

Building AI Chatbot with RAG

TAUTOLOGY - 27/07/2567



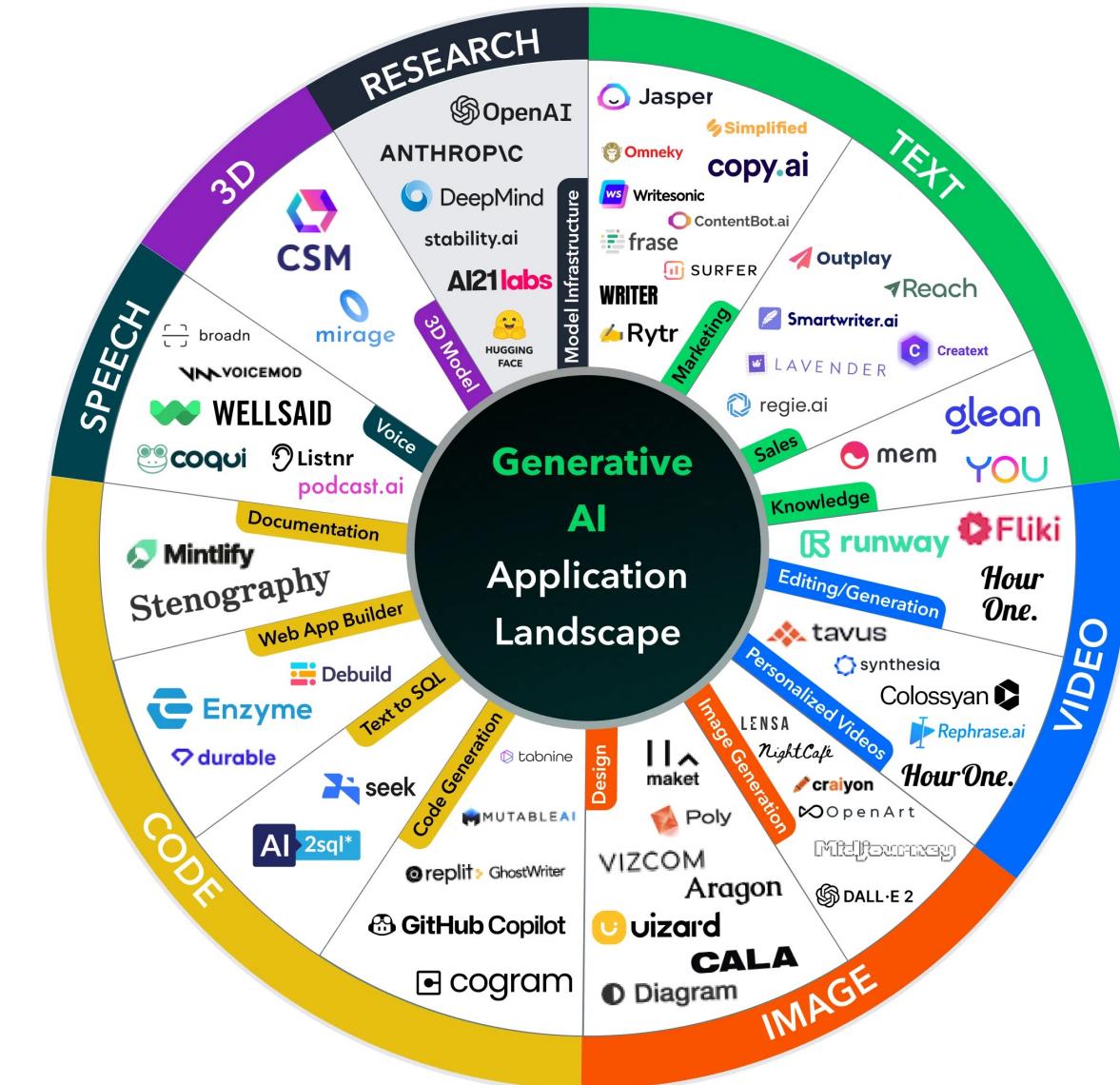
KRIN CHINPRASATSAK

DATASCIENTIST
INSTRUCTOR
SPEAKER

Experienced data scientist, instructor and speaker,
AI consultant for the public and private sectors,
CEO & Co-founder of MADEBYAI and QuantMetric



