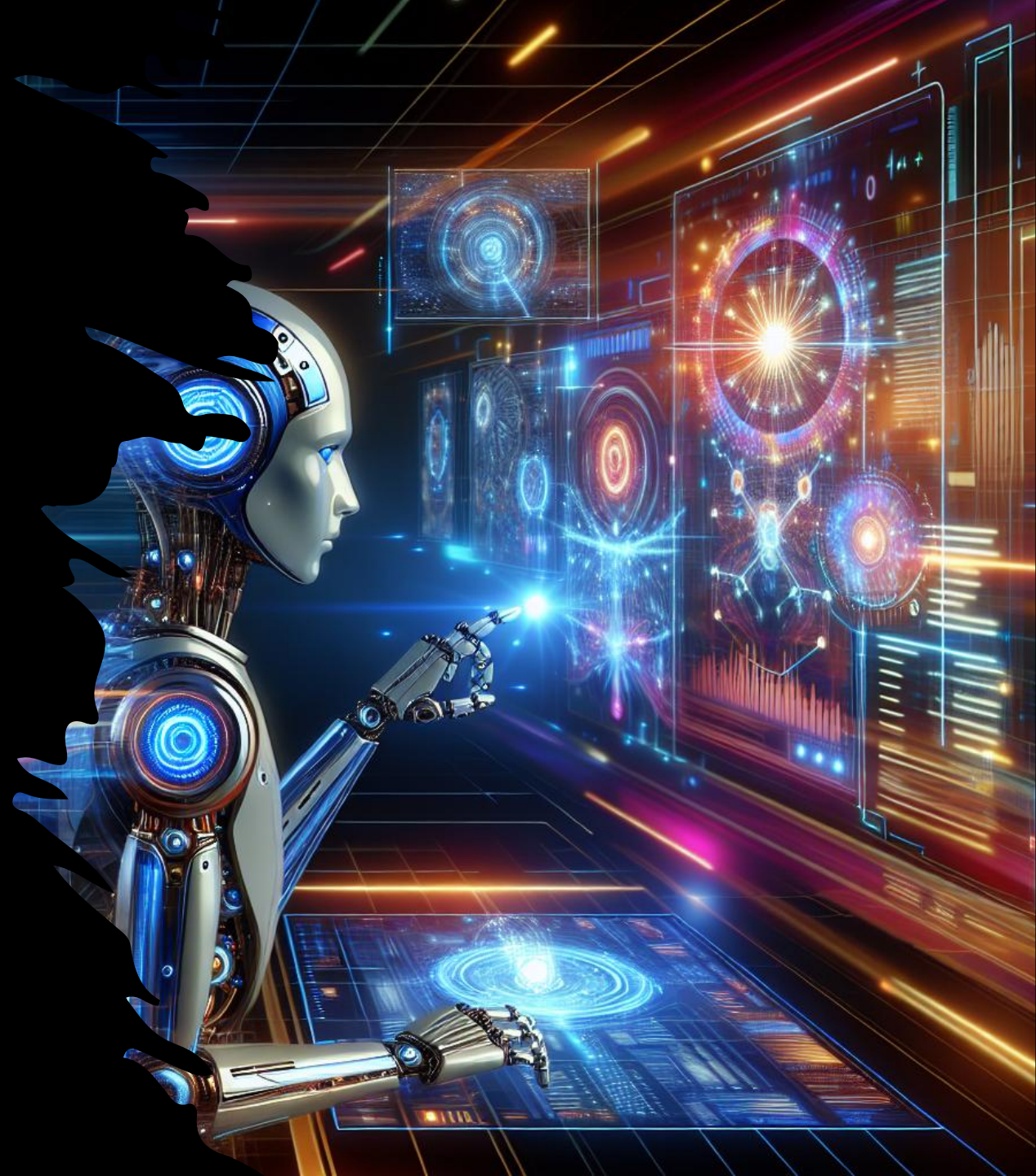


Thai LLM

Potentials for Research and Innovation

Chai Wutiwiwatchai

NECTEC¹
a member of NSTDA



Revolution of Thai LLM

2005 – The Cloud

2010 – IBM Watson

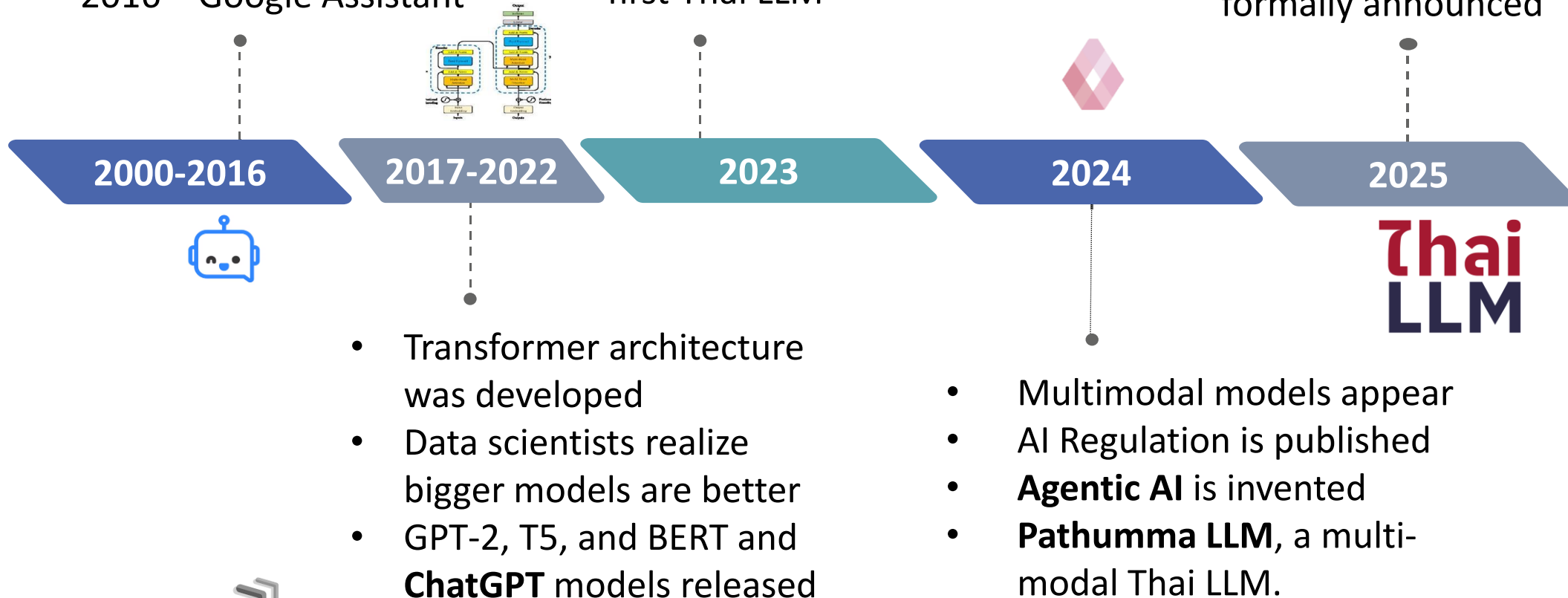
2011 – Siri

2014 – Cortana & Alexa

2016 – Google Assistant

- **RAG** (Retrieval Augmented Generation) is invented
- **OpenThaiLLM** launched first Thai LLM

- Reasoning and Voice models
- **Agentic AI** is the main key trend, with lots of innovation
- **ThaiLLM** development is formally announced

