Facilitating, not replacing:

The role of AI & ML in crafting *tools* for astronomy research

Kartheik Iyer @ COPAG session (AAS 245 | 12 Jan '25)
Hubble Fellow @ Columbia

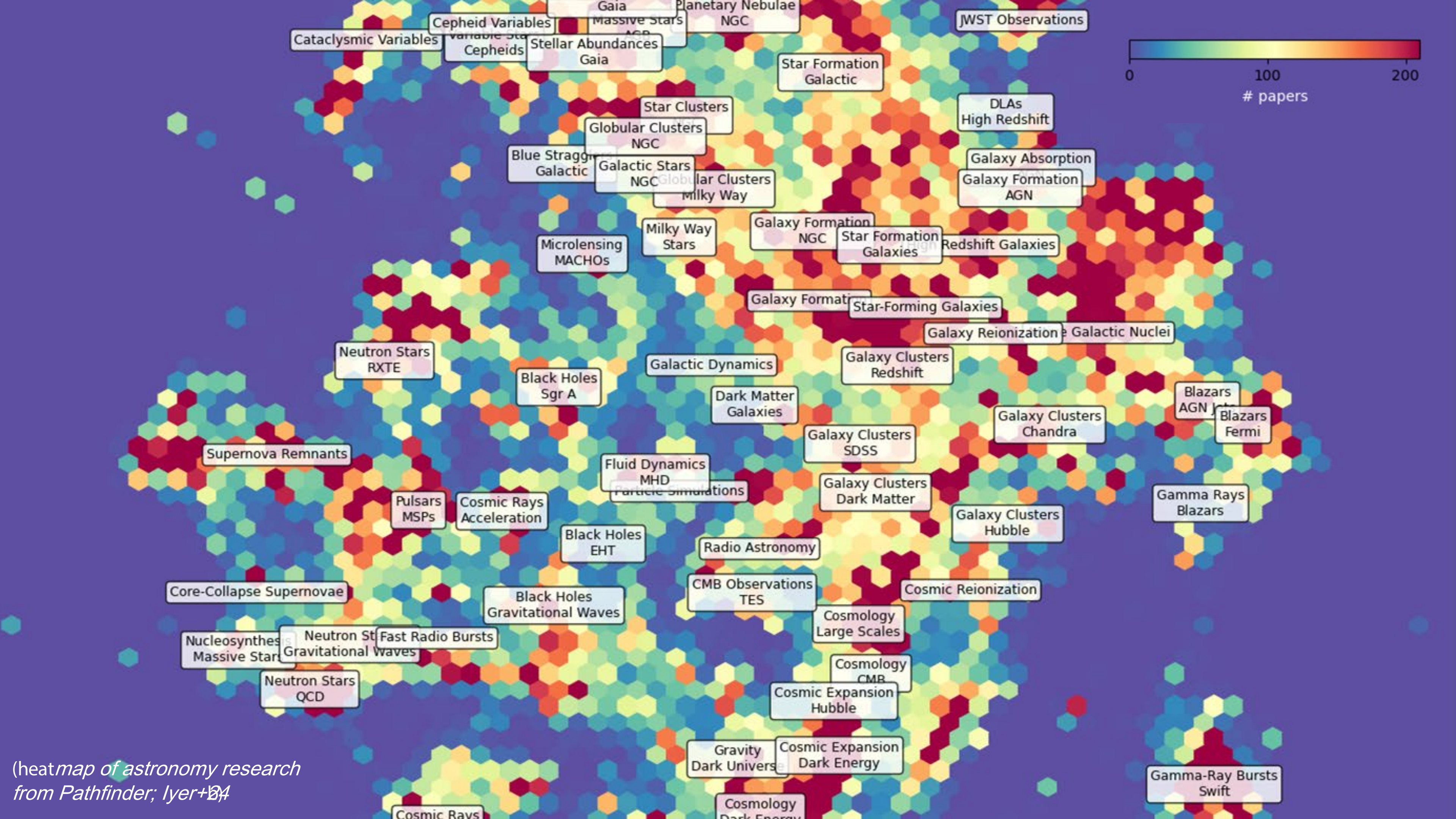
Facilitating, not replacing:

The role of AI & ML in crafting *tools* for astronomy research

(*map of astronomy research from Pathfinder; Iyer+24*)

Kartheik Iyer @ COPAG session (AAS 245 | 12 Jan '25)
Hubble Fellow @ Columbia

(*map of astronomy research from Pathfinder; Iyer+24)*



(heatmap of astronomy research from Pathfinder; Iyer+24)

Ask it questions at

https:// pfdr.app

(*map of astronomy research from Pathfinder; Iyer+24*

Facilitating, not replacing:

The role of AI & ML in crafting *tools* for astronomy research

So what's the deal with AI?

'*artificial intelligence*' is an umbrella term for
a wide variety of semi-related software tools and techniques

Facilitating, not replacing:

The role of AI & ML in crafting *tools* for astronomy research

Is it better to call it ML?

‘*machine learning*’ uses mathematical models to teach machines how to perform specific tasks and learn from data

yes (*for the purpose of this talk*) & **no** (*at a more semantic level;* link) -
ML algorithms are generally task-specific, with a well-defined loss function
AI is used to define a broader collection of methods that enable generalization

Facilitating, not replacing:
The role of AI & ML in crafting *tools* for astronomy research

What do ML models do?

'*machine learning*' uses mathematical models to teach machines how to perform specific tasks and learn from data, to

Learn mappings between (x,y)
Organize data by similarity
Sample from a distribution

Facilitating, not replacing:

The role of AI & ML in crafting *tools* for astronomy research

What do ML models do?

'*machine learning*' uses mathematical models to teach machines how to perform specific tasks and learn from data, to

Learn mappings between (x,y) -> regression, classification, denoising, emulators, machine translation

Organize data by similarity -> clustering, outliers/ anomalies, object detection,

Sample from a distribution -> inference, generative models, accelerating simulations

(*with lots of overlaps, e.g. VAEs do both i and iii, GMMs do ii) and iii*

Facilitating, not replacing:

The role of AI & ML in crafting *tools* for astronomy research

What do ML models do?

'*machine learning*' uses mathematical models to teach machines how to perform specific tasks and learn from data, to

Learn mappings between (x,y) -> regression, classification, denoising, emulators…

Organize data by similarity -> clustering, outliers/ anomalies, object detection…

Sample from a distribution -> inference, generative models, accelerating simulations…

(*possible hidden 4th class:* ***learning representations****, ask me about this later!*)

(*with lots of overlaps, e.g. VAEs do both i and iii, GMMs do ii and iii*)

Kartheik Iyer @ COPAG session (AAS 245 | 12 Jan '25)
Hubble Fellow @ Columbia

(map of astronomy research
from Pathfinder; Iyer+24)

> automated Classification using light curves and spectra (e.g. Davison+22, Burhanudin+23), anomaly detection (e.g. Pruzhinskaya+19), real-time event alerts (e.g. ZTF), emulation (AI, Chen+24)...

> PLAsTICC

> AI-driven spatio-temporal engine for finding gravitationally lensed type Ia supernovae (Kodi Ramanah+22)

(map of astronomy research from Pathfinder; Iyer+24)

Asteroid dynamics (e.g., Carruba+22)

> **stellar classification** (e.g. Sedaghat+21), **detection of stellar streams and galactic structure mapping** (e.g. Sengupta+24, Bonaca & Price-Whelan '24), **dark matter and galaxy dynamics** (Putney+24), **abundance patterns** (Almannaei+24, Pérez-Couto+24).

