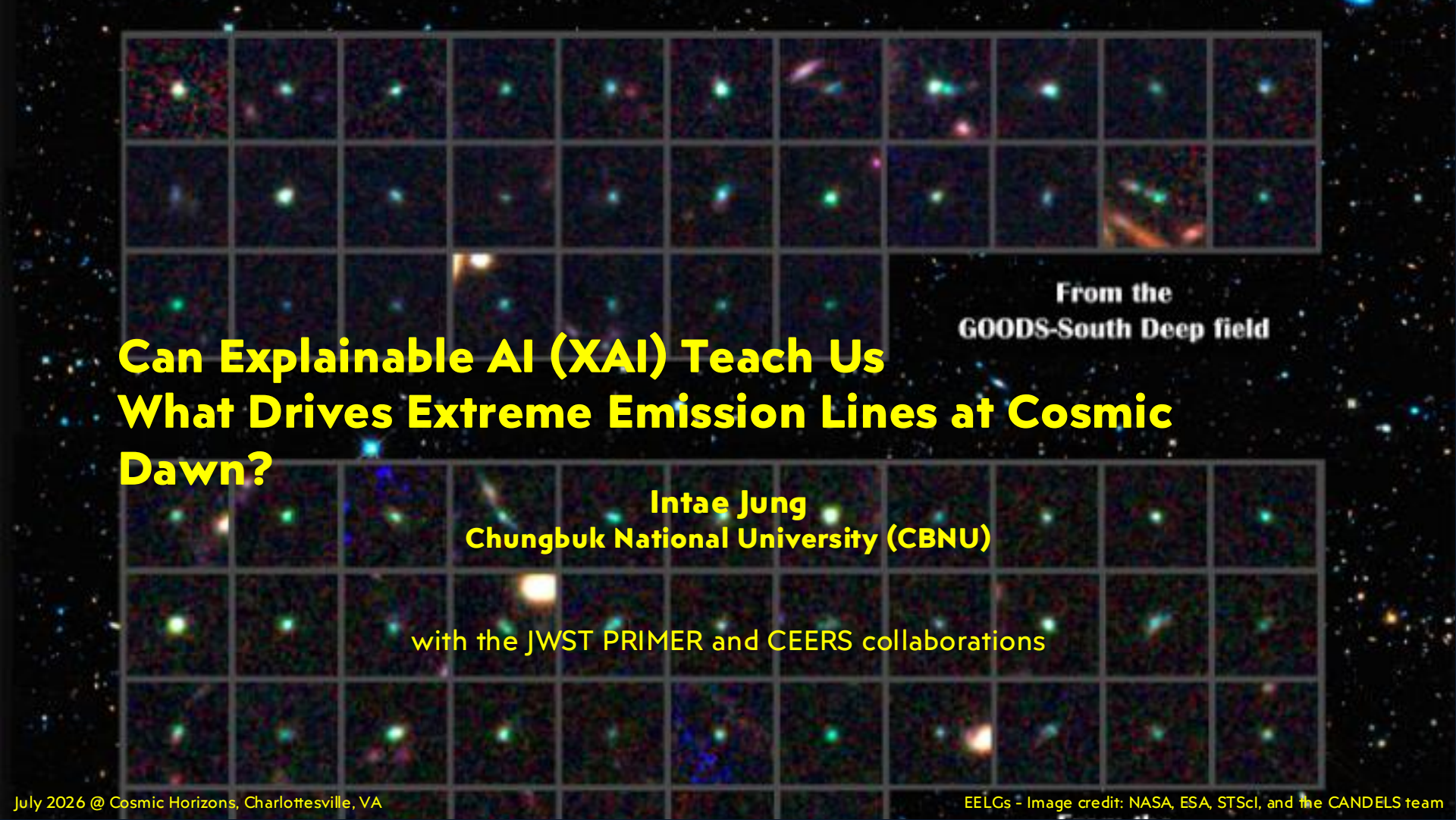


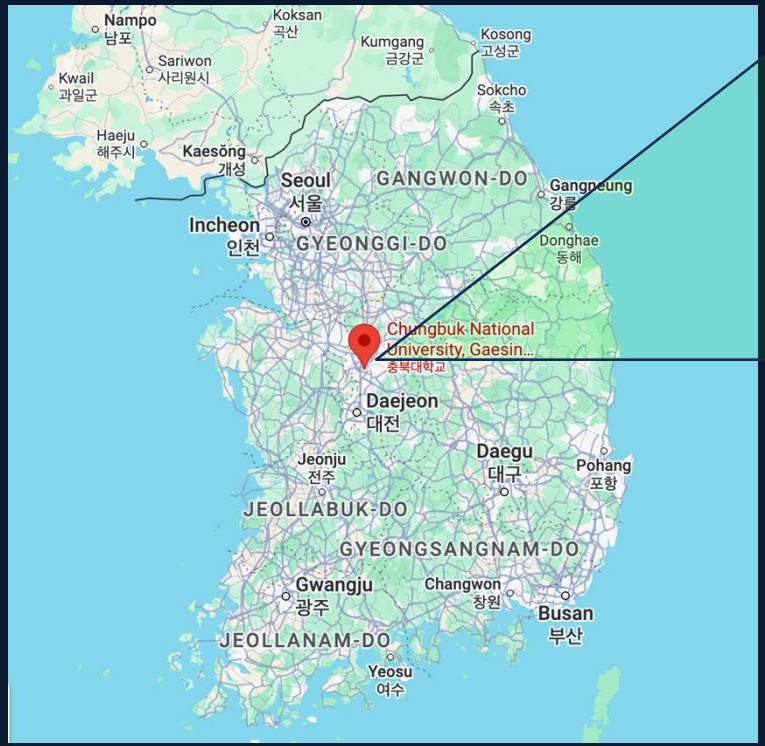
Can Explainable AI (XAI) Teach Us What Drives Extreme Emission Lines at Cosmic Dawn?

