Foundation Models for Healthcare

Akshay Chaudhari
Tanveer Syeda-Mahmood
James Zou

Stanford MEDICINE

Stanford AIMI

BIODS 271: Foundation Models for Healthcare

Stanford University, Spring 2025

Akshay Chaudhari
akshaysc@stanford.edu
Office Hours TBD

James Zou
jamesz@stanford.edu
Office Hours TBD

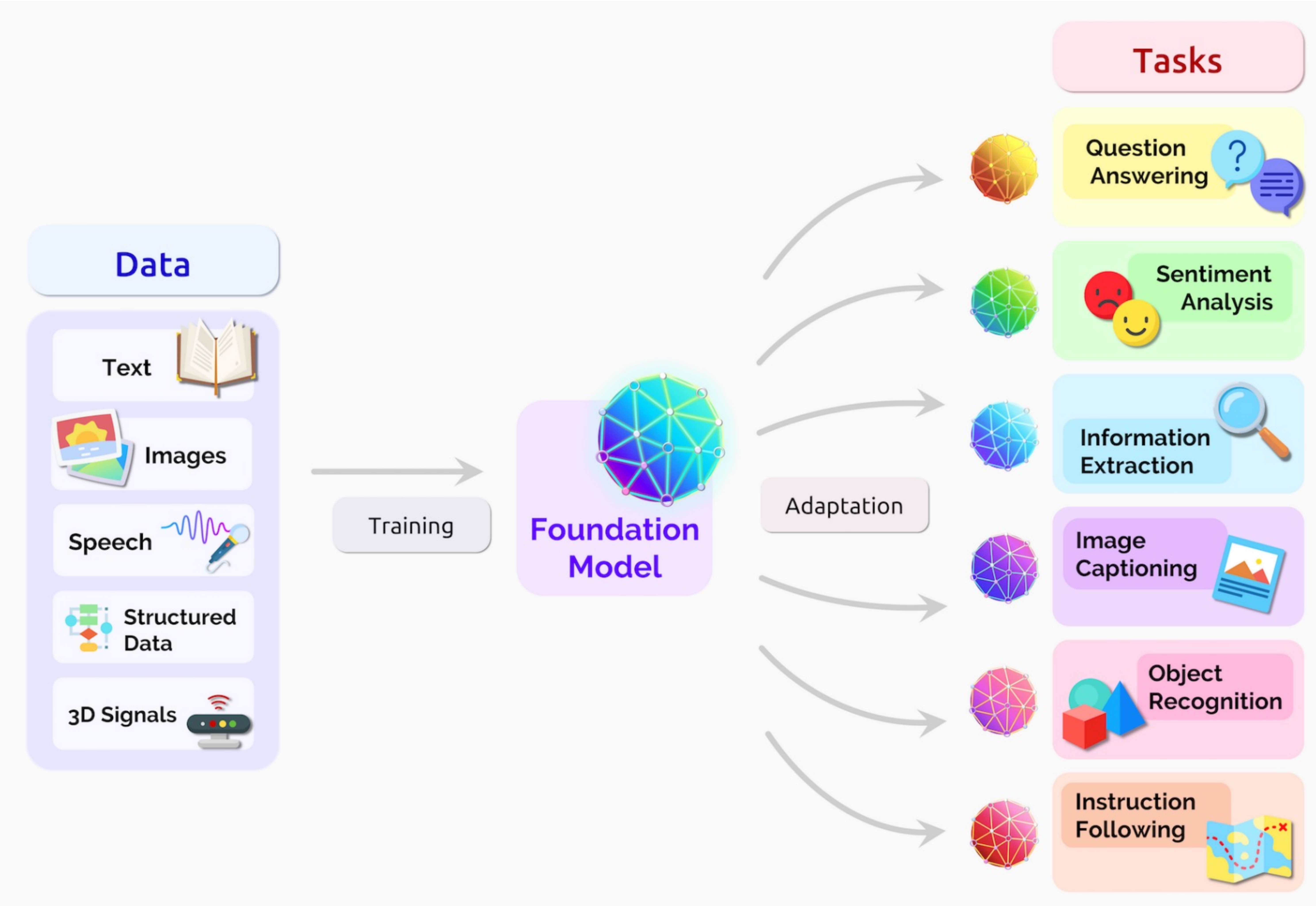
Tanveer Syeda-Mahmood
stf@us.ibm.com
Office Hours TBD

Maya Varma
mvarma2@stanford.edu
Office Hours Tuesdays 7-8pm

Nitya Thakkar
nityat@stanford.edu
Office Hours TBD

https://web.stanford.edu/class/biods271/

Foundation Models



Bommasani et al. On the Opportunities and Risks of Foundation Models. arXiv 2021.

What to Expect from Course

- Provide background and discussion of (some) of the latest large-scale foundation models
- Discuss applications, exciting opportunities, and challenges for adaptation to healthcare
- Educate in this rapidly moving area and help enable high-quality research and applications

What to Expect from Students

- Active participation in course content - discussions, readings, and assignment
- Working towards a high-quality final project
- Flexibility to adapt to a fast-changing field!!

To What Extent Do You Use FMs?

A. I have used FMs occasionally
B. I use FMs during my day-to-day activities
C. I am a power-user of FMs
D. I exclusively vibe code

What are Your Common FM Uses?

A. Coding
B. General writing
C. Image generation/manipulation
D. Agentic workflows
E. Other

(Rough) Course Outline – Part 1

Week 1

Mar 31:	BACKGROUND	**Course Overview & Introduction to Foundation Models**
Apr 02:	BACKGROUND	**Introduction to LLMs**

Week 2

Apr 07:	APPLICATIONS	**AI Scientist Agents**
Apr 09:	OPPORTUNITIES	**Current Advances in LLMs and Agents**

Week 3

Apr 14:	BACKGROUND	**Evaluations of LLMs and agents**
Apr 16:	APPLICATIONS	**Generative AI for Drug Discovery**

(Rough) Course Outline – Part 2

Week 4

Apr 21:	BACKGROUND **Self-Supervised Learning for Vision** HOMEWORK #1 DUE
Apr 23:	BACKGROUND **Self-Supervised Vision-Language Models**

Week 5

Apr 28:	APPLICATIONS **Generative VLMs in Health**
Apr 30:	APPLICATIONS **Foundation Models for Imaging**

Week 6

May 05:	BACKGROUND **Foundation Model Adaptation and Evaluation** HOMEWORK #2 DUE

(Rough) Course Outline – Part 3

May 07:	BACKGROUND	Improving LLM Performance

Week 7

May 12:	OPPORTUNITIES	Training and Deployment Considerations
May 14:	BACKGROUND	Mixture of Expert Models

Week 8

May 19:	OPPORTUNITIES HOMEWORK #3 DUE	Deployment Considerations
May 21:	OPPORTUNITIES	Inference Scaling and Reasoning

