"Larger than Larger" Seminar Series

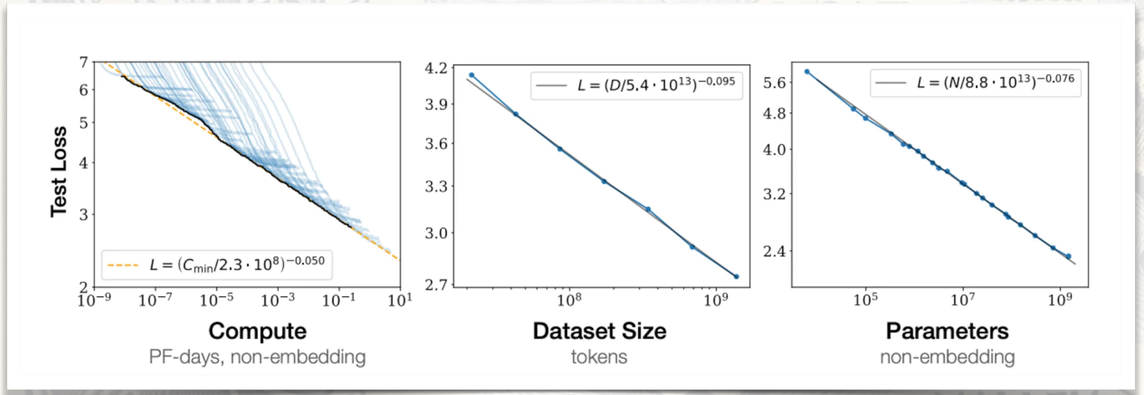
Large AI Models at the Frontiers of Experimental High-Energy Physics

Episode I:

Large Jet Models

LARGER THAN LARGER

LARGER AI MODELS



Large *Hadron Collider*

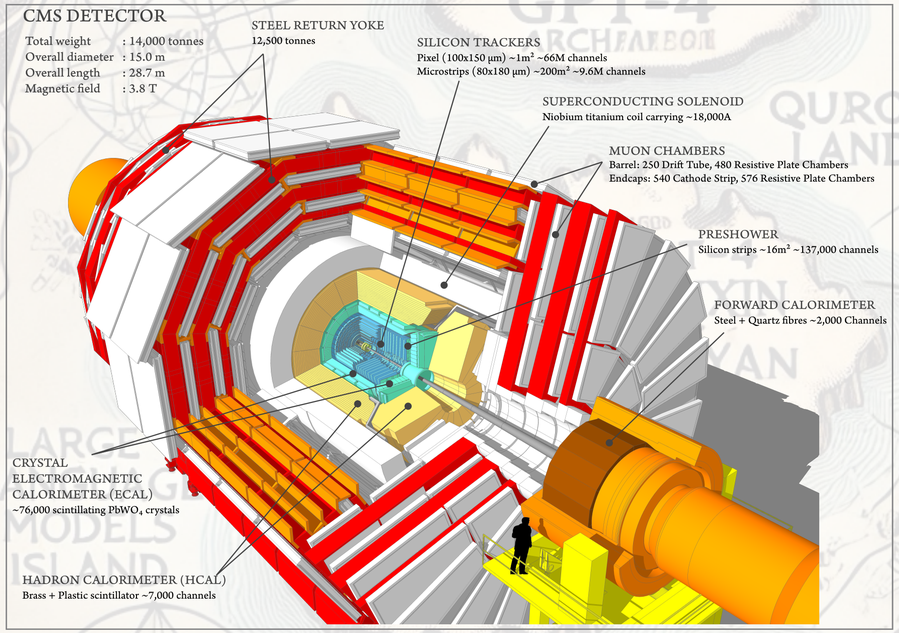
Very ***Large*** *Telescope*

"LARGER SCIENCE": LARGER FACILITIES LARGER DATA

• • •

Large *High Altitude Air Shower Observatory*

Big Experiment Produces Big Data

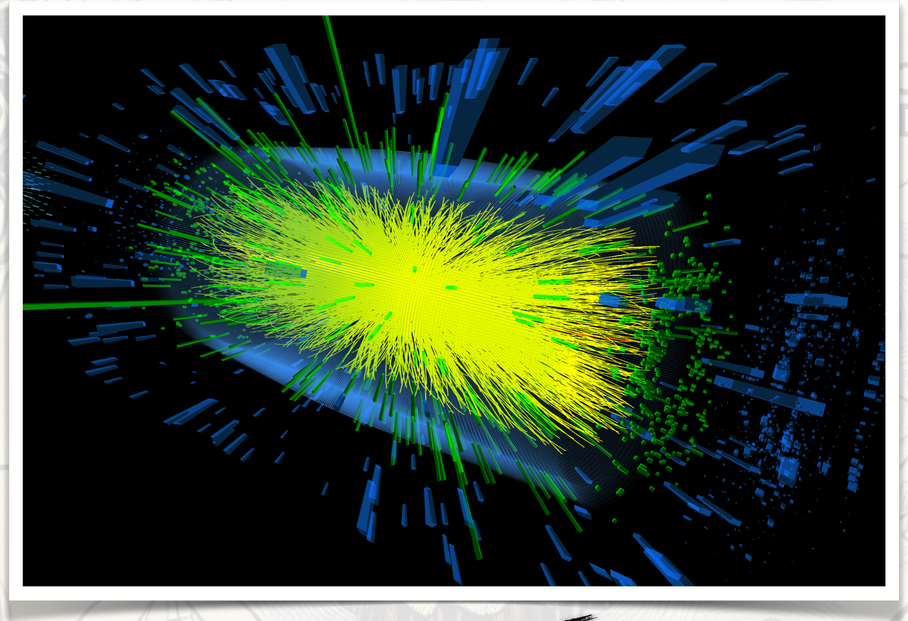


Biggest Collider Ever: the LHC

Big Collider needs Big Detectors

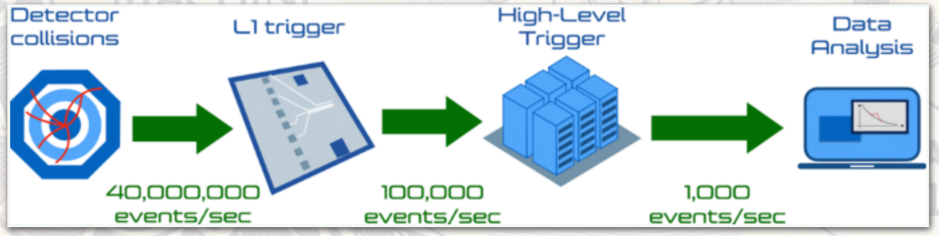
(Biased) Example Above:
The *Compact Muon Solenoid*
Heaviest Detector Ever

