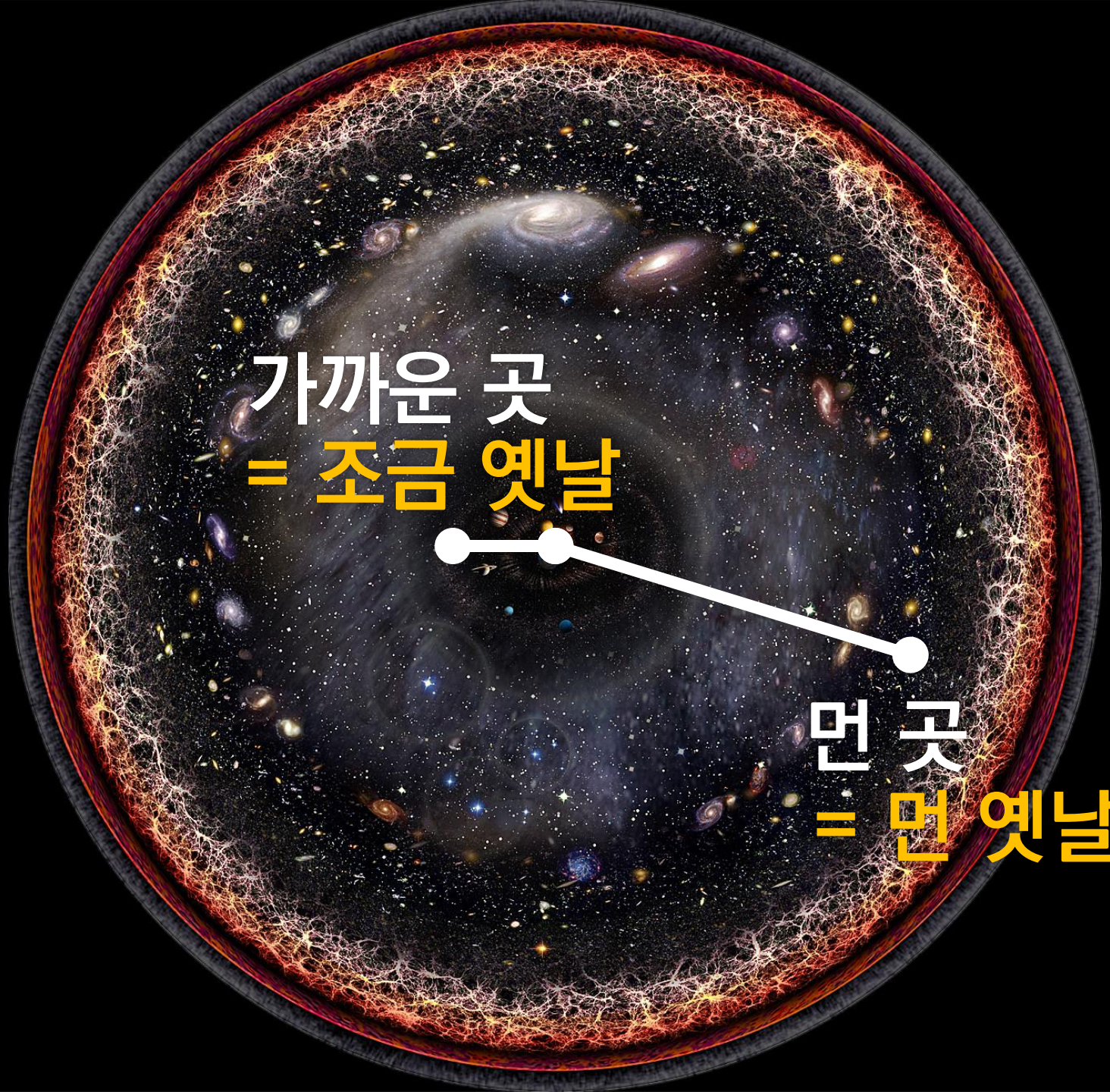
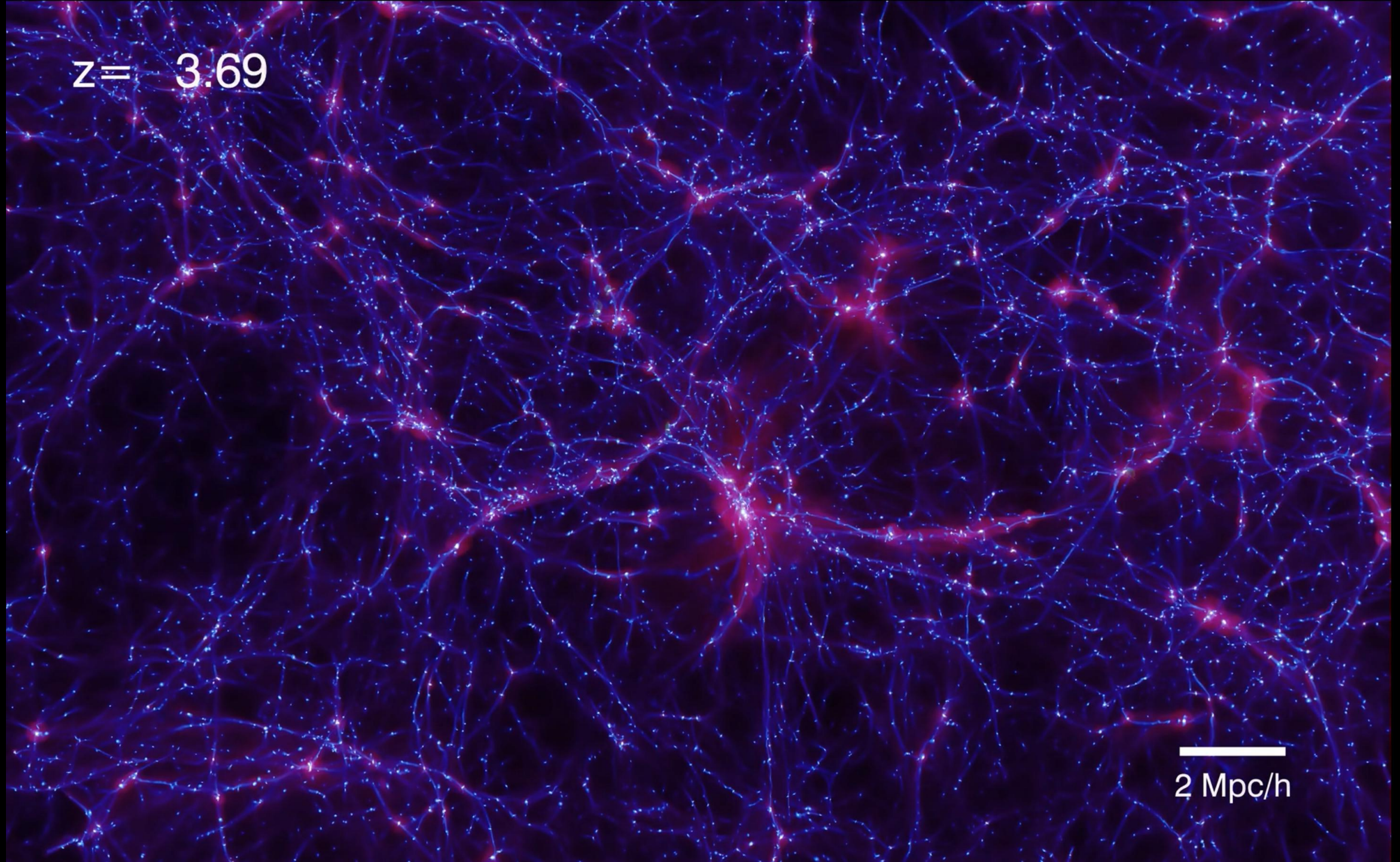
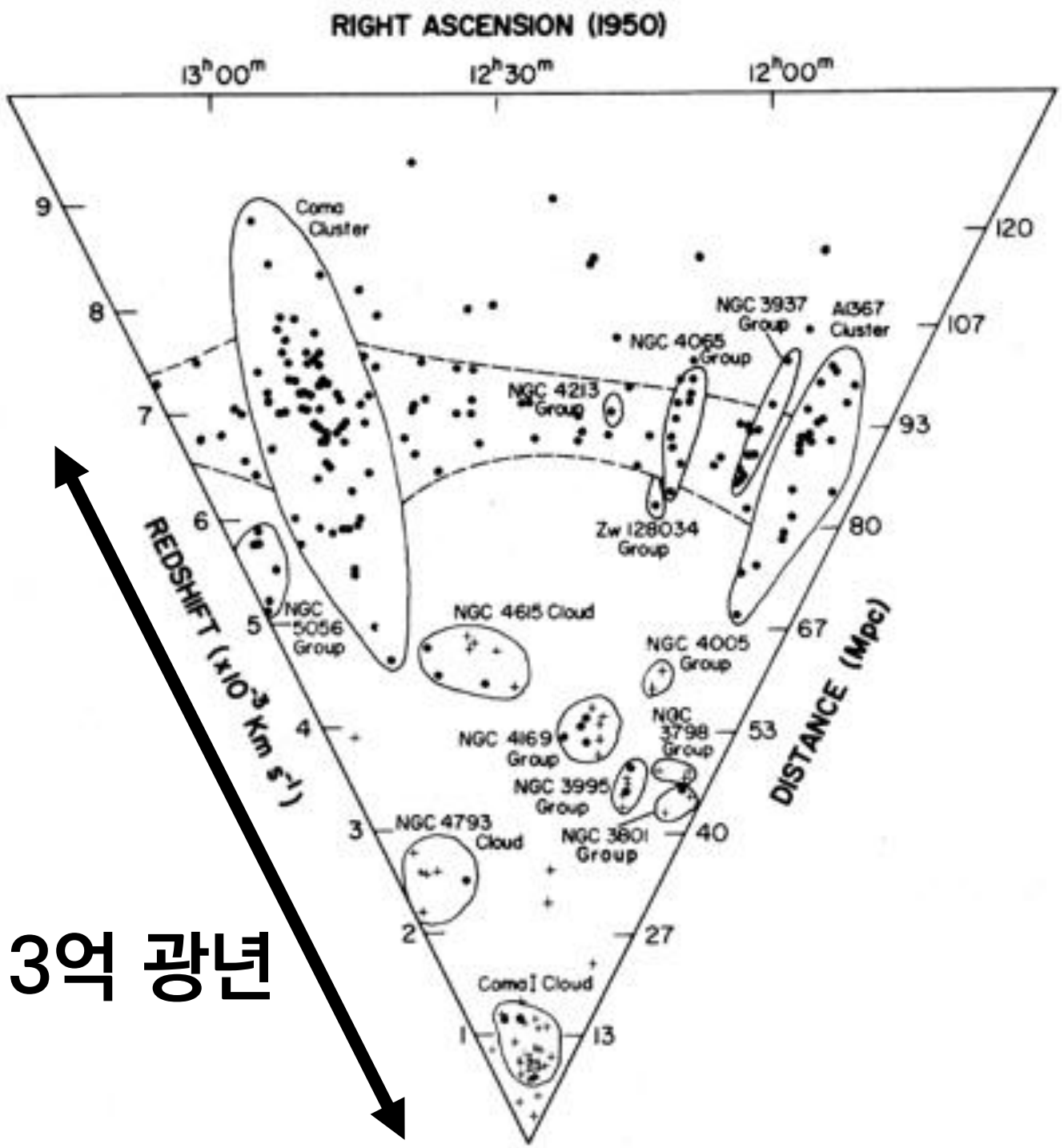
(한국) 천문우주 과학에서의 AI 활용, 그리고 도전

