A satellite in space, likely the Hubble Space Telescope, is shown in the foreground. The Earth's horizon is visible in the background, with a bright sun or star in the upper left corner. The satellite has a large circular lens and various instruments.

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

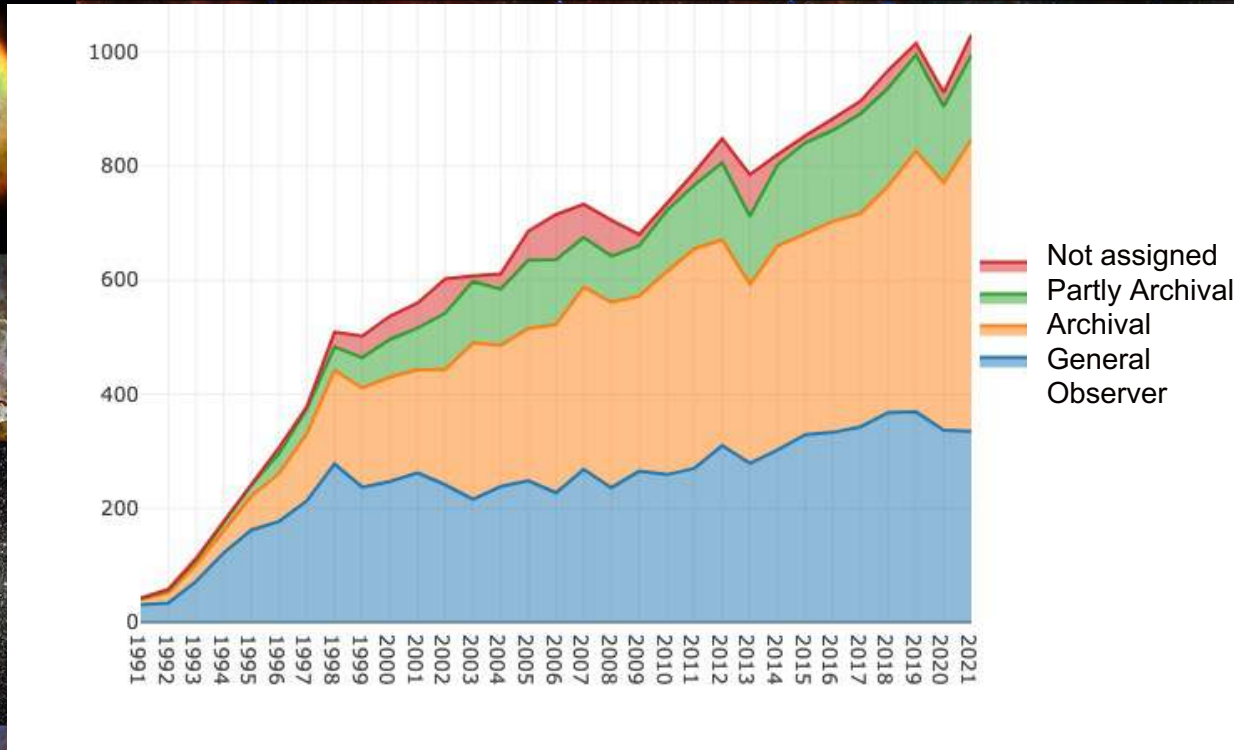
ESA PSIDA2022, 22 June 2022

Collaborators: Pablo García Martín¹, Marcel Popescu², Bruno Merín³, Max Mahlke⁴, Benoît Carry⁴, Ross Thomson⁵, Samet Karadag⁵, Javier Durán³, Deborah Baines³, Elena Racero³, Fabrizio Giordano³, Guido de Marchi³, René Laureijs³

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

Growing number of publications using archival data

Hubble Space Telescope publications by observation type



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

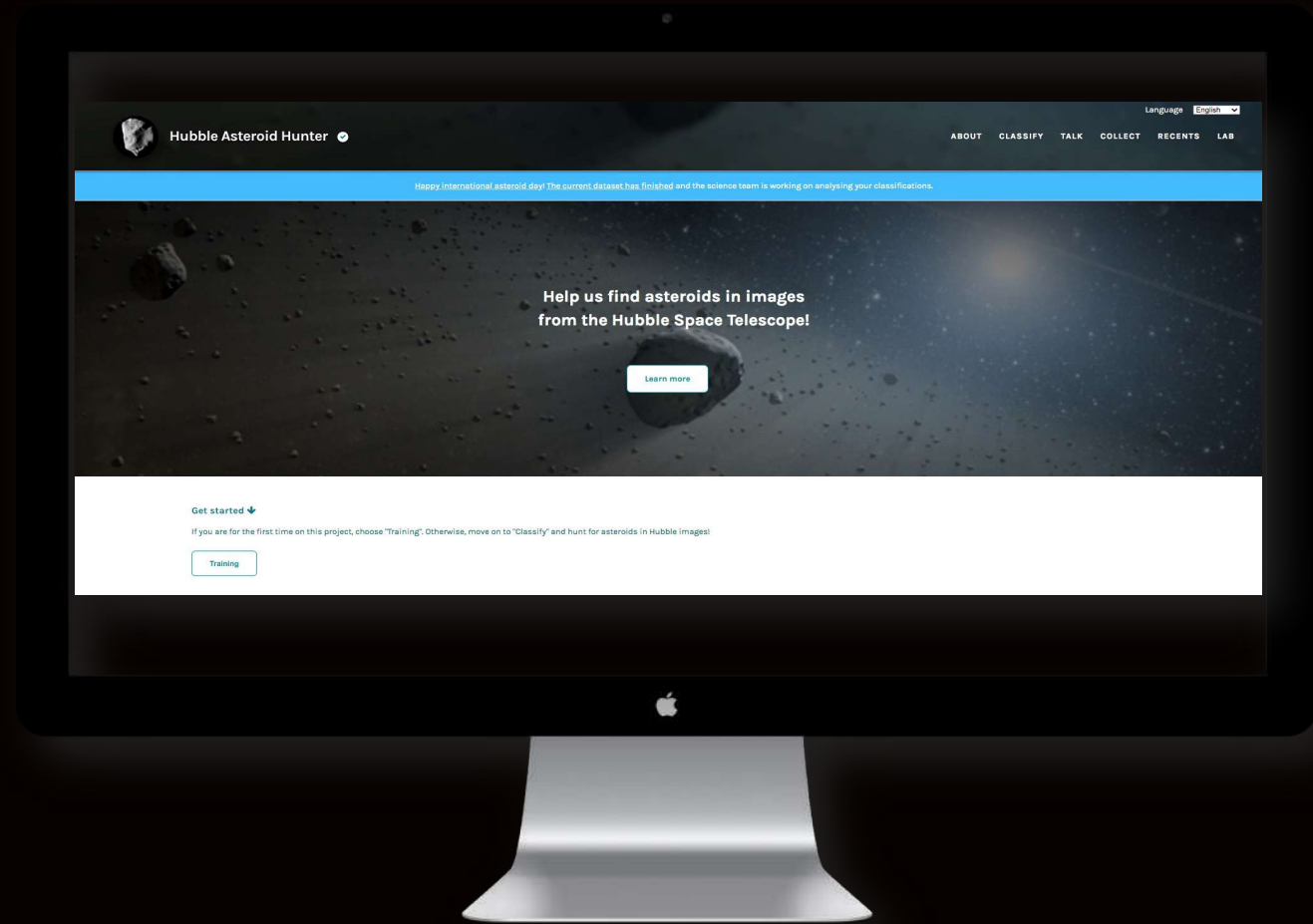
AI and crowdsourcing for scientific exploration of archives



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



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





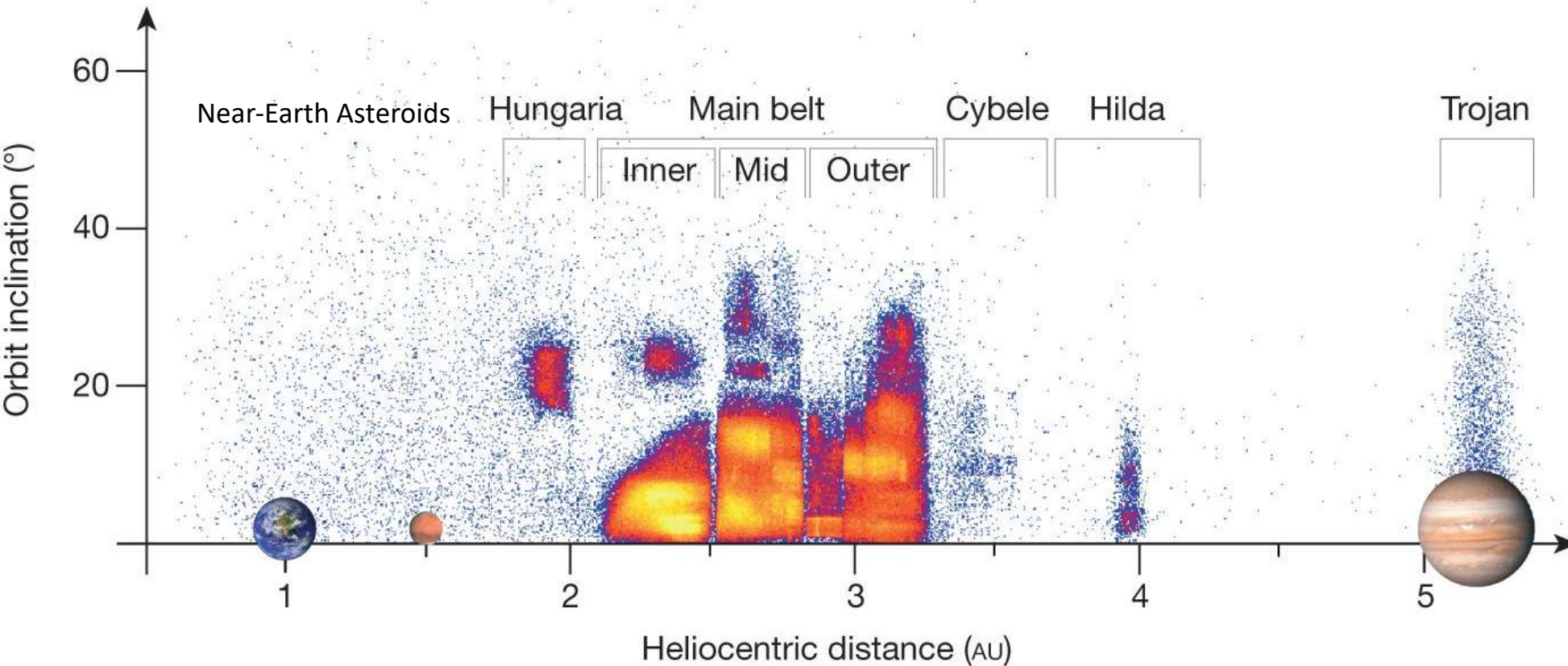
Identifying asteroid trails in the Hubble Space Telescope images





