



# The Transformative Power of AI and Open Challenges

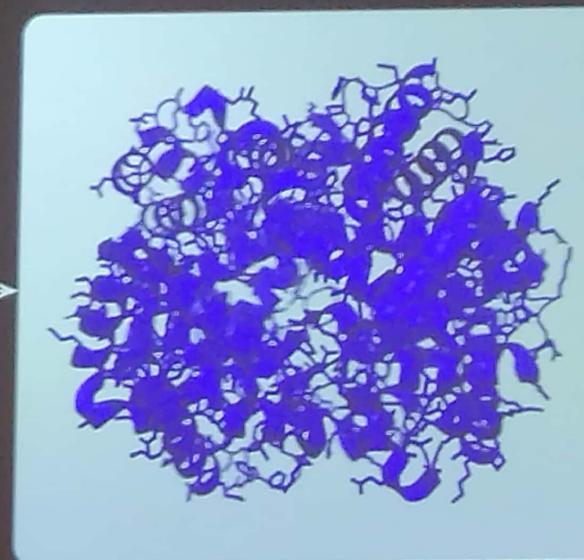
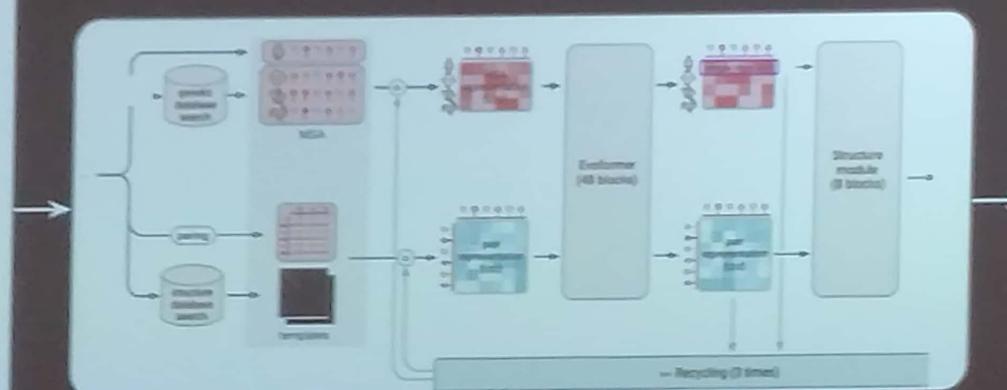
Manish Gupta  
Google DeepMind

Acknowledgement: Work done mostly by my colleagues at Google DeepMind

## Understanding Proteins

卷之三

AlphaFold Neural Network



## Nobel Prize in Chemistry, 2024

# AlphaFold is accelerating progress on a range of important problems

See more AlphaFold impact stories at [deepmind.google/impact/](https://deepmind.google/impact/)



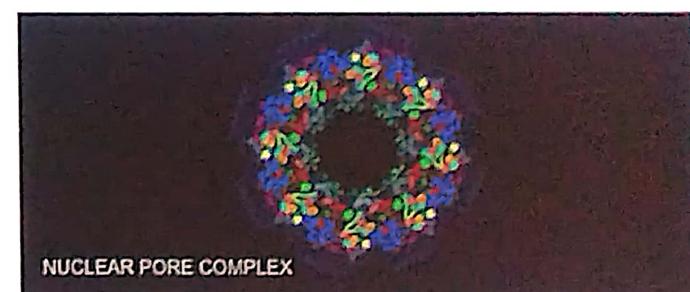
## Plastic pollution

Designing plastic eating enzymes  
McGeehan et al. (Portsmouth)



## Antibiotic resistance

Blocking antibiotic resistance mechanisms  
Sousa & Mitchell (Colorado)



## Structural biology

Determined structure of nuclear pore complex  
Fontana et al., Mosalaganti, et al. (Science)



## Neglected diseases

Accelerating drug discovery in Leishmaniasis  
Benjamin Perry (DNDi)



## Malaria vaccine

Designing a more effective malaria vaccine  
Matthew Higgins (Oxford)



## Drug delivery

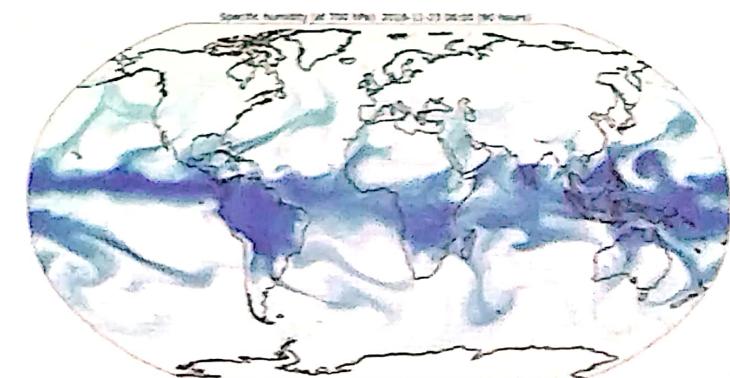
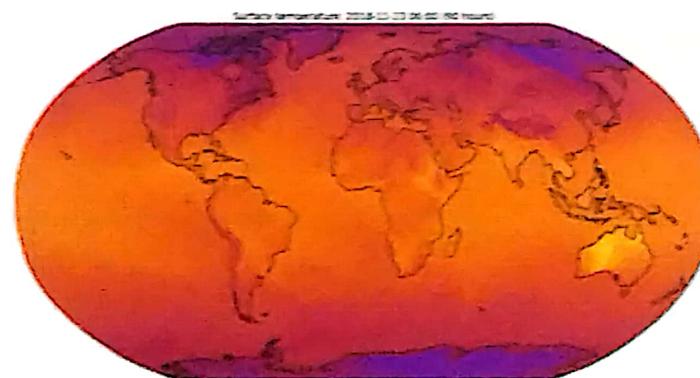
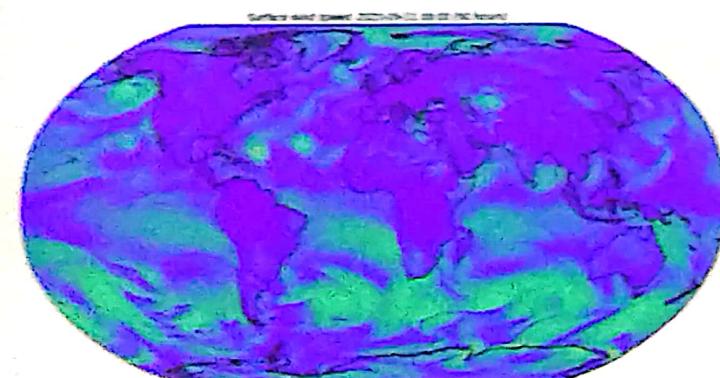
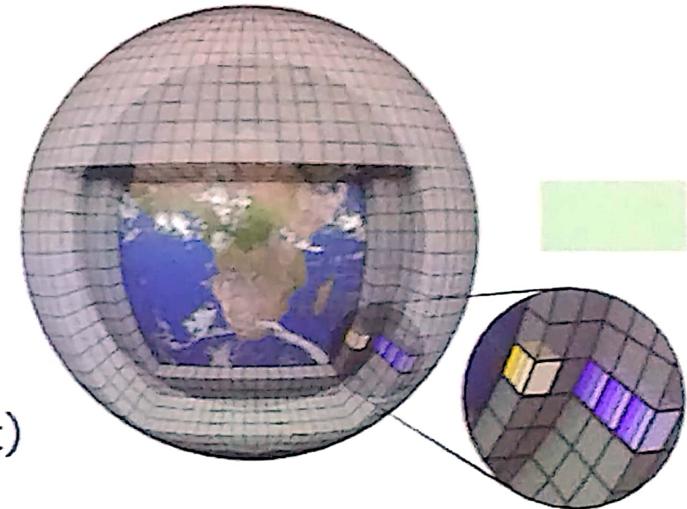
Molecular protein syringe for payload delivery  
Feng Zhang (Broad Institute)



Scanned with OKEN Scanner

# What is medium-range global weather forecasting?

- Predicting the state of the atmosphere
  - Up to 15 days
  - Everywhere on earth (in 3D)
  - For multiple variables
- Traditional approach uses physics-based solvers:
  - Slow (hours on supercomputer for a 15 days forecast)
  - Require more compute to improve in quality