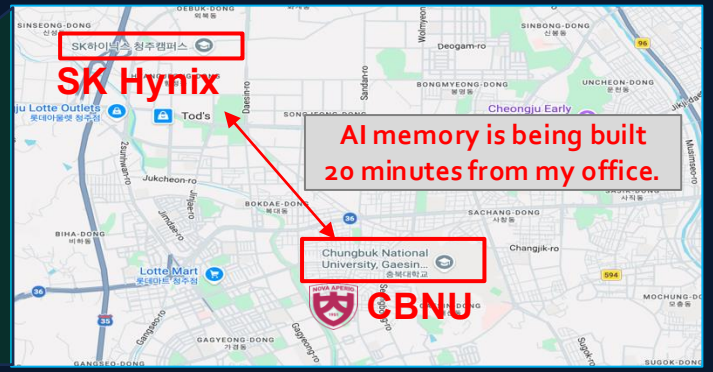
Intae Jung
Chungbuk National University (CBNU)

with the JWST PRIMER and CEERS collaborations

EELGs - Image credit: NASA, ESA, STScI, and the CANDELS team

Why do we care about AI?



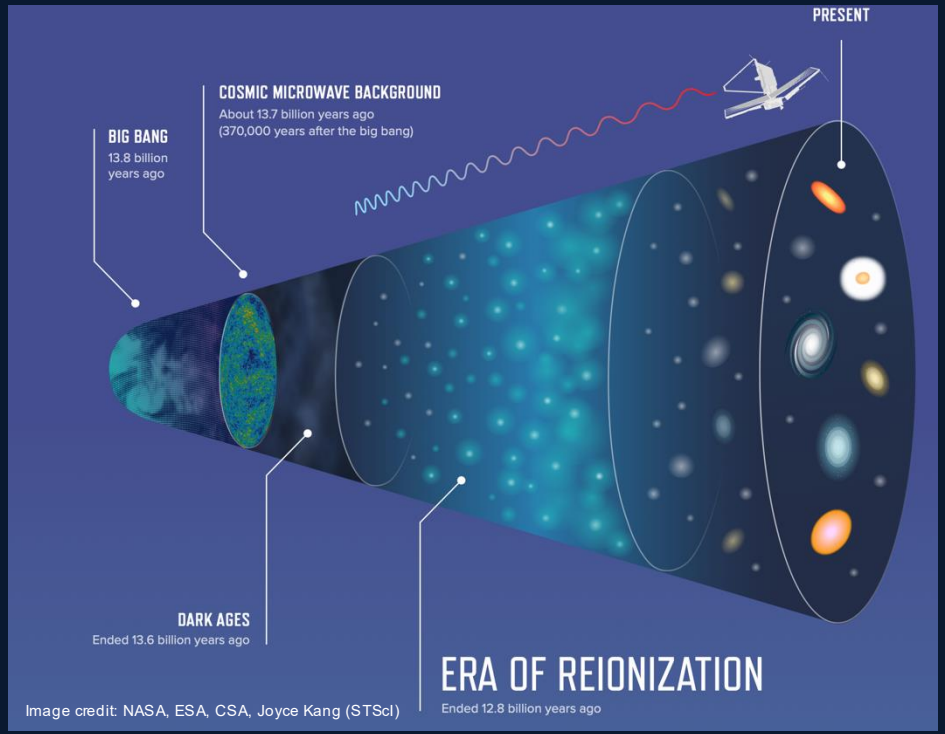


SK Hynix Stock Price over the year (~1000% ↑)

I may be late to the AI stock market, but hopefully not too late to use AI for astronomy and my reionization science...

Reionization in Cosmic History

From the Big Bang to Reionization



The Epoch of Reionization:

- The last major phase transition in the IGM.
- The era of first luminous sources.

First stars,

First star clusters,

First galaxies,

First protoclusters,

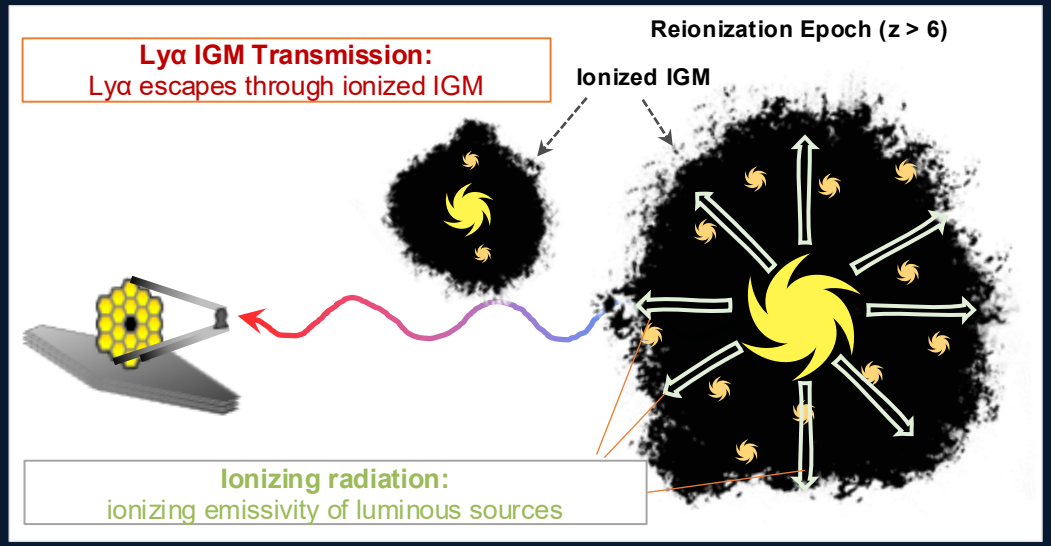
First ionized bubbles,

and

First black holes.

