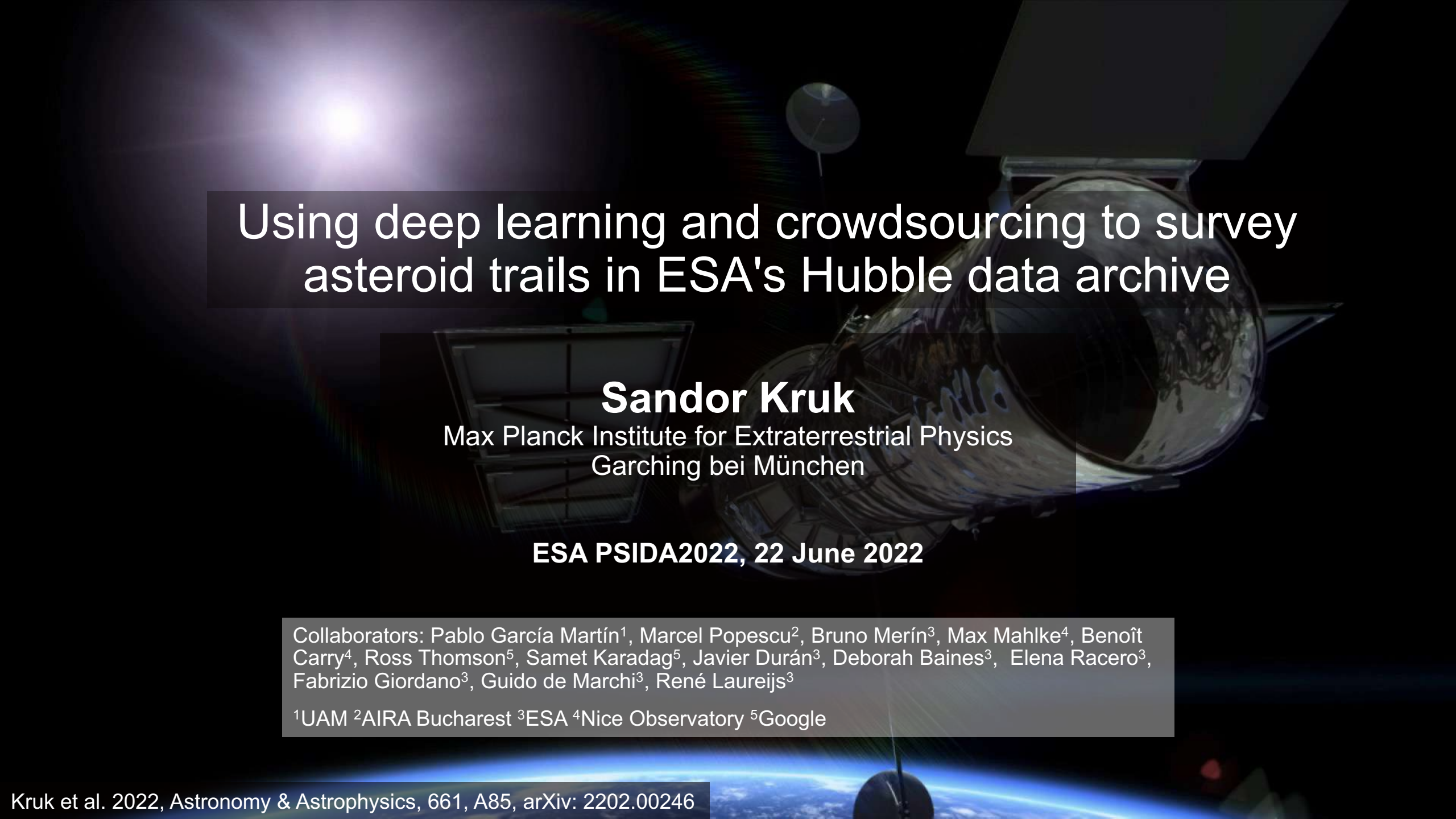


Using deep learning and crowdsourcing to survey asteroid trails in ESA's Hubble data archive

Sandor Kruk

Max Planck Institute for Extraterrestrial Physics
Garching bei München

ESA PSIDA2022, 22 June 2022

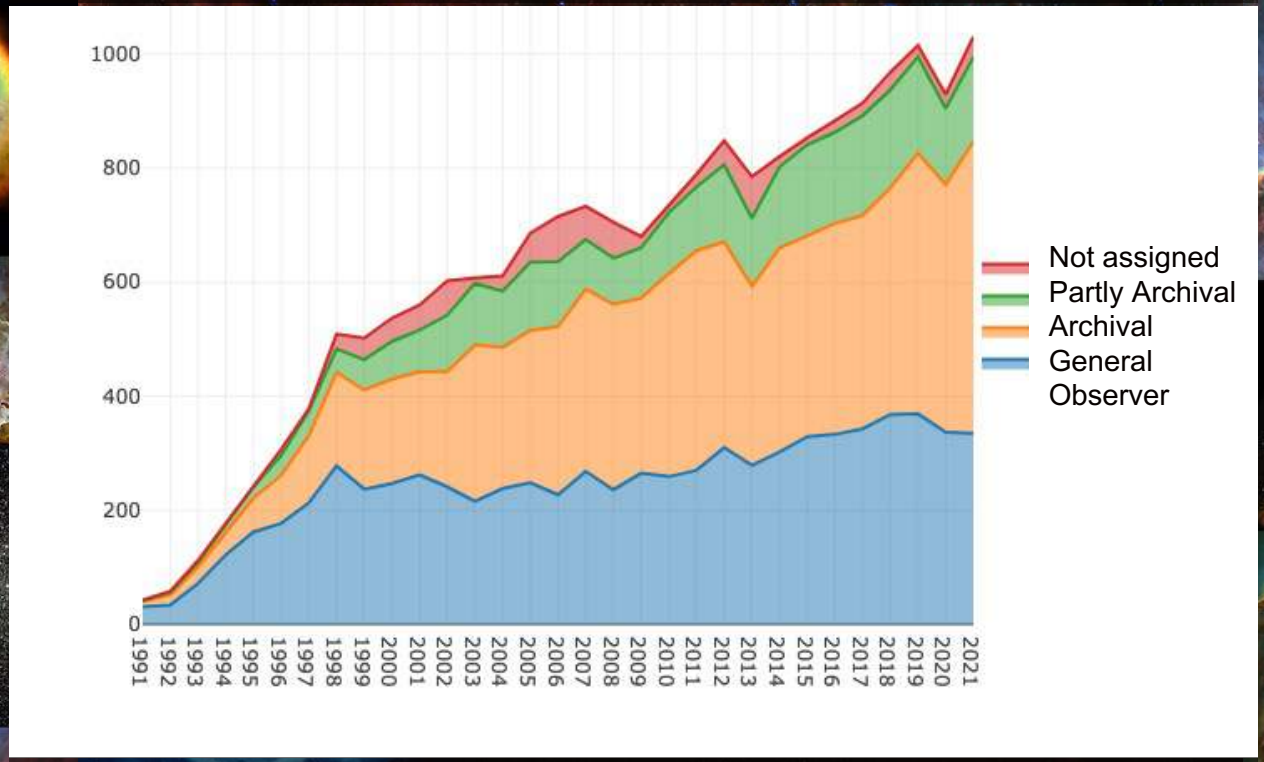
Collaborators: Pablo García Martín[1], Marcel Popescu[2], Bruno Merín[3], Max Mahlke[4], Benoît Carry[4], Ross Thomson[5], Samet Karadag[5], Javier Durán[3], Deborah Baines[3], Elena Racero[3], Fabrizio Giordano[3], Guido de Marchi[3], René Laureijs[3]

[1]UAM [2]AIRA Bucharest [3]ESA [4]Nice Observatory [5]Google

Kruk et al. 2022, Astronomy & Astrophysics, 661, A85, arXiv: 2202.00246

Growing number of publications using archival data

Hubble Space Telescope publications by observation type

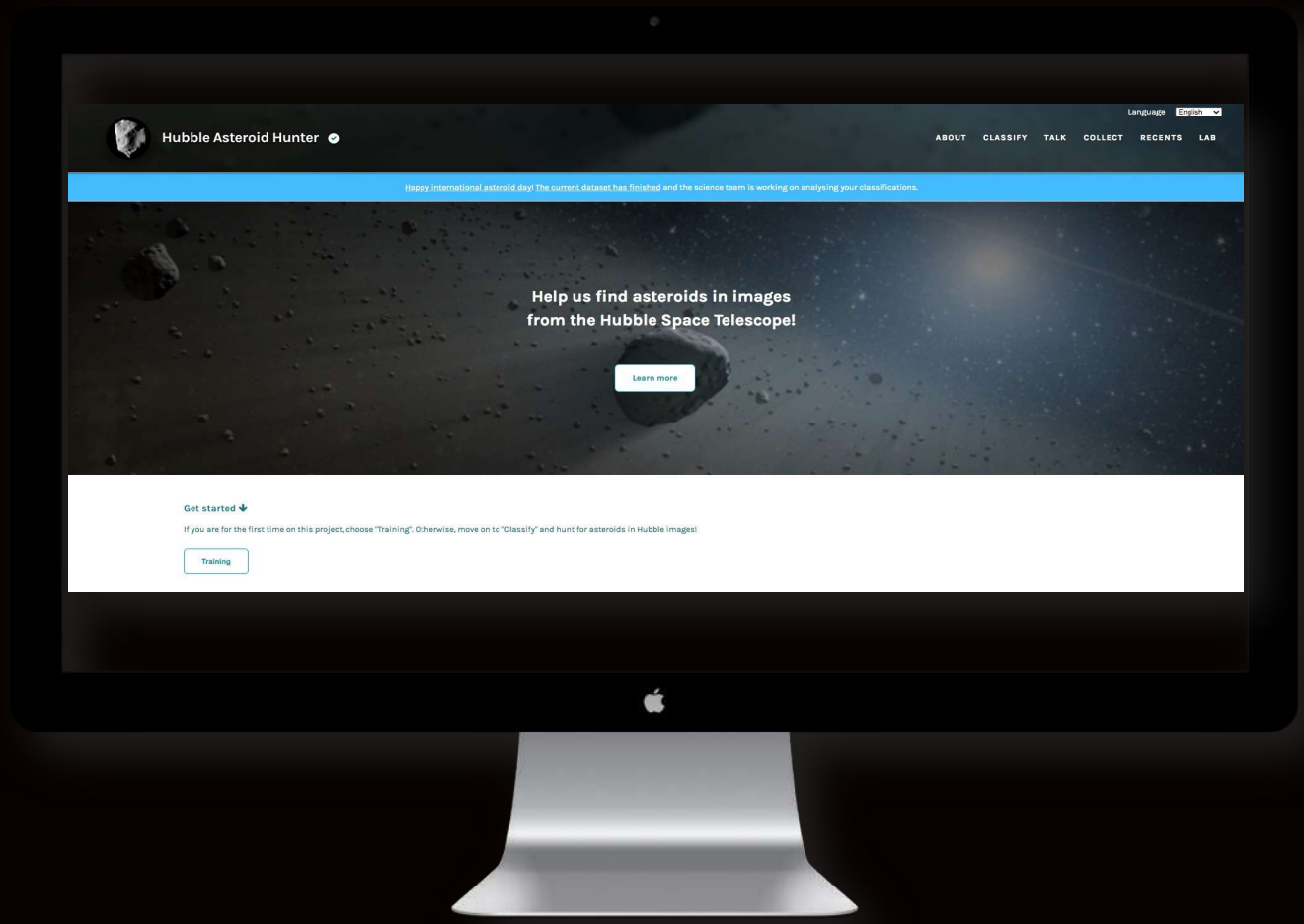


https://archive.stsci.edu/hst/bibliography/pubstat.html

AI and crowdsourcing for scientific exploration of archives

We need novel techniques for data analysis ➡ AI/Machine Learning

We need AI-ready datasets and labels for training and validation ➡ crowdsourcing



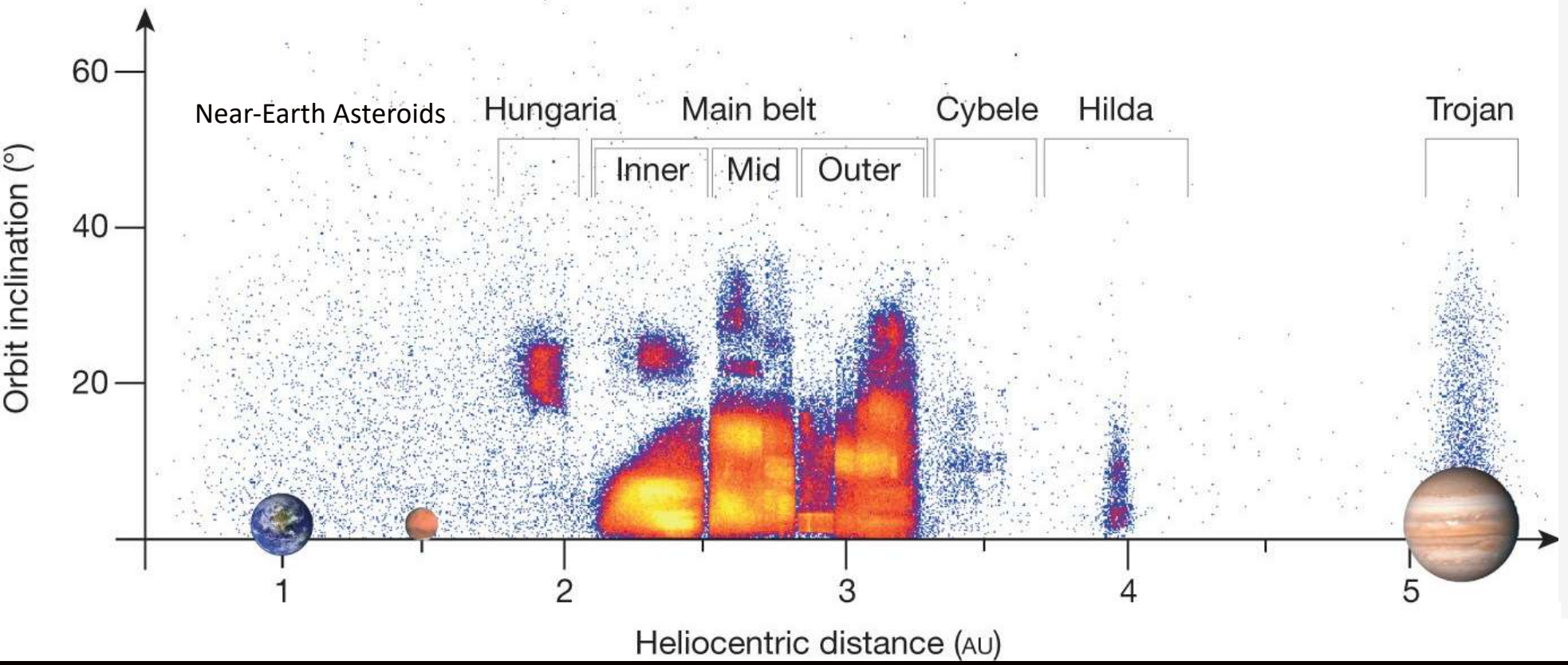
Identifying asteroid trails in the Hubble Space Telescope images