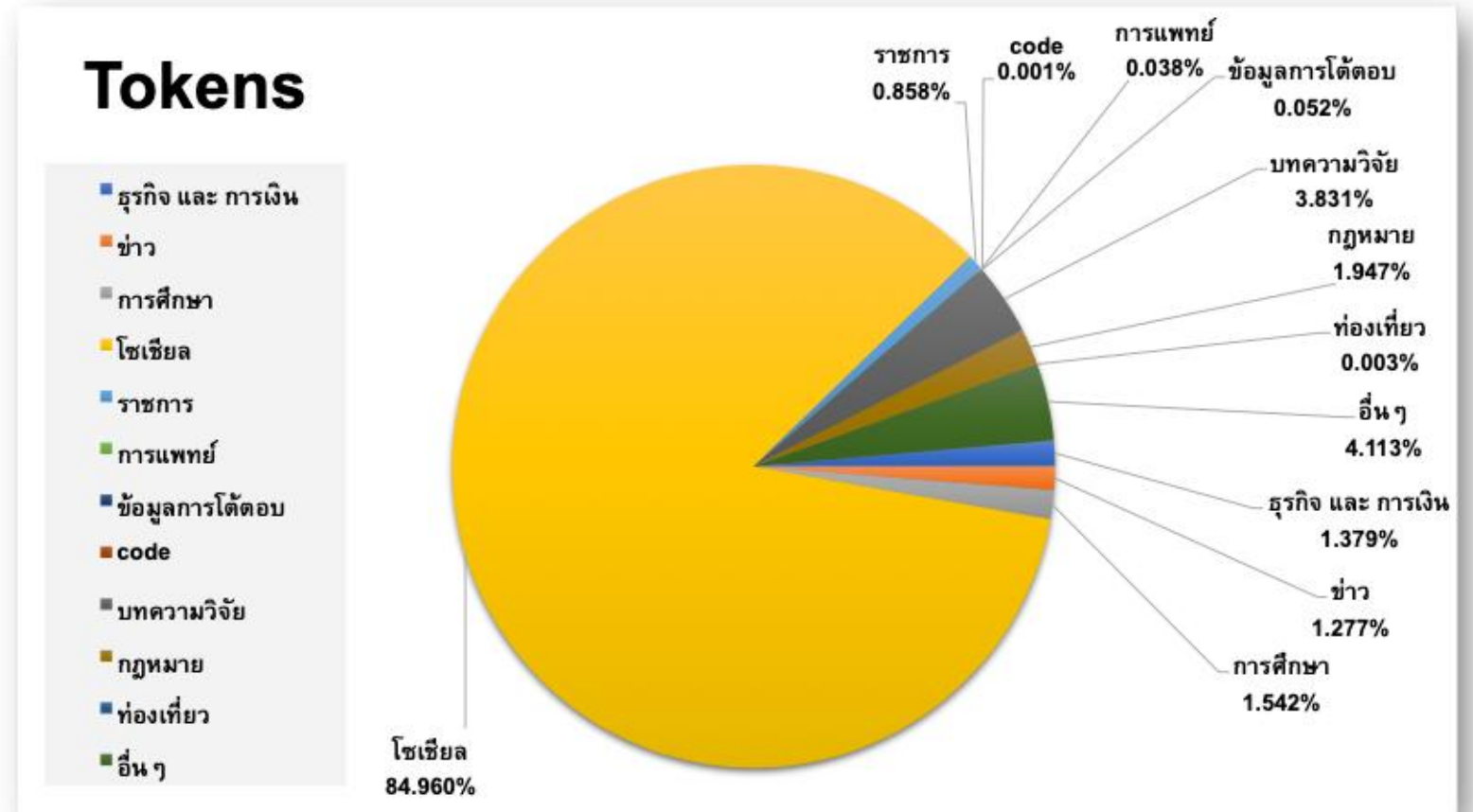


Thai LLM Foundation



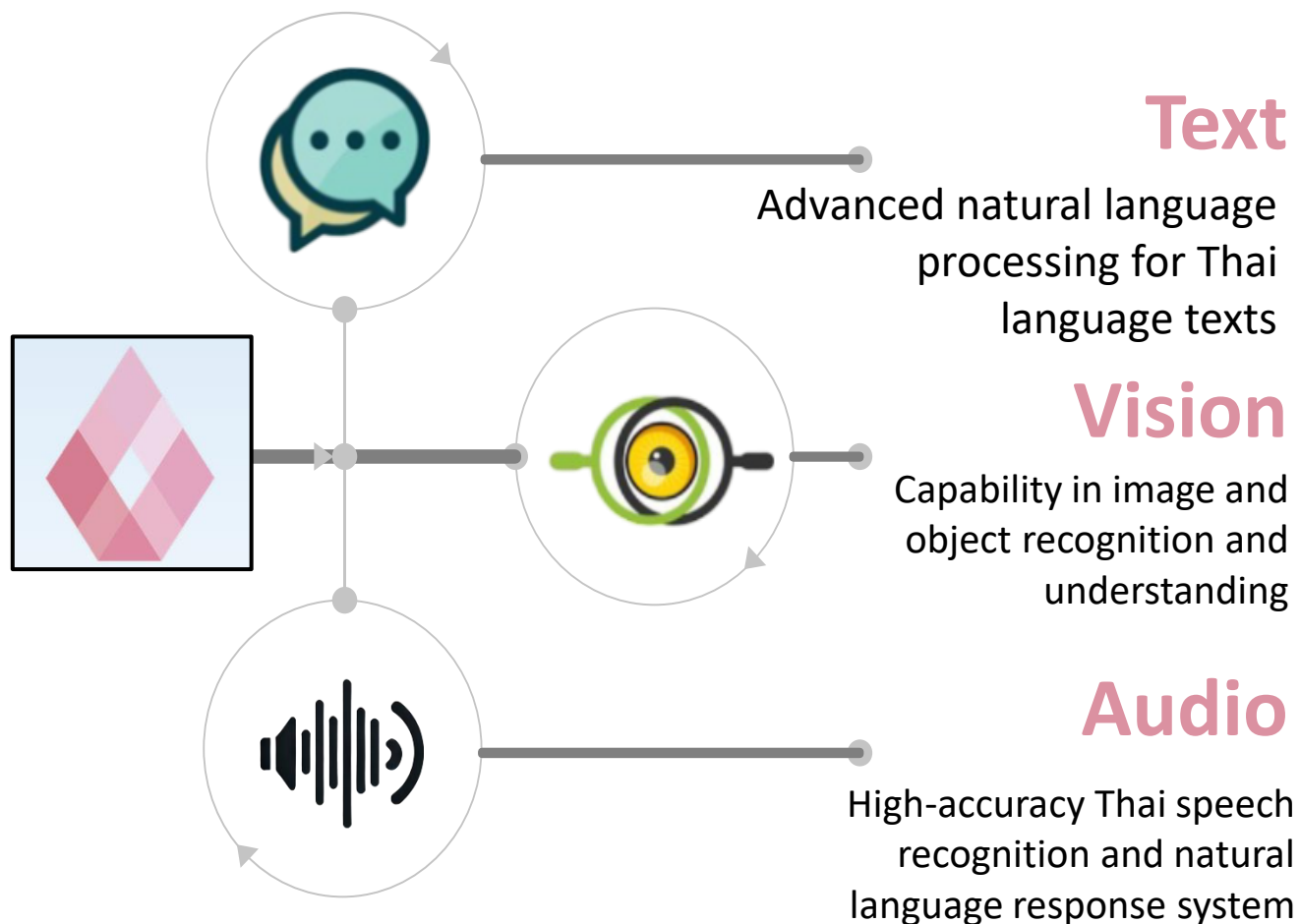
Open-source Foundation Model

Initiated in 2022, currently available in three sizes: 7 billion, 30 billion, and 100 billion parameters (equivalent to the number of neurons in the human brain).



NECTEC's Pathumma LLM

Smart communication with multimodal support: image, speech, and text



Key Features

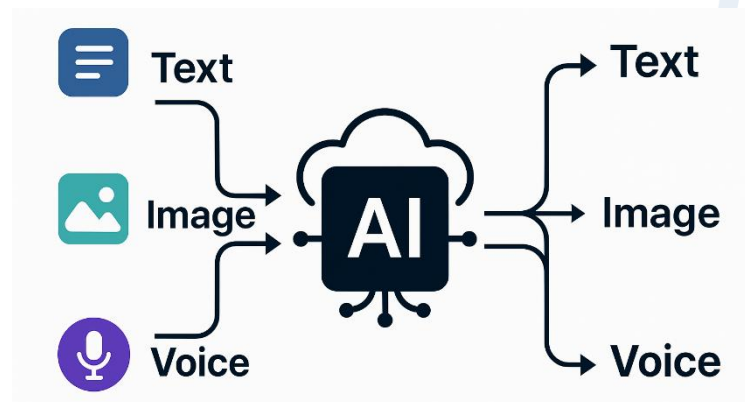
Text: Understanding Thai language context