AI business opportunities



ref: <https://www.rapidops.com/blog/generative-ai-tools/>

ref: <https://medium.com/@stanfordgseit/a-new-class-of-ai-tools-9055131bfc53>



The Artificial Creativity Landscape

Linguistic Creativity

ALL-Purpose Text Generation



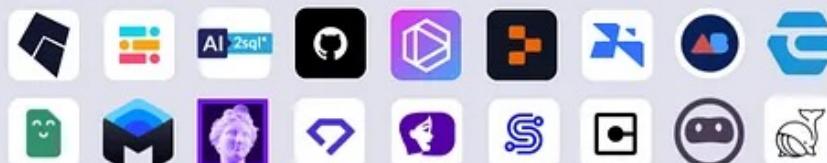
Marketing & Sales Copy



Text Optimization



Coding & Web Development



Musical Creativity

Sound & Music Generation

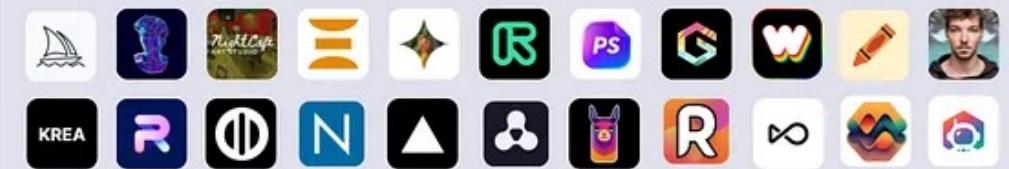


Voice Synthesis



Visual & Artistic Creativity

Image Generation



Video Generation



Design

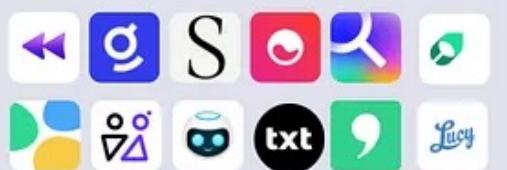


Scientific Creativity



Other

KnowLedge Management & Search



Customer Support & Automation



Characters

Gaming



Erotica



Market Intelligence



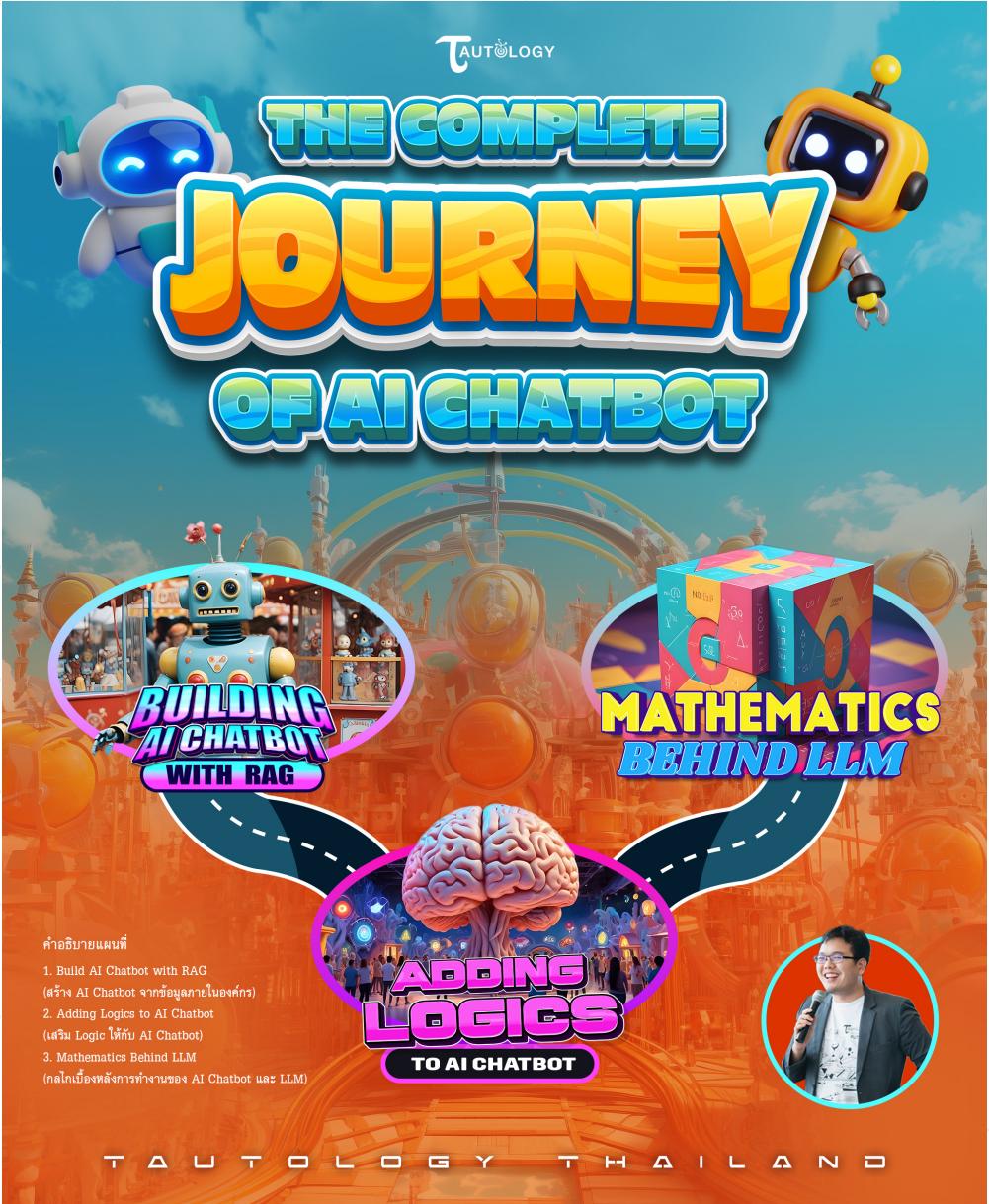
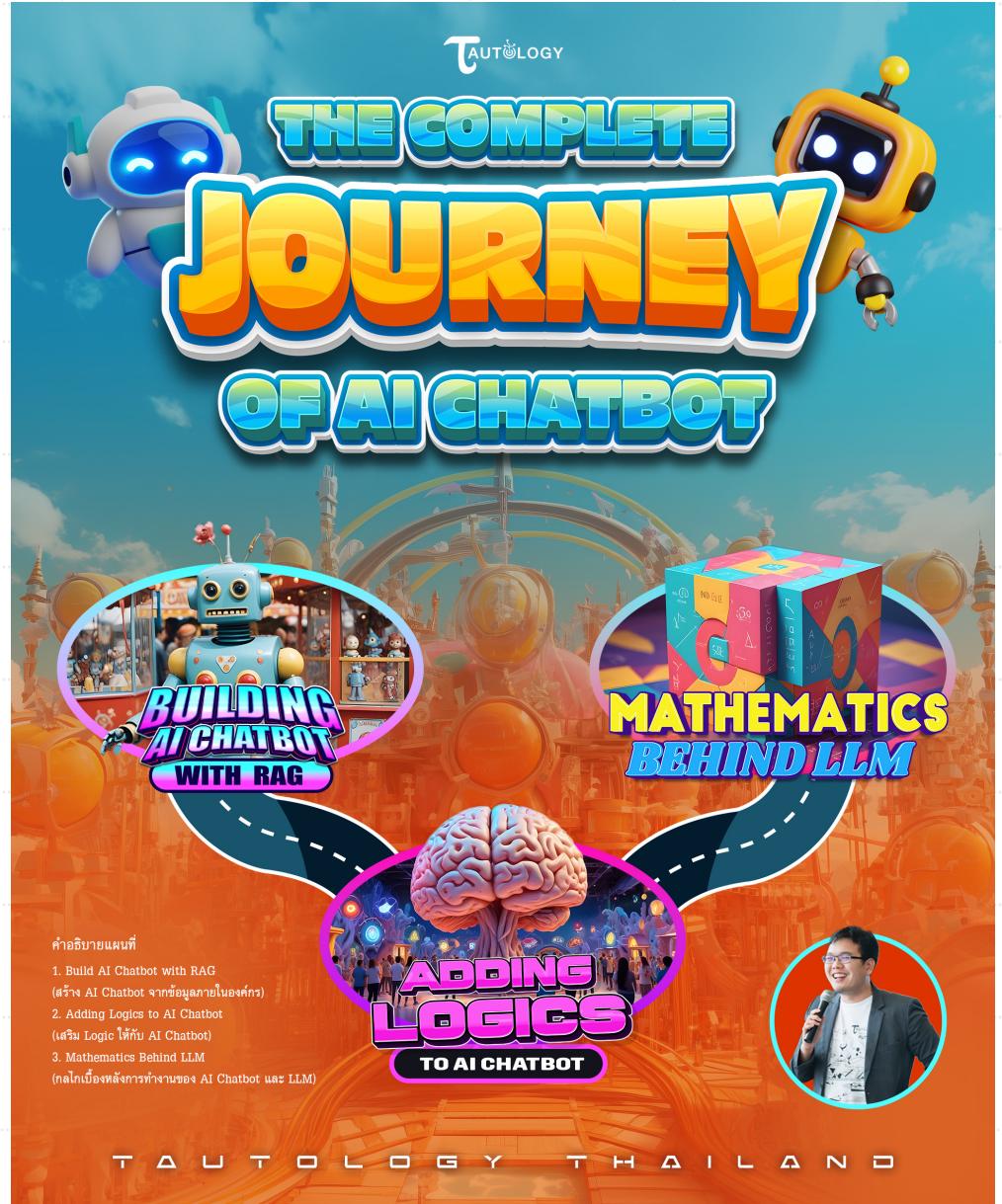


“มี startup หลายรายรับกระแสโดดเข้ามากว่าโอกาสในกระแส AI เร็วเกินไป

สุดท้ายเจ้าใหญ่ทำการปรับปูรุ่งช่องให้วิเพียงเล็กน้อย AI startup เหล่านั้นก็ปิดตัวลงทันที”