3 of hundreds of weather variables modeled

# Outline

- AI Trends
  - Deep Learning (dominant paradigm within ML)
  - Foundation Models (new paradigm within DL)
- AI for Synthesis and Creativity
- Inclusive AI
- Some Challenges and Takeaways

# Deep Learning

## Modern Reincarnation of Artificial Neural Networks

Collection of simple trainable mathematical units, organized in layers, that work together to solve complicated tasks

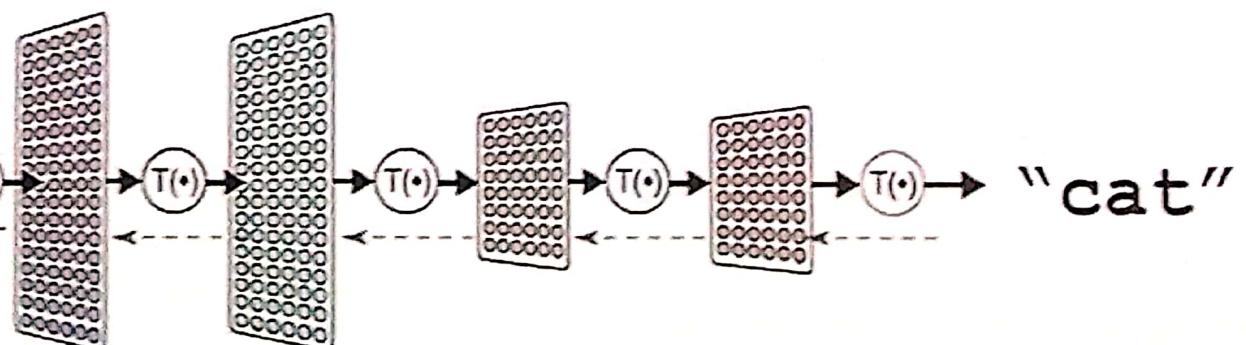
## What's New

new network architectures,  
new training math, scale



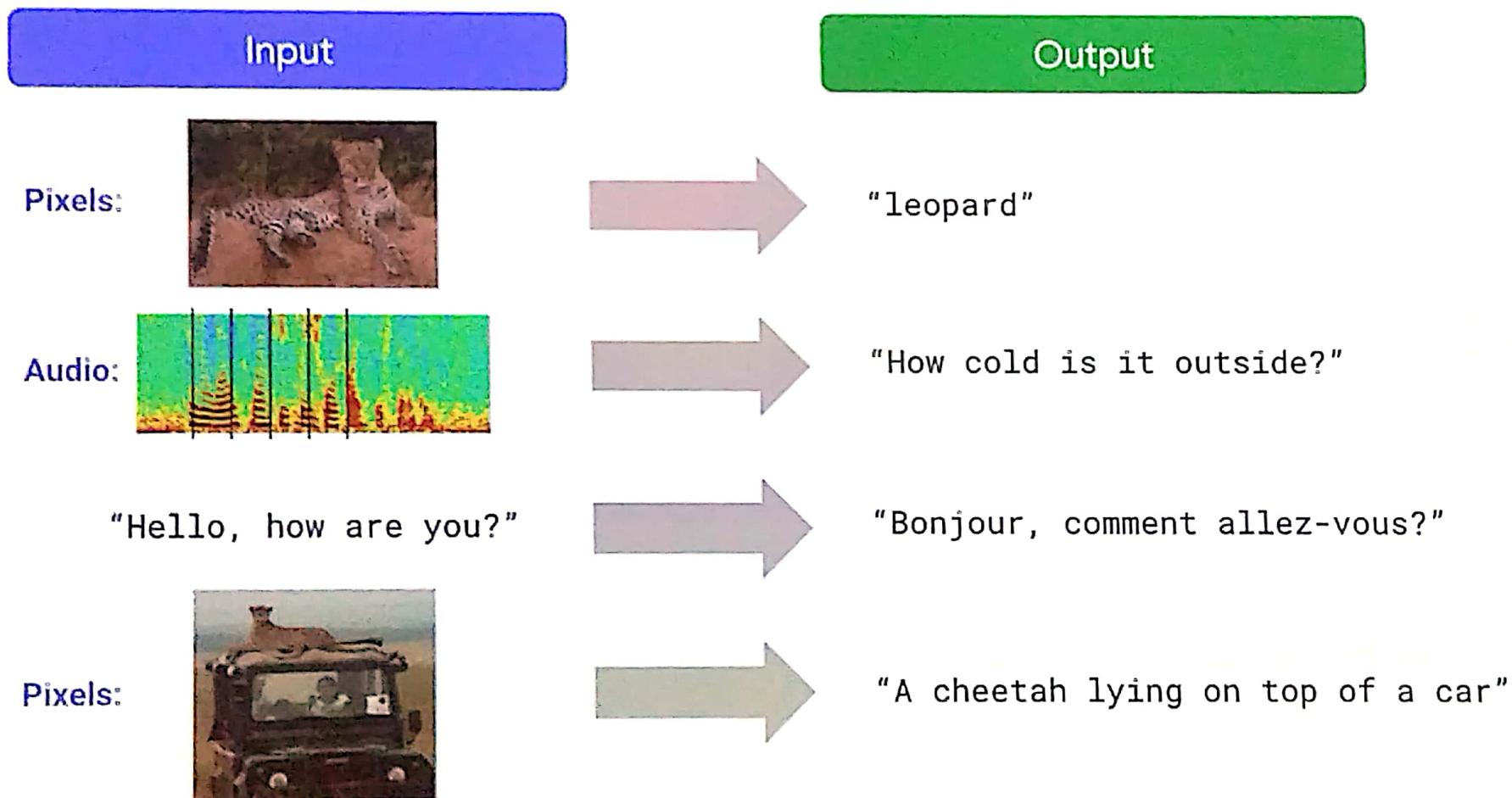
## Key Benefit

Learns features from raw, heterogeneous, noisy data  
No explicit feature engineering required



Google

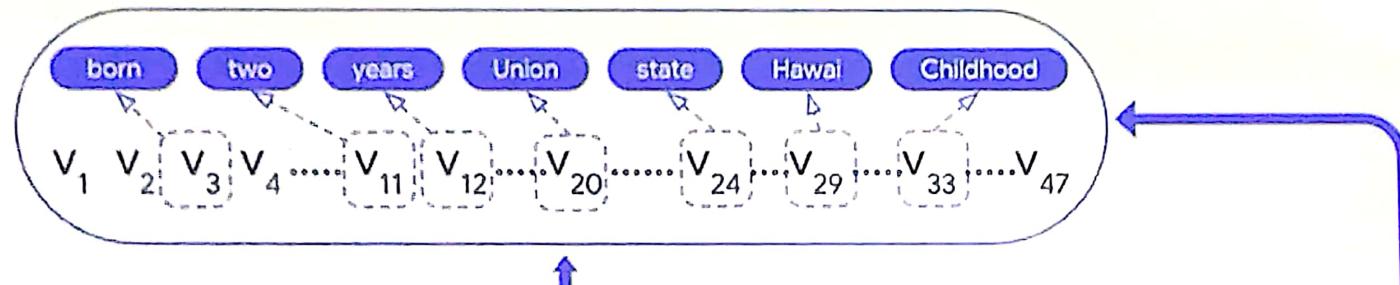
# Deep Learning Phase 1: A decade of amazing progress in what computers can do



Google

## Large Models using Self-Supervision and Transformer Architecture: BERT (2018)

Predict  
masked words



Encode

 BERT (Google AI's deep learning language model)

Masked  
words

Obama was \_\_\_\_ in 1961 in Honolulu , Hawaii , \_\_\_\_ after the territory was admitted to the \_\_\_\_ as the 50th \_\_\_\_ . Raised largely in \_\_\_\_ , he also spent one year of his \_\_\_\_\_ in Washington state and four years in Indonesia.