홍성욱
한국천문연구원 우주진화연구센터

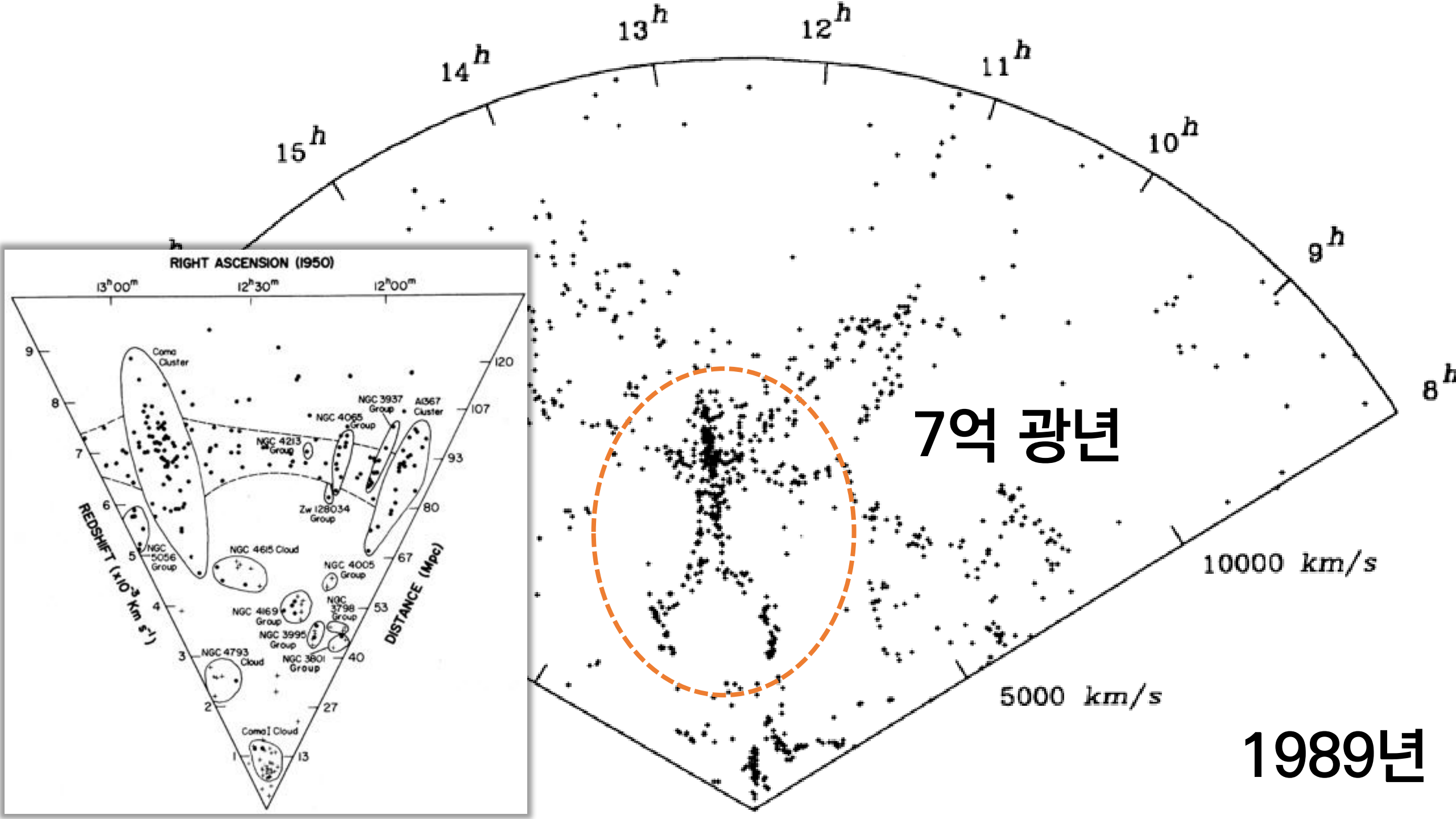

가까운 곳
= 조금 옛날
먼 곳
= 먼 옛날


z= 3.69
2 Mpc/h

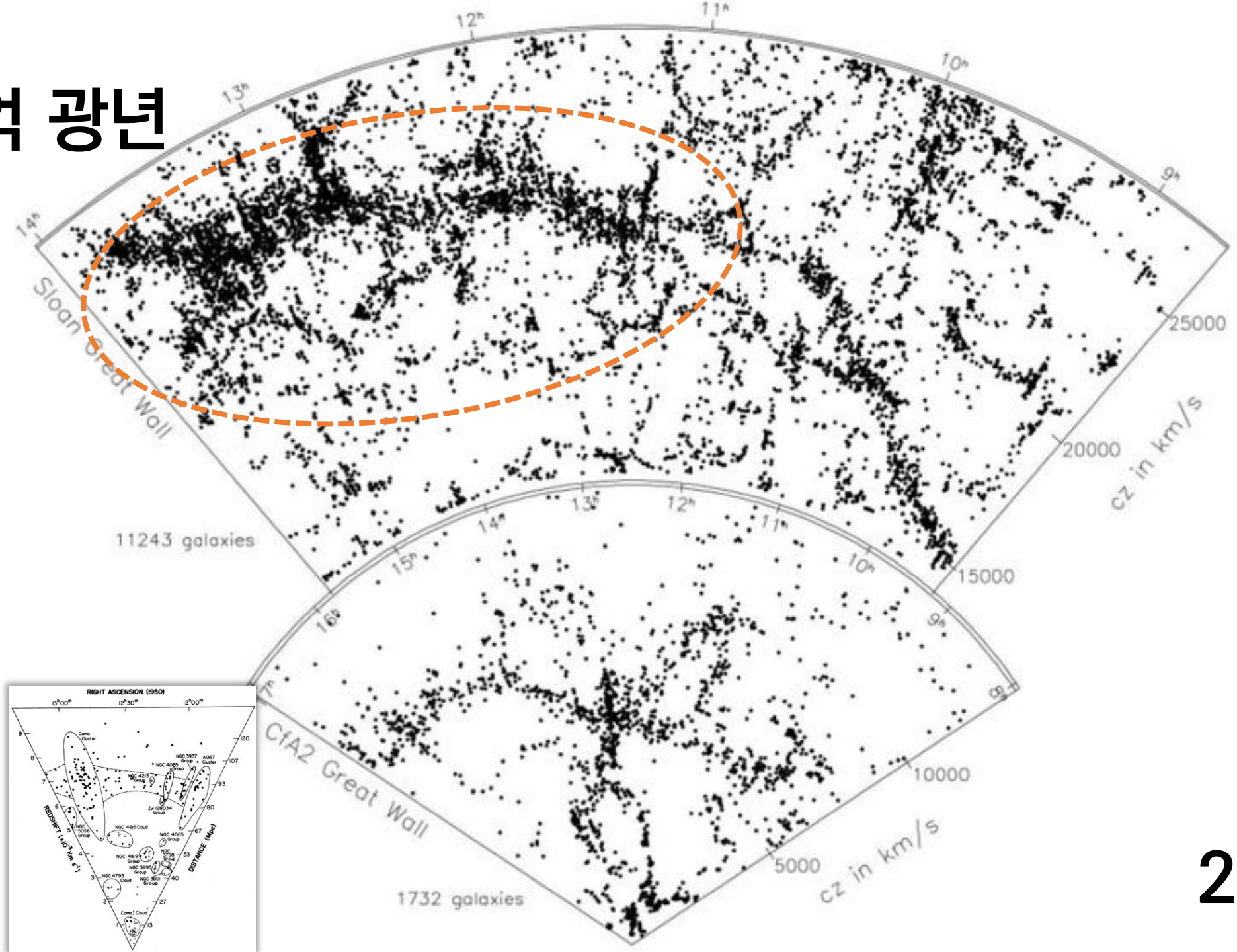

RIGHT ASCENSION (1950)
13h00m
12h30m
12h00m
9
8
7
6
5
4
3
2
1
120
107
93
80
67
53
40
27
13
Coma Cluster
NGC 3937 Group
A1367 Cluster
NGC 4065 Group
NGC 4213 Group
Zw 128034 Group
NGC 5056 Group
NGC 4615 Cloud
NGC 4005 Group
NGC 4169 Group
NGC 3798 Group
NGC 3995 Group
NGC 4793 Cloud
NGC 3801 Group
Coma I Cloud
REDSHIFT (x10-3 Km s-1)
DISTANCE (Mpc)