We're developing new AI product





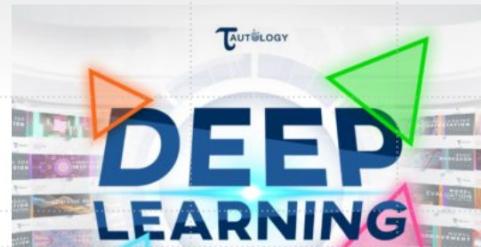
Line Group



THE BEST PLACE FOR
INNOVATION



POPULAR COURSES



REGISTER

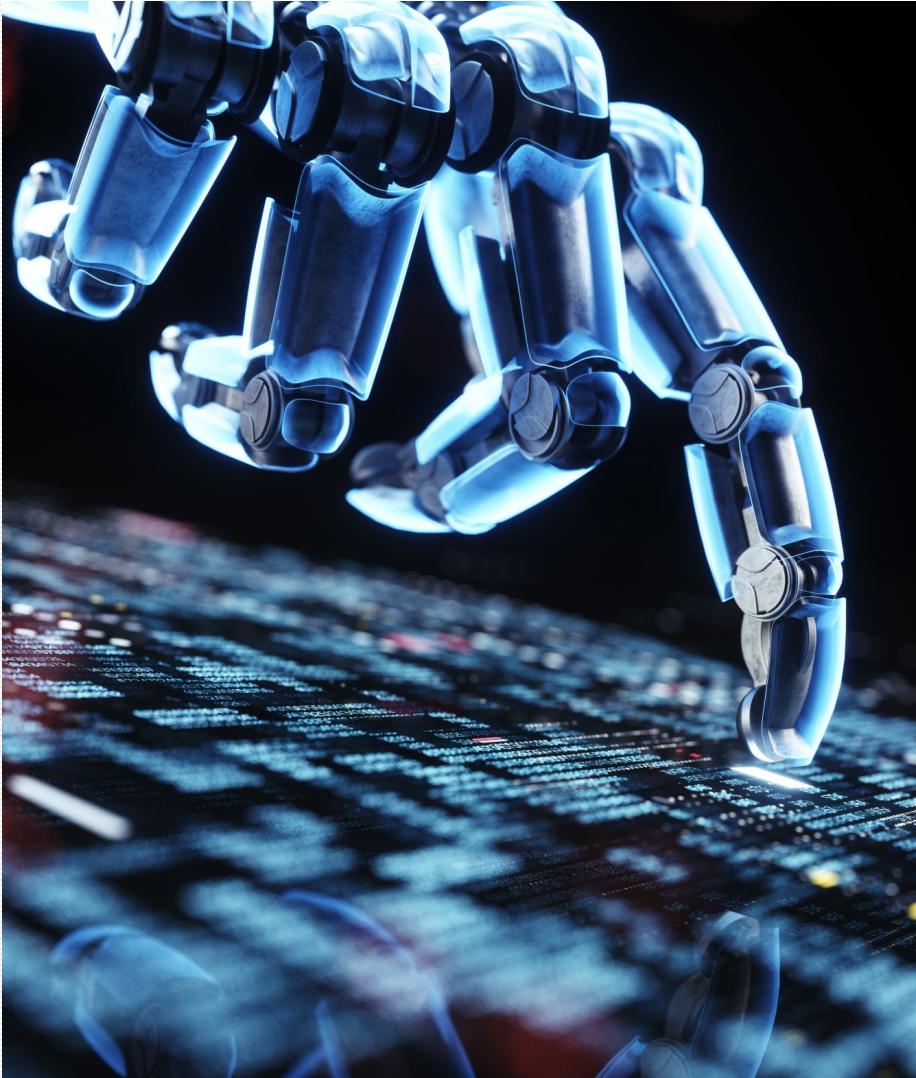
Home • Register

Username *

E-mail Address *

Password *

<https://madebyai.io/register/>



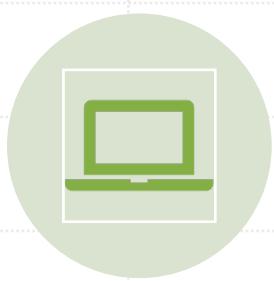
Agenda

- Setting up AI chatbot
- Chatbot without RAG
- Chatbot with RAG (pdf)
- Chatbot with RAG (json)
- General chatbot with RAG
- Connect with Line
- Connection with FB messenger
- Deployment

Setting up AI chatbot



Setting up chatbot



GOOGLE
COLAB



API KEYS



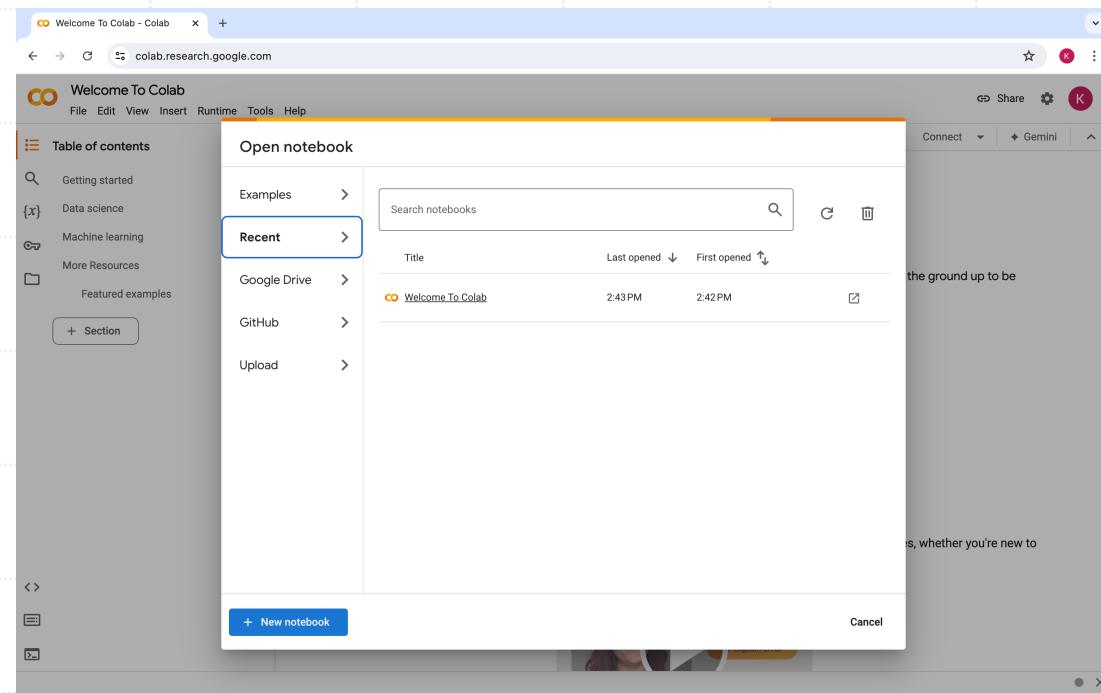
API PRICING



WHY NOT
LLAMA?

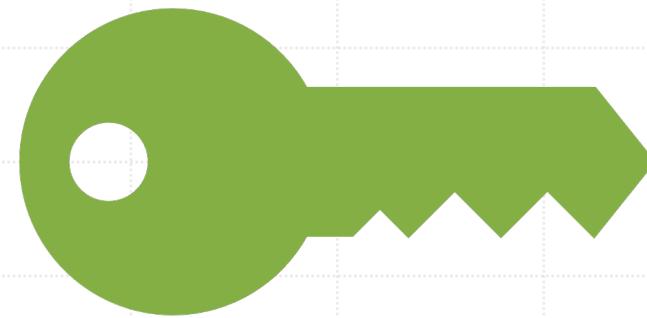
Google Colab

<https://colab.research.google.com/>



API keys

way_to_get_API_keys.pdf



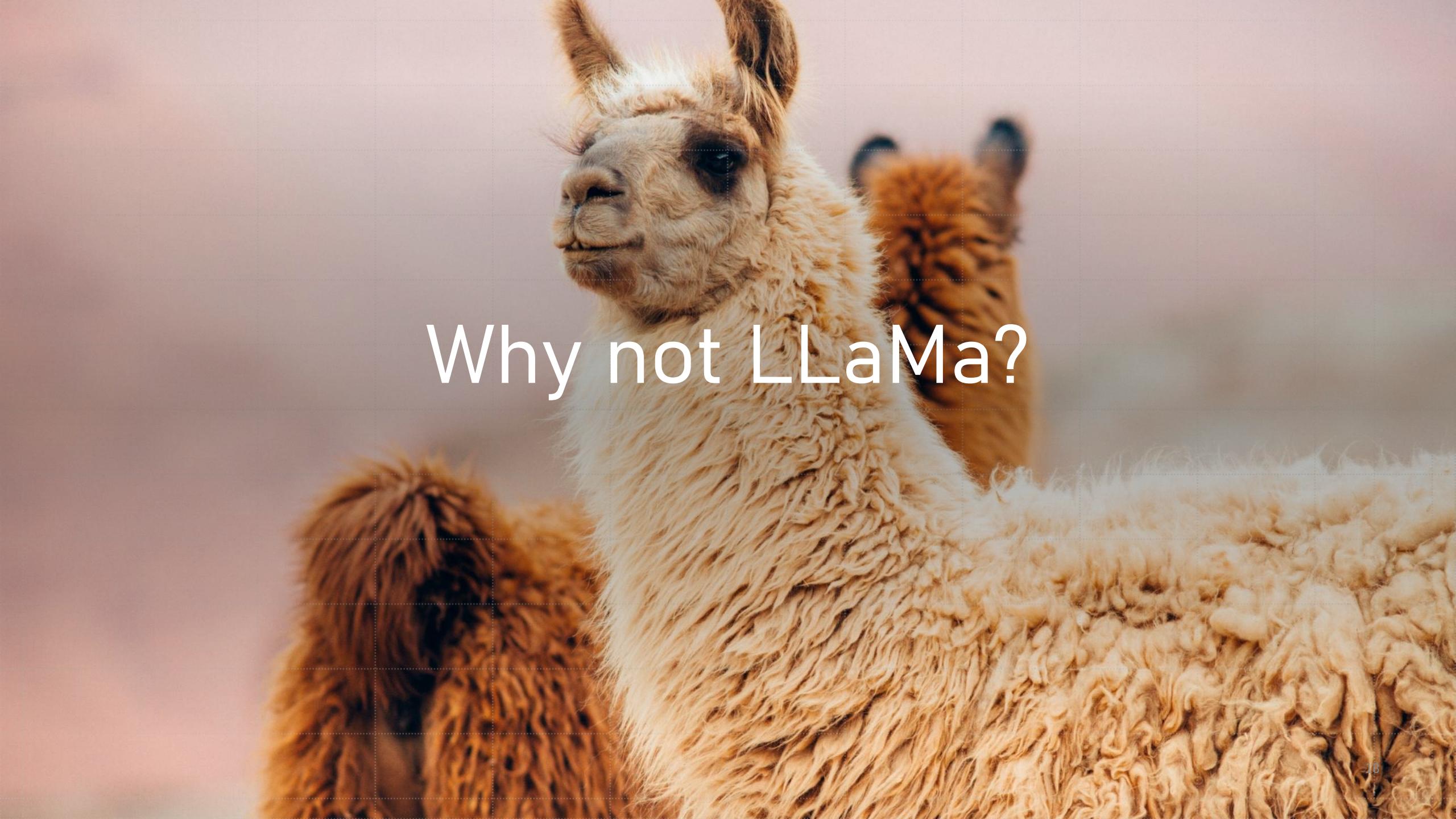


Pricing

- <https://openai.com/api/pricing/>
- <https://ai.google.dev/pricing>
- <https://www.anthropic.com/pricing#anthropic-api>



SMALL LANGUAGE MODEL
VS
LARGE LANGUAGE MODEL

A close-up photograph of two llamas. One llama is in sharp focus in the foreground, facing left, showing its profile and thick, light-colored, wavy fur. Behind it, another llama's head and neck are visible, slightly blurred, showing darker, more textured brown fur. The background is a soft, warm-toned gradient.