Abell 370, Credit: NASA, ESA/Hubble, B.Sunnquist and J. Mack


DON'T
LOOK
UP

Asteroids across the Solar System

Distribution of asteroid classes in the Solar System



DeMeo and Carry, 2014

Minor Planet Center

Near-Earth Objects Discovered

THIS MONTH:	55
THIS YEAR:	1305
ALL TIME:	29230

Minor Planets Discovered

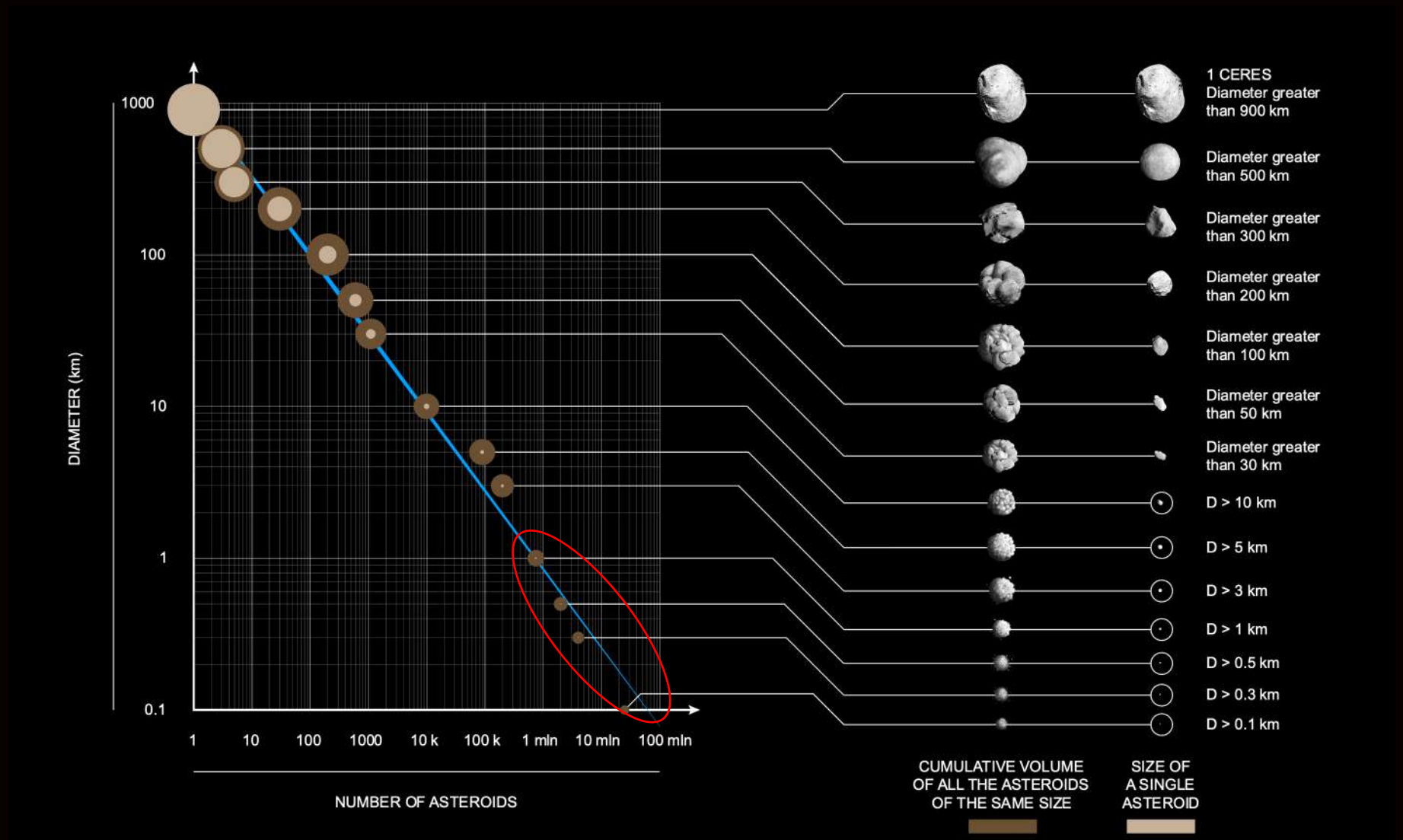
THIS MONTH:	89
THIS YEAR:	6100
ALL TIME:	1207472

Comets Discovered

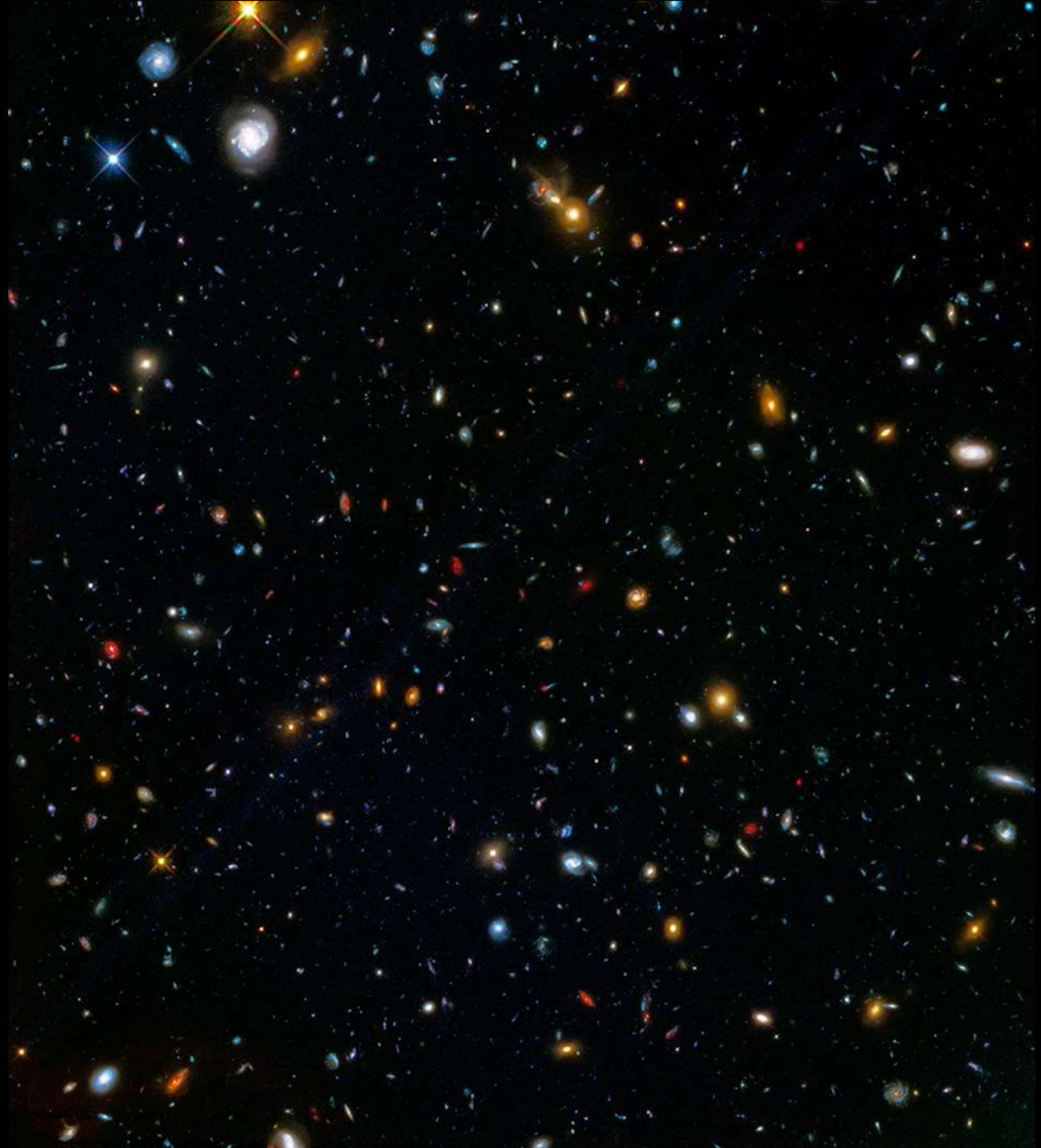
THIS MONTH:	1
THIS YEAR:	20
ALL TIME:	4407

The number of discovered asteroids, as of 21 June 2022. Source: MPC https://www.minorplanetcenter.net/

Distribution of asteroid sizes in the Solar System



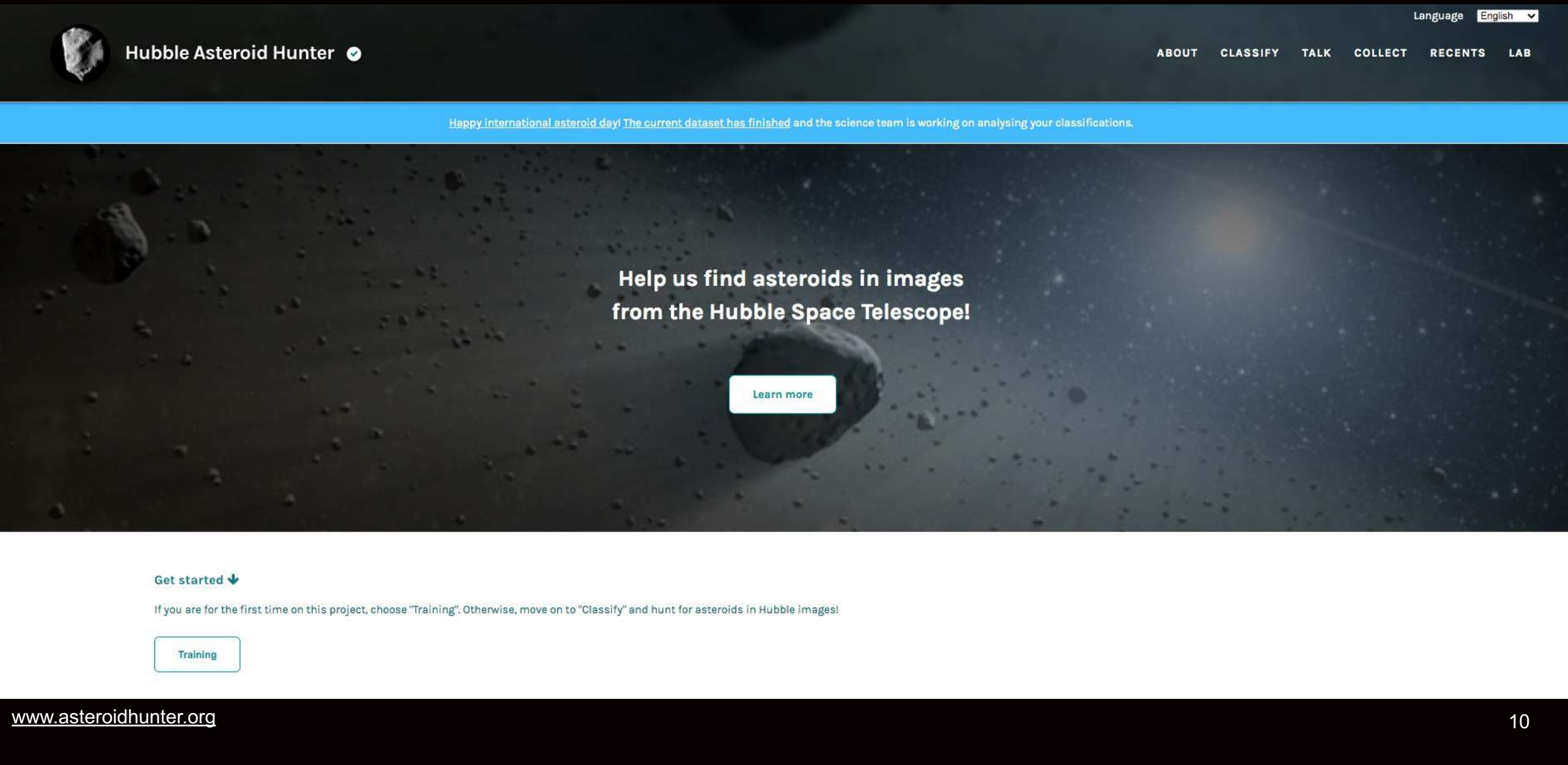
Present-day main belt and NEO populations – based on Bottke et al. 2015 (Credit: Marco Colombo, Density Design Research Lab)



Abell 370 Parallel Field
Credit: NASA, ESA/Hubble

Citizen Science Project

ZOONIVERSE



www.asteroidhunter.org

TASK TUTORIAL

Is there an asteroid trail visible in the images?

Yes

No

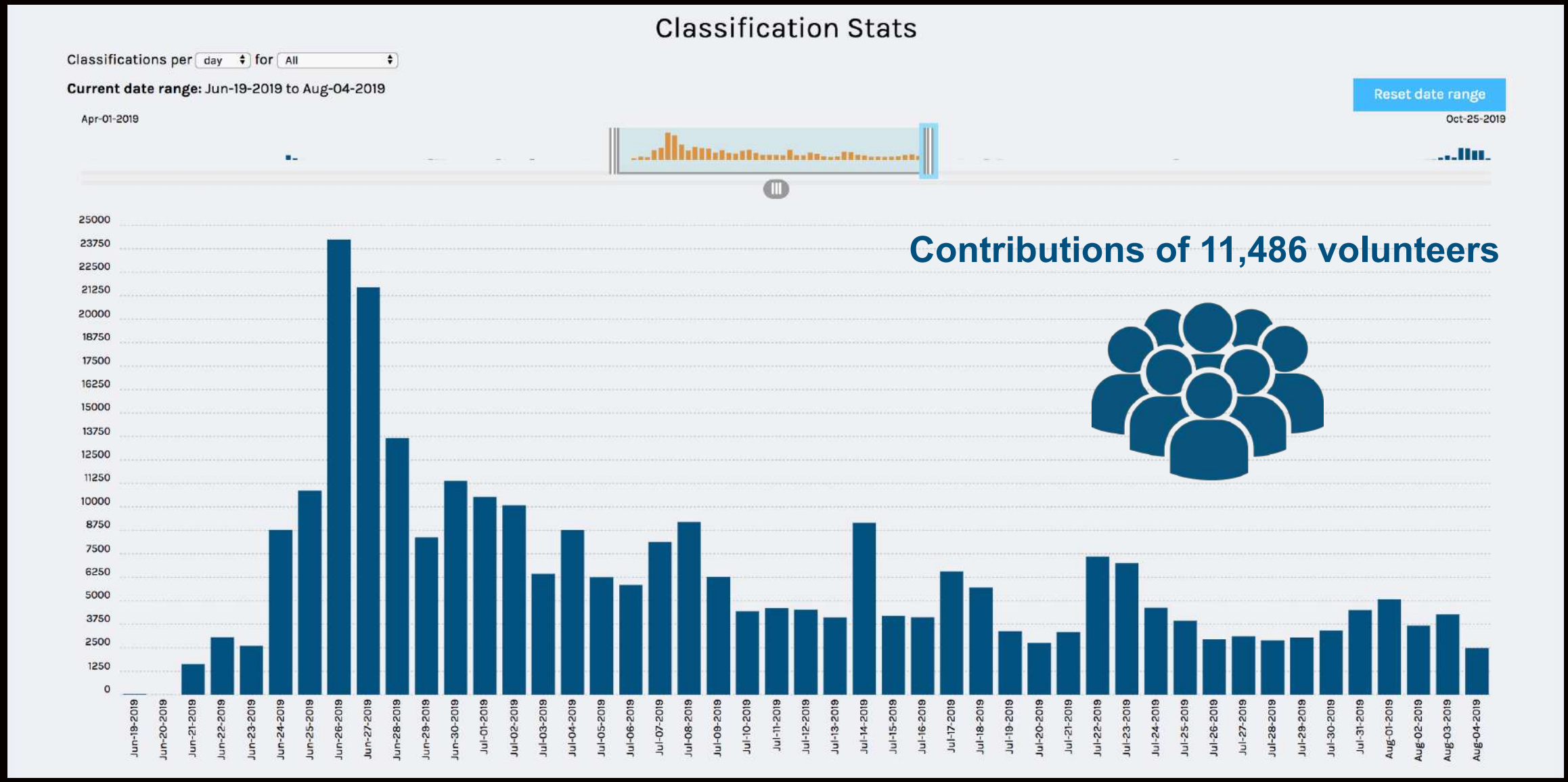
Impossible to tell

NEED SOME HELP WITH THIS TASK?

