Crowdsourced data analysis in EOSC

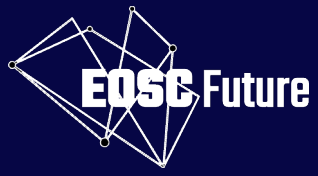
OR: How to increase EOSC use by many factors of ten using citizen science

Prof. Stephen Serjeant; Dr James Pearson; Dr Hugh Dickinson

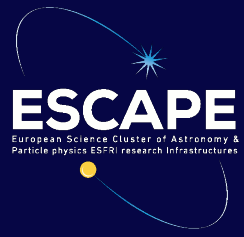
The Open University

16th November 2022

REINFORCE
REsearch INfrastructures FOR Citizens in Europe



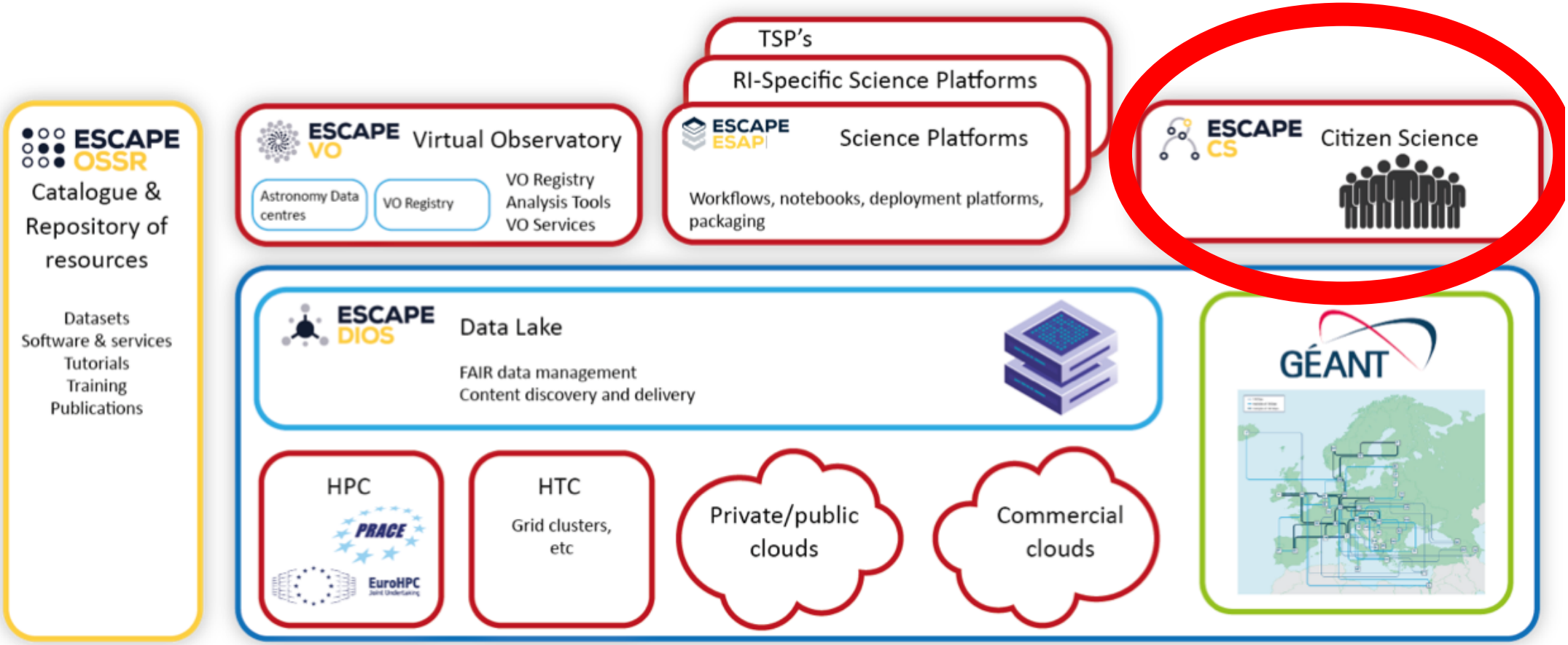
EOSC Future



ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

The Open University

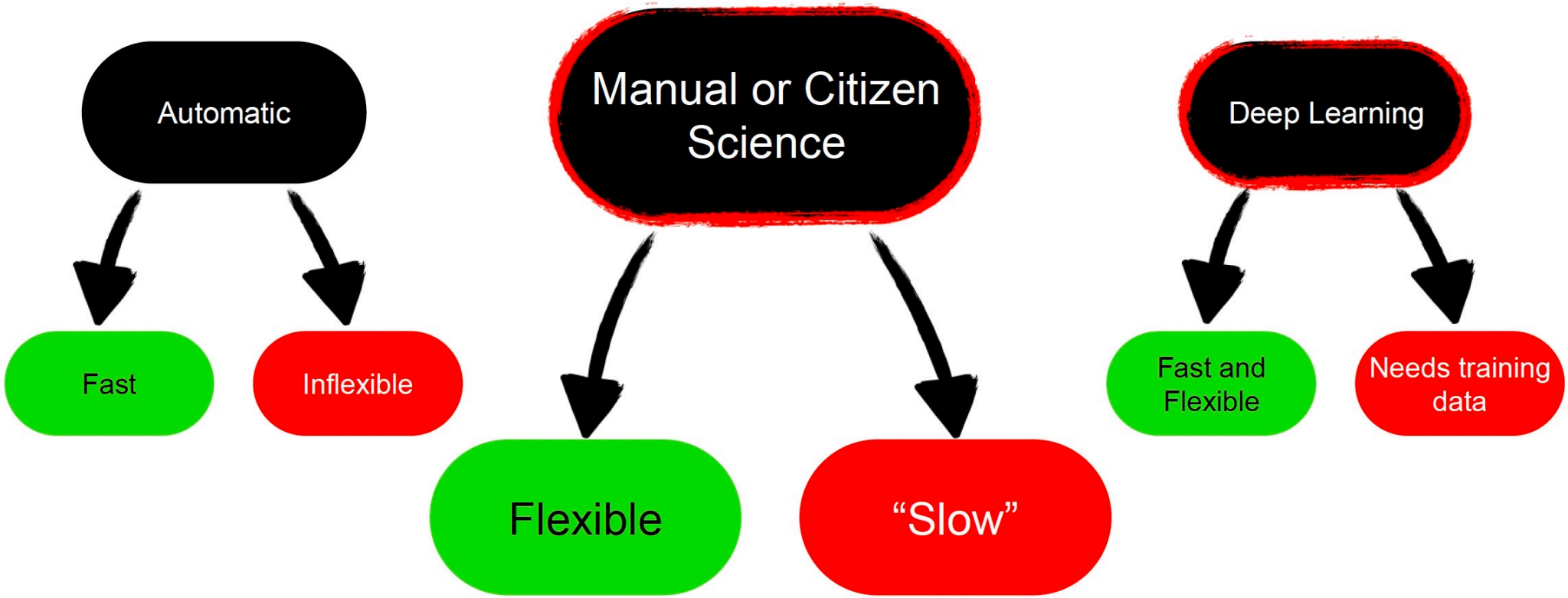
ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 824064.
REINFORCE has received funding from the European Union's Horizon 2020 project call H2020-SwafS-2018-2020 under Grant Agreement no. 872859
The EOSC Future project is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020 - Grant Agreement Number 101017536.