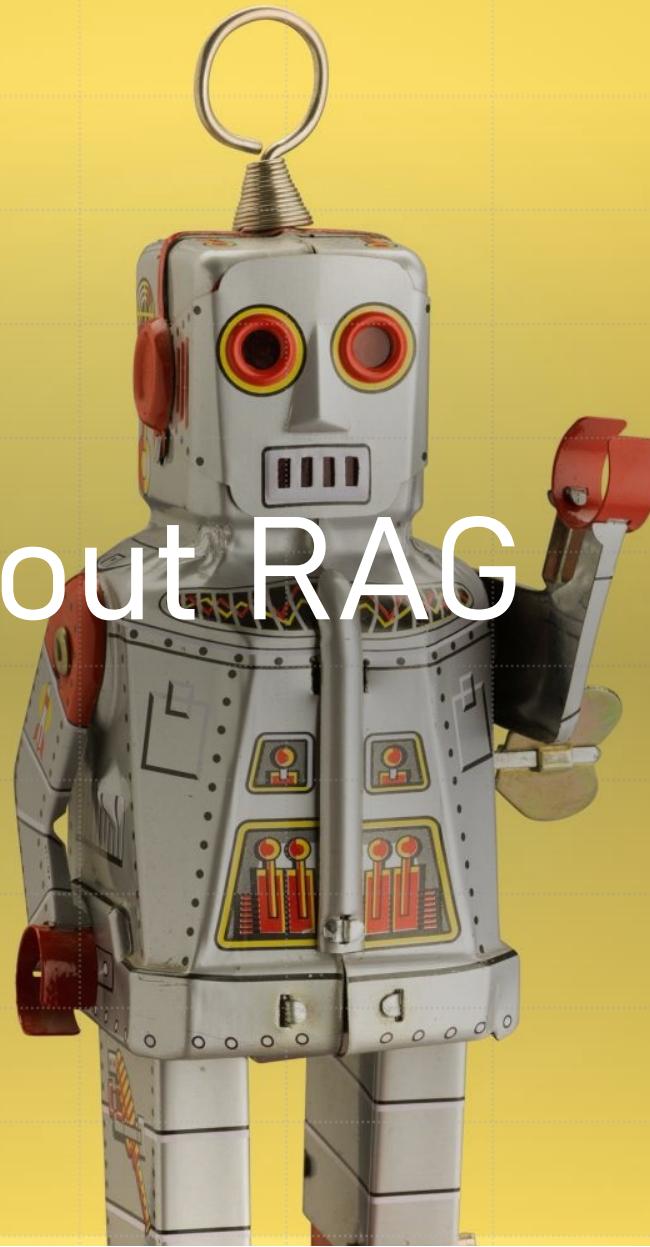
Why not LLaMa?

LLaMa3.1

405B requires 8xH100
8xH100 = \$64/hour
= \$1536/day
= \$46080/month
≈ 1.7 ล้านบาท/เดือน



Chatbot without RAG



Chatbot without RAG

- ChatGPT_without_RAG.ipynb
- Claude without RAG.ipynb
- Gemini without RAG.ipynb



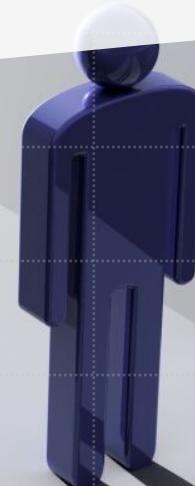
Chatbot with RAG (pdf)



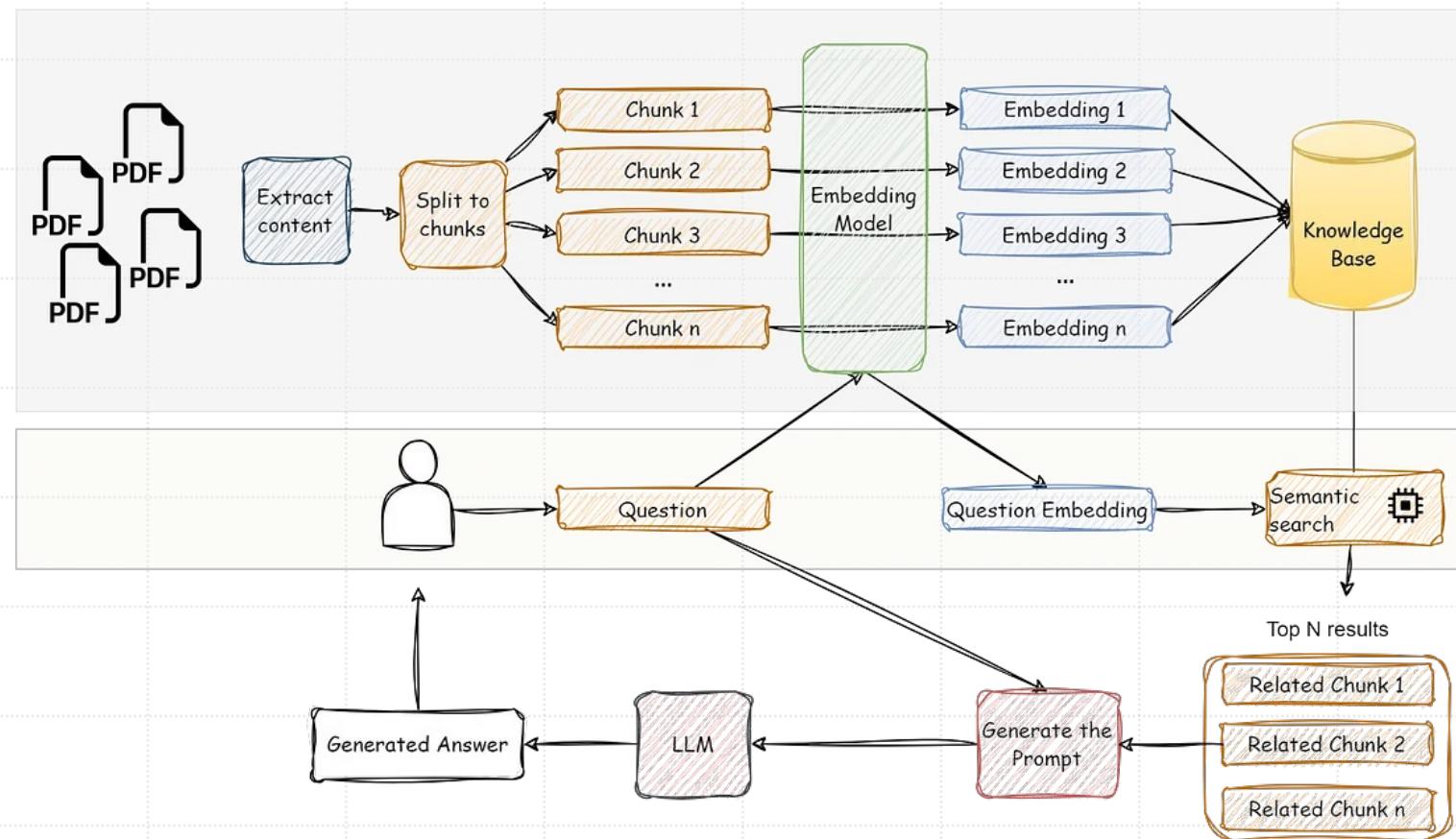
What is RAG?