Big Experiment Produces Big Data

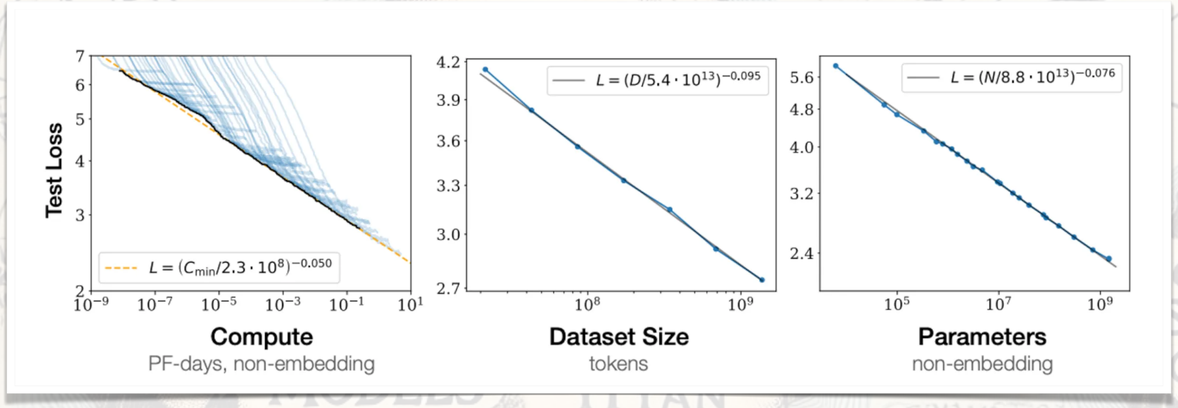


Biggest Collider ever: the LHC

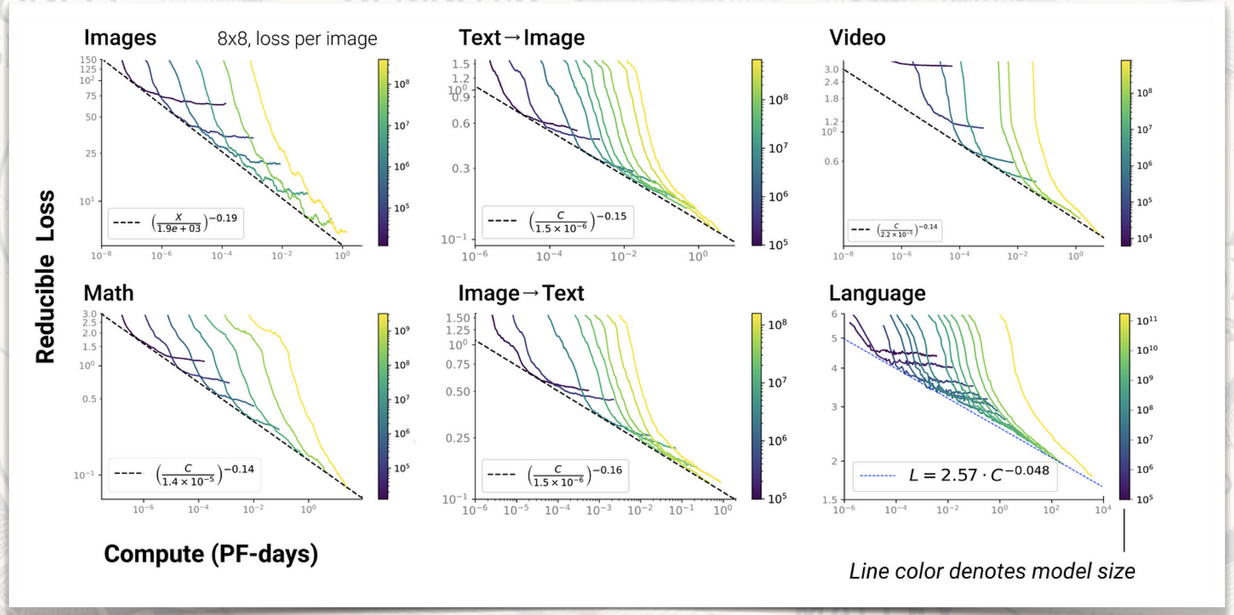
Big Detectors record
Big amount of Data

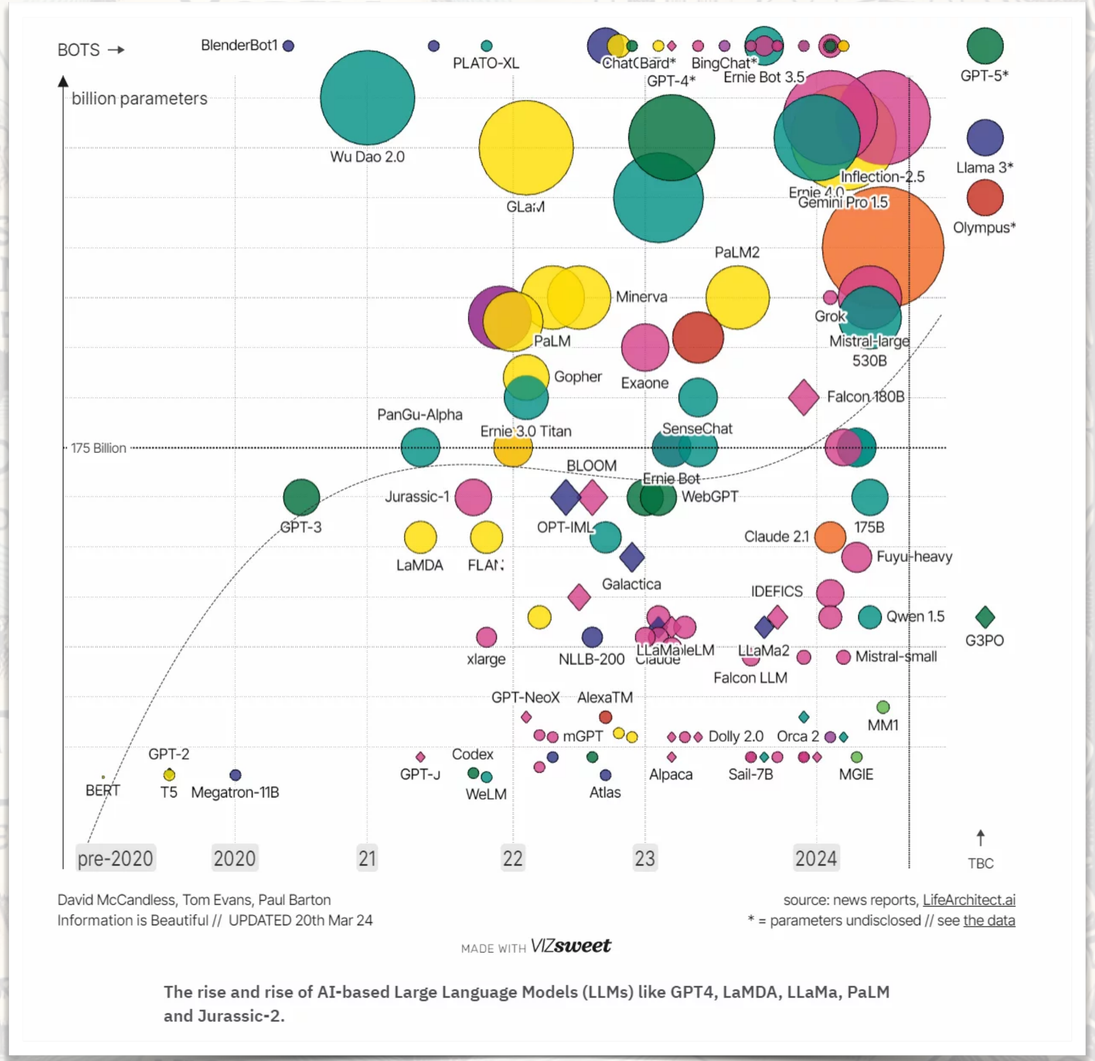


Big Data Feeds Up Big Models



Big Models are trained with Big Data…



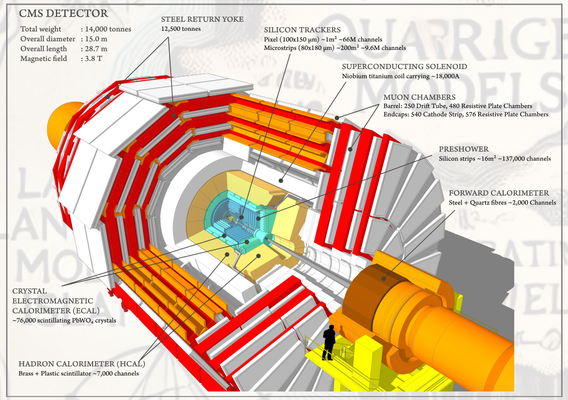


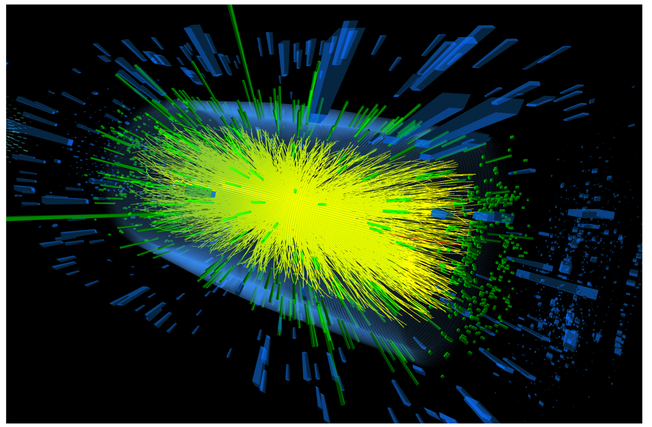
The rise and rise of AI-based Large Language Models (LLMs) like GPT4, LaMDA, LLaMa, PaLM and Jurassic-2.