(Rough) Course Outline – Part 4

Week 9

May 28:	APPLICATIONS **EMR Foundation Models**

Week 10

Jun 2:	OPPORTUNITIES **Bias and Health Equity**
Jun 4:	PROJECT PRESENTATIONS DUE **Final Presentations**

Week 11

Jun 11:	PROJECT REPORTS DUE

Course Grading

- 10% Class participation
- 45% Homework (3 assignments)
- 45% Final Project

Classic ML Models

Input → **ML Model** → Output

Input

- Structured
- Time series
- Text
- Image
- Audio
- Video
- Table
- Domain-specific
- Multi-D
- Multimodal

Output

- Label
- Regression
- Text
- Image
- Audio
- Video
- Table
- Domain-specific

Function

- Clustering
- Classification
- Prediction
- Regression
- Synthesis

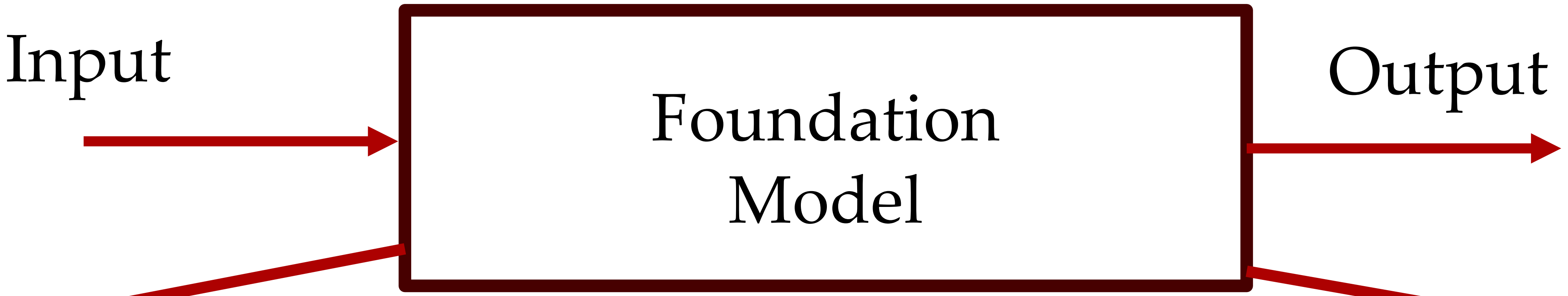
Tasks

- Recognition
- Detection
- Segmentation
- Captioning
- Image/text/synthesis
- Audio/Video/multimodal synthesis
- Question answering
- Text generation
- Autocompletion
- Translation
- Summarization
- Sentiment analysis
- Navigation
- Search/IR
- Recommendation
- ...

Application

- Customer service
- Social Media
- Marketing
- Code generation
- Automated reporting
- Domain-specific

Foundation Models

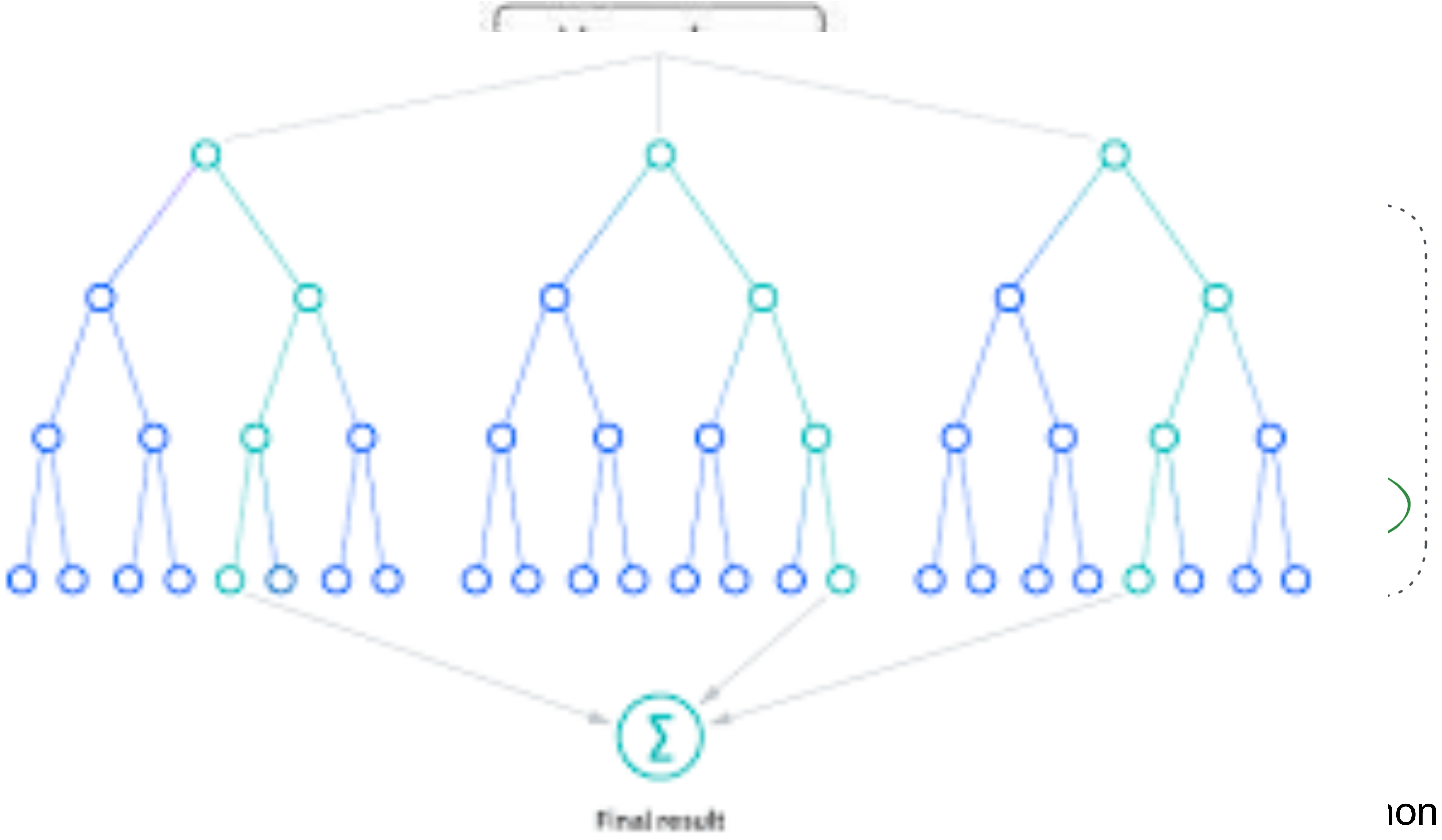


Input	Output	Function	Tasks	Application

• Trained on a variety of data input	• Different types of output based on downstream use	• Defined based on the scope of ability, range of uses, breadth of tasks or types of output and input
		• Encoder
		• Decoder
		• Encoder-Decoder
			• A single model can accomplish multiple tasks:
			• Sentence completion
			• Sentiment classification
			• Summarization
			• ….
			• Works for arbitrary input
			• Can be fine-tuned
				• Same model can serve multiple applications
				  • Customer service
				  • Code generation
				  • …

Are these Foundation Models?

