Intelligenz Welt

Nova 2.5 Sonic

KI-Modellportrait · Quellenstand 2026-10-31

NS

Kein geprüftes Modell- oder Anbieterlogo; Monogramm-Fallback (AWS-Logo auf Wikimedia Commons als urheberrechtlich geschützt und markenrechtlich markiert eingestuft, kein PD-Textlogo) Monogramm-Darstellung, kein Logo-Asset übernommen.

Speech-to-Speech-Modell von Amazon für Echtzeit-Sprachdialoge mit Werkzeugnutzung in Amazon Bedrock.

Quellen geprüft am 31. Oktober 2026 · Redaktion: Philipp Paulik

Entwickler: Amazon Web Services →

DAS MODELL EINFACH ERKLÄRT

Echtzeit-Sprachdialoge mit Werkzeugnutzung als verwalteter Bedrock-Dienst.

Nova 2.5 Sonic ist Amazons Speech-to-Speech-Modell für natürliche Sprachdialoge in Echtzeit. Es verarbeitet Sprache und Text in derselben Sitzung und erzeugt gesprochene und textuelle Antworten. Die Modellkarte nennt 256.000 Token Kontext und bis zu 64.000 Ausgabe-Token; der Zugang läuft ausschliesslich über die bidirektionale Streaming-API von Amazon Bedrock.

Offizielle Dokumentation: Nova 2.5 Sonic

Quellen geprüft: 31.10.2026. Herstellerangaben, kein eigener Modelltest.

Wofür eignet sich Nova 2.5 Sonic?

AWS positioniert Nova 2.5 Sonic für Sprachagenten etwa im Kundenservice: Bestellung nachschlagen, Berechtigung prüfen und Aktion auslösen in einem Gespräch. Das ist Anbieterberatung, kein hier ermitteltes Urteil. Prüfe einen kompletten Ablauf mit echten Aufnahmen: Dialekte, Hintergrundgeräusche, Unterbrechungen, Werkzeugaufrufe und die Kosten je Gesprächsminute.

Aufgabe und Modalitaeten

Nova 2.5 Sonic verbindet Sprachverstehen und Spracherzeugung in einem Modell, ohne die klassische Kette aus Transkription, Textmodell und Sprachsynthese. Sprache und Text gehen hinein, Sprache und Text kommen heraus; Bild, Video und Einbettungen sind laut Modellkarte nicht unterstützt. Cross-Modal-Eingabe erlaubt den Wechsel zwischen Sprechen und Tippen in derselben Sitzung.

Offizielle Dokumentation: Nova 2.5 Sonic

Zugang nur über bidirektionales Streaming

Unterstützt wird ausschliesslich InvokeModelWithBidirectionalStream am bedrock-runtime-Endpunkt. Die Sitzung bleibt als offener Kanal bestehen; Ereignisse wie sessionStart, promptStart und audioInput steuern den Ablauf. Als Service-Tier steht nur Standard zur Verfügung; Priority, Flex und Reserved sind laut Modellkarte nicht verfügbar.

Offizielle Dokumentation: Nova 2.5 Sonic

Regionen und Kontingente

Verfügbar ist das Modell In-Region in us-east-1 (N. Virginia), us-west-2 (Oregon), eu-north-1 (Stockholm) und ap-northeast-1 (Tokio); Geo- und Global-Inference werden nicht unterstützt. Das Standardkontingent liegt bei 20 gleichzeitigen Streaming-Sitzungen je Konto und Region, eine Erhöhung geht über das AWS-Kontoteam.

Offizielle Dokumentation: Nova 2.5 Sonic

Sprachen und Stimmen

Offiziell unterstützt sind Englisch (US, britisch, indisch, australisch), Spanisch, Deutsch, Französisch, Italienisch, Portugiesisch (Brasilien) und Hindi, mit insgesamt 18 Stimmen. AWS rät von anderen Sprachen ab, obwohl das Modell mehrsprachig trainiert wurde. Polyglotte Stimmen und Sprachwechsel mitten im Gespräch sind vorgesehen.

Offizielle Dokumentation: Nova 2.5 Sonic

Sicherheit und Wasserzeichen (Herstellerangaben)

Laut Servicekarte weicht das Modell über 95 Prozent schädlicher Prompts auf einem internen Datensatz aus und lehnt Versuche des Stimmklonens zu 100 Prozent ab; die Filter sind nicht konfigurierbar. Alle erzeugten Audiodaten erhalten ein unhörbares Wasserzeichen. Für Ausgaben allgemein verfügbarer Nova-Modelle nennt AWS eine IP-Freistellung in Abschnitt 50.10 der Service Terms. Diese Angaben sind Herstellerangaben, keine unabhängige Prüfung.

Offizielle Dokumentation: Nova 2.5 Sonic

Kontextfenster ist nicht Ausgabelimit.

256.000

Token Kontextfenster

64.000

maximale Ausgabe-Token, synchron



Zwei unterschiedliche Grenzen auf gemeinsamer Nullskala, nicht addieren. Offizielle Dokumentation: Nova 2.5 Sonic

Was kostet Nova 2.5 Sonic über die API?