Asteroids across the Solar System

Distribution of asteroid classes in the Solar System



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

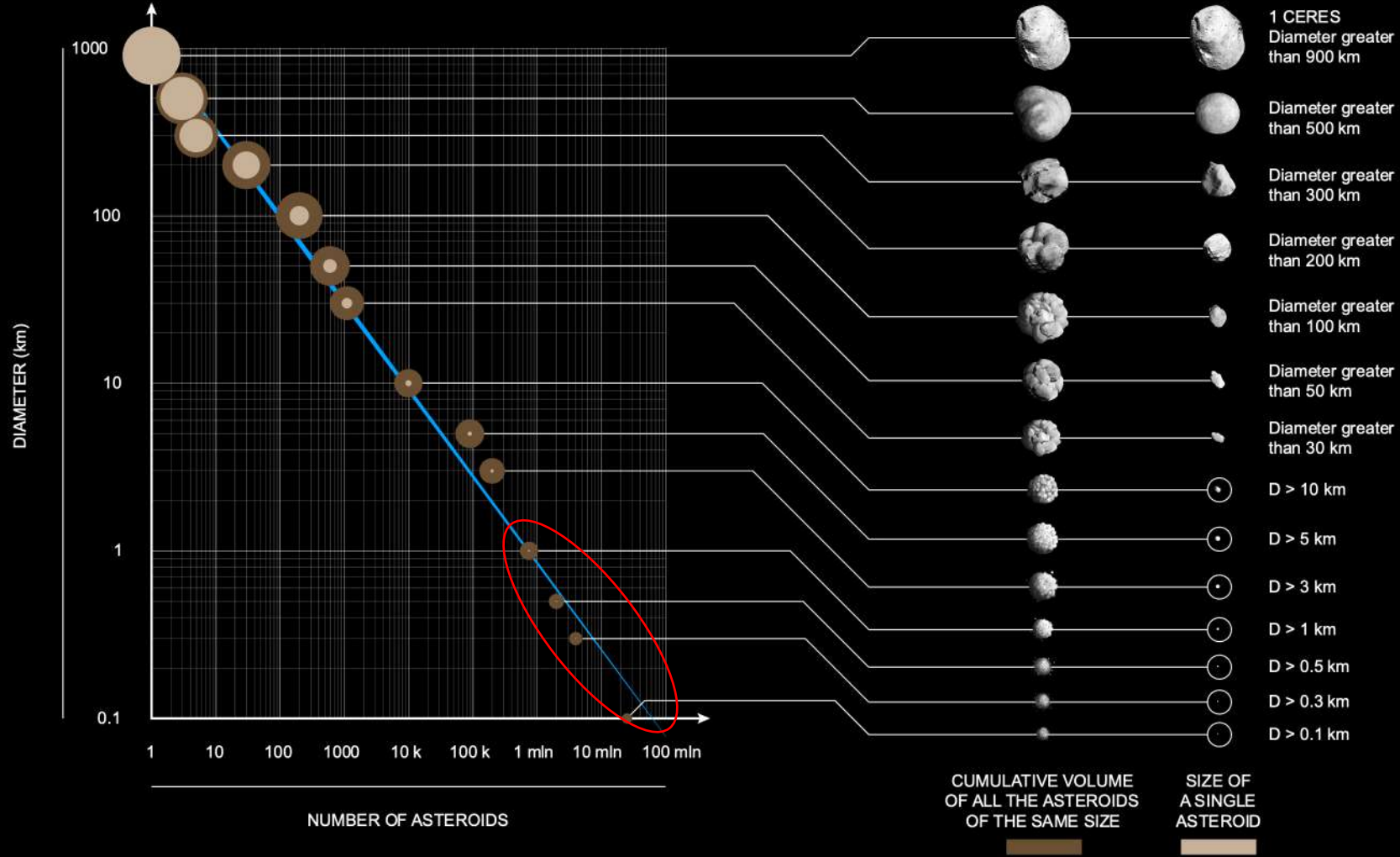
THIS MONTH:	1
THIS YEAR:	20
ALL TIME:	4407

DeMeo and Carry, 2014

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

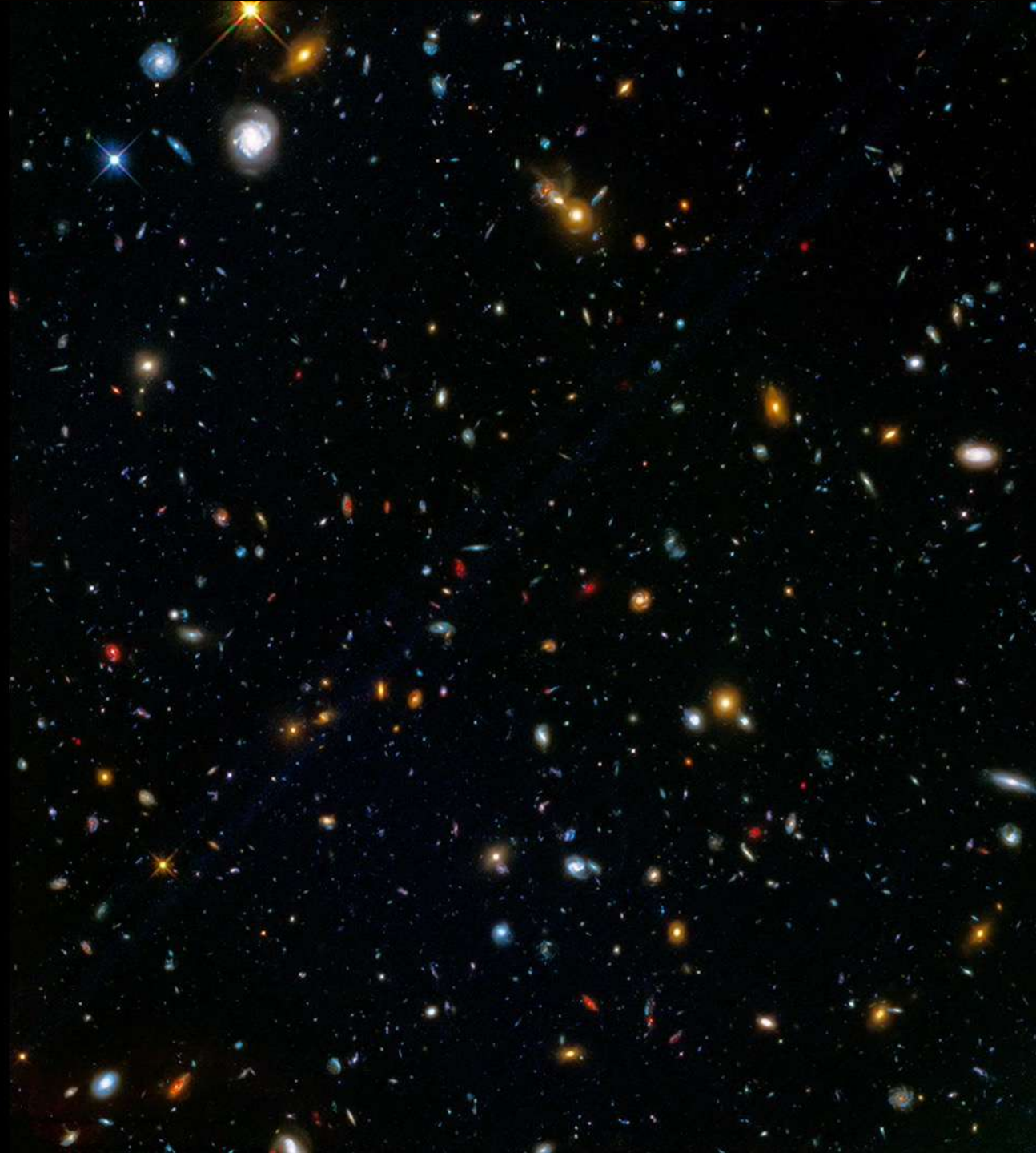


Distribution of asteroid sizes in the Solar System



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



[Happy international asteroid day!](#) [The current dataset has finished](#) and the science team is working on analysing your classifications.

Help us find asteroids in images
from the Hubble Space Telescope!

[Learn more](#)

Get started

If you are for the first time on this project, choose "Training". Otherwise, move on to "Classify" and hunt for asteroids in Hubble images!

[Training](#)



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



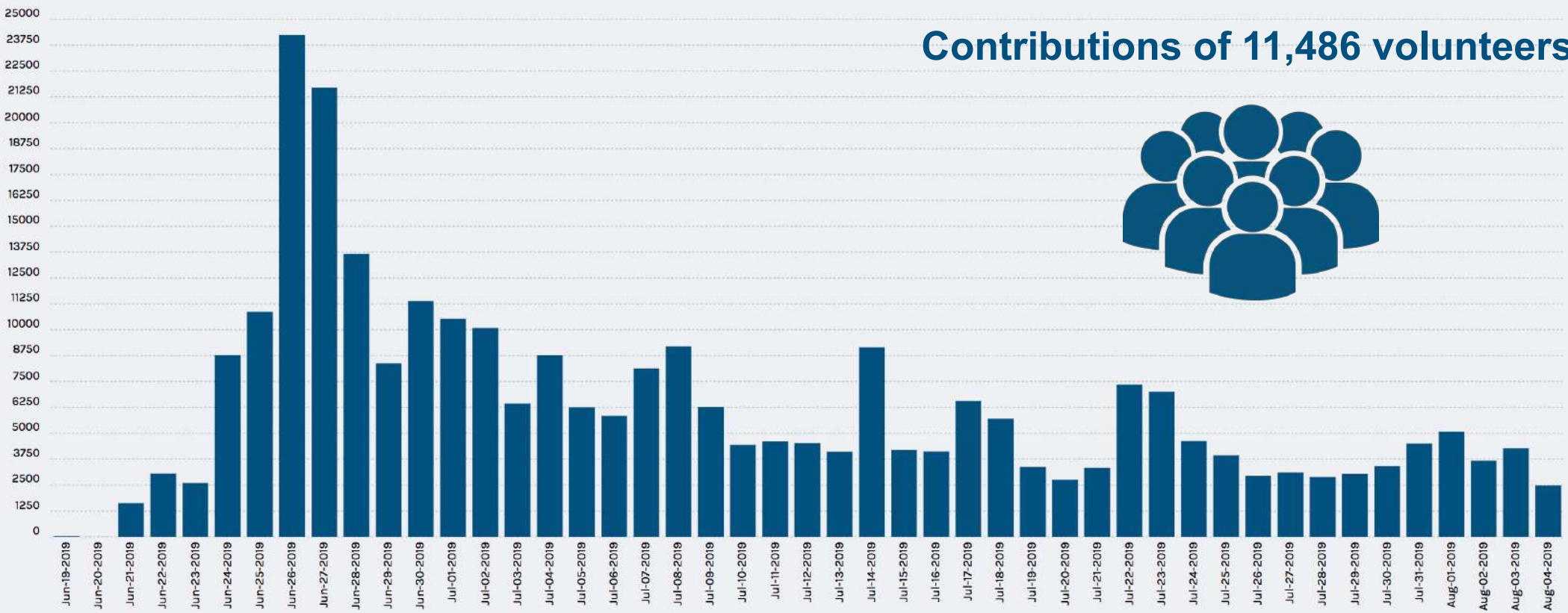
Volunteer participation in the citizen science project

Classification Stats

Classifications per for

Current date range: Jun-19-2019 to Aug-04-2019

[Reset date range](#)