3억 광년

1978년


RIGHT ASCENSION (1950)
13h00m
12h30m
12h00m
REDSHIFT (x10-3 Km s-1)
DISTANCE (Mpc)
Coma Cluster
NGC 3937 Group
A1367 Cluster
NGC 4065 Group
NGC 4213 Group
Zw 128034 Group
NGC 4415 Cloud
NGC 5056 Group
NGC 4005 Group
NGC 4169 Group
NGC 3798 Group
NGC 3995 Group
NGC 4793 Cloud
NGC 3801 Group
Coma I Cloud
7억 광년
10000 km/s
5000 km/s
1989년

13억 광년


Sloan Great Wall
11243 galaxies
cz in km/s
CfA2 Great Wall
1732 galaxies
cz in km/s

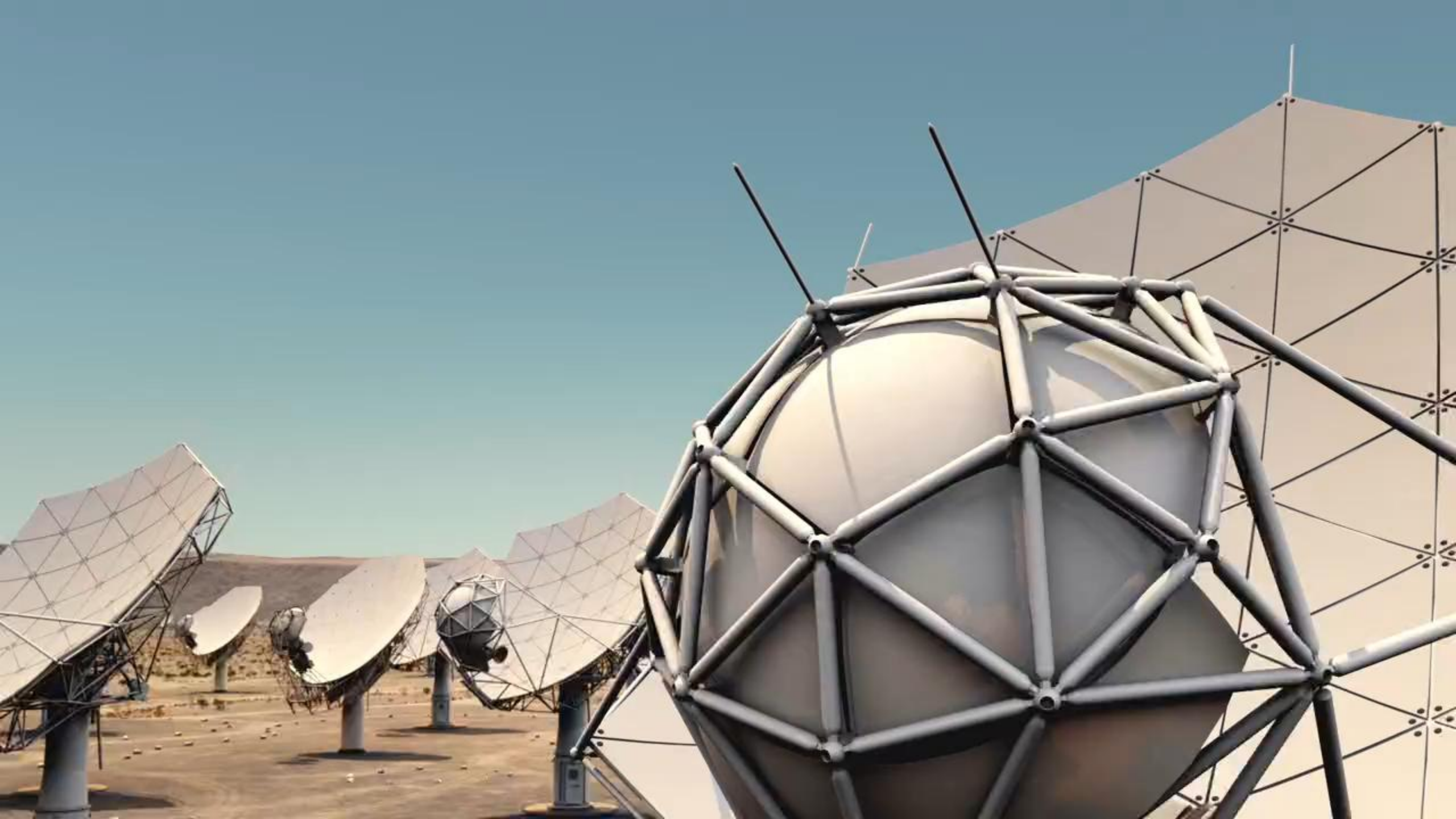
RIGHT ASCENSION (1950)
REDSHIFT (x10³ Km s⁻¹)
DISTANCE (Mpc)
Coma Cluster
NGC 4065 Group
NGC 4065 Cloud
NGC 4005 Group
NGC 4793 Cloud
Coma I Cloud