Original  
words

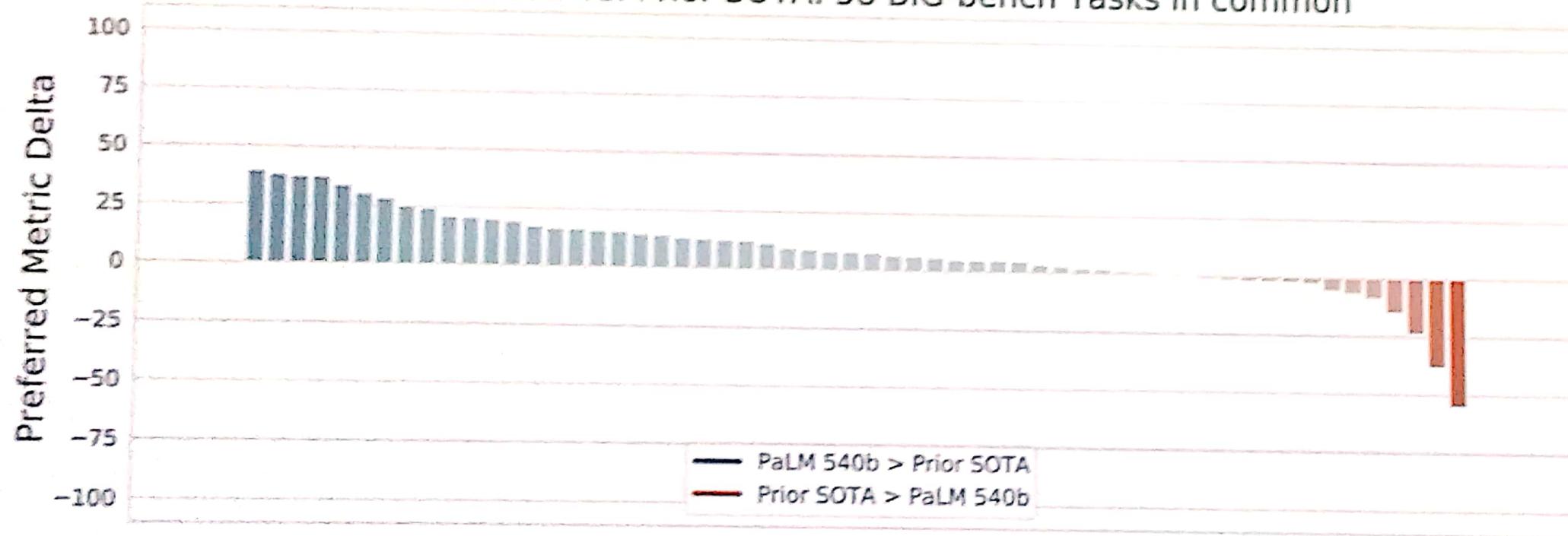
Obama was born in 1961 in Honolulu , Hawaii , two years after the territory was admitted to the Union as the 50th state . Raised largely in Hawaii , he also spent one year of his childhood in Washington state and four years in Indonesia.

Masks  
(3, 11, 12, 20, 24, 29, 33)

# Large language models: PaLM

Google Research

PaLM 540B vs. Prior SOTA: 58 BIG-bench Tasks in common



*performance comparison between the PaLM 540B parameter model and the prior state-of-the-art (SOTA) on 58 tasks from the Big-bench suite. (See [paper](#) for details.)*

# Chain of Thought

## Standard prompting

Input:

Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

A: The answer is 9.

...

Q: John takes care of 10 dogs. Each dog takes .5 hours a day to walk and take care of their business. How many hours a week does he spend taking care of dogs?

A:

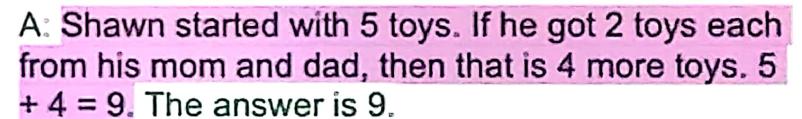
Model output:

The answer is 50. 

## Chain of thought prompting

Input:

Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

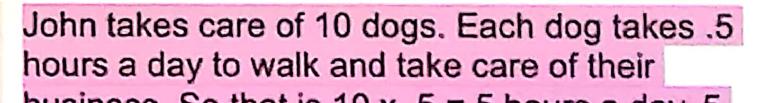
A: Shawn started with 5 toys. If he got 2 toys each from his mom and dad, then that is 4 more toys.  $5 + 4 = 9$ . The answer is 9.

...

Q: John takes care of 10 dogs. Each dog takes .5 hours a day to walk and take care of their business. How many hours a week does he spend taking care of dogs?

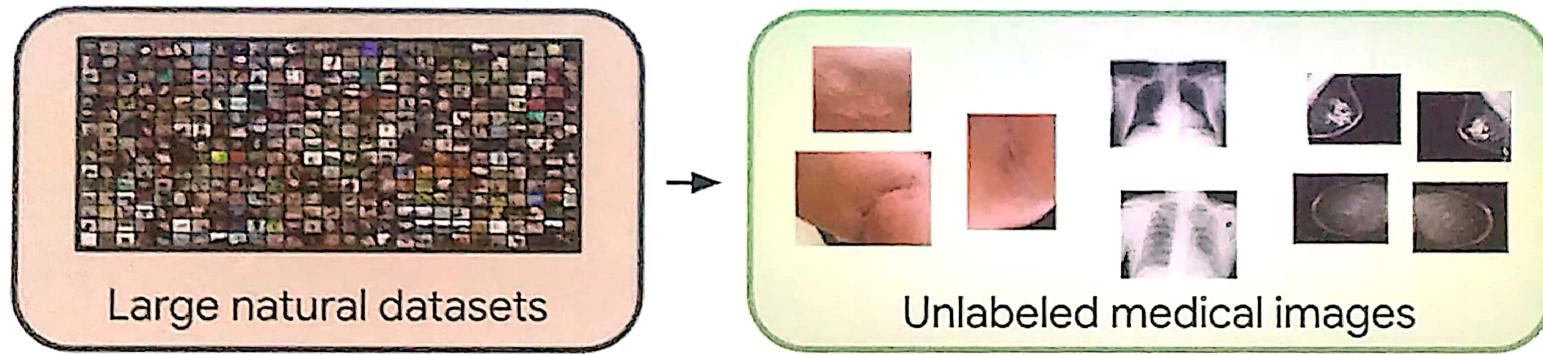
A:

Model output:

John takes care of 10 dogs. Each dog takes .5 hours a day to walk and take care of their business. So that is  $10 \times .5 = 5$  hours a day.  $5$  hours a day  $\times$  7 days a week = 35 hours a week.  
The answer is 35 hours a week. 

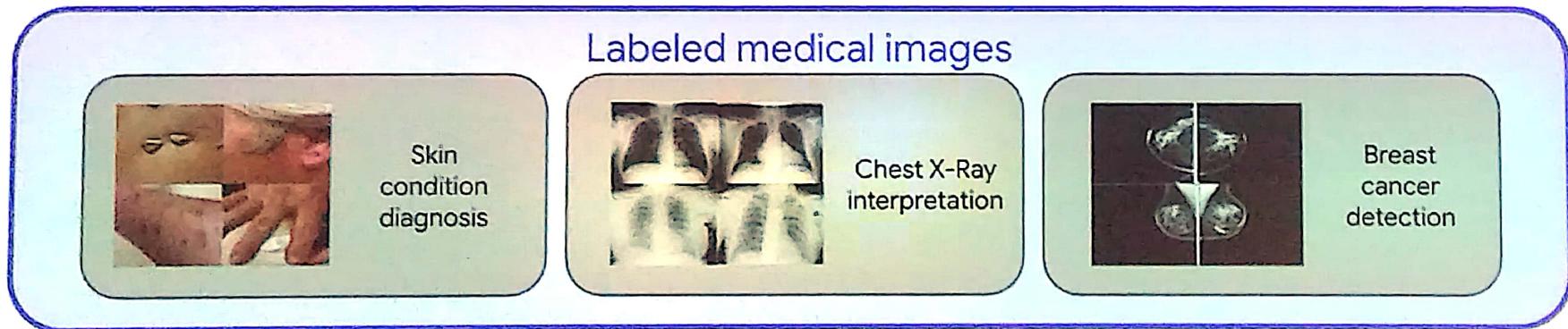
*Chain of Thought Prompting Elicits Reasoning in Large Language Models*, Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Ed Chi, Quoc Le, and Denny Zhou, <https://arxiv.org/abs/2201.11903>