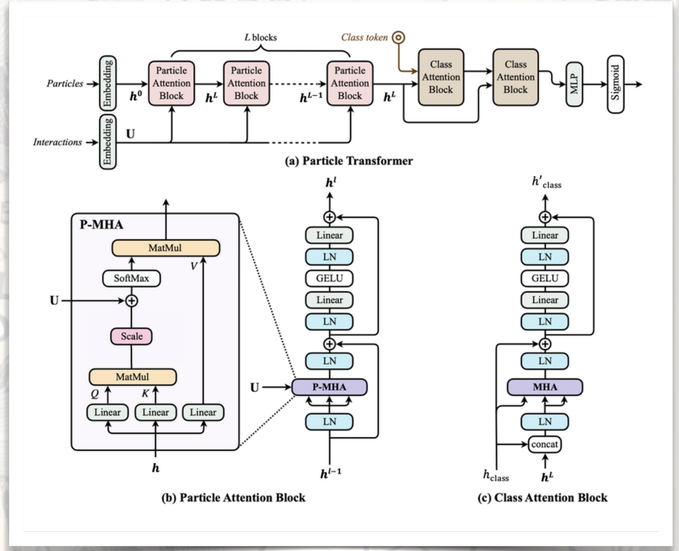
So…

Larger

Larger

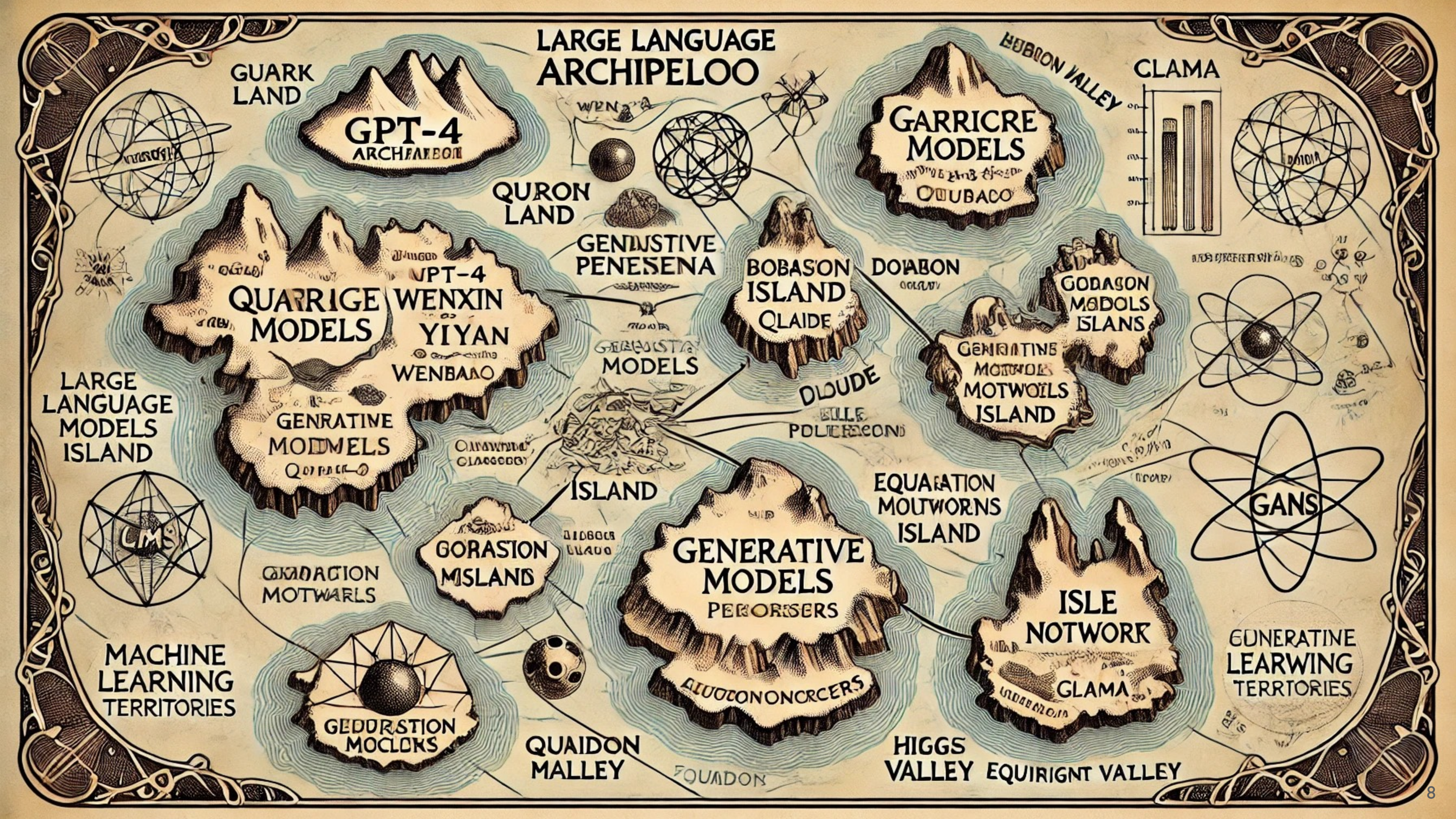






Larger than Larger

WE HAVE AN AMBITIOUS PLAN!


LARGE LANGUAGE
ARCHIPELOO
GUARK
LAND
GPT-4
QURON
LAND
GENISTIVE
PENESENA
QUARRIGE
MODELS
VPT-4
WENXIN
YIYAN
WENBAAO
GENRATIVE
MODELS
LARGE
LANGUAGE
MODELS
ISLAND
GARRICRE
MODELS
BOBASON
ISLAND
QLAIDE
DOBABON
GODASON
MODOLS
ISLANS
GENERATIVE
MOTWOILS
ISLAND
DLOUDE
ISLAND
GORASION
MSLANDS
GENERATIVE
MODELS
PERORSERS
EQUARATION
MOUTWORNS
ISLAND
GANS
ISLE
NOTWORK
GLAMA
MACHINE
LEARNING
TERRITORIES
GEDORSTION
MOCLOKS
QUAIDON
MALLEY
HIGGS
VALLEY
EQUIRIGNT VALLEY
GENERATINE
LEARWING
TERRITORIES

We have an ambitious plan!

(Bi-)Monthly Event

Spoiler: Next Event on LLM4HEP in March (Stay tuned)!!!

February skipped for AI+HEP in East Asia

AI+HEP in East Asia

2025年2月24日至28日
IBS
Asia/Seoul 时区

输入您的搜索词

- 概览
- 征集摘要
- 注册
- 参会人名单
- Maps and Directions
- Visa Information
- Code of Conduct