The Open University

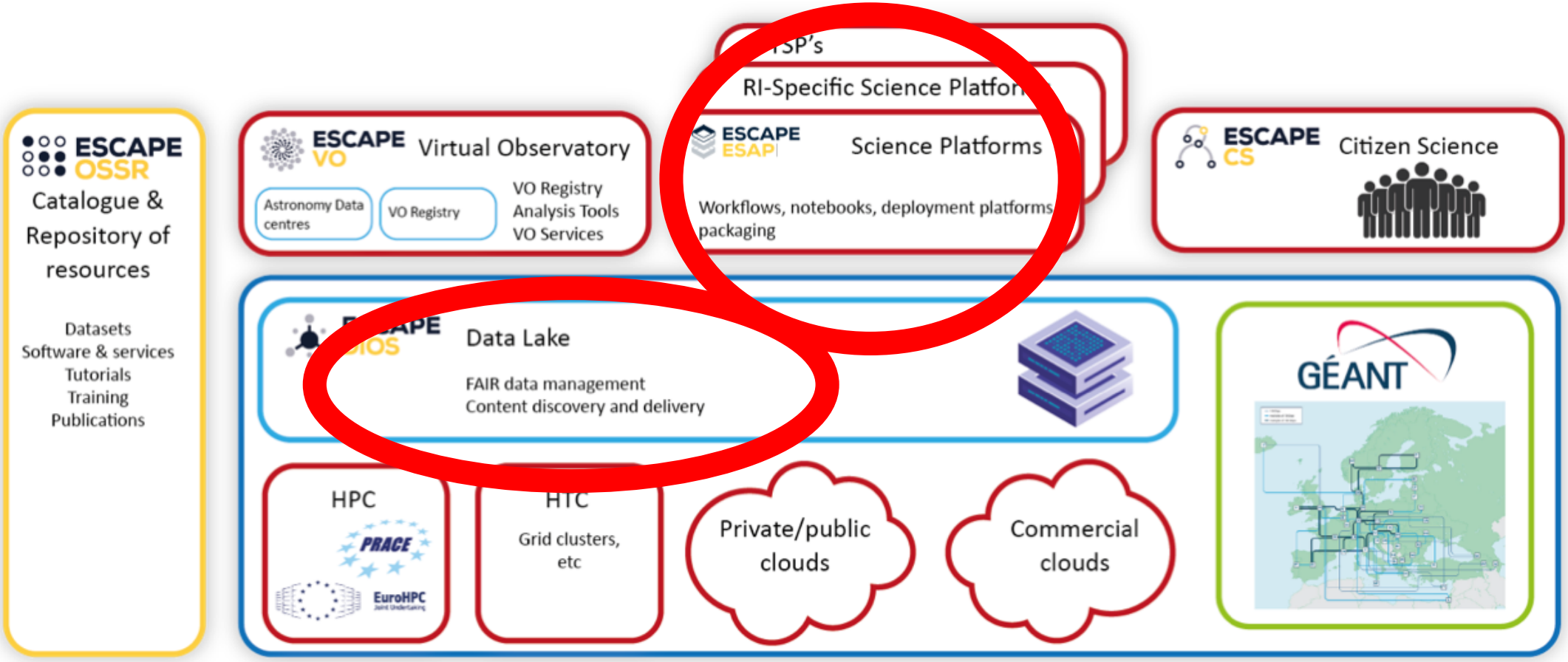
Improve access to data and tools through citizen science crowdsourcing experiments for most of the facilities in the ESCAPE remit.