Done & Talk Done

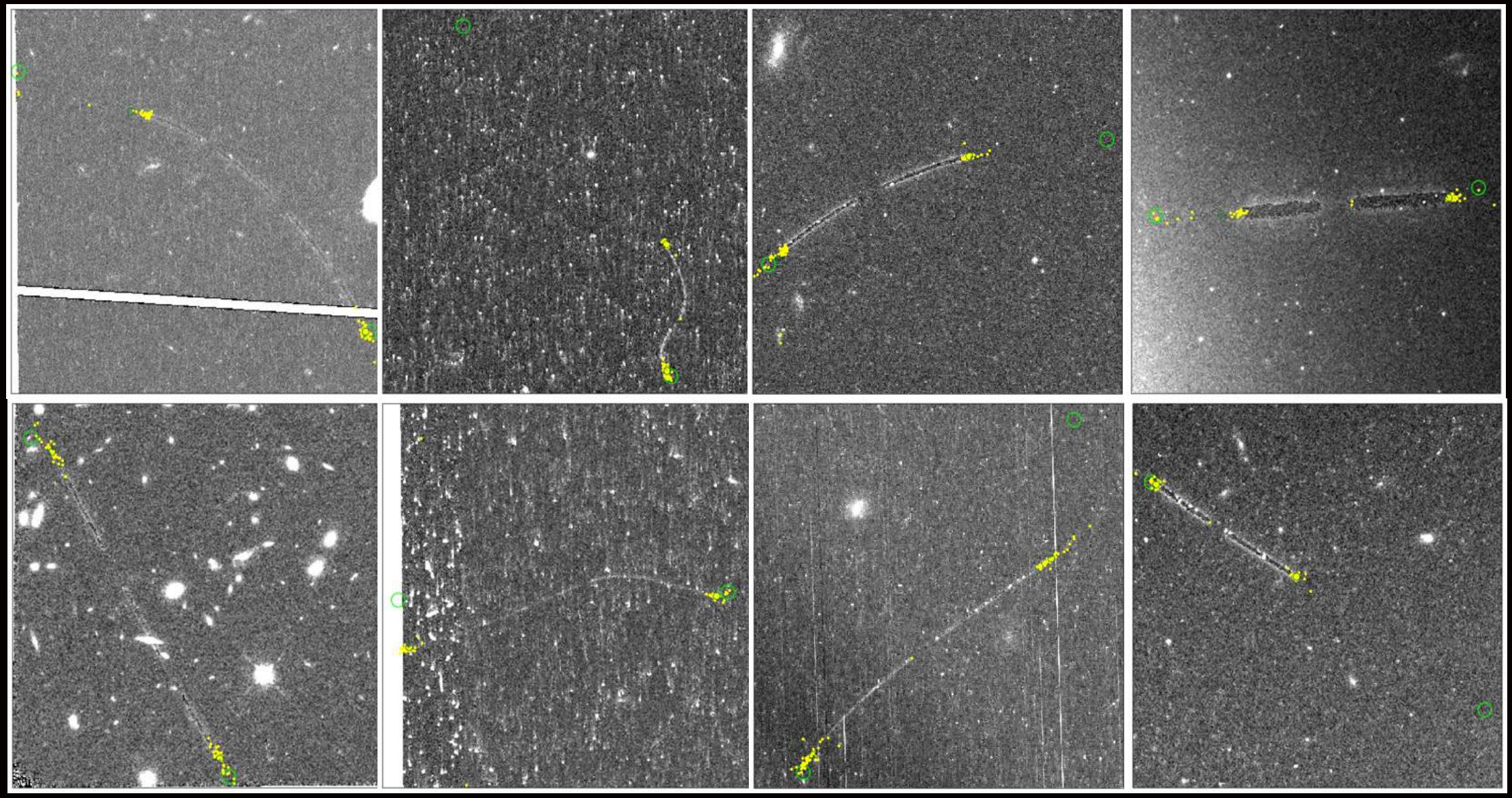
FIELD GUIDE

Volunteer participation in the citizen science project



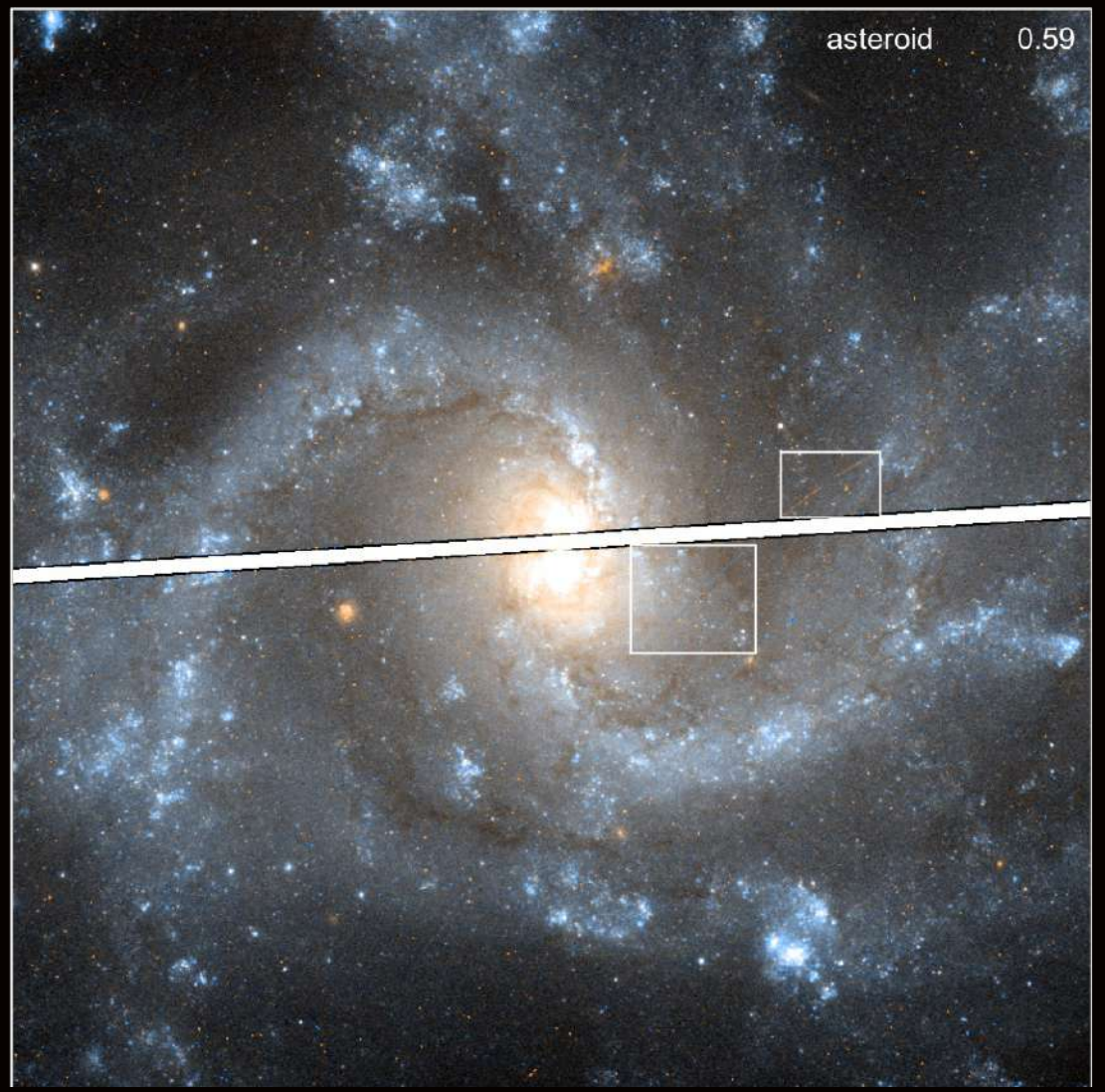
Asteroid trails identified by the volunteers in images

1,488 asteroid trails identified by 11,486 volunteers in the citizen science project from 2019-2020



Scaling up with AutoML object detection to identify trails

Classifying the entire HST archive
(2002 – 2021) of 37,324 images
(x4 = 150k cutouts)



Asteroid 2002 LX55 moving between two Hubble exposures in front of galaxy NGC 5468, detected with AutoML.

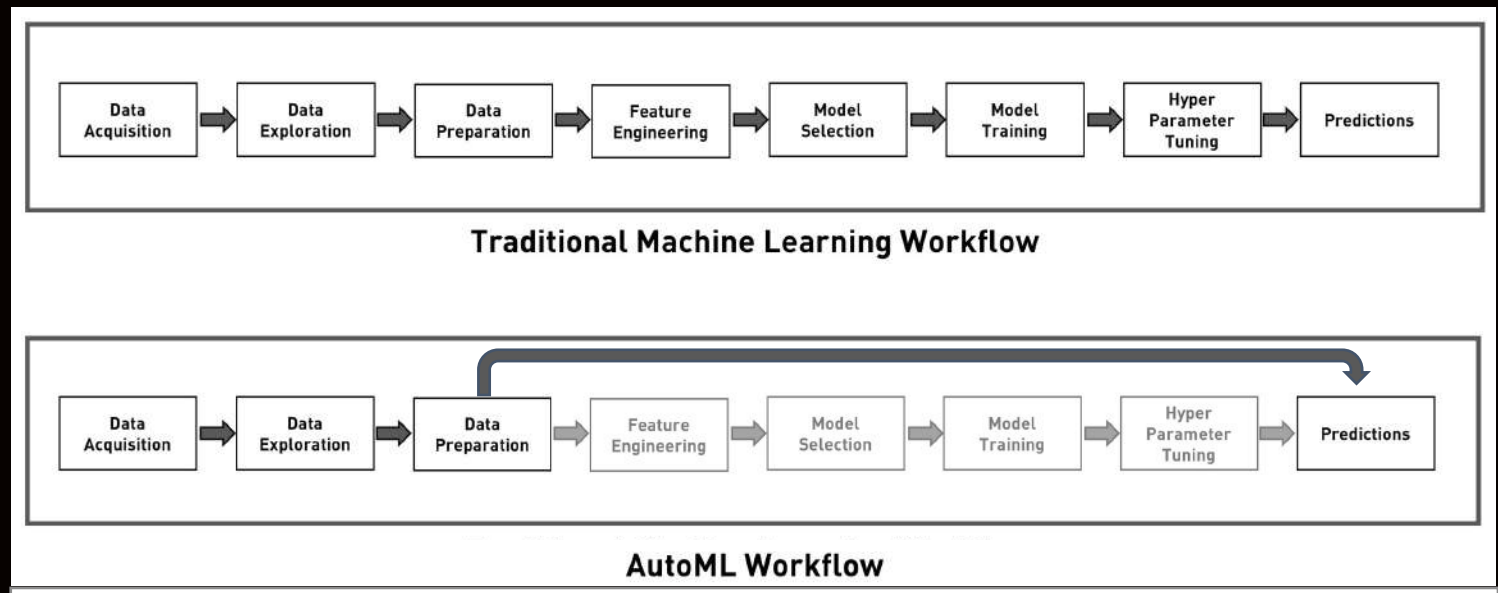
Scaling up with AutoML object detection to identify trails

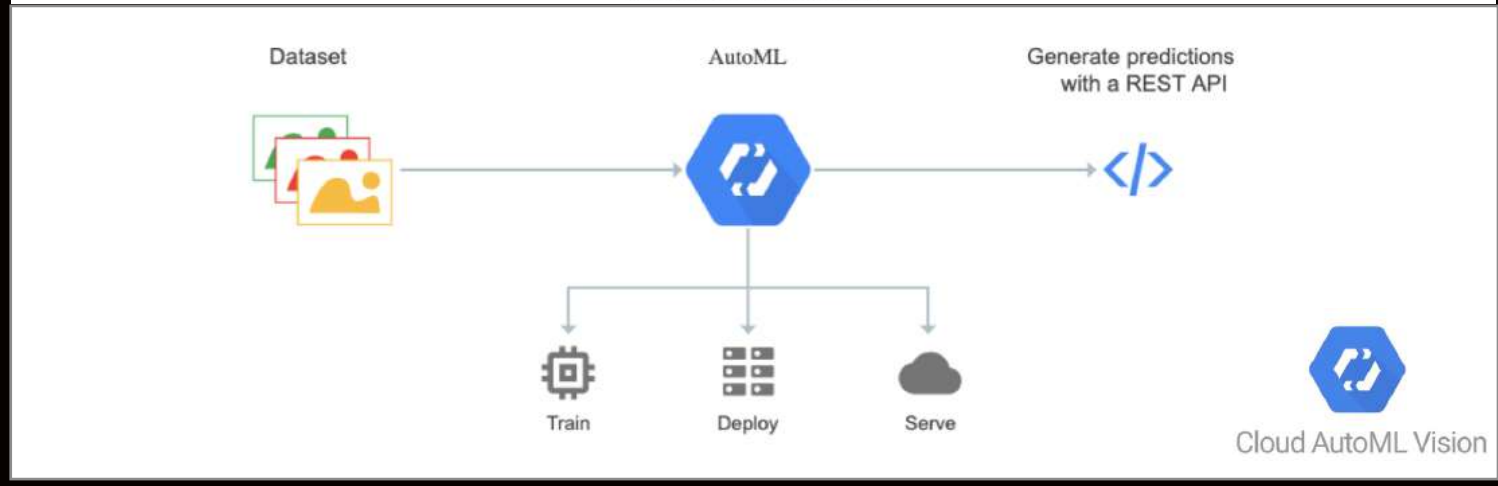
* In collaboration with Google

Classifying the entire HST archive (2002 – 2021) of 37,324 images (x4 = 150k cutouts)

Use automated machine learning (AutoML) on Google Cloud

Automating ML pipelines





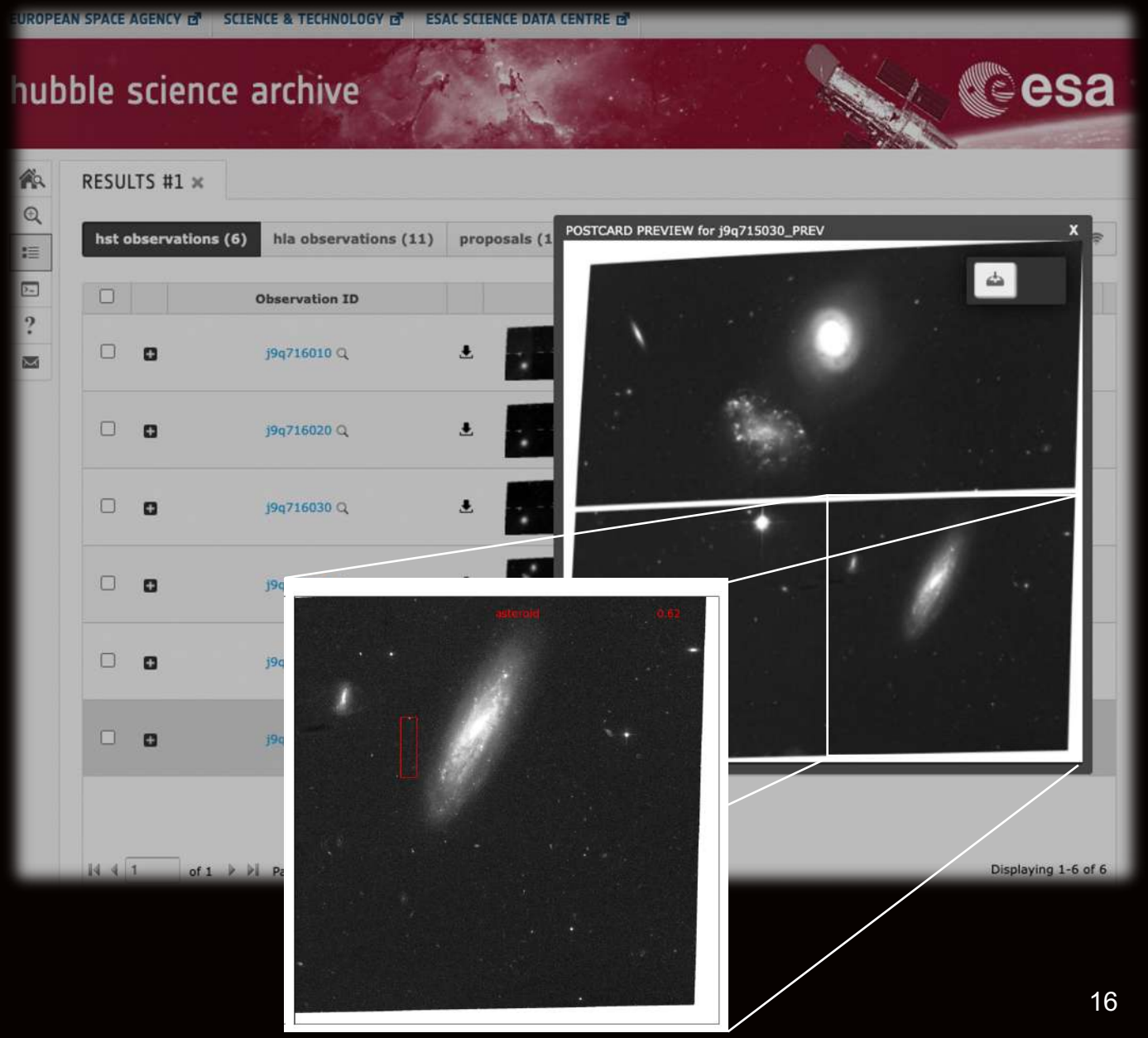
How AutoML works – using Neural Architecture Search