# How can we utilise unlabelled images?

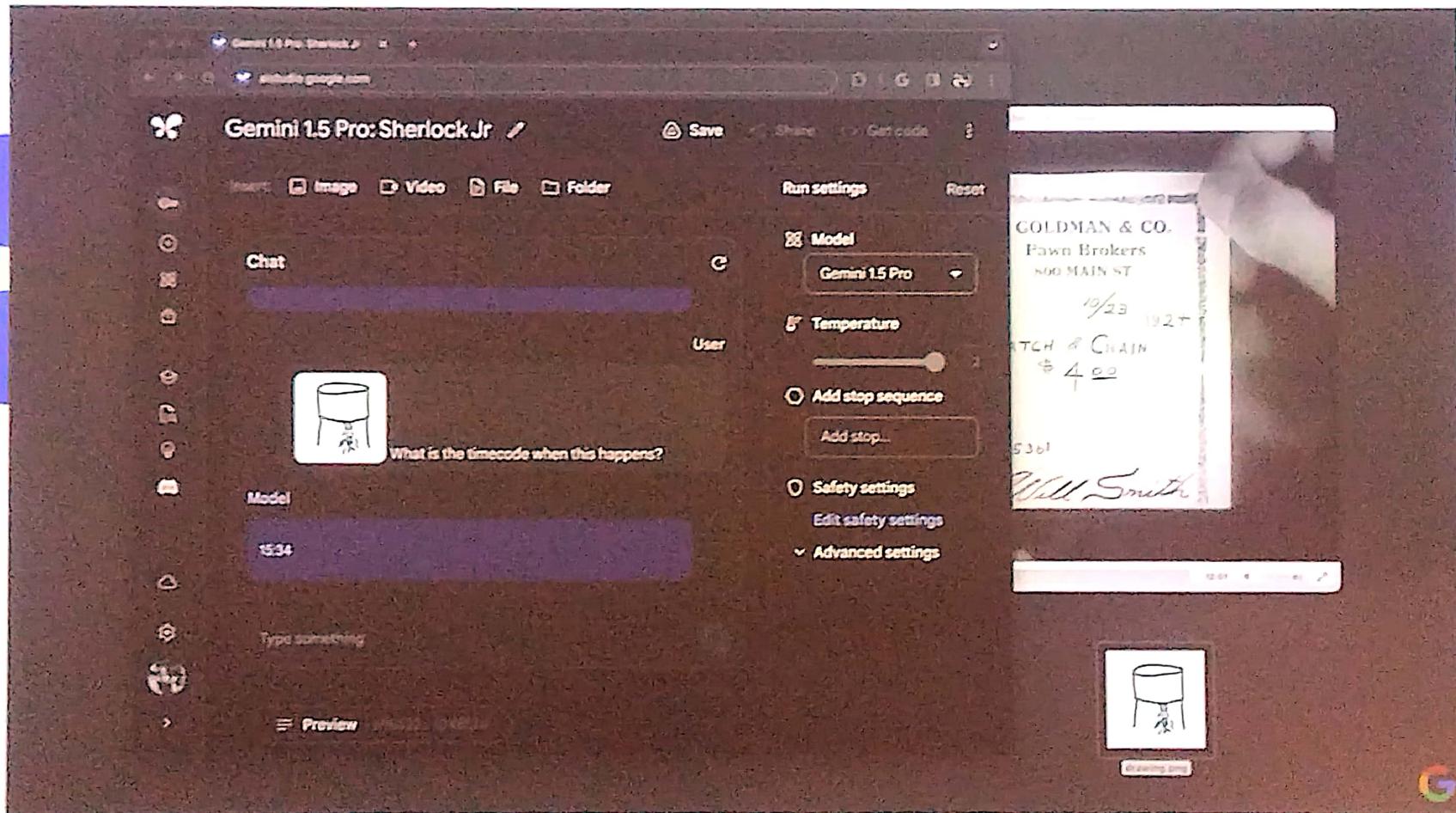


Big Transfer (BiT)

Self and semi supervised learning (SSL)



# Gemini 1.5: Long Context Window in Action

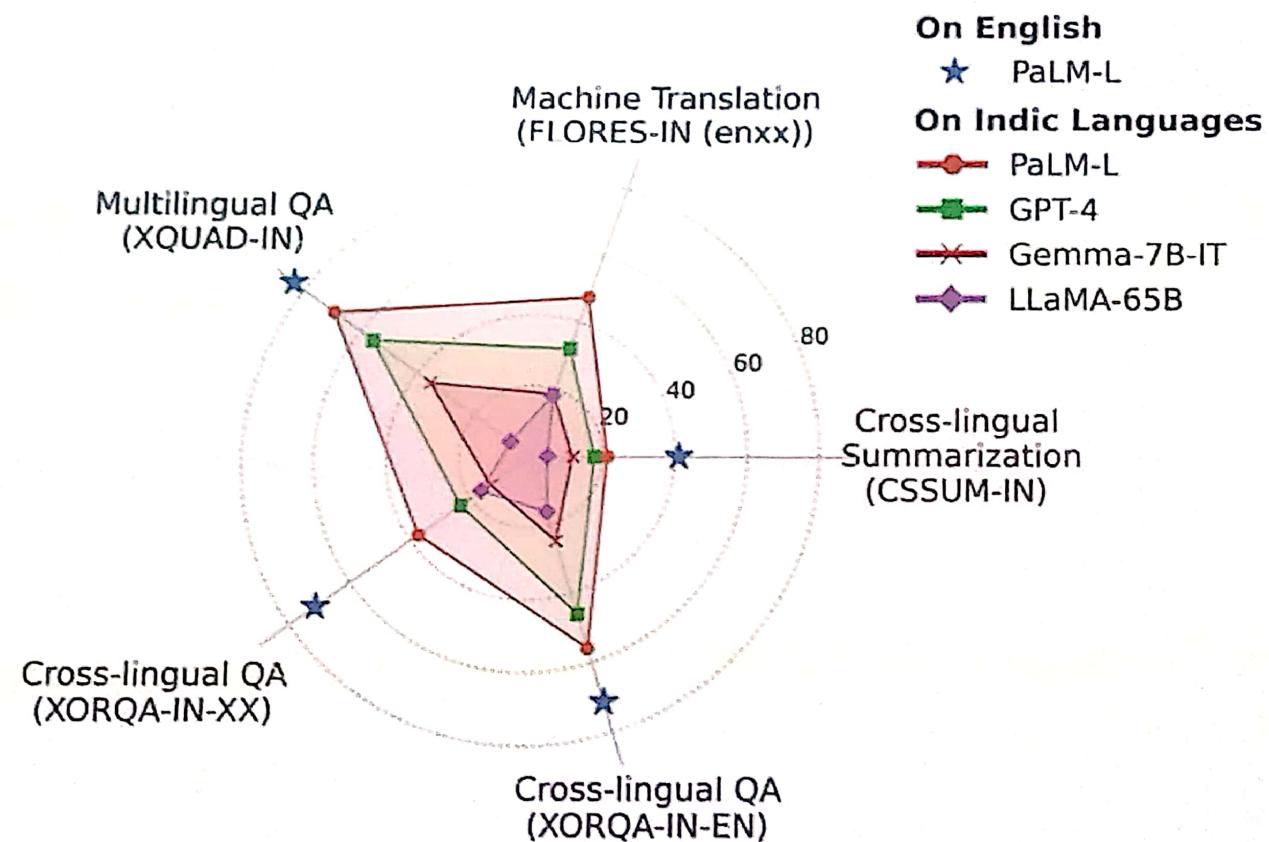


Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context, [arxiv.org/abs/2403.05530](https://arxiv.org/abs/2403.05530)

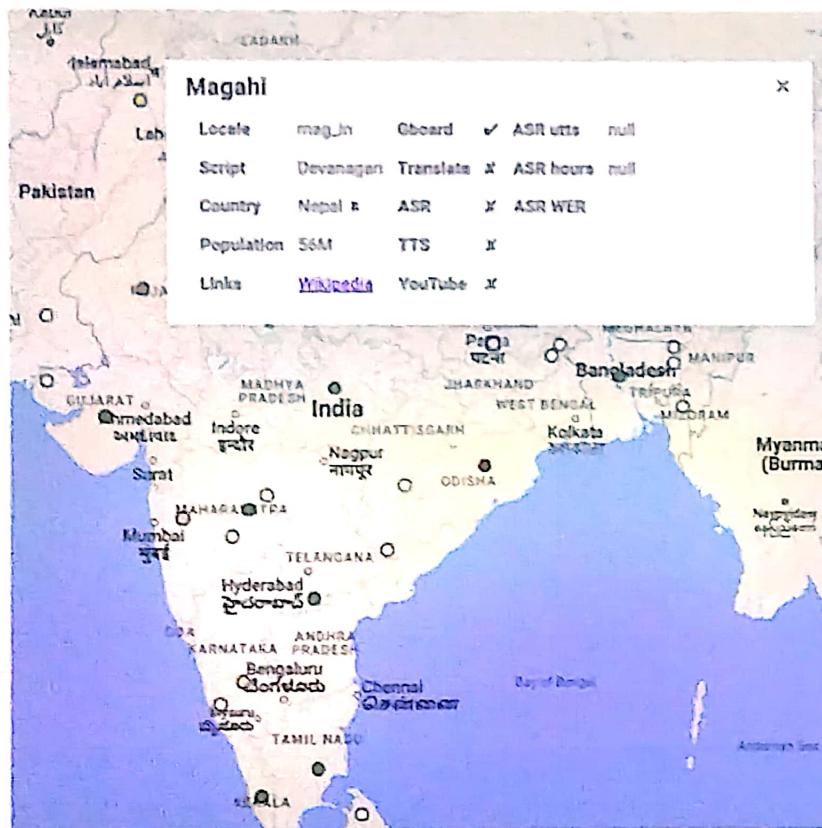
# IndicGenBench (ACL 2024)