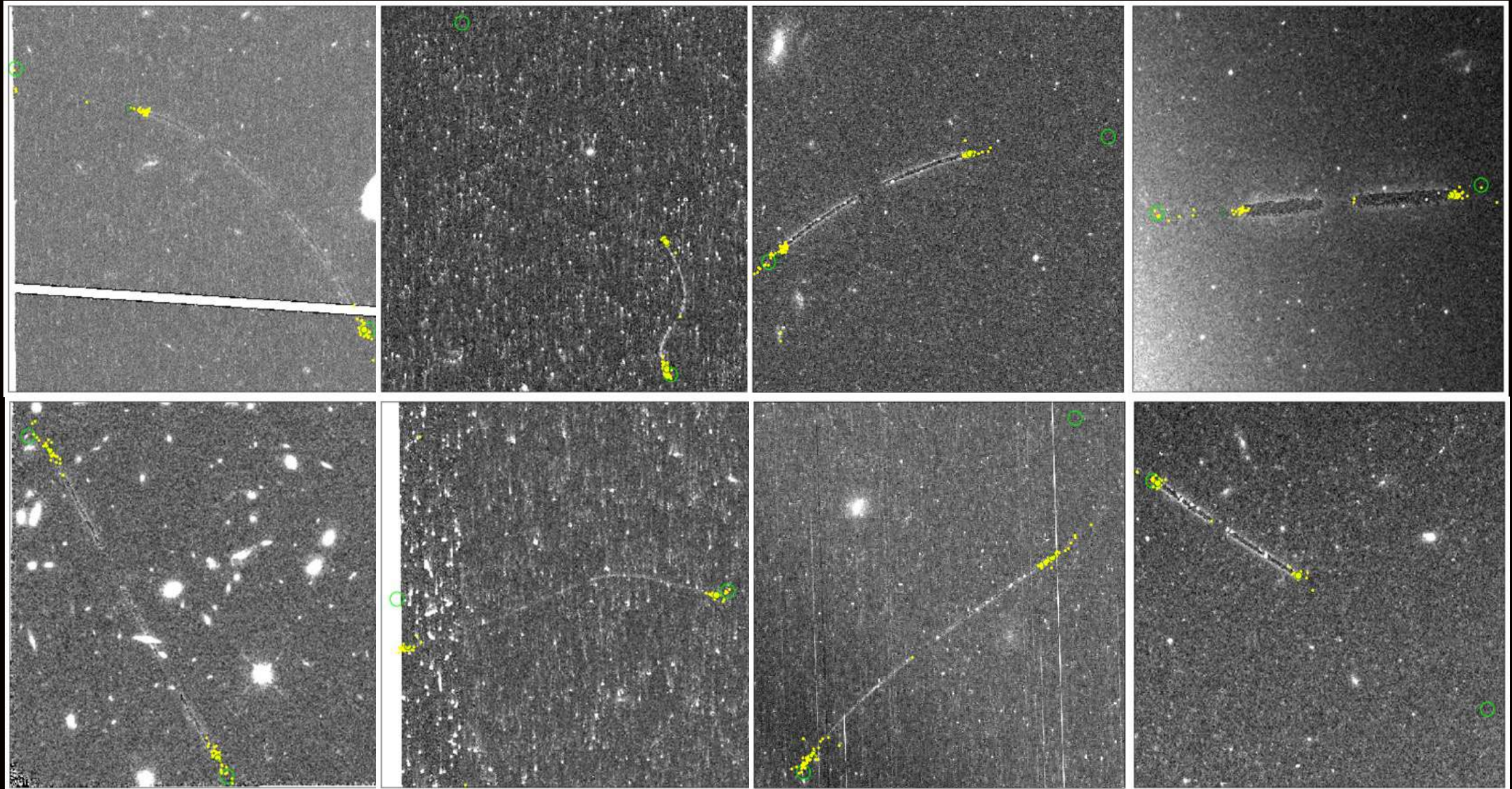
Contributions of 11,486 volunteers





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

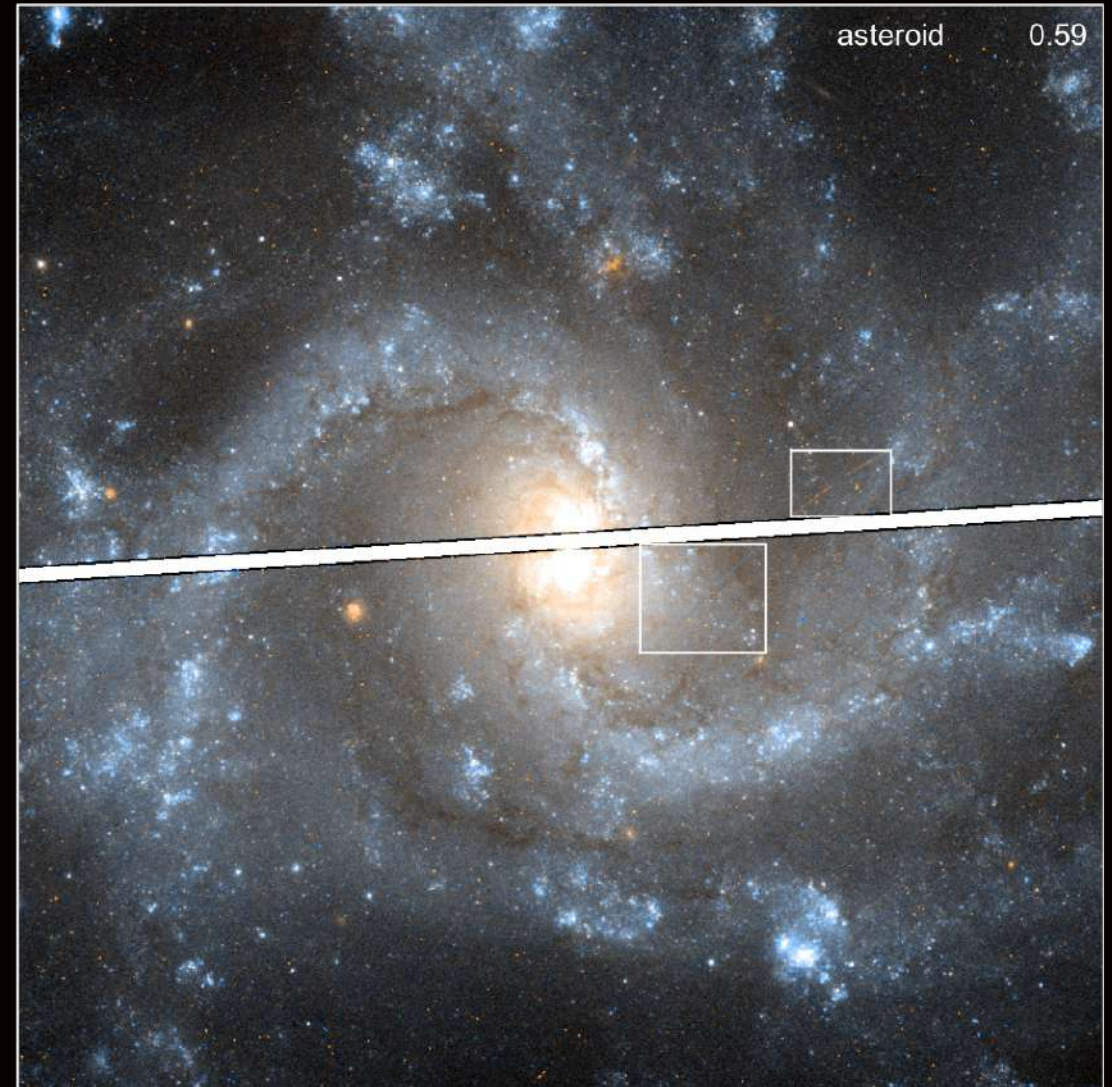




Scaling up with AutoML object detection to identify trails



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.

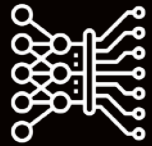


Scaling up with AutoML object detection to identify trails

* In collaboration with 

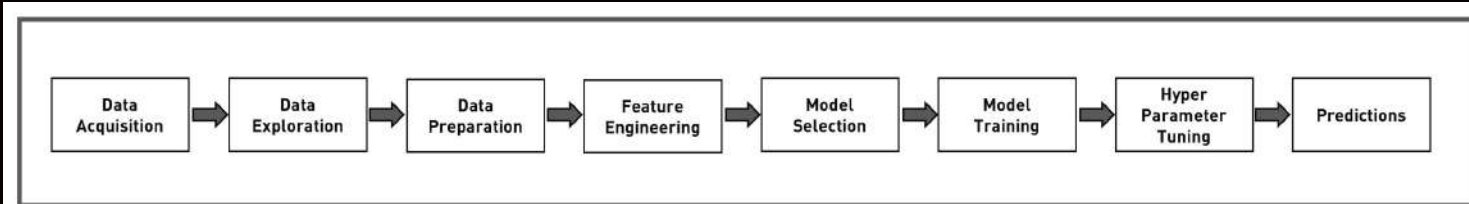


Classifying the entire HST archive (2002 – 2021) of 37,324 images (x4 = 150k cutouts)

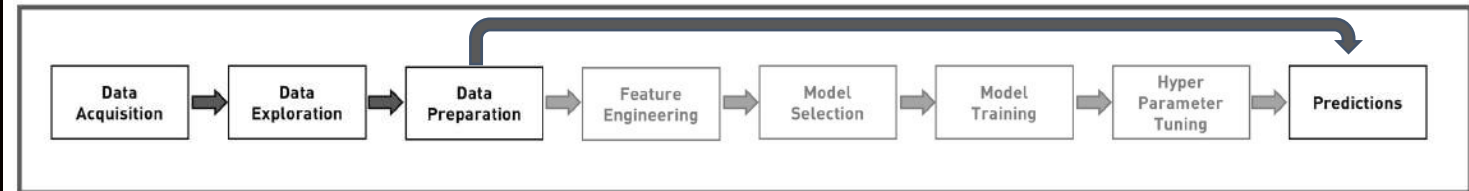


Use automated machine learning (AutoML) on Google Cloud

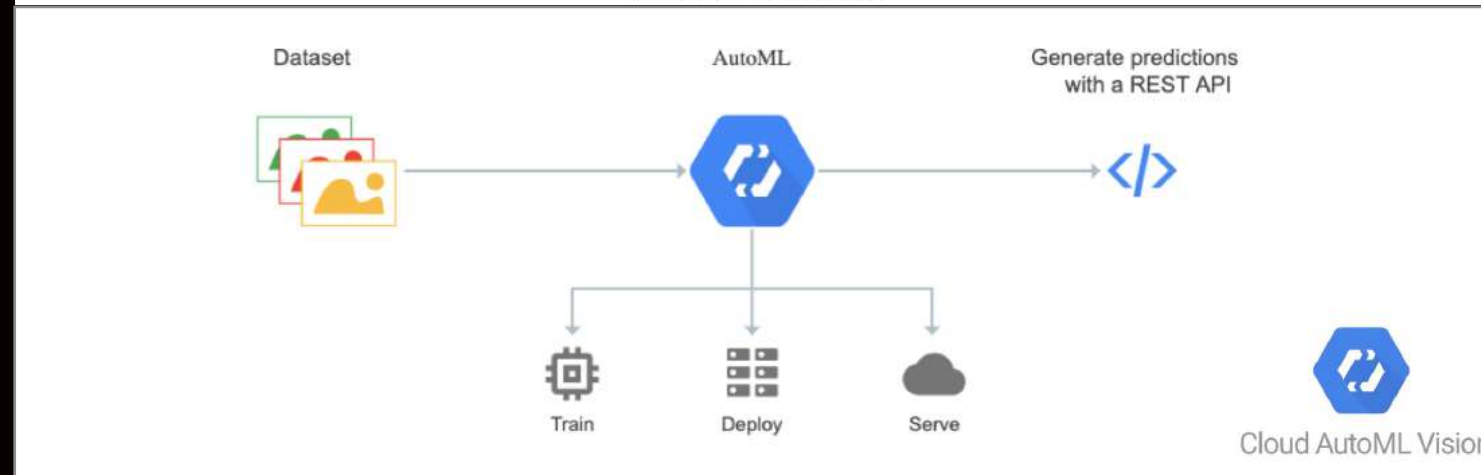
Automating ML pipelines



Traditional Machine Learning Workflow



AutoML Workflow



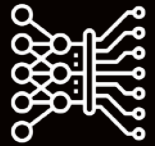
How AutoML works – using Neural Architecture Search



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

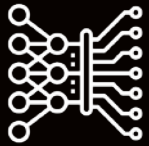
The screenshot shows the Hubble Science Archive website. At the top, there are logos for the European Space Agency, Science & Technology, and ESAC Science Data Centre. The main header reads "hubble science archive" and "esa". Below this is a navigation menu with "RESULTS #1 x" and tabs for "hst observations (6)", "hia observations (11)", and "proposals (1)". A table lists observation IDs such as j9q716010, j9q716020, and j9q716030. A "POSTCARD PREVIEW for j9q715030_PREV" window is open, showing a large image of a galaxy with a red box highlighting a region. A zoomed-in view of this region shows a bright, elongated object labeled "asteroid" with a confidence score of 0.62.