Contact

sunghak.lim@ibs.re.kr
sunghak.lim@rutgers.edu

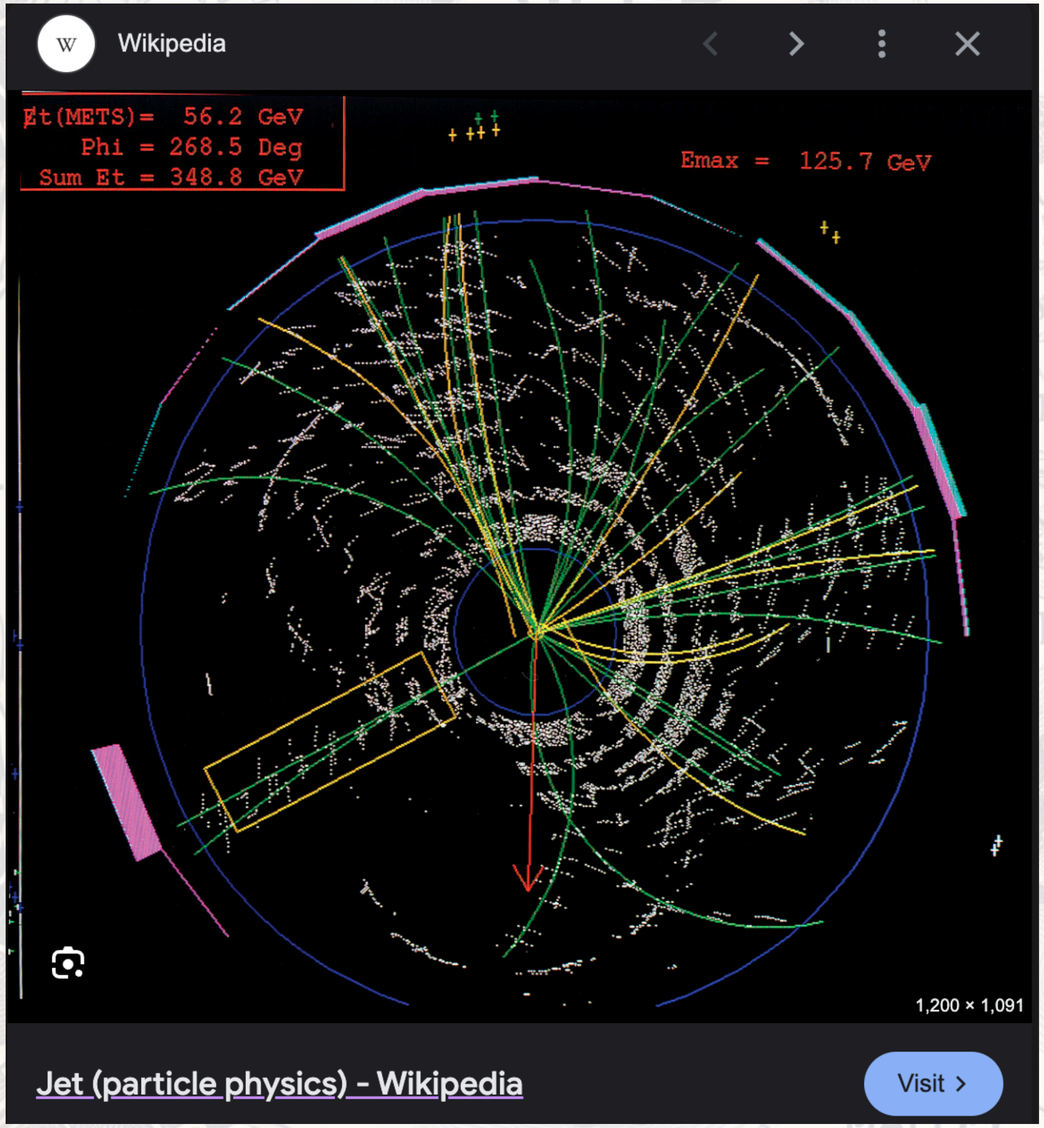
Notes:

- Please ignore any emails from 3rd party companies, as we do not have any contract.
- Connections to this website from overseas may be slow due to the IBS firewall. Sorry for the inconvenience.

This regional workshop aims to connect researchers in East Asia working in the interdisciplinary field of Artificial Intelligence and High Energy Physics (AI+HEP). The main topics covered include machine learning for particle theory, phenomenology and experiments, astrophysics and cosmology, as well as HEP tools for AI theory.

The workshop will have invited plenary talks, contributed presentations, and ample time for discussions. Both domain experts and those who are interested in exploring the field are welcome to participate, especially postdocs and graduate students. The goal is to foster a regional research community and to stimulate more collaborations.

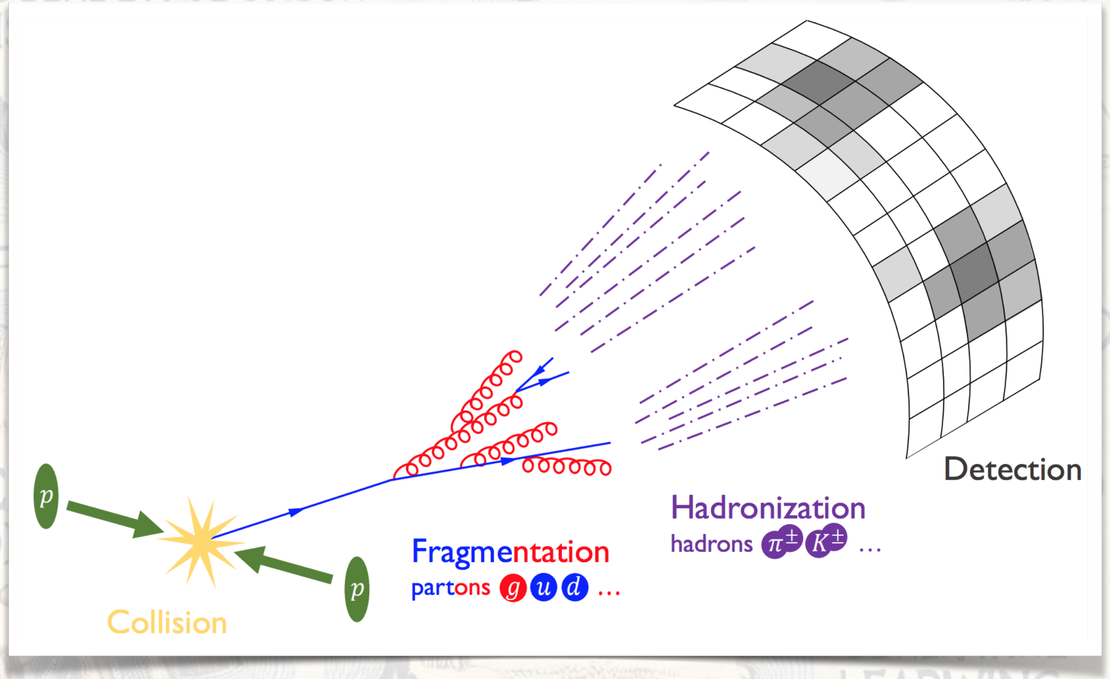
Today's Topic: Jet



Jet is

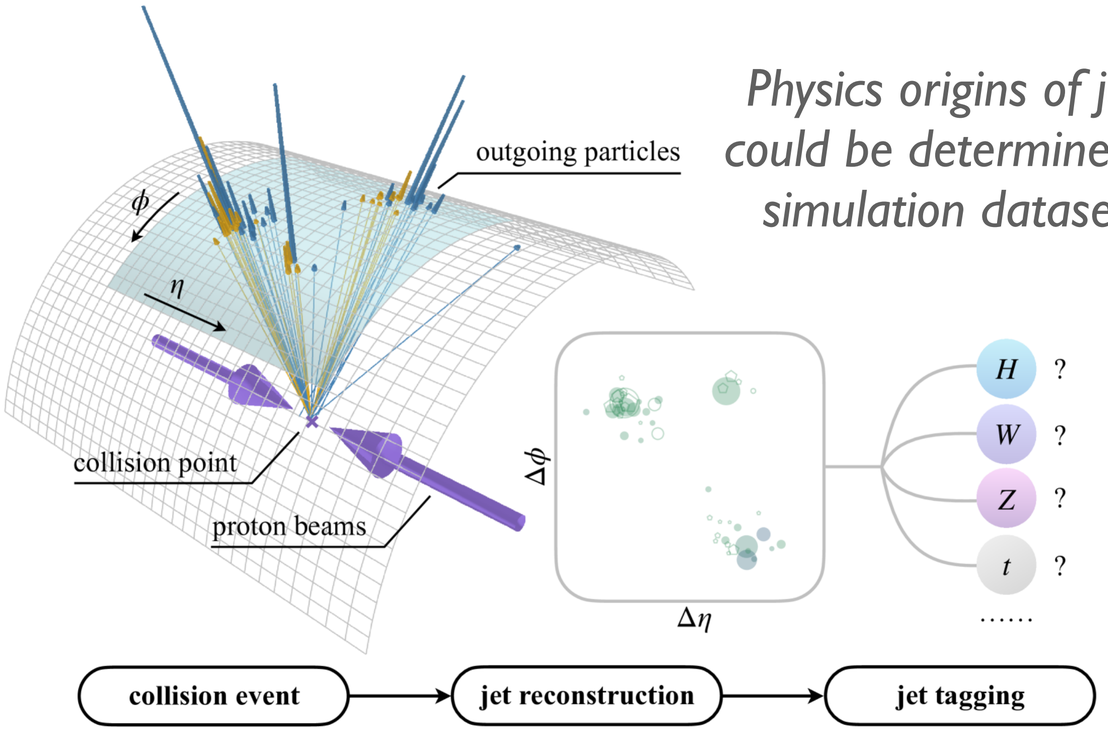
A spiral of collinear "particles"