## Performance gap in Indic languages vs. English

All Open-source and Proprietary LLMs perform much better in English as compared to Indic languages



# Morni: Multimodal Representation for India



Population: 1.4B  
Languages: 60 (1M+), 125 (100K+)  
Google ASR: 11 languages

## Morni Vision

Build the best multimodal representation for  
100+ Indic languages and accelerate  
development of *inclusive* and *equitable* Indic  
language technologies

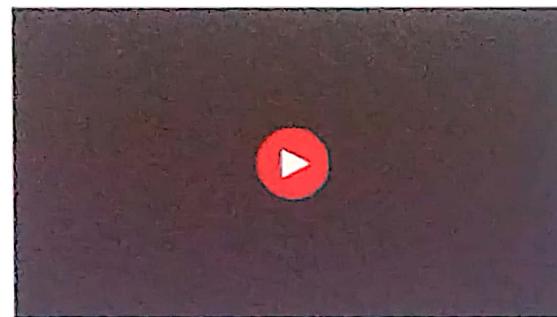
Google



Scanned with OKEN Scanner

# Vaani: Capturing the Speech Landscape in India

A given language varies across region, collecting data anchored on language doesn't capture on the ground diversity



[vaani.iisc.ac.in](http://vaani.iisc.ac.in)



## Vaani approach

Collect image-prompted speech data anchored on region, rather than language, while ensuring diversity

773  
districts

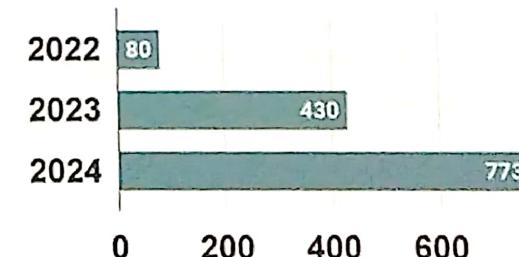
154K  
speech

15.4K  
transcribed

Open  
Source

PoC: partha@, dineshtewari@

## Districts Coverage Plan



Google



Scanned with OKEN Scanner

# Vaani Phase 1 & 2 (ongoing): Dataset Released

**143 Districts**

# of Districts from  
which Speech Data is  
being Collected from  
over 140K speakers

**>27k hr Audio**

Quality controlled  
Image prompted  
Speech Data has  
been OS

**>1150 hr**

**Transcription**

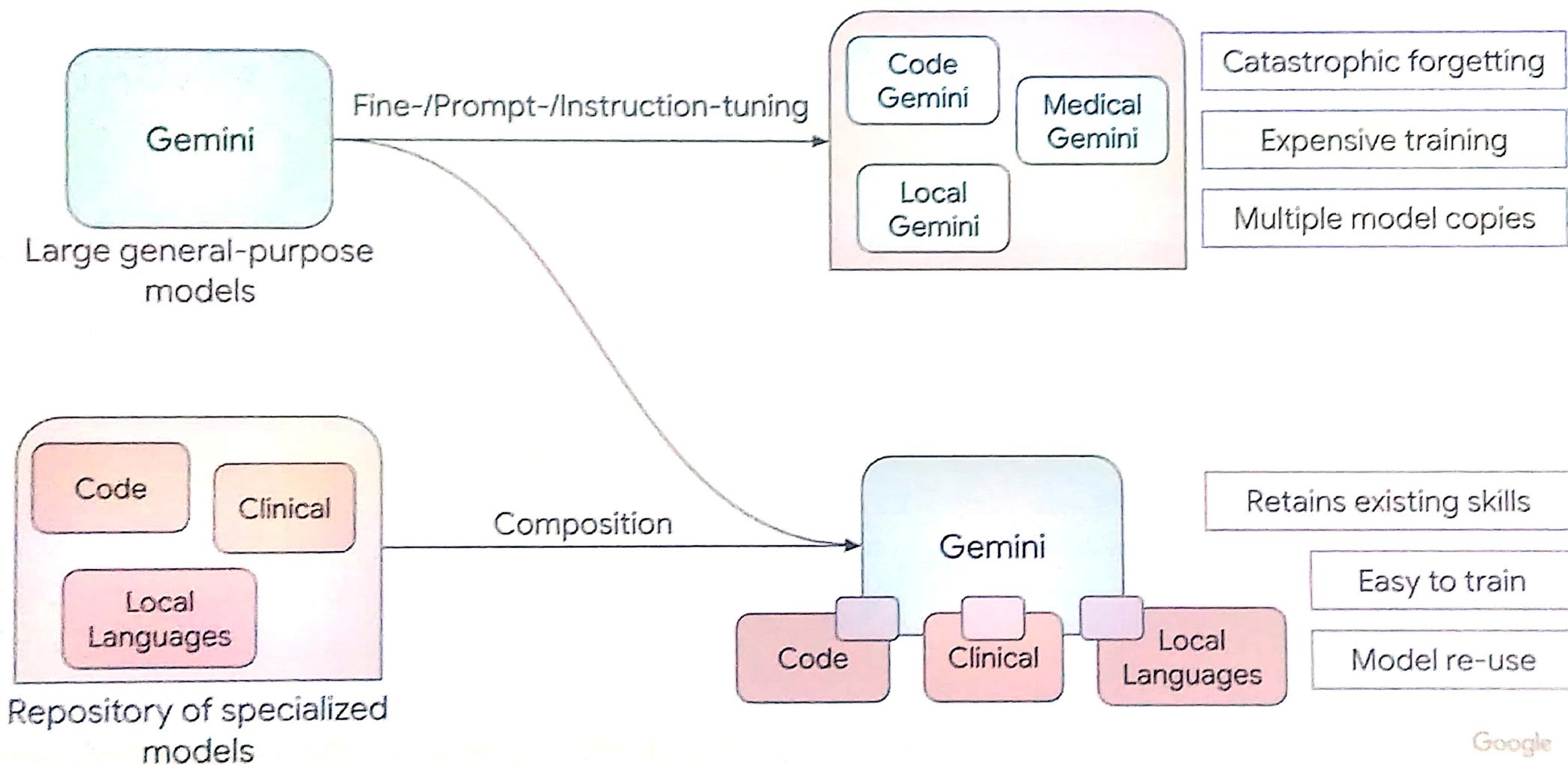
of Transcribed Indic  
Language data

**103 Languages**

# of Languages for  
which speech data  
has been collected

First recorded digital data for 22 of those languages, first speech data for 35 of them.

