



**ĐẠI HỌC**  
**BÁCH KHOA HÀ NỘI**  
HANOI UNIVERSITY  
OF SCIENCE AND TECHNOLOGY

**BKAI**

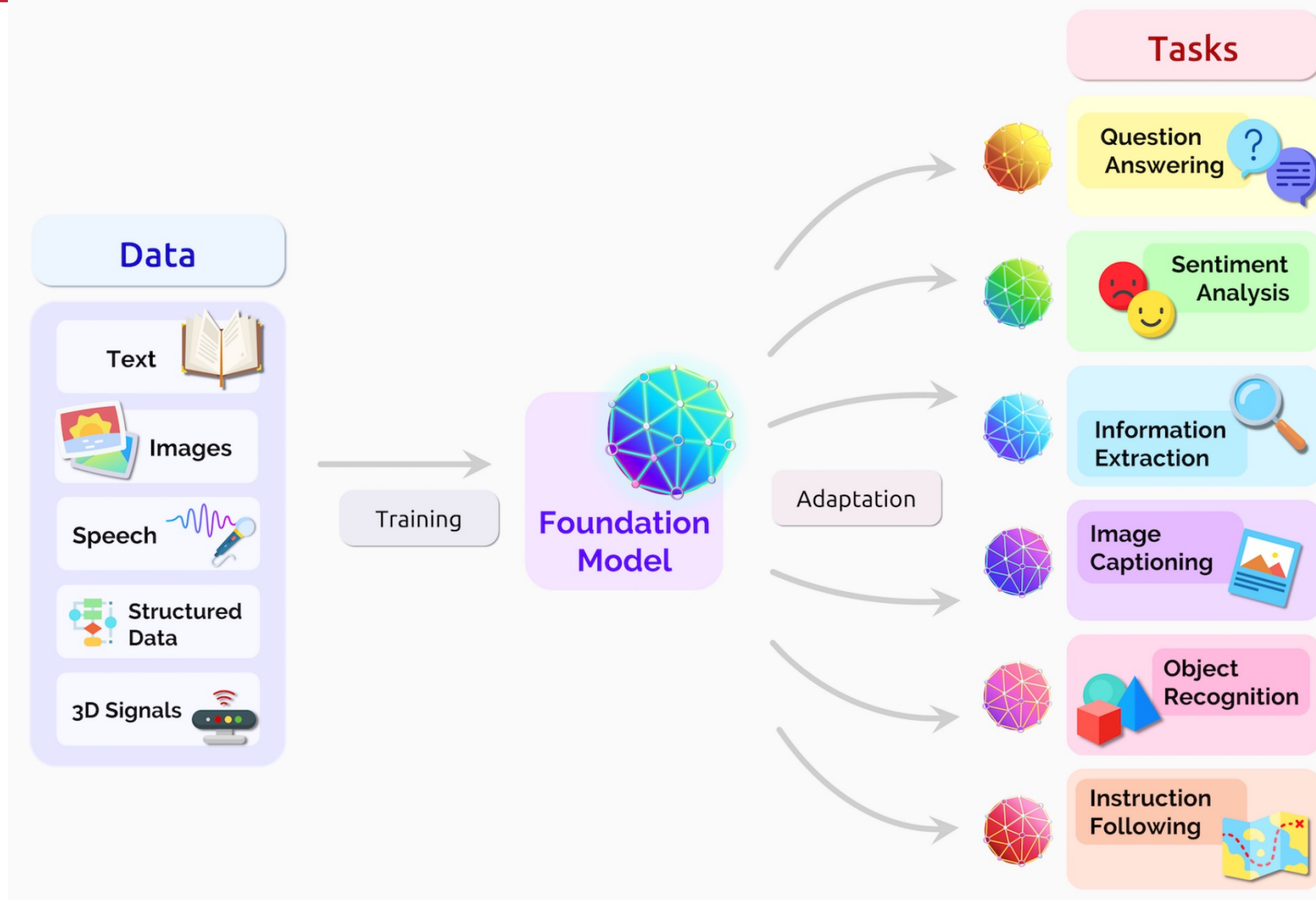
**Foundation Models Labs**

**ONE LOVE. ONE FUTURE.**

1. Giới thiệu về mô hình nền tảng
2. Tình hình nghiên cứu thế giới
3. Thông tin PTN và kế hoạch hoạt động

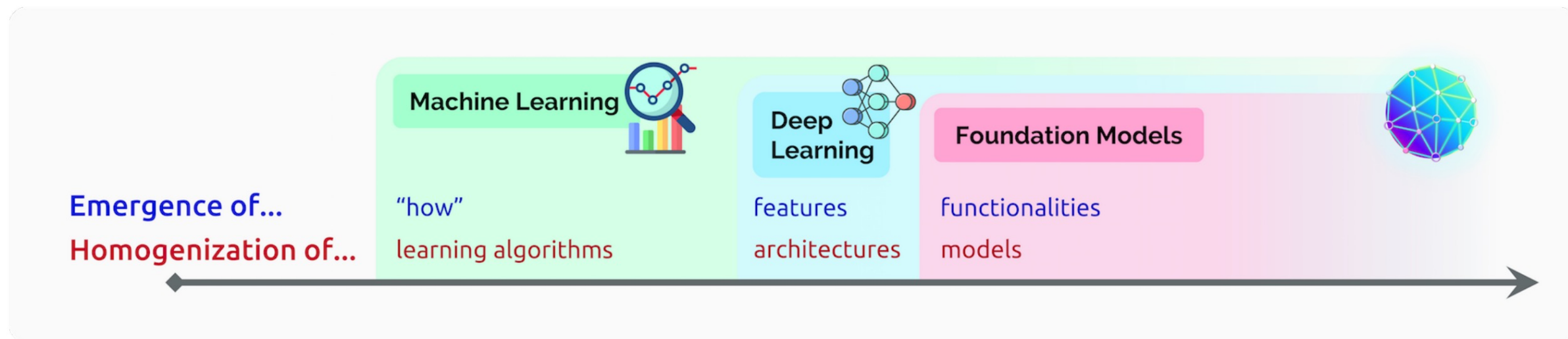
# Mô hình nền tảng là gì?

