

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

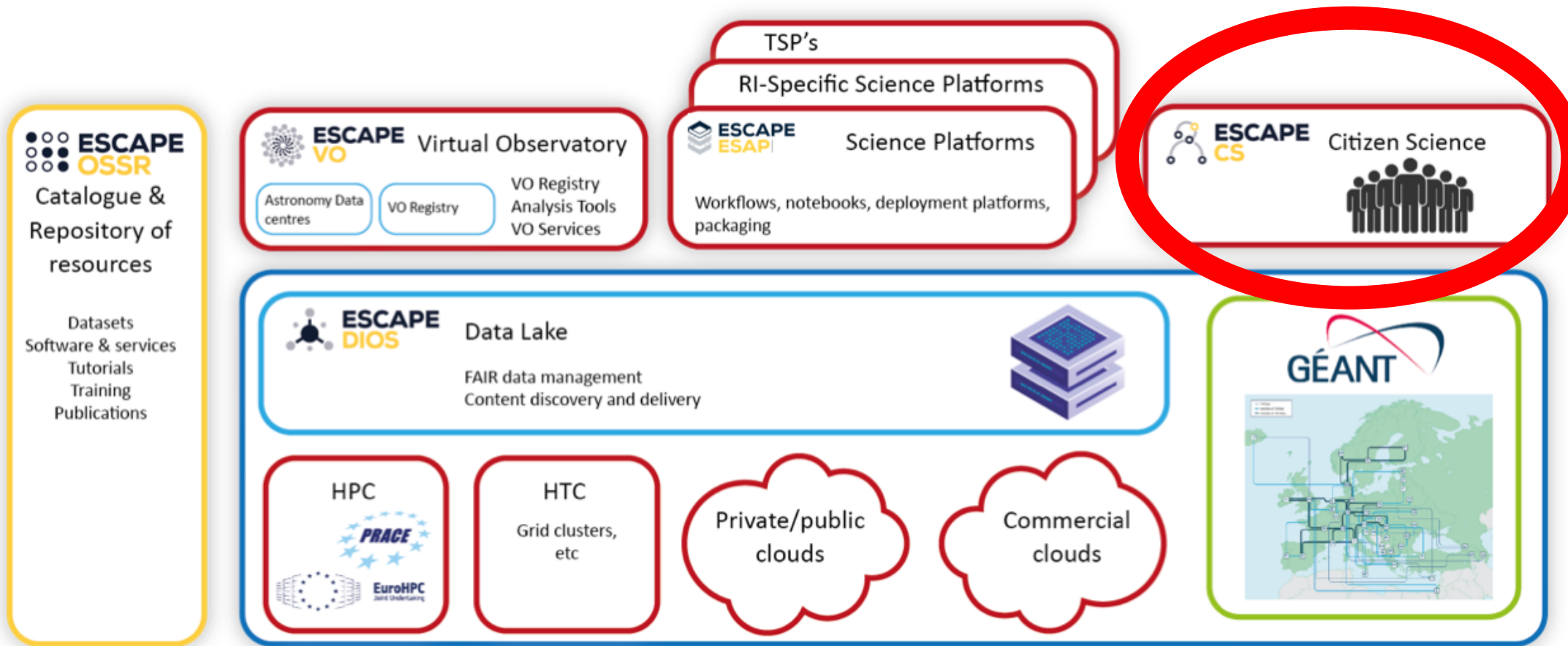
The Open University
16th November 2022



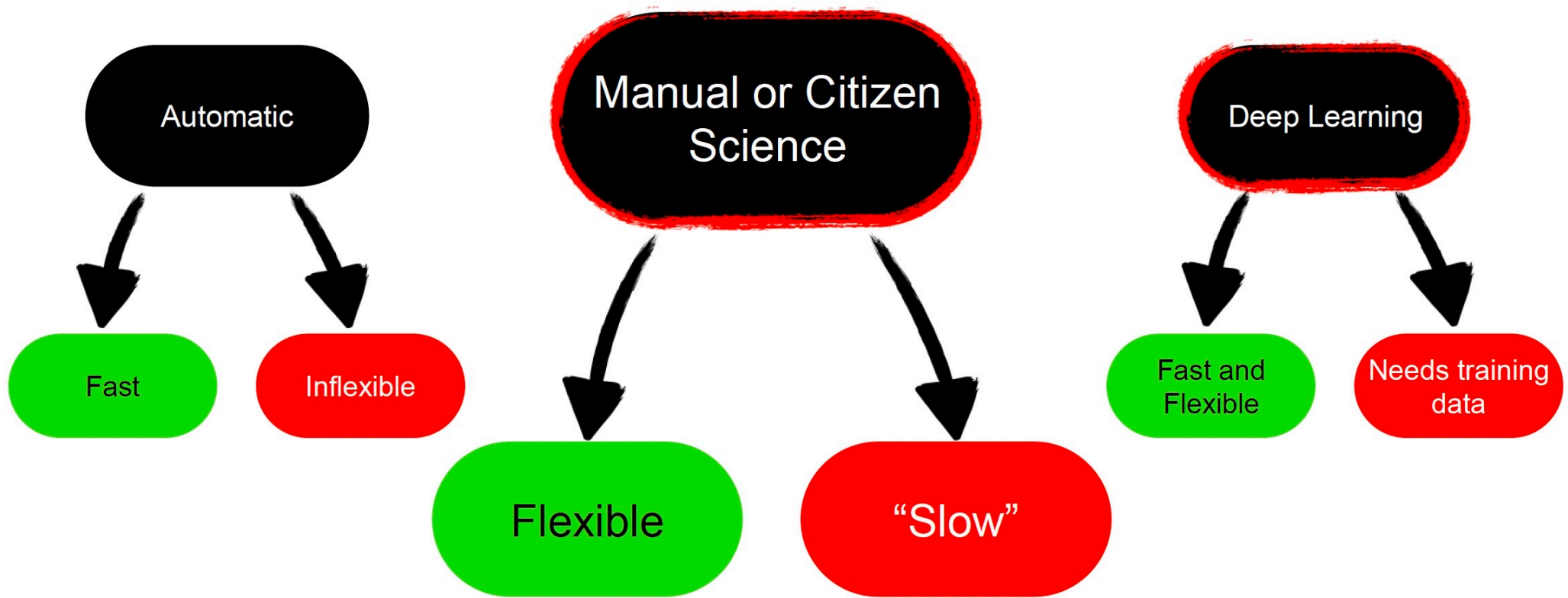
REINFORCE
REsearch INFrastructures FOR Citizens in Europe



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.







ESCAPE OSSR
 Catalogue & Repository of resources

- Datasets
- Software & services
- Tutorials
- Training
- Publications

ESCAPE VO Virtual Observatory

- Astronomy Data centres
- VO Registry
- VO Registry
- Analysis Tools
- VO Services