Scaling up with AutoML object detection to identify trails

Classifying the entire HST archive (2002 – 2021) of 37,324 images (x4 = 150k cutouts)

Use automated machine learning (AutoML) on Google Cloud

Scalable: training and batch classification on Google Cloud: ~10 hours

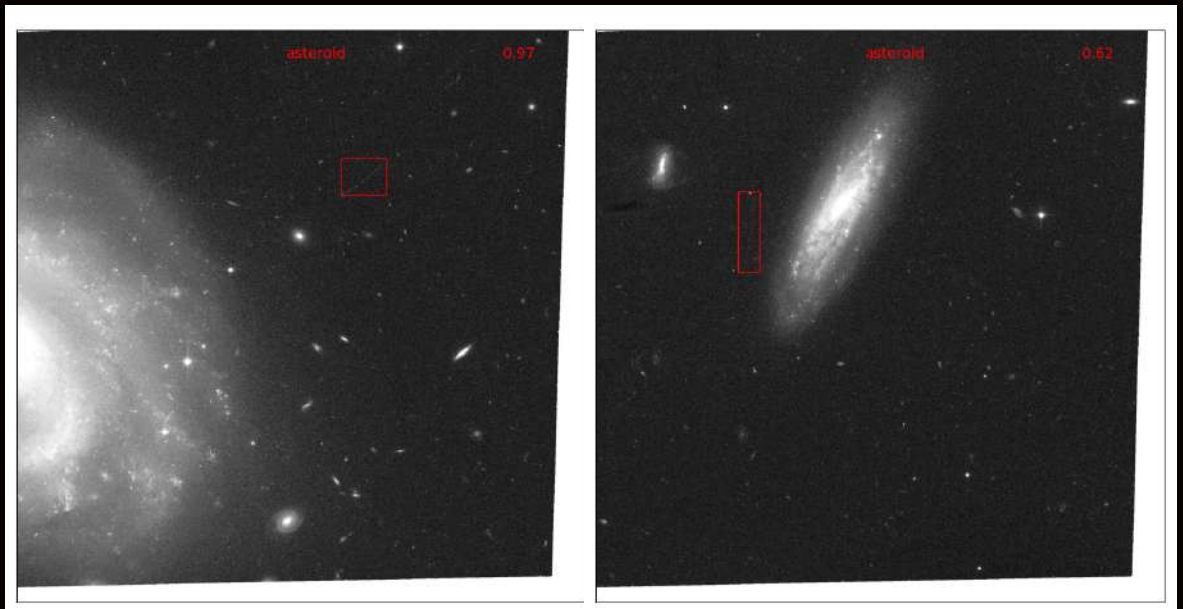


Scaling up with AutoML object detection to identify trails

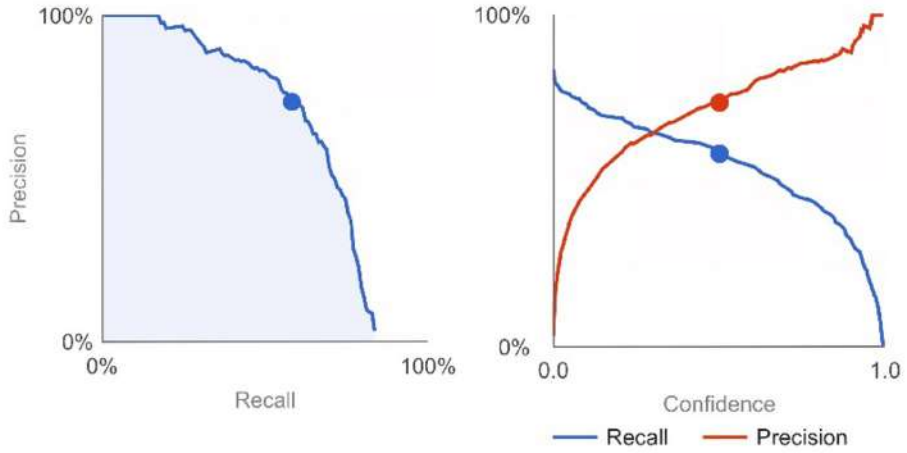
Classifying the entire HST archive (2002 – 2021) of 37,324 images (x4 = 150k cutouts)

Use automated machine learning (AutoML) on Google Cloud

Scalable: training and batch classification on Google Cloud: ~10 hours

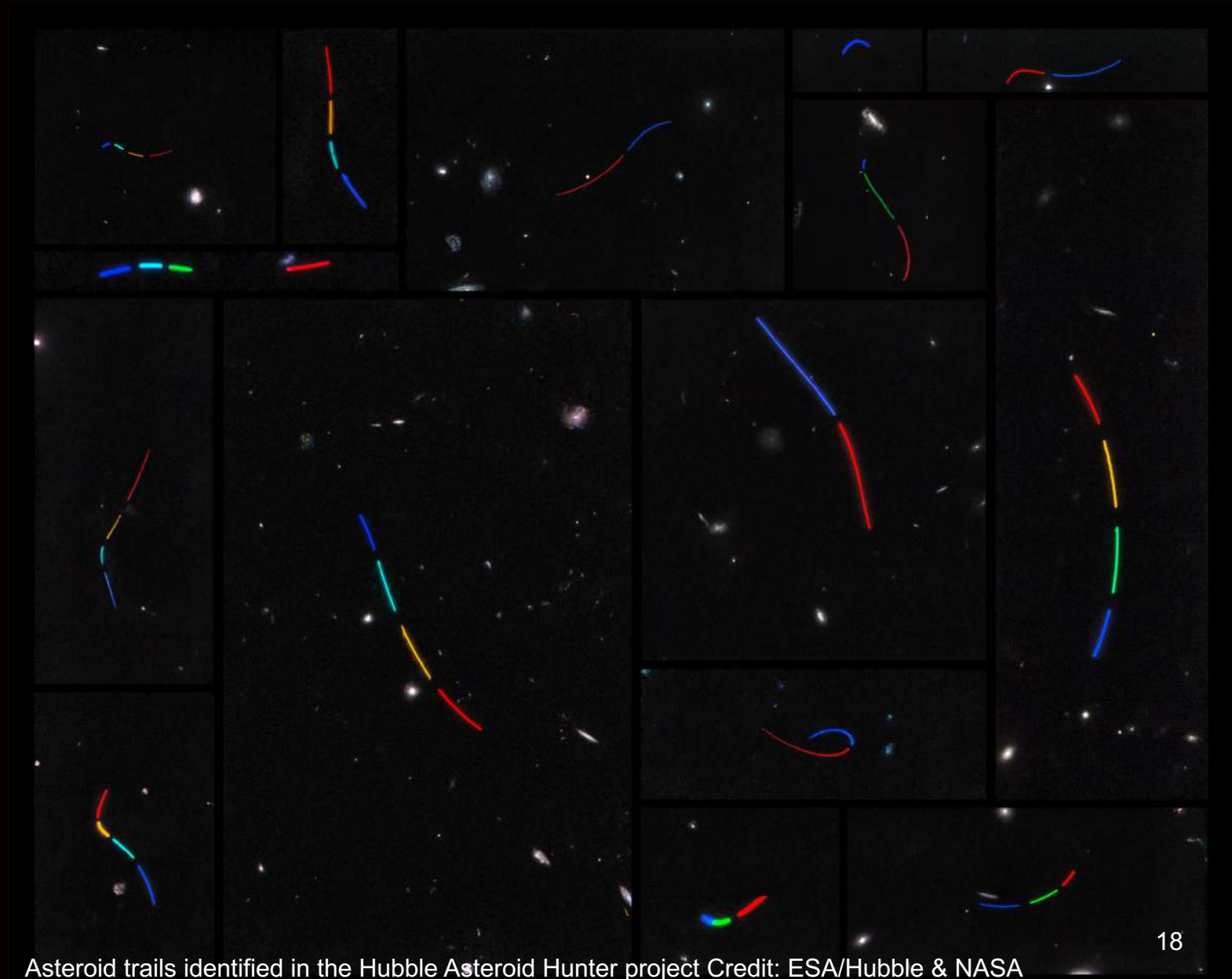


Performance in identifying asteroid trails



Results: Asteroids detected in the Hubble images

2487 asteroid trails recovered by citizen scientists and by AutoML



Asteroid trails identified in the Hubble Asteroid Hunter project Credit: ESA/Hubble & NASA

Results: Asteroids detected in the Hubble images

2487 asteroid trails recovered by citizen scientists and by AutoML

1701 asteroids validated by the team

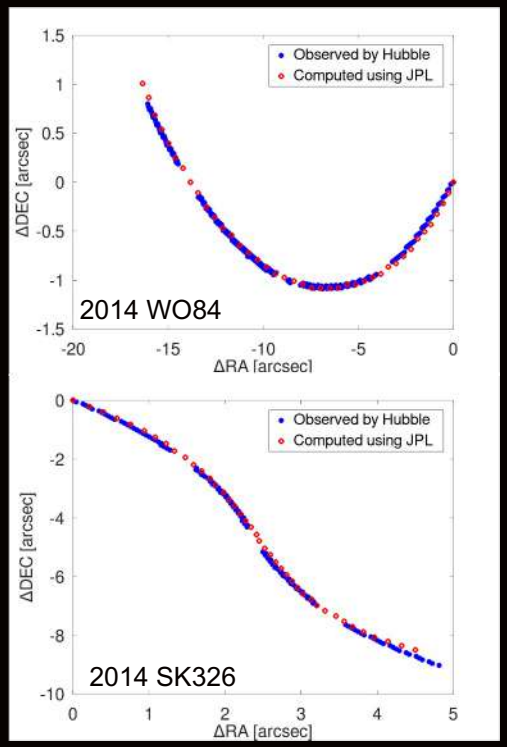
Asteroids moving between individual Hubble exposures

Matching trails with known asteroids from Minor Planet Center

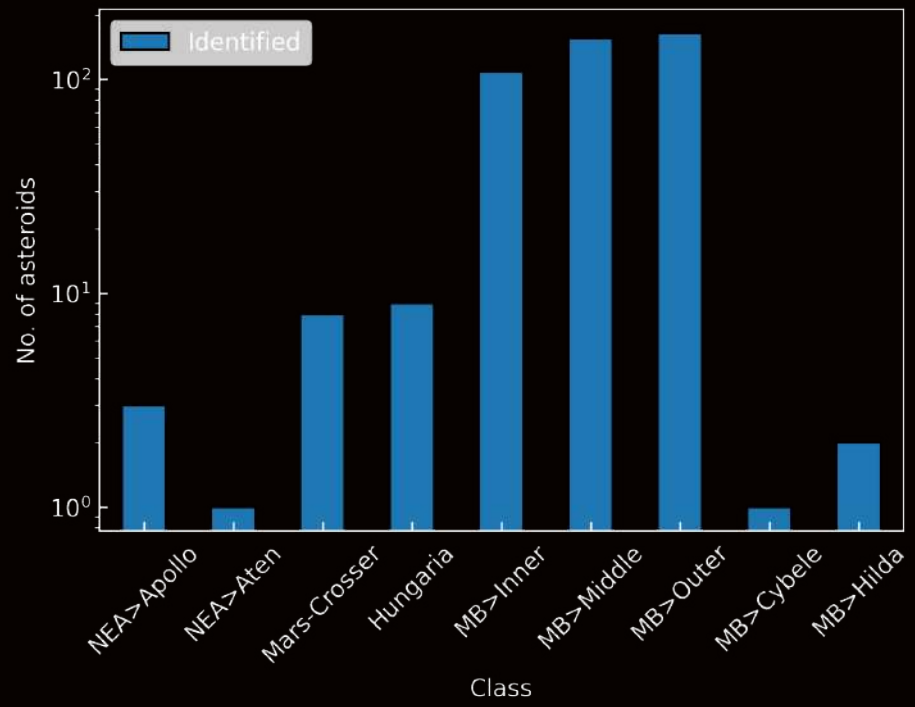
2487 asteroid trails recovered by citizen scientists and by AutoML

1701 asteroids validated by the team

670 asteroids matched with known objects. 95% are Main Belt.



Class distribution of 670 known asteroids



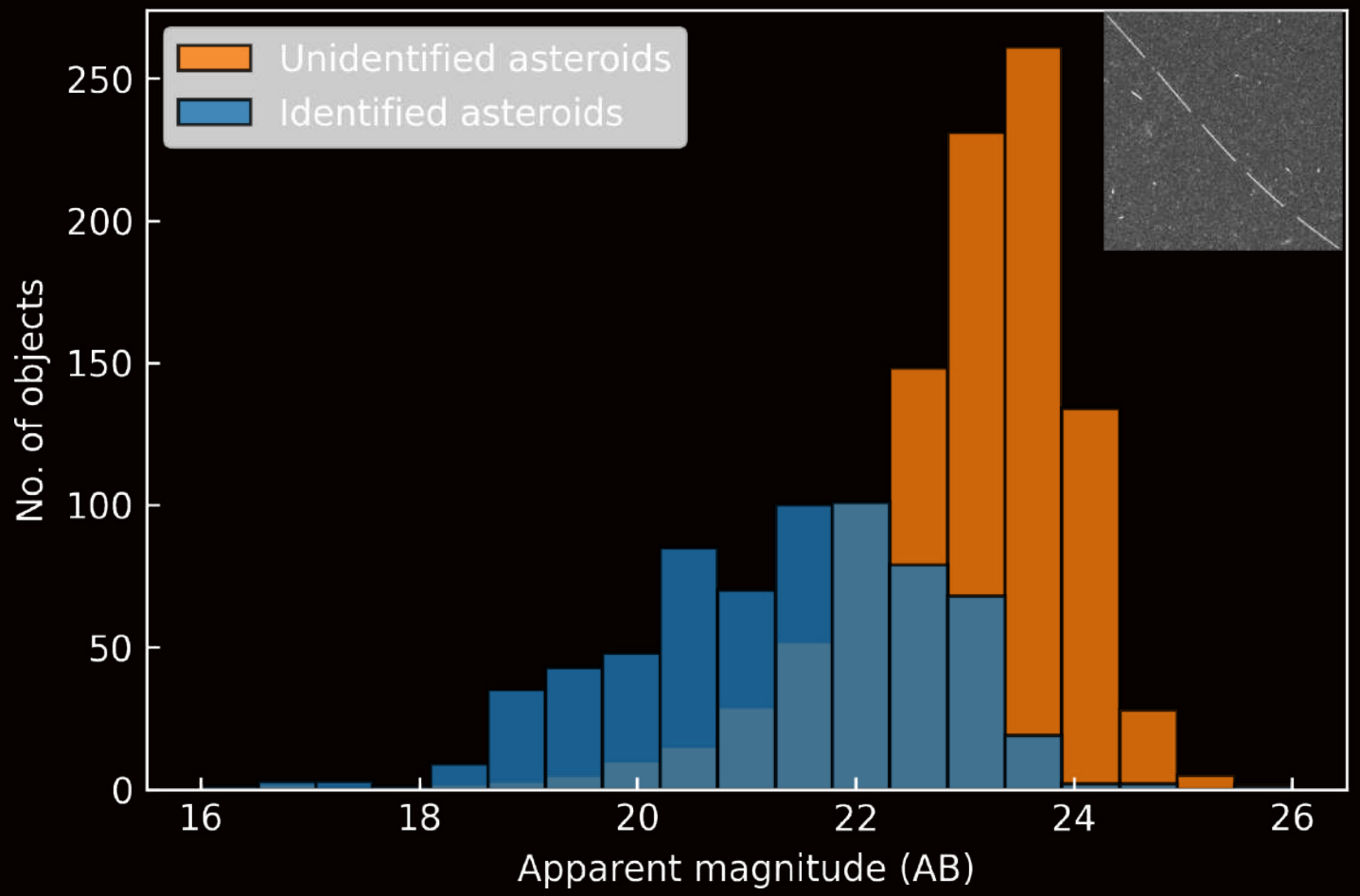
Results: magnitude distribution of detected asteroids

2487 asteroid trails recovered by citizen scientists and by AutoML

1701 asteroids validated by the team

670 asteroids matched with known objects. 95% are Main Belt.