- Mô hình lớn (nhiều tham số)
- Huấn luyện tự giám sát (self-supervised learning) **một lần** trên tập dữ liệu không nhãn khổng lồ (chi phí tốn kém)
- Hiệu chỉnh (adapt) sang các bài toán cụ thể **nhANH chóng, độ chính xác cao, chi phí rẻ (thậm chí ~ 0)**

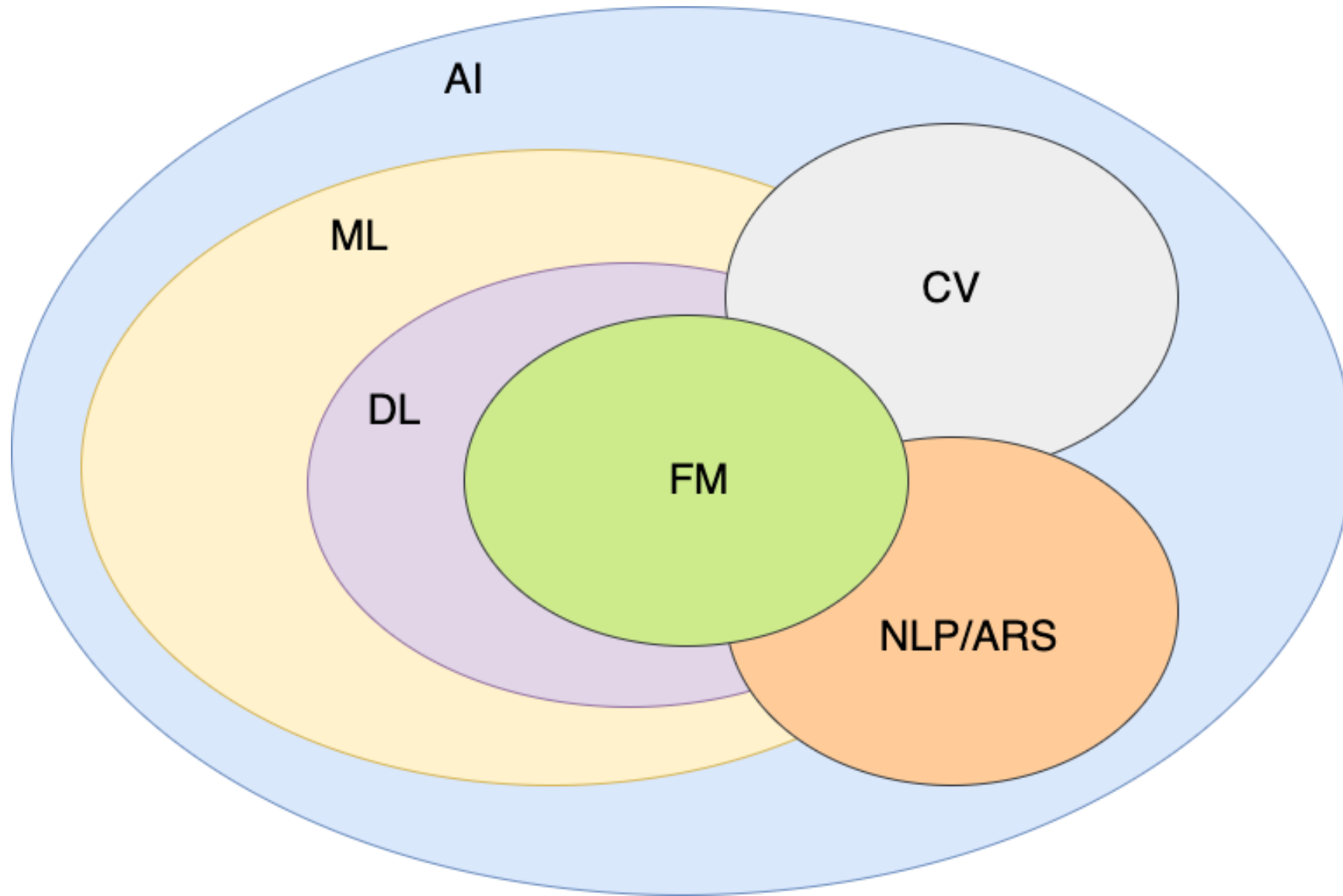


# Tại sao cần mô hình nền tảng?

- **Lập trình truyền thống:** Input + Program = Output
- **Học máy:** Input + Output = Program
  - Mỗi bài toán riêng cần tập dữ liệu riêng, phương pháp học máy riêng, thuật toán học riêng, trích xuất đặc trưng riêng
- **Học sâu:** Mỗi bài toán cần tập dữ liệu riêng, phương pháp học máy chung (mạng nơ-ron) nhưng kiến trúc riêng, thuật toán học chung
- **Mô hình nền tảng:** dữ liệu chung, mô hình chung, thuật toán chung → **Đồng nhất hoá AI**

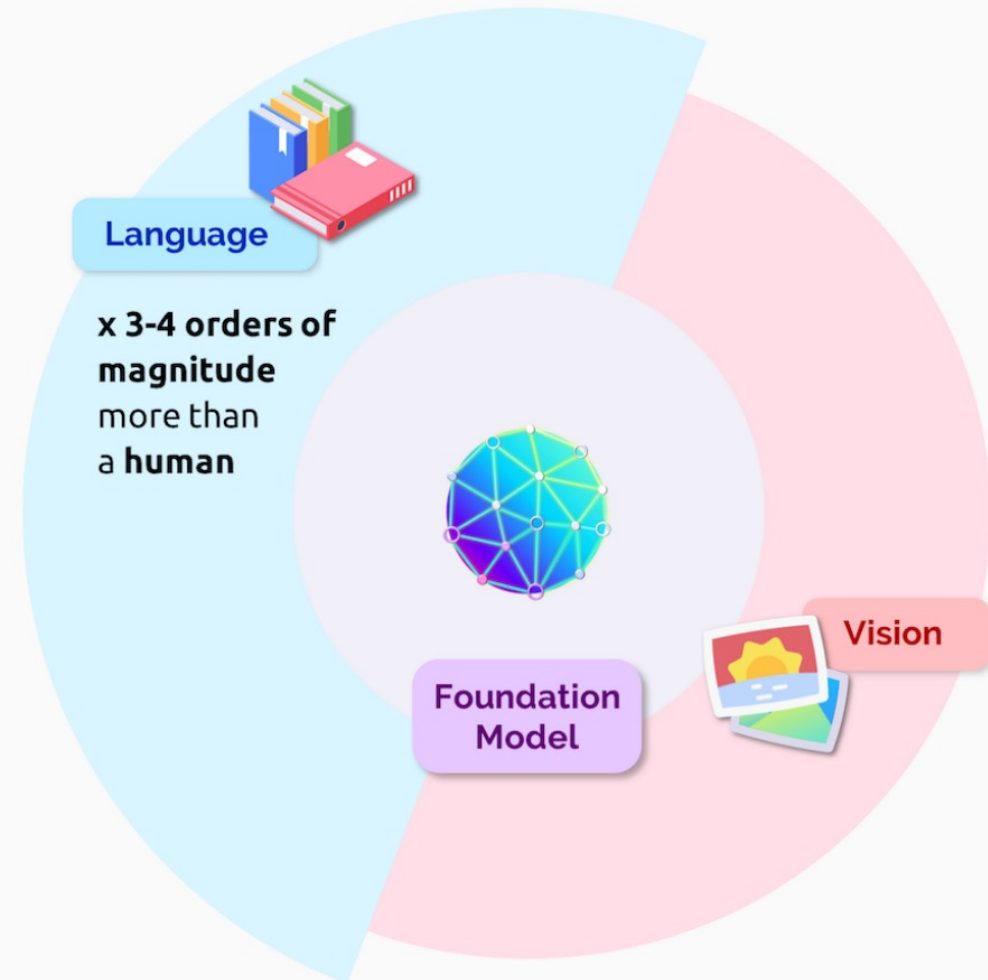
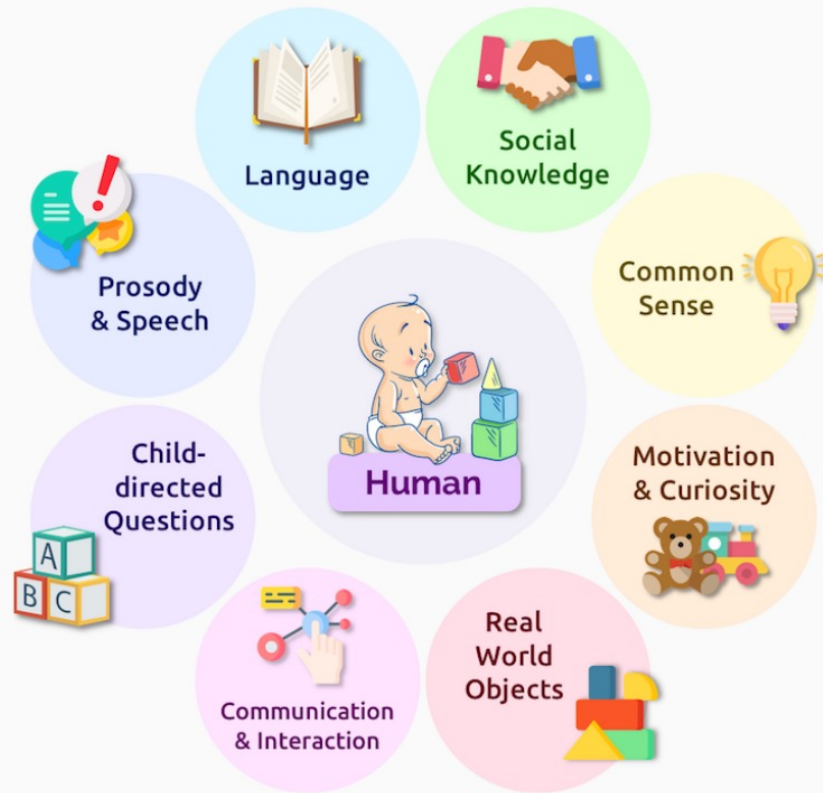


