

MULTIMODAL AI: INTEGRATING VISION AND VOICE IN AUTOMATION

The next frontier of Aura Visual Studio is Multimodal interaction. This means agents that can "see" and "hear."

VISION-DRIVEN INTAKE

Imagine a property management bot where a tenant uploads a photo of a broken pipe. The AI uses GPT-4 Vision to identify the issue, determine the urgency, and notify the correct plumber—all within a single chat window.

VOICE LOGIC

By integrating Twilio and ElevenLabs with our Voiceflow backends, we are architecting agents that don't just text—they speak. This is the future of automated discovery calls for the high-end editorial market.