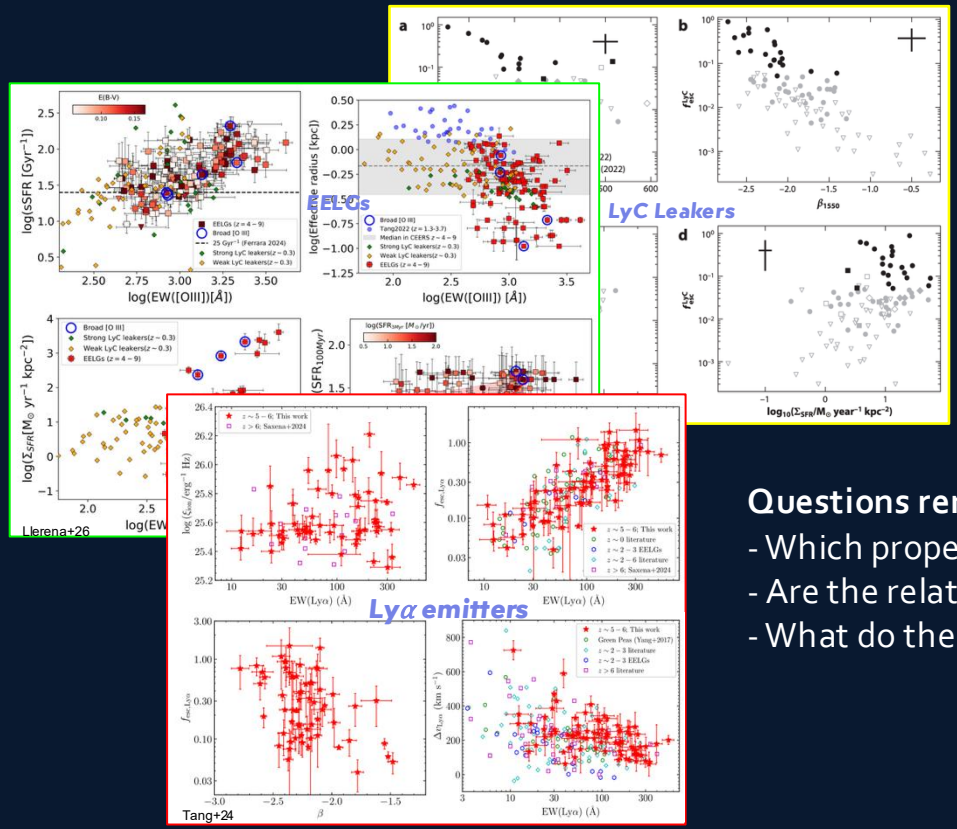
Physical Components of Reionization

Ionization state of the IGM ←→ Ionizing sources



Lyα emitters: key to probing ionization state of the IGM
LyC leakers: key to understanding the ionizing photon budget
EELGs: likely to be strong LyC leakers, the dominant source of ionizing photons

Many pairwise correlations, but no joint explanation

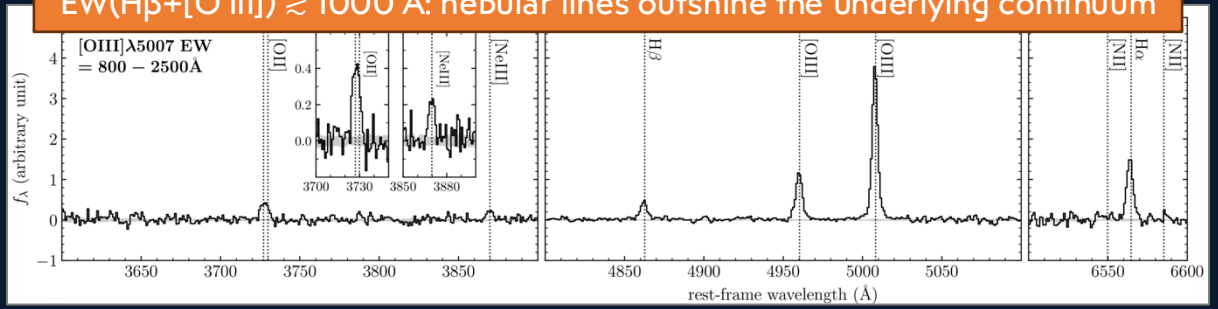


Questions remaining:

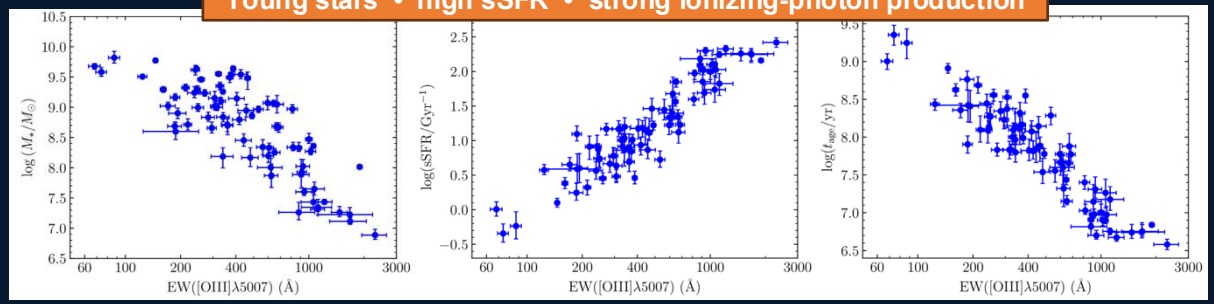
- Which properties add information after conditioning on the others?
- Are the relationships nonlinear or population-dependent?
- What do the unexplained extremes tell us is missing?