newsbiscuit.com

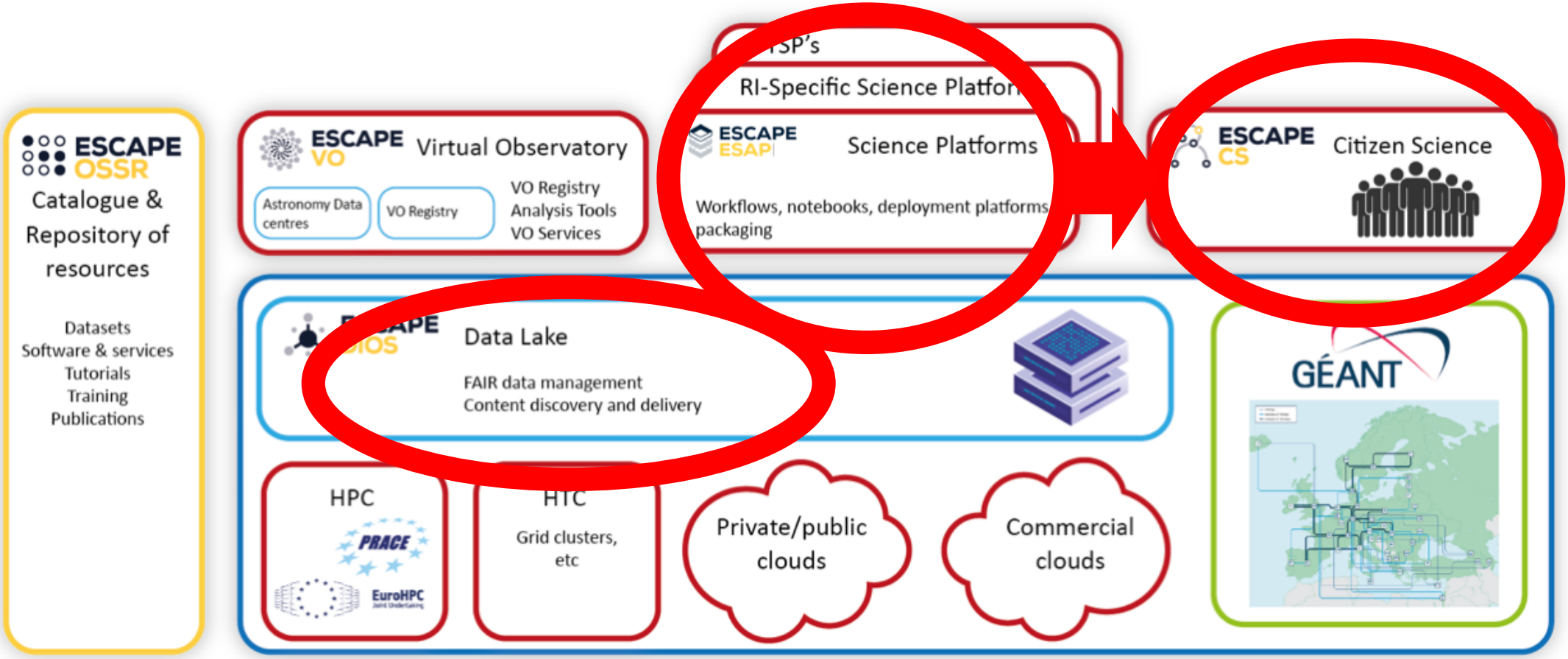


The Open University

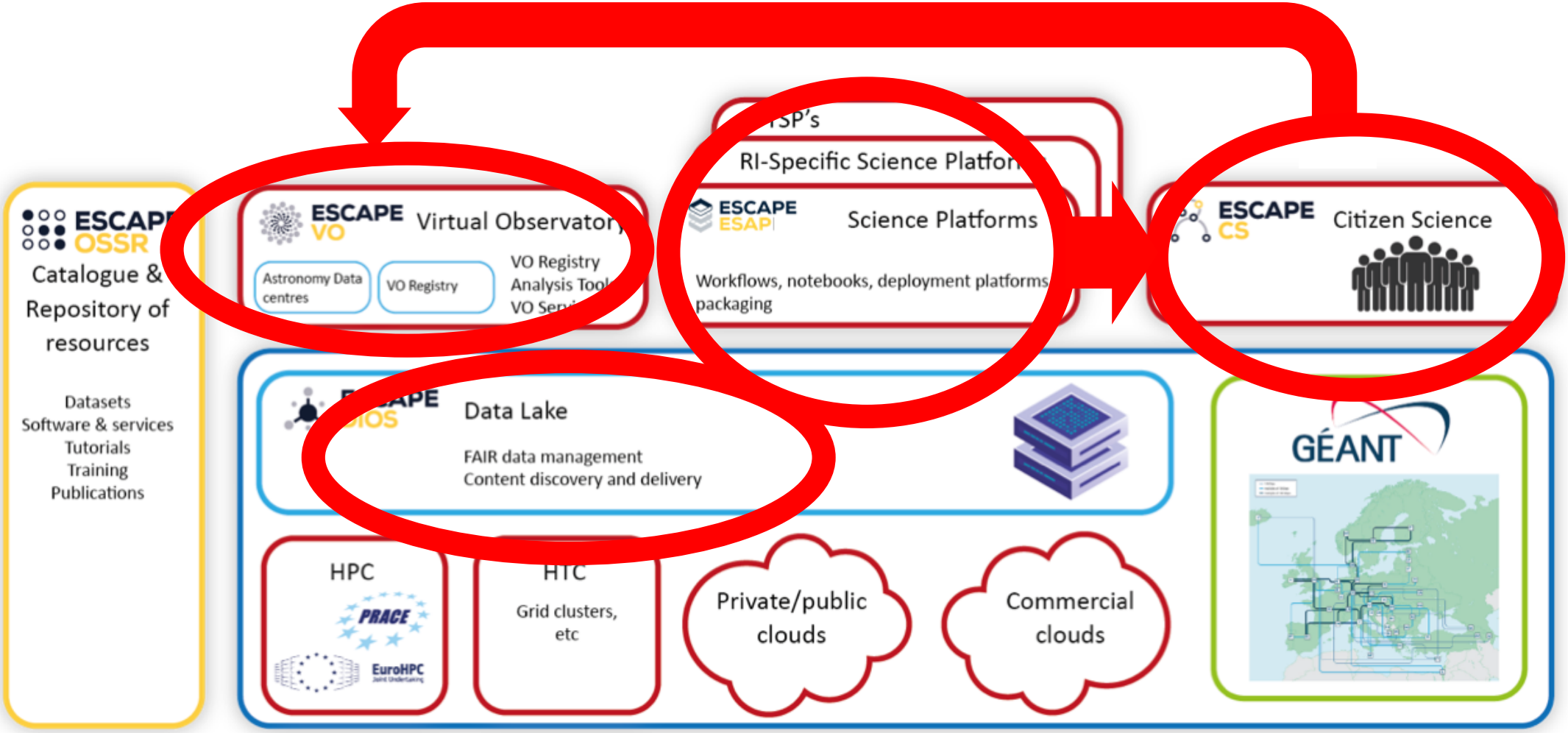
Image credit: Hugh Dickinson



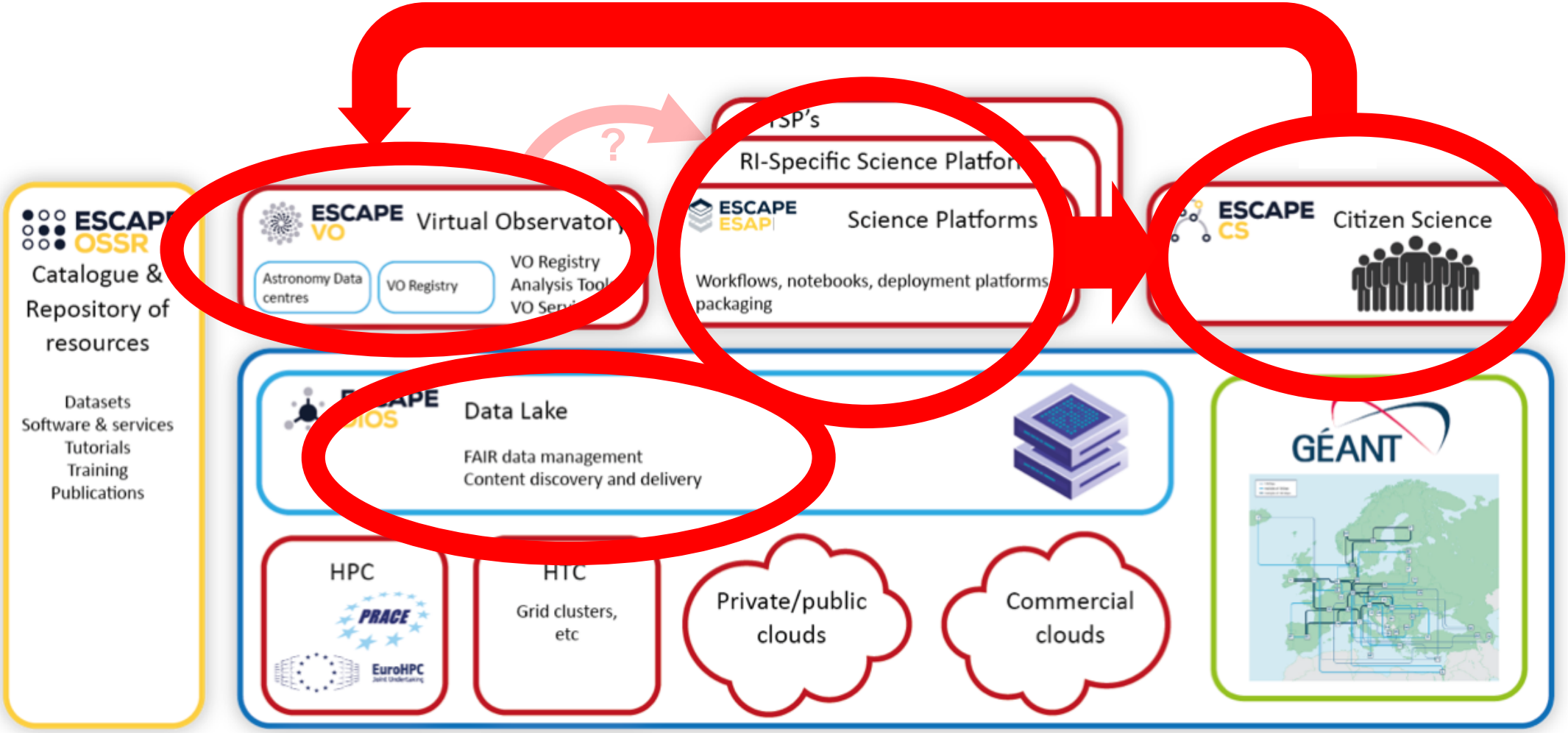
The Open University



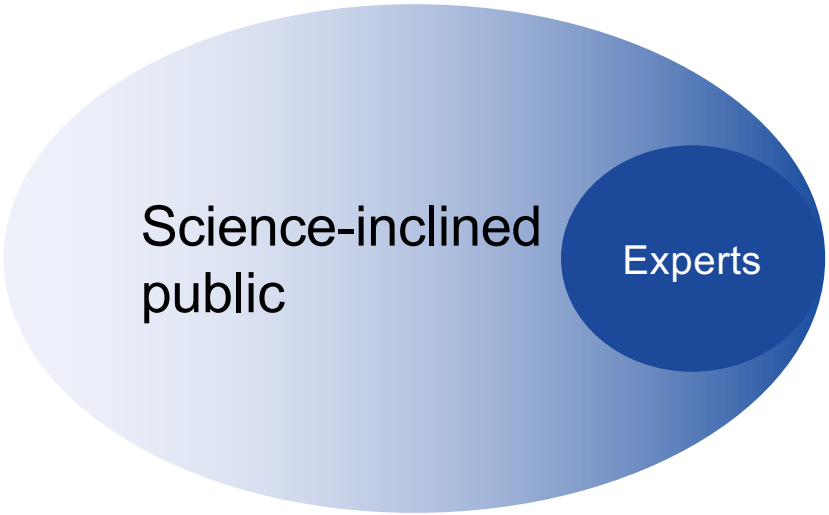
The Open University

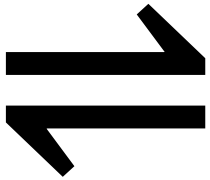


The Open University



The Open University





EUROPEAN OPEN SCIENCE CLOUD

The Open University

Zooniverse: Advanced Project Building

Description: Demonstrates techniques for advanced Zooniverse project management using Python.
Link: https://git.astron.nl/astron-sdc/escape-wp5/workflows/zooniverse-advanced-project-building
Author: Hugh Dickinson
Runtime Platform: Python
Keywords: jupyter-noteboo

Zooniverse: Integrating Machine Learning

Description: Demonstrates techniques to integrate your Zooniverse project with machine learning.
Link: https://git.astron.nl/astron-sdc/escape-wp5/workflows/zooniverse-integrating-machine-learning

- Tutorial Jupyter notebooks.
- Recorded walkthroughs of these tutorials.
- Speech-to-text documentation of these recordings, for greater accessibility.

Zooniverse: Advanced

Description: Demonstrates how to use the Zooniverse Caesar engine to aggregate your data.
Link: https://git.astron.nl/astron-sdc/escape-wp5/workflows/zooniverse-advanced-aggregation-with-caesar.git
Author: Hugh Dickinson
Runtime Platform: Python
Keywords: jupyter-notebook

Description: Shopping Cart and Zooniverse Example
Link: https://git.astron.nl/astron-sdc/escape-wp5/workflows/muon-hunters-example