A natural consequence of QCD

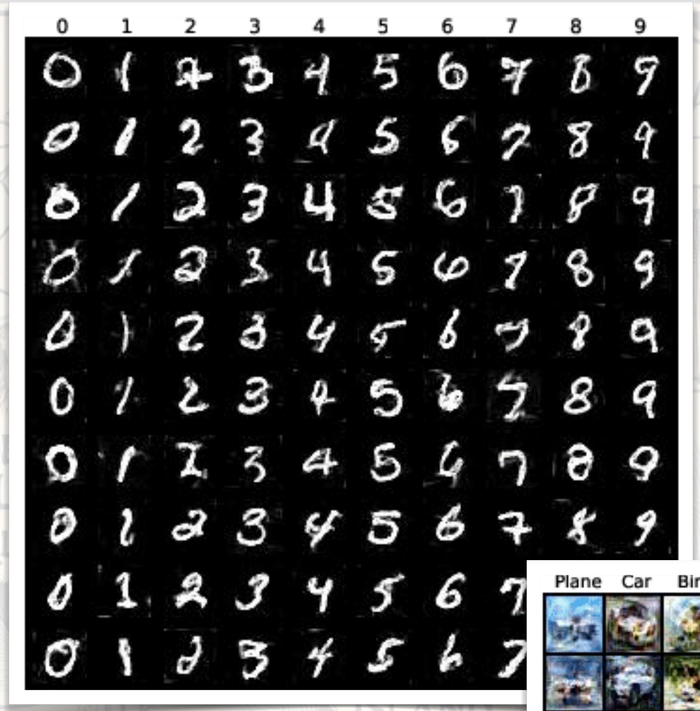


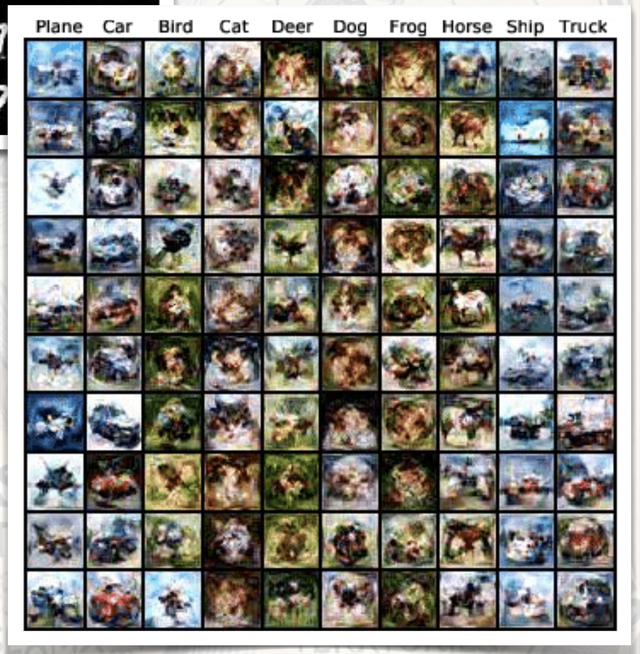
And a perfect ML playground!

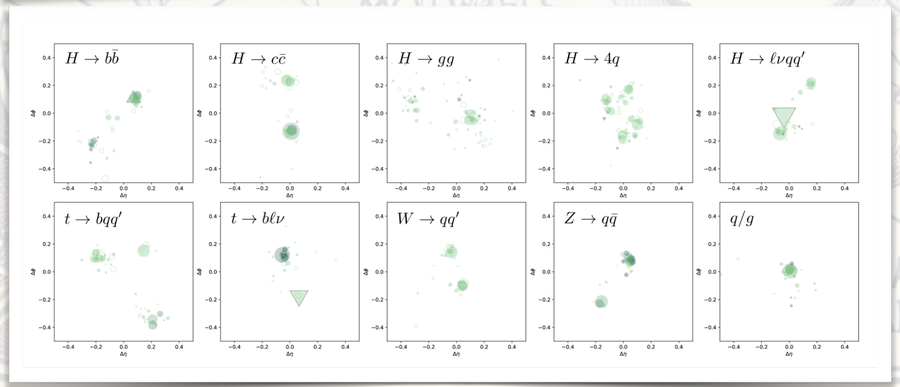
JETS CAN BE CLASSIFIED



Physics origins of jets could be determined in simulation datasets

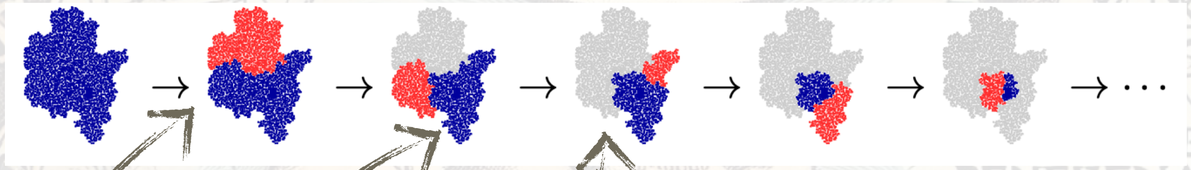


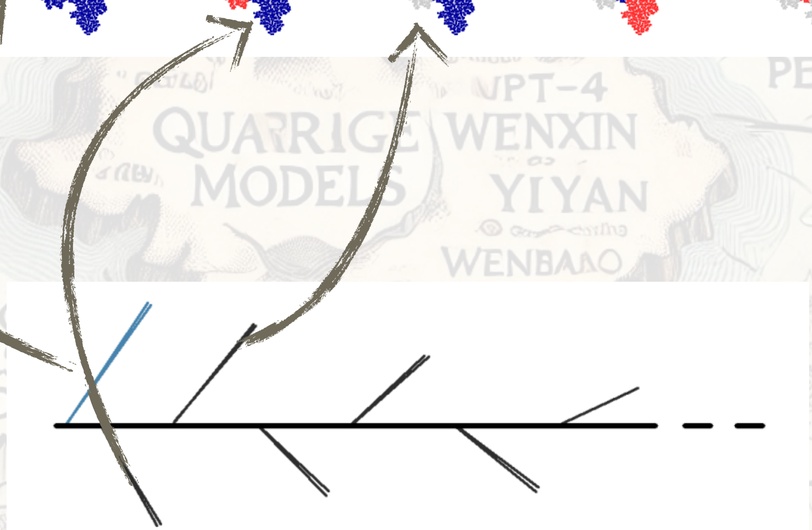




AS OBJECTS AND DIGITS

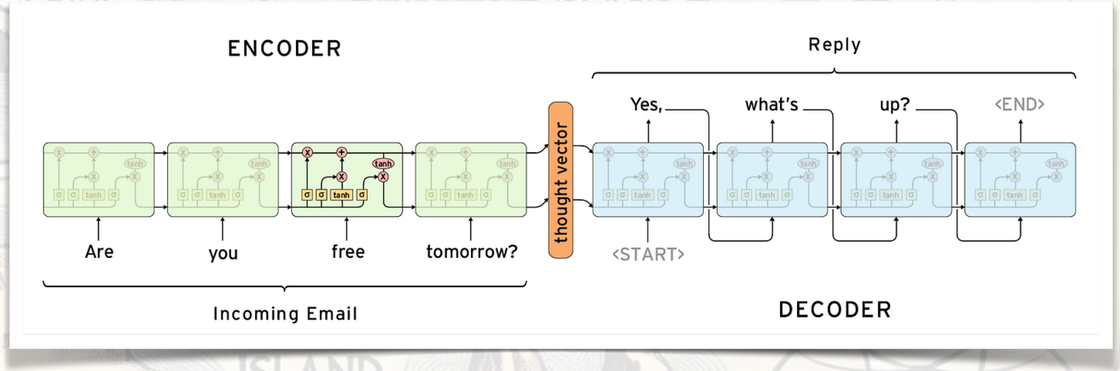
JETS IS A SEQUENCE





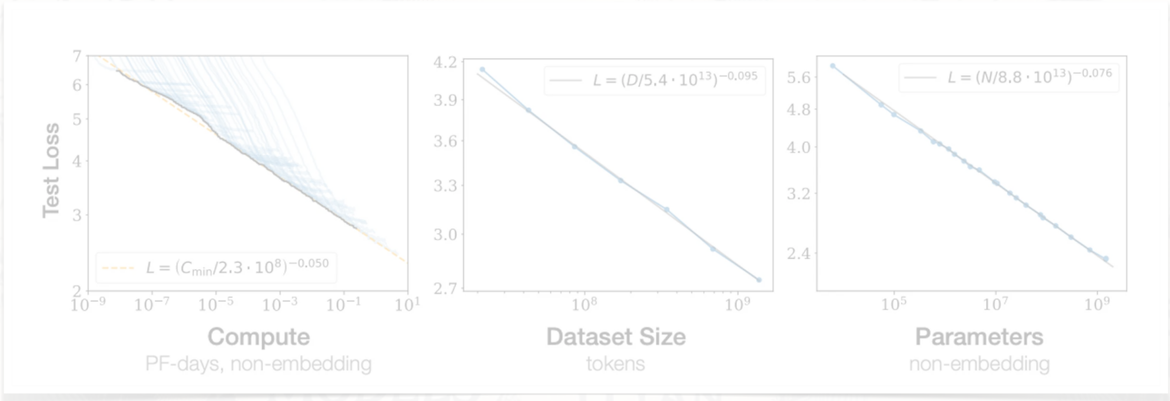
Due to certain QCD quantities, a jet could be decluttered and then form a sequence