# Mối liên quan giữa Foundation Models và các lĩnh vực khác

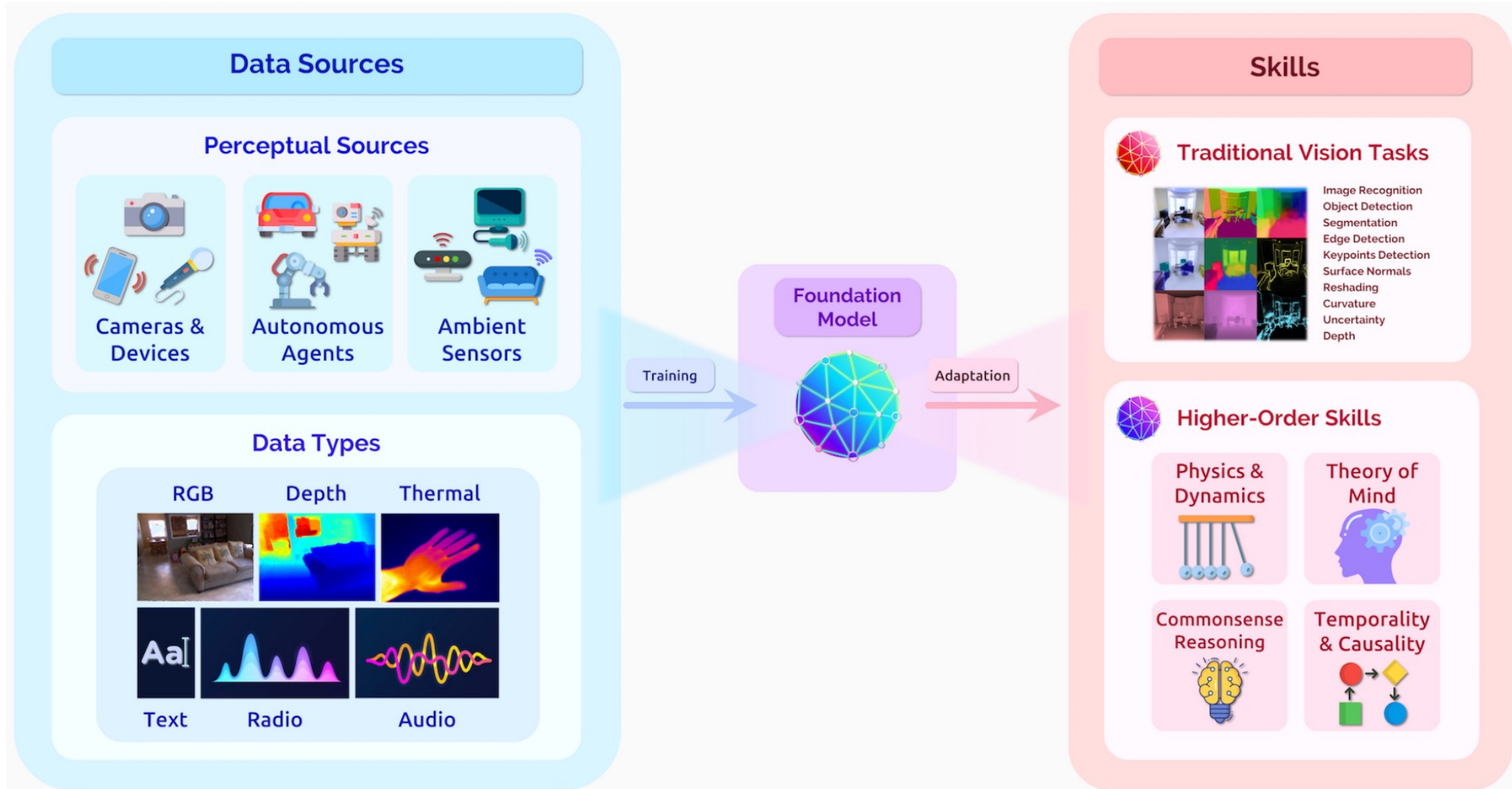


# Một số loại mô hình nền tảng: Large Language Models

## Language Acquisition



# Một số loại mô hình nền tảng: Visual Foundation Models



# Một số loại mô hình nền tảng: Foundation Models for Robotics

## Data Sources (2.3.2)

Robotic Interaction



Videos of Humans



Simulation



Natural Language

"Pick up the cup. Turn on the stove."

Training

Foundation Model



Adaptation

## Tasks (2.3.1)

Intuitive, multi-modal task specification

"Make a sandwich"  
*input*



Reward Function  
*output*



Fast adaptation for task learning

Policy in  
Kitchen A  
*input*



"Open Fridge"



Policy in  
Kitchen B  
*output*

"Open Fridge"



Adapts to new tasks, environments, and embodiments.



# Một số loại mô hình nền tảng khác

Google Research

Philosophy

Research Areas

Publications

People

Resources

Outi

## Universal Speech Model (USM): State-of-the-art speech AI for 100+ languages

MONDAY, MARCH 06, 2023

Posted by Yu Zhang, Research Scientist, and James Qin, Software Engineer, Google Research



ibm.com

<https://research.ibm.com> > publications · Dịch trang này

## Foundation Model for Material Science - IBM Research

viết bởi S Takeda · 2023 — In this paper, we propose building an FM for **material science**, which is trained with massive data across a wide variety of material domains and data...