# CALM: Composition of Language Models



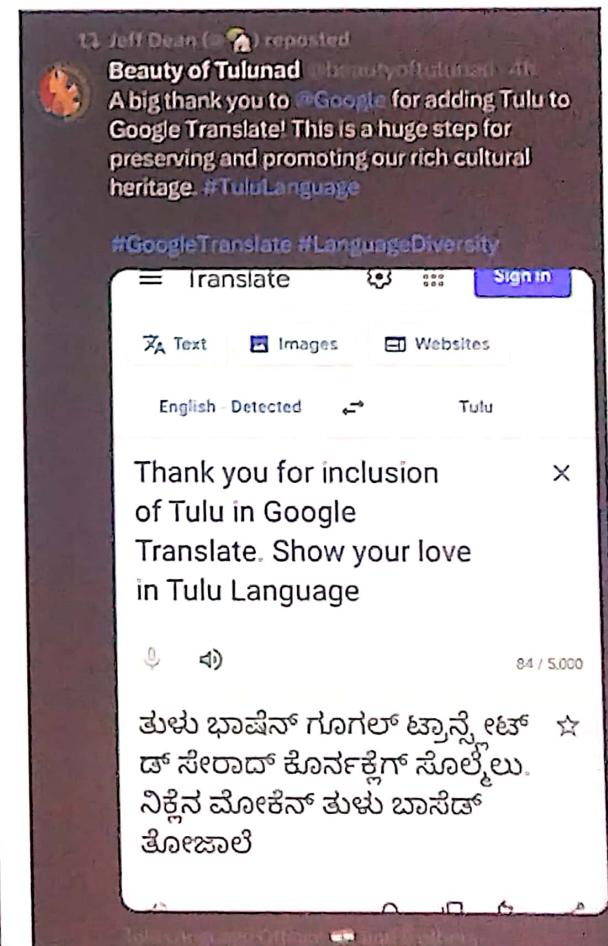
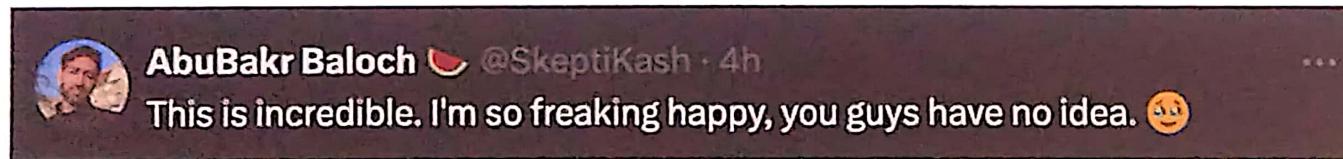
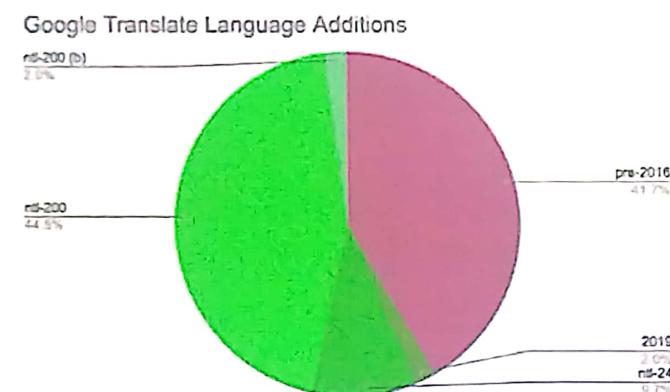
Expanded Translate coverage to 614M more people,  
enabled by PaLM2

110+ languages worldwide

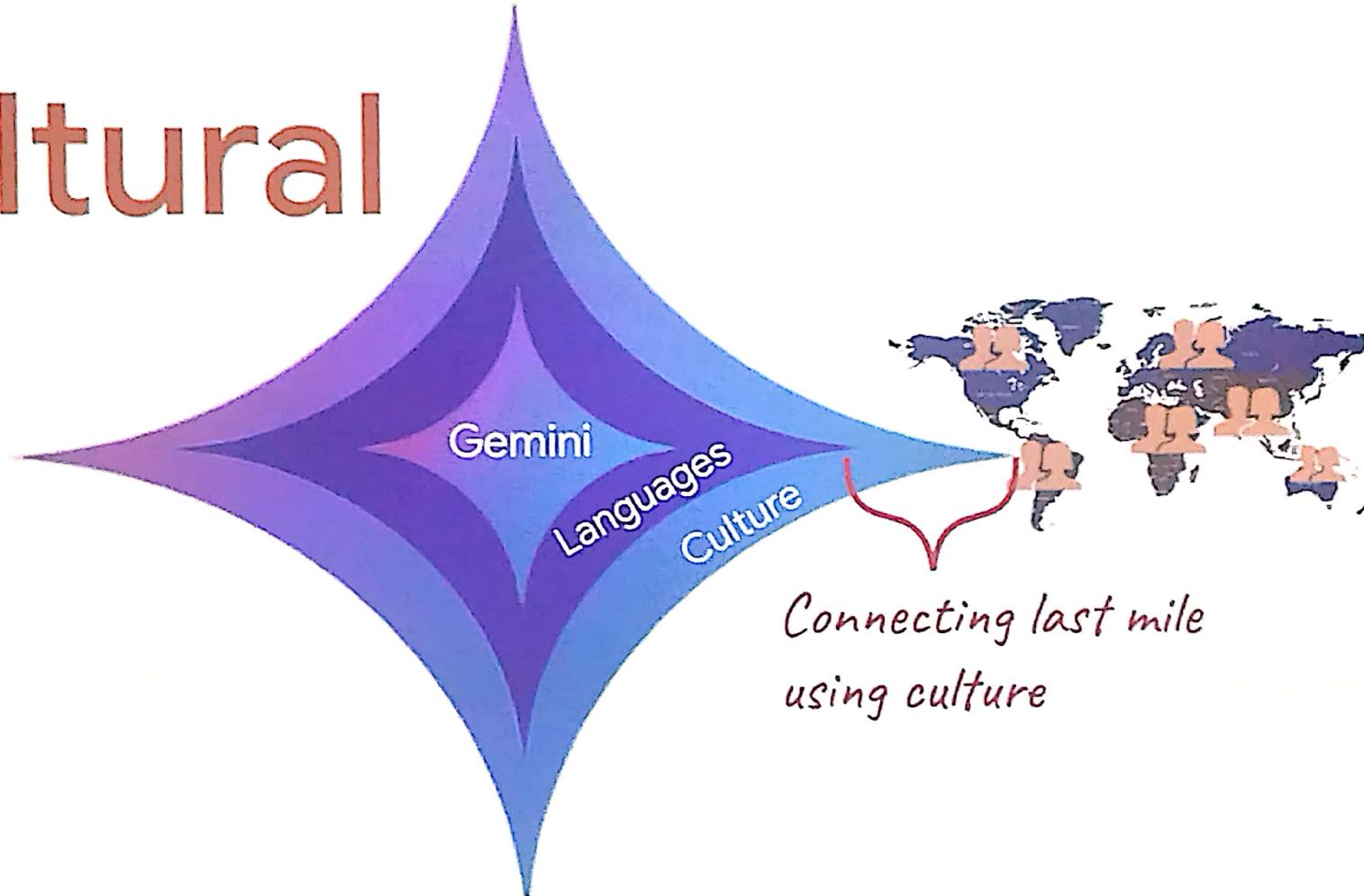
Largest expansion ever

7 new Indics covered (total 28)

Awadhi, Marwadi, Santali, Tulu, Khasi, Kokborok



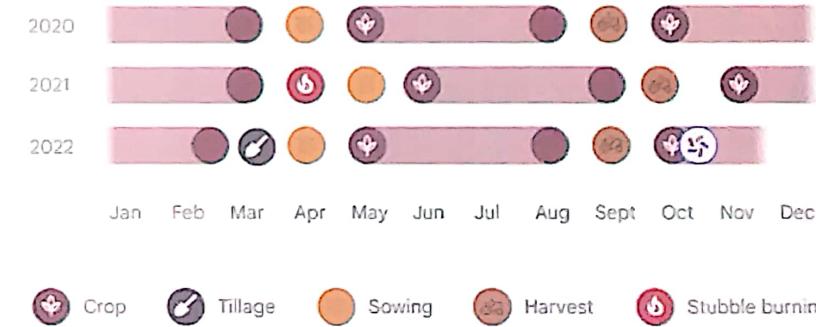
# Multicultural Models



# Agricultural Landscape Understanding & Field Event Detection (from Satellite Imagery)



([demo](#))



Foundational layer of an agri-stack - use cases like farmer loans, crop insurance, incentives to avoid stubble burning, carbon modeling to incentivize regenerative agricultural practices