> stellar foundation models (Koblischke & Bovy '24)

Searching for axions (e.g., Day & Krippendorf '20)

(map of astronomy research from Pathfinder; Iyer+24)

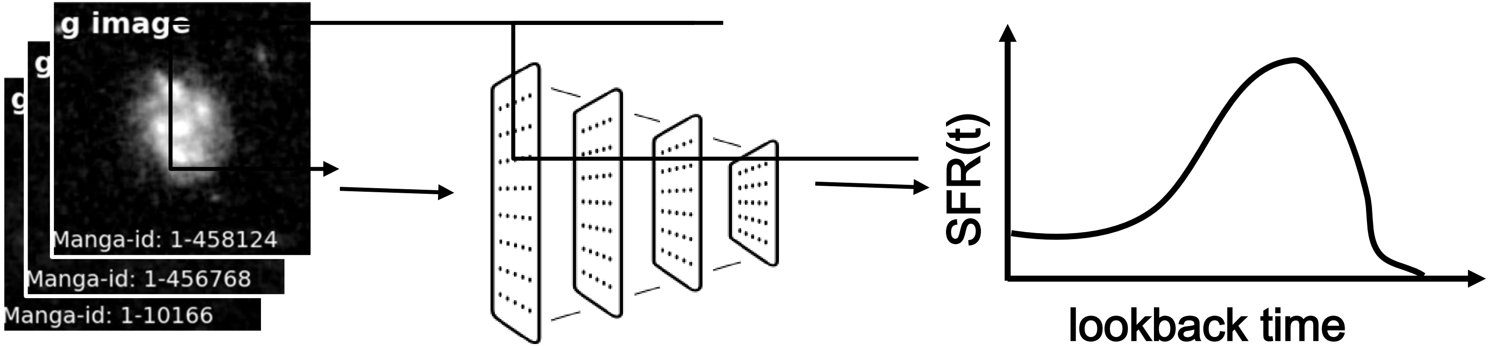
> **better photo-z (so many!), morphologies** (e.g. Zoobot, Walmsley+23), **predicting spectral properties from images** (e.g. Wu+20, Alfonzo+24), **field-level or population-level inference** (e.g. Villaescusa-Navarro+23, Iyer & Speagle+24, Matthews+24, Lovell+24), **spectral latent spaces** (e.g. Portillo+20), **accelerated hydro simulations & emulators** (Kamdar+16, Alsing+21, Nguyen+23)

(map of astronomy research from Pathfinder; Iyer+24)

Juan Alfonzo (Tohoku U)
juanpabloalfonzo@astr.tohoku.ac.jp

What is the link between *galaxy morphology* and *star formation histories*?

(Katachi | Alfonzo, Iyer et al. (2024) | ArXiv: 2404.05146)



→ Train a network to predict SFH parameters from images

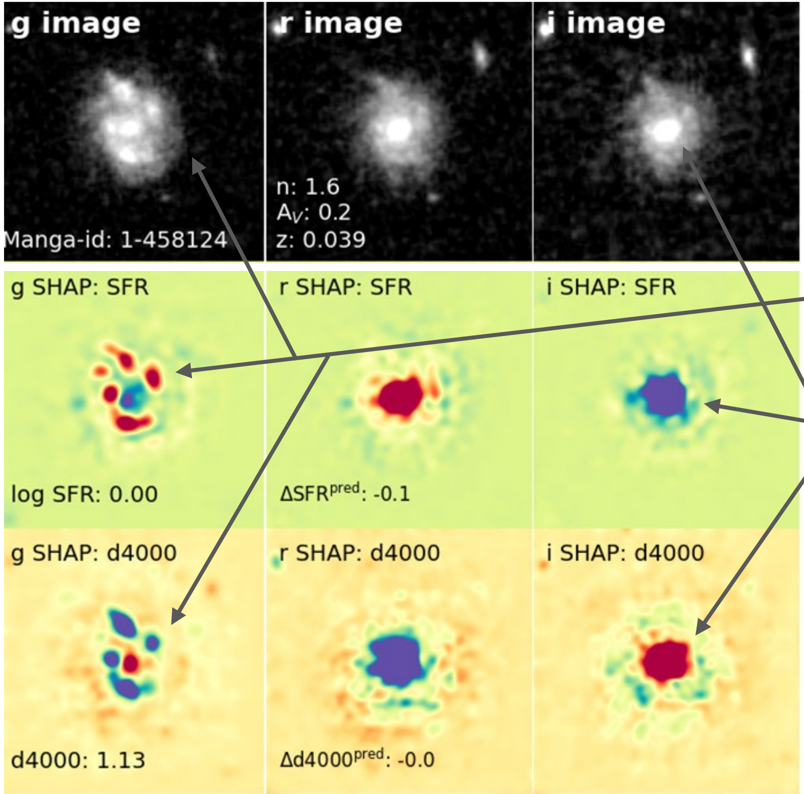
← Use XAI methods like SHAP to understand
which parts of the image drive the prediction

Katachi

Alfonzo, KI et al *2024 (ApJ) Katachi: https://arxiv.org/abs/2404.05146*

decoding the *SFH-morphology connection* with SHAP

(Katachi | Alfonzo, Iyer et al. (2024) | ArXiv: 2404.05146)



SHAP: SHapley Additive Analysis (Lundberg & Lee '17)

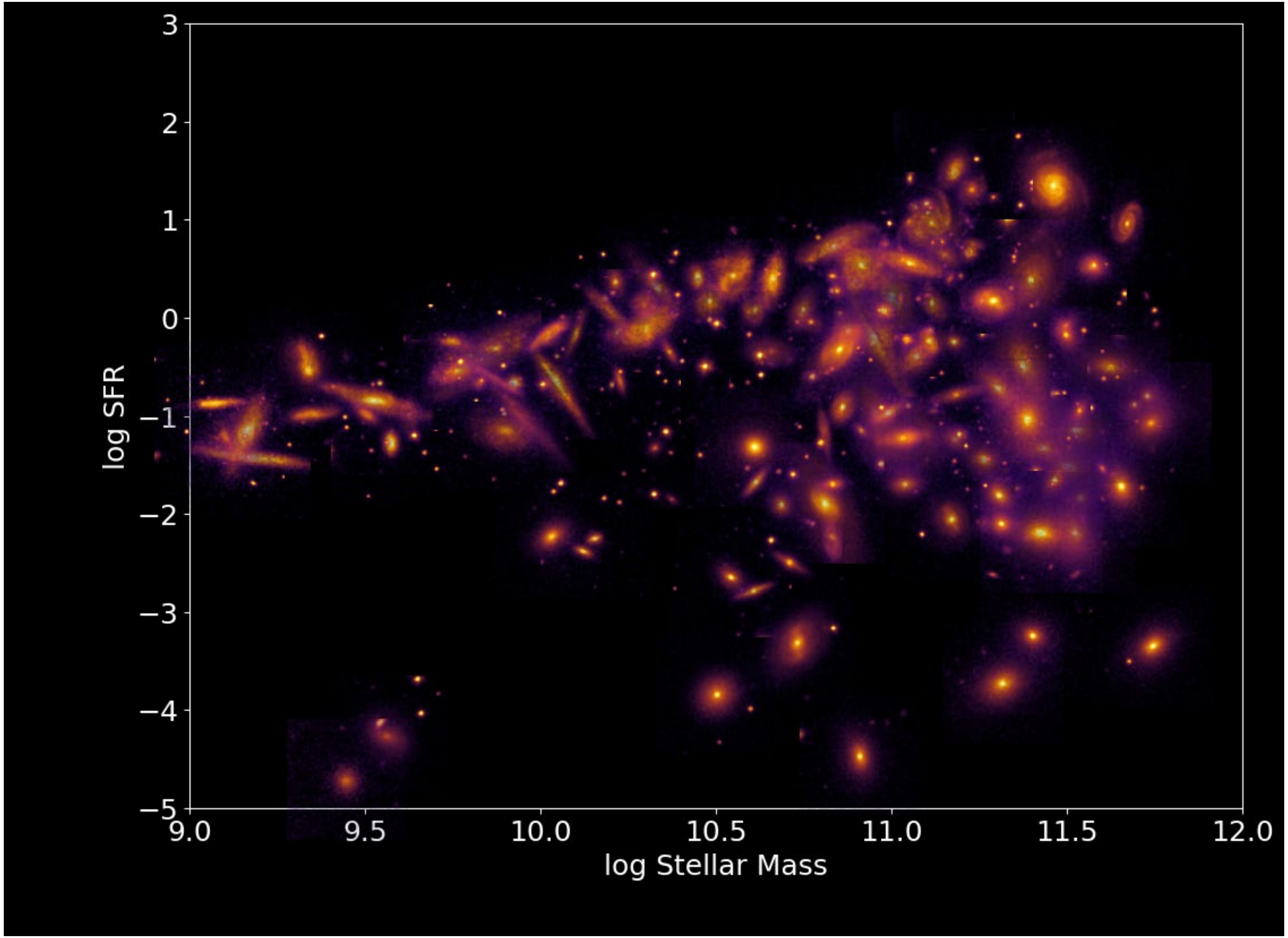
Red: more flux here increases parameter estimate

Blue: more flux here decreases parameter estimate

Star forming clumps in the bluer band drive the network to **higher SFR** and **lower t50** values