RAG stands for Retrieval-Augmented Generation

Vector search → Semantic understanding

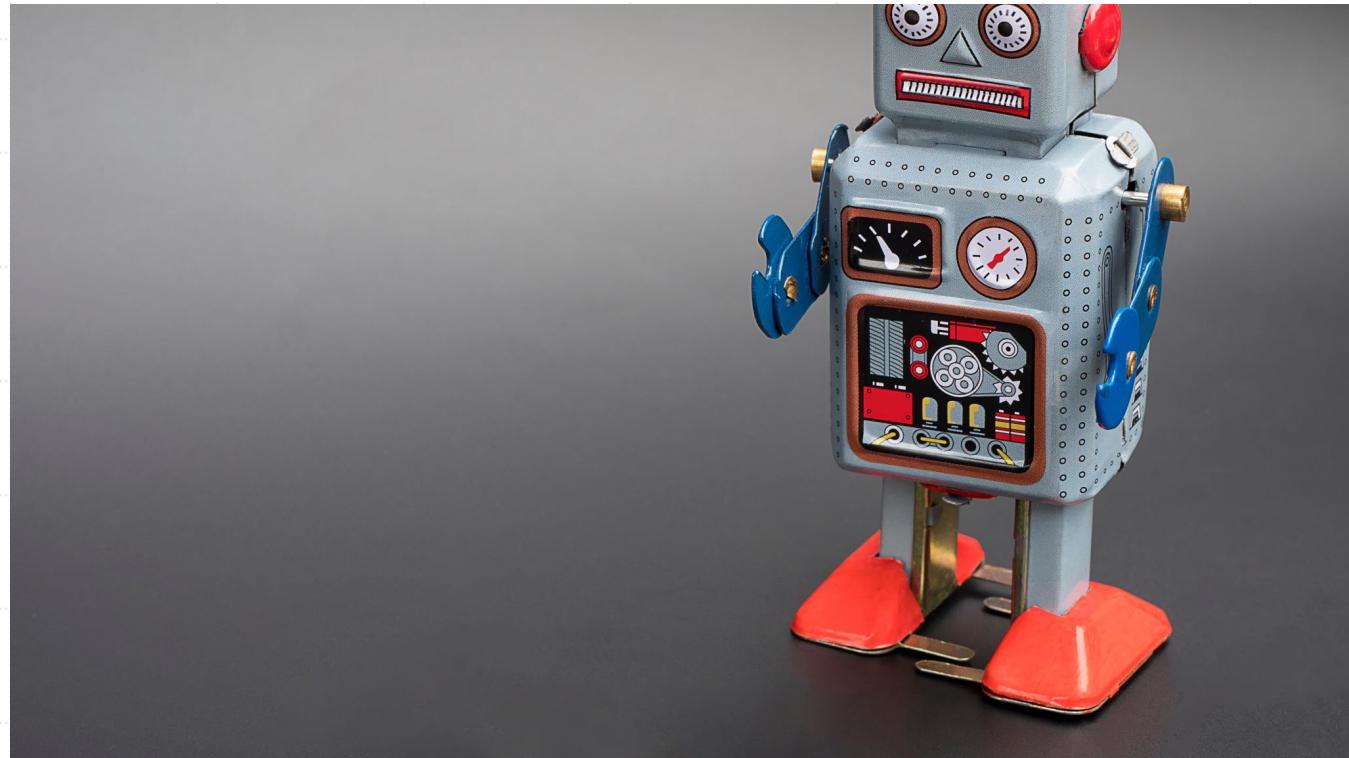


Chatbot with RAG (pdf)

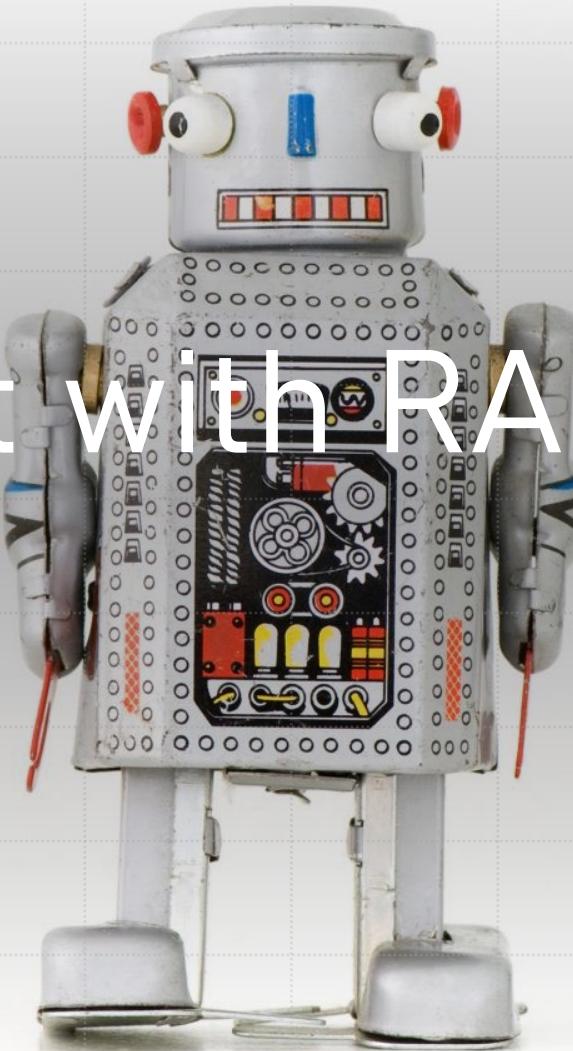


Chatbot with RAG (pdf)

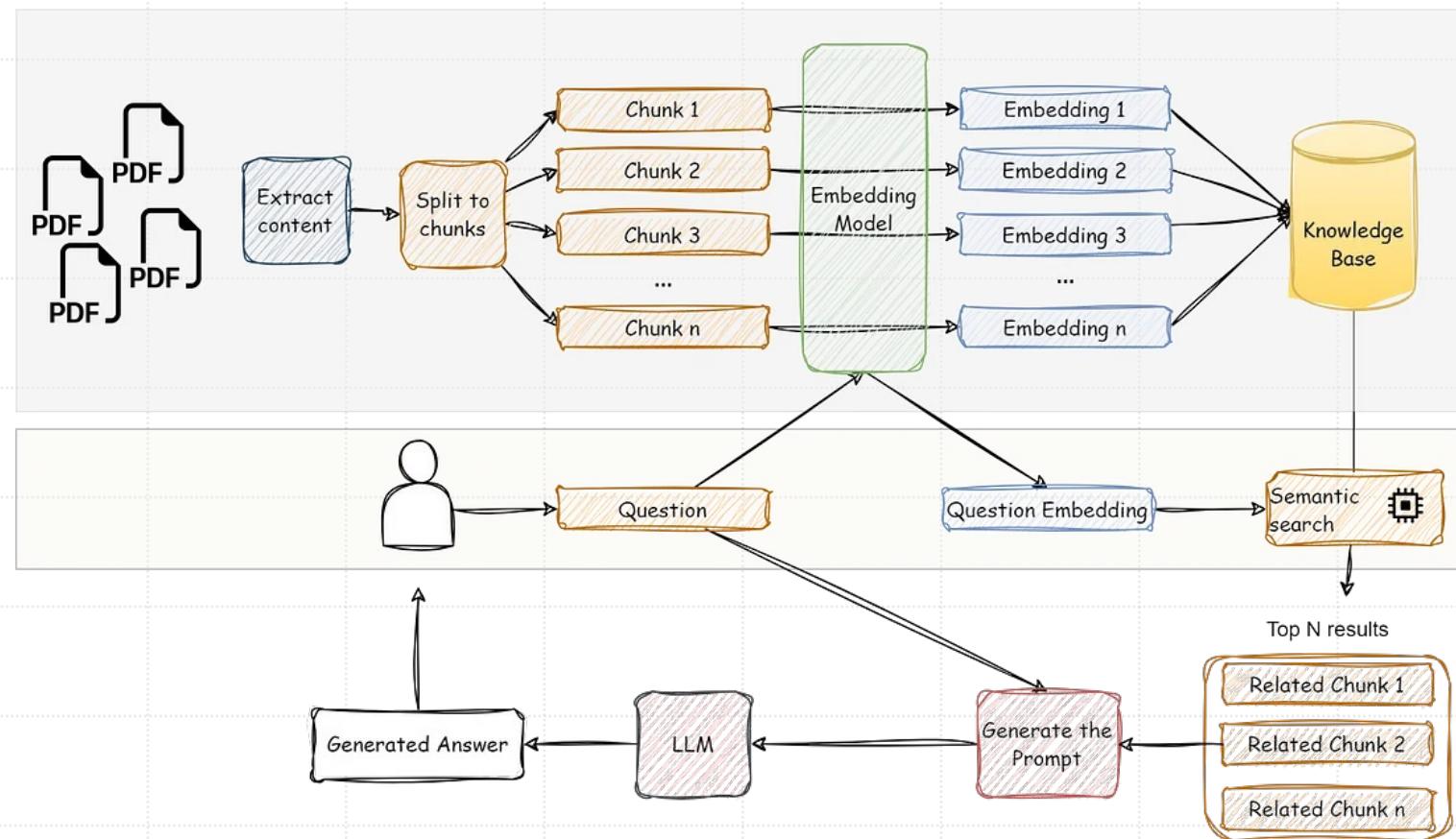
- ChatGPT_with_RAG_(pdf).ipynb
- Claude_with_RAG_(pdf).ipynb
- Gemini_without_RAG_(pdf).ipynb



