

NVIDIA Nemotron On-Device-Spracherkennung

Nemotron ist Hedys neue On-Device-Spracherkennungs-Engine: vollständig lokal, genauer als frühere lokale Engines, und sie kennzeichnet, wer was gesagt hat.

Veröffentlicht von Julian Pscheid · 23. Juni 2026

[Diesen Artikel online lesen: https://www.hedy.ai/de/post/nemotron-on-device-speech-engine/](https://www.hedy.ai/de/post/nemotron-on-device-speech-engine/)



Sechs Kolleginnen und Kollegen unterschiedlicher ethnischer Herkunft in einer aktiven Diskussion an einem Konferenztisch; eine Person spricht und gestikuliert, während die anderen zuhören

Kurzantwort Nemotron ist eine neue On-Device-Spracherkennungs-Engine in Hedy, basierend auf dem Nemotron-Sprachmodell von NVIDIA. Sie läuft vollständig auf Ihrem Gerät, transkribiert genauer als frühere lokale Engines und kennzeichnet, wer was gesagt hat, sodass ein Gespräch mit mehreren Personen anschließend leicht lesbar ist. Sie finden sie unter Einstellungen im Bereich Speech & AI, in den Versionen English Only und Multilingual. Nemotron ergänzt Whisper und Parakeet und ersetzt Parakeet künftig.

Zwei Probleme begleiten lokale Transkription schon lange. Die Genauigkeit blieb hinter der Cloud zurück, und das Transkript eines Gesprächs mit mehreren Personen kam als ein ungeteilter Textblock zurück, der später schwer zu lesen war. Nemotron hilft bei beidem und läuft vollständig auf Ihrem Gerät.

Nemotron ist das On-Device-Spracherkennungsmodell von NVIDIA (<https://huggingface.co/nvidia/nemotron-3.5-asr-streaming-0.6b>). NVIDIA hat es im Juni 2026 veröffentlicht, und Hedy gehört zu den ersten Apps, die es in ein veröffentlichtes Produkt integrieren, statt es nur als Demo zu zeigen. Es steht nun neben Whisper und Parakeet als eine der Spracherkennungs-Engines, die Sie in Hedy auswählen können.

Ein deutlicher Schritt nach vorn bei lokaler Genauigkeit

Lokale Transkription war schon immer die richtige Wahl für Menschen, die ihre Gespräche lieber auf ihrer eigenen Hardware behalten. Der Haken war die Genauigkeit: On-Device-Modelle lagen hinter der Cloud zurück, manchmal deutlich spürbar. Nemotron verkleinert diese Lücke. Für alltägliche Meetings und Anrufe ist das erzeugte Transkript ein deutlicher Schritt nach vorn gegenüber dem, was lokale Modelle bisher leisten konnten, und nichts verlässt dafür Ihr Gerät.

Wenn Sie ein umfassenderes Bild davon möchten, wie On-Device-Transkription in Hedy passt, lesen Sie unseren Überblick zu lokaler KI für Meetings (</de/post/local-ai-meetings-hedy-3-2/>) und den technischen Deep Dive zu Hedy 3.2 (</de/post/local-ai-engineering-deep-dive-hedy-3-2/>) .

Nemotron unterscheidet Ihre Sprecher

Die größere Änderung ist, wie ein Transkript mit mehreren Personen anschließend aussieht. Ältere lokale Transkription gab eine Textwand zurück, ohne zu zeigen, wer gesprochen hatte. Nemotron trennt die Sprecher und kennzeichnet sie: Speaker 1, Speaker 2, Speaker 3 und so weiter. Ein Meeting mit Hin und Her liest sich wie ein Meeting mit Hin und Her.

Wo diese Labels erscheinen, hängt von Ihrer Plattform ab:

- Auf iPhone und Mac erscheinen die Labels live, während Personen sprechen und das Gespräch zwischen ihnen wechselt.
- Auf Windows und Android werden sie während der Verarbeitung am Ende der Sitzung hinzugefügt. Sie sehen sie also, sobald Ihr Transkript fertig ist, nicht in Echtzeit.

Zwei Versionen: English Only und Multilingual

Nemotron gibt es in zwei Varianten, und Sie wählen die aus, die zu Ihrer Arbeitsweise passt. Die Version English Only ist auf rein englische Gespräche abgestimmt. Die Version Multilingual verarbeitet eine breite Auswahl an Sprachen mit einem einzigen Modell und ist die richtige Wahl, wenn Ihre Meetings zwischen Sprachen wechseln oder nicht auf Englisch stattfinden.

Sie finden beide unter Einstellungen , im Bereich Speech & AI , neben Whisper. Der Wechsel der Engine braucht nur ein paar Fingertipps. Einen direkten Vergleich aller Optionen finden Sie unter So vergleichen sich Hedys Spracherkennungs-Engines (</de/help/speech-recognition-providers-in-hedy/>) .

Abschied von Parakeet

Nemotron ersetzt Parakeet. Es erledigt dieselbe Aufgabe besser, mit genauerer Transkription. Wenn Sie heute Parakeet verwenden, wechseln Sie zu Nemotron -- Parakeet wird eingestellt, und Nemotron ist der Ort, an dem diese Arbeit weitergeht. Whisper bleibt genau dort, wo es ist: als die am breitesten kompatible Engine und als gute Standardoption, wenn Sie ein älteres Gerät nutzen oder sie einfach bevorzugen.

Was Sie dafür brauchen

Auf iPhone und iPad benötigt Nemotron ein iPhone 12 oder neuer oder ein iPad ab dieser Generation. Auf älteren Apple-Geräten wird die Option nicht angezeigt. Es läuft auch auf Mac, Windows und Android.

Öffnen Sie Einstellungen !' Speech & AI , wählen Sie Nemotron in English Only oder Multilingual, und Ihre nächste Sitzung wird auf Ihrem Gerät mit Sprecherlabels transkribiert. Danach leichter zu lesen, und nichts davon musste Ihr Gerät verlassen.

Hedy AI · Live-KI-Coaching für wichtige Gespräche

Hedy kostenlos testen: <https://www.hedy.ai/de/downloads/>

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