Then a sequence could be treated the same as a sentence

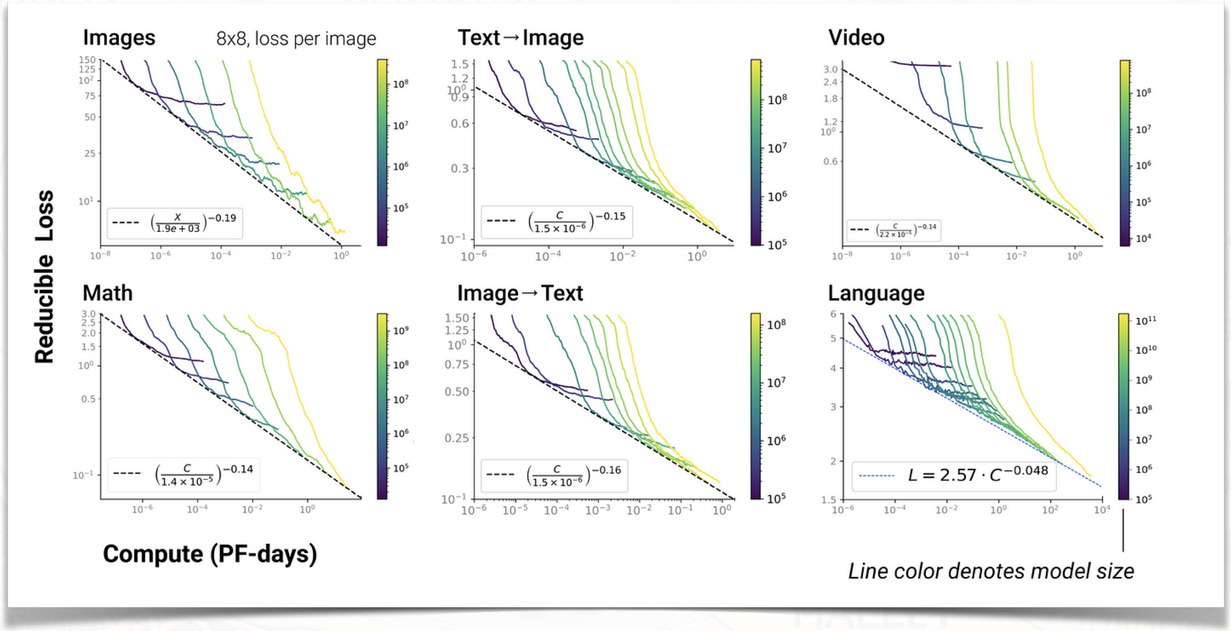


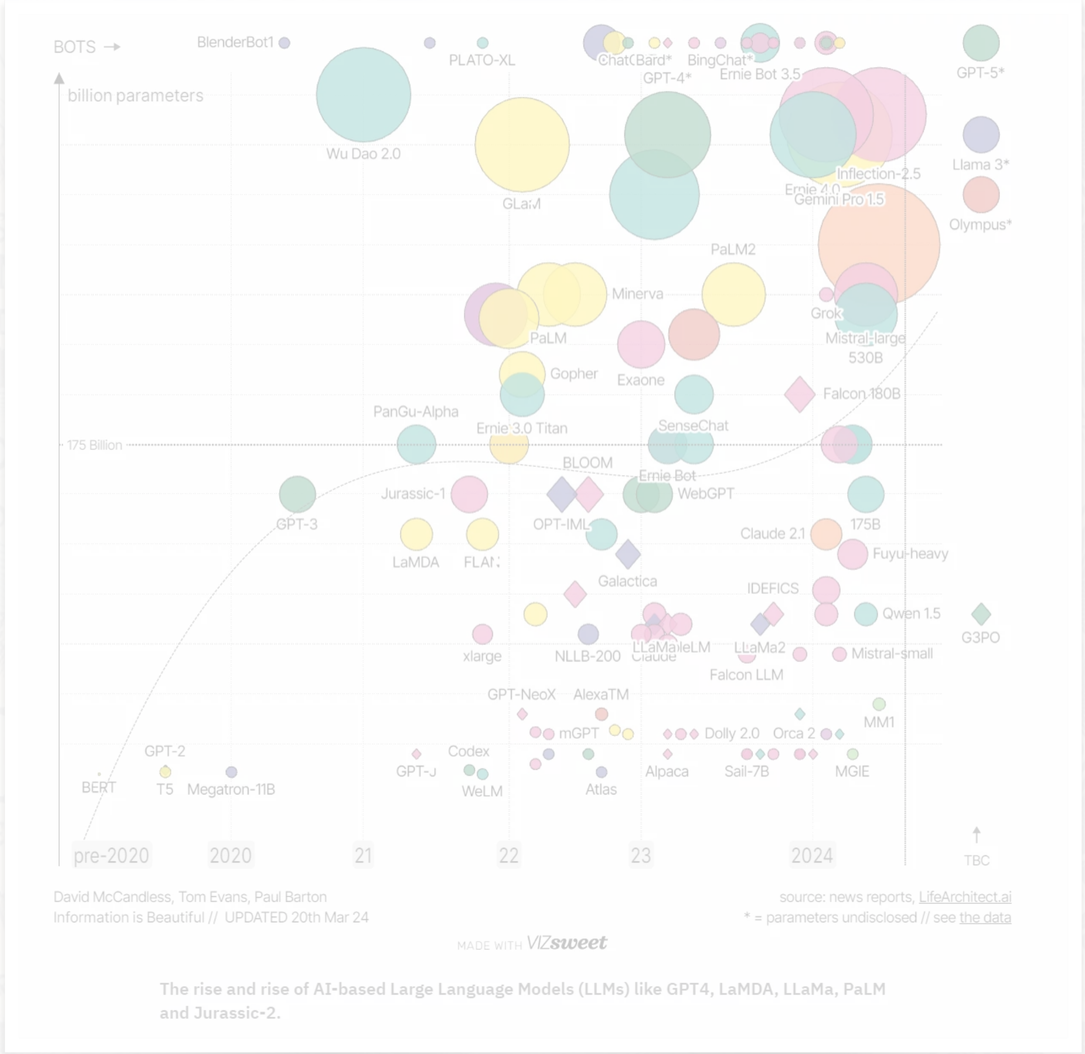
AS NATURAL LANGUAGE

Big Data Feeds Up Big Models