Application: Extreme Emission Line Galaxies (EELGs)

EW(Hβ+[O III]) ≳ 1000 Å: nebular lines outshine the underlying continuum



Young stars • high sSFR • strong ionizing-photon production

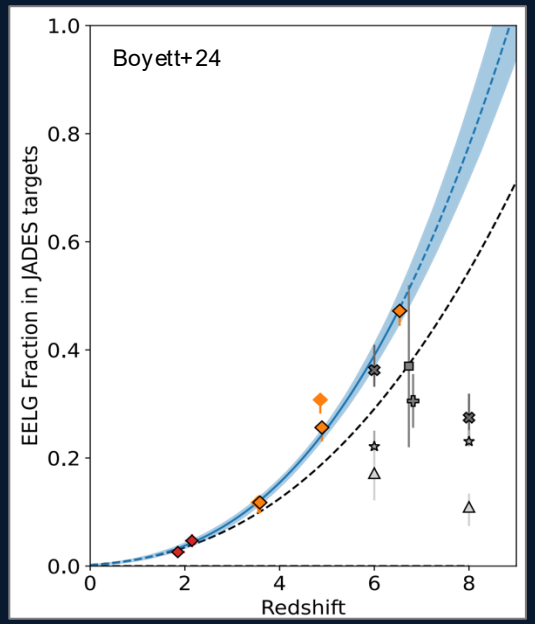


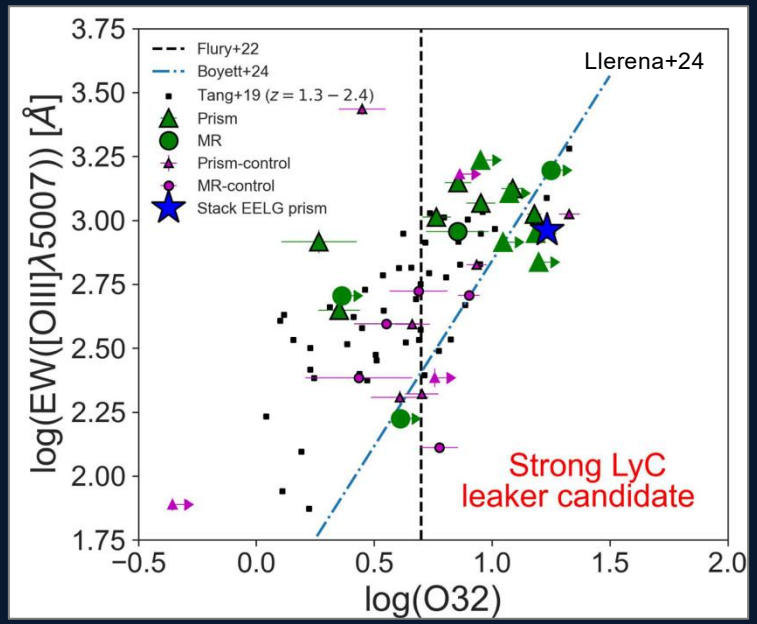
Tang+2019

+ Green Peas, Blueberries, Purple Grapes … (e.g., Cardamone+09, Amorin+10, Izotov+11, van der Wel +11, Yang+16,17, …)

EELGs connect bursty growth to reionization

Promising LyC Leaker Candidates during the Epoch of Cosmic Reionization



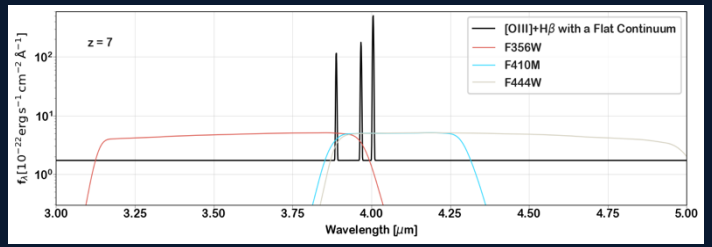


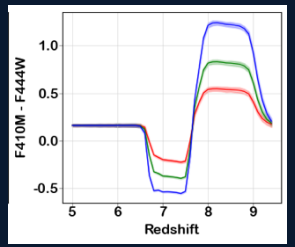
- JWST shows their fraction increases rapidly toward high redshift (left).
- These galaxies may play a key role in cosmic reionization (right).