[About](#) [Platforms](#) [Pipeline](#) [Partnerships](#)

## Foundation Models for Proteins: Cyrus, OpenFold and the future of biologics

October 3, 2022



nature

Explore content

About the journal

Publish with us

[nature](#) > [perspectives](#) > article

Perspective | [Published: 12 April 2023](#)

## Foundation models for generalist medical artificial intelligence



ACL Anthology

[FAQ](#)

[Corrections](#)

[Submissions](#)

Search...

## Foundation Models of Scientific Knowledge for Chemistry: Opportunities, Challenges and Lessons Learned

Sameera Horawalavithana, Ellyn Ayton, Shivam Sharma, Scott Howland, Megha Subramanian, Scott Vasquez, Robin Cosbey, Maria Glenski, Svitlana Volkova

Stanford University



Google AI Why AI Responsibility Discover Build

## Foundation Models

Learn more about Google's foundation models that include text-to-image, text-to-code and speech-to-text.

Meta

Meta Research | Programs | Requests for Proposals | Research Areas | Publications

Our People | Blog

## LLaMA: Open and Efficient Foundation Language Models

research.ibm.com/artificial-intelligence

IBM Research Focus areas | Publications Collaborate Careers Events About

## What's Next in AI is *foundation models at scale*

microsoft.com  
https://www.microsoft.com > project > f... · Dịch trang này

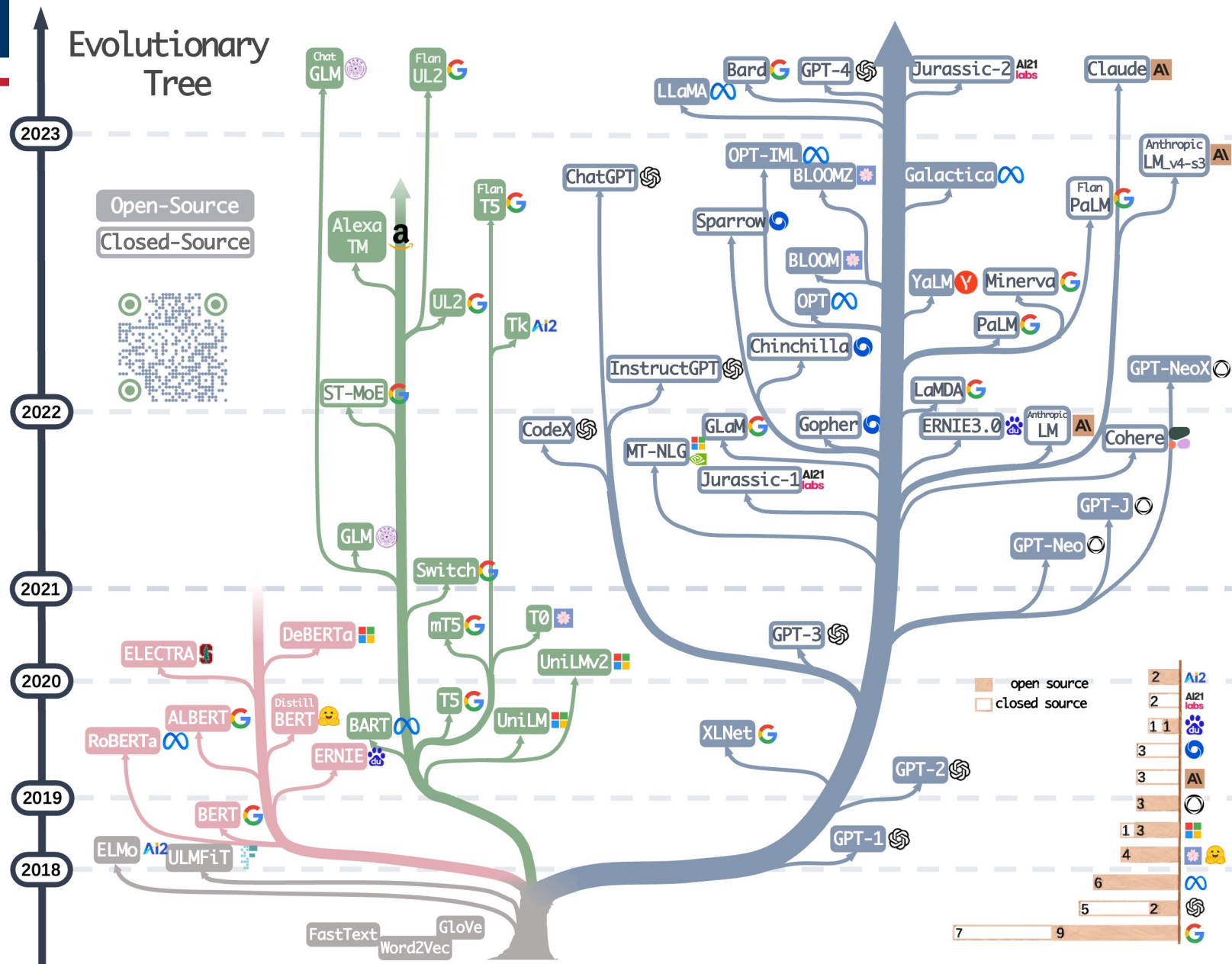
## Foundation Models - Microsoft Research

**Foundation Models.** Larges-scale Pre-training across Tasks, Languages and Modalities.  
Overview · People · Large-scale Self-supervised Pre-training across ...

# Tình hình thế giới: LLMs

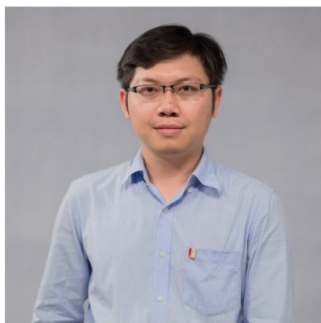
- Dữ liệu cho tiếng Việt còn hạn chế
- Các thuật toán chưa tối ưu cho tiếng Việt
- Chi phí đắt đỏ

➔ Tính cấp thiết trong việc nghiên cứu **các mô hình nền tảng cho dữ liệu Việt** và ứng dụng trong **các bài toán mang tính đặc thù cho Việt Nam**

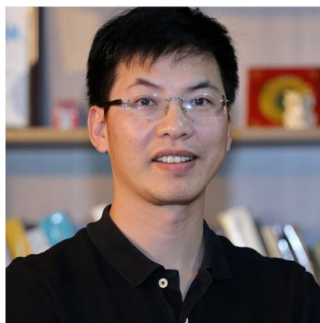


# Thông tin PTN

- Tên: Foundation Models Labs
- Địa điểm: Tầng 10 B1 (trực thuộc BKAI)