- Decision Trees
- Clustering
- Support Vector Machines
- Random Forests



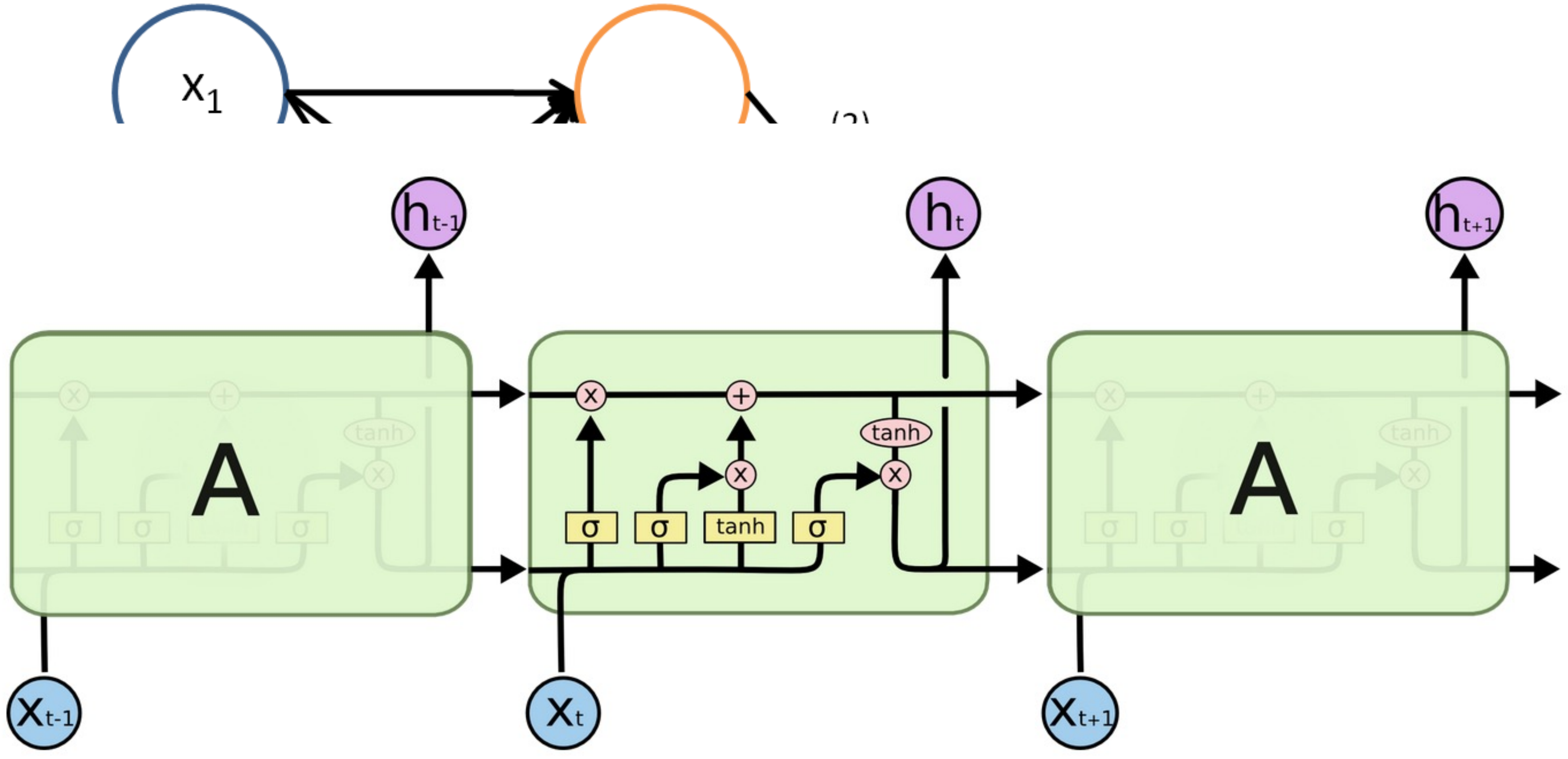
ion

https://www.sciencedirect.com/topics/computer-science/support-vector-machine

clustering

Are these Foundation Models?

- Multi-layer perceptron
- RNN
- LSTM



The repeating module in an LSTM contains four interacting layers.

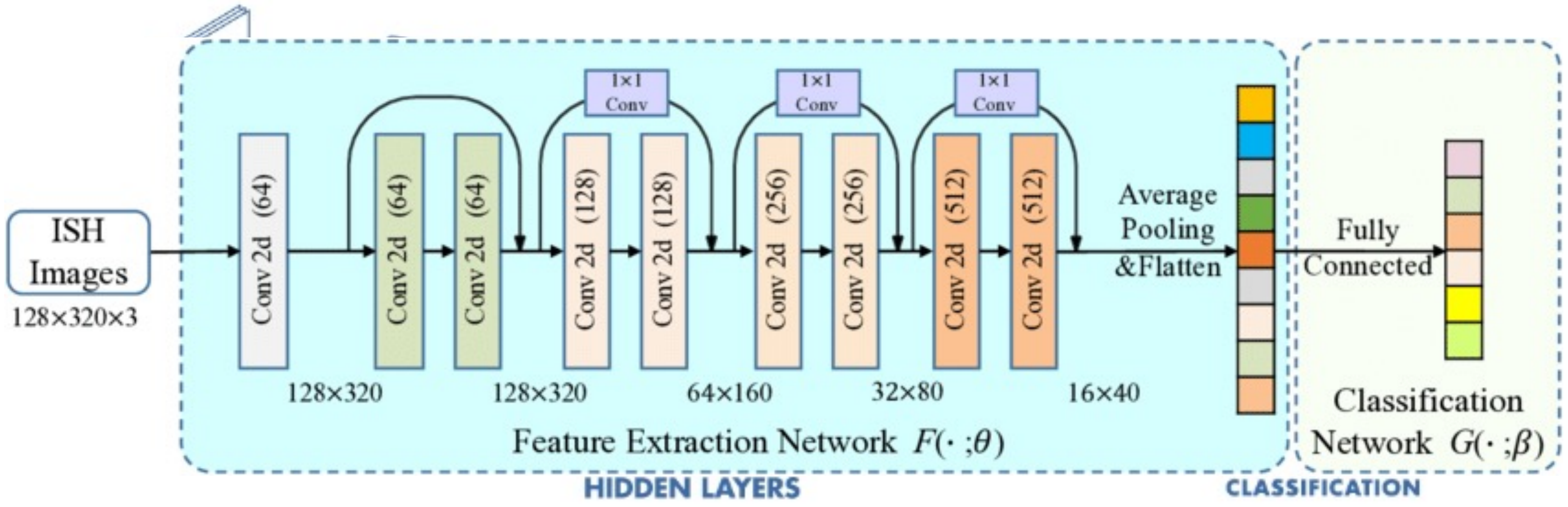
https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Layer L_1 Layer L_2

http://deeplearning.stanford.edu/tutorial/supervised/MultiLayerNeuralNetworks/

Are these Foundation Models?