Scaling up with AutoML object detection to identify trails



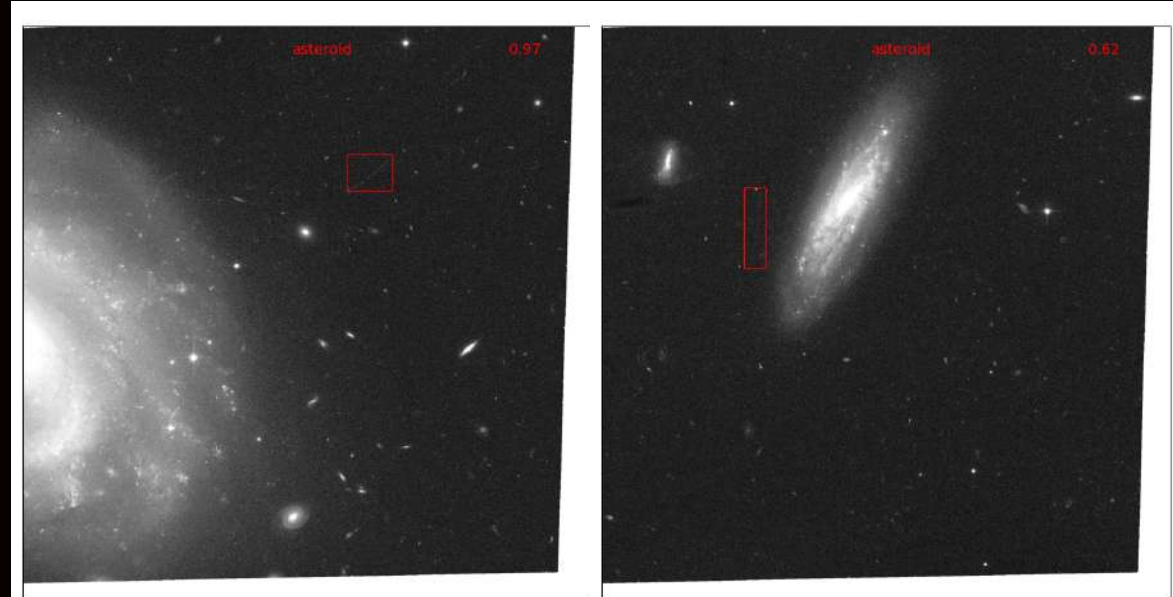
Classifying the entire HST archive (2002 – 2021) of 37,324 images (x4 = 150k cutouts)



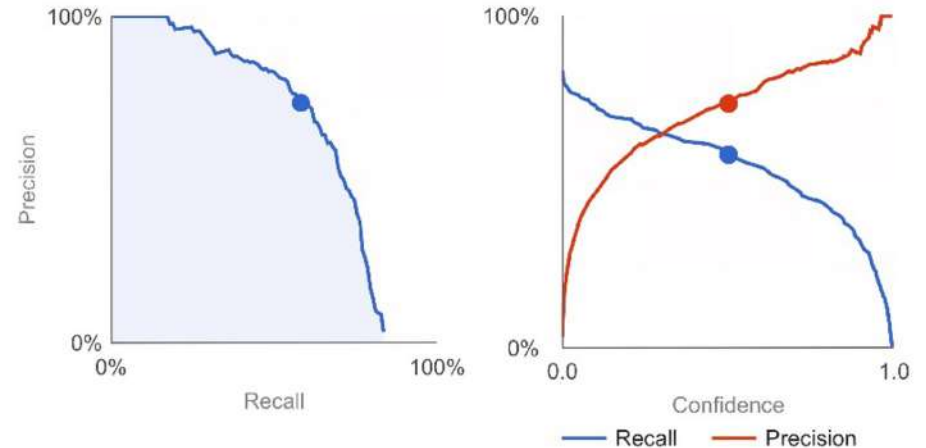
Use automated machine learning (AutoML) on Google Cloud



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





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



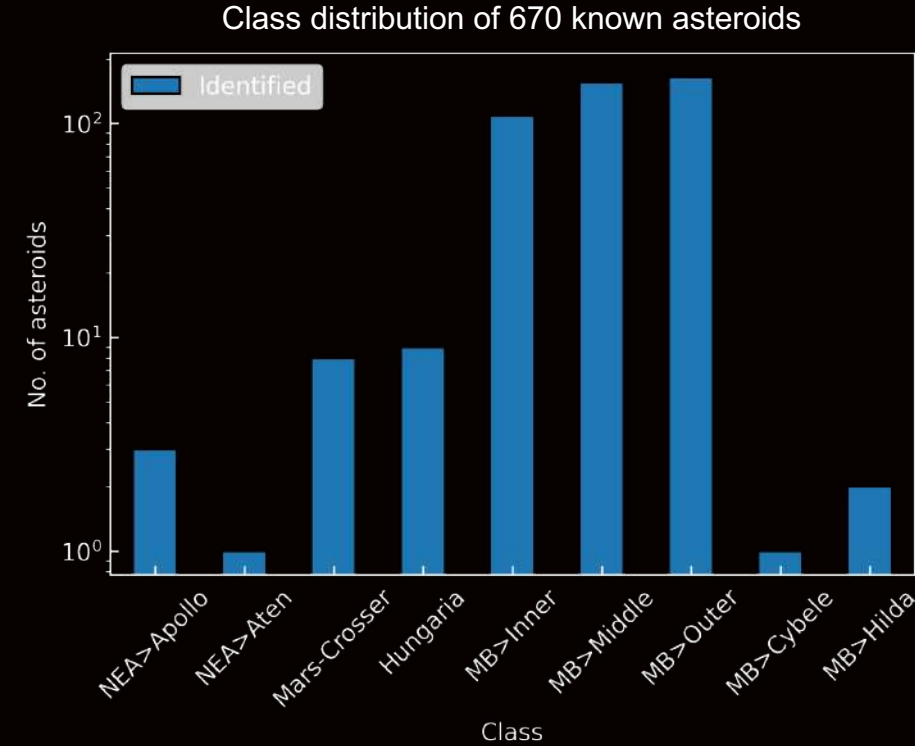
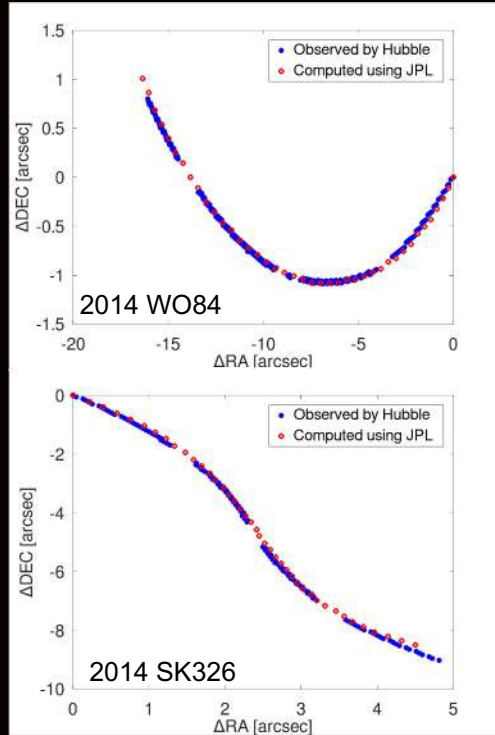
2487 asteroid trails recovered by citizen scientists and by AutoML



