



สถาบันไอทีจีเนียส
IT GENIUS INSTITUTE

COURSE ID
AIC-89

AI / VOICE AI

COURSE OUTLINE · รายละเอียดหลักสูตรอบรม

Voice AI Agents for Customer Service

สร้าง Voice Agent ภาษาไทยสำหรับคอลเซ็นเตอร์และงานบริการ

● 12 ชั่วโมง

● 2 วัน

● พื้นฐาน เขียน Python ได้

● ระดับ Intermediate



ระยะเวลา: 2 วัน

12 ชั่วโมง (2 วัน วันละ 6 ชม.) · 09:00-16:00 น.



รูปแบบ: Classroom / Hybrid

Onsite หรือ Online ผ่าน Microsoft Teams · Public / In-house



รายละเอียดหลักสูตร

www.itgenius.co.th

ลูกค้าจำนวนมากยังเลือกโทรหาธุรกิจเมื่อมีเรื่องด่วน แต่คอลเซ็นเตอร์มักได้รับสายไม่ทันในช่วงพีค และพนักงานต้องตอบคำถามเดิมซ้ำทั้งวัน Voice AI รุ่นใหม่คุยภาษาไทยได้เป็นธรรมชาติขึ้นมากและตอบได้แทบทันที แต่การสร้าง Voice Agent ที่ใช้กับลูกค้าจริงต่างจากแชทบอตข้อความ เพราะต้องคุมความหน่วง รู้ว่าเมื่อไรลูกค้าพูดจบ ให้ลูกค้าพูดแทรกได้ อ่านตัวเลขและชื่อเฉพาะภาษาไทยให้ถูก และรู้ว่าเมื่อไรต้องส่งต่อให้พนักงาน

หลักสูตรนี้พาผู้เรียนสร้าง Voice AI Agent ภาษาไทยสำหรับงานบริการลูกค้าด้วย Python เริ่มจากเลือกกระหว่าง Pipeline แบบ STT-LLM-TTS กับโมเดล Speech-to-Speech แบบ Real-time ทดลองตัวเลือก STT และ TTS ภาษาไทย สร้าง Agent ด้วย Gemini Live API และ OpenAI Realtime API ปรับ Turn-taking และการพูดแทรก จากนั้นเชื่อมข้อมูลธุรกิจด้วย Function Calling ออกแบบการส่งต่อพนักงาน พร้อมสรุปเคส เชื่อมเบอร์โทรผ่าน SIP วาง Guardrail และ PDPA แล้วทดสอบด้วยบทสนทนาหลายสถานการณ์ ปิดท้ายด้วย Capstone Voice Agent ของตนเอง หลักสูตรนี้เป็นสายนักพัฒนา ต่างจาก AI for Customer Service ที่เน้นการใช้ AI ในระดับธุรกิจ (อบรม 2 วัน วันละ 6 ชั่วโมง รวม 12 ชั่วโมง ระดับ Intermediate)

เครื่องมือ



Gemini Live



STT Speech-to-Text



11 ElevenLabs



Py Python



Ag AI Agents

ราคาค่าอบรม · PUBLIC / IN-HOUSE

7,900 บาท / ท่าน

ราคายังไม่รวมภาษีมูลค่าเพิ่ม (VAT 7%) · รวมเอกสารประกอบการอบรม Lab Guide โค้ด Python ตัวอย่างครบทุก Lab ชุดทดสอบบททดสอบ และ Certificate · ผู้เรียนเตรียม Notebook ชุดหูฟังพร้อมไมค์ และบัญชี API ของตนเอง · รับผิดชอบ In-house ขอใบเสนอราคาได้

สมัครผ่านเว็บไซต์ ลด 10%



เพิ่มเพื่อน
LINE

☎ 02-570-8449

☎ 088-807-9770

LINE @itgenius

🌐 www.itgenius.co.th

01 วัตถุประสงค์ Objectives

- เลือกสถาปัตยกรรม Voice Agent ที่เหมาะกับงานบริการและงบประมาณ
- เลือกและปรับ Speech-to-Text และ Text-to-Speech ภาษาไทยให้รับมือกับตัวเลขและชื่อเฉพาะได้
- สร้าง Voice Agent ภาษาไทยด้วย Gemini Live API และ OpenAI Realtime API
- ปรับ Turn Detection และการพูดแทรกให้บทสนทนาเป็นธรรมชาติ
- เชื่อม Voice Agent กับข้อมูลธุรกิจผ่าน Function Calling อย่างปลอดภัย
- ออกแบบการส่งต่อพนักงานพร้อมสรุปบริบทให้ลูกค้าไม่ต้องเล่าซ้ำ
- เข้าใจการเชื่อมต่อโทรผ่าน SIP และเสียงบนเว็บผ่าน WebRTC
- วาง Guardrail, PDPA และชุดทดสอบก่อนเปิดใช้กับลูกค้าจริง