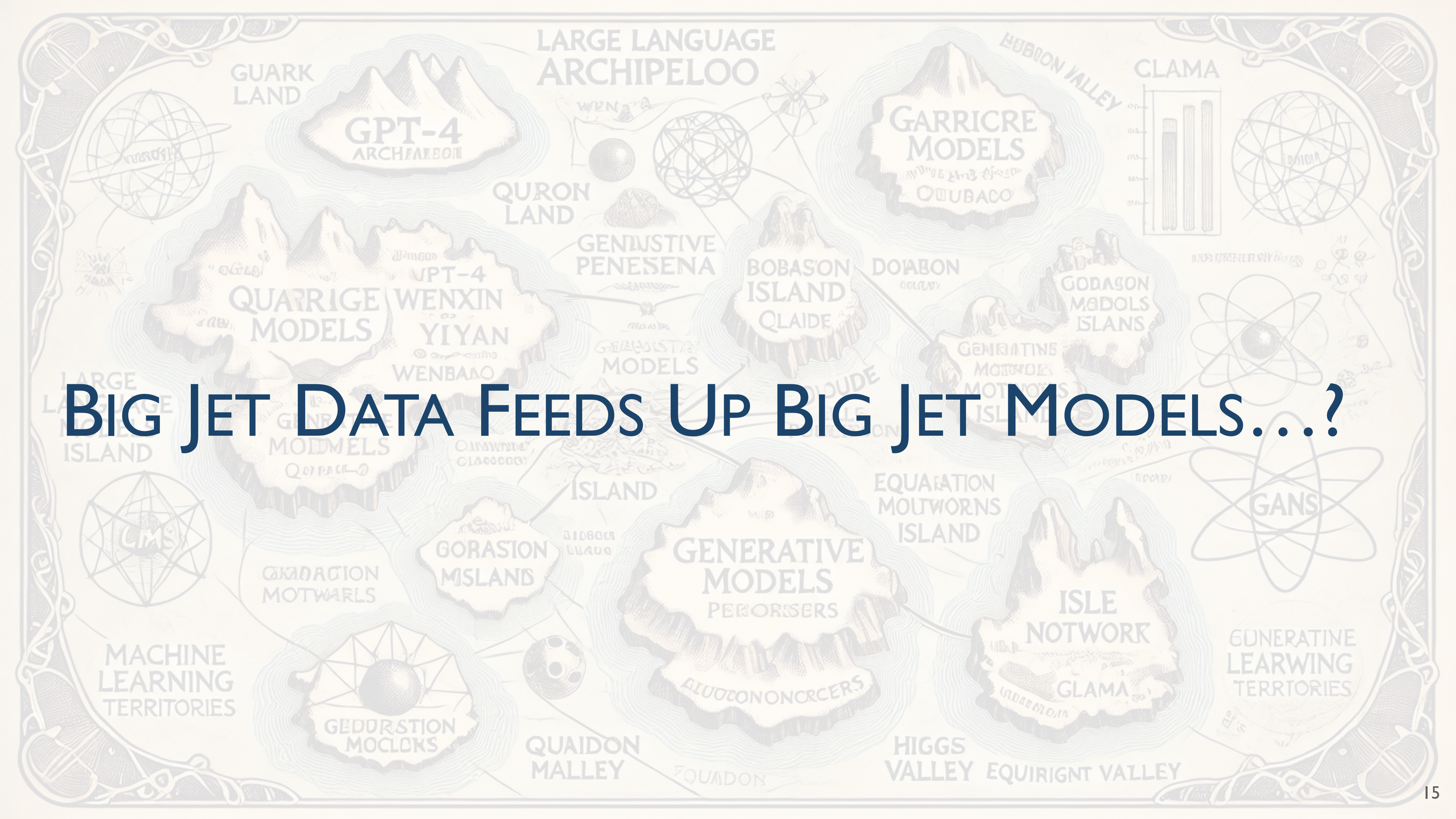


Big Models are trained with Big Data…



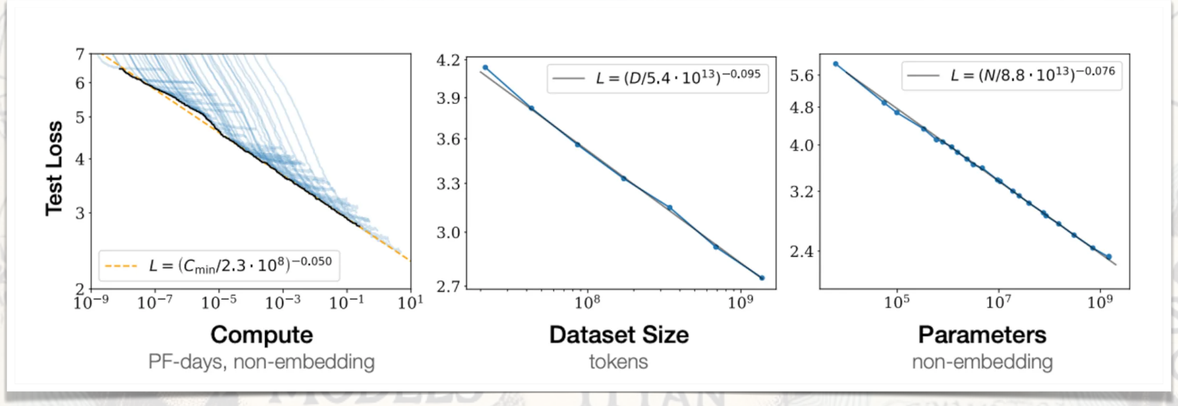


The rise and rise of AI-based Large Language Models (LLMs) like GPT4, LaMDA, LLaMa, PaLM and Jurassic-2.