1701 asteroids validated by the team



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





Results: magnitude distribution of detected asteroids



2487 asteroid trails recovered by citizen scientists and by AutoML



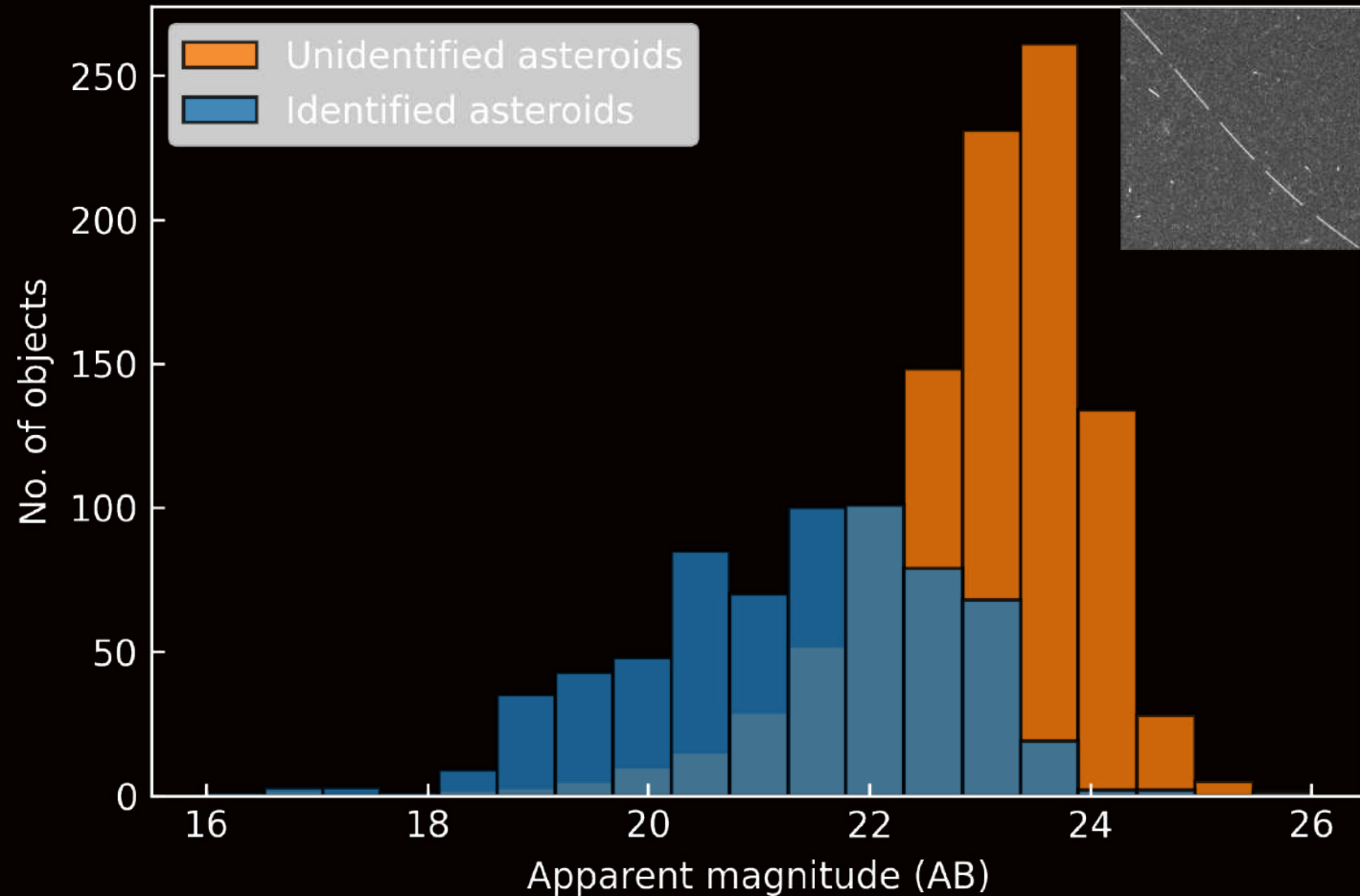
1701 asteroids validated by the team



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



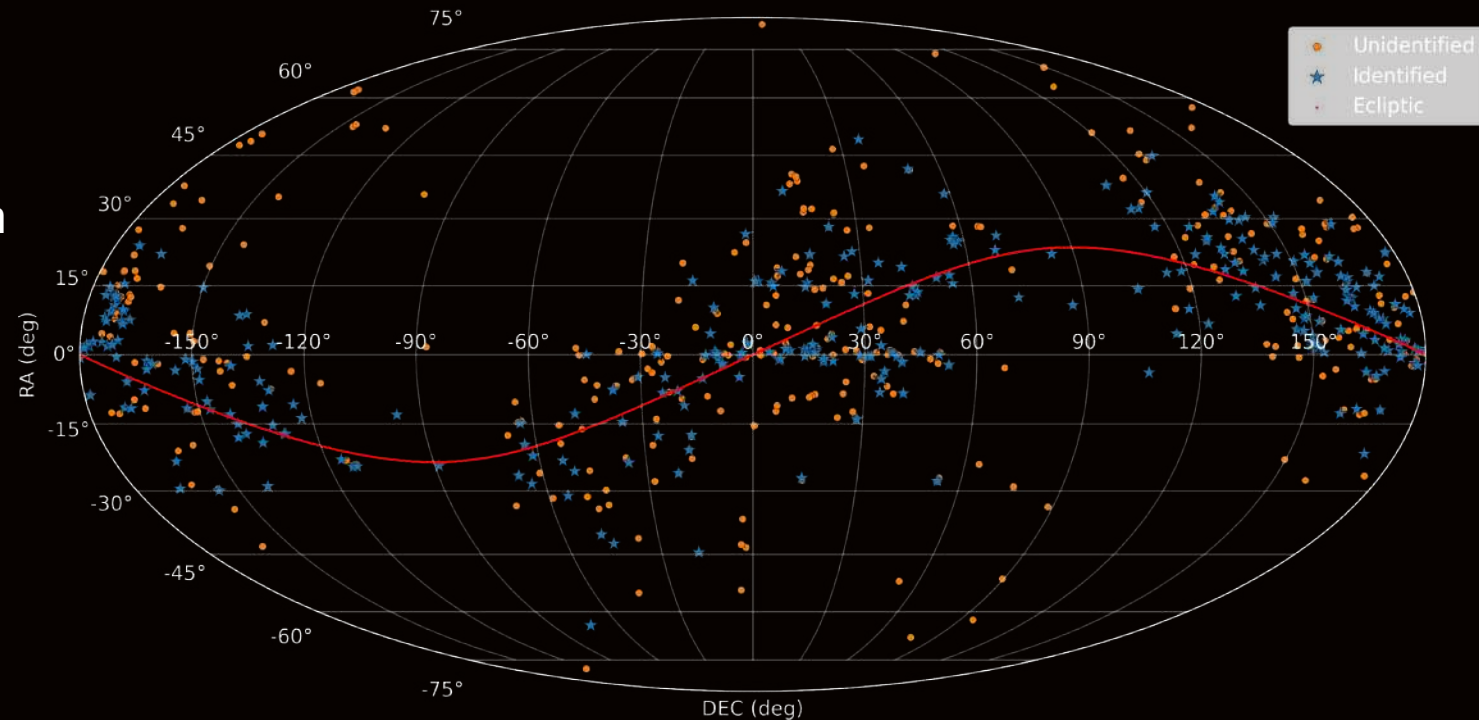
1701 asteroids validated by the team



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

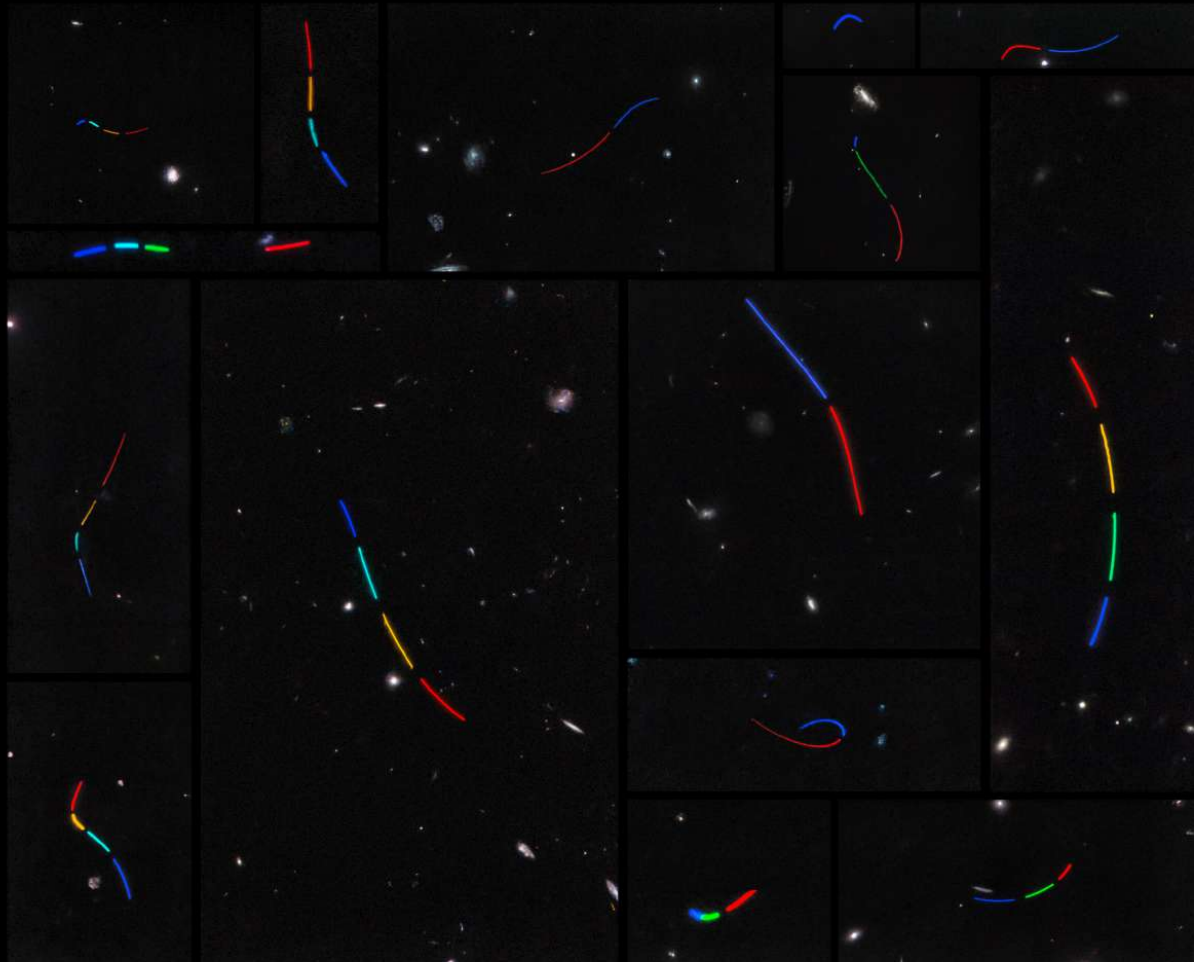


Likely previously unknown small size Main-Belt asteroids

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



See Pablo's poster



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



Hubble Asteroid Hunter: Analysing asteroid trails in HST images

Pablo García Martín¹, Sandor Kruk², Marcel Popescu³, Bruno Merin⁴, Ross Thomson⁵, Robin W. Evans⁶, Karl R. Stapelfeldt⁷

- 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 Cuizil de Argint, 840357 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 15208, United States
- 6 - Bantlin 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.

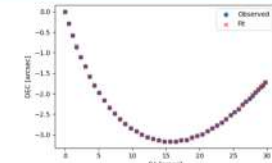


Figure 2: Real (red dots) and fit (red) for its best parallax-distance solution (red)

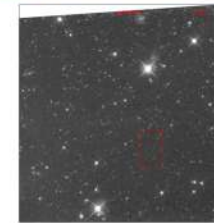


Figure 1: Examples of asteroid trails identified by Hubble Asteroid Hunter project

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 3: HST sees asteroids as trails (fixed target pointing)



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.

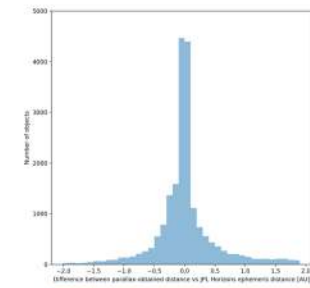


Figure 4: Difference between parallax-calculated distance vs JPL ephemeris distance for 21,280 known objects

Do real science!

Learn about galaxies while contributing to research

You can help us!

There are hundreds of billions of galaxies scattered throughout our Universe, each with its own story to tell. How do we get to know them? Astronomers take pictures of the sky with high-powered telescopes like the Dark Energy Camera Legacy Survey (DECaLS) in Chile.

It takes a lot of people to study these galaxies—far too many for one small team.

People just like you

can help us learn about the universe.



How the research process works

Other people look at a picture and tell us what you see.

They help us figure out what's in the picture.

They help us learn about the past, present, and future of our Universe.

What's in the picture?

Help scientists sort galaxy images on the web. You'll be able to identify galaxies like the Smooth Galaxy, Galaxy with Features, and Ring Galaxy.

SMOOTH GALAXY



Smooth galaxies are common and are usually round or oval. They might be a bright spot at the center. Smooth galaxies are older and probably don't have stars anymore.

GALAXY WITH FEATURES



Galaxies with features might be spiral arms, bulges, or irregular shapes. They might be younger and have more stars. They might be in the process of merging with another galaxy.

HOW TO HELP?



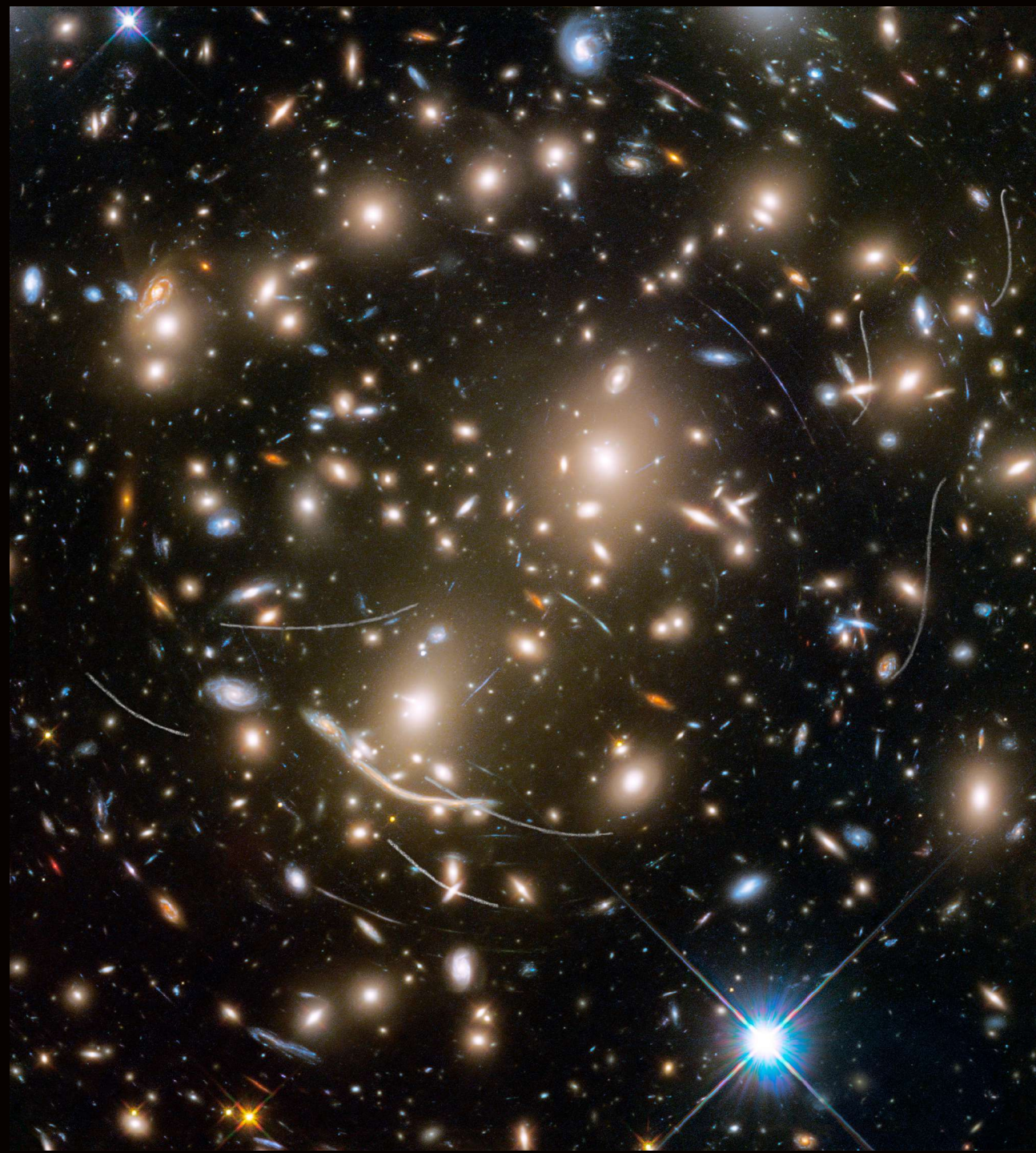
ON THE INTERNET

Visit www.zooniverse.org to help with the project and many more to learn how to help with other research, history, biology, and more.