DR. DINH VIET SANG  
HEAD



DR. TRAN VIET HUNG  
CO-HEAD



ASSOC. PROF. LE THANH  
HUONG  
MEMBER



DR. TRAN VAN DANG  
MEMBER



DR. NGUYEN PHI LE  
MEMBER



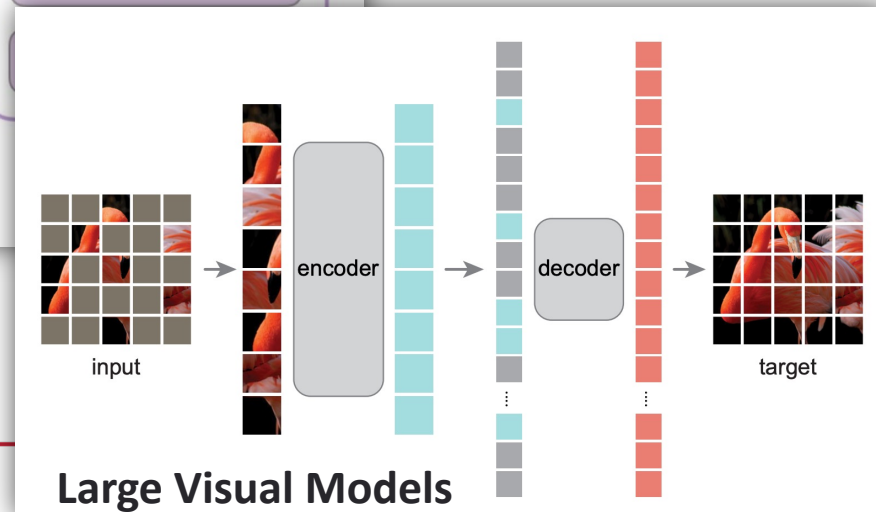
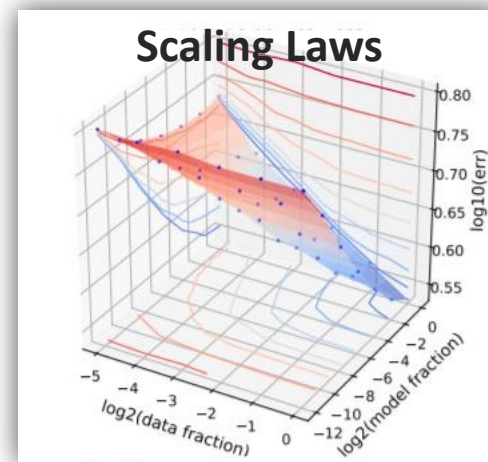
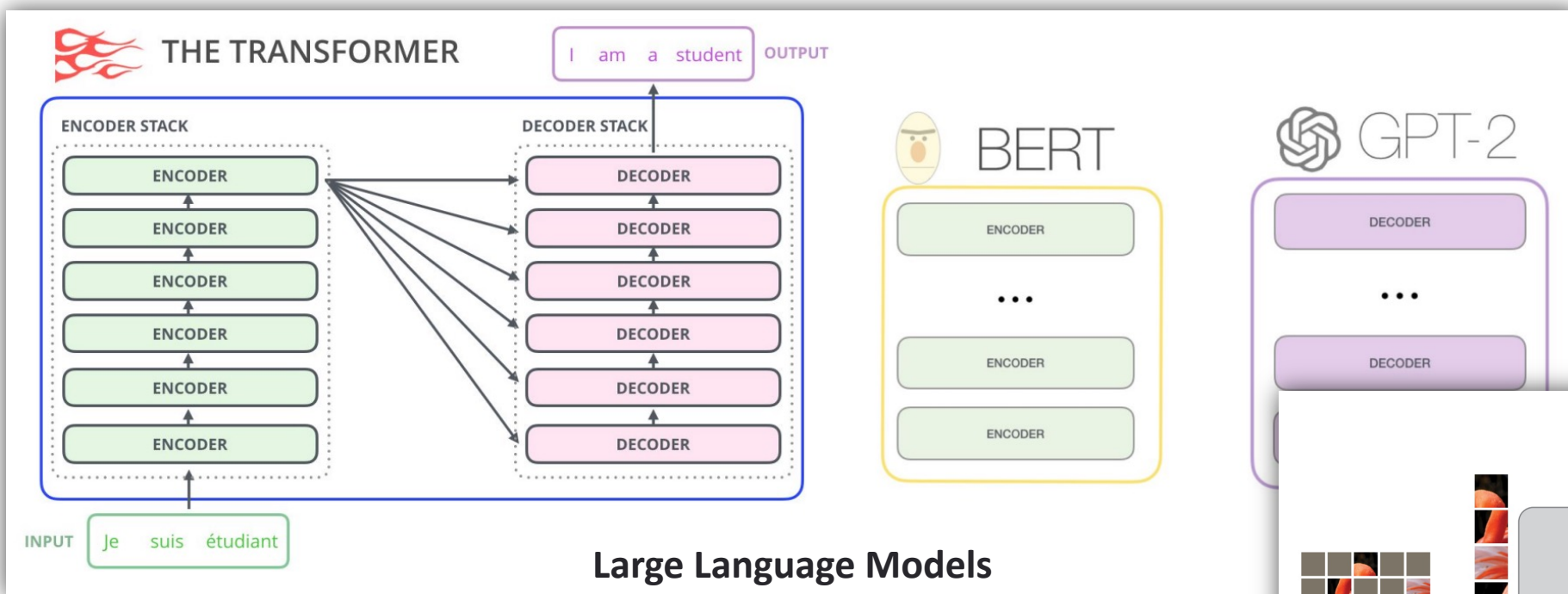
DR. DANG TUAN LINH  
MEMBER