02 กลุ่มเป้าหมาย Target Audience

- นักพัฒนา Python และ AI Engineer ที่ต้องสร้าง Voice Agent สำหรับงานบริการลูกค้า
- ทีม IT ของคอลเซ็นเตอร์และ Contact Center ที่ต้องการนำ Voice AI มาใช้
- นักพัฒนาที่เคยสร้างแชทบอทข้อความแล้วต้องการขยายสู่ช่องทางเสียง
- Solution Architect ที่ต้องเลือกแพลตฟอร์มและออกแบบการเชื่อมต่อระบบโทรศัพท์
- Product Owner สายเทคนิคที่ดูแลโครงการ Voice Bot หรือ Callbot

03 จุดเด่นของหลักสูตร Highlights

- **เน้นภาษาไทยจริง** STT, TTS ตัวเลข ชื่อเฉพาะ และไทยปนอังกฤษ
- **Real-time Voice API** Gemini Live API และ OpenAI Realtime API
- **คุยได้เป็นธรรมชาติ** Turn Detection, Barge-in และ Latency
- **ส่งต่อพนักงานได้** โอนสายพร้อมสรุปเคสให้คนรับต่อ
- **พร้อมใช้อย่างรับผิดชอบ** Guardrail, PDPA และการแจ้งว่าเป็น AI

04 ความรู้พื้นฐาน Prerequisites

- เขียน Python ได้ในระดับใช้งาน เช่น Function, async และติดตั้ง Package
- เข้าใจพื้นฐานการใช้งาน LLM และการเขียน Prompt
- ใช้ Command Line และเรียก REST API ขึ้นพื้นฐานได้
- โน้ตบุ๊กที่ติดตั้ง Python และ VS Code ได้ พร้อมชุดหูฟังที่มีไมโครโฟน

05 สิ่งที่ต้องเตรียมและเครื่องมือ Setup & Tools

รายการ	รายละเอียด
อุปกรณ์	Notebook เชื่อมต่ออินเทอร์เน็ต ชุดหูฟังพร้อมไมค์ ติดตั้ง Python และ VS Code ได้
Real-time Voice	Gemini Live API และ OpenAI Realtime API
STT และ TTS	บริการเสียงบนคลาวด์ ElevenLabs และโมเดลเปิด Typhoon ASR Real-time
เฟรมเวิร์ก	Python, Pipecat หรือ LiveKit Agents และ FastAPI
ข้อมูลตัวอย่าง	ระบบคำสั่งซื้อและนัดหมายจำลองพร้อม FAQ สำหรับทุก Lab

หมายเหตุ: ผู้เรียนใช้บัญชีของตนเองหรือของบริษัท Gemini API, OpenAI API, ElevenLabs และผู้ให้บริการ Telephony มีค่าใช้จ่ายตามการใช้งาน (บางบริการมี Free Tier หรือแพ็คเกจทดลอง) การเชื่อมต่อโทรจริงต้องใช้บริการ SIP Trunk หรือ Telephony ที่ผู้เรียนสมัครเอง ในห้องจิงสาริตและทดลองผ่านเบราว์เซอร์เป็นหลัก หลักสูตรนี้เป็นสายนักพัฒนา ต่อยอดจาก AI for Customer Service ที่เน้นการใช้ AI ในงานบริการระดับธุรกิจ ชื่อเครื่องมือและฟีเจอร์ของ AI เปลี่ยนเร็ว เนื้อหาจะอิงเวอร์ชันล่าสุด ณ วันอบรม

06 ภาพรวมการเรียนรู้ 2 วัน Big Picture

วัน	หัวข้อหลัก
Day 1 Section 1-6	ออกแบบสถาปัตยกรรม Voice Agent, เลือก Speech-to-Text และ Text-to-Speech ภาษาไทย, สร้าง Agent ด้วย Gemini Live API และ OpenAI Realtime API, คุม Turn-taking และการพูดแทรก และออกแบบบทสนทนาสำหรับเสียง
Day 2 Section 7-12	เชื่อมข้อมูลธุรกิจด้วย Function Calling, ส่งต่อพนักงานพร้อมสรุปเคส, เชื่อมเบอร์โทรผ่าน SIP และ WebRTC, ทำ Guardrail และ PDPA, ทดสอบและวัดผล แล้วปิดท้ายด้วย Capstone Voice Agent