Speech: Speech-to-text, emotion analysis, and gender detection

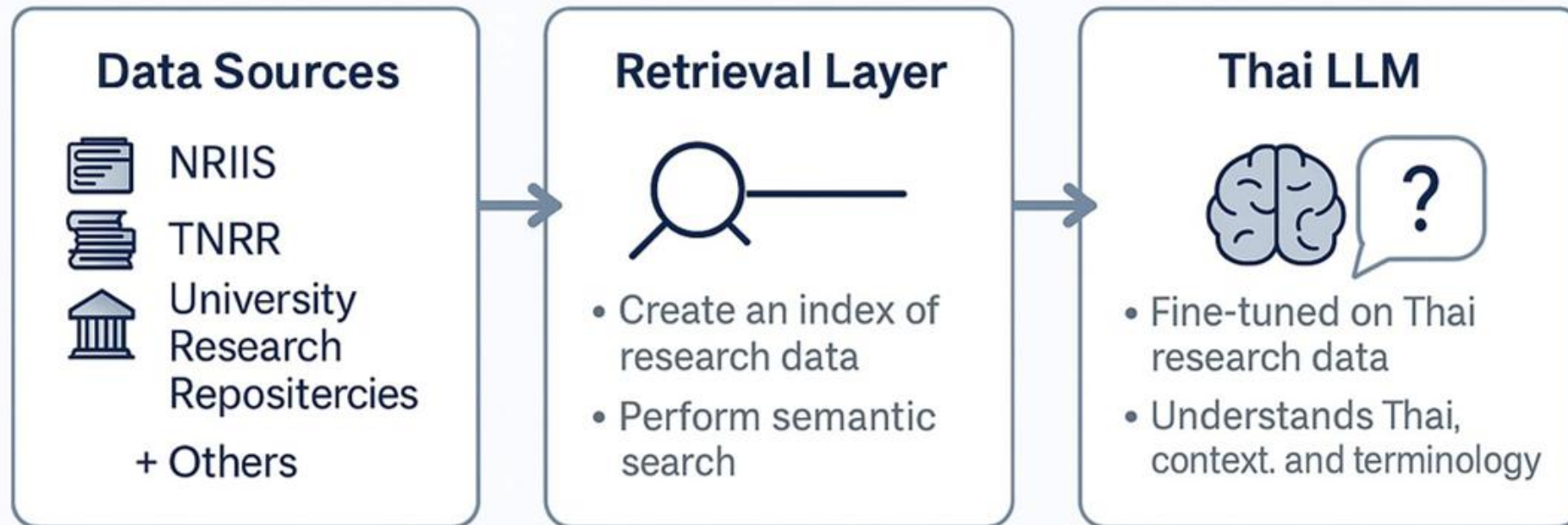
Image: Detailed image description

Multimodal: Supports queries via speech, image, and text

Deep Context: Capable of understanding complex questions



Use Case: Thai LLM-based Research QA System



Advantages: Comprehensive, understands Thai research terms, natural answers

Limitations: Currently focused on NRIIS and TNRR

Use Case: Thai LLM-based Research QA System



Thai Large Language Model-based Research Q&A System

ระบบสืบค้น คลังข้อมูลงานวิจัยไทย ด้วย Thai LLM

TNRR Thai National Research Repository

- ◆ รองรับการค้นหาโดยใช้ ภาษาธรรมชาติ
- ◆ รองรับการค้นหาข้ามภาษา
- ◆ รองรับการค้นหา คำที่มีความหมายใกล้เคียงกัน
- ◆ รองรับการค้นหา คำที่สะกดผิด



NRIIS Chatbot

ใช้สำหรับ ค้นหาข่าวทุน
และ ทุนวิจัย ในระบบ NRIIS



ระบบข้อมูลสารสนเทศวิจัยและนวัตกรรมแห่งชาติ
National Research and Innovation Information System

TNRR DocChat

- ◆ ใช้สำหรับ ถามตอบงานวิจัย
- ◆ สามารถ สรุปสาระสำคัญของเอกสาร
- ◆ ตั้งประเด็นคำถามที่น่าสนใจให้โดยอัตโนมัติ
- ◆ สามารถพูดคุยตอบโต้หรือตั้งคำถามกับเอกสารที่กำหนดได้อย่างเป็นธรรมชาติ