DR. DINH THI HA LY  
MEMBER

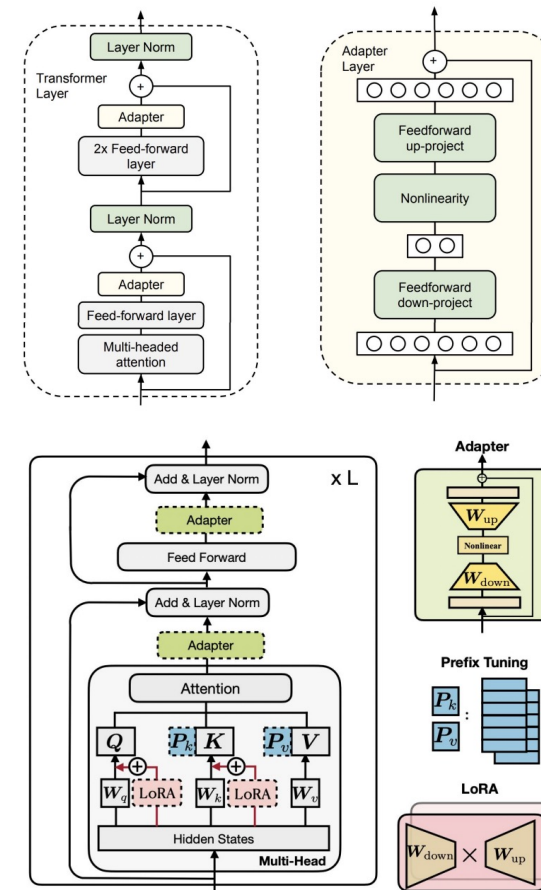
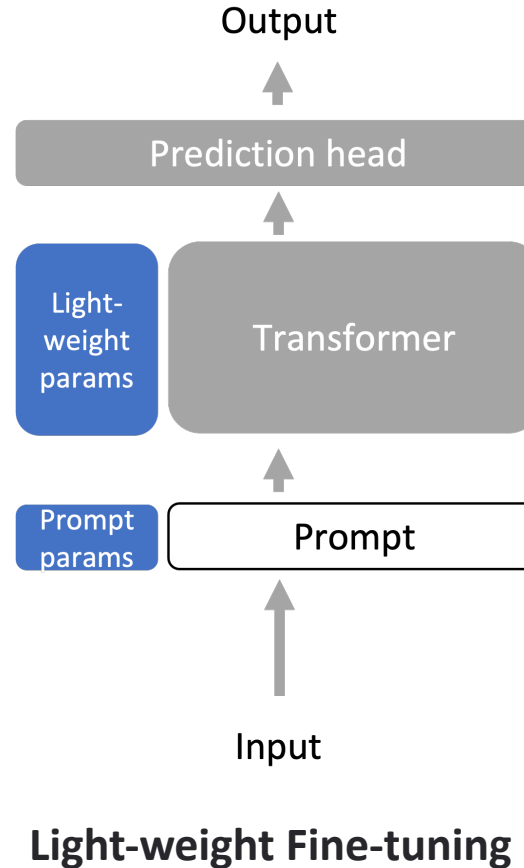
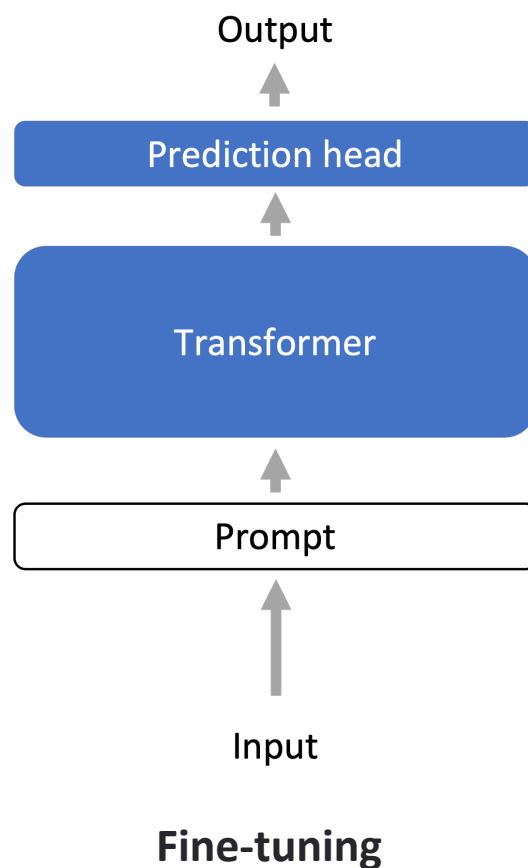
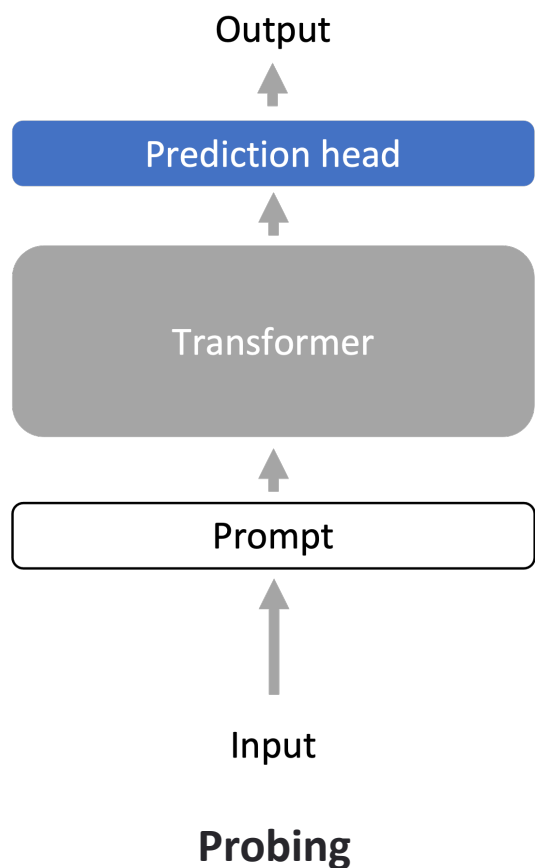
# Định hướng hoạt động

- Nghiên cứu phát triển các mô hình nền tảng cho các loại dữ liệu **ngôn ngữ**, hình ảnh, âm thanh, ... và các nguồn dữ liệu đa thể thức



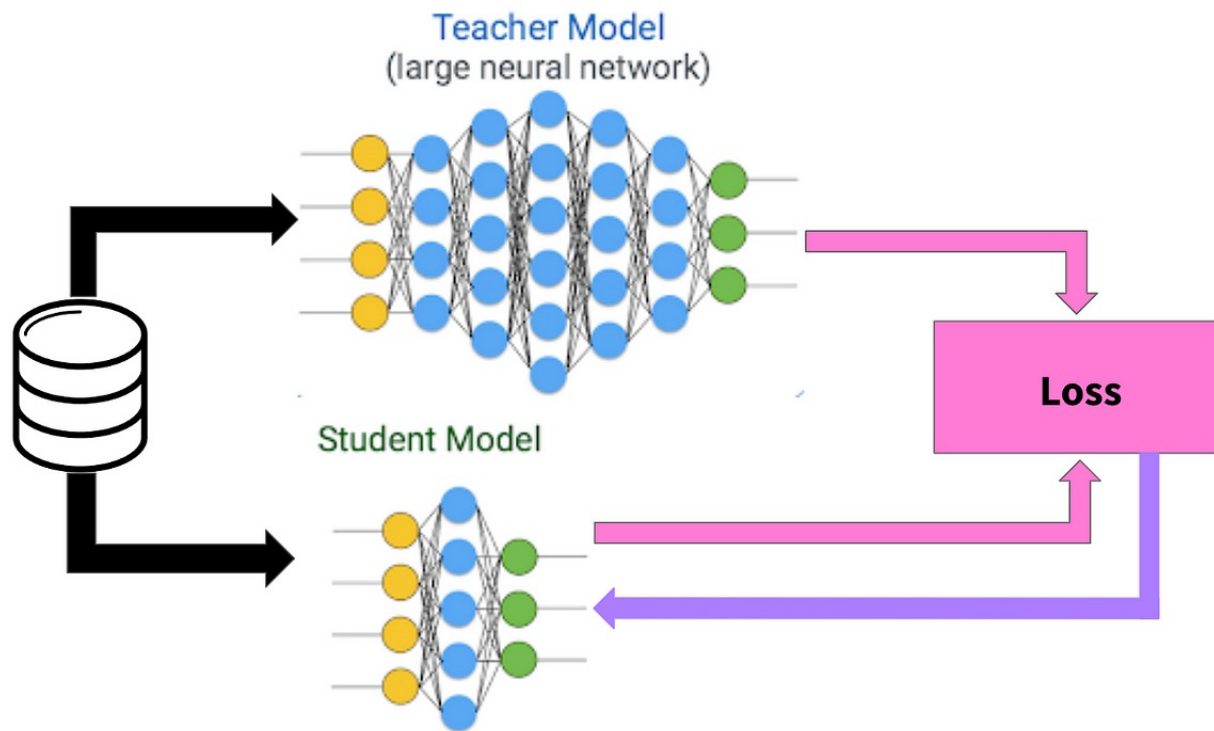
# Định hướng hoạt động

- Hiệu chỉnh (adapt) mô hình cho các tác vụ đặc thù, đặc biệt chú trọng cho dữ liệu tiếng Việt