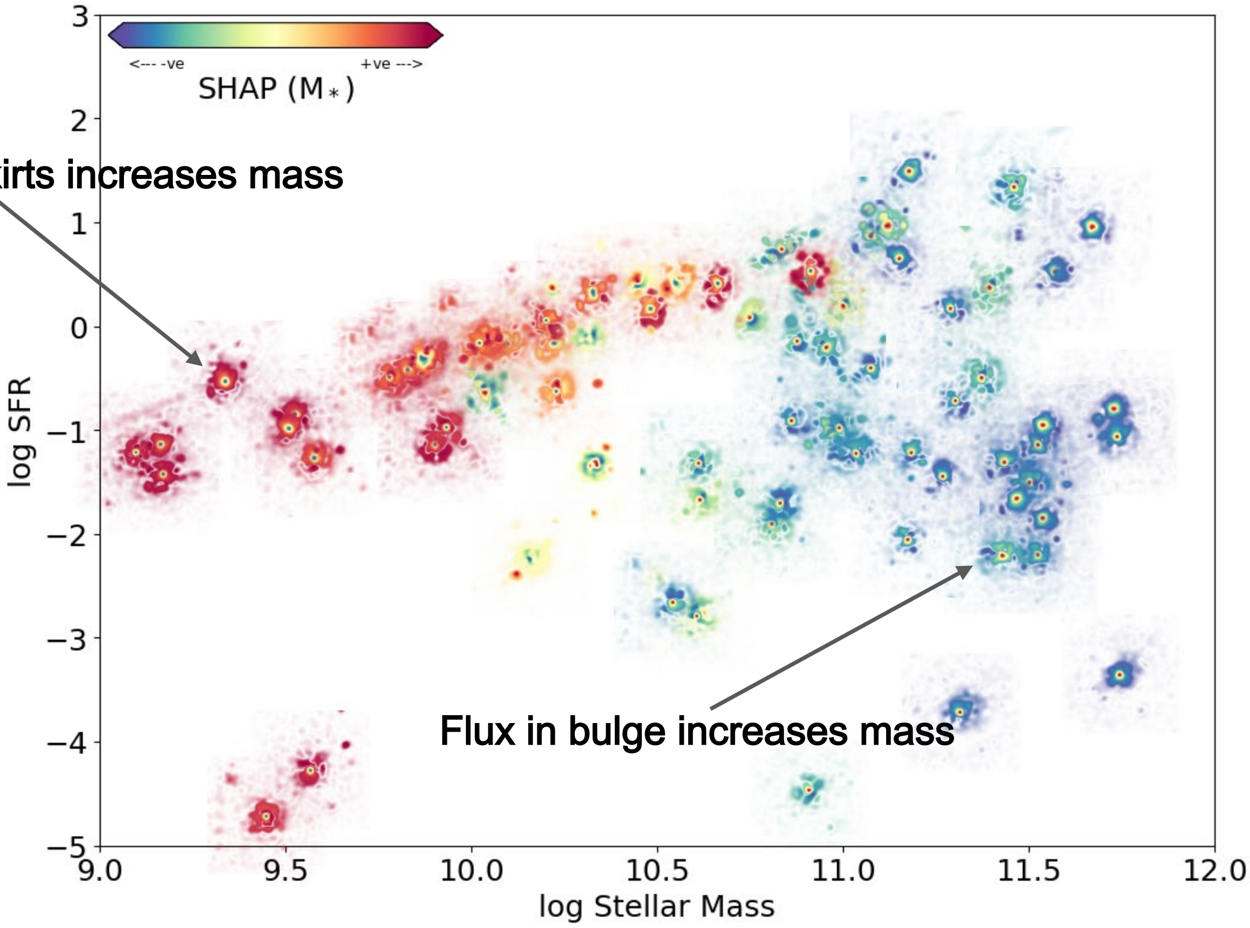
Flux in the bulge in redder bands reduce SFR and **increase t50** values

Alfonzo, KI et al *2024 (ApJ) Katachi: https://arxiv.org/abs/2404.05146*



Katachi

Alfonzo, KI et al *2024 (ApJ) Katachi: https://arxiv.org/abs/2404.05146*



Katachi

Alfonzo, KI et al *2024 (ApJ) Katachi: https://arxiv.org/abs/2404.05146*

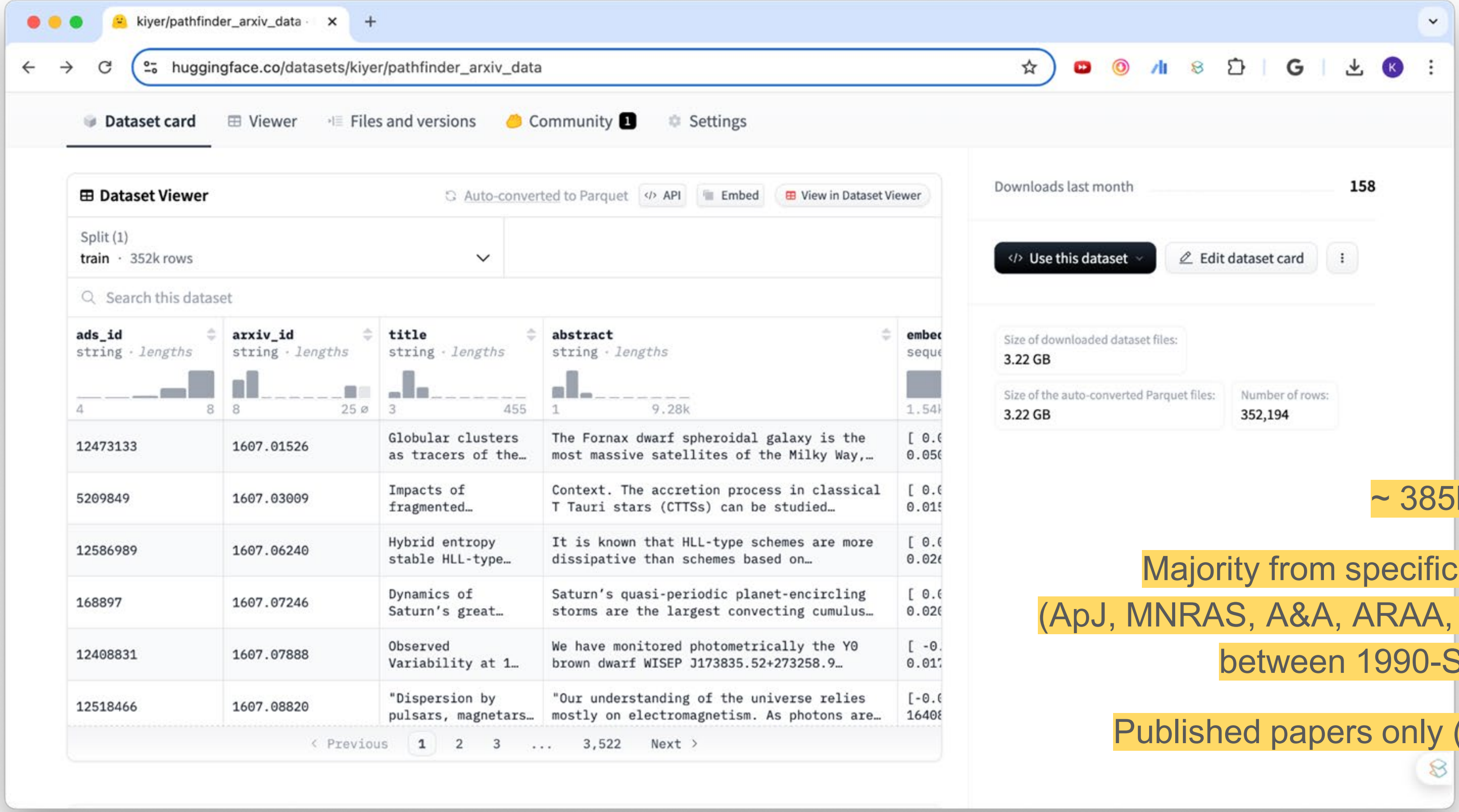
> Methods matter!
e.g. likelihood-free inference
(LFI, also SBI or ILI)