1031 unidentified asteroid trails
-- previously unknown asteroids?



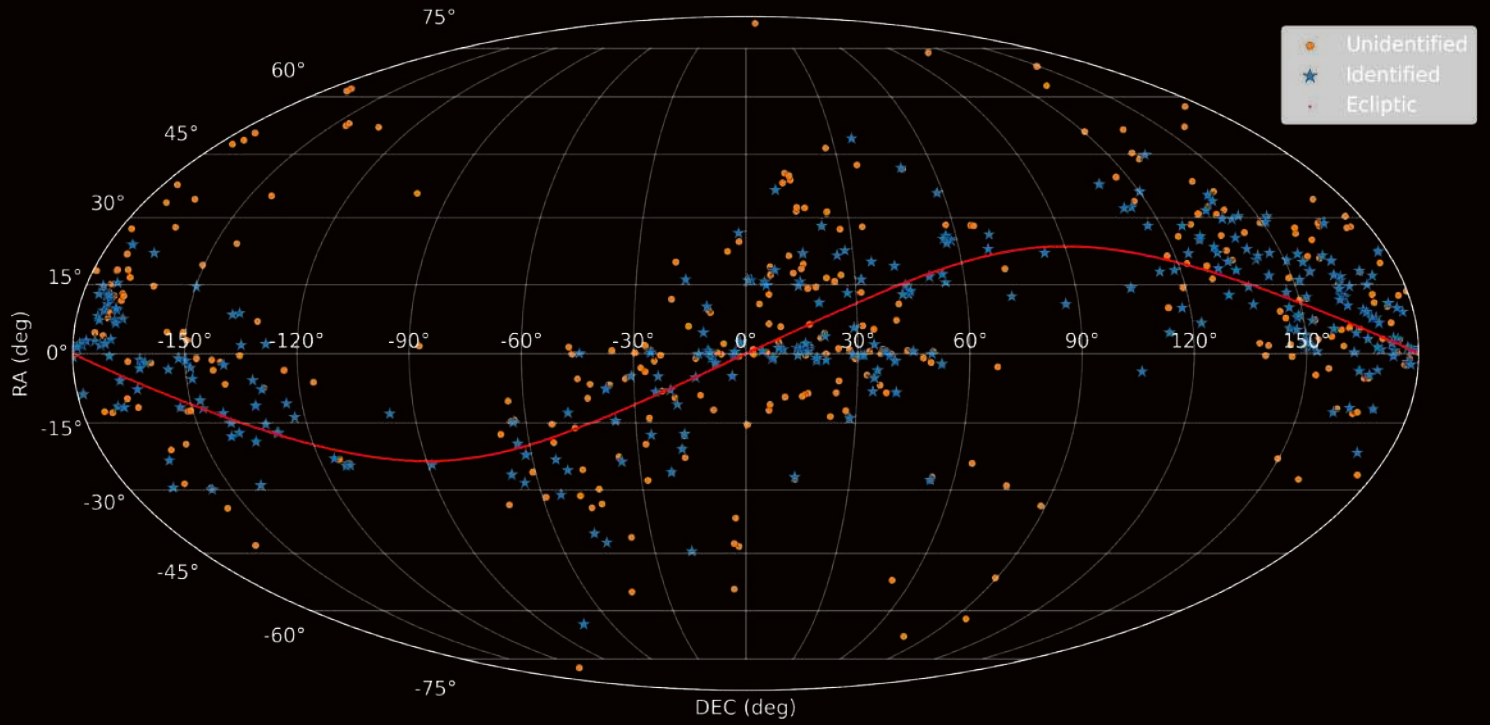
Results: sky distribution of detected asteroids

2487 asteroid trails recovered by citizen scientists and by AutoML

1701 asteroids validated by the team

670 asteroids matched with known objects. 95% are Main Belt.

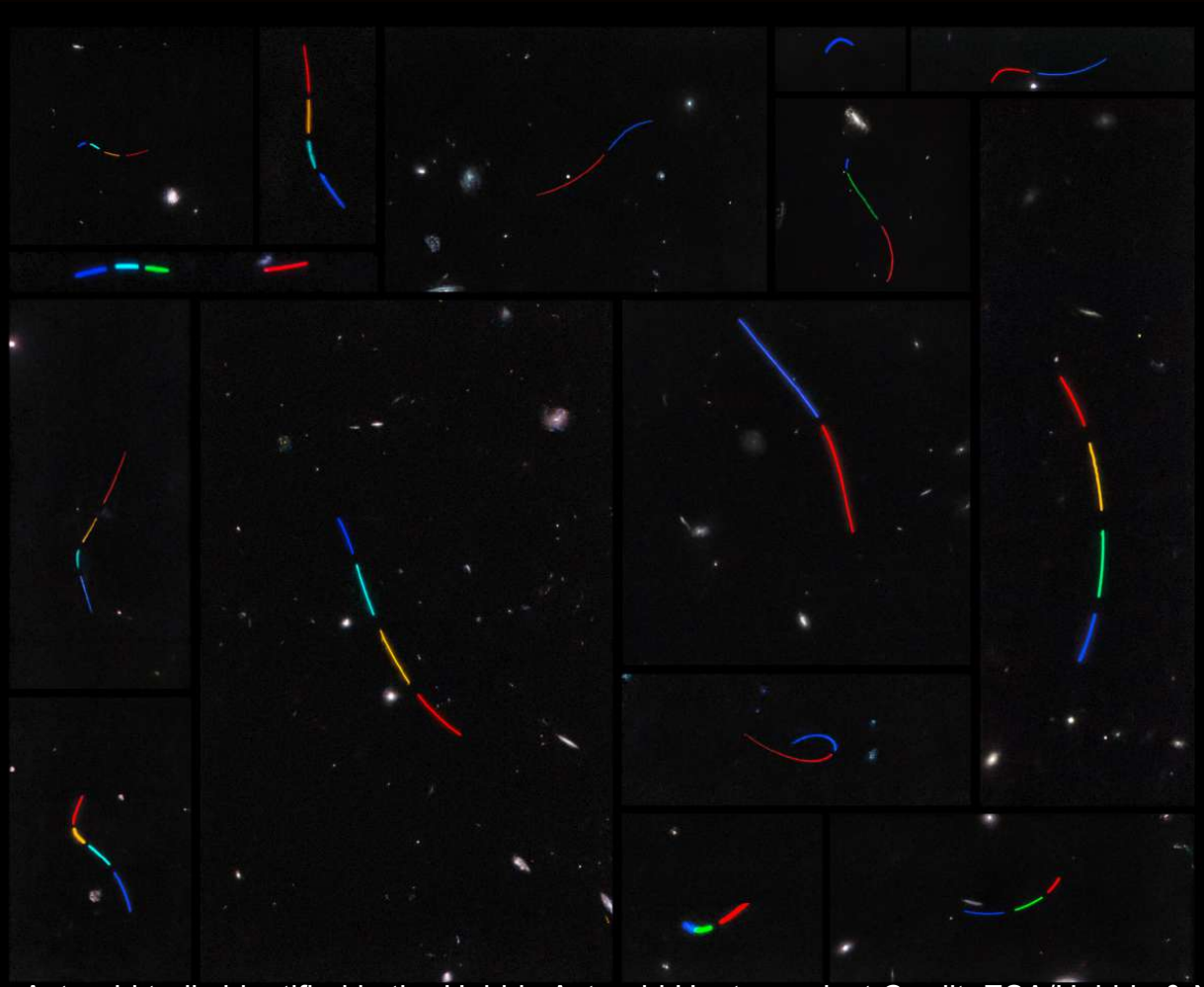
1031 unidentified asteroid trails
-- previously unknown asteroids?



Follow-up work: determining asteroid sizes and orbits

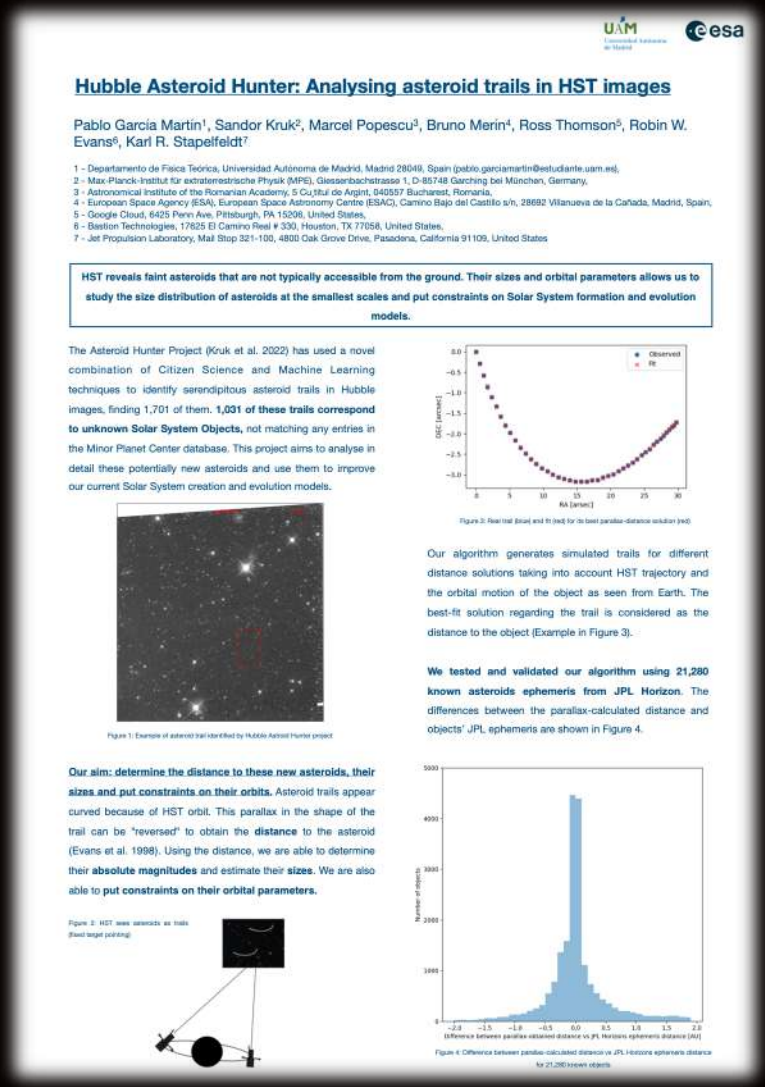
Likely previously unknown small size Main-Belt asteroids

*studied further in García-Martín, Kruk et al. in prep.



Asteroid trails identified in the Hubble Asteroid Hunter project Credit: ESA/Hubble & NASA

See Pablo's poster



Hubble Asteroid Hunter: Analysing asteroid trails in HST images

Pablo García Martín[1], Sandor Kruk[2], Marcel Popescu[3], Bruno Merín[4], Ross Thomson[5], Robin W. Evans[6], Karl R. Stapelfeldt[7]

1 - Departamento de Física Teórica, Universidad Autónoma de Madrid, Madrid 28049, Spain (pablo.garciamartin@estudiante.uam.es),
2 - Max-Planck-Institut für extraterrestrische Physik (MPE), Giessenbachstrasse 1, D-85748 Garching bei München, Germany,
3 - Astronomical Institute of the Romanian Academy, 5 Cuțitul de Argint, 040557 Bucharest, Romania,
4 - European Space Agency (ESA), European Space Astronomy Centre (ESAC), Camino Bajo del Castillo s/n, 28692 Villanueva de la Cañada, Madrid, Spain,
5 - Google Cloud, 6425 Penn Ave, Pittsburgh, PA 15206, United States,
6 - Bastion Technologies, 17625 El Camino Real # 330, Houston, TX 77058, United States,
7 - Jet Propulsion Laboratory, Mail Stop 321-100, 4800 Oak Grove Drive, Pasadena, California 91109, United States

HST reveals faint asteroids that are not typically accessible from the ground. Their sizes and orbital parameters allows us to study the size distribution of asteroids at the smallest scales and put constraints on Solar System formation and evolution models.

The Asteroid Hunter Project (Kruk et al. 2022) has used a novel combination of Citizen Science and Machine Learning techniques to identify serendipitous asteroid trails in Hubble images, finding 1,701 of them. **1,031 of these trails correspond to unknown Solar System Objects,** not matching any entries in the Minor Planet Center database. This project aims to analyse in detail these potentially new asteroids and use them to improve our current Solar System creation and evolution models.

Figure 1: Example of asteroid trail identified by Hubble Asteroid Hunter project

Our aim: determine the distance to these new asteroids, their sizes and put constraints on their orbits. Asteroid trails appear curved because of HST orbit. This parallax in the shape of the trail can be "reversed" to obtain the **distance** to the asteroid (Evans et al. 1998). Using the distance, we are able to determine their **absolute magnitudes** and estimate their **sizes**. We are also able to **put constraints on their orbital parameters.**

Figure 2: HST sees asteroids as trails (fixed target pointing)

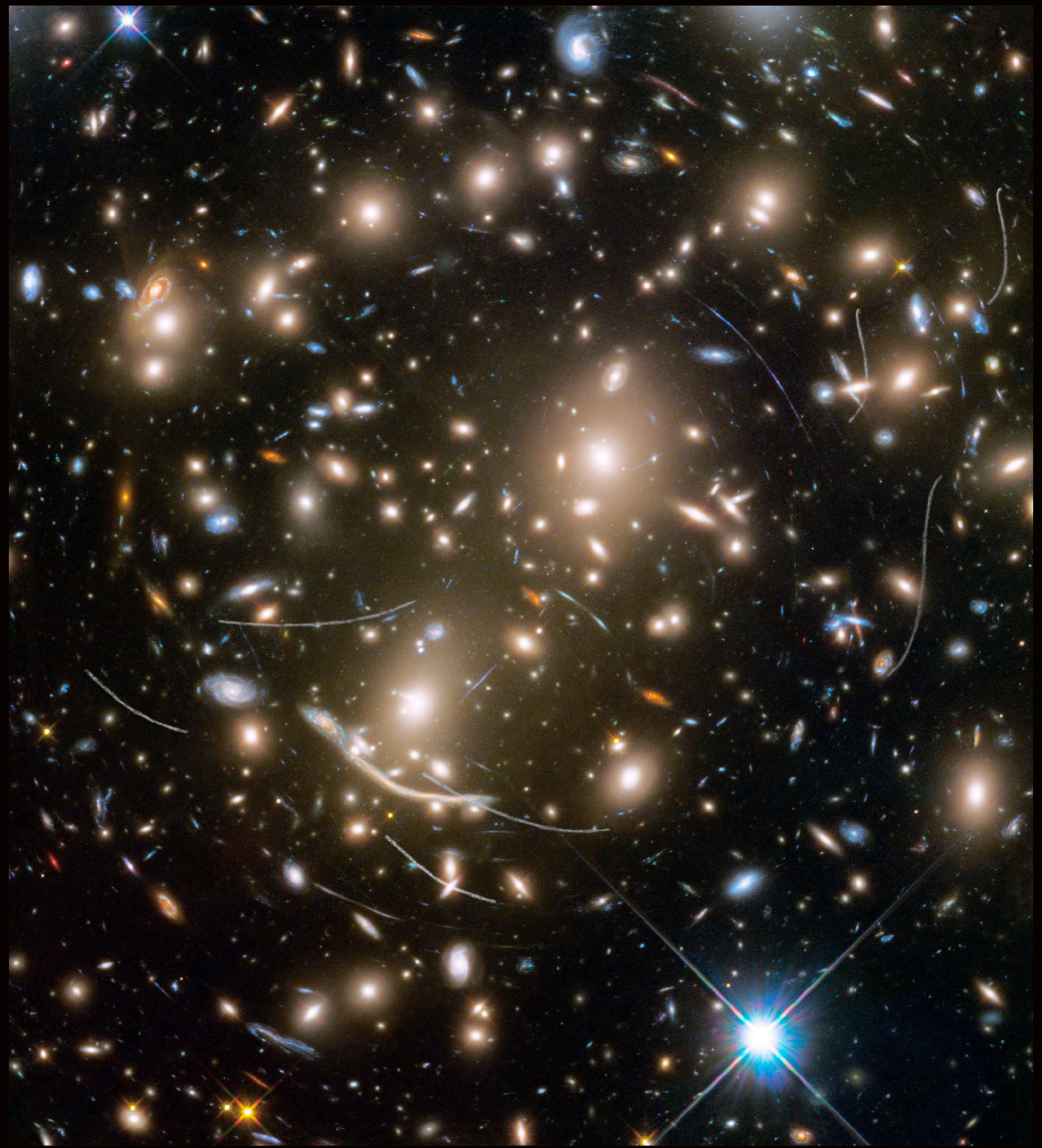
Figure 3: Real trail (blue) and fit (red) for its best parallax-distance solution (red)

Our algorithm generates simulated trails for different distance solutions taking into account HST trajectory and the orbital motion of the object as seen from Earth. The best-fit solution regarding the trail is considered as the distance to the object (Example in Figure 3).