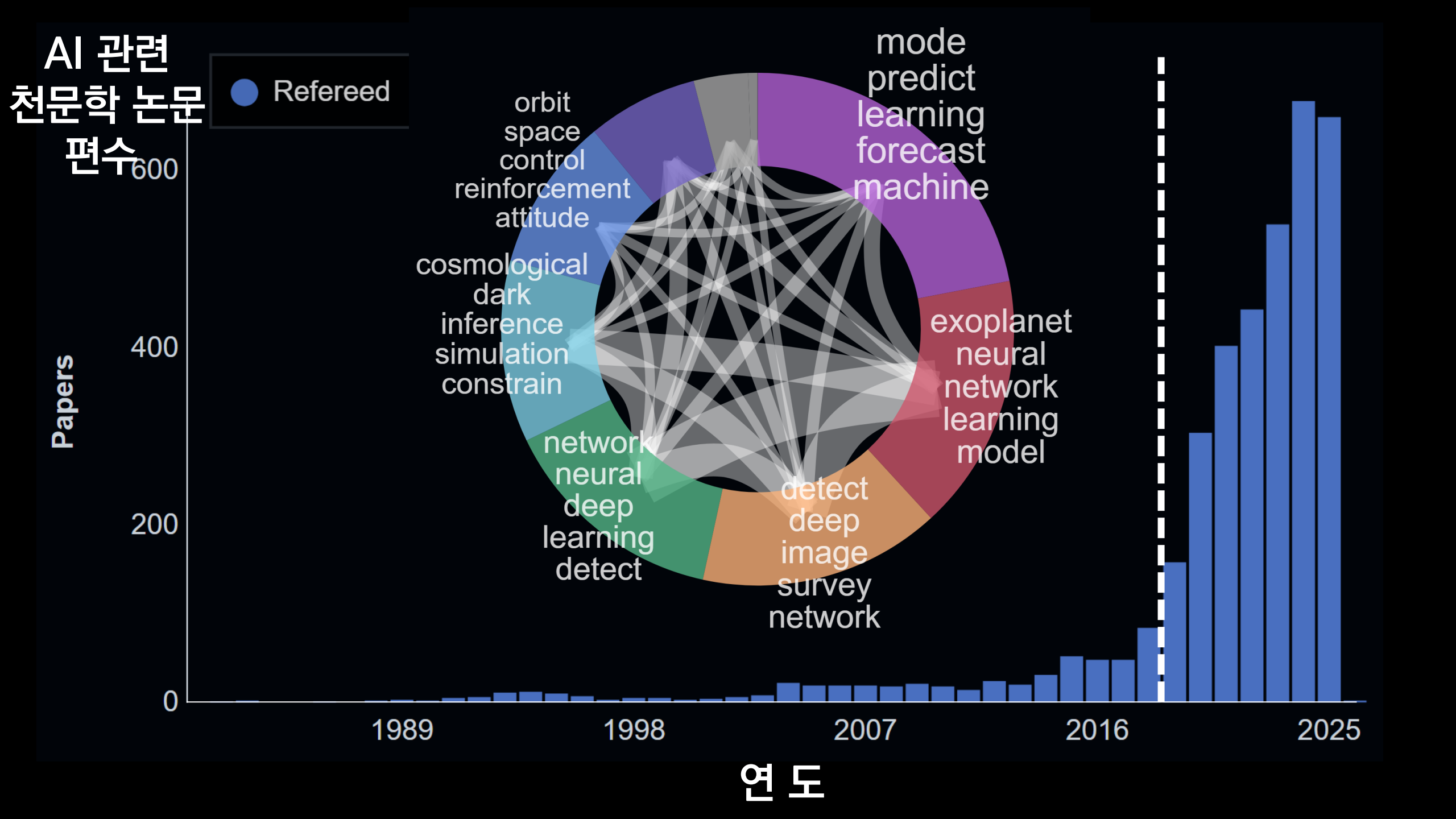
2003년

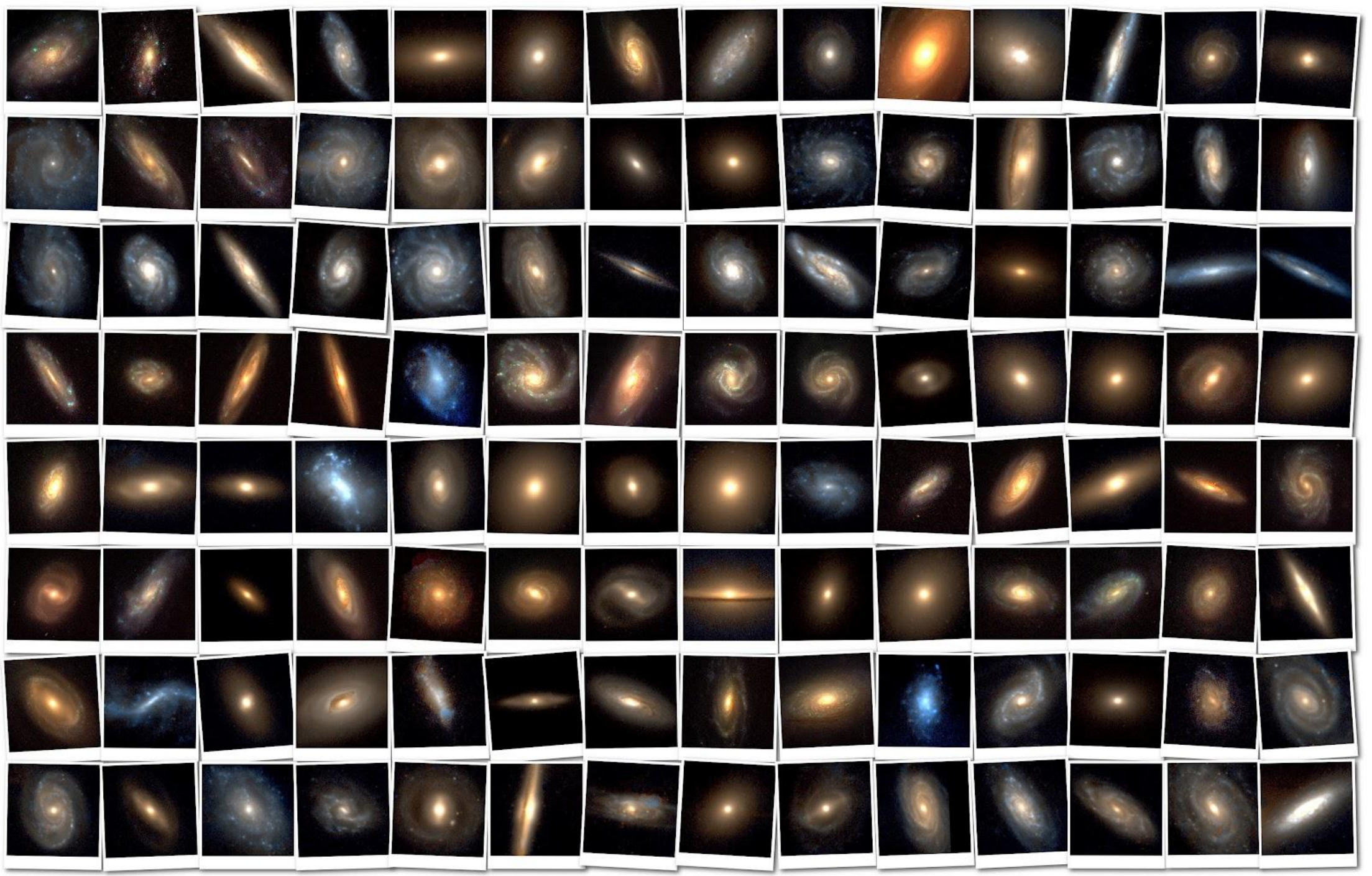
110억년

DESI

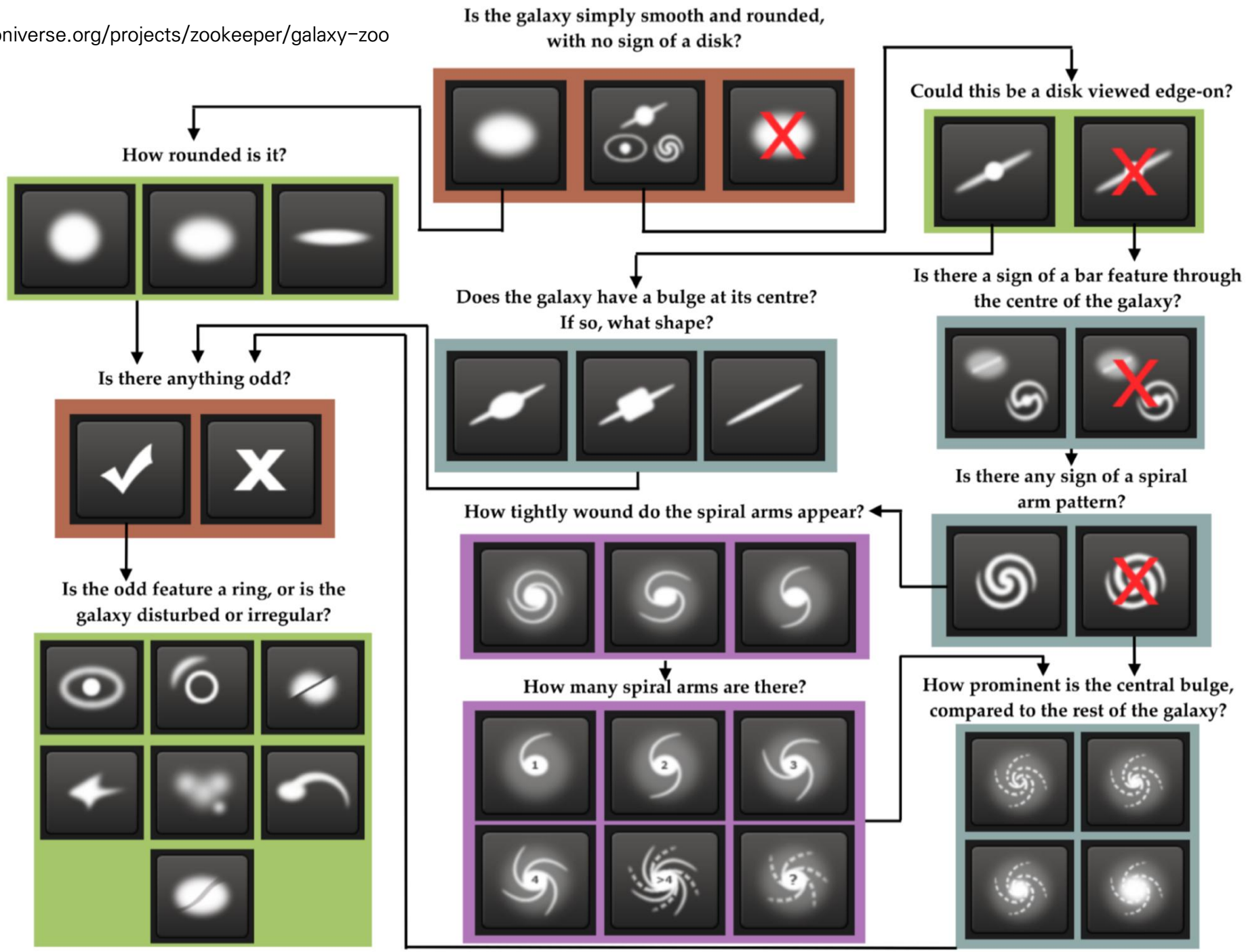
2025년

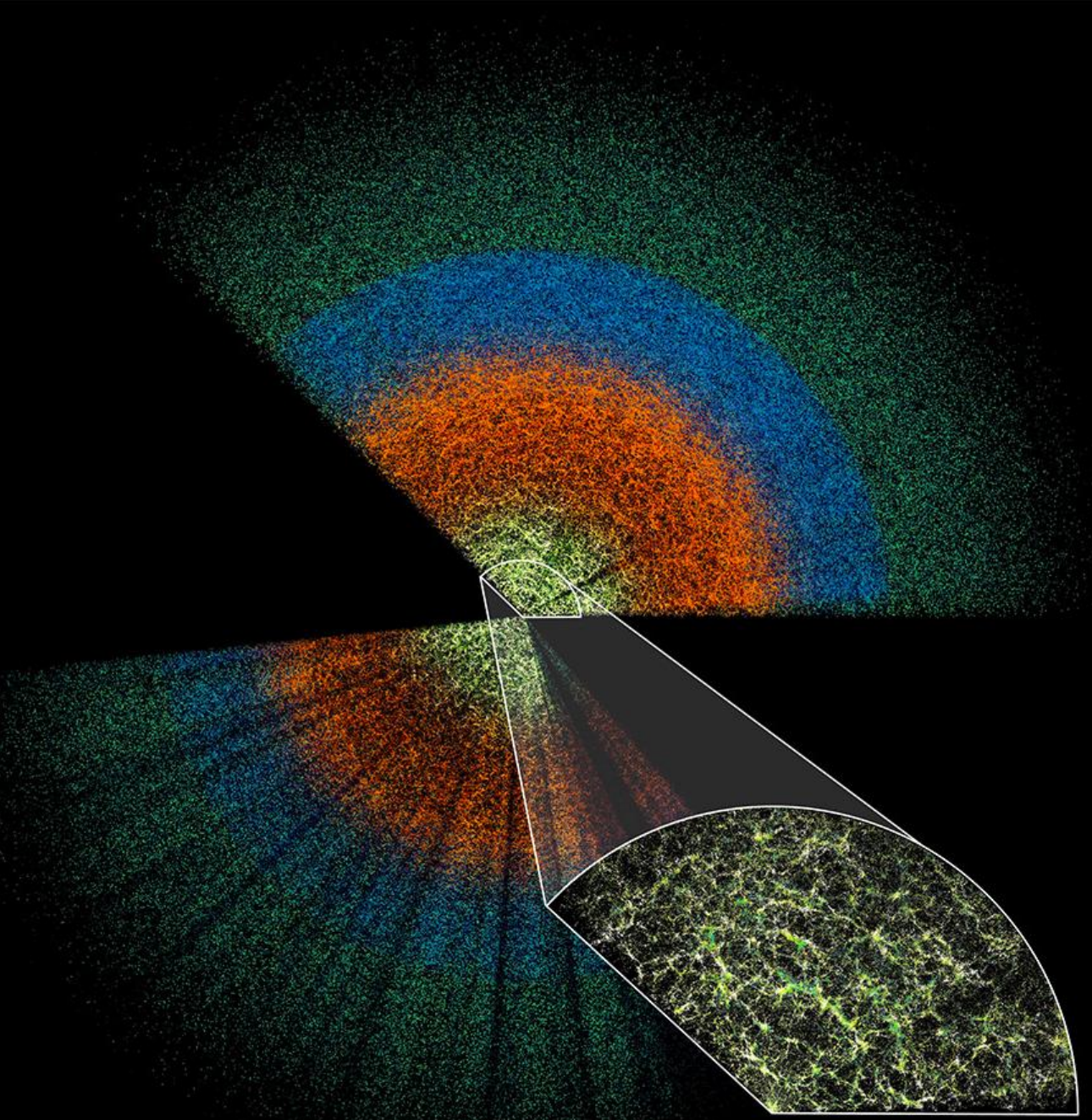



AI 관련
천문학 논문
편수
Refereed
Papers
600
400
200
0
1989
1998
2007
2016
2025
연 도
mode
predict
learning
forecast
machine
exoplanet
neural
network
learning
model
detect