(*map of astronomy research from Pathfinder; Iyer+24*

> Datasets matter!
e.g. Gaia DR2, SDSS IV MaNGA, CAMELS

(map of astronomy research from Pathfinder; Iyer+24

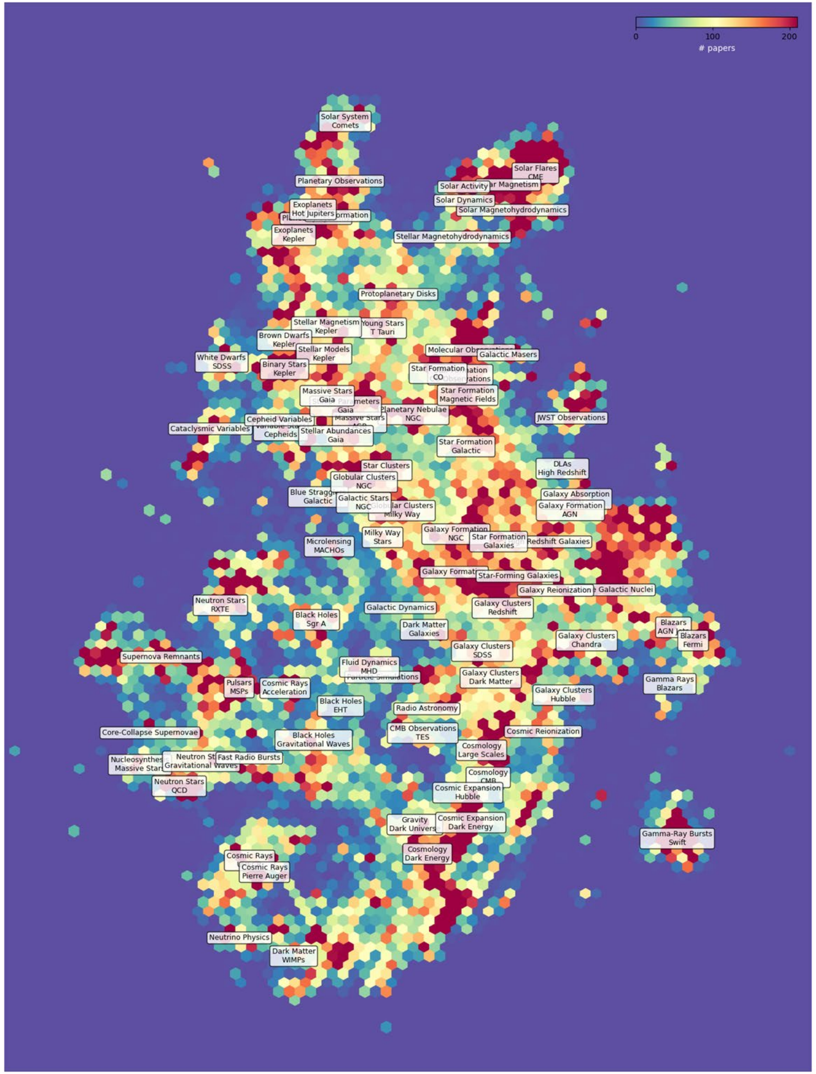
Building the corpus

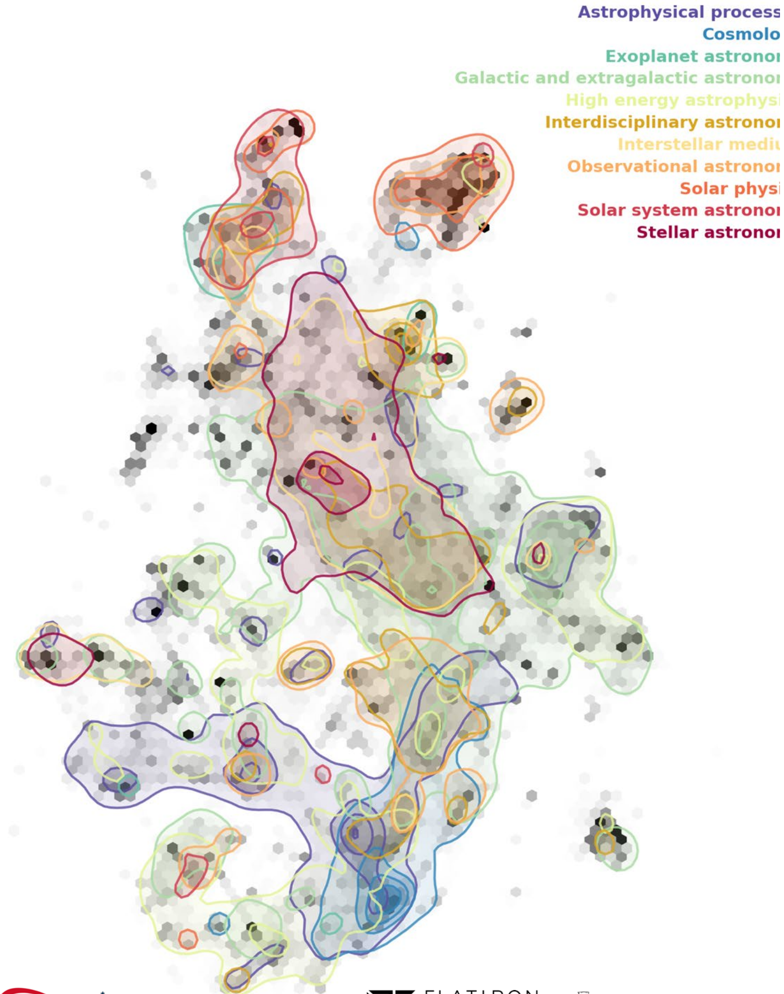


~ 385k papers

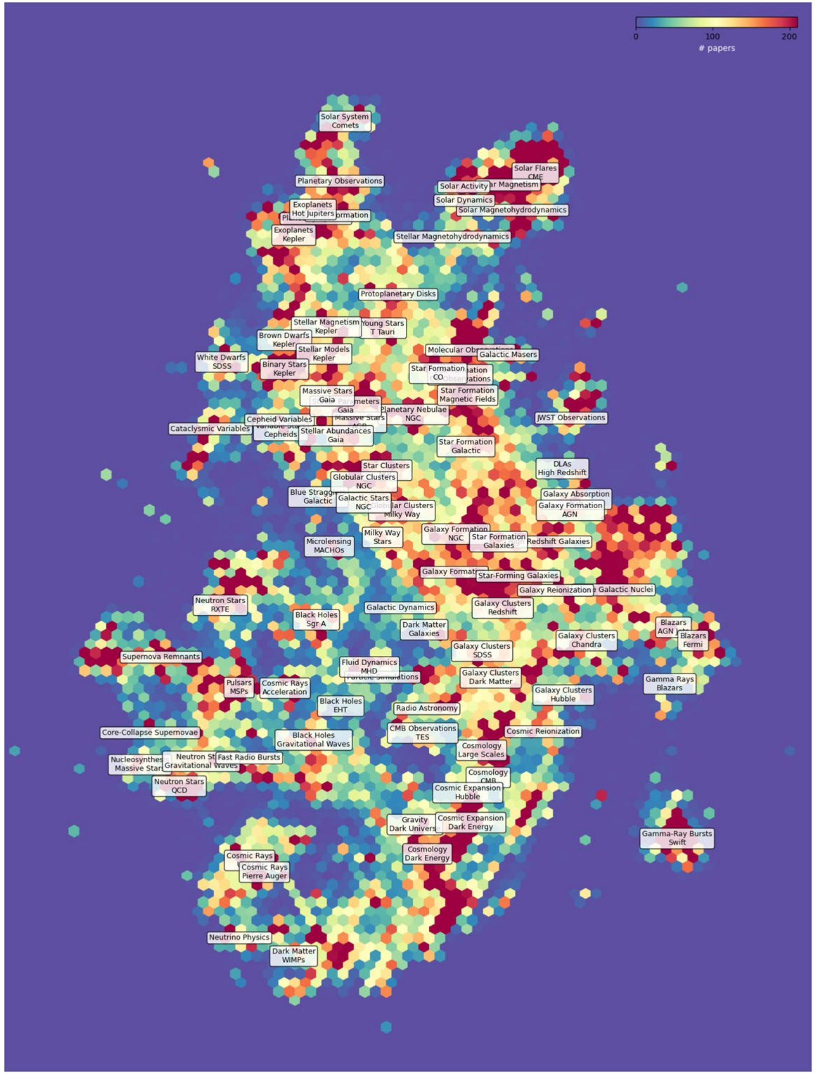
Majority from specific journals (ApJ, MNRAS, A&A, ARAA, Nature), between 1990-Sep 2024