DAY
1

เสียงภาษาไทยและ Voice Agent ตัวแรก

เนื้อหาการอบรม · Section 1-6

รวม
6 ชั่วโมง

1 ภาพรวม Voice AI ในงานบริการ

- งานที่ Voice Agent ทำได้ดี: ตอบคำถามซ้ำ เช็คสถานะ นัดหมาย และคัดกรองก่อนโอนสาย
- งานที่ควรให้พนักงานรับ เช่น เรื่องร้องเรียนอ่อนไหวและการตัดสินใจที่มีผลทางการเงิน
- Pipeline แบบ STT-LLM-TTS เทียบกับโมเดล Speech-to-Speech แบบ Real-time
- จังหวะเวลาของแต่ละขั้น และสิ่งที่ทำให้บทสนทนาฟังไม่เป็นธรรมชาติ
- Lab: เตรียม Python และ API Key แล้วทดลองคุยกับ Voice Agent ตัวอย่าง

2 Speech-to-Text ภาษาไทย

- ตัวเลือก STT ภาษาไทย: บริการบนคลาวด์ ElevenLabs และโมเดลเปิด Typhoon ASR Real-time
- Streaming เทียบกับ Batch และผลต่อความหน่วง
- ปัญหาที่พบบ่อย: ชื่อเฉพาะ รหัสสินค้า เบอร์โทร และคำอังกฤษปนไทย
- Lab: เทียบผลถอดเสียงจากไฟล์เสียงโทรจำลองหลายสำเนียง

3 Text-to-Speech ภาษาไทยและบุคลิกเสียง

- เลือก TTS จากความเป็นธรรมชาติ ความเร็ว ราคา และการรองรับ Streaming
- เตรียมข้อความให้อ่านตัวเลข วันที่ ราคา และคำย่อภาษาไทยถูกต้อง
- การโคลนเสียงต้องได้รับความยินยอมจากเจ้าของเสียง และเงื่อนไขการใช้เชิงพาณิชย์ของแต่ละแพลตฟอร์ม
- Lab: สร้างชุดข้อความตอบรับและทดสอบเสียงหลายแบบ

4 Lab: Real-time Voice API

- Gemini Live API: เชื่อมต่อผ่าน WebSocket ส่งและรับเสียงแบบต่อเนื่อง
- OpenAI Realtime API: Session, Instructions และการเลือกเสียง
- ตั้งค่าภาษาไทย ขอบเขตการตอบ และบุคลิกของ Agent
- Lab: สร้าง Voice Agent คุยภาษาไทยผ่านเบราว์เซอร์

5 Turn-taking และการพูดแทรก

- Voice Activity Detection และการตัดสินใจว่าลูกค้าพูดจบแล้ว
- Barge-in: ให้ลูกค้าพูดแทรกได้และหยุดเสียงของบอททันที
- รับมือความเงียบ เสียงรบกวน และลูกค้าที่พูดยาว
- คู่มือ Pipeline ด้วยเฟรมเวิร์กโอเพนซอร์สอย่าง Pipecat หรือ LiveKit Agents
- Lab: ปรับค่า Turn Detection แล้วทดสอบกับบทสนทนาในห้อง

6 Lab: ออกแบบบทสนทนาสำหรับเสียง

- เขียน Prompt สำหรับการพูด: ประโยคสั้น ถามทีละเรื่อง และไม่อ่านรายการยาว
- ยืนยันตัวตนและทวนข้อมูลก่อนดำเนินการ
- ขอบเขตที่บอทห้ามตอบ และวิธีปฏิเสธอย่างสุภาพ
- Lab: บทสนทนาเช็คสถานะคำสั่งซื้อพร้อมกรณียกเว้น



DAY
2

เชื่อมระบบ ส่งต่อพนักงาน และพร้อมใช้งานจริง

เนื้อหาการอบรม · Section 7-12

รวม
6 ชั่วโมง

7 Lab: Function Calling เชื่อมข้อมูลธุรกิจ

- ให้ Agent เรียก Tool เช็กคำสั่งซื้อ จองนัดหมาย และเปิด Ticket
- ตอบจากคู่มือและ FAQ ด้วย RAG แบบคำตอบสั้นที่เหมาะสมกับเสียง
- พูดย่อระหว่างรอผลจาก Tool เพื่อไม่ให้ลูกค้ารอเฉยๆ
- ตรวจสอบสิทธิ์ก่อนเปิดเผยข้อมูลของลูกค้า

8 Lab: ส่งต่อพนักงาน (Handover)

- เชื้อนไขส่งต่อ: ลูกค้าขอคุยกับคน อารมณ์ไม่ดี เรื่องเกินขอบเขต หรือบอกรอไม่มั่นใจ
- สรุปบริบทให้พนักงานรับต่อโดยลูกค้าไม่ต้องเล่าซ้ำ
- โอนสายไปยังเบอร์หรือ SIP และส่งเป็น Ticket หรือนัดโทรกลับเมื่อไม่มีคนว่าง
- Lab: สร้าง Tool ส่งต่อพนักงานพร้อมบทสรุปบทสนทนา