deep
image
survey
network
network
neural
deep
learning
detect
cosmological
dark
inference
simulation
constrain
orbit
space
control
reinforcement
attitude



https://www.zooniverse.org/projects/zookeeper/galaxy-zoo

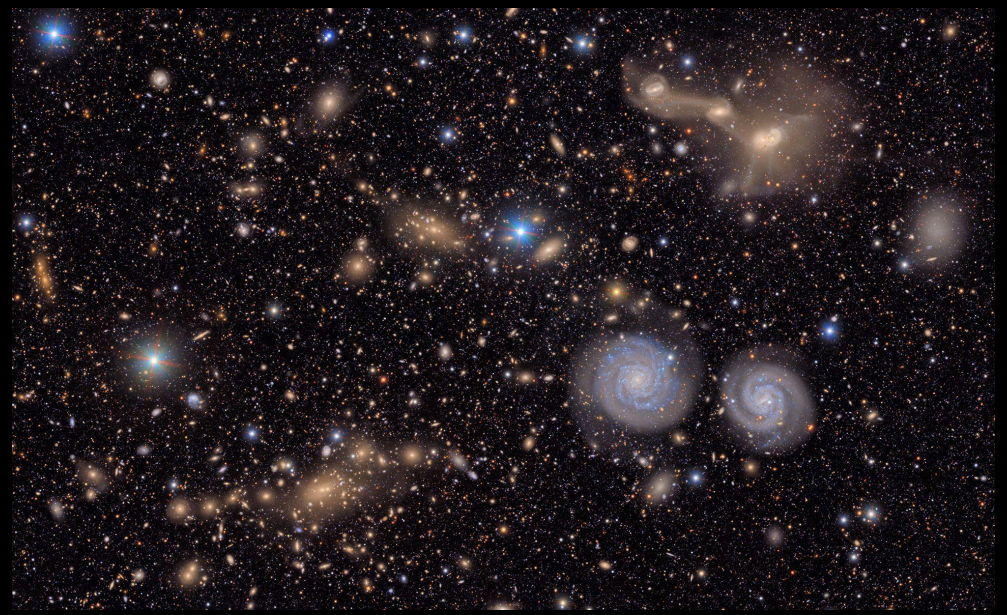


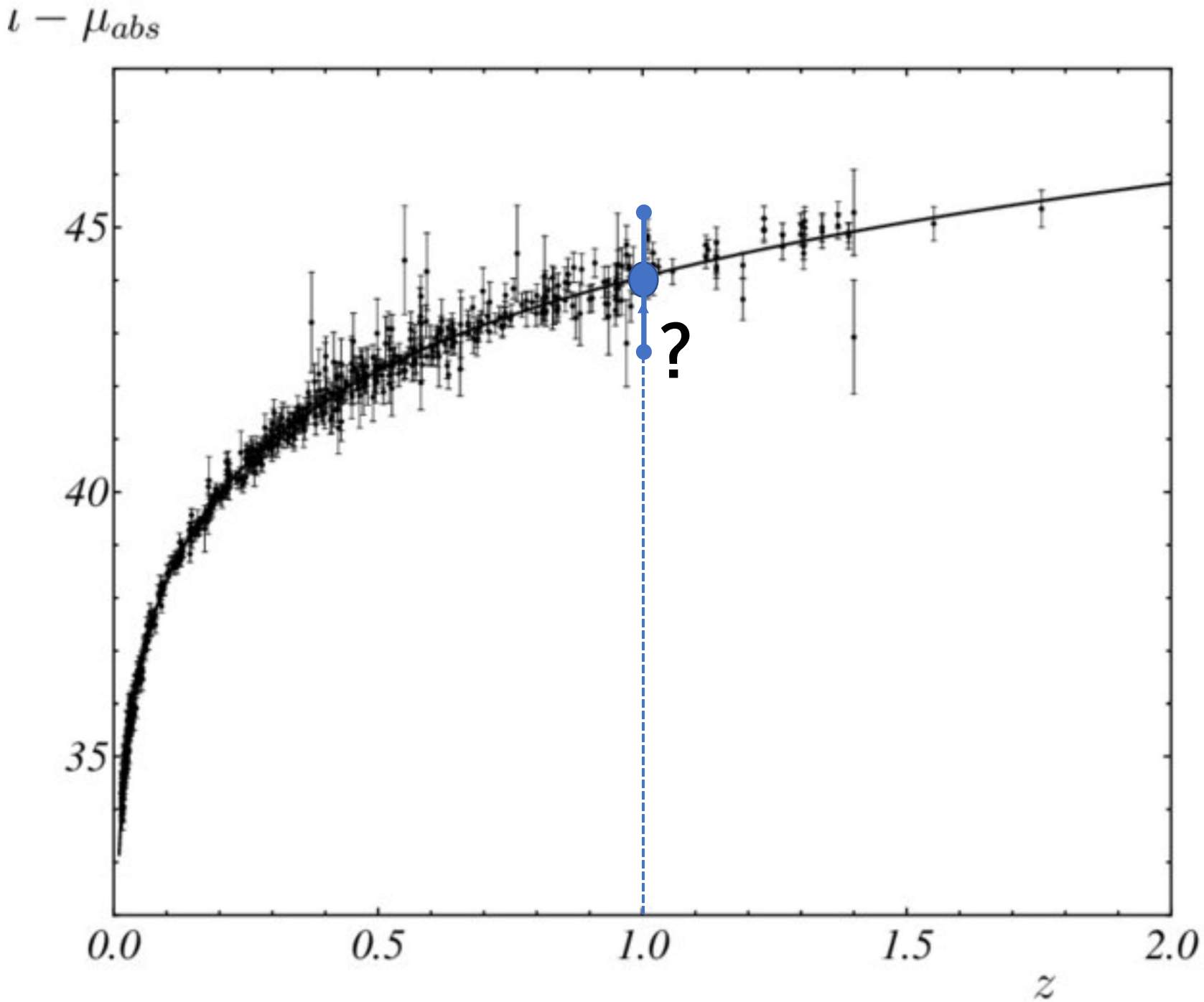


별과 먼 은하 구분

은하의 거리 측정

- 원래는 스펙트럼이 필요
- 영상 자료 밖에 없다면?