Author:
Runtime Platform: Python
Keywords: jupyter-notebook

The Open University

Image credit: James Pearson

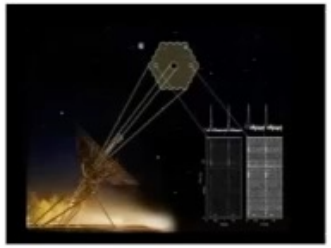
ESAP-GUI | Edit « Practice Project — Zooniv

sdc-dev.astron.nl/esap-gui/

ESCAPE ESAP ESFRI Science Analysis Platform

Archives | Multi Query | Interactive Analysis | Batch Analysis | Asynchronous Jobs | IVOA-SAMP | Logout James Pearson

WSRT-Apertif



Apertif Surveys

Data from the Apertif surveys include imaging and time-domain data. The time-domain products consist of high-time resolution filterbank data in the PSRFITS standard. The imaging data products include the raw observations in the measurement set (MS) standard format. In addition processed data

ASTRON VO

ASTRON VO Virtual Observatory

ASTRON Virtual Observatory

The Virtual Observatory defines a set of standards that can be used to download astronomical data. The ASTRON VO contains several image surveys, which are images in the FITS format. Since the VO is currenty under development, more data types will be

Zooniverse

Zooniverse Classification Database

The Zooniverse is the world's largest and most popular platform for people-powered research. This research is made possible by volunteers — more than a million people around the world who come together to assist

Virtual Observatory (VO)

IVOA

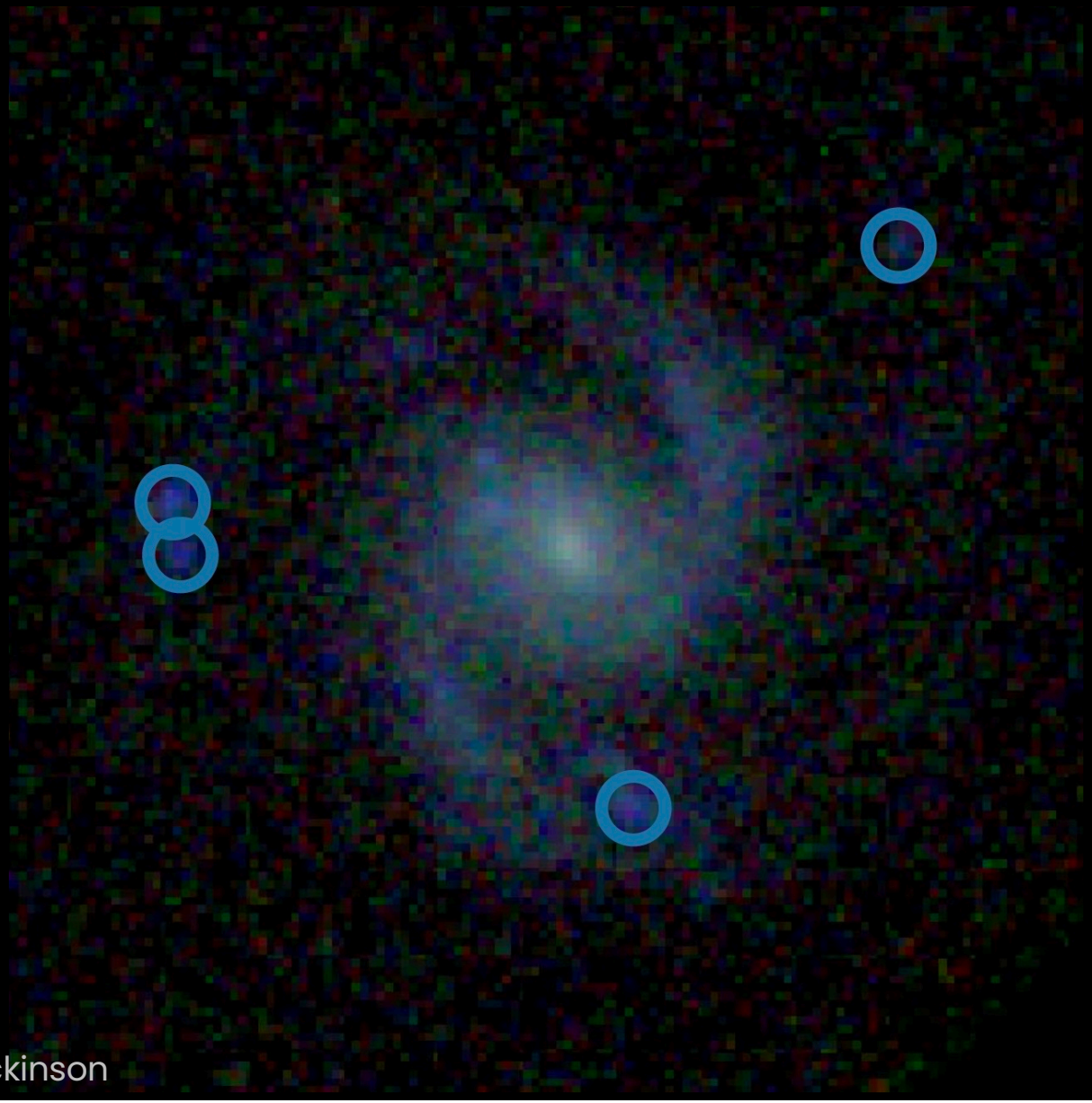
Virtual Observatory (VO)

The Virtual Observatory defines a set of standards that can be used to download astronomical data.