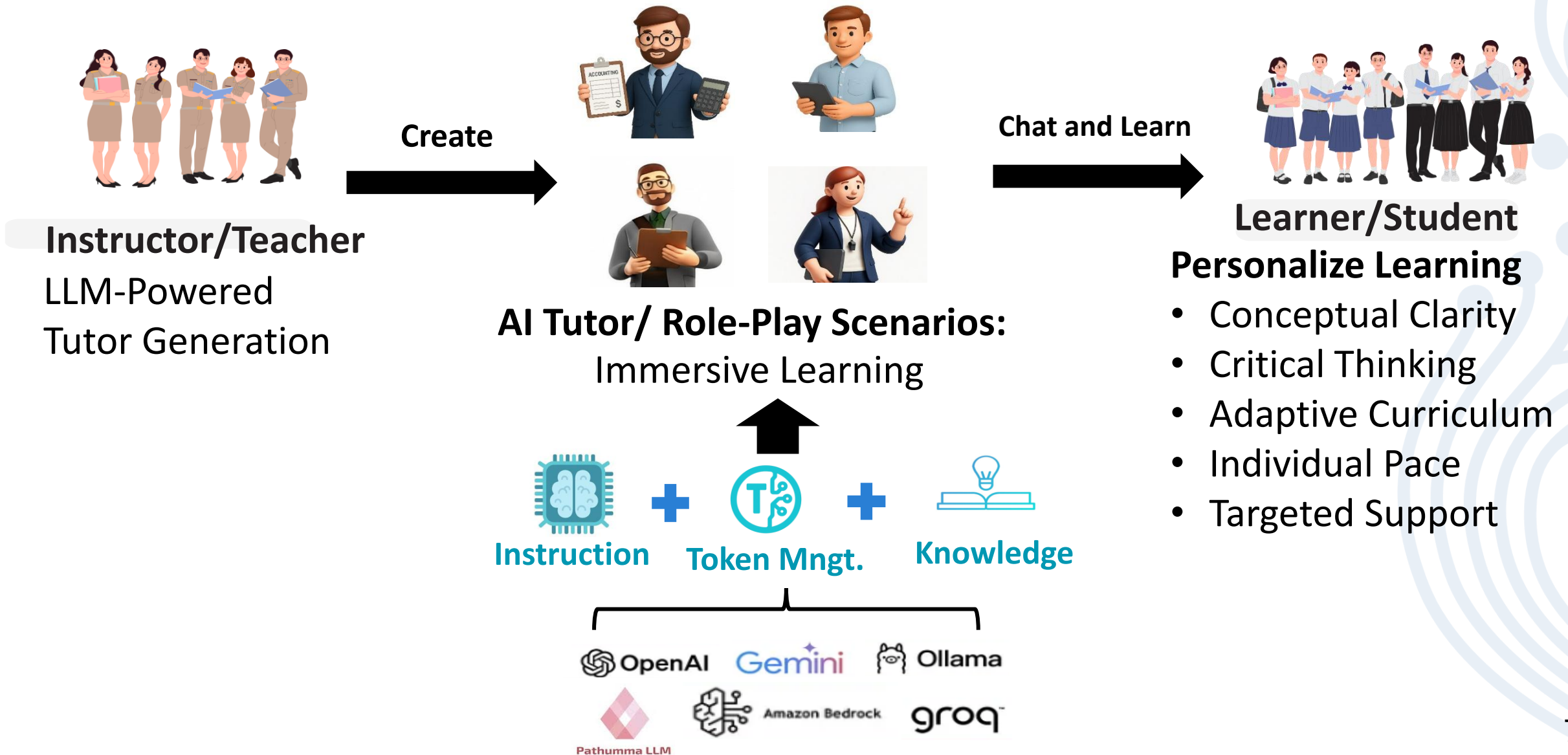
จุดเด่น ของเทคโนโลยี

1. เข้าใจภาษาไทยสำหรับงานวิจัยได้ลึกซึ้ง
ด้วยโมเดลที่ fine-tune จากข้อมูล งานวิจัยไทย โดยเฉพาะ

2. เชื่อมโยงแหล่งข้อมูลหลากหลาย
เริ่มจาก NRIIS และ TNRR และพร้อมขยายต่อ

3. ถาม-ตอบ แบบโต้ตอบธรรมชาติ
โต้ตอบกับผู้ถามอย่างเป็นธรรมชาติ

Use Case: AI-Assisted Active Learning Platform



Use Case: Thai LLM for Public Sectors



Development of a Portal Web for Government Data Integration using Agentic AI

A single access point for citizens to reach services and information from all government agencies.



Development of LLM capability assessment tools for target groups

Expanding the prompt set development process to evaluate Trustworthy AI performance, aimed at establishing baseline criteria for future Sandbox testing

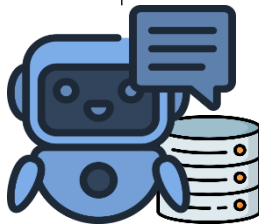
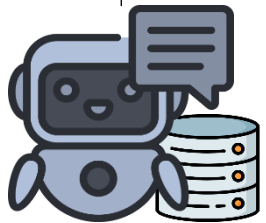


Use Case: Thai LLM for Public Sectors

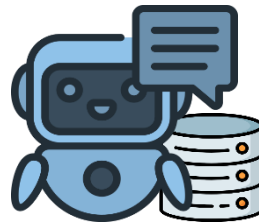


สำนักงานคณะกรรมการพัฒนาระบบราชการ
Office of the Public Sector Development Commission

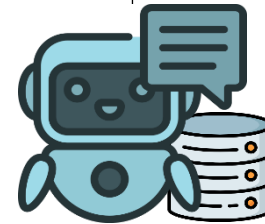
Citizens access data from multiple government agencies that operate on **separate systems**, and complex searches often require retrieving information from **several agencies simultaneously**.