결과에 대한
통계적 신뢰도를
알기 어려움

Machine Learning

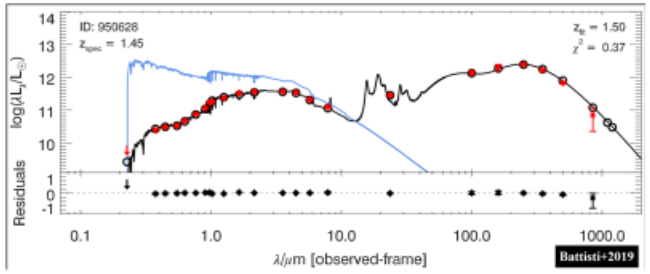
- All results should be lead by physics
  - Software is no replacement for physical understanding
- ML is a useful technique in three cases:
  - No known or understandable physical mechanism exists:
    - Dark energy
  - Physics known, but computing the required quantities is challenging:
    - Photometric redshifts
    - Large numbers of simulations - non-linear emulators
  - Paradigm testing: looking for things not predicted by any physics or models
    - Unknown unknown
    - Model-independent tests

Techniques (examples):

1. Classification/catagoriszation
   - Support Vector Machines
   - Naive Bayes
   - Perceptron/CNN
   - Decision Trees
2. Clustering
   - Hierarchical
   - k-means
   - DBSCAN
   - Density
3. Interpolation/Regression
   - Gaussian Process Regression
   - Kernel Ridge Regression
   - Multivariate adaptive regression splines

전통적인 학습에서는...

파장	밝기
X.XXXXXX	X.XXXXXX
X.XXXXXX	X.XXXXXX
X.XXXXXX	X.XXXXXX
X.XXXXXX	X.XXXXXX
X.XXXXXX	X.XXXXXX
...	...

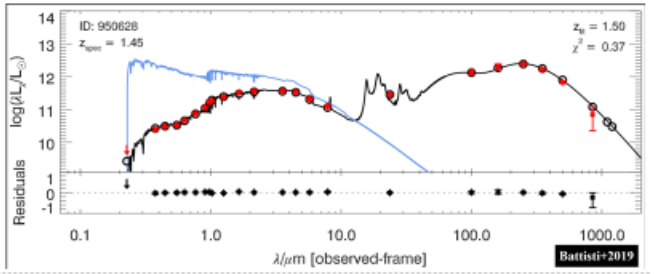


이론이나 경험에서
피팅 함수를 설정함

천체 종류, 적색이동, ...

머신 러닝(기계학습)에서는...

파장	밝기
X.XXXXXX	X.XXXXXX
X.XXXXXX	X.XXXXXX
X.XXXXXX	X.XXXXXX
X.XXXXXX	X.XXXXXX
X.XXXXXX	X.XXXXXX
...	...

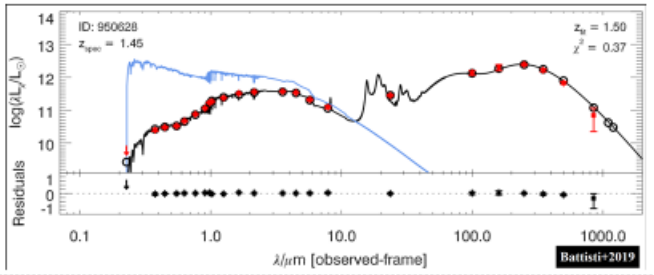


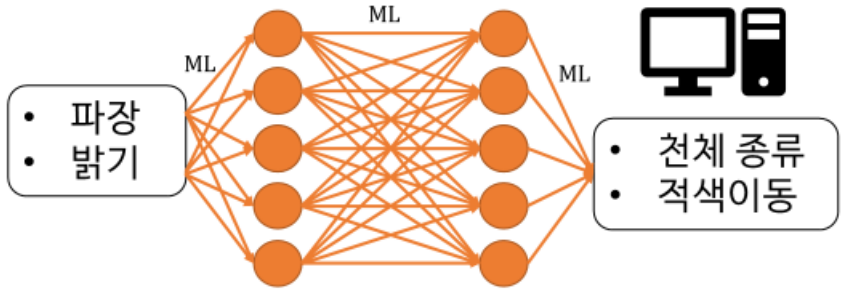
학습 전략을 설정함

천체 종류, 적색이동, ...

딥러닝(심층학습)에서는...

파장	밝기
X.XXXXXX	X.XXXXXX
X.XXXXXX	X.XXXXXX
X.XXXXXX	X.XXXXXX
X.XXXXXX	X.XXXXXX
X.XXXXXX	X.XXXXXX
...	...



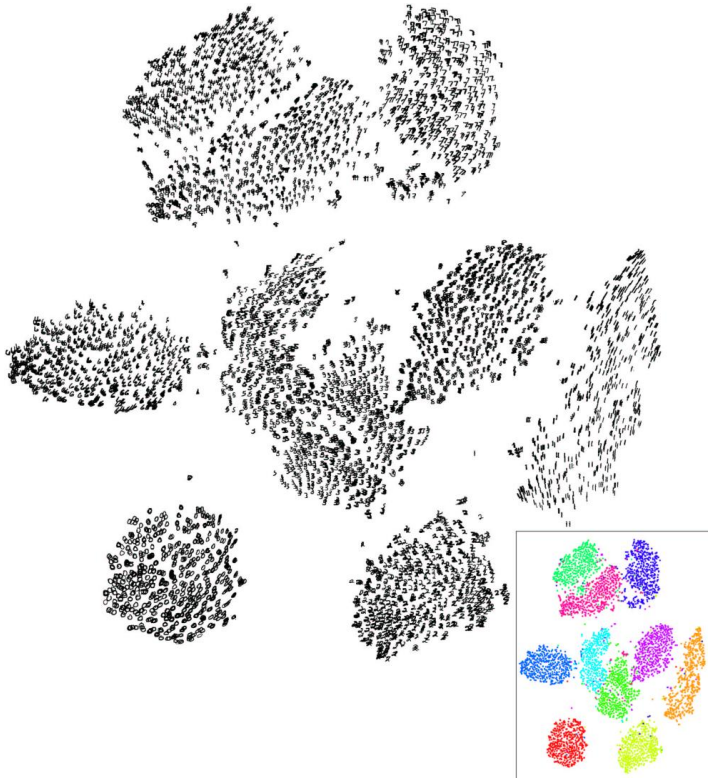


해석 가능한 머신 러닝
(Interpretable Machine Learning)