F410M turns line excess into a statistical sample

Selection #1: Photometric z ~ 7 selection
Phot-z criteria + continuum detection criteria
*Similar to Finkelstein et al. (2024)

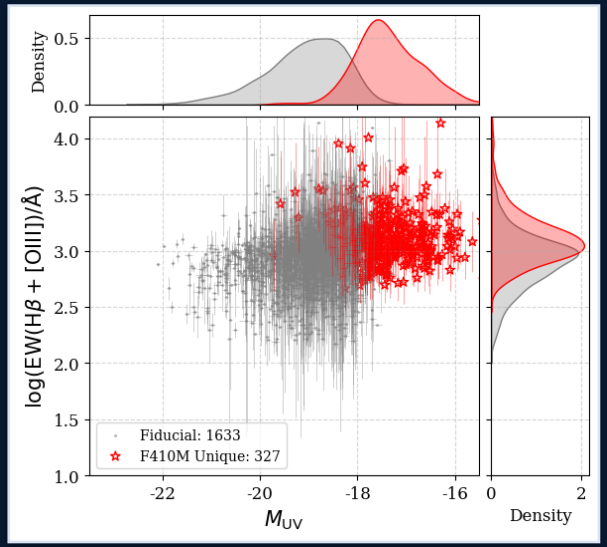
Selection #2: via F410M excess
Phot-z criteria + F410M excess criteria (Bunker+95, Sobral+13)





*Hβ+[OIII] equivalent widths from F410M flux excess

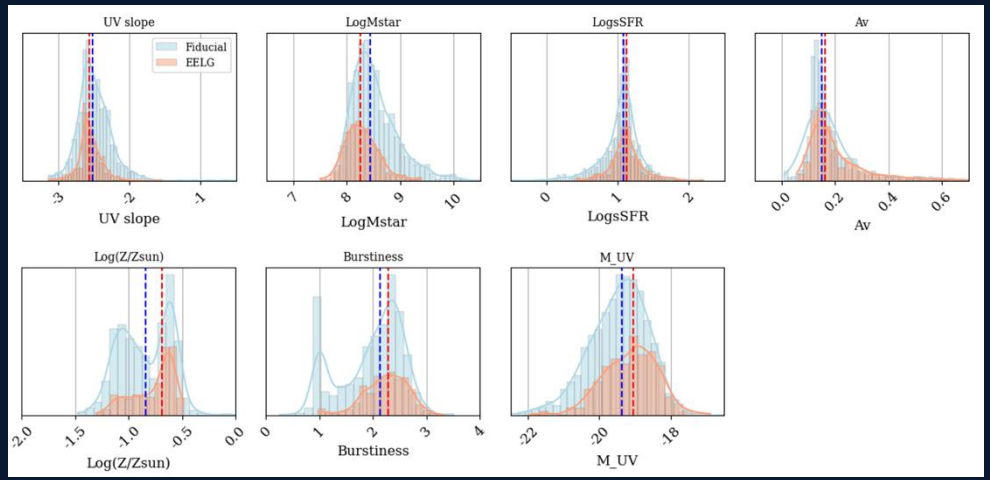
$$\mathrm{EW}_{\mathrm{rest}} = \Delta\lambda_{\mathrm{MB}} \cdot \frac{f_{\mathrm{MB}} - f_{\mathrm{BB}}}{f_{\mathrm{BB}} - f_{\mathrm{BB}}(\Delta\lambda_{\mathrm{MB}}/\Delta\lambda_{\mathrm{BB}})} \cdot \frac{1}{1+z}$$



1,960 galaxies at z~7
816 EELGs > 1000 Å (57 > 3000 Å)

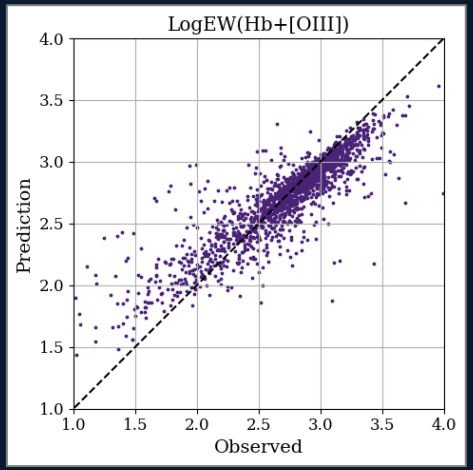
z~7 EELGs: Random Forest Models

Physical Properties of z~7 EELGs



Random Forest models

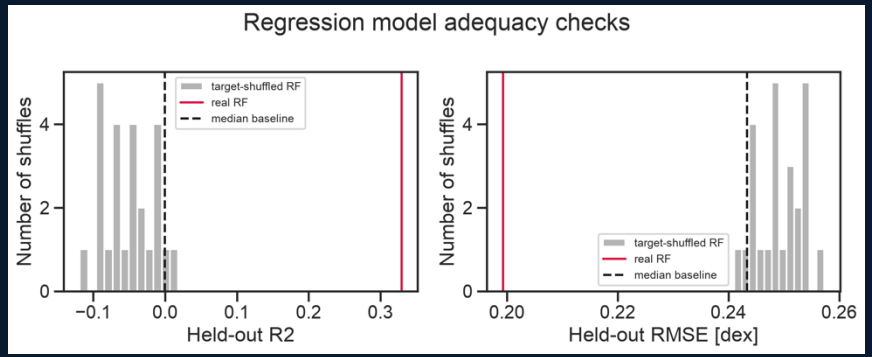
- $P(EW_{H\beta+[OIII]} \mid \boldsymbol{\eta}, M_{UV}, \beta, sSFR, M_{*}, A_{v}, ...)$



***A flexible model can learn the joint relation learns the joint relation. However, prediction is not yet physical understanding.**

Predict:
Can the measured properties recover emission-line strength?

1) Validate



2) Then, Explain

Which features drive the model globally and for individual galaxies?

3) And Test

Do the conclusions survive covariance, ablation, model choice, and measurement coupling?