สำนักงานคณะกรรมการอาหารและยา
Food and Drug Administration



กรมบัญชีกลาง
The Comptroller General's Department



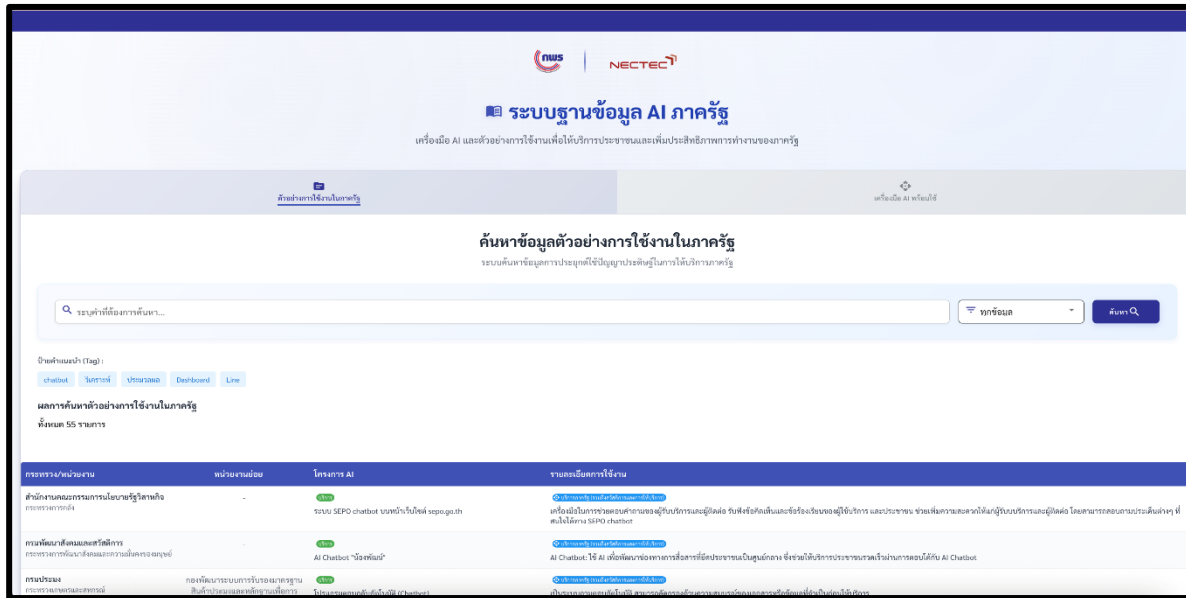
กรมที่ดิน
Department of Lands



กรมการปกครอง กระทรวงมหาดไทย
Department Of Provincial Administration

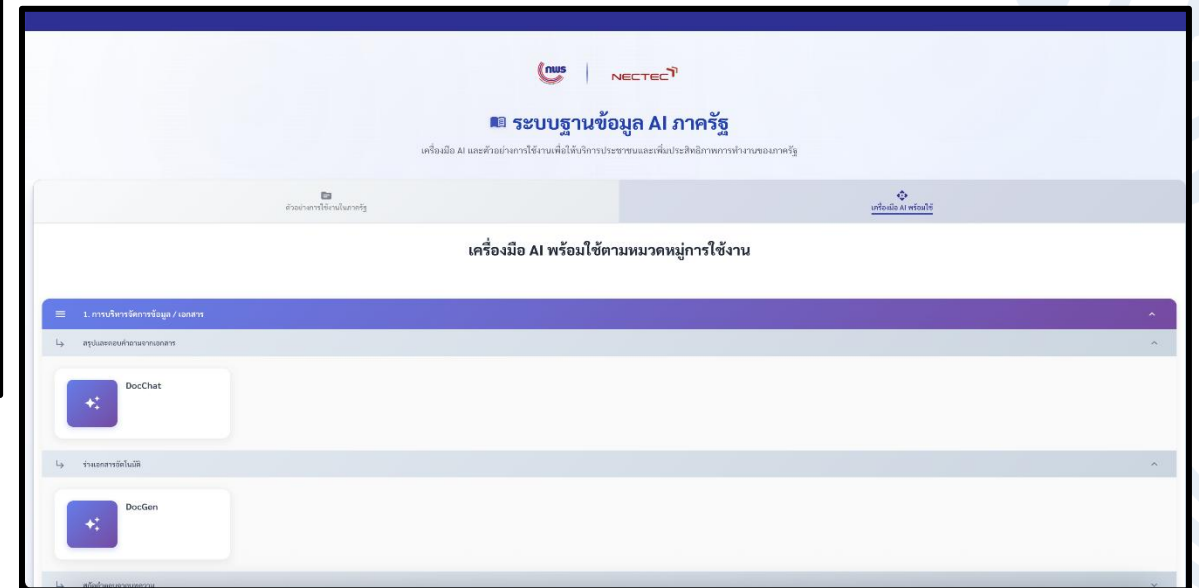
Use Case: Thai LLM for Public Sectors

Government AI Use Case Repository



opdc.ai.in.th

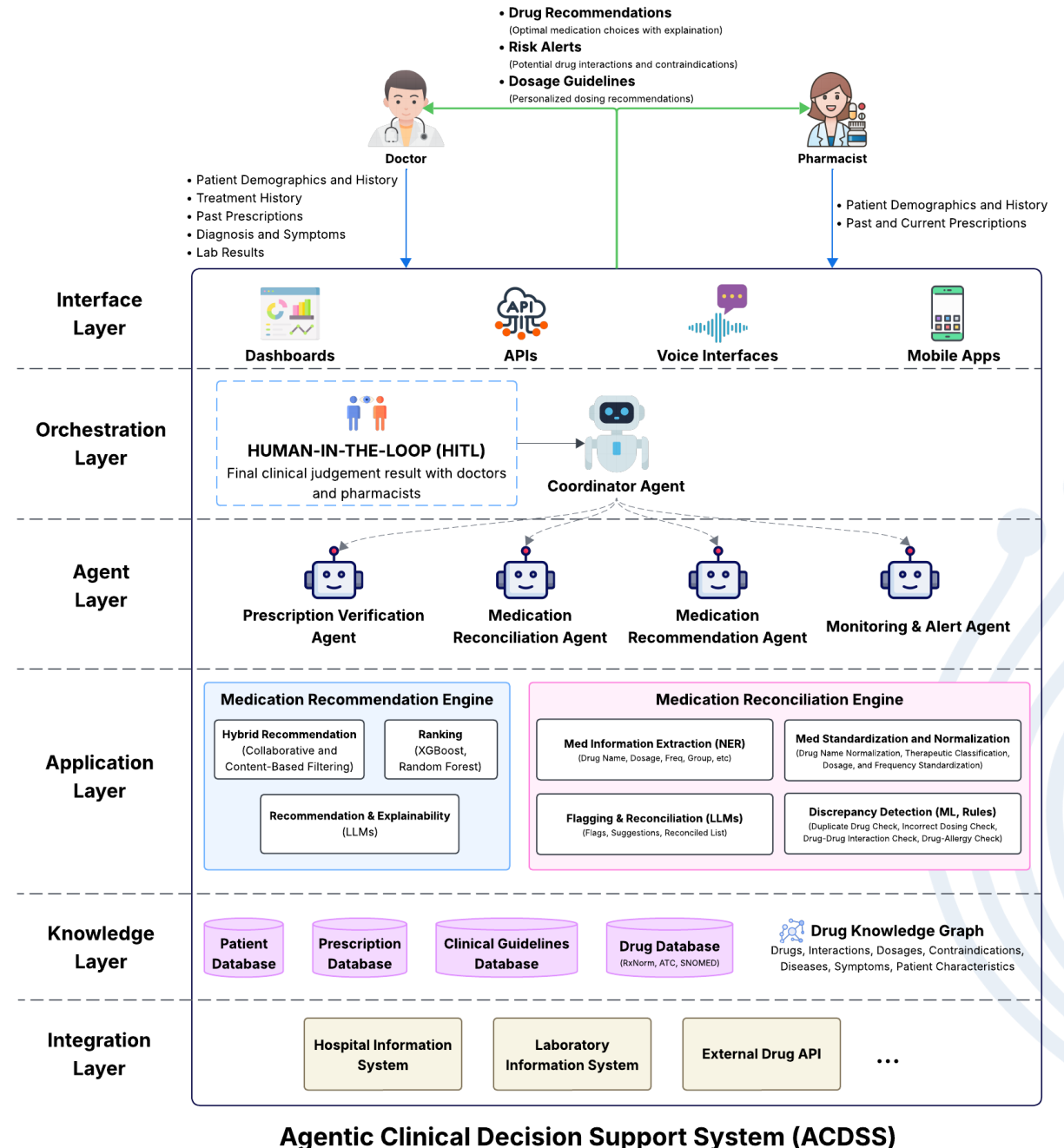
Ready-to-use AI Tools Categorized by Functions



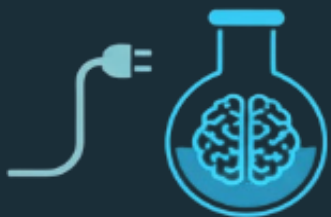
Use Case: Healthcare

Agentic Clinical Decision Support System (ACDSS)

- Drug recommendation
- Voice interface
- Treatment recording
- Prescription verification
- Agent orchestration



Coming Soon



Resource-saved LLM:

Optimizing AI for Efficiency

Unlocking Powerful AI with Less Power & Cost

THE CHALLENGE:

Resource-Hungry AI



Massive Memory
Hundreds of GBs to TBs



High Compute Cost
Expensive GPUs & Energy



Slow & Inefficient
Not for Real-Time Use

4 KEY OPTIMIZATION TECHNIQUES

1. Quantization

Reduces numerical precision (weights). Smaller model, faster computation, lower memory usage.

2. Pruning

Removes redundant parameters. Smaller size, less energy consumption.

3. Knowledge Distillation