Standard-Eingabe

3 USD

je 1 Mio. Texttoken · Basisrate

Standard-Ausgabe

12 USD

je 1 Mio. Texttoken · Basisrate

On-Demand-Preise in USD je 1 Mio. Token, Region us-east-1: Sprach-Eingabe 3,00, Sprach-Ausgabe 12,00, Text-Eingabe 0,33, Text-Ausgabe 2,75. Texttokens fallen etwa für Transkription, Werkzeugaufrufe und Gesprächshistorie an. Regionale Abweichungen (Sprache Eingabe/Ausgabe, Text Eingabe/Ausgabe): us-west-2 3,18/12,73/0,32/2,65; eu-north-1 3,33/13,31/0,36/2,99; ap-northeast-1 3,74/14,96/0,40/3,31. Abrechnung nach Token, nicht nach Gesprächsminuten; kein Abopreis, keine Gesamtkostenzusage. Stand: 31.10.2026.

Preisstand 31.10.2026. Offizielle Dokumentation: Nova 2.5 Sonic

Vor dem Einsatz prüfen

- Modell-ID auf amazon.nova-2-5-sonic umstellen; die API-Oberfläche entspricht laut AWS der von Nova 2 Sonic.
- Eigene Anwendungsfälle erneut testen: AWS empfiehlt ausdrücklich, nicht gleiche Ergebnisse wie beim Vorgänger anzunehmen.
- Verbindungslimit von 8 Minuten je Sitzung beachten und das Erneuerungsmuster für längere Gespräche einplanen.
- Nicht auf Invoke-, Converse- oder Chat-Completions-Endpunkte ausweichen; nur InvokeModelWithBidirectionalStream wird unterstützt.

Häufige Fragen zu Nova 2.5 Sonic

Was ist Nova 2.5 Sonic?

Ein Speech-to-Speech-Modell von Amazon für Echtzeit-Sprachdialoge in Amazon Bedrock, seit dem 5. Oktober 2026 allgemein verfügbar. Die Bedrock-Modell-ID lautet amazon.nova-2-5-sonic.

Was kostet Nova 2.5 Sonic?

In us-east-1 kosten Sprachtokens 3 USD je Million in der Eingabe und 12 USD je Million in der Ausgabe; Texttokens liegen bei 0,33 beziehungsweise 2,75 USD je Million. Die Preise unterscheiden sich je Region. Stand: 31.10.2026.

Wie groß ist das Kontextfenster von Nova 2.5 Sonic?

Die Modellkarte nennt 256.000 Token Kontext und bis zu 64.000 Ausgabe-Token. Zusätzlich gilt nach dem Nova-2-Leitfaden zur gemeinsamen Streaming-API ein Verbindungslimit von 8 Minuten je Sitzung, das ein Erneuerungsmuster erfordert.

Welche Sprachen und Stimmen unterstützt Nova 2.5 Sonic?

Offiziell unterstützt sind Englisch (US, britisch, indisch, australisch), Spanisch, Deutsch, Französisch, Italienisch, Portugiesisch (Brasilien) und Hindi mit insgesamt 18 Stimmen. Für andere Sprachen rät AWS vom Einsatz ab.

Kann man Nova 2.5 Sonic wie ein normales Chat-Modell aufrufen?

Nein. Der einzige unterstützte Weg ist die bidirektionale Streaming-API (InvokeModelWithBidirectionalStream). Invoke, Converse, Chat Completions und Responses werden laut Modellkarte nicht unterstützt.

Gibt es Nova 2.5 Sonic als offene Gewichte?

Nein. Das Modell ist ein proprietärer, verwalteter Dienst von AWS. Die Nutzung folgt dem AWS Customer Agreement und den AWS Service Terms.

Alle Antworten beziehen sich auf die verlinkte Herstellerdokumentation und den angegebenen Stand. Kein unabhängiger Benchmark, keine Garantie für Verfügbarkeit oder Ergebnisqualität.

Quellen & Stand

1. Amazon Bedrock: Modellkarte Nova 2.5 Sonic
   Geprüft: 2026-10-31
2. AWS AI Service Card: Amazon Nova 2.5 Sonic
   Geprüft: 2026-10-31
3. AWS What's New: Amazon Nova 2.5 Sonic allgemein verfuegbar (05.10.2026)
   Geprüft: 2026-10-31
4. AWS: Amazon Bedrock Pricing
   Geprüft: 2026-10-31
5. AWS Price List API: Amazon Bedrock, On-Demand-Preise (regional, USD)
   Geprüft: 2026-10-31

Redaktionelle Zusammenfassung mit KI-Unterstützung, verantwortet von Philipp Paulik. Kein eigener Leistungstest. Änderbare Anbieterangaben gelten für den angegebenen Quellenstand.

Intelligenz Welt · Redaktion: Philipp Paulik · Quellenstand: 2026-10-31
Vollständiges Webportrait: https://intelligenz-welt.de/ki-modelle/nova-2-5-sonic/
Statischer Export, kein Live-Datenblatt. Preise, Zugriff und Anbieterangaben können sich ändern. Quellen und ihre jeweiligen Abrufstände stehen oben. Impressum · Transparenz