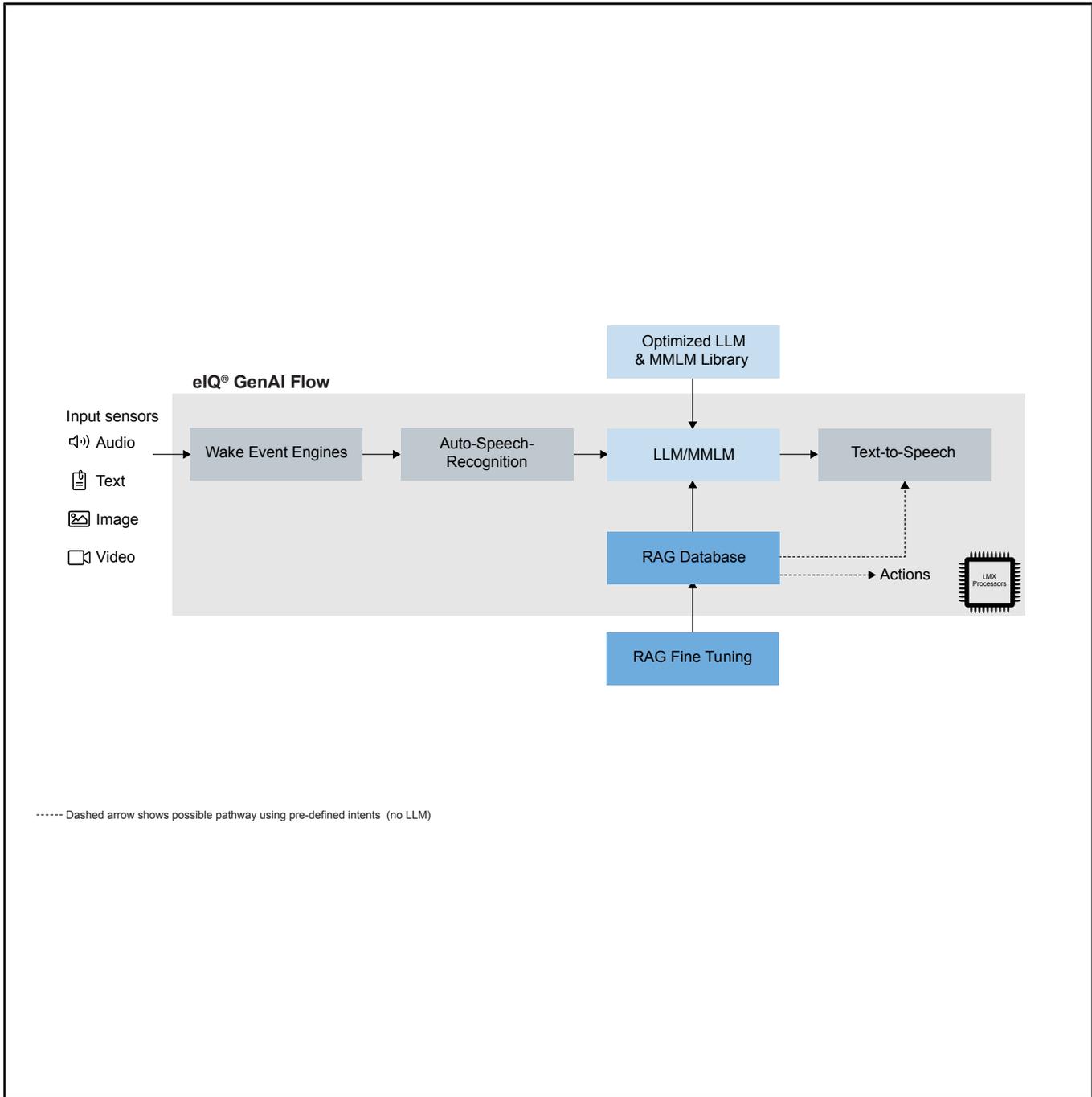# eIQ GenAI Flow助力简化和优化边缘生成式AI

GEN-AI-FLOW

Last Updated: Nov 25, 2025

eIQ Gen-AI Flow是一款模块化、端到端软件管道，经过优化，可将大语言模型和小语言模型(LLM和SLM)部署到恩智浦的应用处理器上，例如i.MX 95、93和8M Plus。该管道包含唤醒词、语音转文本、文本转语音和生成式AI模型，这些模型可在器件的CPU和/或NPU上运行。Gen-AI Flow还集成了检索增强生成(RAG)功能，可安全、便捷地利用PDF等定制私有知识源对大语言模型进行精细调优。

将生成式**AI**引入边缘 **Block Diagram**

eIQ® GenAI Flow

Input sensors
- ◁ᐤ Audio
- 📄 Text
- 🖼 Image
- 🎞 Video

Wake Event Engines → Auto-Speech-Recognition → LLM/MMLM → Text-to-Speech

Optimized LLM & MMLM Library

RAG Database

RAG Fine Tuning

Actions

i.MX Processors

------ Dashed arrow shows possible pathway using pre-defined intents (no LLM)

View additional information for eIQ GenAI Flow助力简化和优化边缘生成式AI.

www.nxp.com