Chatbot with RAG (json)

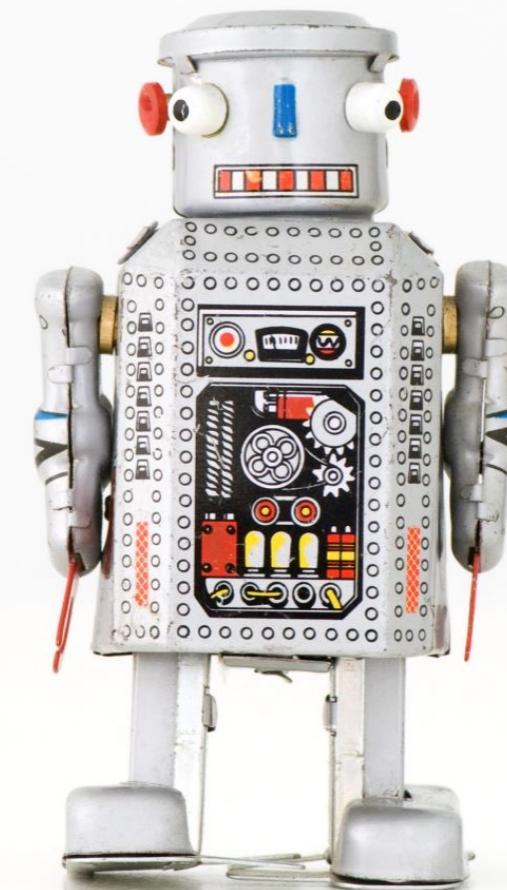


Chatbot with RAG (pdf)



Chatbot with RAG (json)

- ChatGPT_with_RAG_(json).ipynb
- Claude_with_RAG_(json).ipynb
- Gemini_without_RAG_(json).ipynb

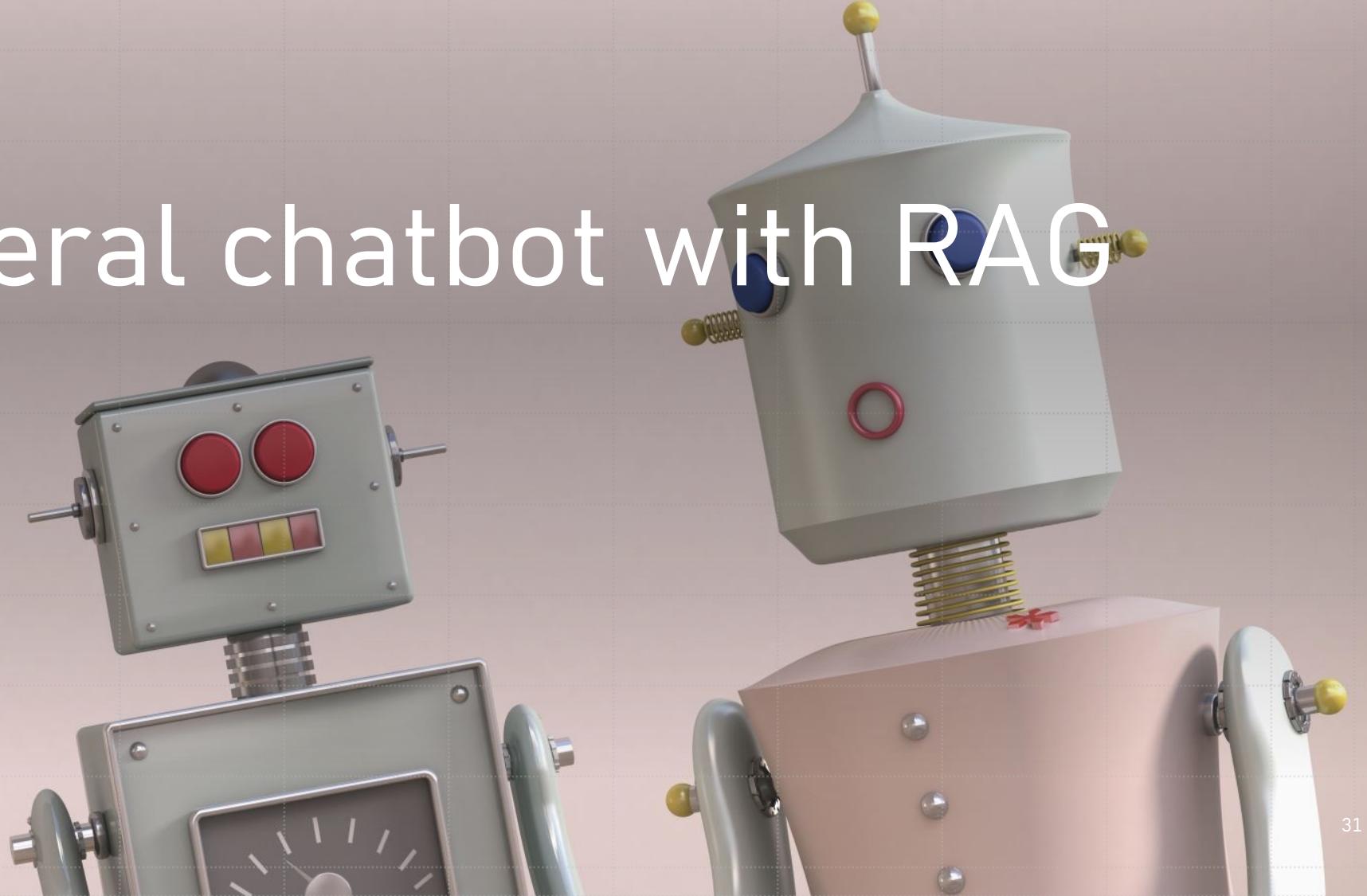


A blue pen with a silver clip and a translucent blue barrel lies diagonally across a white sheet of paper. The paper features several horizontal blue bars of varying lengths, suggesting a bar chart or a grid pattern. The background is a light grey.

What if company data?