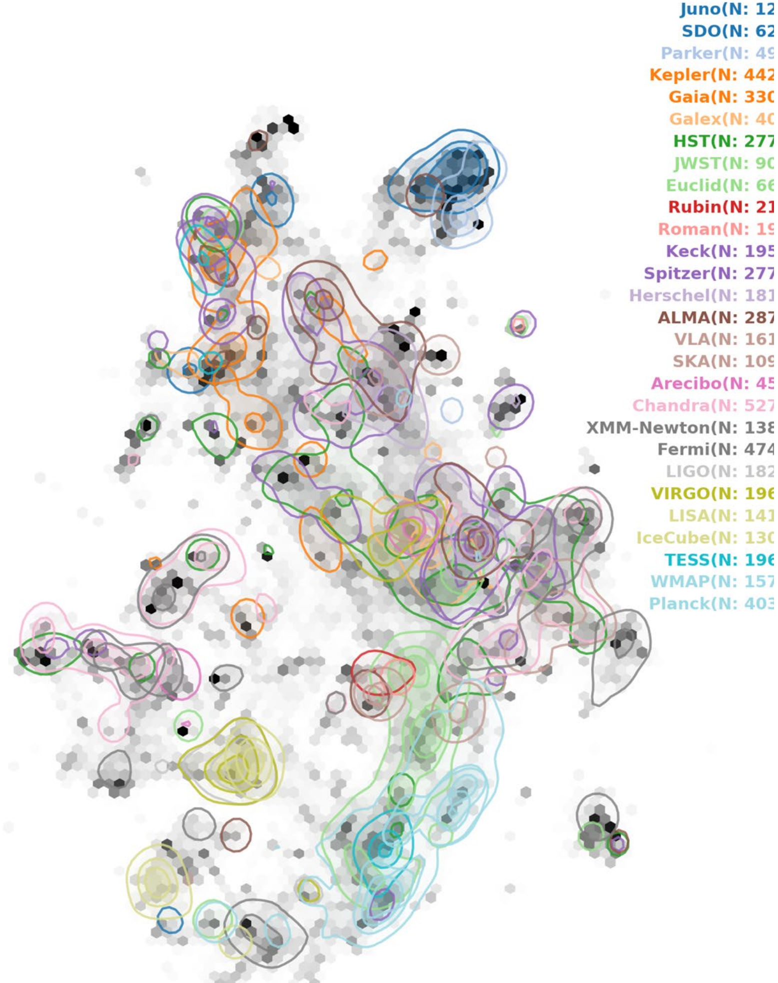
Published papers only (for now)





NHFP COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK FLATIRON INSTITUTE Center for Computational Astrophysics universeTBD astrophysics data system





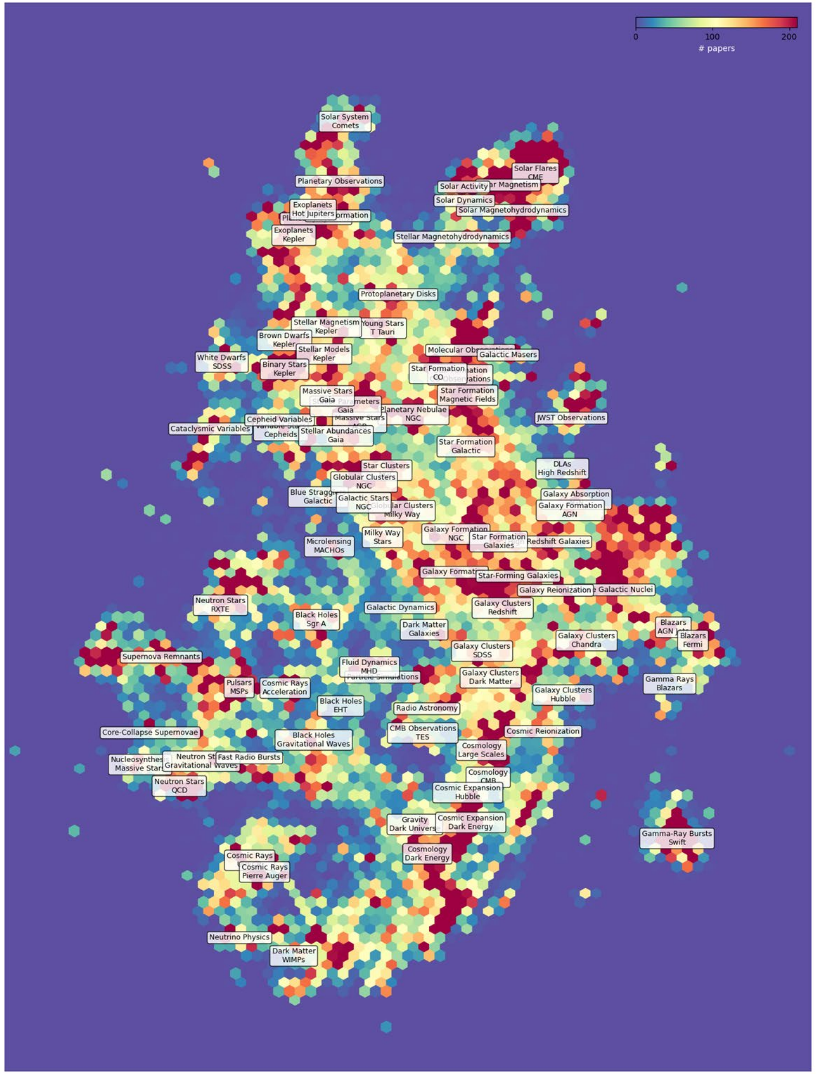
NHFP

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

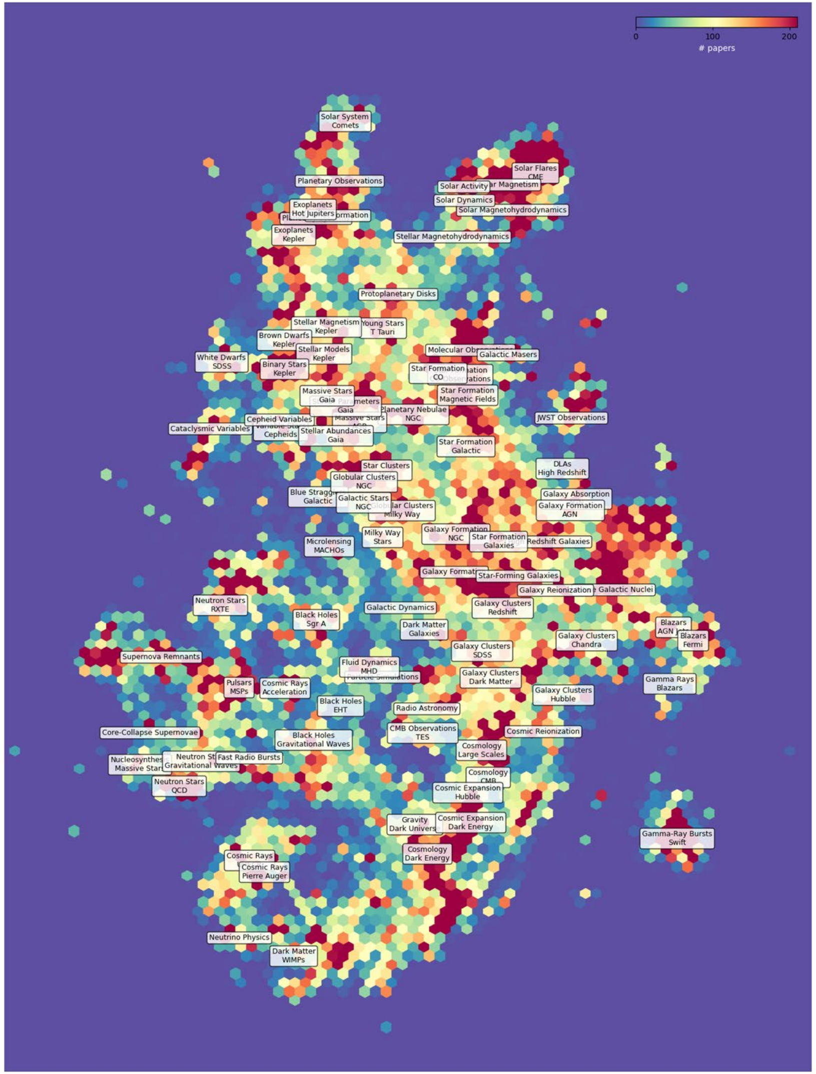
FLATIRON INSTITUTE Center for Computational Astrophysics