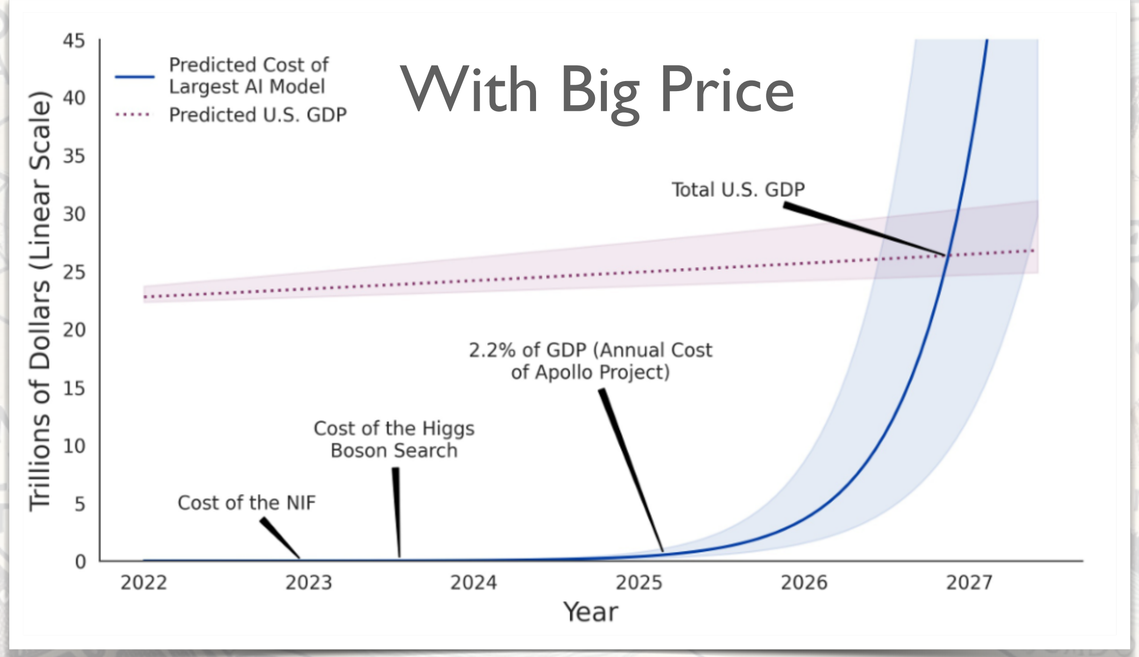


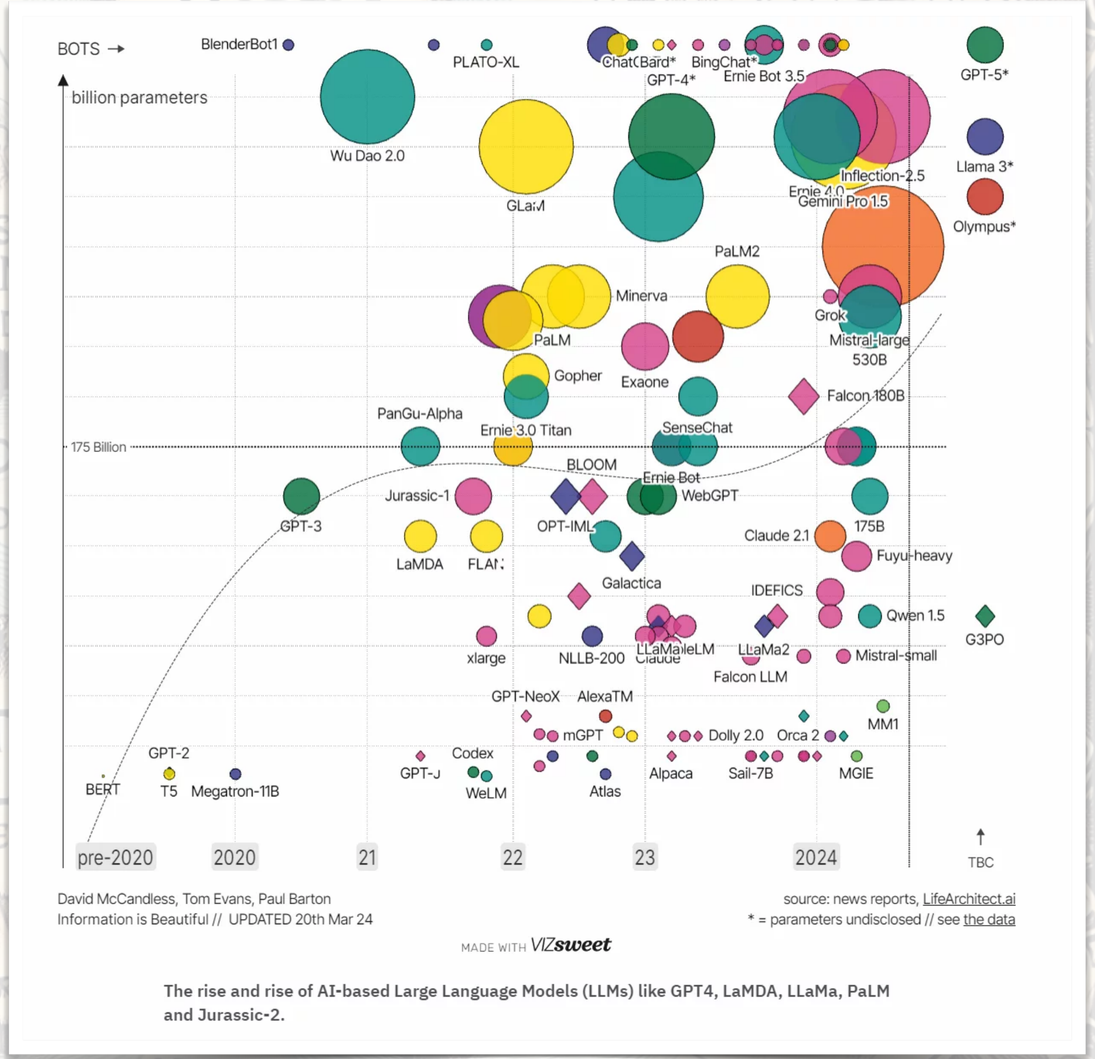
BIG JET DATA FEEDS UP BIG JET MODELS…?

Big Data Feeds Up Big Models Costly!



Big Models are trained with Big Data…



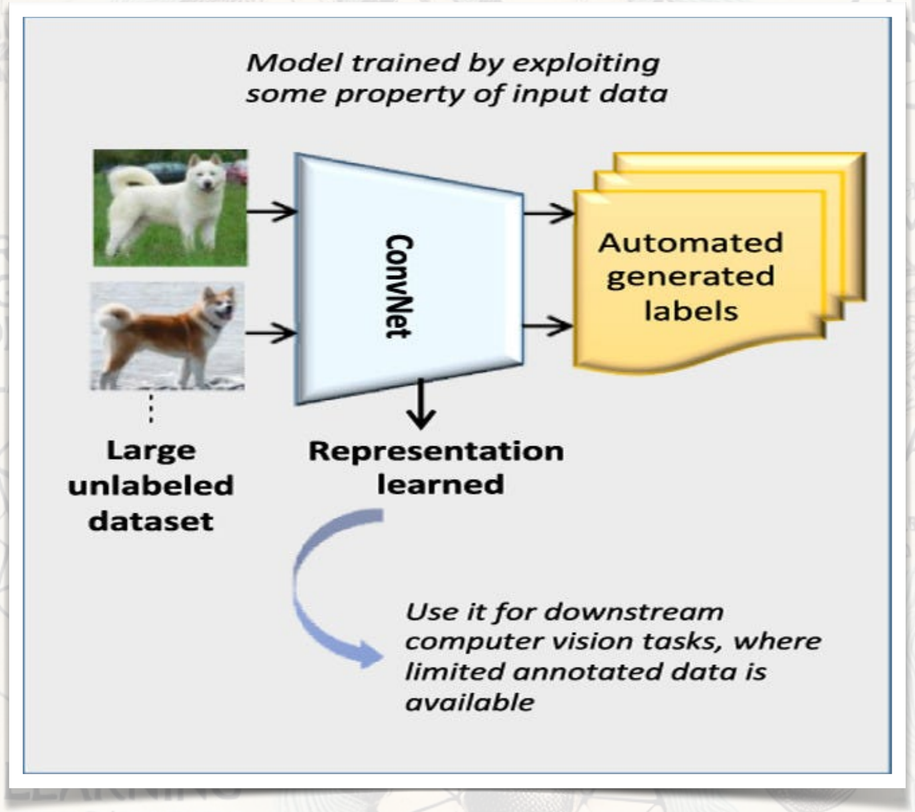


The rise and rise of AI-based Large Language Models (LLMs) like GPT4, LaMDA, LLaMa, PaLM and Jurassic-2.

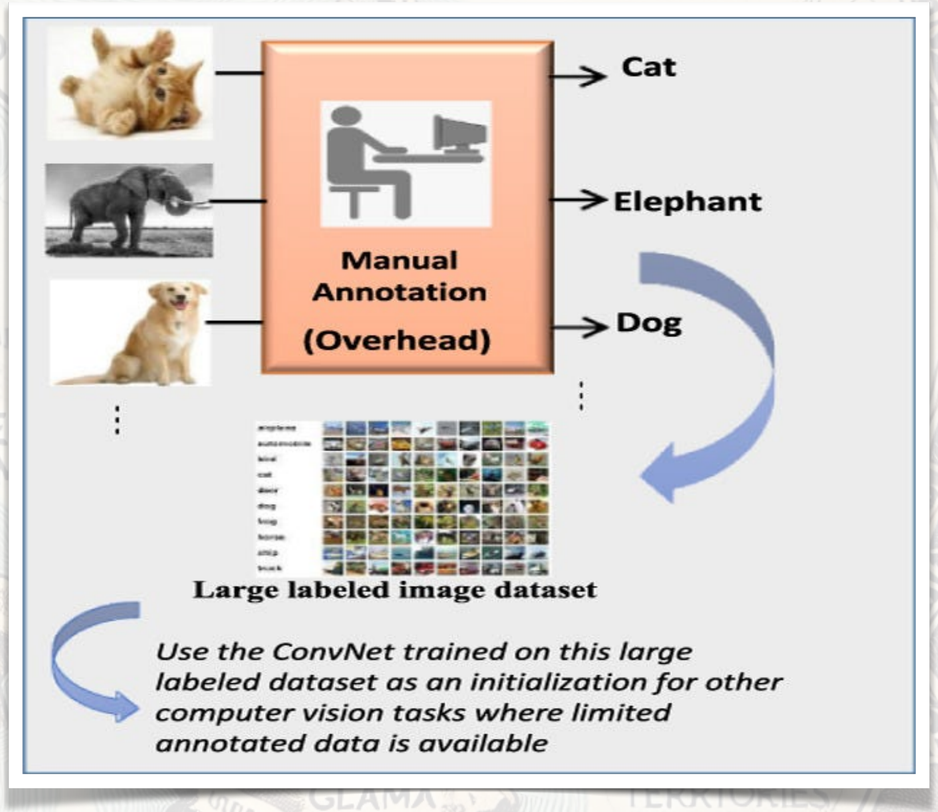
Big Data Feeds Up Big Models via…

In general, a smart way of training is needed to save money!

You want a more inclusive model



Approach A:
Self-Supervised Learning



Approach B:
Large Supervised Learning

Big Data Feeds Up Big Models via…

In general, a smart way of training is needed to save money!
You want a more inclusive model by digesting as much information as possible

Zihan Zhao (赵梓涵)
UCSD / CMS

Shudong Wang (王书栋)
IHEP / ATLAS

20min break

…

Congqiao Li (李聪乔)
PKU / CMS

Approach A:
Self-Supervised Learning

Approach B:
Large Supervised Learning