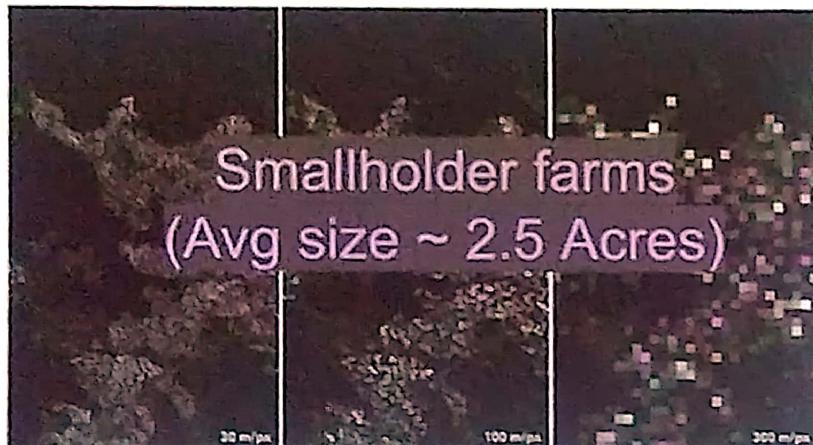
Google



Scanned with OKEN Scanner

# GeoFMs & Types of Earth Observation resolutions

Spatial

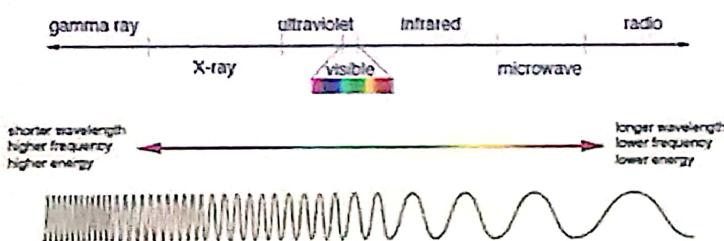


Temporal



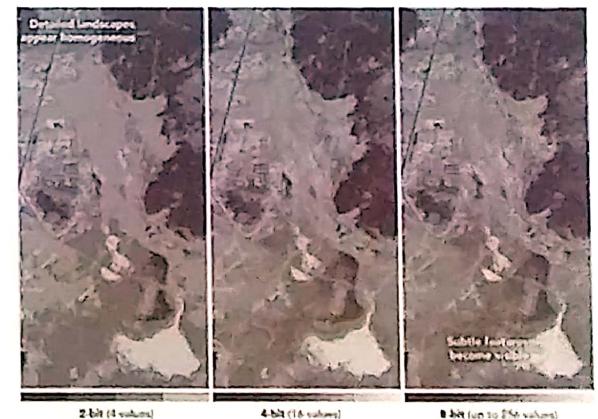
Understanding & interpreting data across all of these resolutions is the core challenge for GeoFMs

Spectral



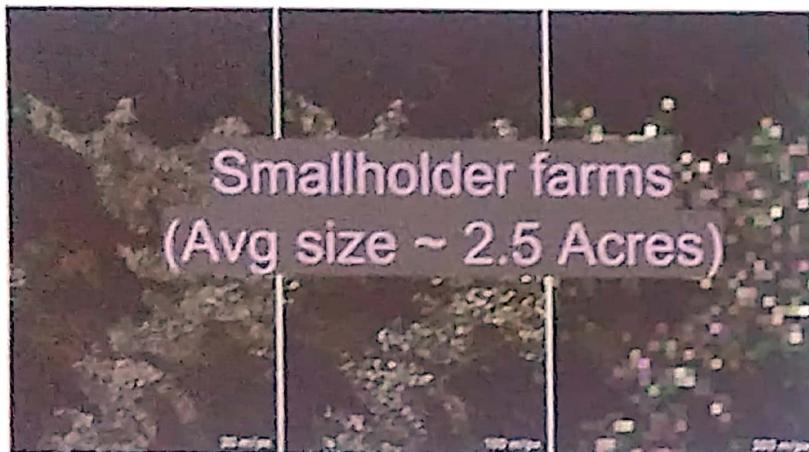
AnthroKrishi pushes on Spatial, Temporal and Spectral simultaneously.

Radiometric



# GeoFMs & Types of Earth Observation resolutions

Spatial

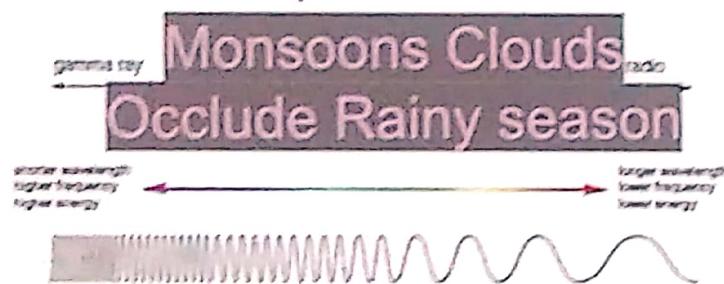


Understanding & interpreting data across all of these resolutions is the core challenge for GeoFMs

Temporal

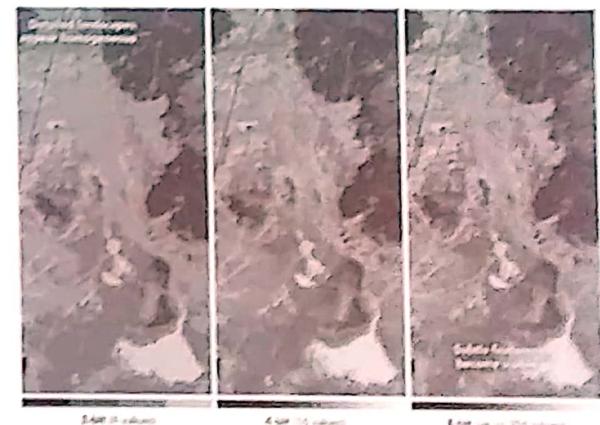


Spectral

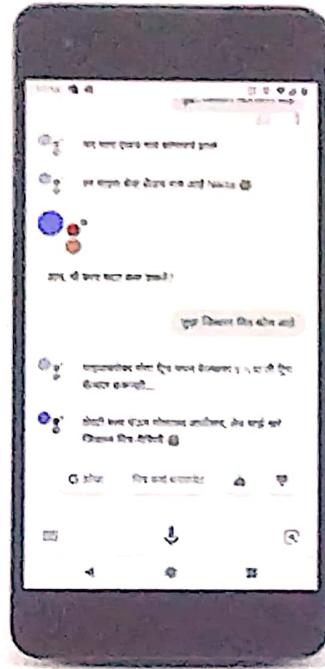


AnthroKrishi pushes on Spatial, Temporal and Spectral simultaneously.