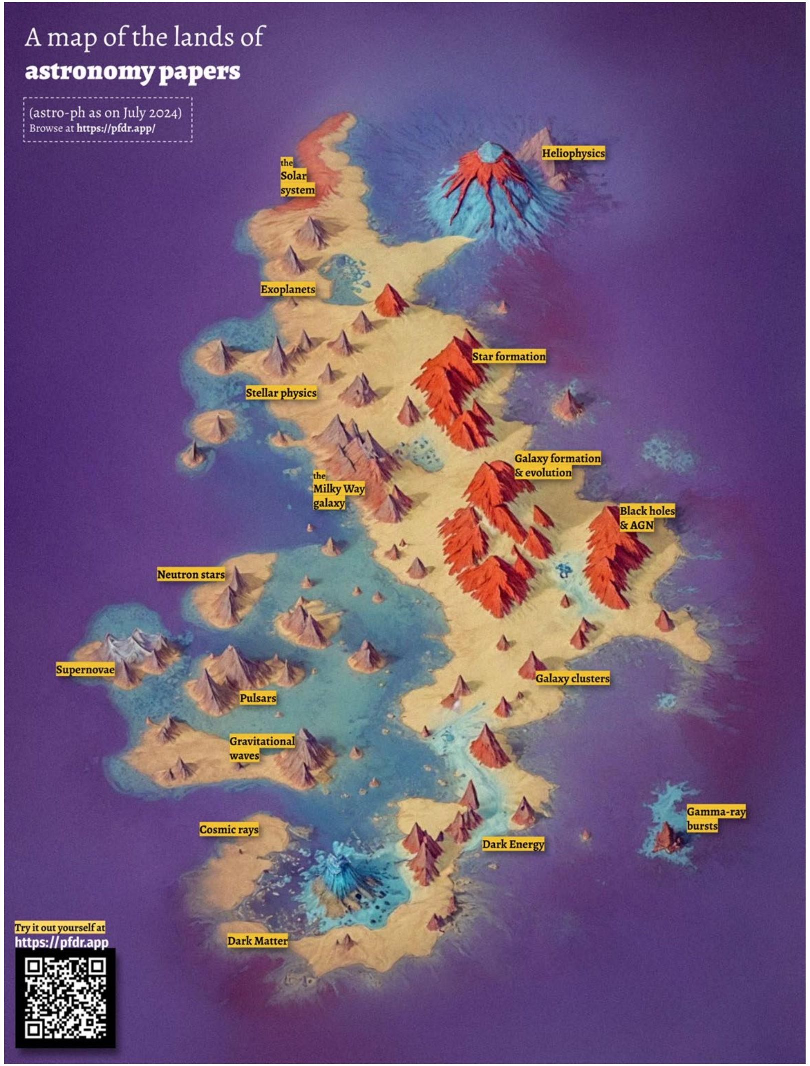
universeTBD

astrophysics data system



NHFP COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK FLATIRON INSTITUTE Center for Computational Astrophysics universeTBD astrophysics data system


0
100
200
papers
Solar System
Comets
Solar Flares
CME
Planetary Observations
Solar Activity
Solar Dynamics
Solar Magnetohydrodynamics
Exoplanets
Hot Jupiters
Exoplanets
Kepler
Stellar Magnetohydrodynamics
Protoplanetary Disks
Stellar Magnetism
Kepler
Young Stars
T Tauri
Brown Dwarfs
Kepler
White Dwarfs
SDSS
Stellar Models
Kepler
Binary Stars
Kepler
Galactic Masers
Star Formation
CO
Massive Stars
Gaia
Star Formation
Magnetic Fields
Planetary Nebulae
NGC
JWST Observations
Cataclysmic Variables
Cepheid Variables
Cepheids
Stellar Abundances
Gaia
Star Formation
Galactic
Star Clusters
Globular Clusters
NGC
DLAs
High Redshift
Galaxy Absorption
Galactic Stars
NGC
Galaxy Formation
AGN
Microlensing
MACHOs
Milky Way
Stars
Galaxy Formation
NGC
Star Formation
Galaxies
Redshift Galaxies
Star-Forming Galaxies
Galaxy Reionization
Galactic Nuclei
Neutron Stars
RXTE
Black Holes
Sgr A
Galactic Dynamics
Dark Matter
Galaxies
Galaxy Clusters
Redshift
Blazars
Blazars
Fermi
Galaxy Clusters
Chandra
Supernova Remnants
Galaxy Clusters
SDSS
Fluid Dynamics
MHD
Galaxy Clusters
Dark Matter
Pulsars
MSPs
Cosmic Rays
Acceleration
Gamma Rays
Blazars
Galaxy Clusters
Hubble
Black Holes
EHT
Radio Astronomy
Core-Collapse Supernovae
CMB Observations
TES
Cosmic Reionization
Black Holes
Gravitational Waves
Cosmology
Large Scales
Fast Radio Bursts
Gravitational Waves
Neutron Stars
QCD
Cosmic Expansion
Hubble
Gravity
Cosmic Expansion
Dark Energy
Gamma-Ray Bursts
Swift
Cosmology
Dark Energy
Cosmic Rays
Cosmic Rays
Pierre Auger
Neutrino Physics
Dark Matter
WIMPs


A map of the lands of
astronomy papers
(astro-ph as on July 2024)
Browse at https://pfdr.app/
the Solar system
Heliophysics
Exoplanets
Star formation
Stellar physics
Galaxy formation & evolution
the Milky Way galaxy
Black holes & AGN
Neutron stars
Supernovae
Galaxy clusters
Pulsars
Gravitational waves
Cosmic rays
Gamma-ray bursts
Dark Energy
Try it out yourself at
https://pfdr.app
Dark Matter

Case studies ¿Qué es una galaxia quieta?

Broad questions + multilingual questions (eg from Karl's demo)

Una galaxia quieta, en el contexto de la astronomía, se refiere a un tipo de galaxia que presenta características de baja actividad en comparación con otras galaxias más activas, como las galaxias activas o los cuásares. Estas galaxias suelen tener una emisión de radio débil y, a menudo, se les asocia con un bajo nivel de formación estelar. En términos más técnicos, las galaxias quietas pueden ser clasificadas como "galaxias radio-quietas", que son aquellas que no emiten significativamente en el espectro de radio, a diferencia de sus contrapartes "radio-loud" (radio-ruidosas).

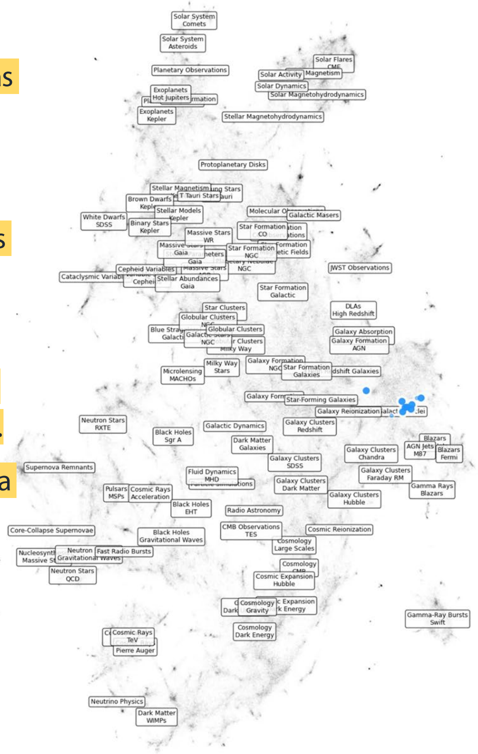
Características de las Galaxias Quietas:

Baja Emisión de Radio: Las galaxias quietas tienen una emisión de radio que es significativamente menor que la de las galaxias activas. Esto se debe a que no tienen un núcleo galáctico activo que emita radiación intensa.

Formación Estelar Moderada: Estas galaxias suelen tener tasas de formación estelar más bajas, lo que significa que no están formando nuevas estrellas a un ritmo elevado.

Estructura y Morfología: Muchas galaxias quietas son espirales o elípticas, y su morfología puede influir en su actividad. Por ejemplo, las galaxias espirales tienden a ser más propensas a la formación estelar que las elípticas.

Ejemplos: Un ejemplo de galaxia quieta podría ser una galaxia elíptica que no muestra signos de actividad en su núcleo, a diferencia de un cuásar que es extremadamente brillante y activo.



