://www.hedy.ai/es/post/nemotron-on-device-speech-engine/>