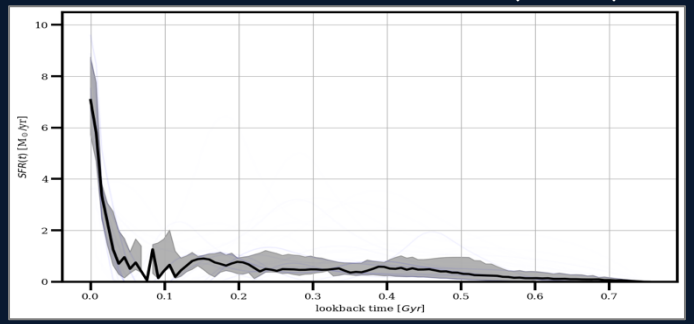
XAI becomes scientifically useful only after prediction quality and interpretation robustness are tested.

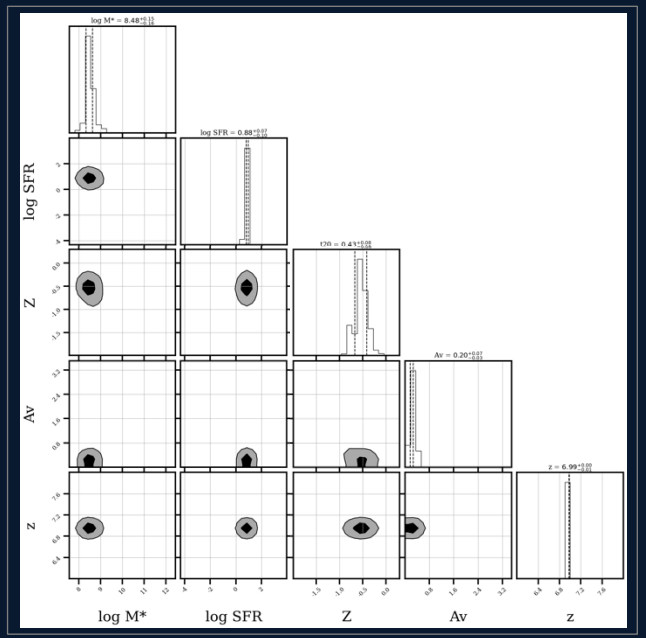
The predictors are inferred, not directly observed

Galaxies Properties from SED fitting (Iyer & Gawiser 2017)

- Recent SFH: burstiness, sSFR
- Galaxy scale: M_*, M_{UV}
- Population state: β_{UV}, A_V, Metallicity
- Environment: neighbor distance, pair flag

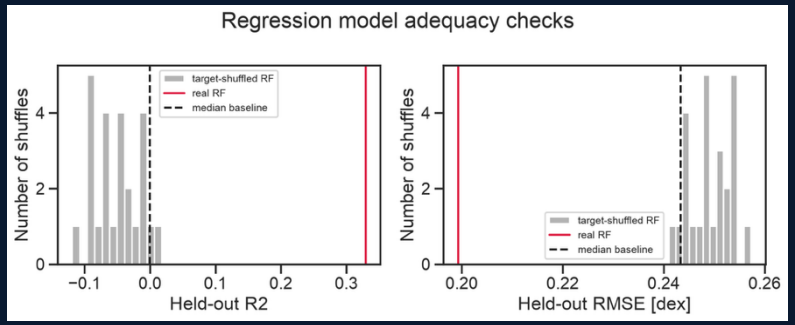
Star Formation Histories (SFHs)





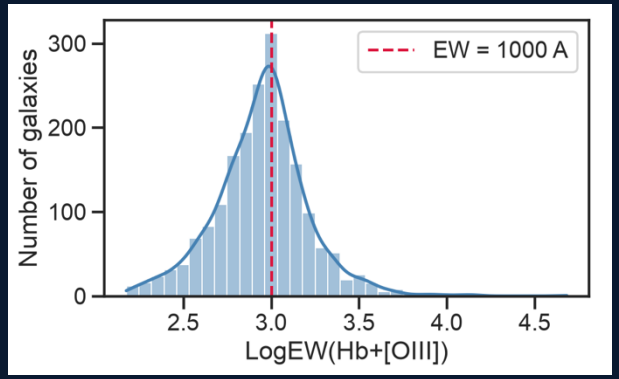
The model learns population-level signal, not precise EWs

Held-out Random Forest regression



Population checks

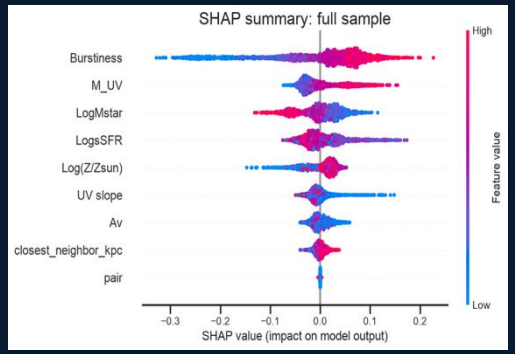
Better than baselines; EELG ROC–AUC = 0.731.
Use rankings and trends—not precision predictions.

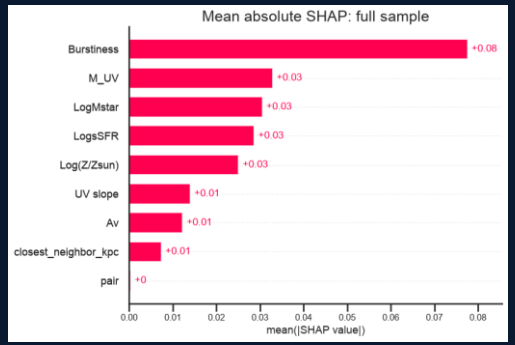


Use the model to generate population-level hypotheses, not to claim an exact EW for an individual galaxy.