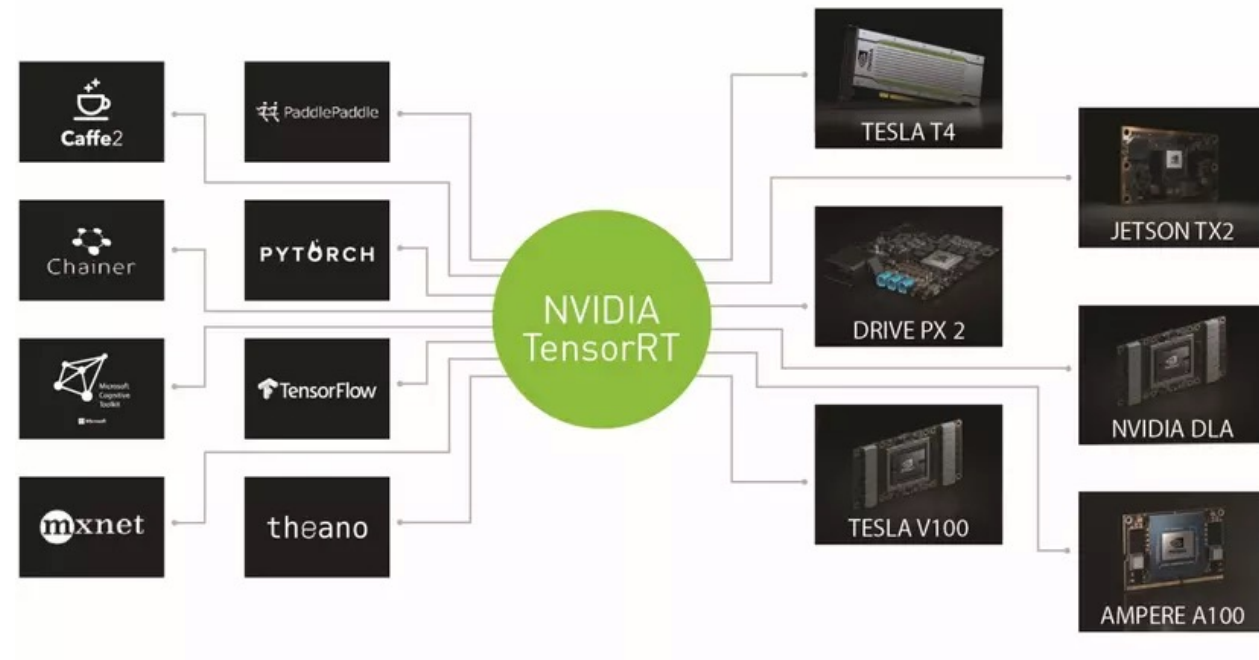


# Định hướng hoạt động

- Làm giảm kích thước và tăng tốc, tối ưu mô hình
- Xử lý dữ liệu độc hại (toxic data), chống thiên vị (debiasing)



Knowledge Distillation



Model Optimization

# Định hướng hoạt động

- Nghiên cứu Prompt Engineering
- Triển khai các khoá đào tạo về Foundation Models và Prompt Engineering

## Chain of Thought (CoT)

(a) Few-shot

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?  
A: The answer is 11.

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?  
A:

(Output) The answer is 8. ✗

(b) Few-shot-CoT (Wei et al)

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?  
A: Roger started with 5 balls. 2 cans of 3 tennis balls each means he bought 6 more balls. 5 + 6 = 11. The answer is 11.

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?  
A:

(Output) The juggler can juggle 16 balls. Half of the balls are golf balls. So there are  $16 / 2 = 8$  golf balls. Half of the golf balls are blue. So there are  $8 / 2 = 4$  blue golf balls. The answer is 4. ✓

(c) Zero-shot

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?  
A: The answer (arabic numerals) is

(Output) 8 ✗

(d) Zero-shot-CoT (KoJima et al)

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?  
A: **Let's think step by step.**

(Output) There are 16 balls in total. Half of the balls are golf balls. That means that there are 8 golf balls. Half of the golf balls are blue. That means that there are 4 blue golf balls. ✓

### In-Context Learning

The diagram illustrates In-Context Learning with GPT-3. It shows three types of prompts: Zero-shot, One-shot, and Few-shot. Each prompt is followed by an arrow pointing to a GPT-3 robot icon, which then produces an output. The outputs are: 'fromage' for the Zero-shot prompt, 'loutre de mer' for the One-shot prompt, and 'menthe poivrée' for the Few-shot prompt.

**Zero-shot**  
The model predicts the answer given only a natural language description of the task. No gradient updates are performed.

1 Translate English to French: ← task description  
2 cheese => ..... ← prompt

**One-shot**  
In addition to the task description, the model sees a single example of the task. No gradient updates are performed.

1 Translate English to French: ← task description  
2 sea otter => loutre de mer ← example  
3 cheese => ..... ← prompt

**Few-shot**  
In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.

1 Translate English to French: ← task description  
2 sea otter => loutre de mer ← examples  
3 peppermint => menthe poivrée ← examples  
4 plush girafe => girafe peluche ← examples  
5 cheese => ..... ← prompt

fromage

GPT-3





# Định hướng hoạt động

- Nghiên cứu phát triển các thư viện, hệ thống, ứng dụng dựa trên các mô hình nền tảng