Visit Virtual Observatory (VO) Archives

The Open University

Image credit: Hugh Dickinson

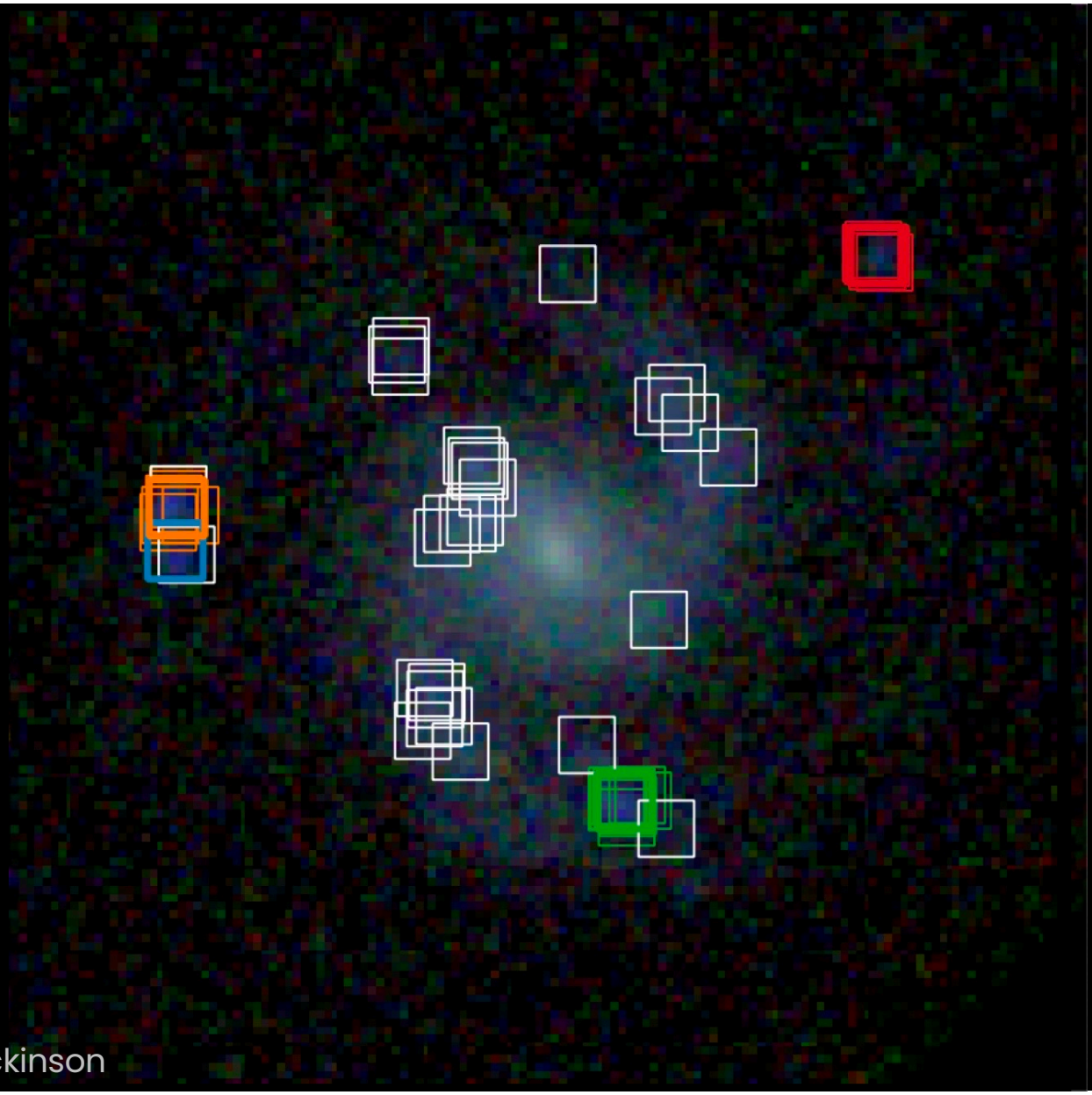


The Open University

Image credit: Hugh Dickinson

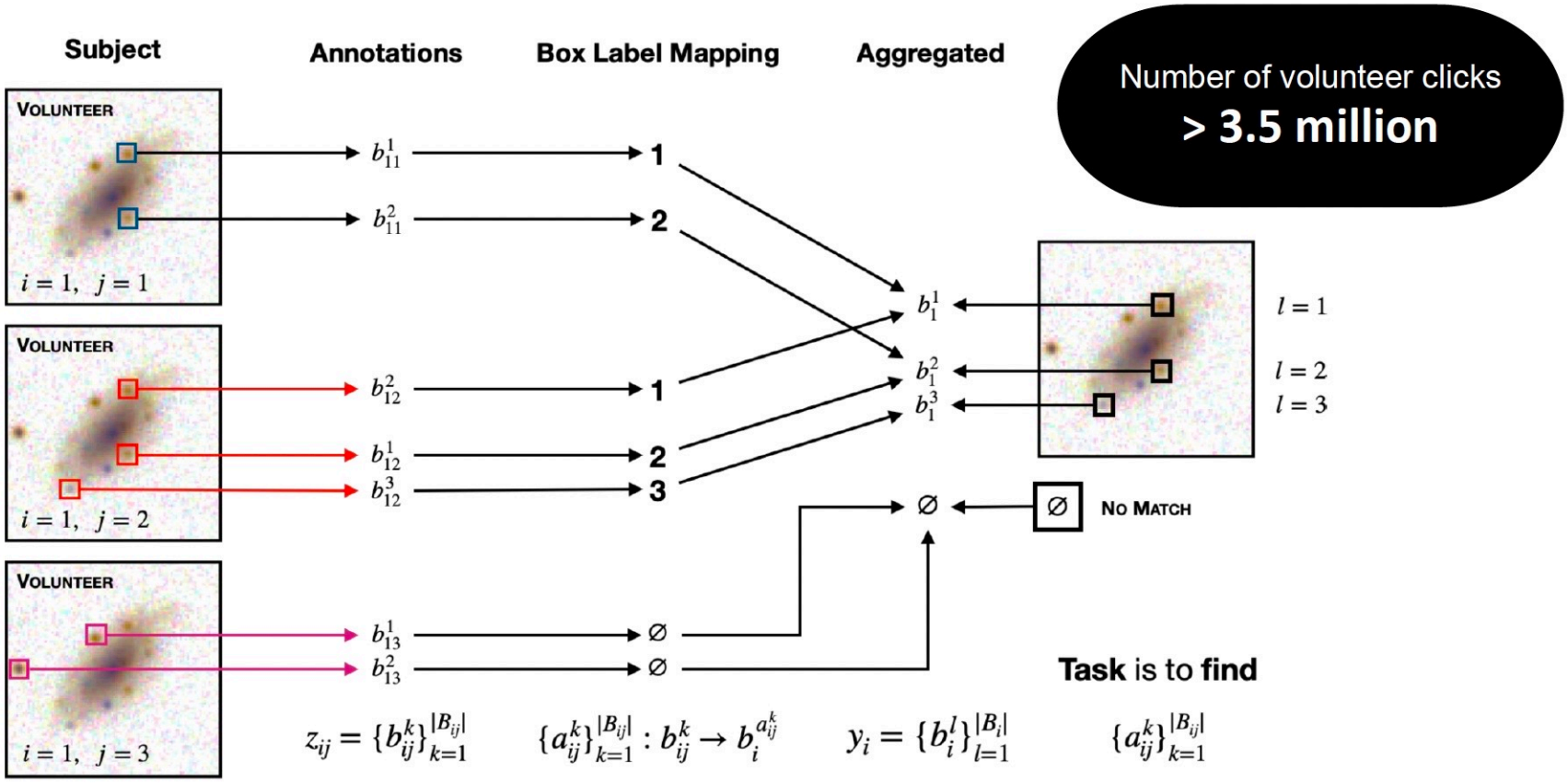
The Open University

Image credit: Hugh Dickinson



The Open University

Image credit: Hugh Dickinson

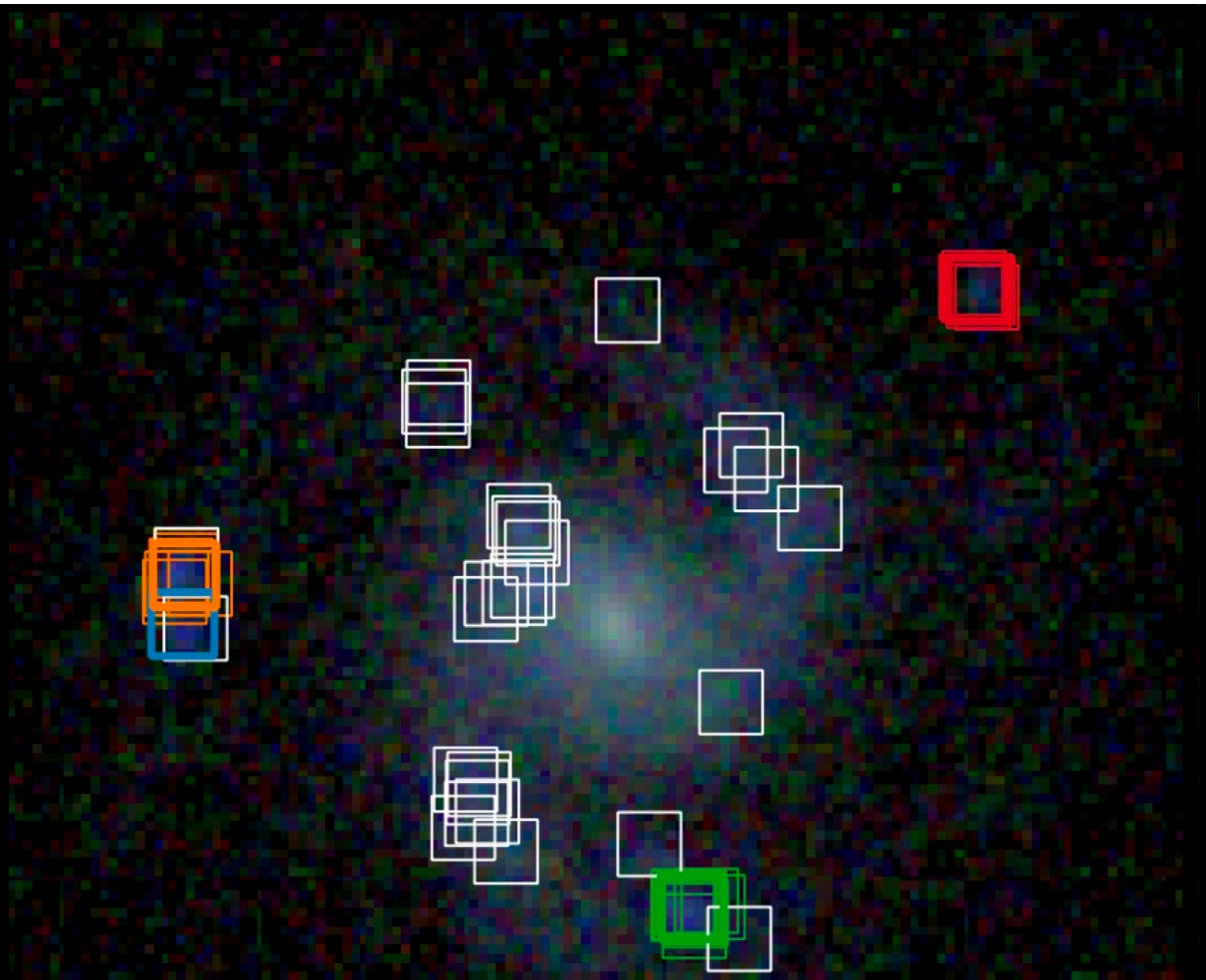