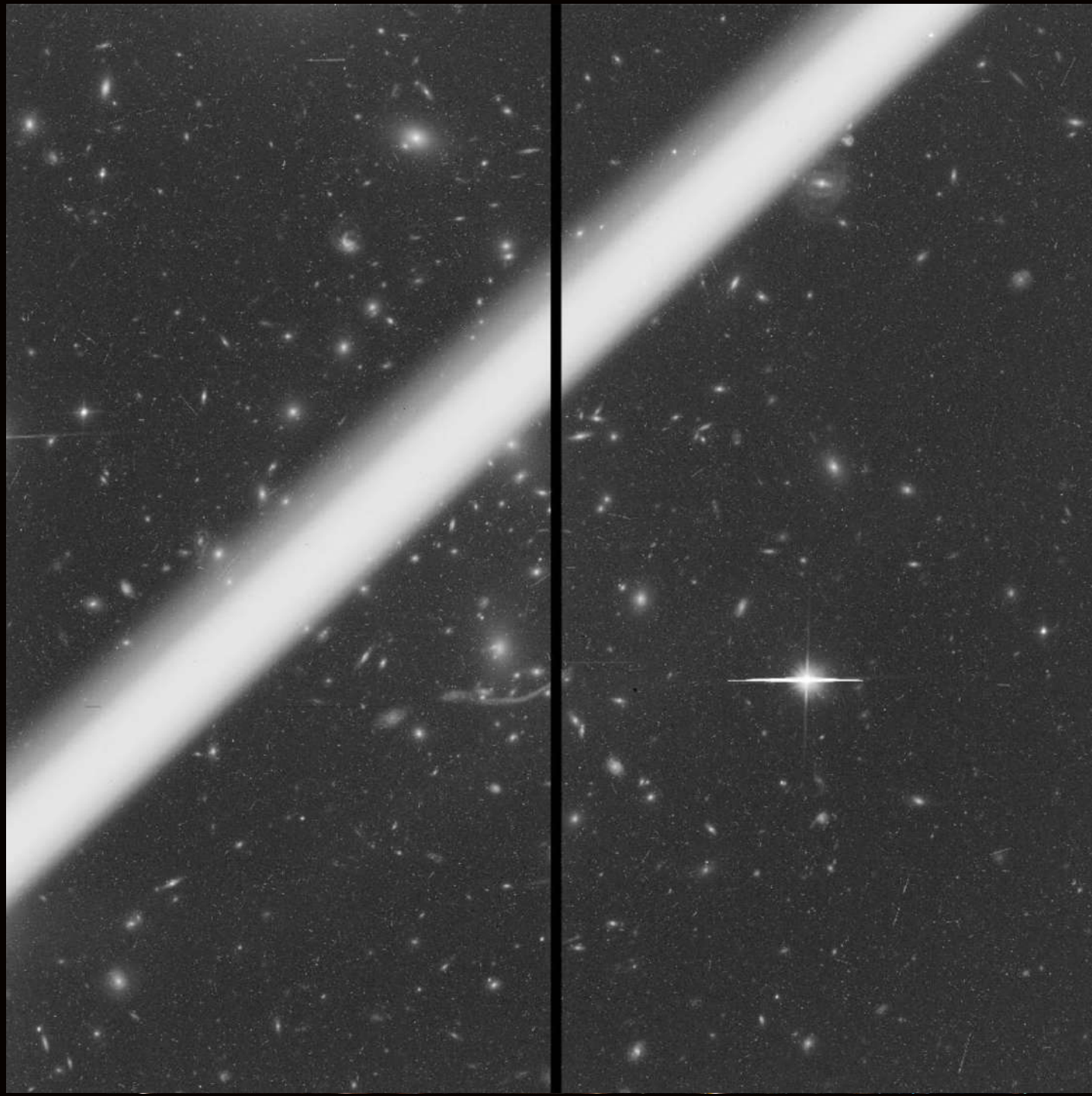
ON YOUR SMARTPHONE

Get the Zooniverse app for iOS or Android.

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.

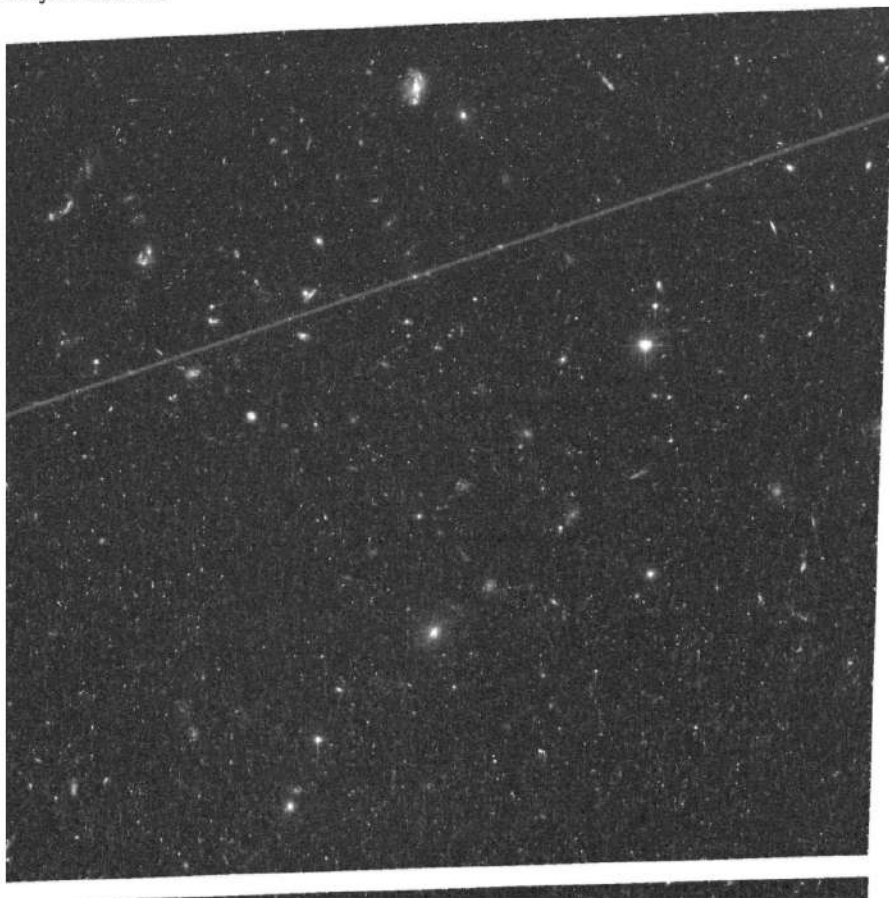
Satellite ID by J. McDowell
Image credit: Judy Schmidt



Hubble Asteroid Hunter Talk

Search or enter a #tag

Subject 37652046



Comments:



Fauxwise
@Fauxwise

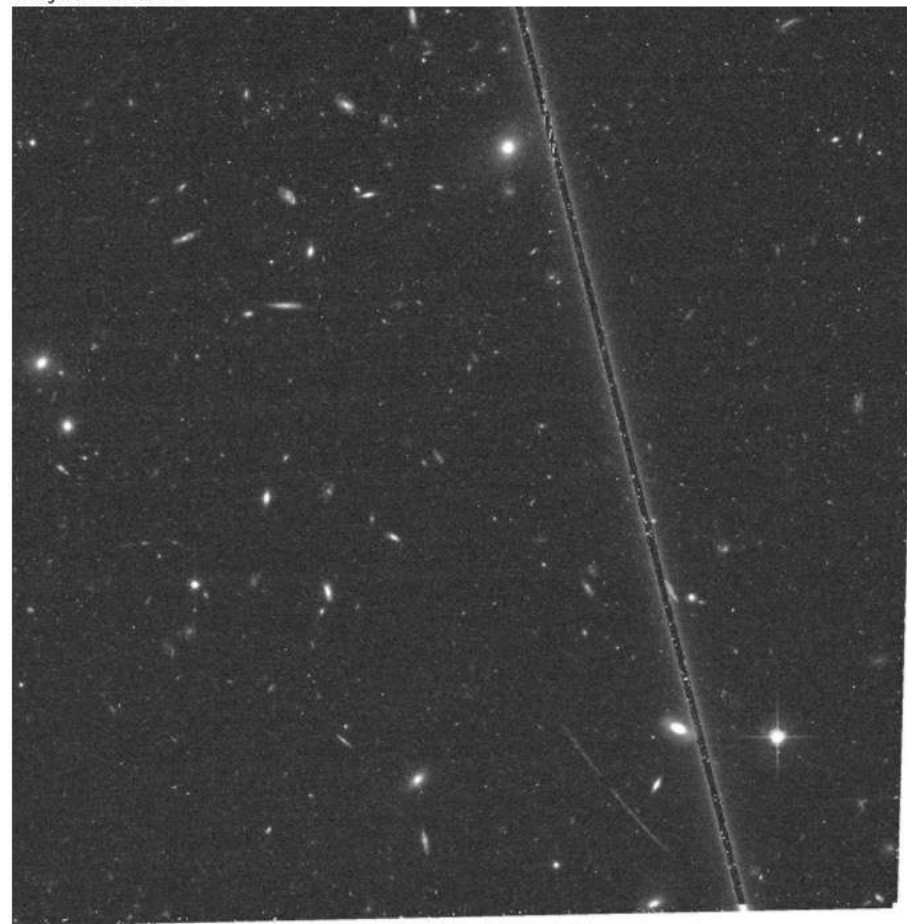
December 18th 2019, 11:51 am

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

Hubble Asteroid Hunter Talk

Search or enter a #tag

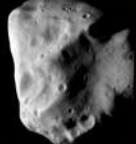
Subject 42950351



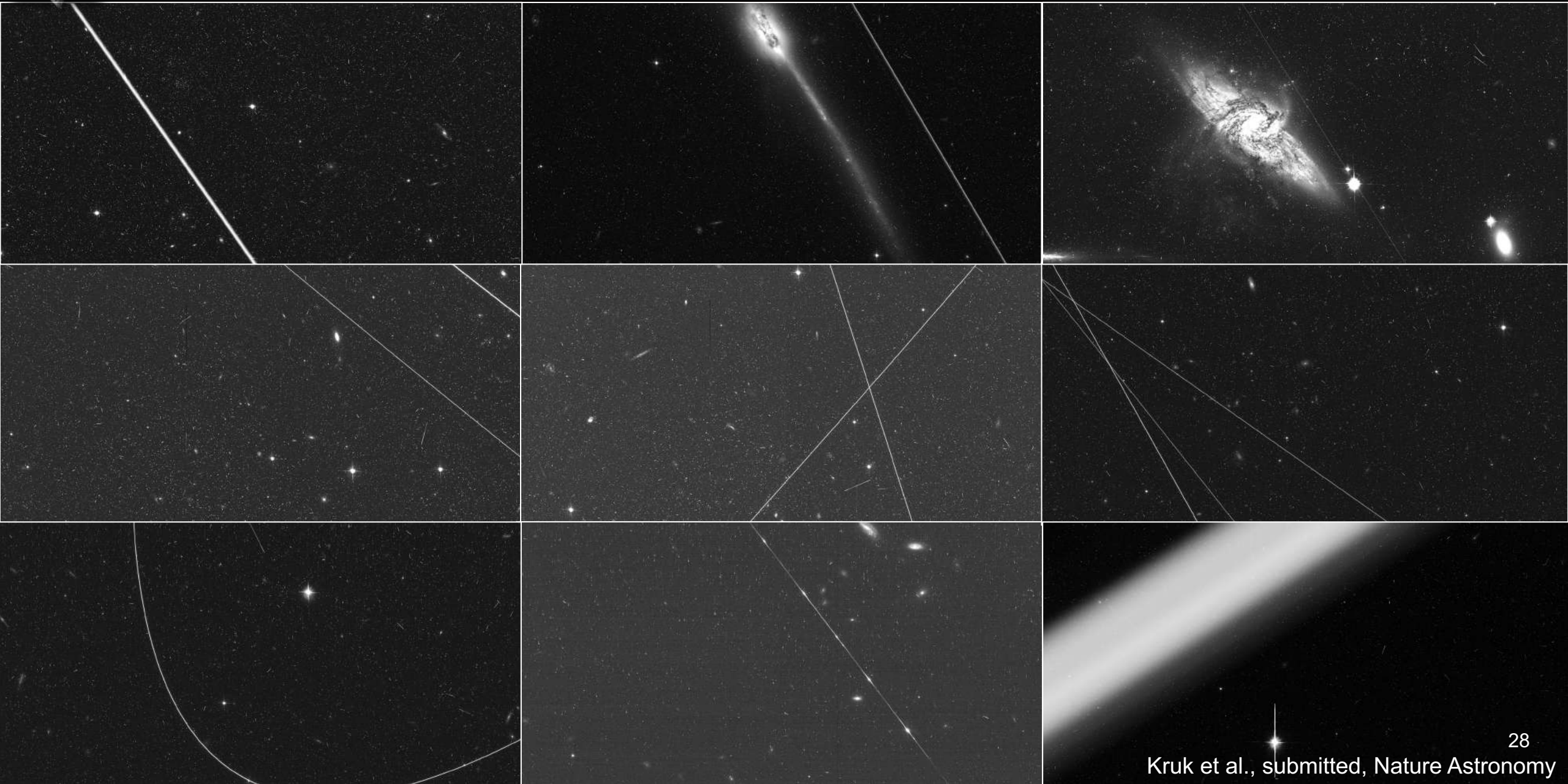
Travagline
@Travagline

April 28th 2020, 10:01 pm

[#asteroid](#) and [#satellite](#)



Satellite trails identified with AutoML in Hubble exposures





Serendipitous findings: new strong gravitational lenses

Happy international asteroid day! The current dataset has finished and the science team is working on analysing your classifications.