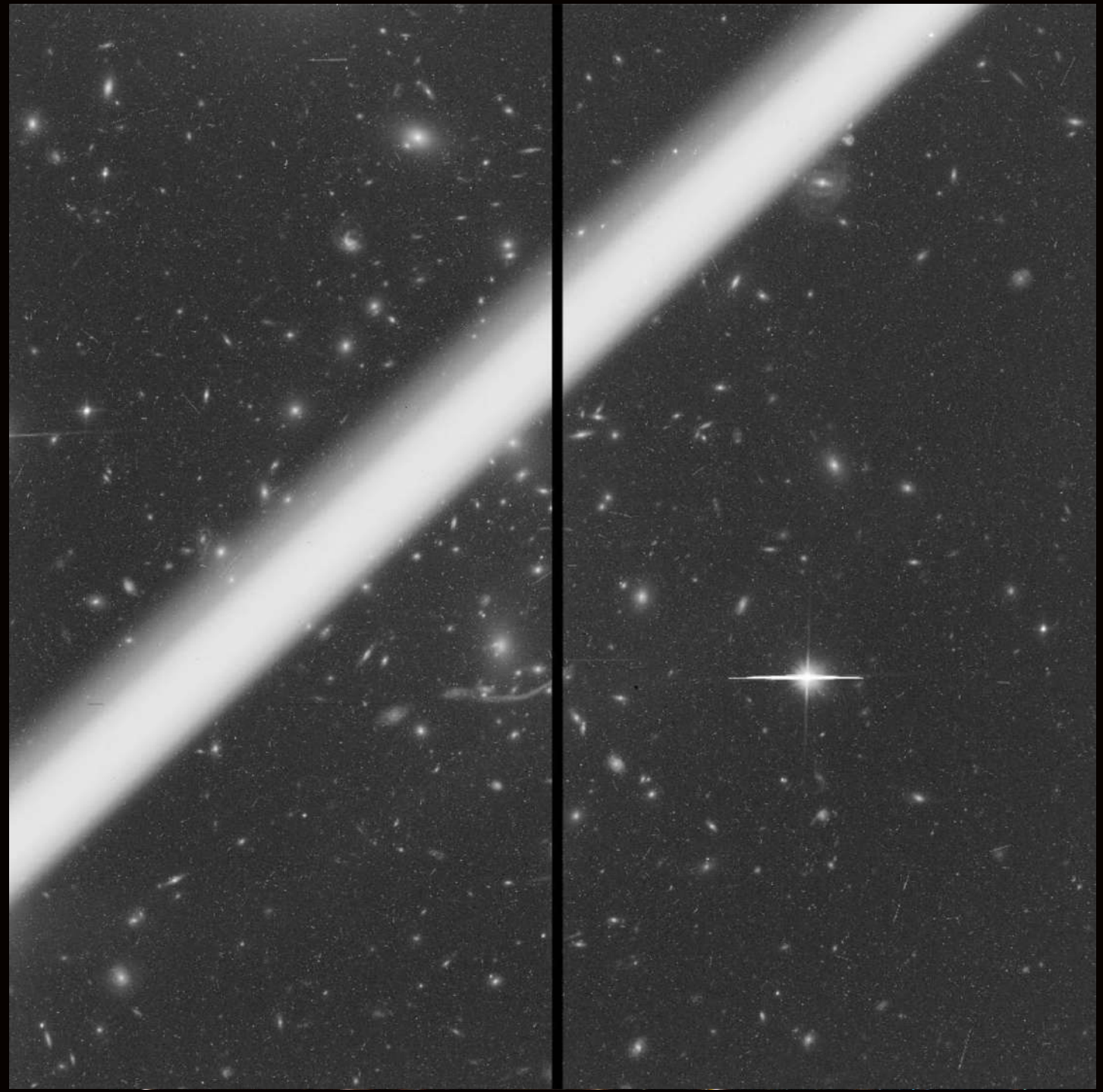
We tested and validated our algorithm using 21,280 known asteroids ephemeris from JPL Horizon. The differences between the parallax-calculated distance and objects' JPL ephemeris are shown in Figure 4.

Figure 4: Difference between parallax-calculated distance vs JPL Horizons ephemeris distance for 21,280 known objects

Citizen scientists
make new discoveries



Asteroids observed serendipitously in image of Frontier Fields cluster Abell 370
Credit: NASA, ESA/Hubble



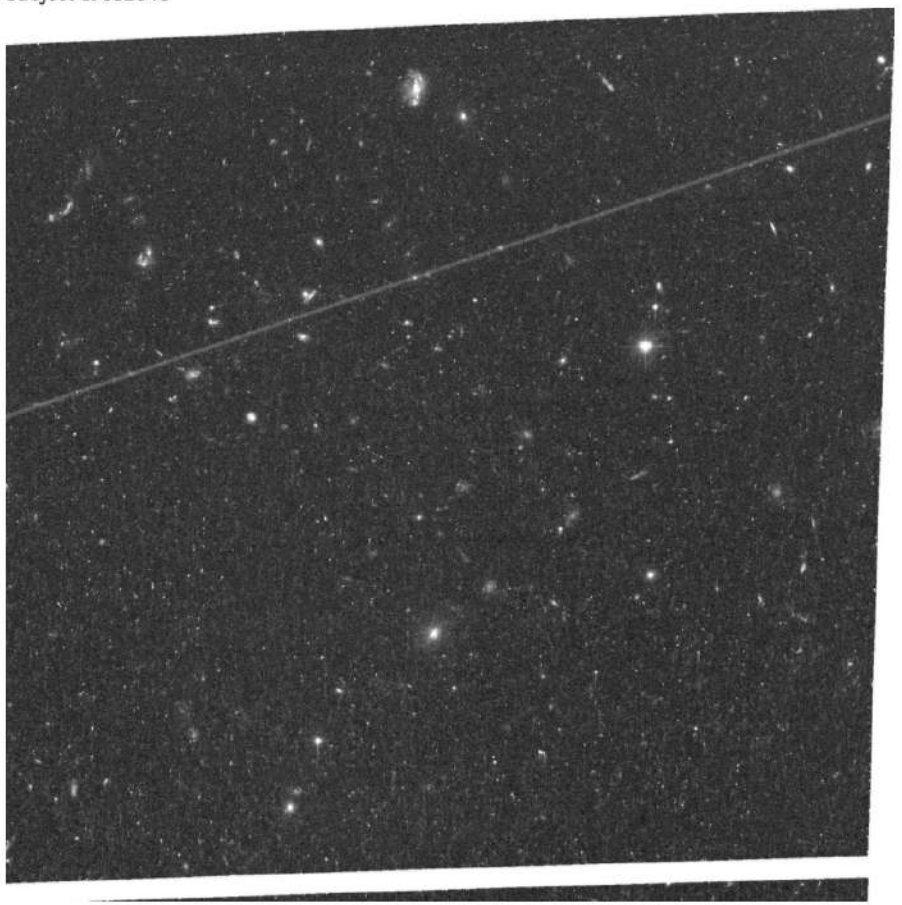
Frontier Fields cluster Abell 370
Trail probably Chinese Long March 4C Y33 third stage passing 34km above HST.

Satellite ID by J. McDowell
Image credit: Judy Schmidt

Hubble Asteroid Hunter Talk

Search or enter a #tag

Subject 37652046



Comments:

December 16th 2019, 11:51 am

Not clear what the grey line across the image could be. Perhaps a common #satellite ?