github.com/ou-astrophysics/BoxAggregator

arXiv Dickinson et al (2022) - arxiv.org/abs/2210.03684

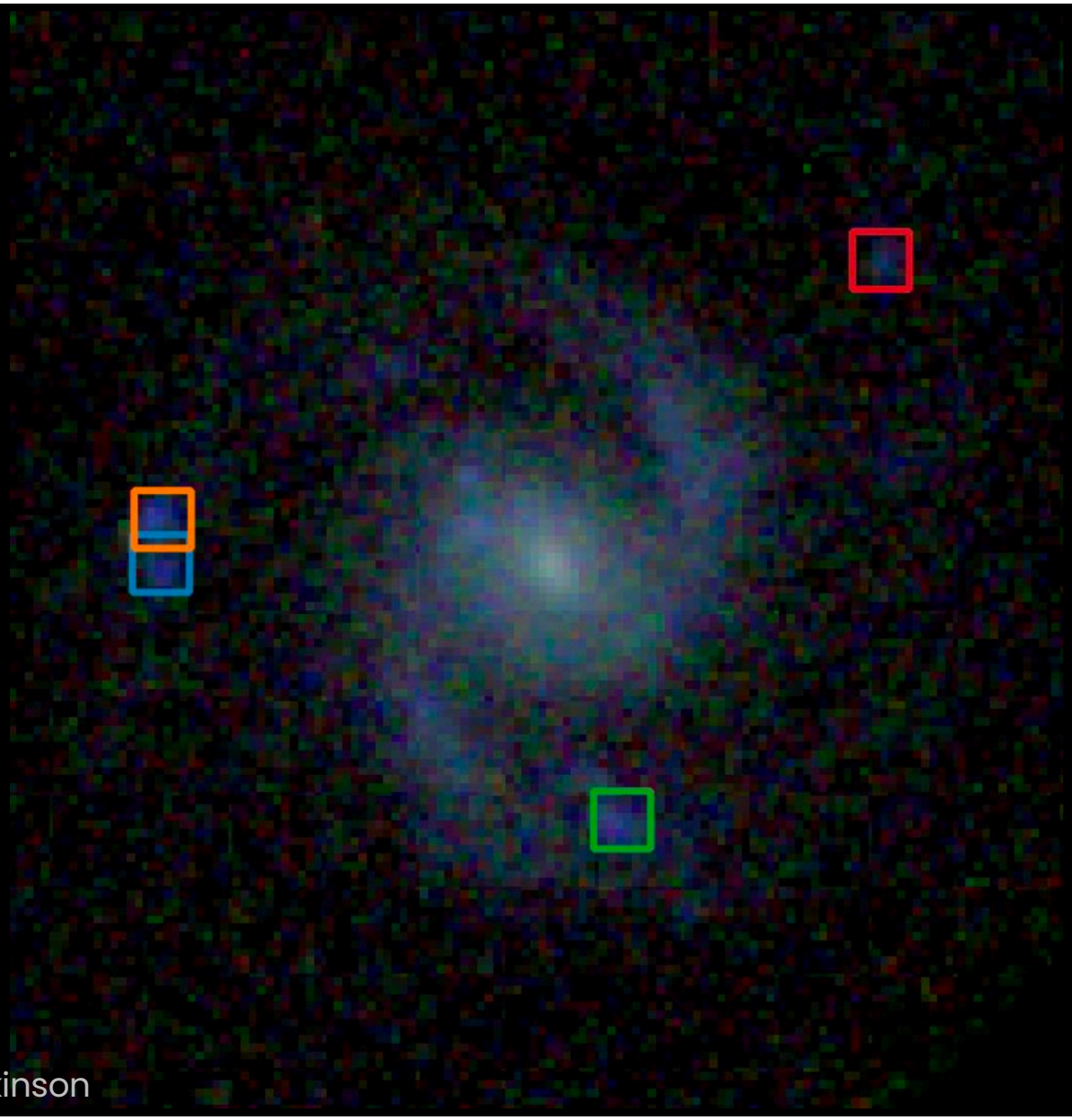
The Open University

Image credit: Hugh Dickinson



The Open University

Image credit: Hugh Dickinson



The Open University

Image credit: Hugh Dickinson

Number of clumpy galaxies:

~35,000

Number of potential clumps:

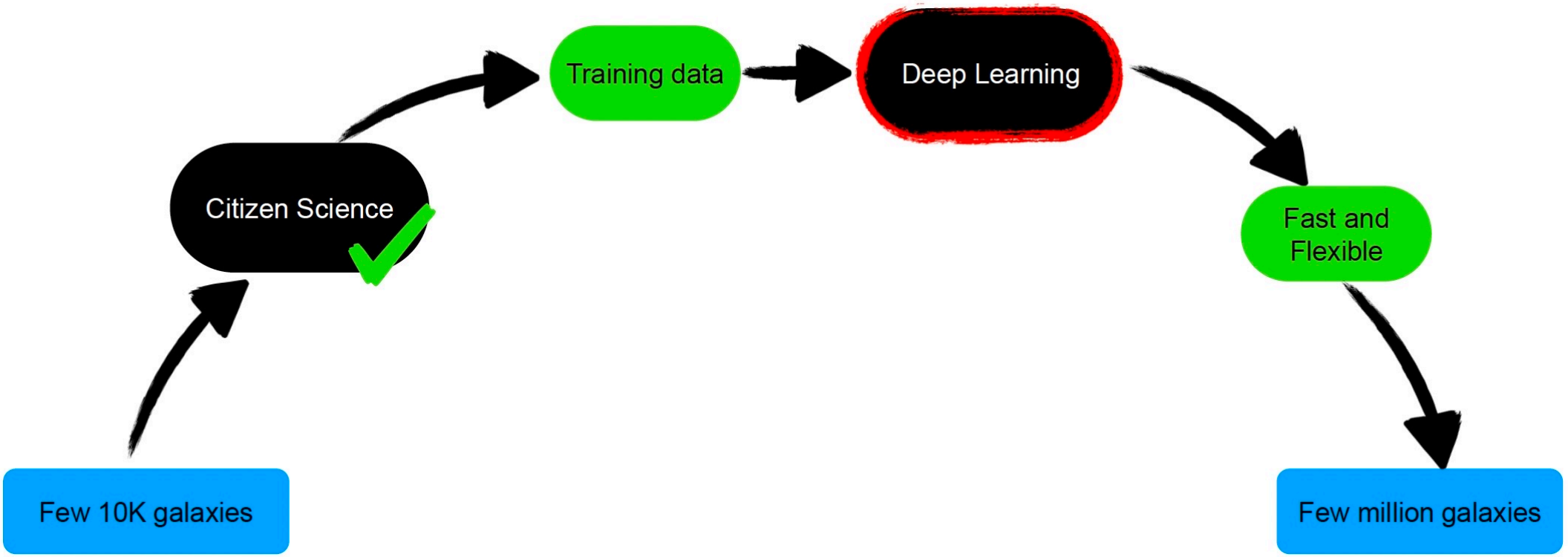
~100,000

First catalogue released!

arXiv Adams et al (2022) - arxiv.org/abs/2201.06581

The Open University

Image credit: Hugh Dickinson



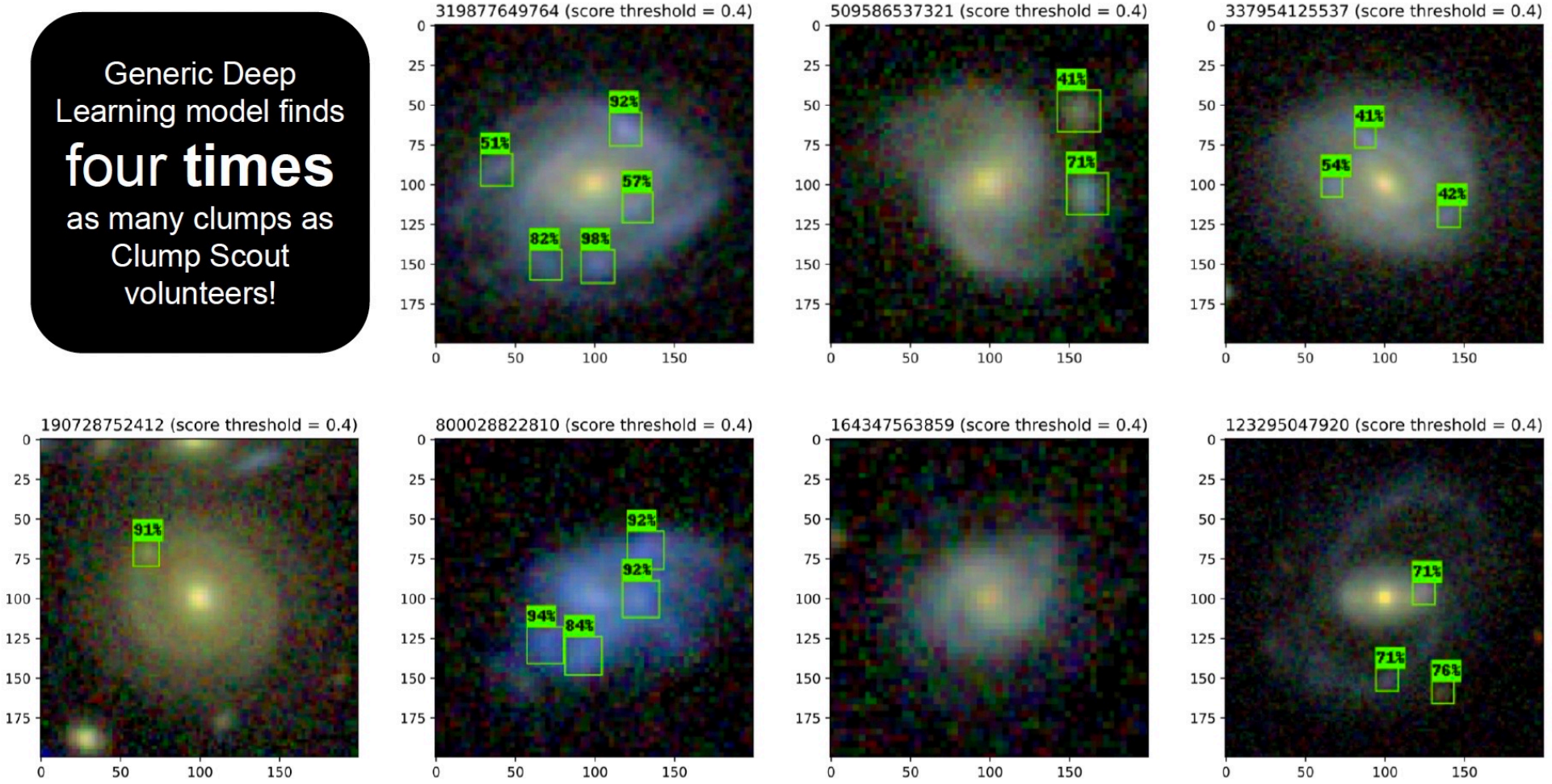
The Open University

Image credit: Hugh Dickinson

Generic Deep Learning model finds

four times

as many clumps as Clump Scout volunteers!