Burstiness is the clearest individual predictor

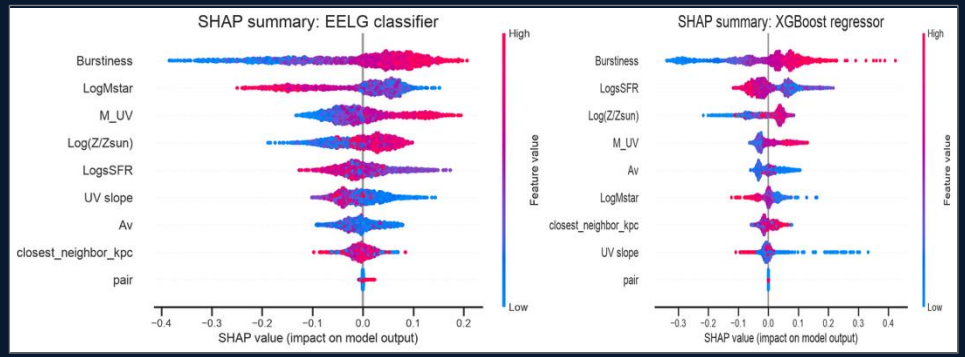
RF regressor: continuous EW





Cross-model check

Across regression and classification, SHAP ranks burstiness first. SHAP explains model behavior; it does not establish causality.

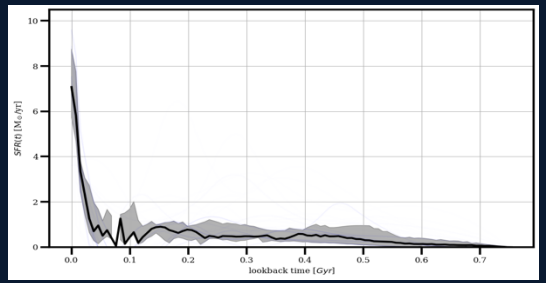


MUV, stellar mass, sSFR, and metallicity add secondary structure; present environment proxies add little.

*SHAP identifies what the model uses; it does not establish causality.

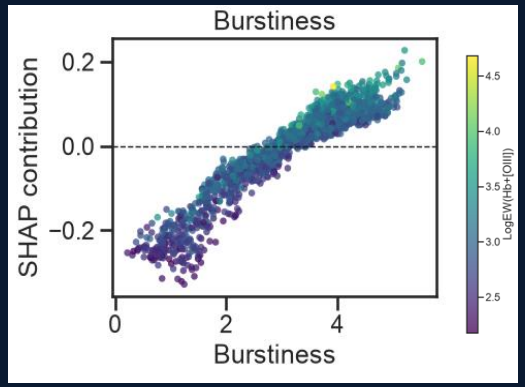
The learned pathway is bursty, low-mass, and nonlinear

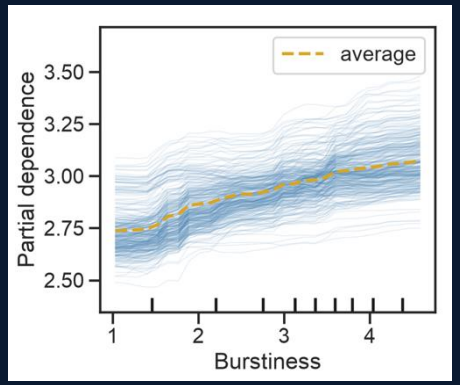
Recent SFH: SFR10 Myr / SFR100 Myr



Directionality

High burstiness raises predicted EW most clearly. Mass and M_{UV} modulate the response; sSFR and metallicity are informative but nonlinear.

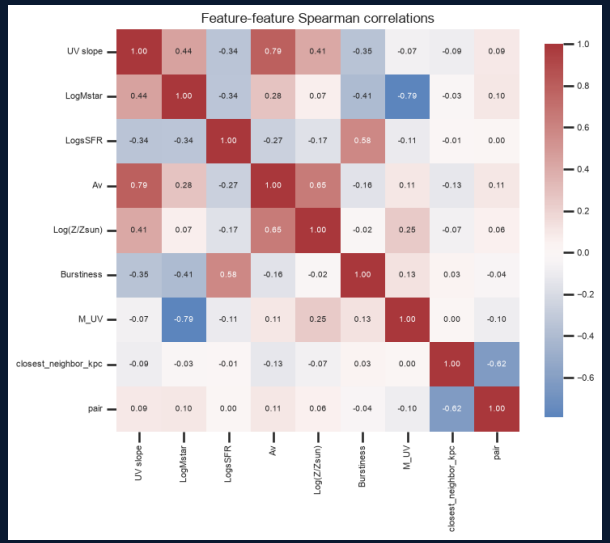




PDP supports a recent-burst pathway while showing that the same feature change need not affect every galaxy identically.