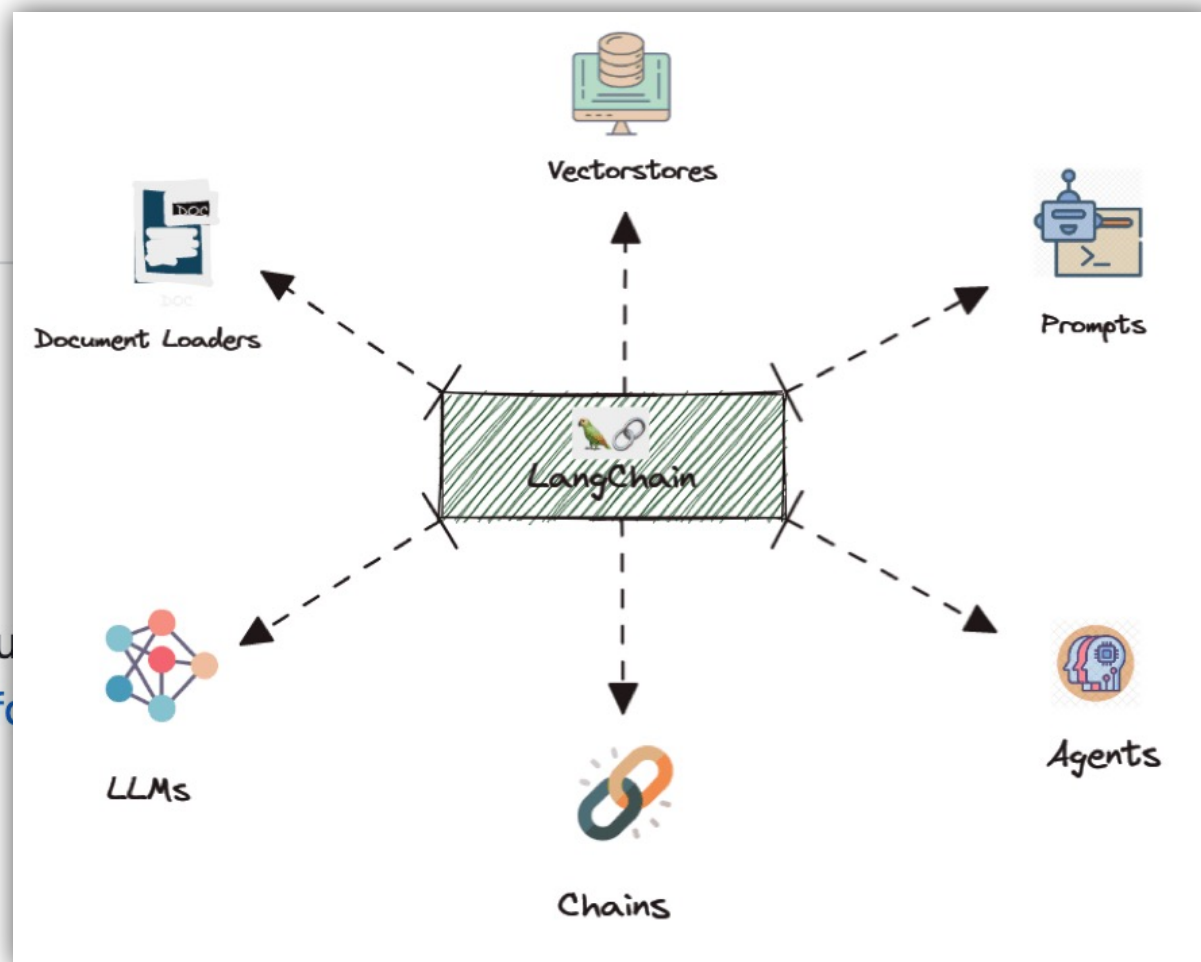


⚡ Building applications with LLMs through composability ⚡

lint passing test passing linkcheck passing License MIT

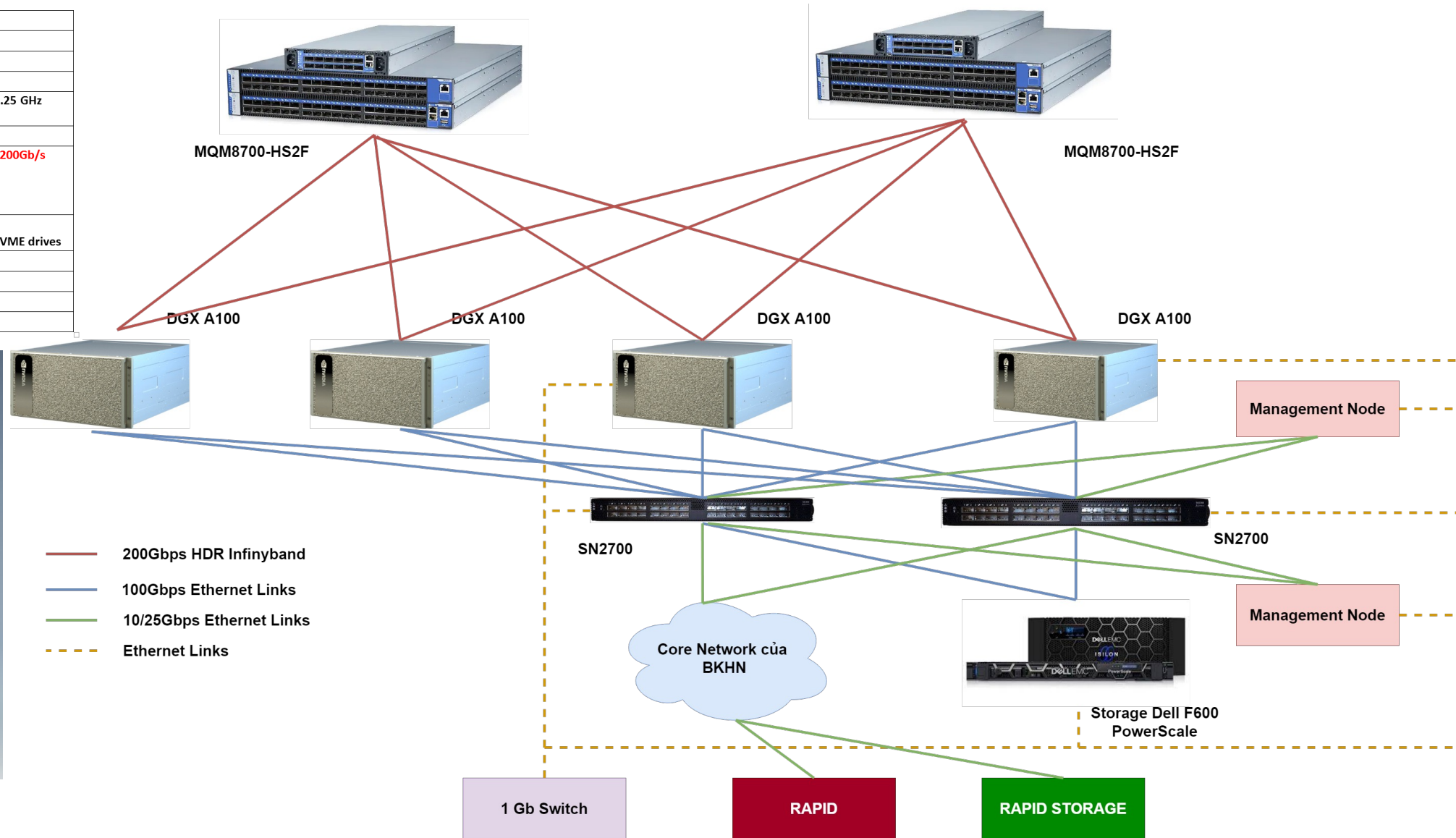
Follow @LangChainAI LangChain

**Production Support:** As you move your LangChains into production, we love to offer more comprehensive support. Please fill out [this form](#) to set up a dedicated support Slack channel.



# Hạ tầng tính toán và lưu trữ

Số GPUs	8x NVIDIA A100 Tensor Core GPUs
GPU Memory	320 GB
Hiệu năng tính toán	5 petaFLOPS AI 10 petaOPS INT8
Số NVSwitches	6
CPU	Dual AMD Rome 7742, 128 cores total, 2.25 GHz (base), 3.4 GHz (max boost)
System Memory	1TB
Networking	8x Single-Port Mellanox ConnectX-6 VPI 200Gb/s HDR InfiniBand 1x Dual-Port Mellanox ConnectX-6 VPI 10/25/50/100/200Gb/s Ethernet
Lưu trữ	OS: 2x 1.92TB M.2 NVME drives Internal Storage: 15TB (4x 3.84TB) U.2 NVME drives
Hệ điều hành	Ubuntu Linux OS
Khối lượng máy	271 lbs (123 kgs)
System Dimensions	Height: 10.4 in (264.0 mm) ~ 6U
Điện năng tiêu thụ	6.5kW max





**HUST**

**THANK YOU !**