conv6

conv9

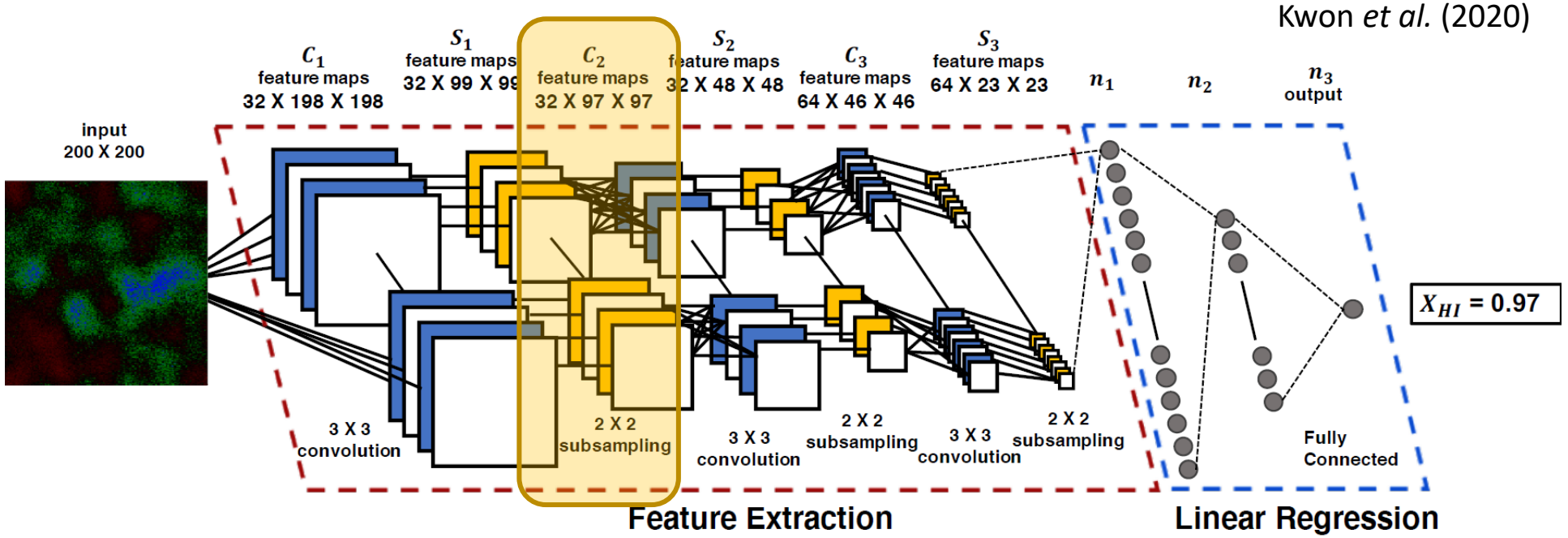
Maximally Activating Images

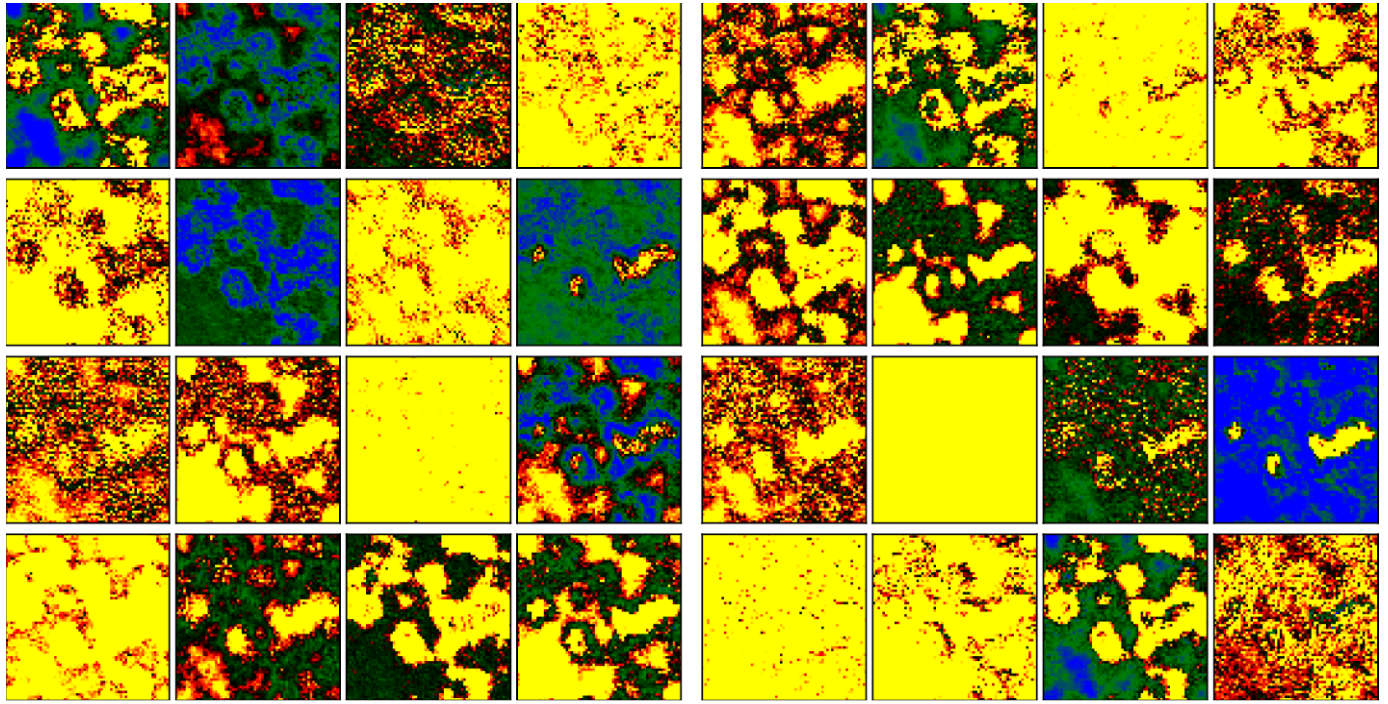


Dataset Visualization

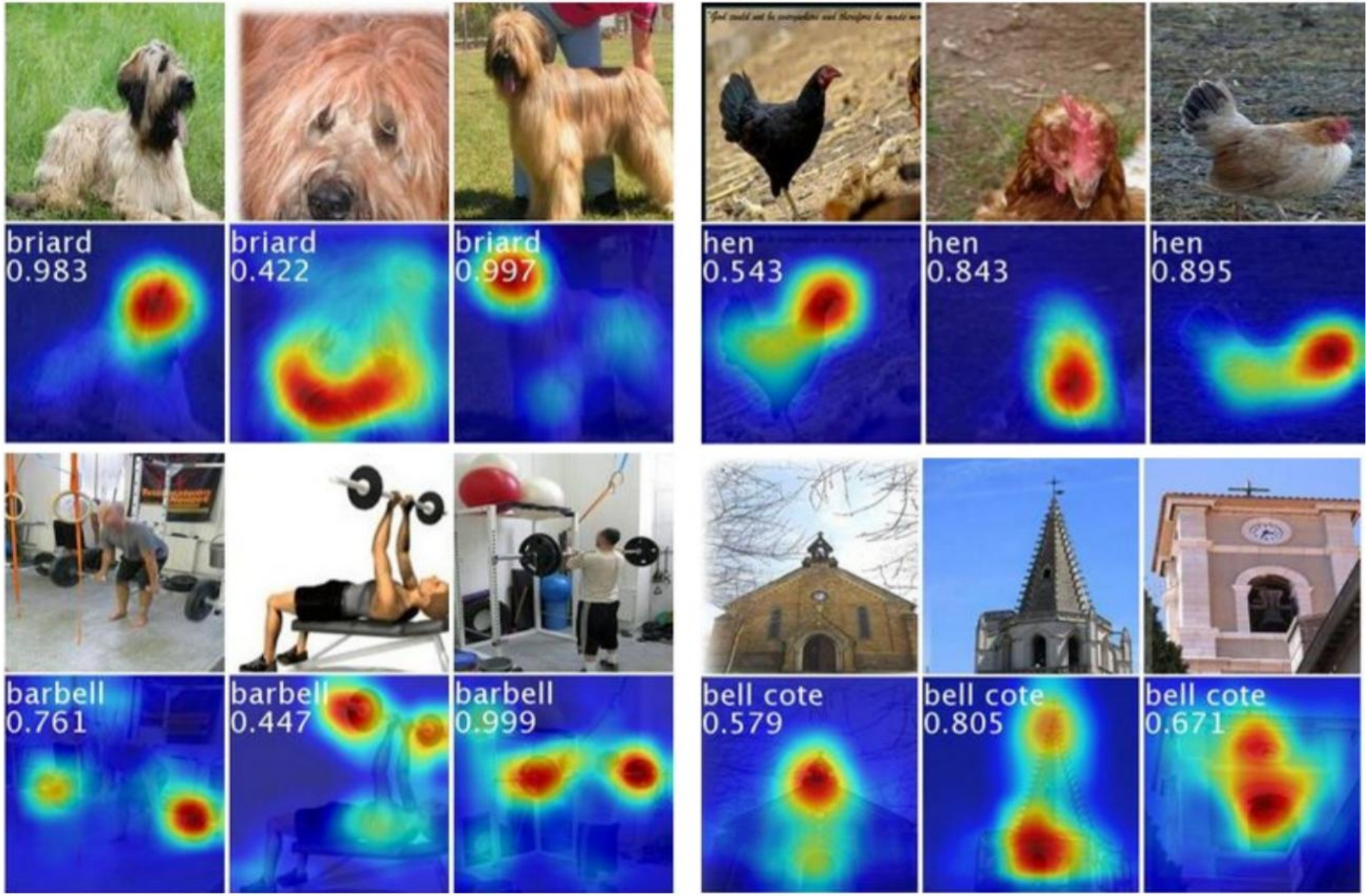
https://www.cognex.com/ko-kr/blogs/deep-learning/research/overview-interpretable-machine-learning-2-interpreting-deep-learning-models-image-recognition

Kwon *et al.* (2020)