- Convolutional neural networks
- ResNet
- DNN

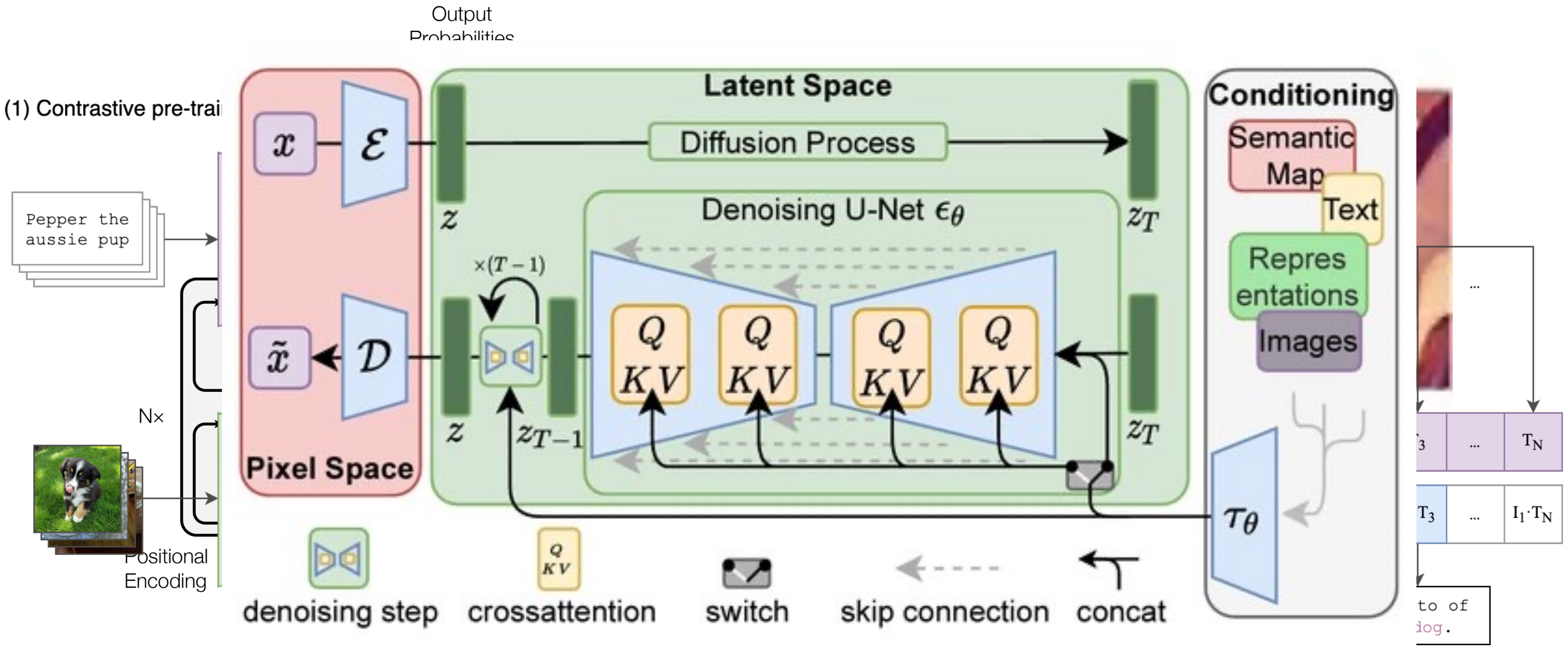


https://www.linkedin.com/pulse/deep-residual-networks-resnet-ayoub-kirouane/

https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939

Are these Foundation Models?

- Auto-encoders
- GANs
- Transformers
- GPT
- CLIP
- Diffusion models



https://openaccess.thecvf.com/content/CVPR2022/papers/Rombach_High-Resolution_Image_Synthesis_With_Latent_Diffusion_Models_CVPR_2022_paper.pdf

https://github.com/openai/CLIP

https://arxiv.org/abs/1706.03762

Foundation vs Traditional ML models

Foundation	Traditional ML
• Trained on broad unlabeled data for a variety of tasks	• Trained on specific datasets and designed for particular tasks
• Very large number of parameters	• Fewer parameters compared to foundation models.
• Fine-tuning for adaptation	• Label supervision
• Self-supervision for labeling	• Easy to implement, interpret, and computationally efficient.
• Reduced development time, less ongoing maintenance	• Risk of overfitting or underfitting
• Dataset bias	• Limited generalization
• lack of domain specificity, potential for misuse	

Which model is more suitable?

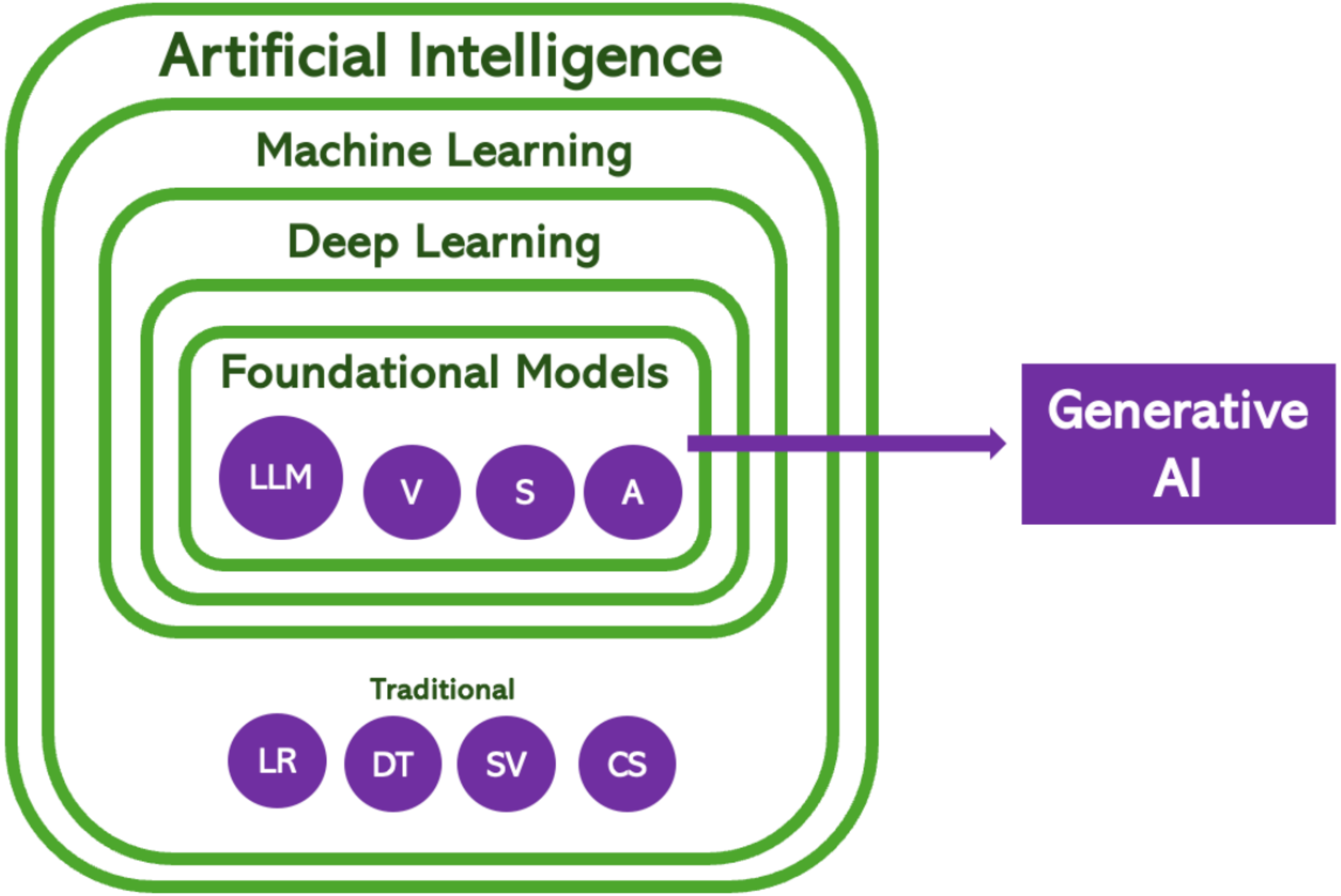
- Recognition
- Detection
- Segmentation
- Synthesis
- Summarization
- Search
- Sentiment analysis