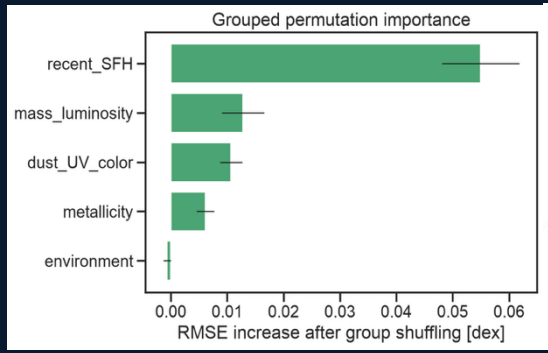
Grouped tests protect the result from covariance

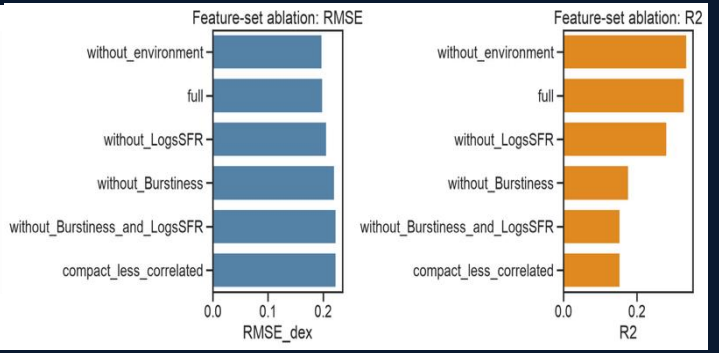
Feature covariance is physical—not a nuisance to ignore



Stress tests

Recent SFH is the dominant feature family. Removing burstiness hurts more than removing sSFR; environment adds little in the current feature set.

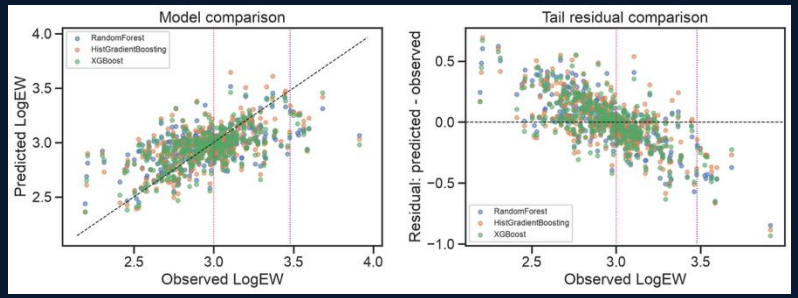




The robust claim is about recent star-formation history as a family, with burstiness as its clearest individual axis.

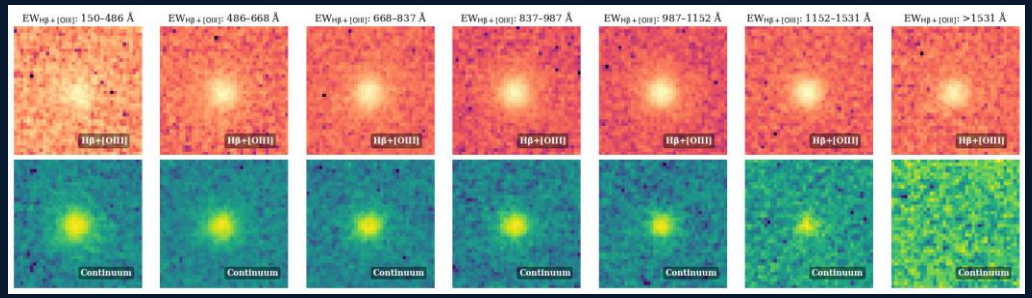
The extreme tail reveals missing observables

RF compresses EW > 3000 Å



What to add next

Burst age / ionization state / line ratios / compact morphology / independent SFH constraints



Boosting reduces the bias but does not solve the extreme tail; failure points toward missing observables, not merely a different algorithm.

XAI tells us what to test next

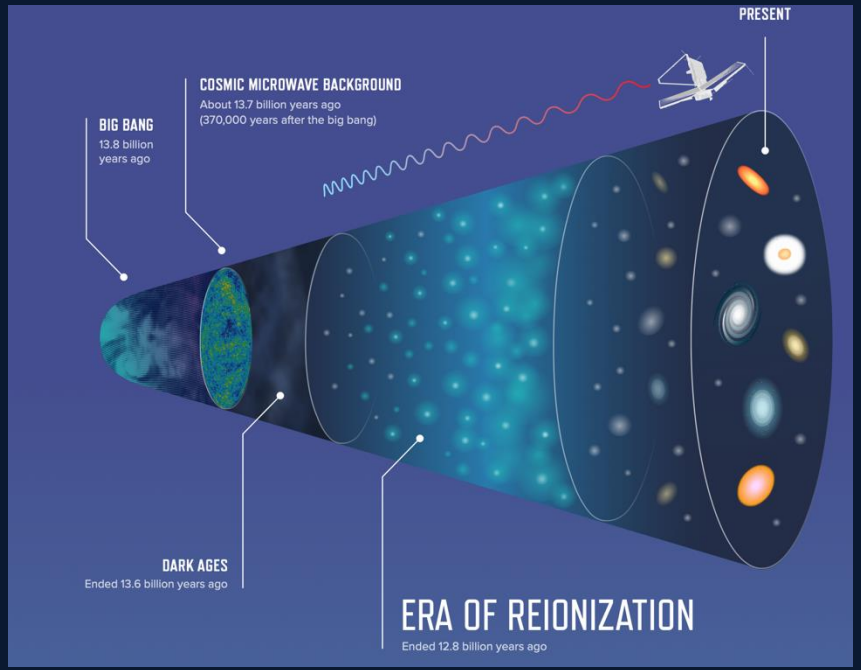


Image credit: NASA, ESA, CSA, Joyce Kang (STScI)

What we can claim

- The model learns real EW signal.
- Recent SFH contains the most stable predictive information.
- Burstiness is the clearest individual predictor.
- Extreme failures point toward missing observables.

What XAI adds

It does not prove the cause;
it tells us what to test next.