Smaller model mimics larger teacher model. New, efficient model with similar performance.

4. Efficient Checkpointing

Saves training progress, provides efficient fine-tuning.

THE IMPACT:

Why Resource-Saving
Matters



Widespread Accessibility:
On-device & IoT



Cost-Efficiency & Speed:
Reduced Infrastructure &
Energy Bills

Coming Soon

AGENTIC TO PHYSICAL AI

Bringing AI Agents into the Real World

Empowering Digital Brains to Control Physical Bodies

1. THE AGENTIC PHASE



- **Digital Brains**
AI models that plan, reason, and make decisions.
- **Simulated Worlds**
Operates in virtual environments.
- **Outputs**
Text, Images, Code, Digital actions

2. THE BRIDGE: PERCEPTION & ACTION

PERCEPTION (Input)



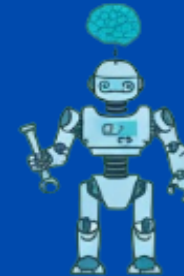
Translating real-world data into digital understanding (e.g. cameras, LIDAR, audio)

ACTION (Output)



Converting digital commands into physical movement (e.g. robot arms, actuators, speaker)

3. THE PHYSICAL AI PHASE



- **Embodied Intelligence**
AI controls robots, drones, & devices
- **Real-World Interaction**
Navigates and manipulates physical environments
- **Applications**
Manufacturing, Healthcare, Logistics, etc.

Future
Beginning
Here!