- Traditional ML
- Deep Learning
- Foundational model
- LLM
- Generative AI

Case study

- In a large hospital system, the new CIO has been tasked to improve revenue cycle management by identifying cases of additional billing:
  - Which ML model to use?
  - Is a model alone sufficient?

Evolution of ML models



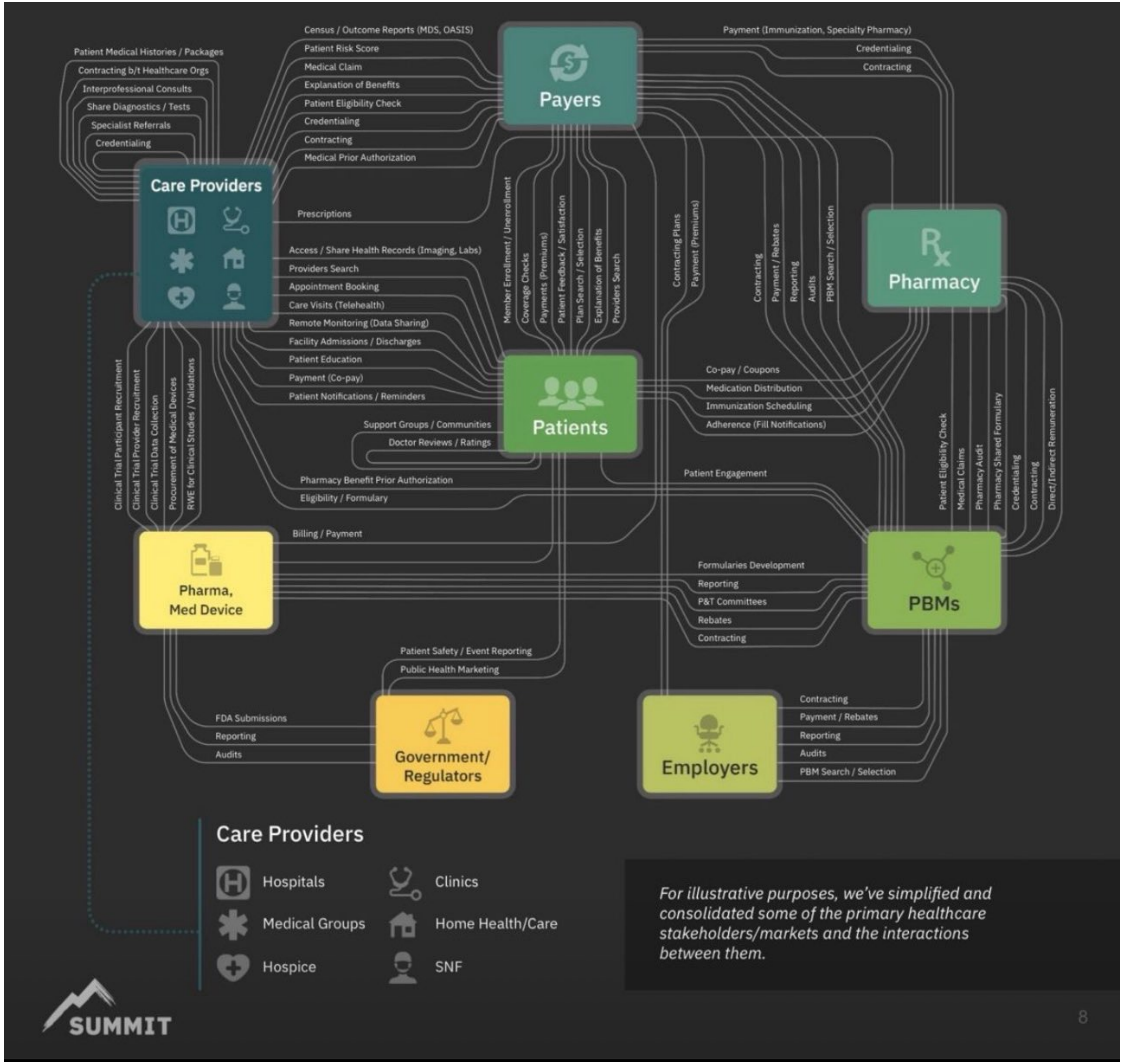
https://www.linkedin.com/pulse/ai-alphabet-soup-from-buzzwords-brilliance-brian-gruttadauria-fxtie/