Project forum "Talk"

Hubble Asteroid Hunter Talk

[Hubble Asteroid Hunter Talk](#) > [Notes](#) > Subject 39619051

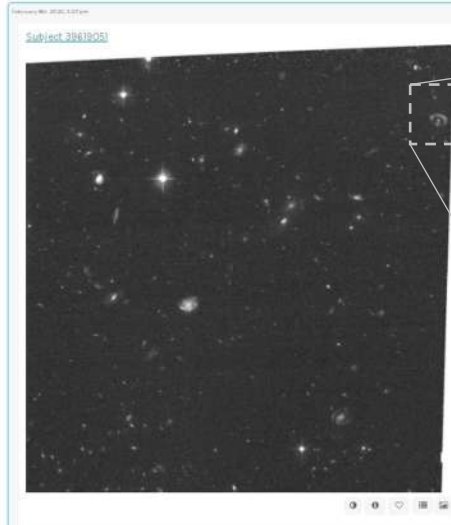
Search or enter a #tag

Subject 39619051

Unsubscribe

You're receiving notifications from this discussion because you've joined it (daily email)

Page 1 of 1



#gravitational_lens #arc (s) and counter images
multi components (or clumps) source lensed
WISEA J034445.19-642133.8
ra, dec 56.1885 -64.35946



main arc radius ~1.90"
J034445.2-642133.8
multi components (clumps) source lensed

#gravitational_lens #arc (s) and counter images
multi components (or clumps) source lensed
WISEA J034445.19-642133.8
ra, dec 56.1885 -64.35946



f606w

main arc radius ~1.90"
J034445.2-642133.8
multi components (clumps) source lensed

Popular Tags:

- [satellite](#)
- [gravitational_lens](#)
- [asteroid](#)
- [cluster_lens](#)
- [dwarf](#)
- [windows_ghost](#)
- [asteroid_trail](#)
- [trail](#)
- [ring](#)
- [odd](#)
- [artifact](#)
- [arc](#)
- [asteroid-trail](#)
- [galaxy](#)
- [nebula](#)
- [dragon_breath](#)
- [satellite-trail](#)
- [gravitational_lens](#)
- [conspiracy](#)
- [lens](#)

1 Active Participants:

[Sandor Kruje](#)

Projects:

- [Zooniverse Talk](#)
- [Next Quest Go: Killdeer](#)
- [Battling Birds: Panama Edition](#)
- [Dix Weather - WW2](#)
- [Zwicky Chemical Factory](#)
- [Bursts from Space](#)
- [Beyond Borders: Transcribing Historic](#)
- [Barnes Land Documents](#)
- [Building Detective For Disaster](#)
- [Crossroads](#)
- [Next Quest Go: Thrushes](#)
- [Next Quest Go: European Sparrows and](#)
- [House Sparrows](#)
- [Citizen ASAS-15](#)

Load more



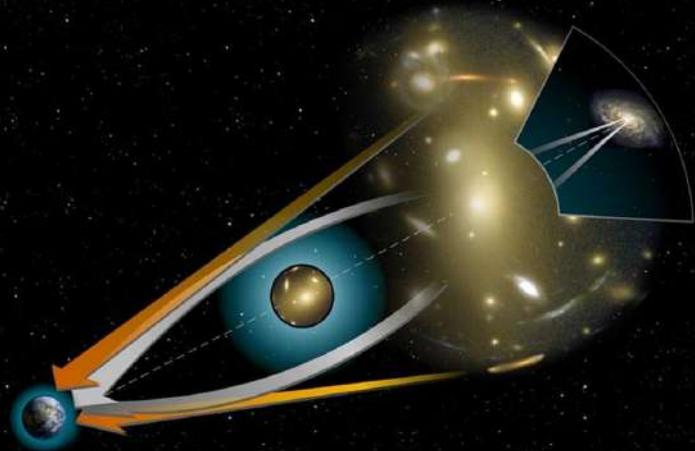
Serendipitous findings: new strong gravitational lenses



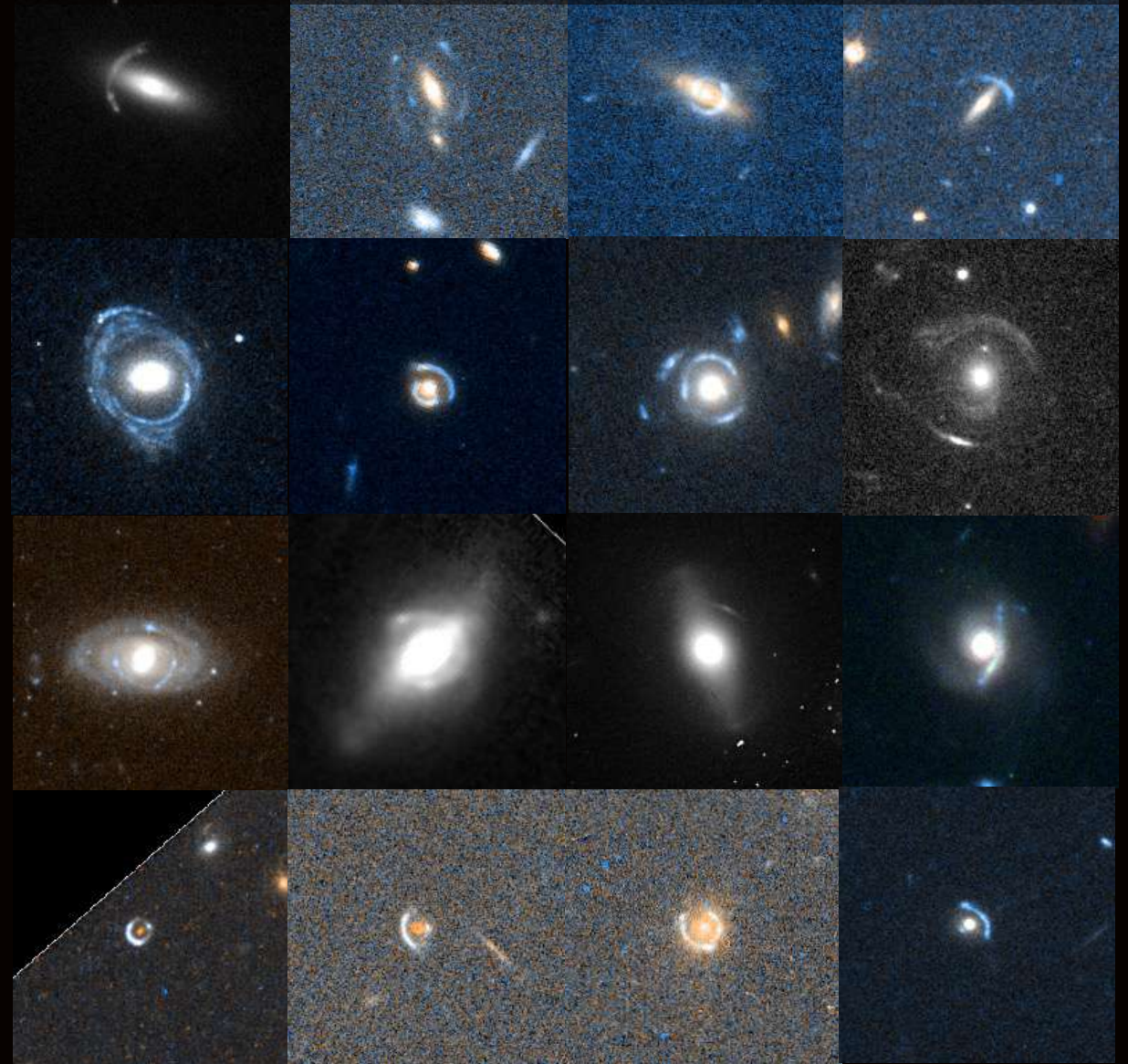
Having human eyes on the data can lead to new discoveries:



Discovery of 198 new strong gravitational lenses



New strong gravitational lenses identified in Hubble Space Telescope images



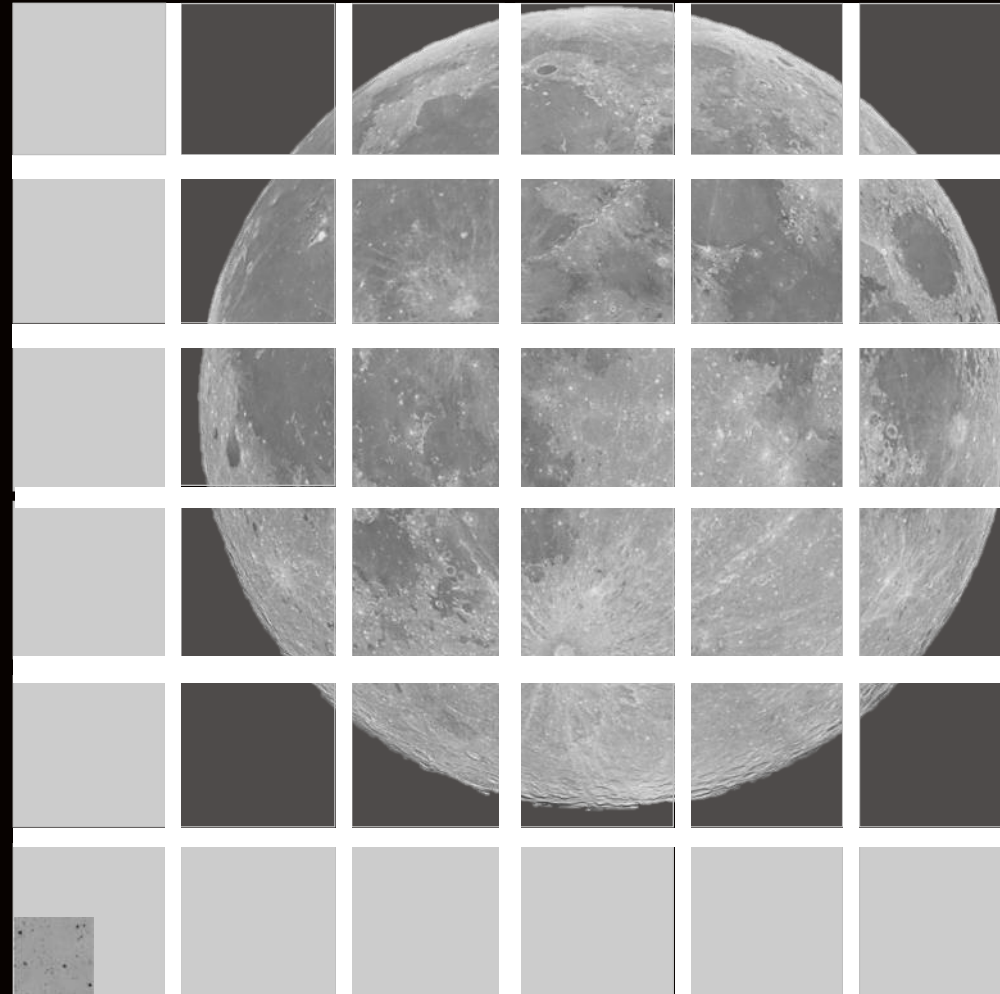
The image shows a detailed 3D rendering of the ESA Euclid mission spacecraft. The satellite is oriented diagonally, showing its large, rectangular solar panel array on the left side, which is covered in a grid of blue solar cells. The main body of the satellite is cylindrical and wrapped in a crinkled, silver thermal insulation material. A large, circular, white structure, likely a sunshield or part of the instrument housing, is visible on the right side. The background is a deep space scene with a gradient from dark blue to purple, filled with numerous stars of varying colors and sizes, some with prominent diffraction spikes. The overall composition is centered, with the spacecraft as the primary focus.