Fauxwise

@Fauxwise

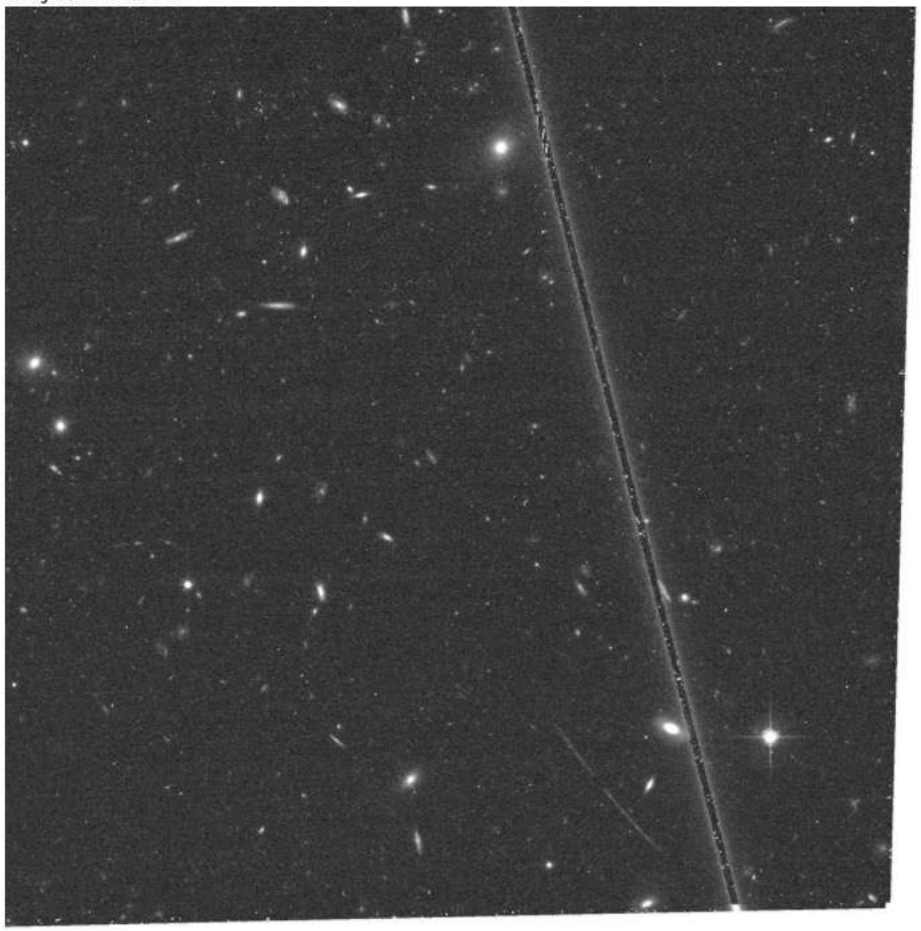
Hubble Asteroid Hunter

Happy international asteroid day! The current d

Hubble Asteroid Hunter Talk

Search or enter a #tag

Subject 42950351

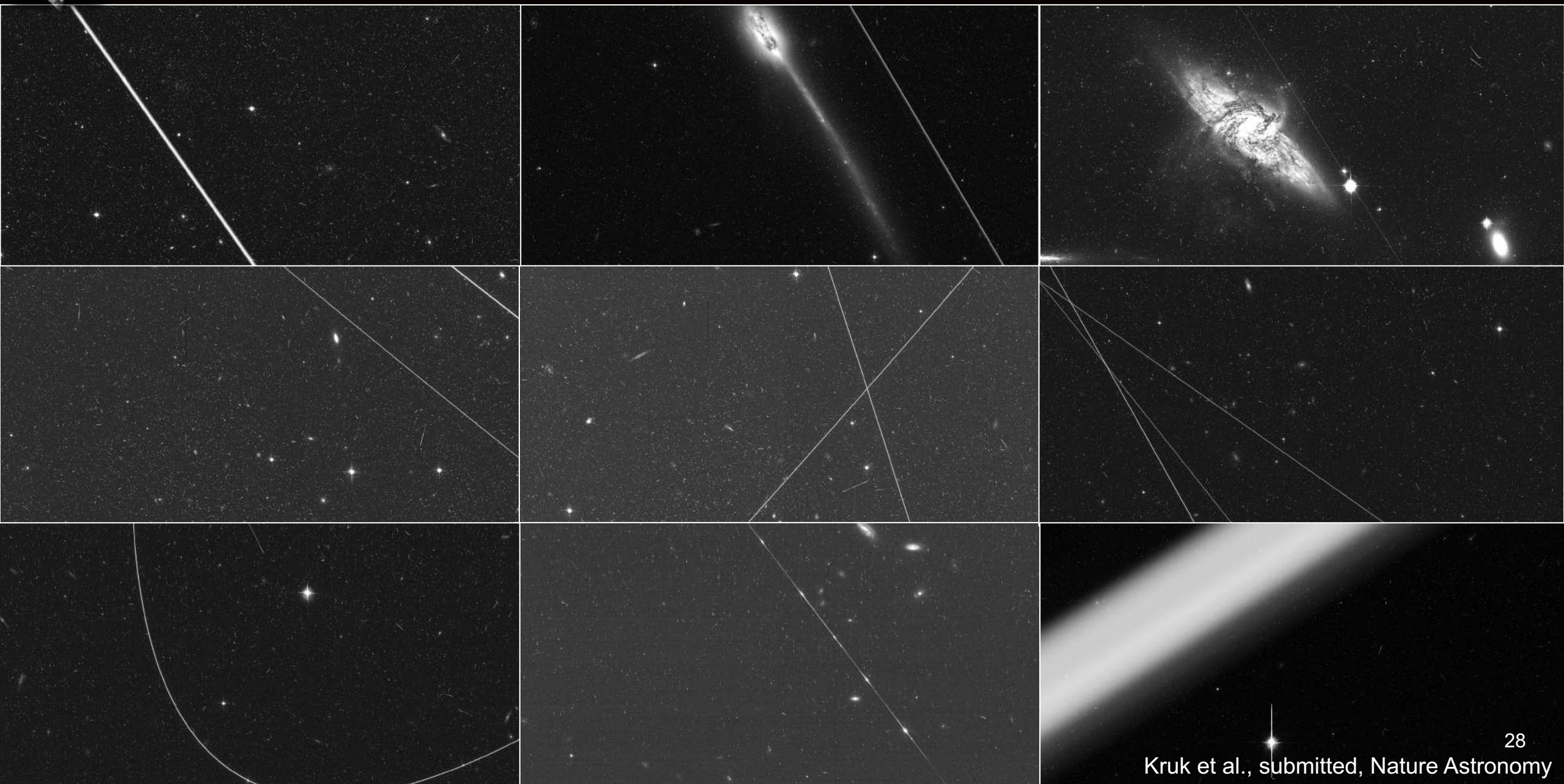


Travaglino

@Travaglino

#asteroid and #satellite

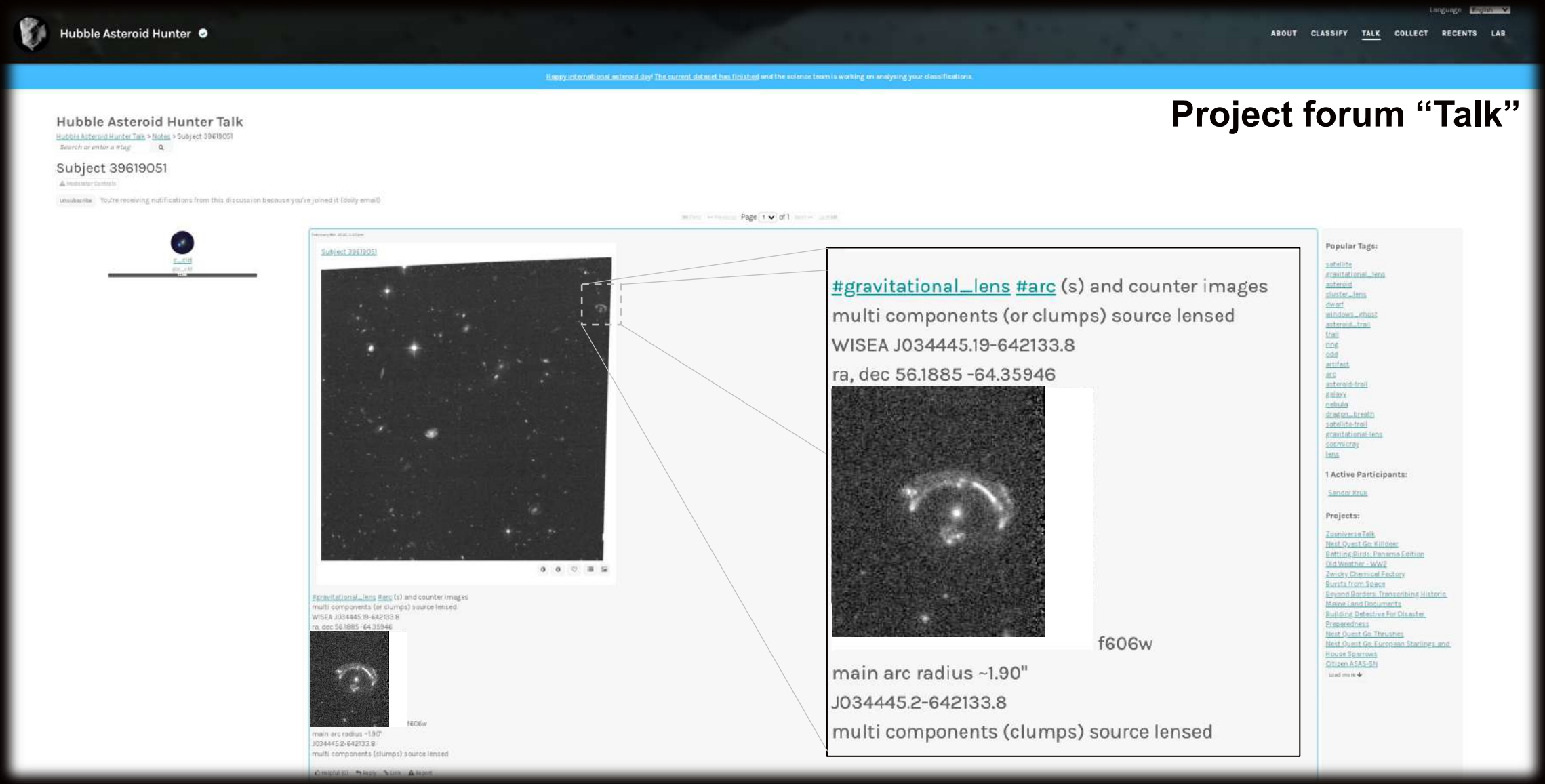
Satellite trails identified with AutoML in Hubble exposures



Kruk et al., submitted, Nature Astronomy

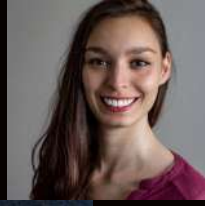
Serendipitous findings: new strong gravitational lenses

Project forum “Talk”



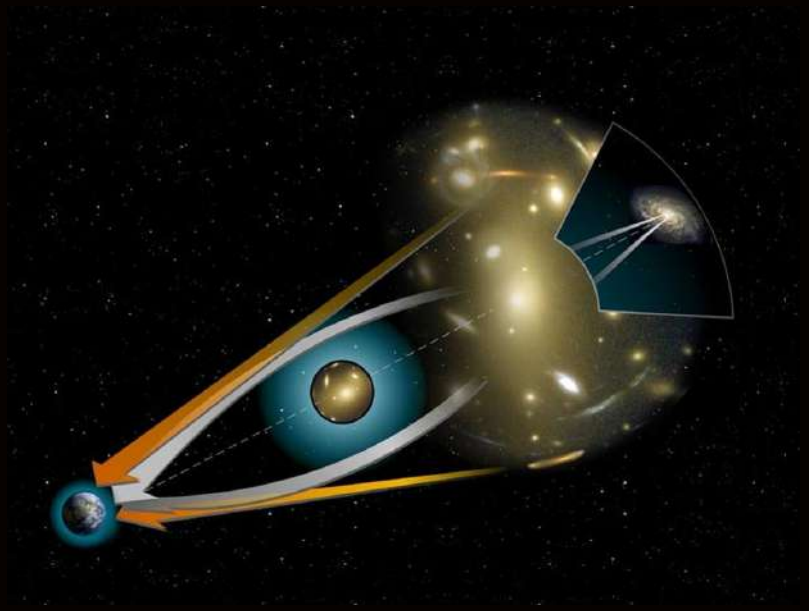
https://www.zooniverse.org/projects/sandorkruk/hubble-asteroid-hunter/talk/

Serendipitous findings: new strong gravitational lenses

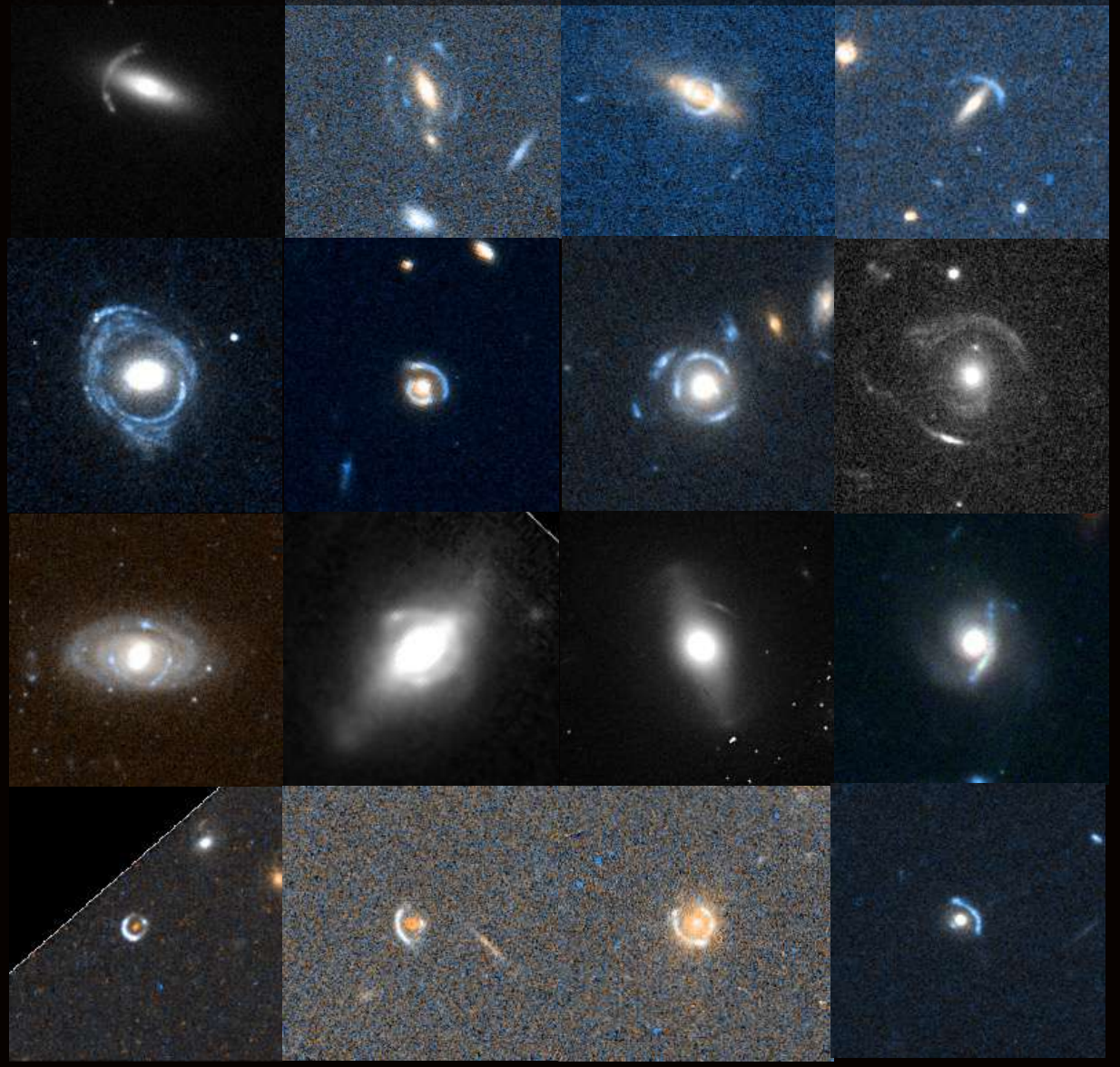


Having human eyes on the data can lead to new discoveries:

- **Discovery of 198 new strong gravitational lenses**



New strong gravitational lenses identified in Hubble Space Telescope images

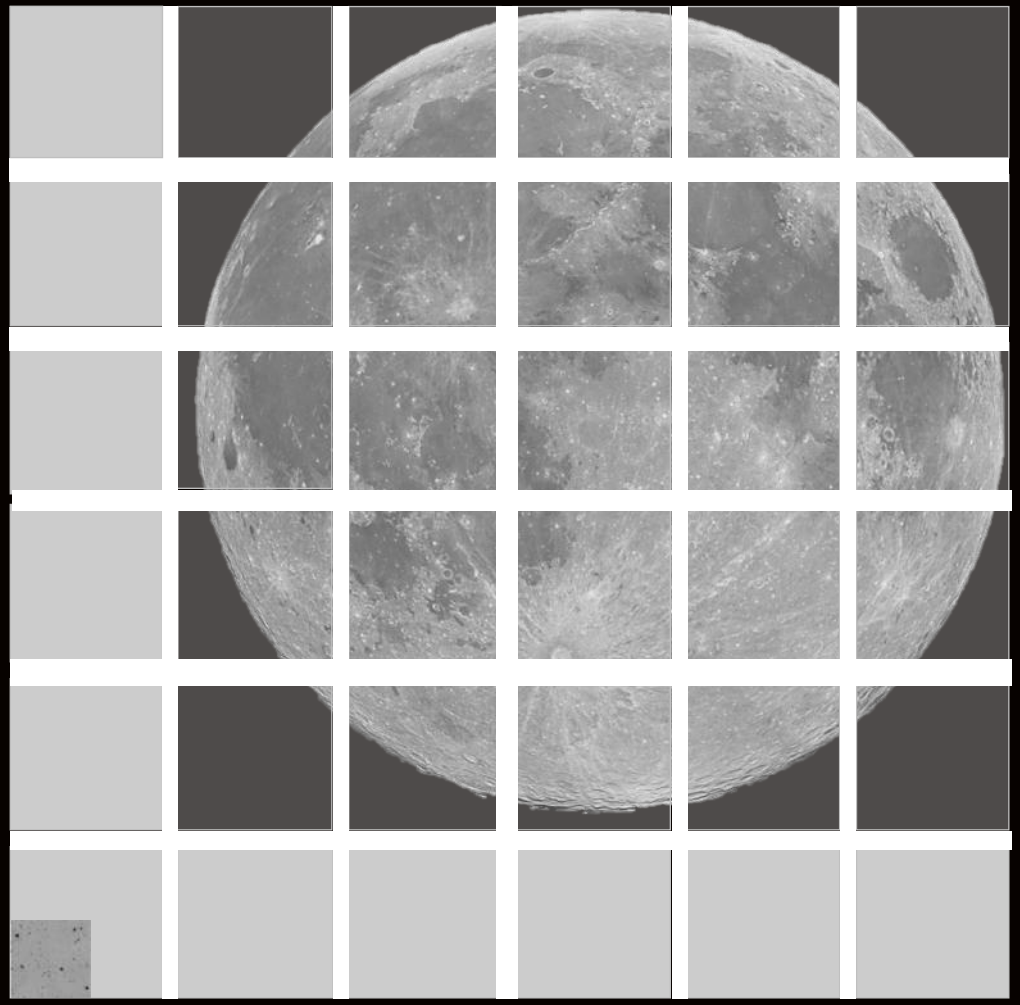


Garvin, Kruk et al. 2022, submitted, A&A


Future: ESA Euclid Mission

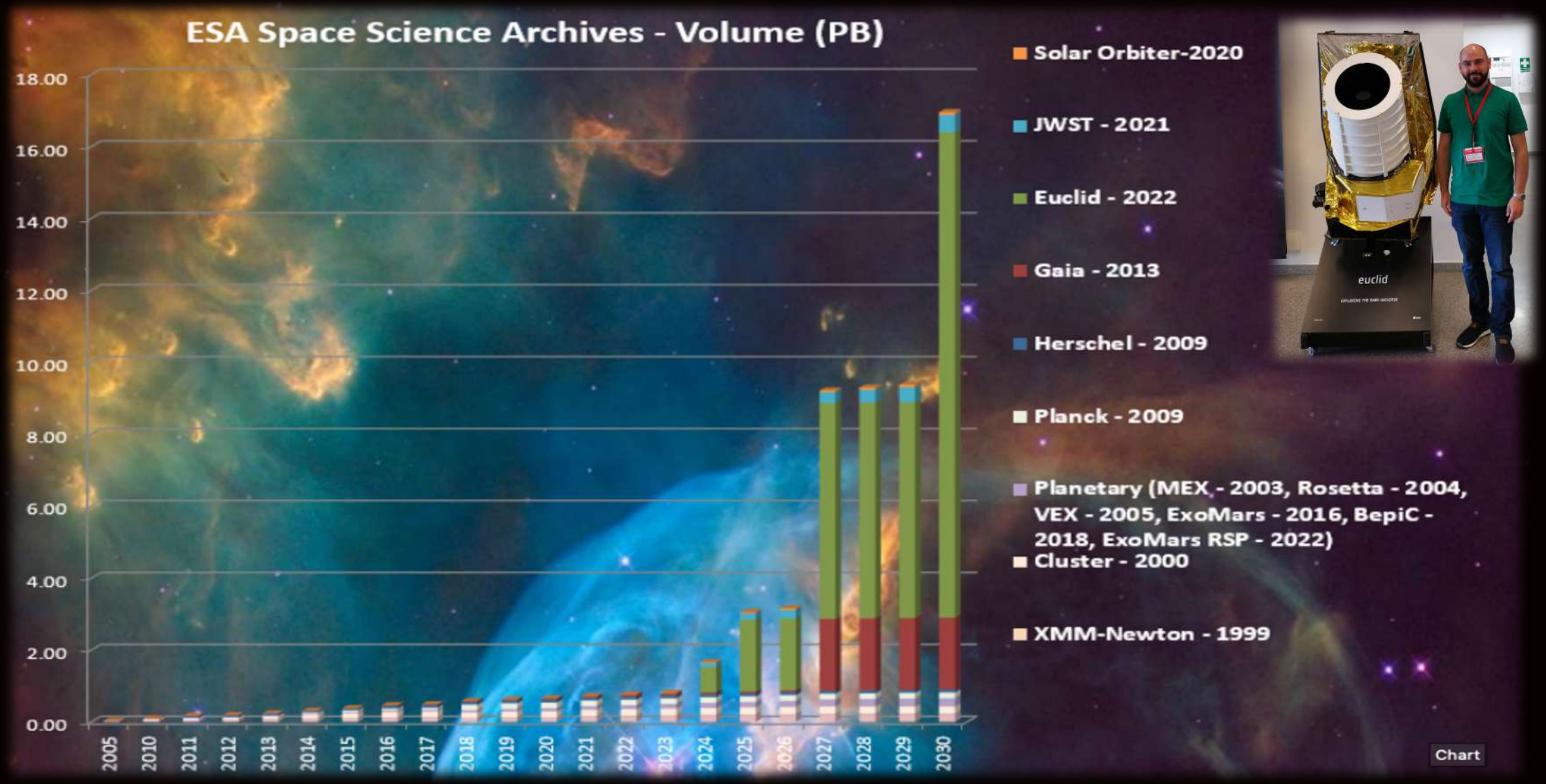
Euclid: providing a high-definition view of 1/3 of the sky

Single Euclid exposure (1/60,000th of the survey)