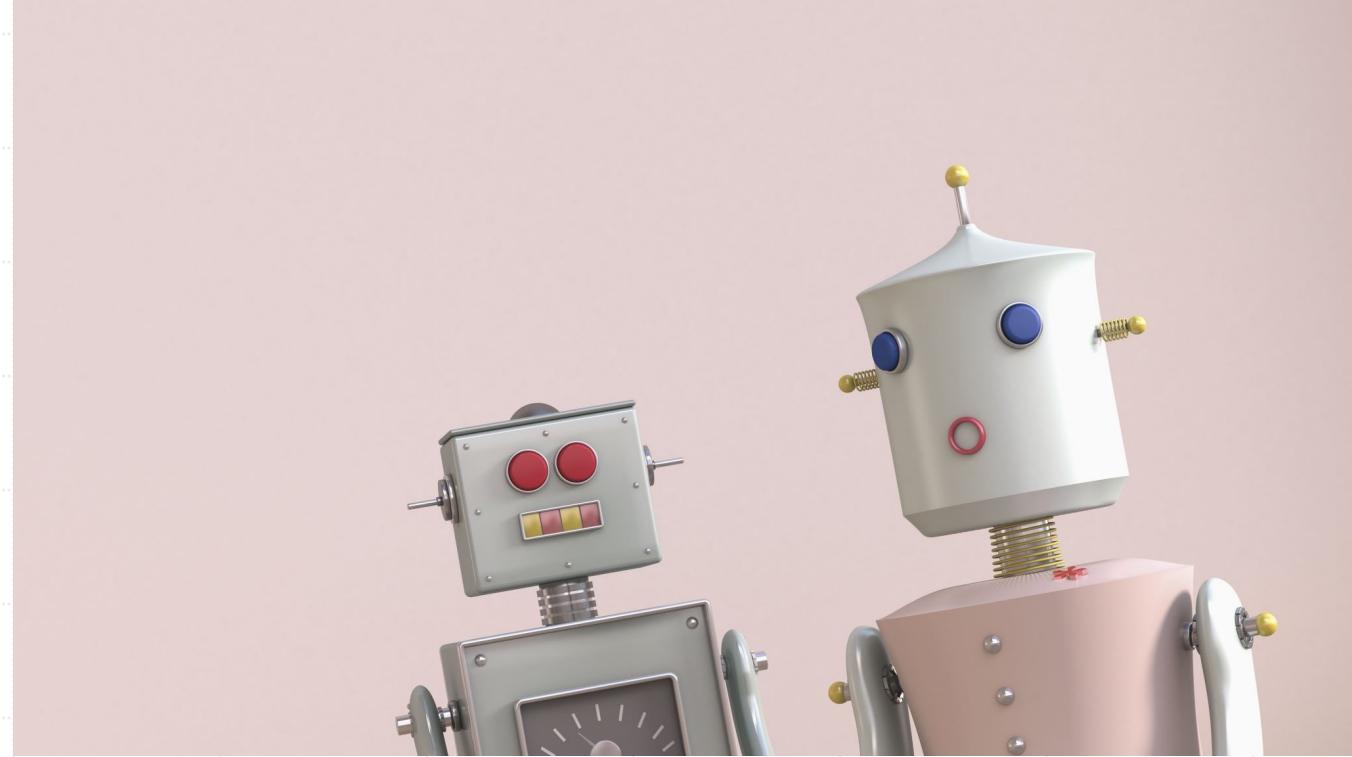
Export it as json

General chatbot with RAG



General chatbot with RAG

- General_ChatGPT_with_RAG.ipynb
- General_Claude_with_RAG.ipynb
- General_Gemini_with_RAG.ipynb





