

亚马逊推出全新Nova Sonic语音到语音模型

北京2025年4月9日 /美通社/ -- 亚马逊宣布推出一款全新的基础模型Amazon Nova Sonic，将语音理解与语音生成统一于单一的模型中，使AI应用程序中的语音对话更贴近真人交流。该模型通过Amazon Bedrock上的新API提供，可简化语音应用开发流程，例如客户服务通话自动化及覆盖旅游、教育、医疗、娱乐等领域的跨行业AI agents。

在过去十余年间，亚马逊始终引领着语音技术的发展，致力于为对话式AI应用赋能——从打造全球领先的个人AI助手Alexa，到开发亚马逊云科技的各类服务，如Lex、Polly和Connect等。但要让语音AI为客户创造更多实际价值，它必须能够理解人类对话的微妙与复杂性。在对话中，文字本身承载意义，但若没有声音语境赋予其深度，仅凭文字往往难以传达完整信息。如何表达与表达什么同样重要，甚至更为关键。从过去直到现在，通过AI实现这一点仍是巨大挑战。

一套捕捉语气、风格与节奏的语音系统

传统语音应用开发需复杂协调多个模型，例如将语音转为文字的语音识别模型、理解并生成回复的大语言模型（LLM）、将文字再转为音频的文本转语音模型。这种分散的方法不仅增加了开发的复杂性，同时也难以保留自然对话中至关重要的声学情境和细微差别，如语气、语调韵律和说话风格等。

Nova Sonic采用全新方法应对这些挑战。它摒弃了使用多个不同模型的方式，而是将理解与生成功能统一于单一模型中。这种整合使模型能根据语气、风格等声学情境以及口语输入调整所生成的语音响应，从而实现更自然的对话。Nova Sonic甚至能理解人类对话的细微之处，包括说话者的自然停顿与犹豫、能在恰当时机做出回应，并能从容应对对话中的插话情况。

基于Amazon Nova Sonic构建的旅游AI agent示例：

<https://soundcloud.com/amazonblog/amazon-nova-sonic-ai-agent-travel>

这段对话展示了客户与虚拟旅游助手讨论夏威夷行程的场景。当客户的语气从兴奋转为担忧费用时，AI的语调随之变得更具安抚性，并实时调取相关价格信息。

该模型还会为用户语音生成文字转录，让开发者能够利用这些文本来调用特定工具与API，从而构建语音AI agents。比如在这个示例中，AI旅游代理可以通过获取最新的航班信息来协助机票预订。这些功能，加上其超快的推理能力，使Nova Sonic支持的语音应用不仅更加自然，而且实用性更强。

基于Amazon Nova Sonic构建的企业AI助手示例：

<https://soundcloud.com/amazonblog/amazon-nova-sonic-ai-agent-travel>

此示例展示了企业客户如何从Nova Sonic基于公司数据生成响应的能力中受益。该助手能够提取报告并以自然对话的方式传递准确信息，同时主动提出相关的后续问题。这种流畅的对话模式支持多轮交互，无需说话者额外设定对话背景。

Nova

Sonic的推出彰显了亚马逊在先进基础模型领域的持续创新，亚马逊始终致力于为每一位亚马逊客户创造实际价值。

立即通过 [Amazon Nova](#)开始体验。欲了解更多信息，请访问[亚马逊在AI领域的创新](#)。

原文地址：<http://www.china-nengyuan.com/news/223897.html>