The Open University

Image credit: Hugh Dickinson

Number of Volunteers engaged:	Number of Galaxies inspected:	Number of clumpy galaxies:	Number of potential clumps:
~15,000	**~80,000**	**~35,000**	**~100,000**

Two papers now out!	First catalogue released!	ML Model Trained!	Third paper in prep!

github.com/ou-astrophysics/BoxAggregator

arXiv Adams et al (2022) - arxiv.org/abs/2201.06581

arXiv Dickinson et al (2022) - arxiv.org/abs/2210.03684

The Open University

Image credit: Hugh Dickinson

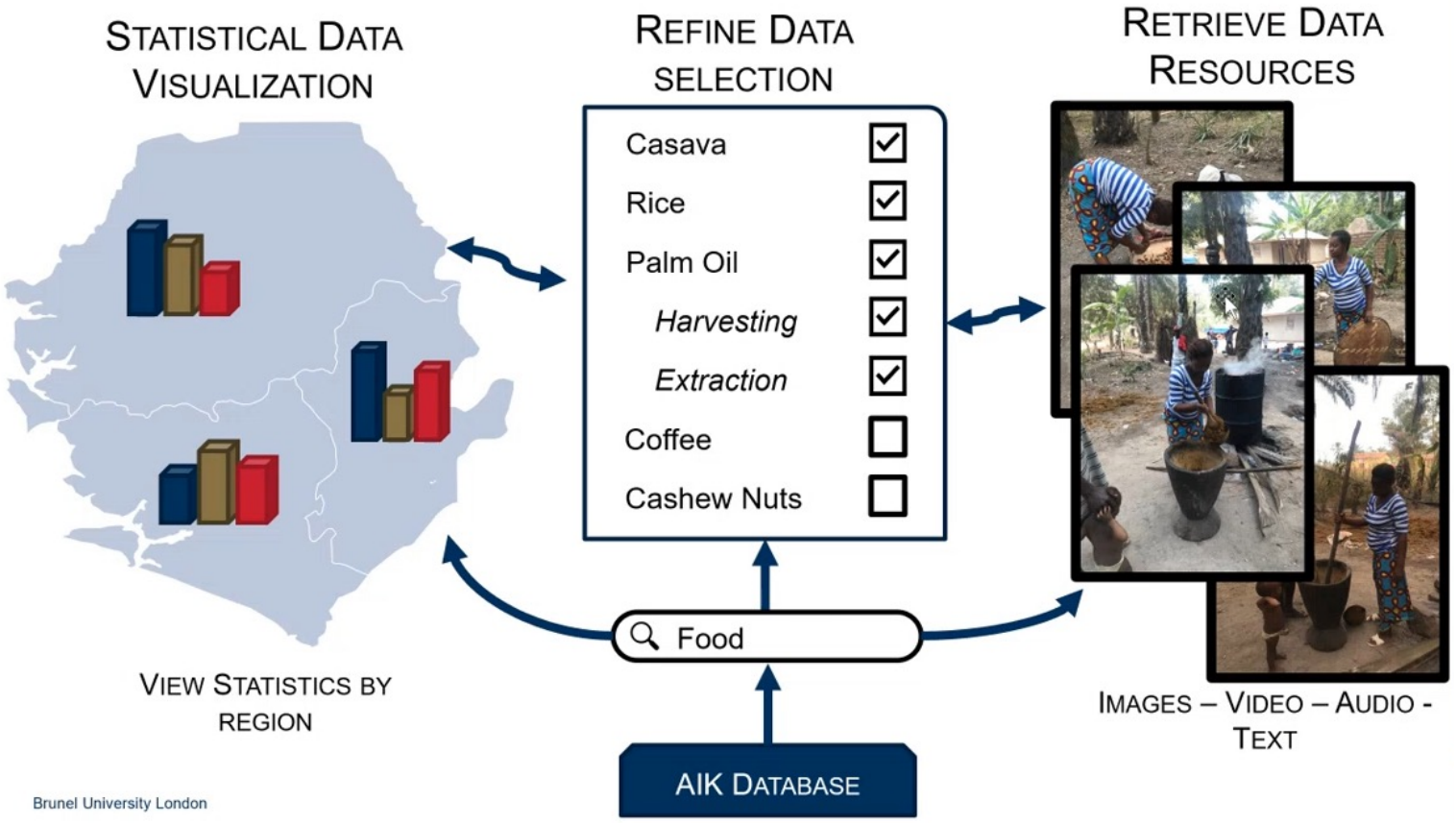
Knitting Patterns in Domestic Magazines

African Indigenous Knowledge (AIK) for Resilient Food Systems

Image credits: James Pearson

The Open University

African Indigenous Knowledge



The Open University

Brunel University London

Knitting patterns

Four Colour Favourite

Enchanting in pastel shades, this attractive little cardigan will fit a two-to-three-year-old

MATERIALS : Of Sirdar Majestic Wool, 3-ply, allow 2 ozs. in white and 1 oz. each of green, rose-pink and blue ; 1 pair each of No. 10 and No. 11 knitting needles ; 4 buttons.

Measurements : To fit a 22-23 inch chest size ; length from shoulder, 12½ inches ; sleeve seam, 8½ inches.

Abbreviations : K., knit ; p., purl ; st., stitch ; sts., stitches ; rep., repeat ; inc., increase (by working twice into same st.) ; dec., decrease (by taking 2 sts. tog.) ; W, white ; G, green ; R, rose ; B, blue.

THE BACK

RIGHT FRONT

LEFT FRONT

THE SLEEVES (both alike)

TO COMPLETE

Completed item

Materials

Title

The information needed to answer the questions can be found in these boxes.

- **What limits the take up of crowdsourced data mining in EOSC?**
  - Trust in the reliability? Skills at aggregating the data? Temptation just to pay Amazon MT?
  - Seeing science results will help
  - Seeing it work up close will help
  - Build multi-disciplinary exemplar experiments following ESCAPE model
  - Create worked examples of plug-and-play notebooks for running projects in EOSC
- **Improve integration with other EOSC services, eg VO, AAI, virtuous circle with ESAP ML (for large projects)**
- **Open data standards in FAIRsharing.org?**
- **Dedicated EOSC task force for citizen science?**
- **Funding for science analysis platform for interdisciplinary citizen science?**

The Open University

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 824064.
REINFORCE has received funding from the European Union's Horizon 2020 project call H2020-SwafS-2018-2020 under Grant Agreement no. 872859
The EOSC Future project is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020 - Grant Agreement Number 101017536.

REINFORCE
REsearch INfrastructures FOR Citizens in Europe

EOSC Future

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

Thank you for listening

https://git.astron.nl/astron-sdc/escape-wp5

Prof Stephen Serjeant

Dr James Pearson

Dr Hugh Dickinson

The Open University

60 SECOND ADVENTURES IN ARTIFICIAL INTELLIGENCE