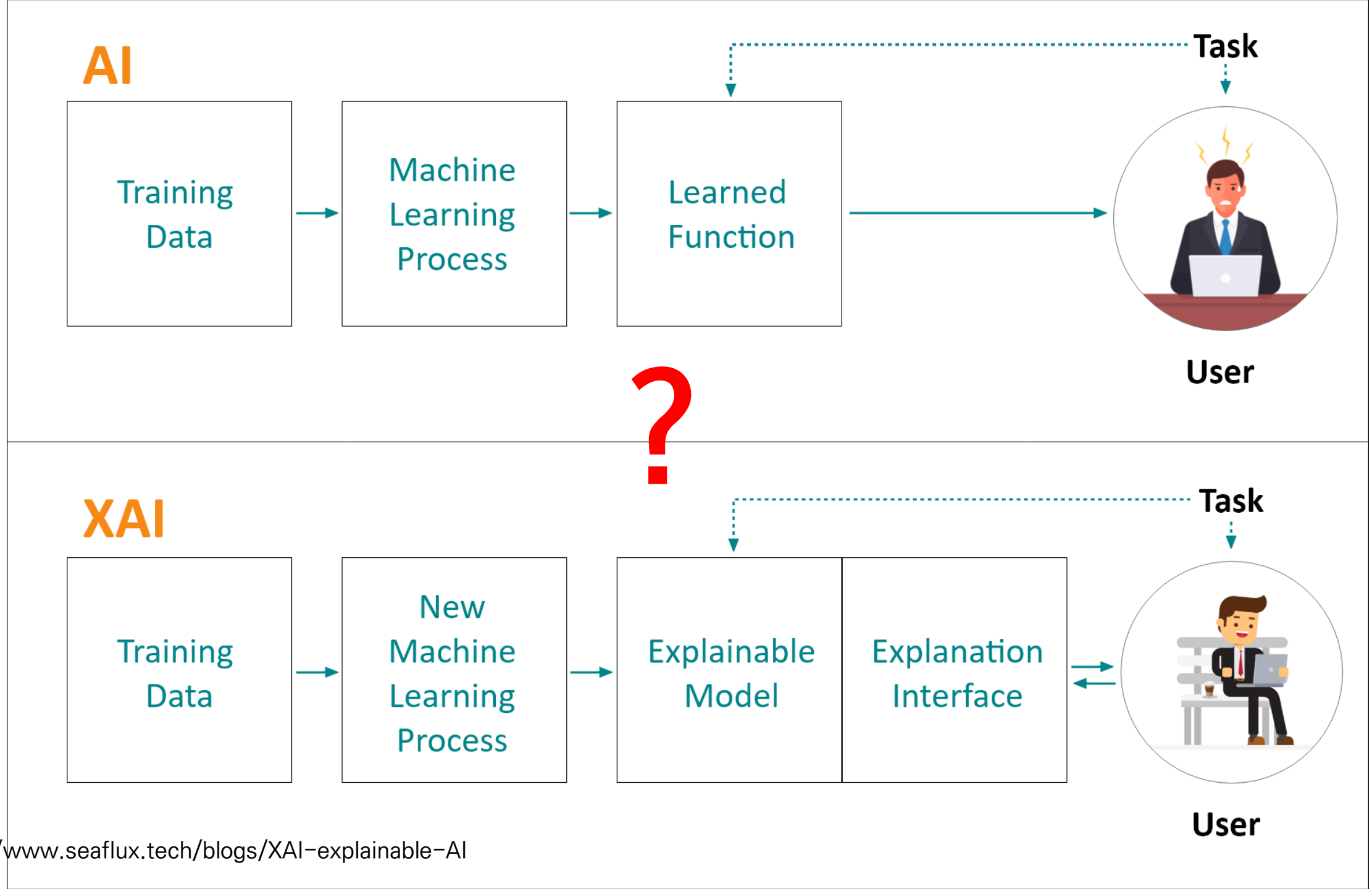


Activation Visualization



Class Activation Map

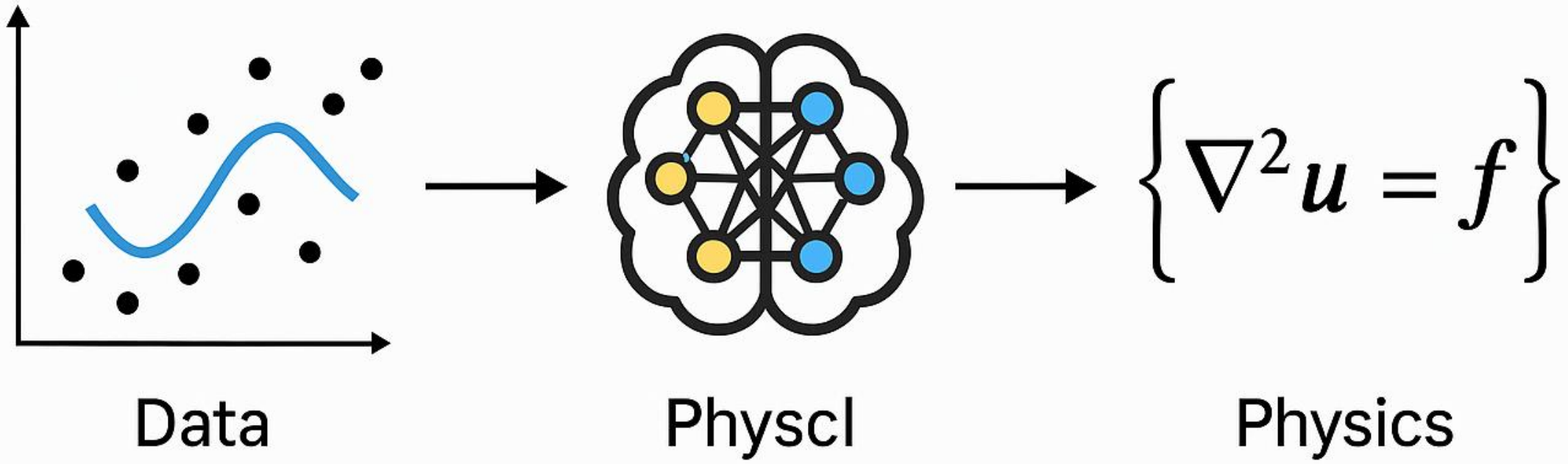
Russakovsky *et al.* (2015)



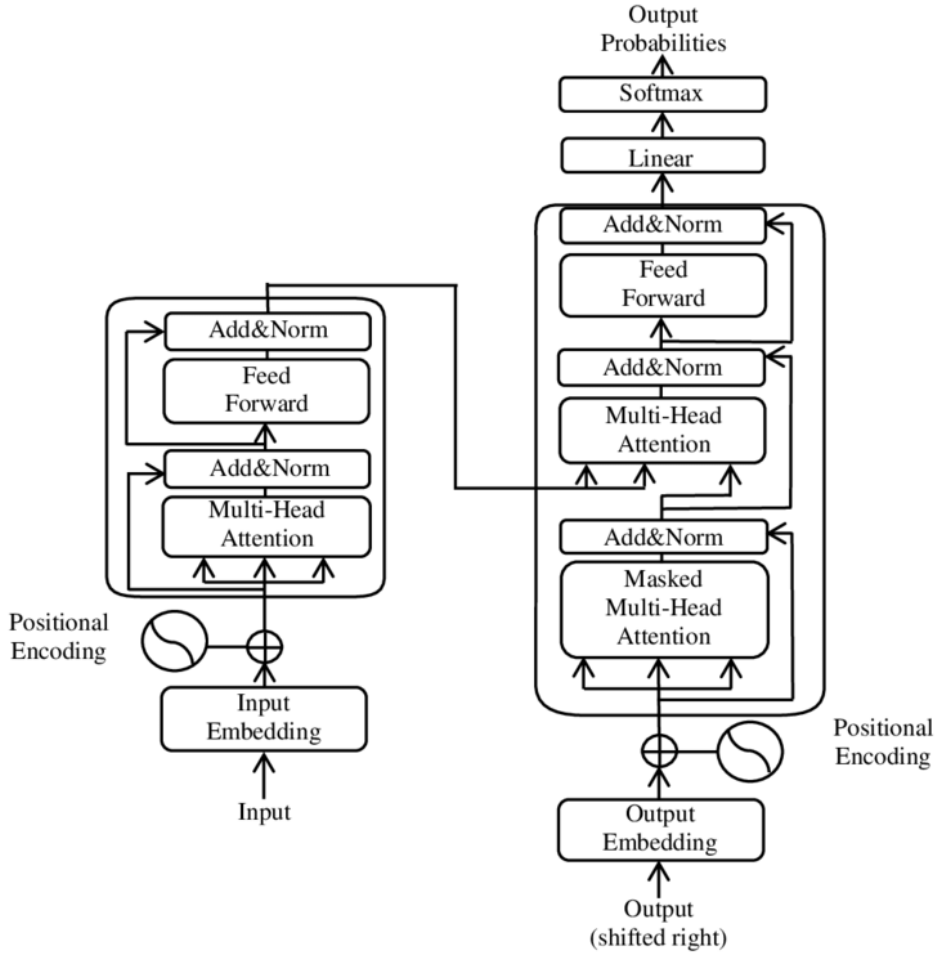
https://www.seaflux.tech/blogs/XAI-explainable-AI

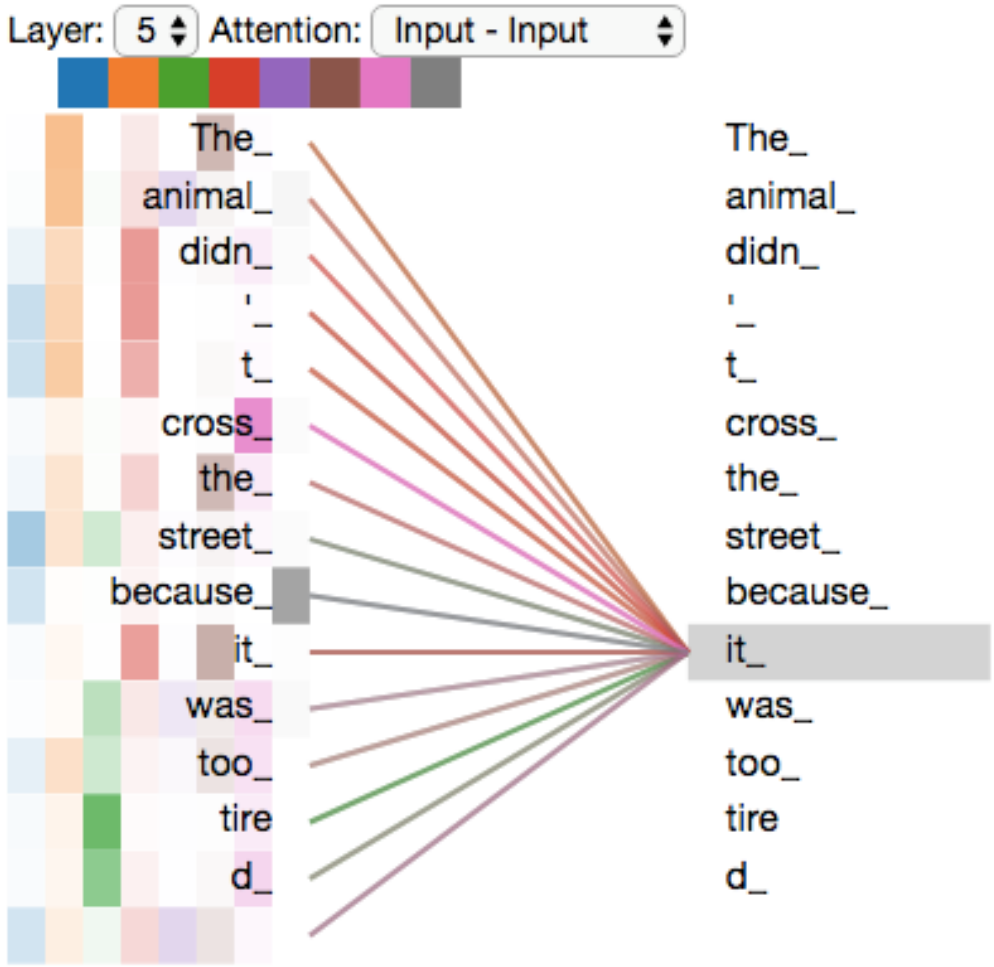
Physics-Informed Artificial Intelligence

?



A brief history... in 2017





Attention Is All You Need

Transformer model → GPT