Radiometric



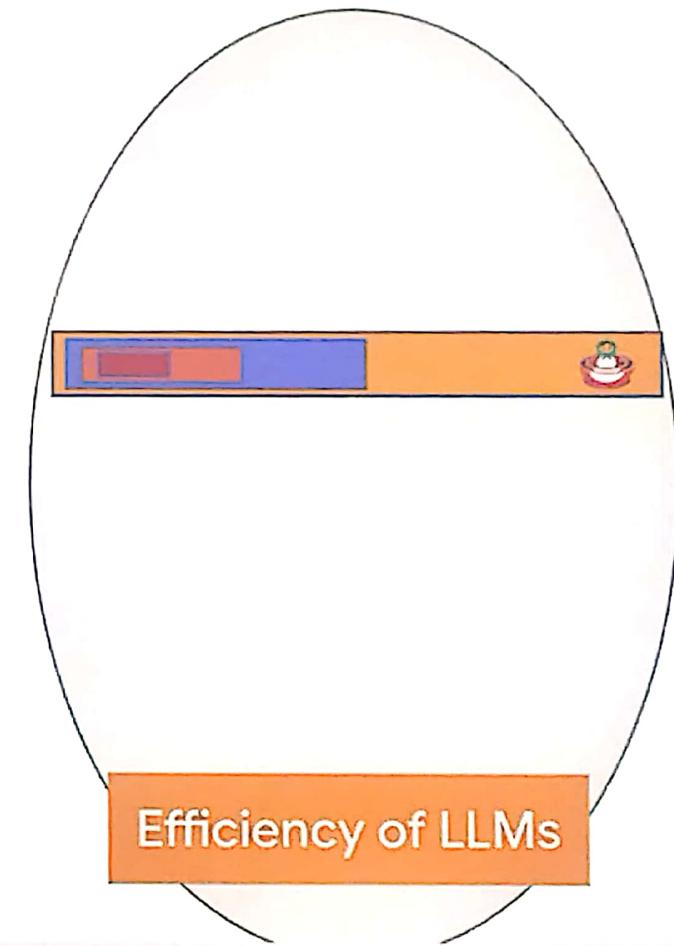
# Inclusive AI



Information in  
Indian Languages

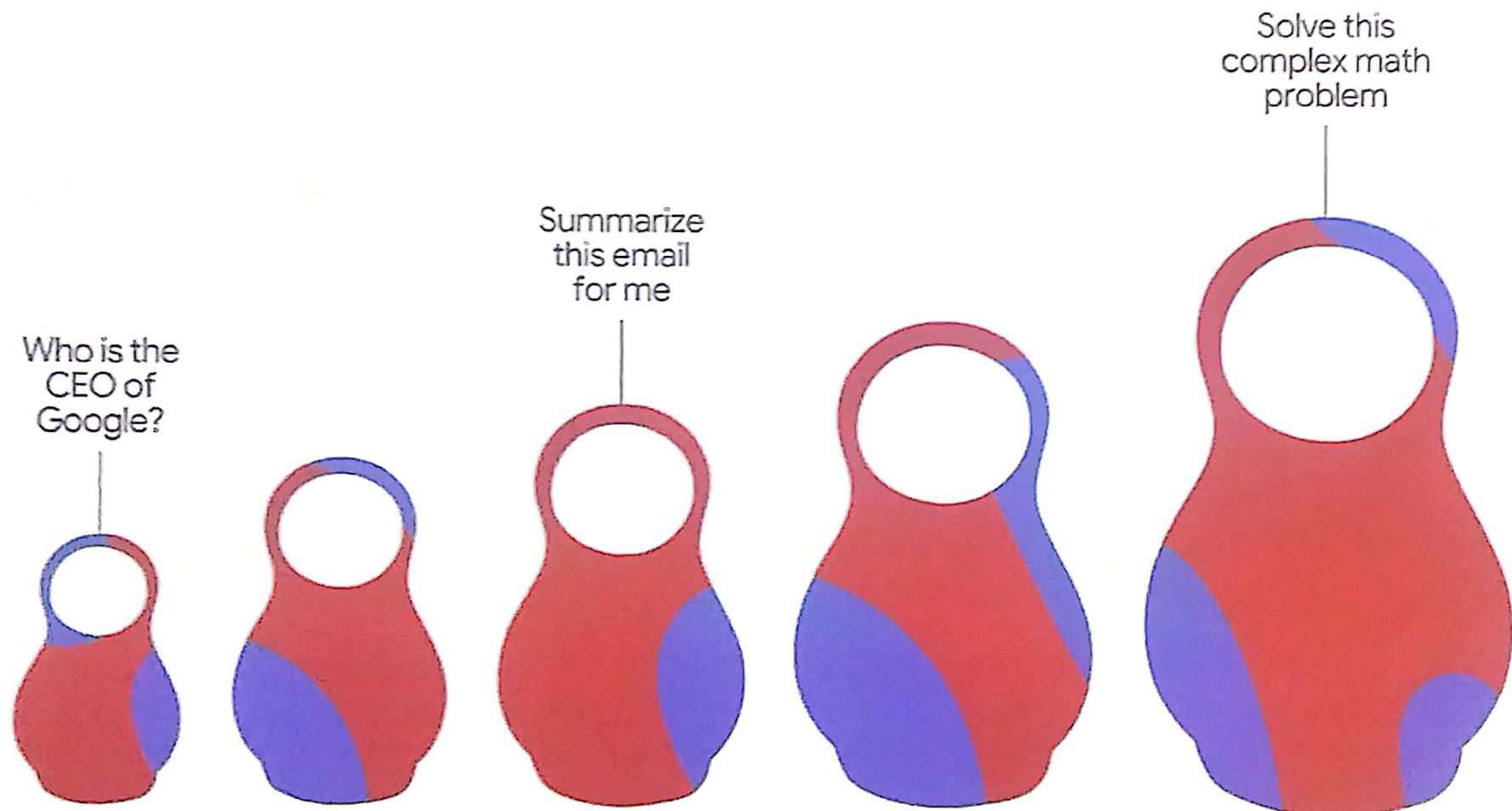


Agriculture

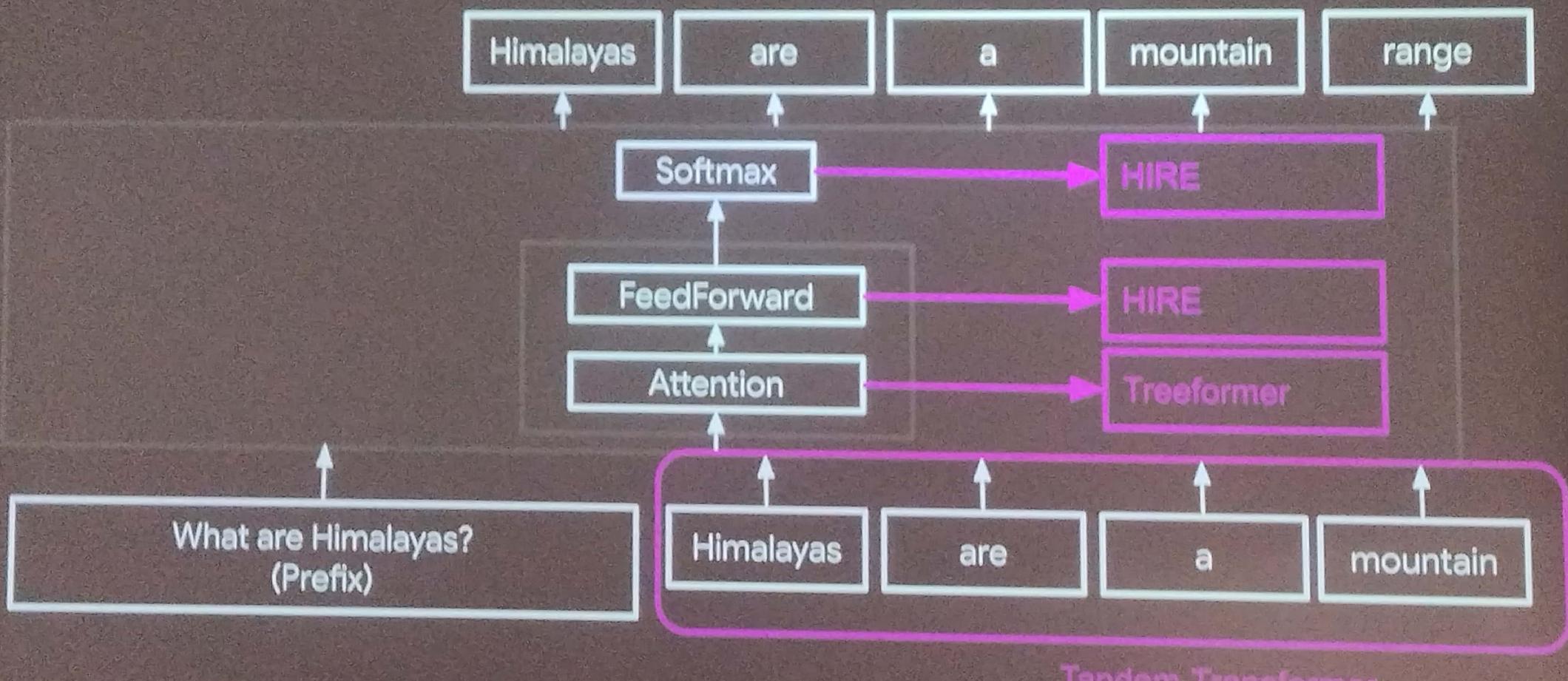


Fundamental Research on AI (Robustness, Fairness, Safety, Explainability)

# Elastic “Matryoshka” Models (Matformers)

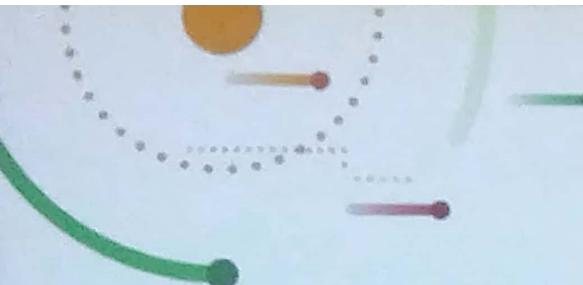


# Inference latency of Generative LLMs



# Some Outstanding Challenges

- Efficiency
- Factuality
- Robustness
- Safety
- Bias and Fairness
- Reasoning and Planning
- Continual Learning



## Takeaways

- AI is here - already impacting every industry and the world at large
- Rise of Foundation Models - Generative AI
- Huge opportunity to develop AI models and solutions to bring benefits to billions of new people - outstanding challenges

**These are exciting times!**