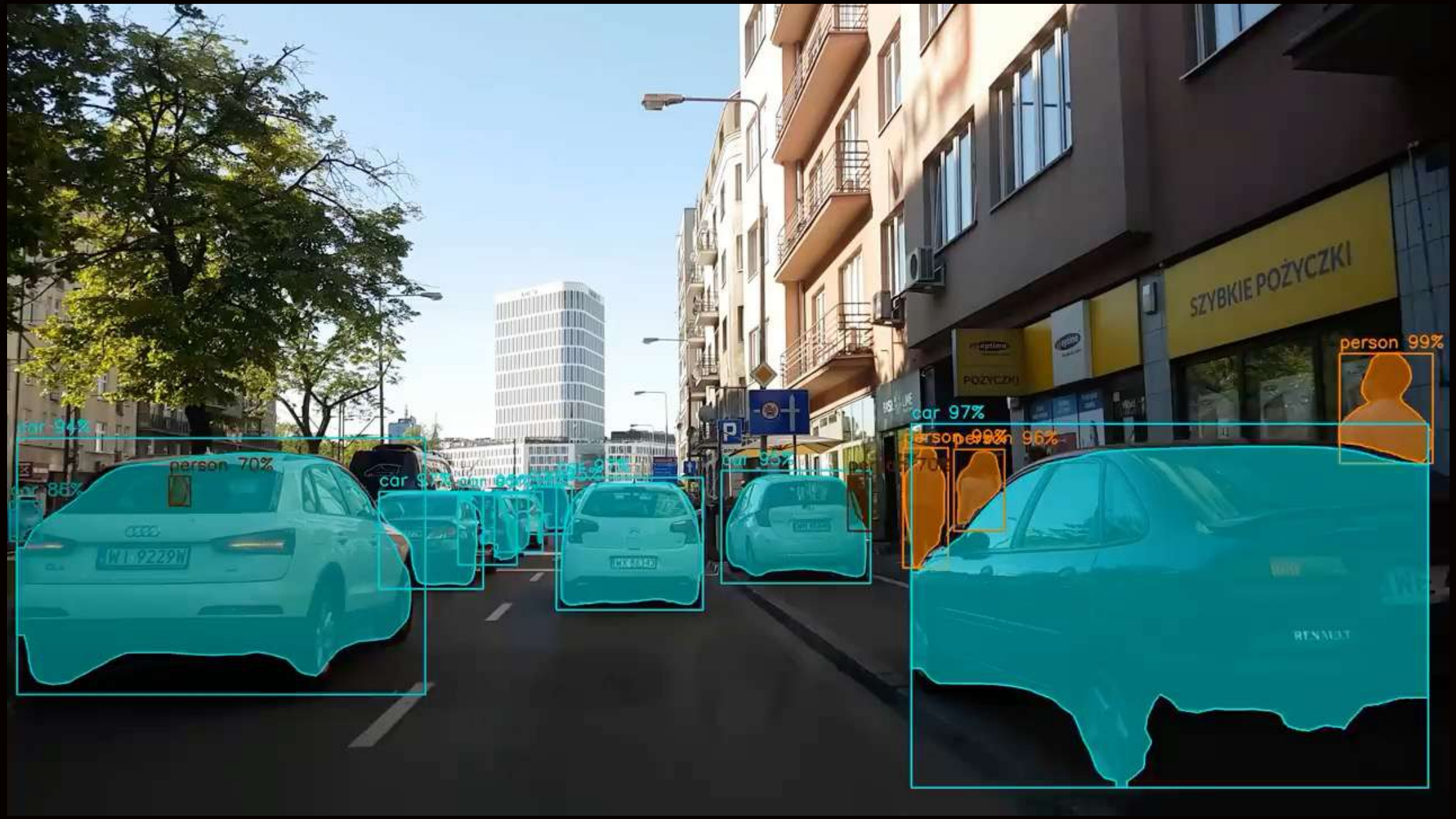
Single Hubble exposure

Euclid will add 10+ PB of data in the next years

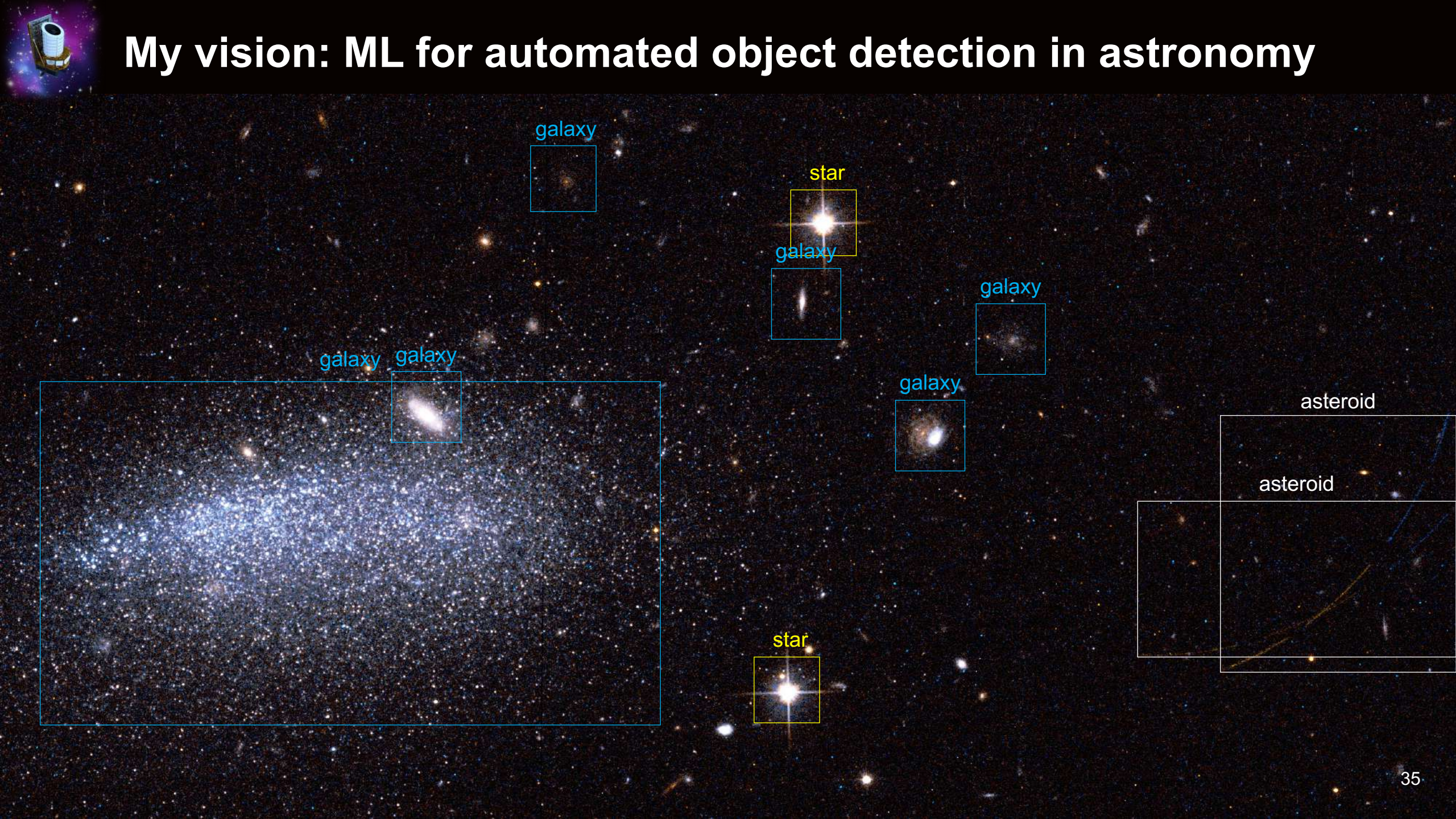


Credit: ESAC Science Data Center

My vision: ML for automated object detection in astronomy



My vision: ML for automated object detection in astronomy



Take home messages:

Artificial intelligence and crowdsourcing

New tools are needed to analyse and mine the increasingly large datasets. Human (crowdsourcing) and machine collaboration is important to avoid the garbage-in garbage-out problem of ML.

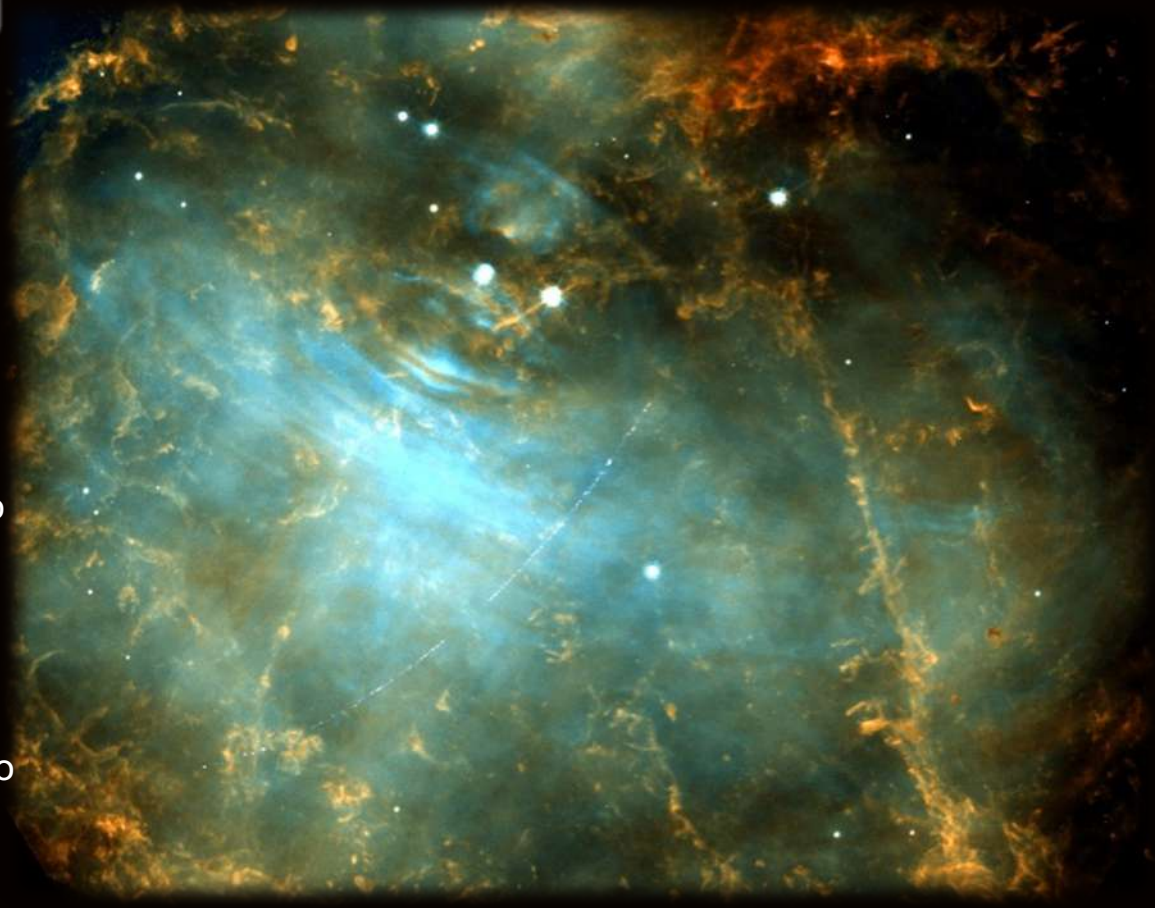
Citizen scientists make new discoveries

Look at the data! Sometimes unexpected things might hide in there. We searched the Hubble archives for asteroids, but also found artificial satellites and strong gravitational lenses.

Data archives are important

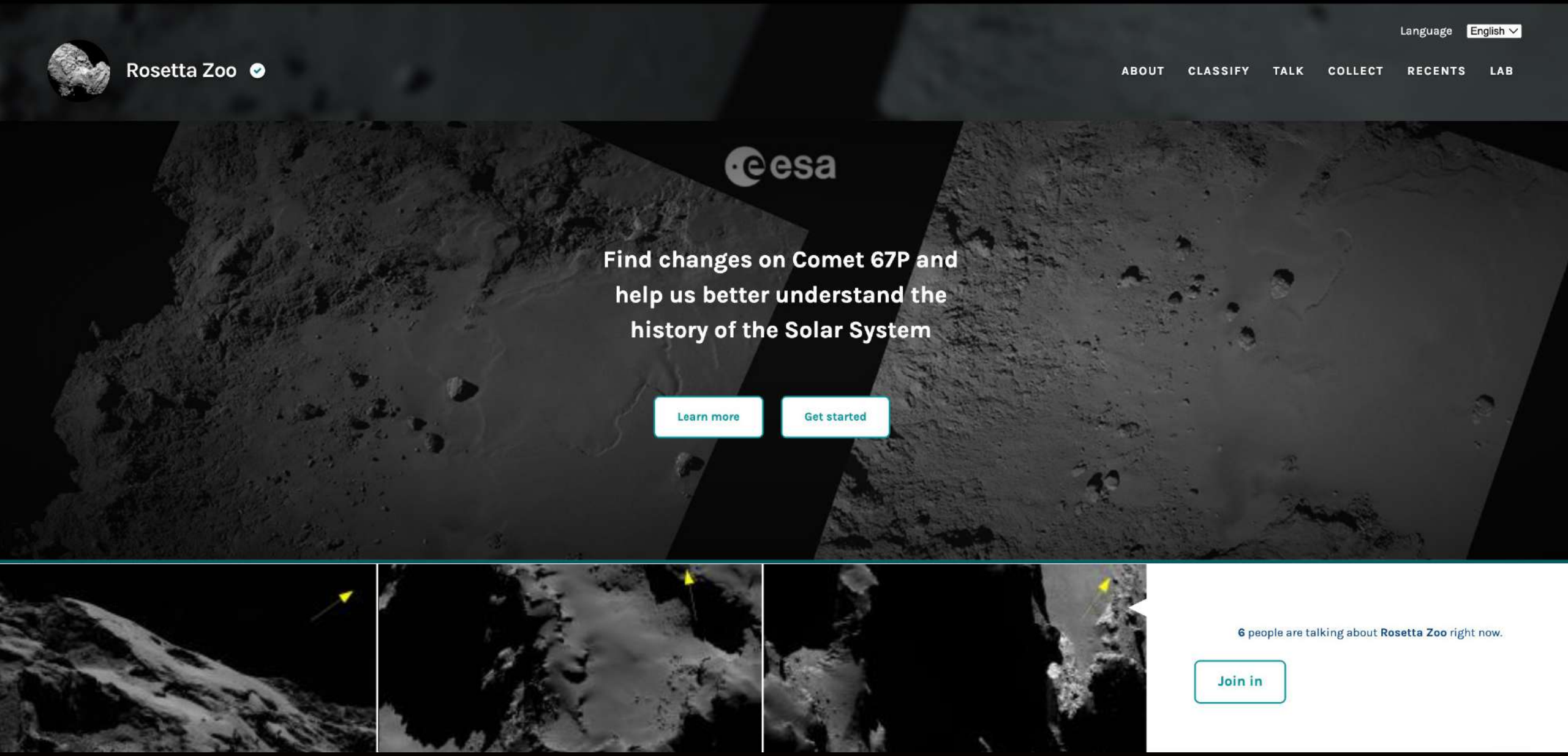
Hubble provides a rich data archive spanning decades. Ideal to survey faint asteroids and strong gravitational lenses.

One astronomer's trash can be another one's treasure



Foreground asteroid passing in front of the Crab Nebula, identified in the Hubble Asteroid Hunter project. Credit: ESA/Hubble & NASA, M. Thévenot

New ESA-Zooniverse citizen science project: Rosetta Zoo



https://www.zooniverse.org/projects/ellenjj/rosetta-zoo

Take home messages:

Artificial intelligence and crowdsourcing

New tools are needed to analyse and mine the increasingly large datasets. Human (crowdsourcing) and machine collaboration is important to avoid the garbage-in garbage-out problem of ML.

Citizen scientists make new discoveries

Look at the data! Sometimes unexpected things might hide in there. We searched the Hubble archives for asteroids, but also found artificial satellites and strong gravitational lenses.

Data archives are important

Hubble provides a rich data archive spanning decades. Ideal to survey faint asteroids and strong gravitational lenses.

One astronomer's trash can be another one's treasure! .