JSALT team


Alina Hyk


Kiera McCormick

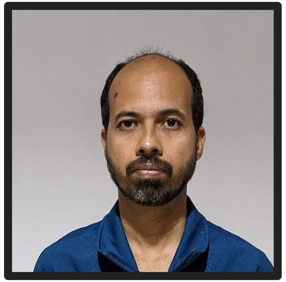

Charlie O'Neill


Christine Ye


Mikaeel Yunus


Elina Baral


Philipp Koehn


Sanjib Sharma


Anjalie Field


Jo Ciucă


Kartheik Iyer


John Wu

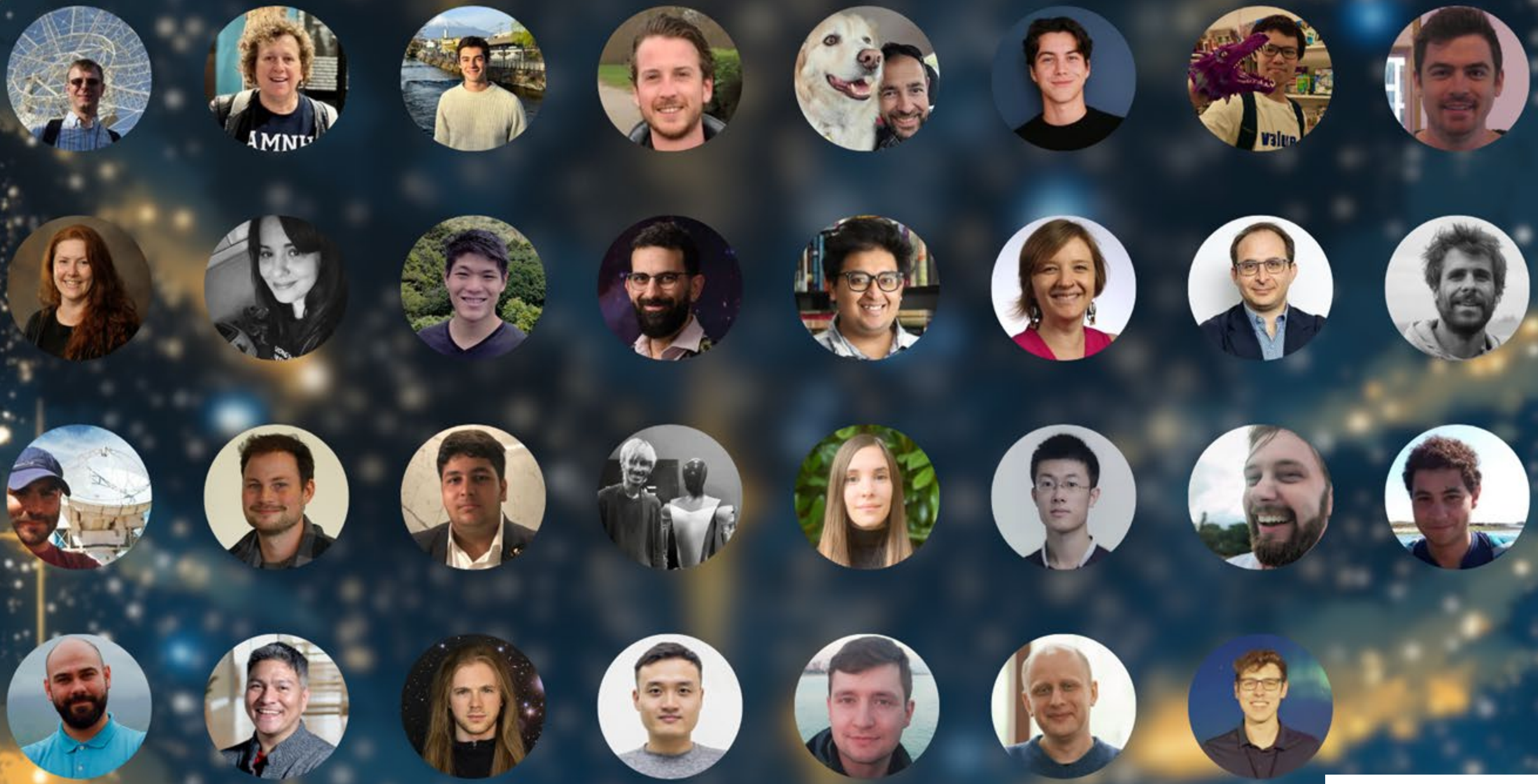
+ thanks to Josh Peek, Mich
Ntampaka, Jenn Kotler, Pao
Garcia, Sanjeev Khudanpur,
JSALT admin/staff team, and
our user base of astronomer

JOHNS HOPKINS
WHITING SCHOOL of ENGINEERING

CENTER FOR LANGUAGE AND SPEECH PROCESSING

STScI | SPACE TELESCOPE SCIENCE INSTITUTE

OUR TEAM



Slide from Maja Jablonska, NLP for space science meeting

universeTBD

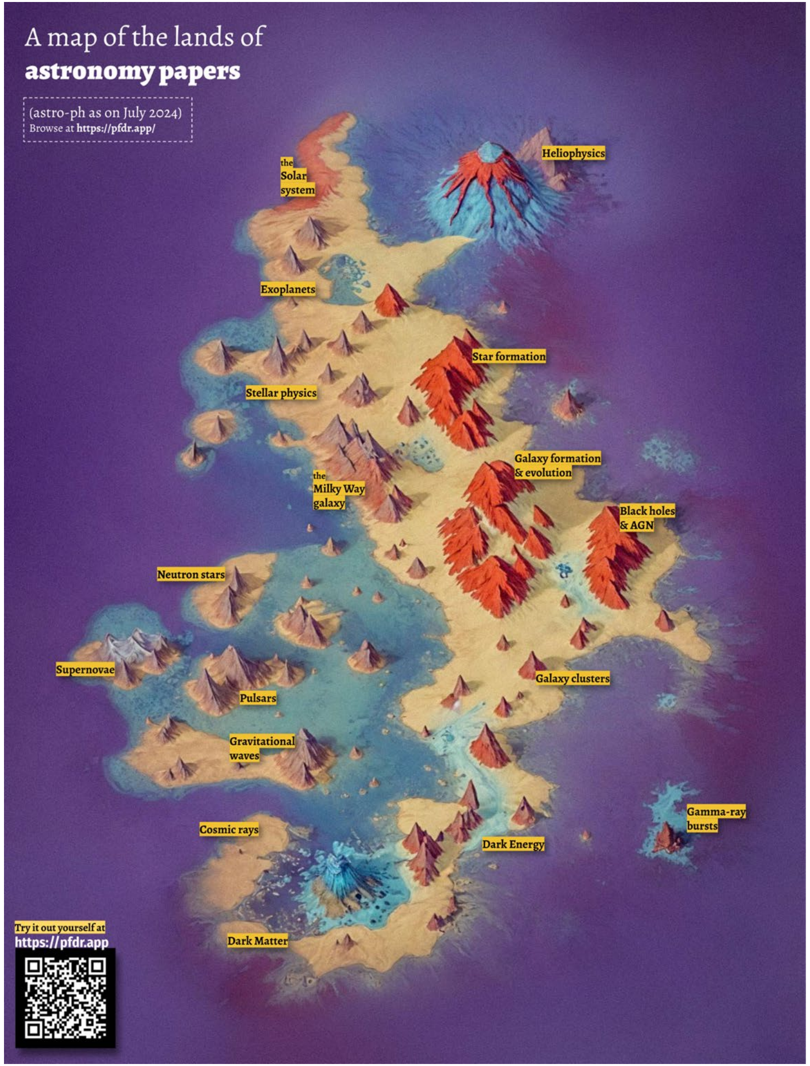
Facilitating, not replacing:

The role of AI & ML in crafting *tools* for astronomy research

Literature & Resources:

- Acquaviva+23: Machine Learning for Physics and Astronomy (textbook)
- Baron '22: Machine Learning in Astronomy: a practical overview (arXiv:1904.07248)
- Acquaviva '21: Teaching Machine Learning for the Physical Sciences: A summary of lessons learned and challenges (arXiv:21
- Djorgovski+22: Applications of AI in Astronomy (arXiv:2212.01493)
- Ntampaka+19: The Role of Machine Learning in the Next Decade of Cosmology (arXiv:1902.10159)
- Huertas-Company & Lanusse '23: The Dawes Review 10: The impact of deep learning for the analysis of galaxy surveys (arXiv:
- Smith & Geach '23: Astronomia ex machina: a history, primer and outlook on neural networks in astronomy (arXiv:2211.03796)
- Fotopoulou '24: A review of unsupervised learning in astronomy (arXiv.2406.17316)
- Buchner & Fotopoulou '24: How to set up your first machine learning project in astronomy (Nature Reviews Physics; link)
- Iyer+24b: pathfinder: A Semantic Framework for Literature Review and Knowledge Discovery in Astronomy (arXiv:2408.01556
- The Multimodal Universe: Enabling Large-Scale Machine Learning with 100TB of Astronomical Scientific Data (arXiv:2412.02527
- Some university-based organisations in the US: AstroAI (Harvard/CfA), PolymathicAI (Flatiron Institute), CosmicAI (UT Austin & more), SkAI (Northwestern & more), SpaceML (https://spaceml.org/, NASA/ESA/FDL)
- UniverseTBD: https://universetbd.org/ & SciX: https://science.data.nasa.gov/science-explorer
- **These slides: https://bit.ly/kiyer_copag_ai_astro**

Kartheik Iyer @ COPAG session (AAS 245 | 12 Jan '25)
Hubble Fellow @ Columbia | kgi2103@columbia.edu



Facilitating, not replacing:
The role of AI & ML in crafting *tools* for astronomy research

Thank you! Brief takeaways

1. Is there hype? *Yes.* Ethical concerns? *Yes.*
2. Should I use it anyway? *Depends, but also probably yes, in parts where your scientific problems involve things that ML is good at.*
3. How do I start? *Lots of resources (prev. slide)*
4. How do I build trust? *Extensive validation, knowledge of assumptions & training data, interpretability tools*
5. What happens next? *Buckle up!*
6. These slides: https://bit.ly/kiyer_copag_ai_astro

NHFP · Columbia University in the City of New York · Flatiron Institute Center for Computational Astrophysics · universeTBD · astrophysics data system

What can Pathfinder currently do?