Benefits of Modern Day FMs

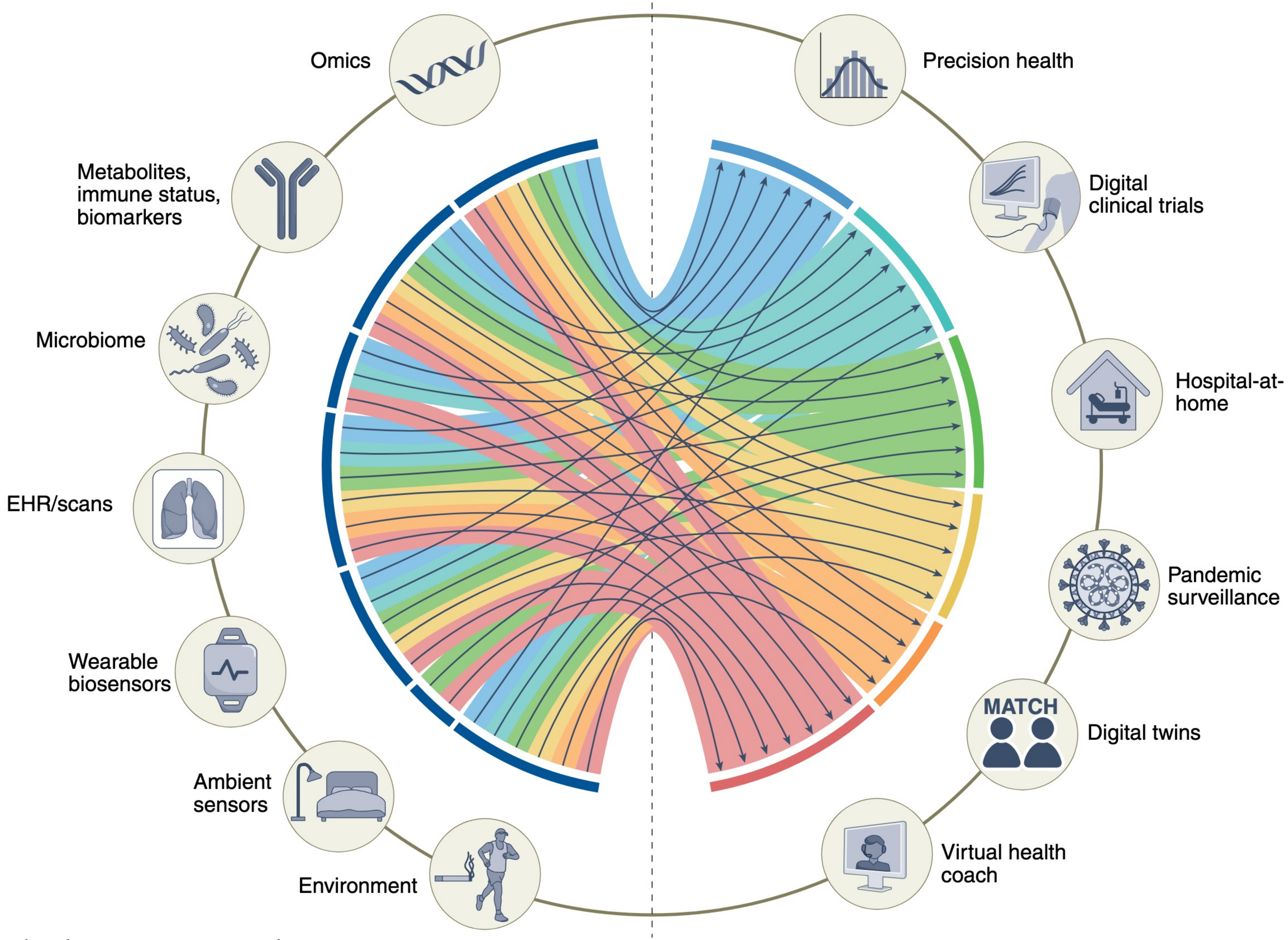
- Can perform generative and discriminative tasks
- Can be conditioned in a variety of ways
- Can be aligned with human preferences
- Can be (easily) adapted

Opportunities in Healthcare

- Vast interconnected set of stakeholders
- Disparate but complementary needs/challenges

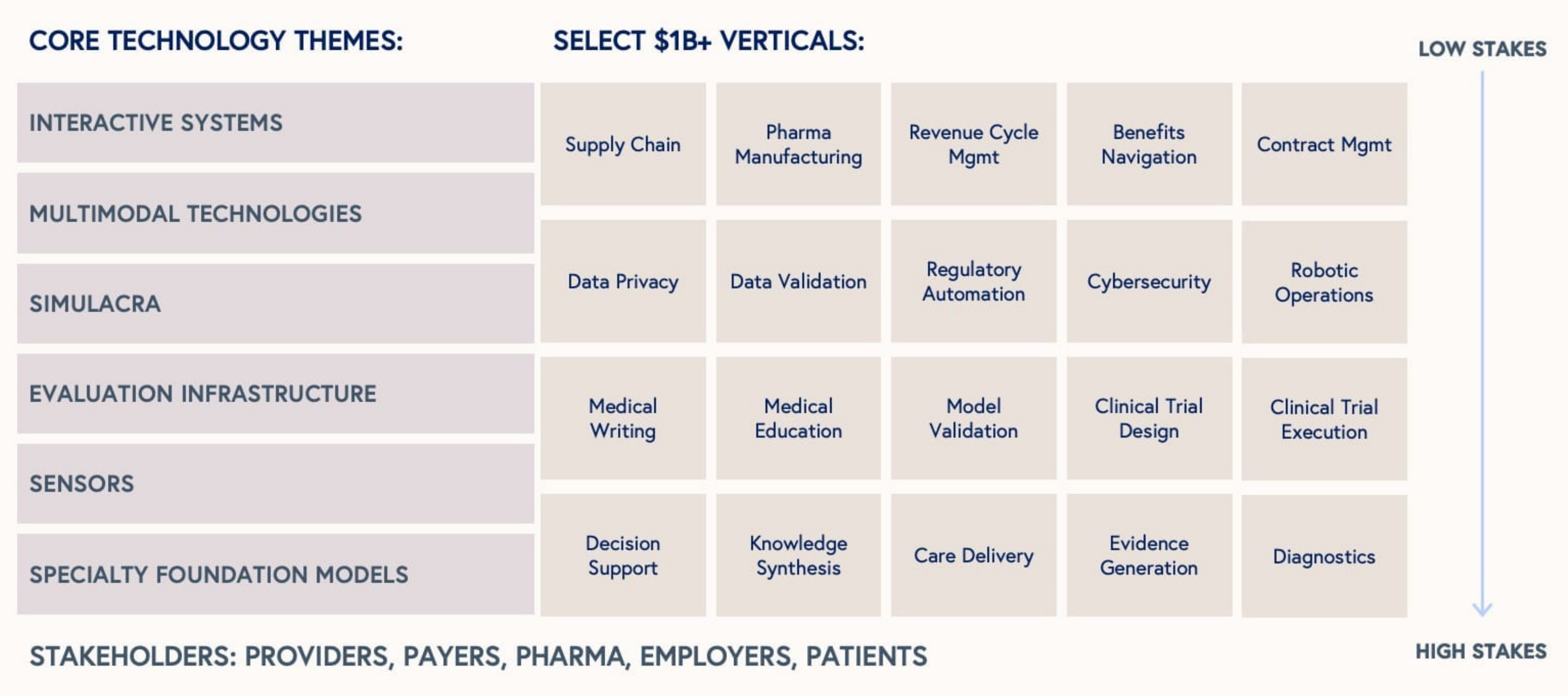


Opportunities in Healthcare



Acosta et al. "Multimodal biomedical AI". Nature Medicine 2022.