Future: ESA Euclid Mission

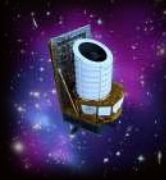


Euclid: providing a high-definition view of 1/3 of the sky

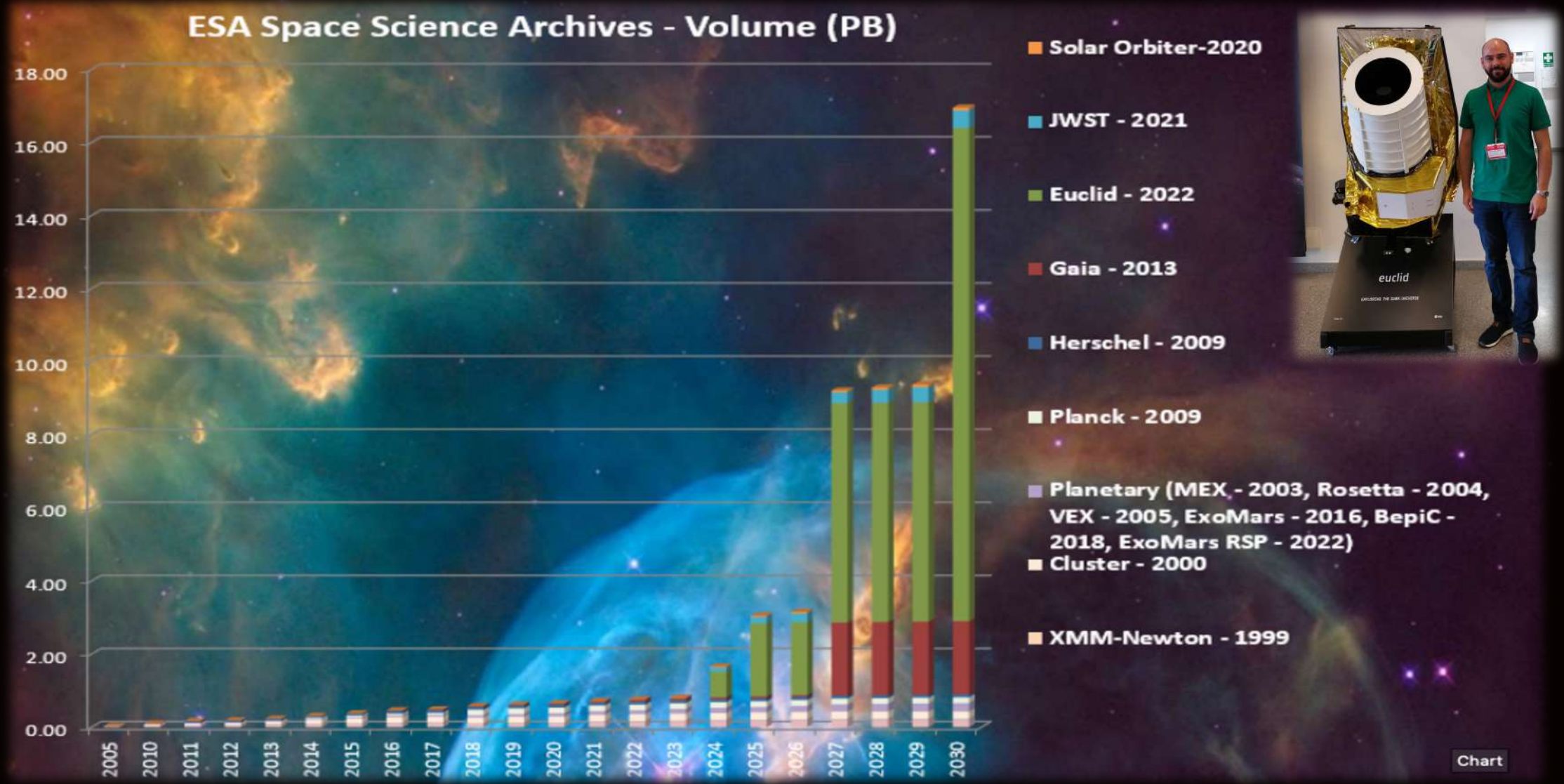
Single Euclid exposure (1/60,000th of the survey)

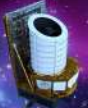


Single Hubble exposure

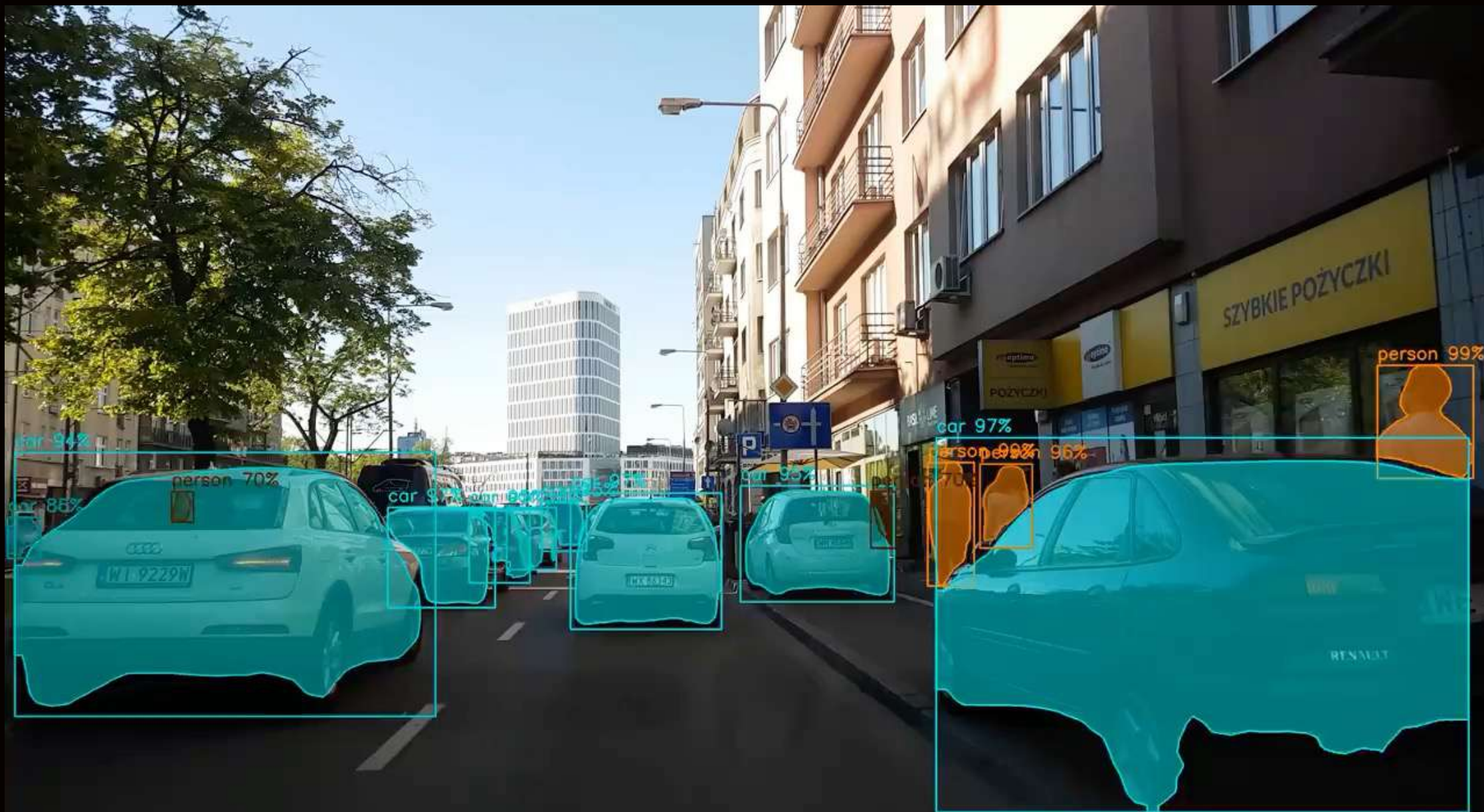


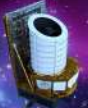
Euclid will add 10+ PB of data in the next years



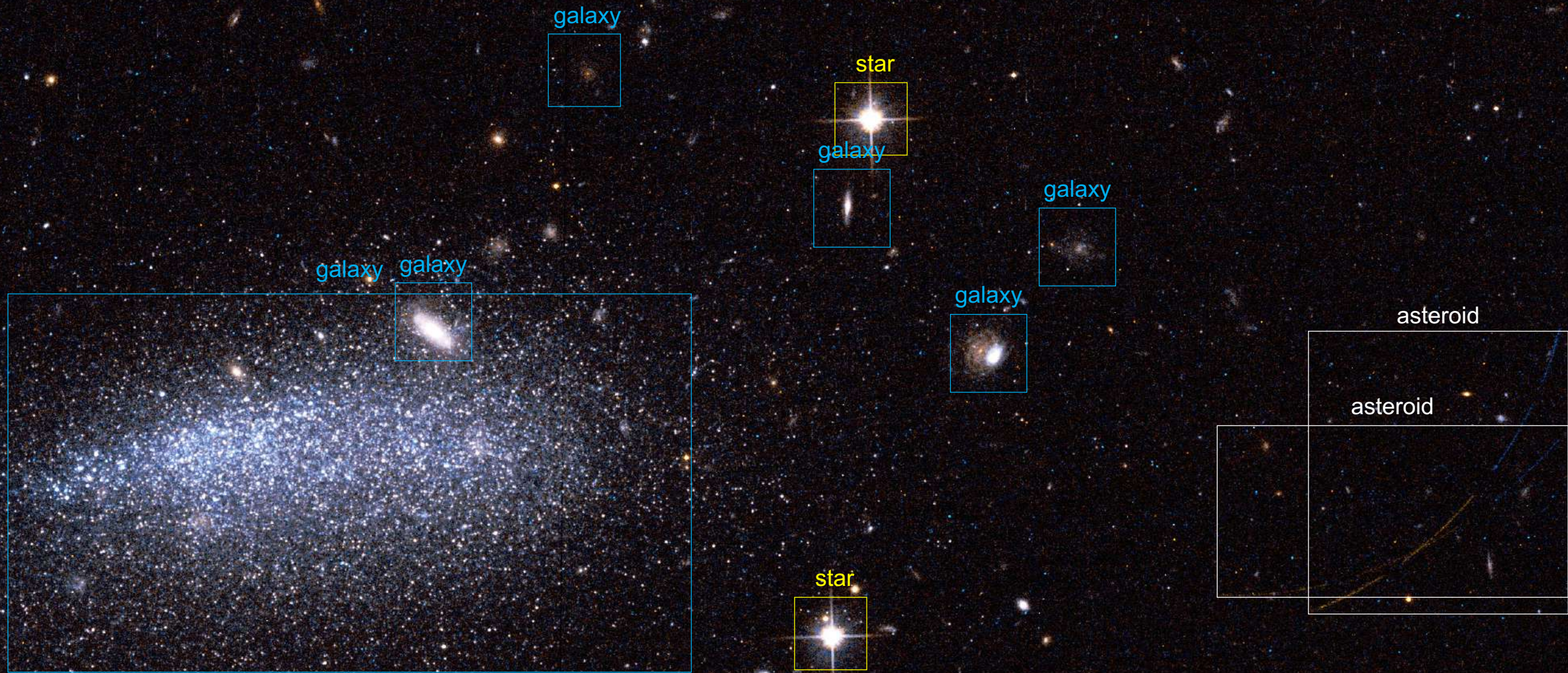


My vision: ML for automated object detection in astronomy





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



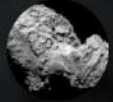
Hubble provides a rich data archive spanning decades. Ideal to survey faint asteroids and strong gravitational lenses.

➔ 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



Rosetta Zoo ✓

Language **English** ▾

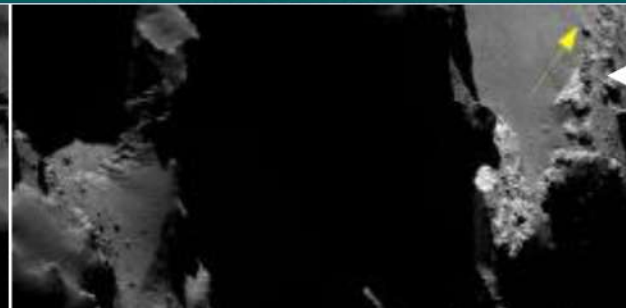
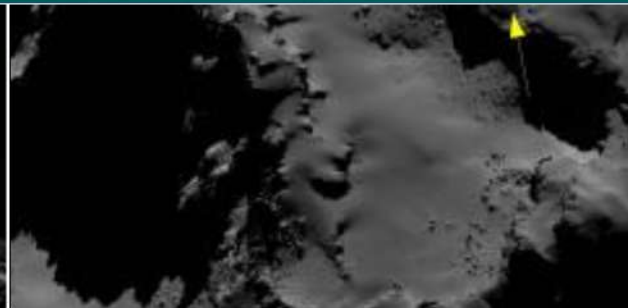
[ABOUT](#) [CLASSIFY](#) [TALK](#) [COLLECT](#) [RECENTS](#) [LAB](#)



Find changes on Comet 67P and help us better understand the history of the Solar System

[Learn more](#)

[Get started](#)



6 people are talking about Rosetta Zoo right now.

[Join in](#)

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!**



 @kruksandor
www.sandorkruk.com