Access pathfinder online at pfdr.app

Pathfinder: Papers in Astronomy Traced and Harnessed For INtelligent Discovery, Exploration, and Research?

https://arxiv.org/abs/2408.01556

> **> Find similar papers** given a natural language query
> **> Synthesize information / answer questions** from paper abstracts
> **> Fine tune search** by keywords/ recency/ citations
> **> Explore** the landscape of astronomy literature
> **> Translate, simplify and contextualize** abstracts
> **> Quantify trends** in research, mission impact, and more

clsp.jhu.edu/evaluating-llm-performance-in-research-astronomy/

JOHNS HOPKINS WHITING SCHOOL of ENGINEERING

CENTER FOR LANGUAGE AND SPEECH PROCESSING

Home About People Publications Seminars Workshops HLT Masters Resources

Abstract | Full Proposal | Group Members | Team Presentations

Evaluating LLM Performance in Research Astronomy

Large language models (LLMs) are being used not only for common-knowledge information retrieval, but also for specialized disciplines such as cutting-edge astronomical research. However, in specialized domains, we lack robust, realistic, and user-oriented evaluations of LLM capabilities. Human evaluations are time-intensive, subjective, and difficult to reproduce, while automated metrics like perplexity or task benchmarks fail to reflect realistic performance. We seek to advance understanding of LLM capabilities for supporting scientific research through user-centric analysis and the development of robust evaluation standards; while we expect outputs from this workshop to be generalizable, we will focus on astronomy. Astronomy has open data and a vibrant and active community that is open to partnering in the design, experimentation, and evaluation processes. The primary goal of the workshop is to develop a quantifiable metric or objective function for evaluating LLMs in astronomy research, thereby taking humans out of the evaluation loop. A secondary goal is to understand how the evaluation criteria for a specialized use case (astronomy) compares to the evaluation criteria for typical English conversations. Our proposal will explore the first step toward a lofty goal: how can AI transform science for the better, by first evaluating what it means to be better.

Subscribe to the CLSP Seminars Mailing List

UPCOMING SEMINARS

SEP 6 Fri — **12:00 pm** Pengtao Xie (UCSD) "Foundation M... @ Hackerman Hall B17

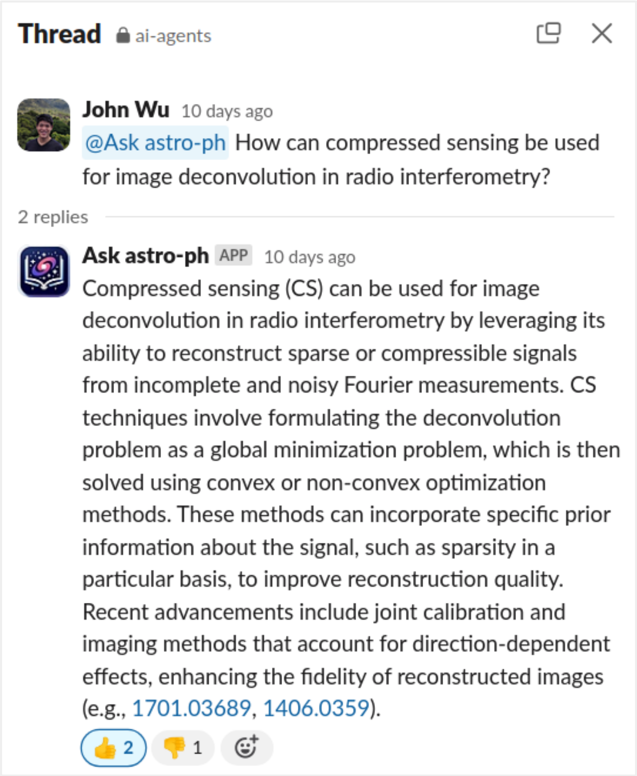
Add ▾ | View Calendar →

Design and data collection for evaluation

Slack chatbot experiment lasted 4.5 weeks, during which we collected some valuable user data (~350 questions and responses) from ~40 astronomers.

37 Users (astronomers)
378 User queries
→ 269 in channel
→ 109 direct message
342 Emoji reactions
263 Feedback/comments
373 Retrievals

Slide credit: JSALT team

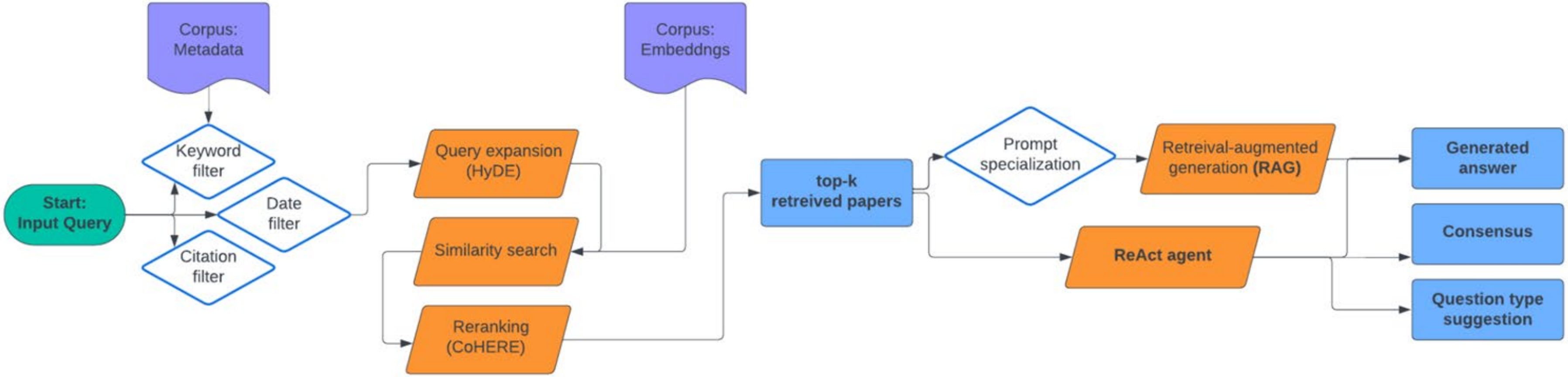


What features do scientists seek or expect in a ChatBot for interacting w academic literature?

- Generation capabilities
  - **Specificity of answer**
  - Fluency of natural language
  - **Extra formatting capabilities (language translation, poems, rap)**
  - General correctness
  - **Interpret or synthesize knowledge**
  - **Critique**
  - **Graceful refusal**
  - **Expression of Confidence**
  - Personalization
  - Surprise Discovery Factor
  - Understanding different technical notations
  - Agent like capabilities
  - Chat with context (multi-turn)
- Retrieval capabilities
  - **Metadata incorporation**
  - **Quality of retrieved citations**
  - **Completeness**

Slide credit: JSALT team

paring down features for scalability…



dense basis

Fast, flexible fitter for galaxy stellar populations

Spectral modeling tools like **dense basis** can be used to observationally constrain galaxy SFHs.

Iyer et al. (2019), dense-basis.readthedocs.io

Pathfinder

LLM-enabled literature search

LLM-assisted literature survey tools like **pathfinder** help keep track of the ever-expanding astronomy literature & democratize research.

Iyer et al. (2024), *accepted in ApJS*, https://pfdr.app

the Firefly Sparkle

star formation on parsec scales at z=8.3

High-z lensed galaxies like the **Firefly Sparkle** offer glimpse of resolved star formation as proto-galaxies assemble.

Mowla & Iyer et al. (2024), *accepted in Nature*

Katachi

Understanding the *Form* of Galaxies

Interpretable ML tools like **Katachi** extend SFH inference to future telescopes & help build models of the SFH-morphology connection.

Alfonzo, Iyer et al. (2024), *published in ApJ*

Suites of cosmological simulations like **CAMELS** can reveal the **SFH-feedback connection** for large populations of galaxies.

Iyer et al. *(in prep.)*

kgi2103@columbia.edu
github.com/kartheikiyer