Opportunities in Healthcare



Cheatham M. "Bessemer Healthcare AI Areas of Focus" from Healthcare AI Thesis Update. 2024.

Challenges of Modern Day FMs

- Data requirements (pre- and post-training)
- Computational demands
- Inference efficiency/costs
- Adequate evaluation

Challenges of Modern Day FMs

🏆 Chatbot Arena LLM Leaderboard: Community-driven Evaluation for Best LLM and AI chatbots

💬 Language | Overview | Price Analysis | WebDev Arena | 👀 Vision | Text-to-Image | Copilot Arena | Arena-Hard-Auto

Total #models: **220**. Total #votes: **2,816,680**. Last updated: 2025-03-25.

Code to recreate leaderboard tables and plots in this notebook. You can contribute your vote at lmarena.ai!

Category: Overall

Apply filter: Style Control | Show Deprecated

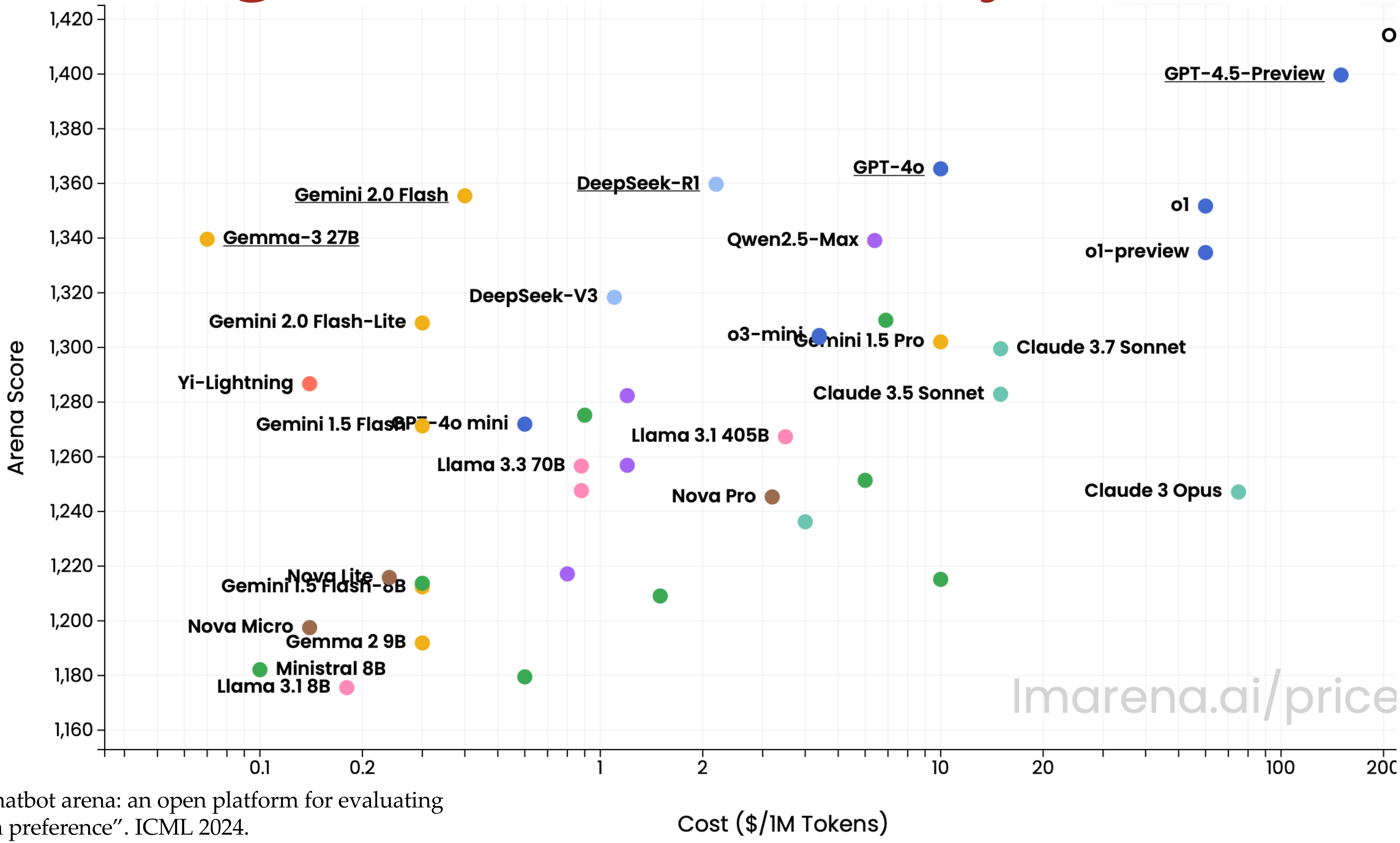
Overall Questions

#models: 220 (100%) **#votes: 2,816,680 (100%)**

Rank* (UB)	Rank (StyleCtrl)	Model	Arena Score	95% CI	Votes	Organization	License
1	1	Gemini-2.5-Pro-Exp-03-25	1443	+11/-8	3474	Google	Proprietary
2	2	ChatGPT-4o-latest (2025-03-26)	1408	+11/-12	2676	OpenAI	Proprietary
2	4	Grok-3-Preview-02-24	1404	+6/-6	10397	xAI	Proprietary
2	2	GPT-4.5-Preview	1398	+6/-7	10907	OpenAI	Proprietary
5	7	Gemini-2.0-Flash-Thinking-Exp-01-21	1381	+4/-5	22987	Google	Proprietary