9 เชื่อมเบอร์โทรและช่องทางอื่น

- เชื่อม Voice Agent กับเบอร์โทรผ่าน SIP Trunk หรือผู้ให้บริการ Telephony
- Realtime API ผ่าน SIP และแพลตฟอร์มสำเร็จรูปอย่าง ElevenLabs Agents
- เสียงบนเว็บและแอปผ่าน WebRTC
- Lab: สาธิตการโทรเข้า Voice Agent ผ่านบริการ Telephony แบบทดลองใช้

10 Guardrail, PDPA และความปลอดภัย

- แจ้งลูกค้าว่ากำลังคุยกับ AI และขอความยินยอมก่อนบันทึกเสียง
- PDPA: เก็บเสียงและ Transcript เท่าที่จำเป็น ปิดบังข้อมูลส่วนบุคคล และกำหนดระยะเวลาเก็บ
- ป้องกัน Prompt Injection และการหลอกลวงข้อมูลของลูกค้าคนอื่น
- ไม่ให้บอกรหัสหรือข้อมูลสำคัญอื่นที่ไม่เกี่ยวข้อง

11 Lab: ทดสอบ วัดผล และปรับปรุง

- ทดสอบด้วยบทสนทนาจำลอง ทั้งลูกค้าโทร เสียงรวม และคำถามนอกเรื่อง
- ตัวชี้วัด: Latency อัตราแก้ปัญหได้เอง อัตราส่งต่อ และความถูกต้องของข้อมูล
- อ่าน Transcript และ Log เพื่อหาจุดที่ผิดพลาด
- Lab: ทดสอบ Agent ของตนแล้วปรับ Prompt และ Tool

12 Workshop: Capstone Voice Agent

- เลือกโจทย์: สอบถามสถานะ นัดหมาย หรือคัดกรองก่อนโอนสาย
- ออกแบบบทสนทนา Tool และเชื้อนไขส่งต่อพนักงาน
- สร้างและทดสอบแบบ End-to-End
- สาธิตสด พร้อมแผนนำร่องและตัวชี้วัด

07 เรียนจบแล้วได้อะไร Outcomes

- ออกแบบ Voice Agent ได้ Pipeline หรือ Speech-to-Speech
- เลือก STT และ TTS ไทยเป็น ตัวเลข ชื่อเฉพาะ และบุคลิกเสียง
- สร้าง Agent แบบ Real-time Gemini Live API และ OpenAI Realtime
- คุยได้เป็นธรรมชาติ Turn-taking และ Barge-in
- ส่งต่อพนักงานได้ลื่น สรุปเคสและโอนสาย
- พร้อมใช้อย่างปลอดภัย Guardrail, PDPA และชุดทดสอบ

08 สิ่งที่คุณเรียนจะได้รับ Deliverables

- เอกสารประกอบการอบรมฉบับสมบูรณ์ครบทั้ง 2 วัน และ Lab Guide ที่ละเอียด
- โค้ด Python ตัวอย่างครบทุก Lab ทั้ง Real-time Voice, Function Calling และ Handover
- ชุดบทสนทนาทดสอบหลายสถานการณ์และ Checklist ก่อนเปิดใช้ Voice Agent
- ประกาศนียบัตรจบหลักสูตร (Certificate of Completion) และช่องทางสอบถามวิทยากรหลังจบหลักสูตร

09 ข้อมูลผู้สอน Instructor

อ.สามิตร ไกยม ผู้สอนและนักพัฒนาที่สร้างระบบ AI สำหรับงานบริการลูกค้าทั้งช่องทางแชตและเสียง สอนผ่านการลงมือทำและฟังผลจริงในห้อง ทั้งเสียงภาษาไทย ความหน่วง และจังหวะการพูด เน้นให้ผู้เรียนออกแบบ Agent ที่รู้ขอบเขตของตัวเองและส่งต่อพนักงานได้ถูกจังหวะ เพื่อให้นำไปทดลองใช้กับลูกค้าจริงได้อย่างมั่นใจและรับผิดชอบ

ราคาค่าอบรม 7,900 บาท/ท่าน · สนใจ IN-HOUSE ขอใบเสนอราคา

7,900 บาท / ท่าน (ยังไม่รวม VAT 7%)

สมัครผ่านเว็บไซต์ ลด 10%

☎ 02-570-8449

☎ 088-807-9770

LINE @itgenius

✉ contact@itgenius.co.th



เพิ่มเพื่อน
LINE