Connect with Line

Line_Console.pdf



Connect with Line



★ Tabby_Channel
เพื่อน 4



แมว



ไลน์

Tabby_CH
@892xdsfb

© LY Corporation
รายงานปีงบประมาณ ดูบัญชีทางการอื่นๆ



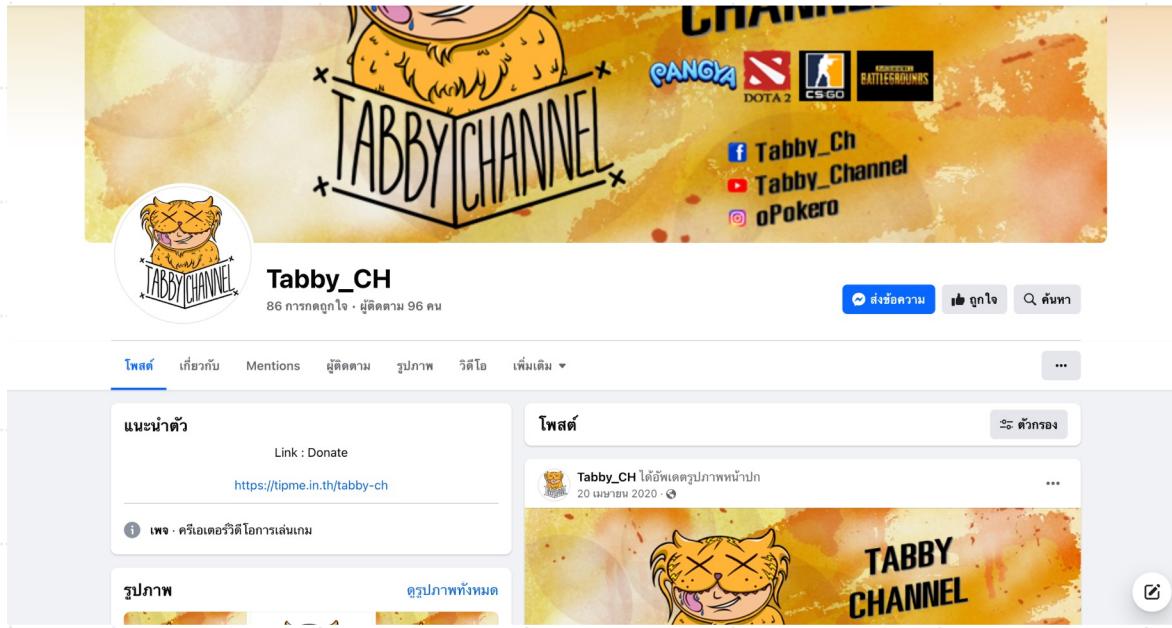
ประเทศไทย

Connect with FB messenger

FB_Messenger_Console.pdf

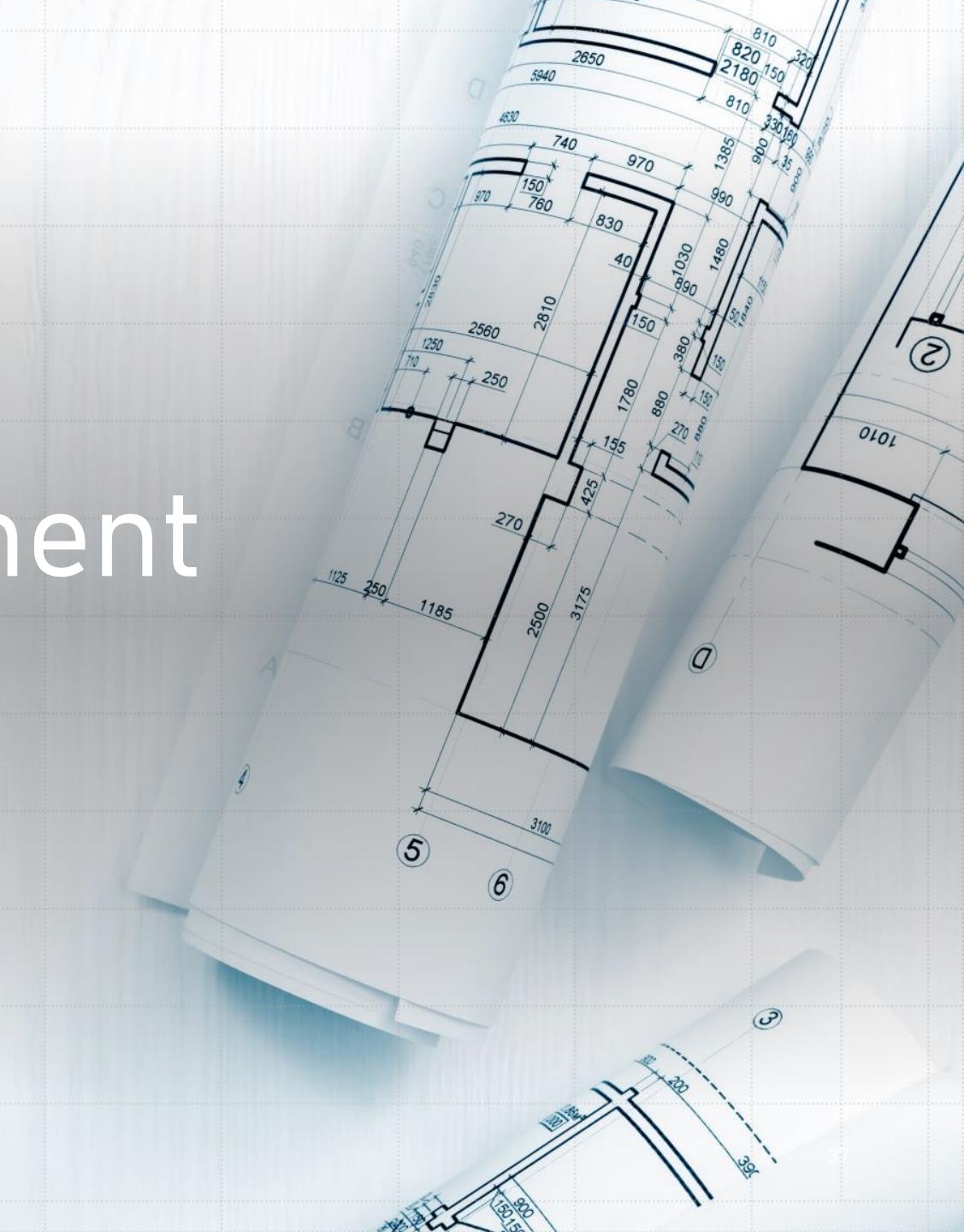


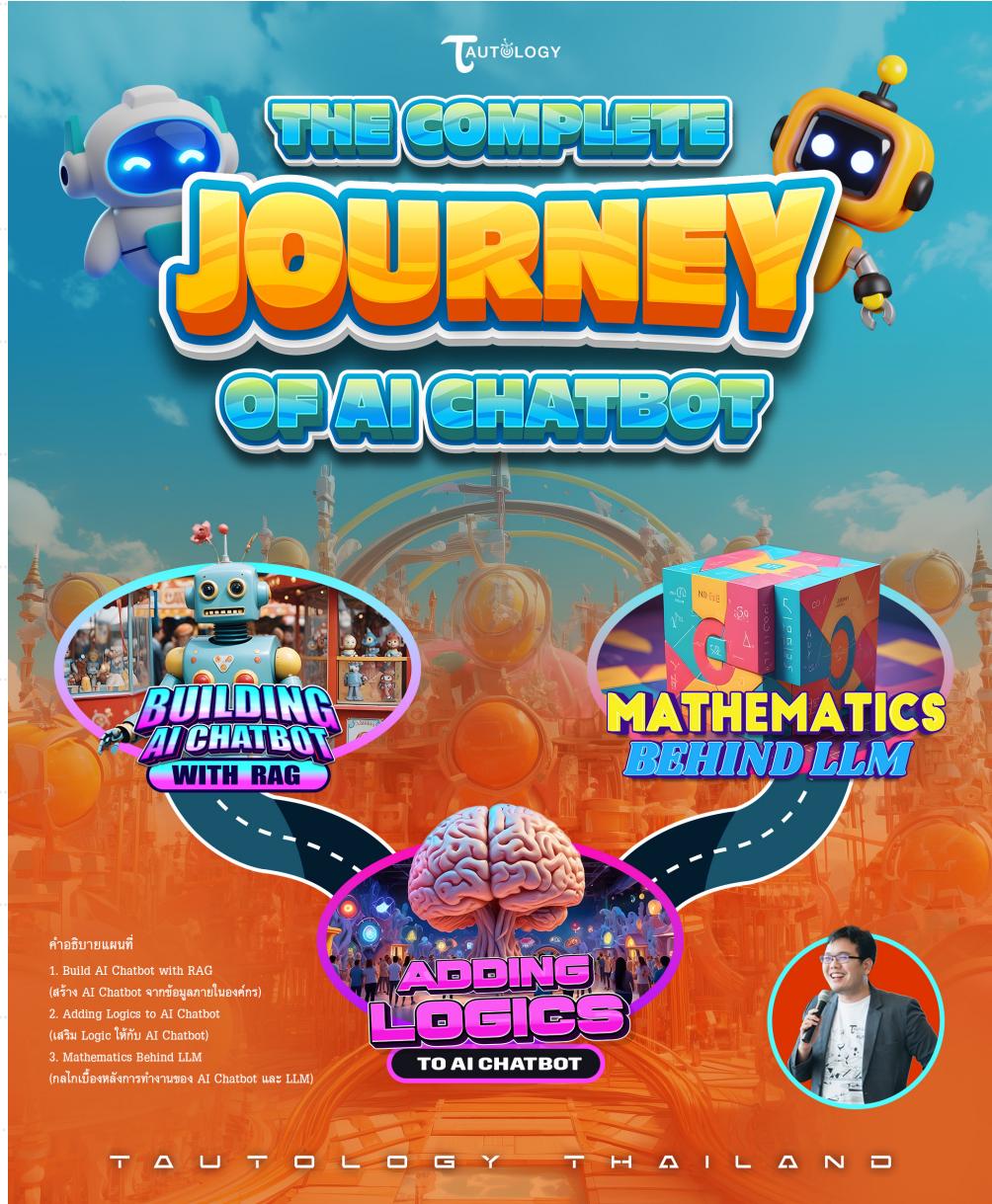
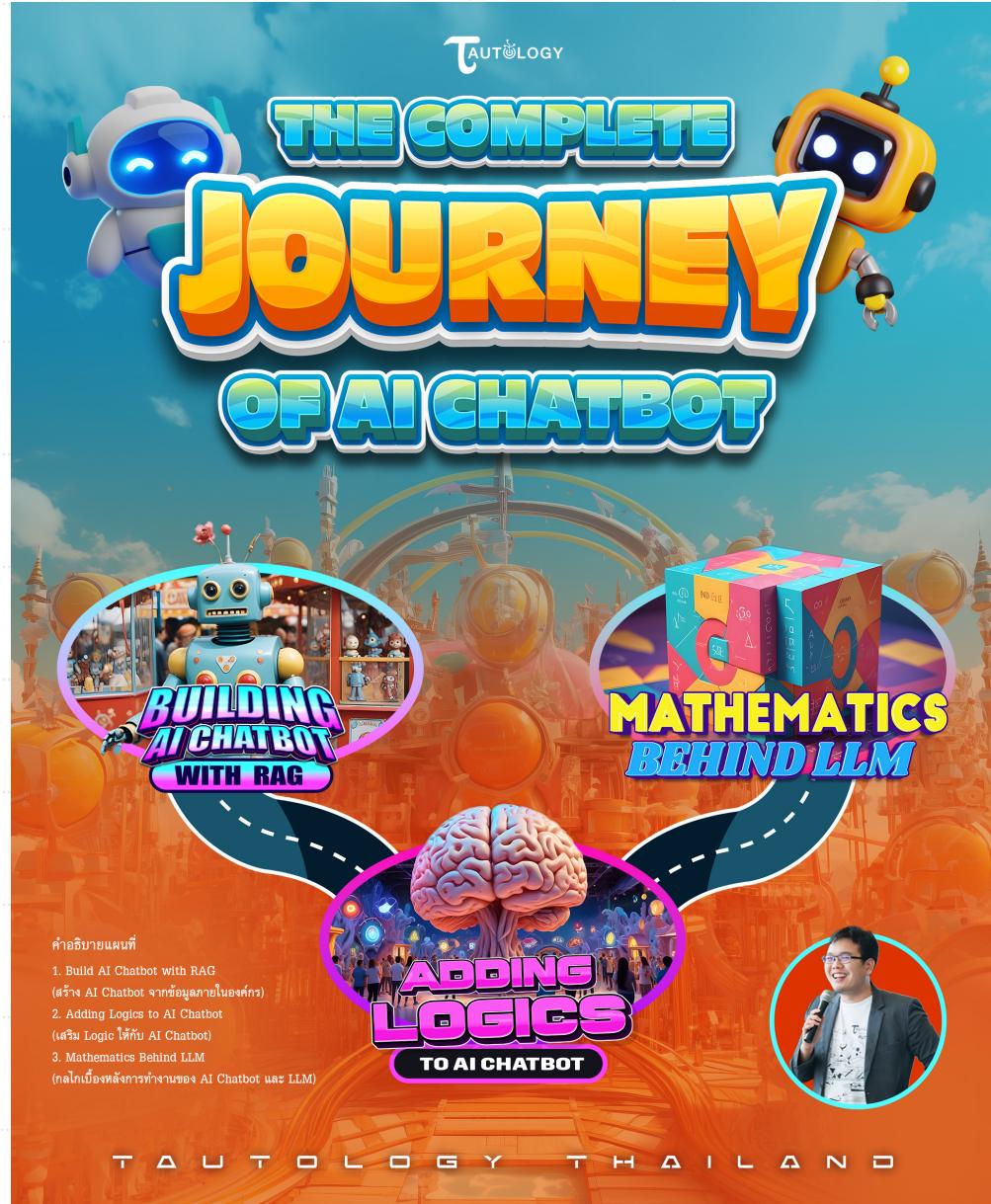
Connect with FB messenger



<https://www.facebook.com/TabbyChannel>

Deployment





THANK YOU