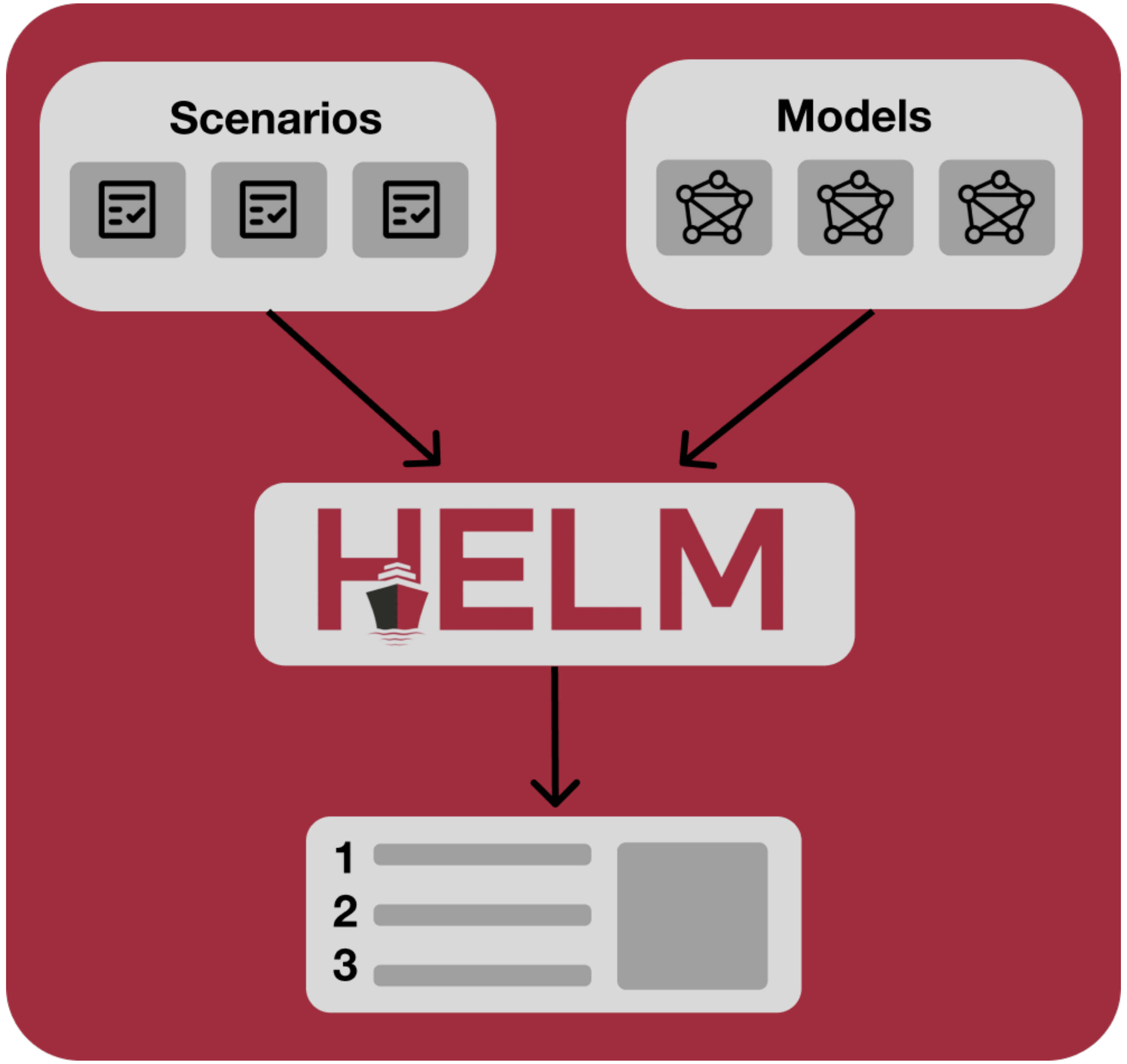
Chiang et al. "Chatbot arena: an open platform for evaluating LLMs by human preference". ICML 2024.

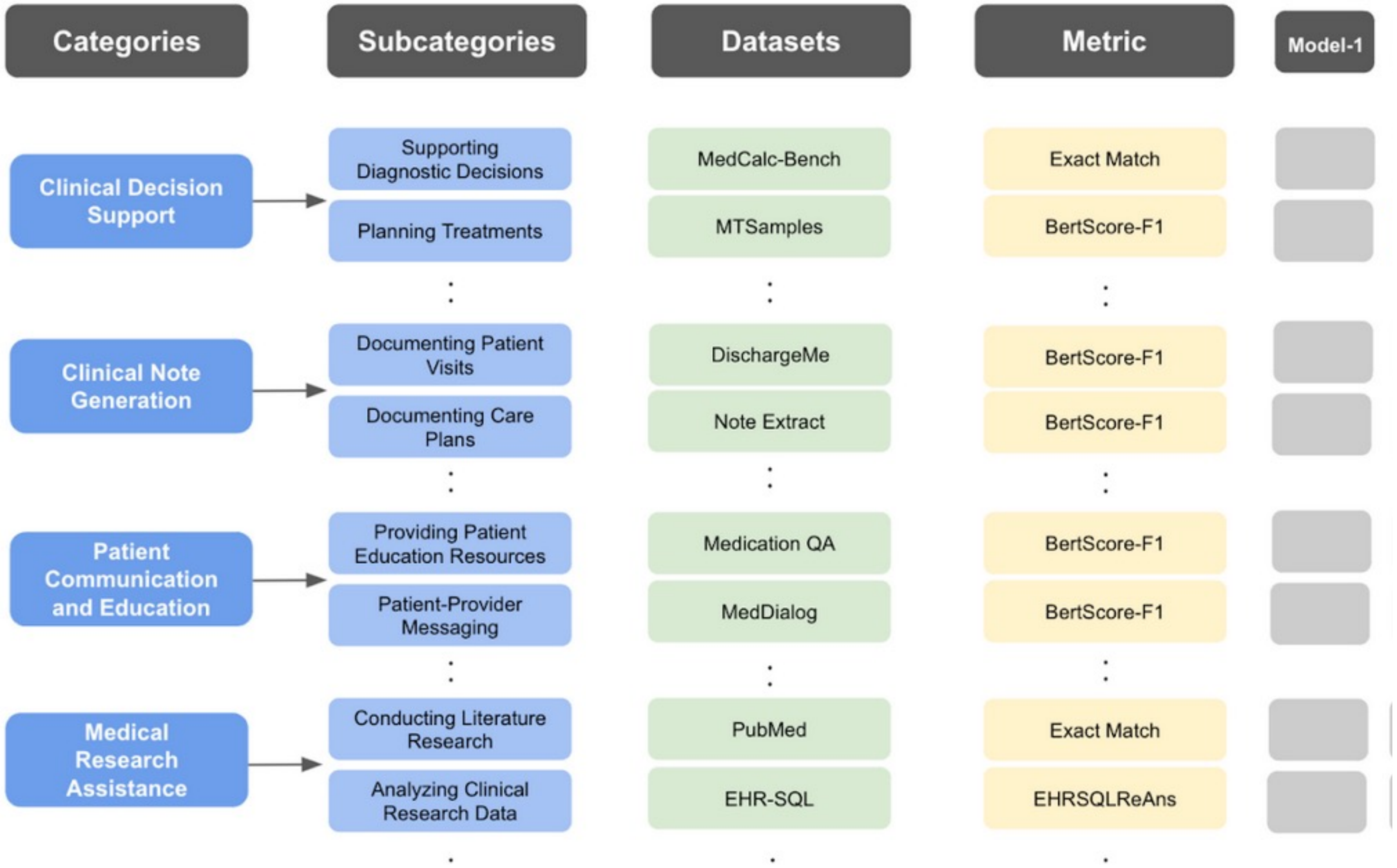
Challenges of Modern Day FMs



Chiang et al. "Chatbot arena: an open platform for evaluating LLMs by human preference". ICML 2024.

Challenges of Modern Day FMs





MedHELM