ESCAPE ESAP Science Platforms

Workflows, notebooks, deployment platforms packaging

ESCAPE ESAP Science Platforms

Workflows, notebooks, deployment platforms packaging

ESCAPE CS Citizen Science

ESCAPE EOS Data Lake

FAIR data management
 Content discovery and delivery

HPC

HTC

Grid clusters, etc

Private/public clouds

Commercial clouds

GÉANT

ESCAPE OSSR
 Catalogue & Repository of resources

Datasets
 Software & services
 Tutorials
 Training
 Publications

ESCAPE VO Virtual Observatory

Astronomy Data centres
 VO Registry
 VO Registry
 Analysis Tools
 VO Services

ISP's
 RI-Specific Science Platform

ESCAPE ESAP Science Platforms

Workflows, notebooks, deployment platforms
 packaging

ESCAPE CS Citizen Science

ESCAPE EOS Data Lake

FAIR data management
 Content discovery and delivery

HPC

PRACE
 EuroHPC
 Joint Undertaking

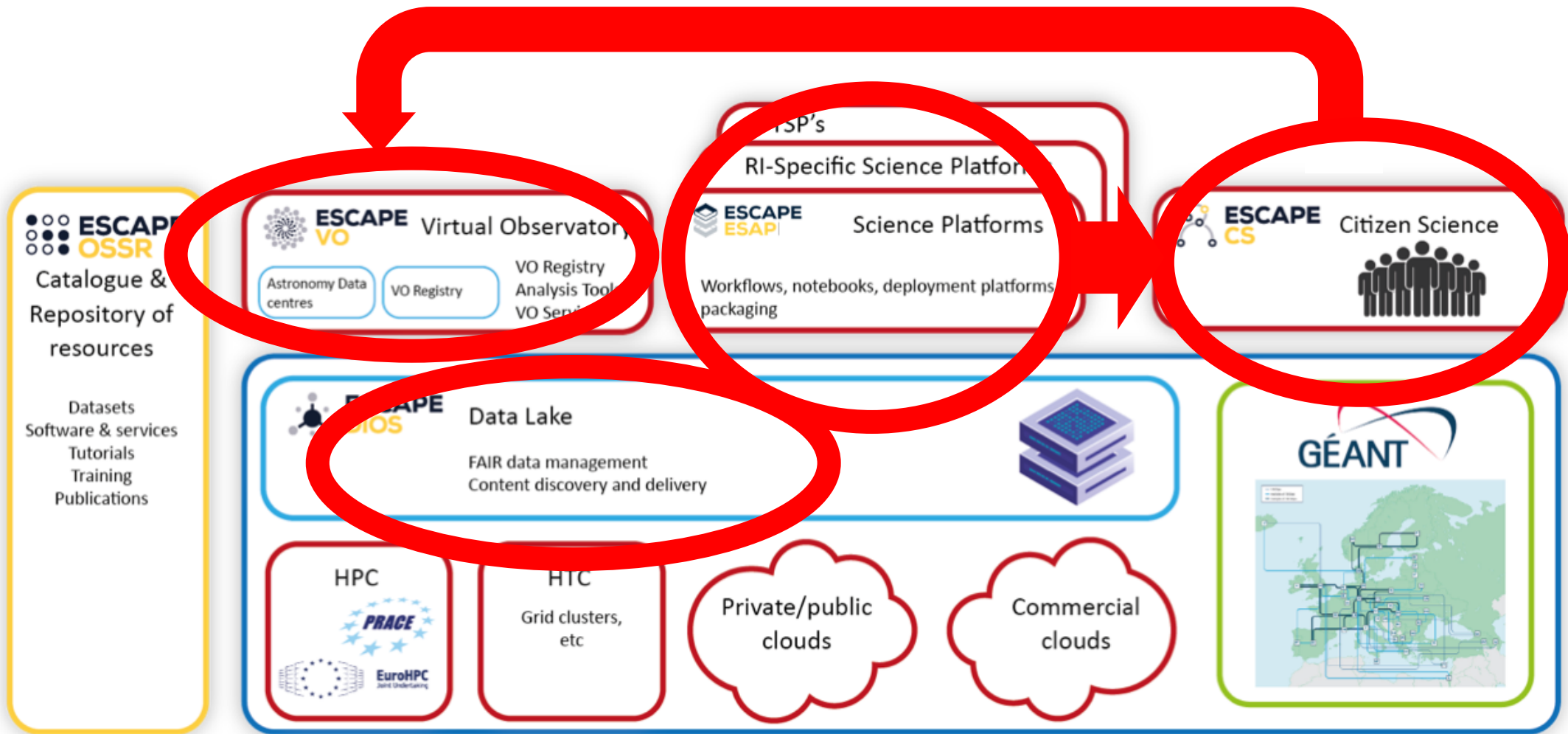
HTC

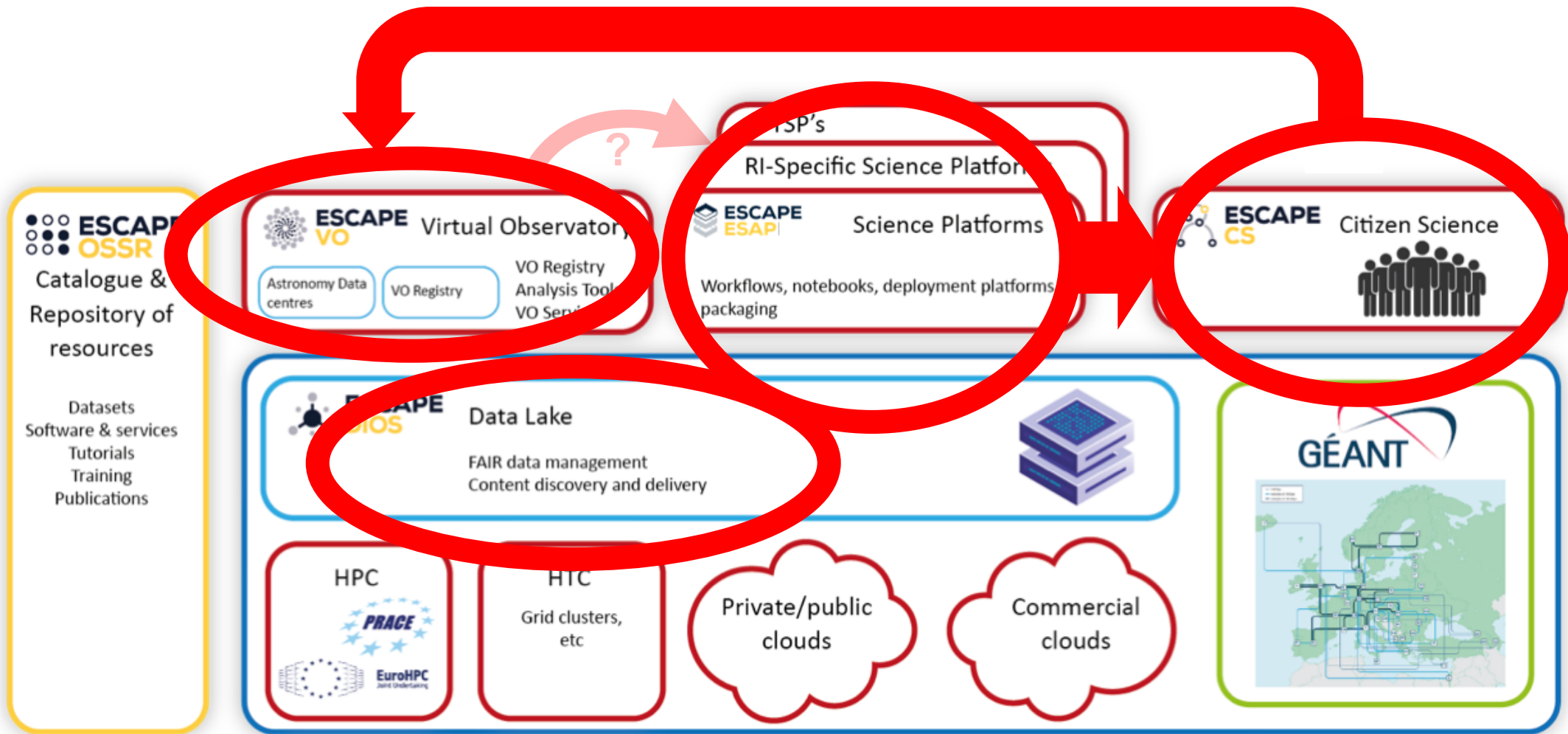
Grid clusters,
 etc

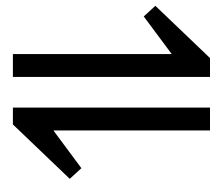
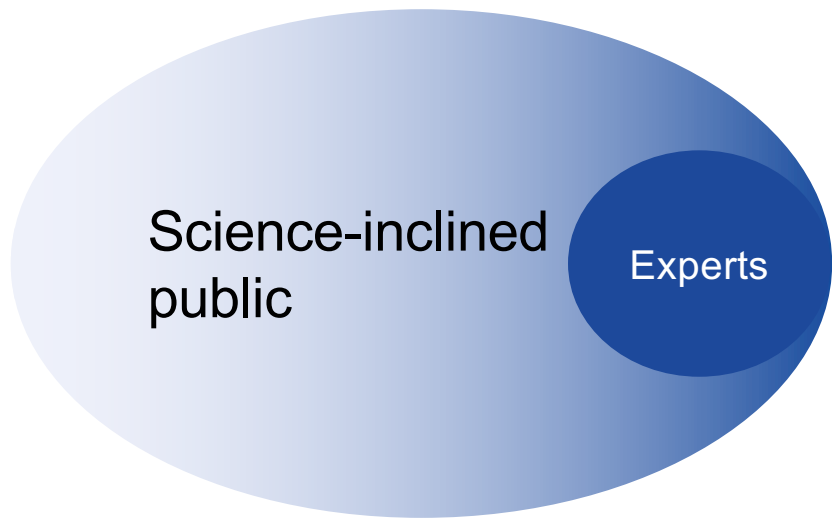
Private/public
 clouds

Commercial
 clouds

GÉANT







**EUROPEAN OPEN
SCIENCE CLOUD**

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-notebook

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

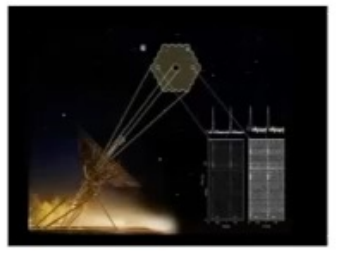
Link: <https://git.astron.nl/astron-sdc/escape-wp5/workflows/muon-hunters-example>

Author:

Runtime Platform: Python

Keywords: jupyter-notebook

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 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 currently 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)

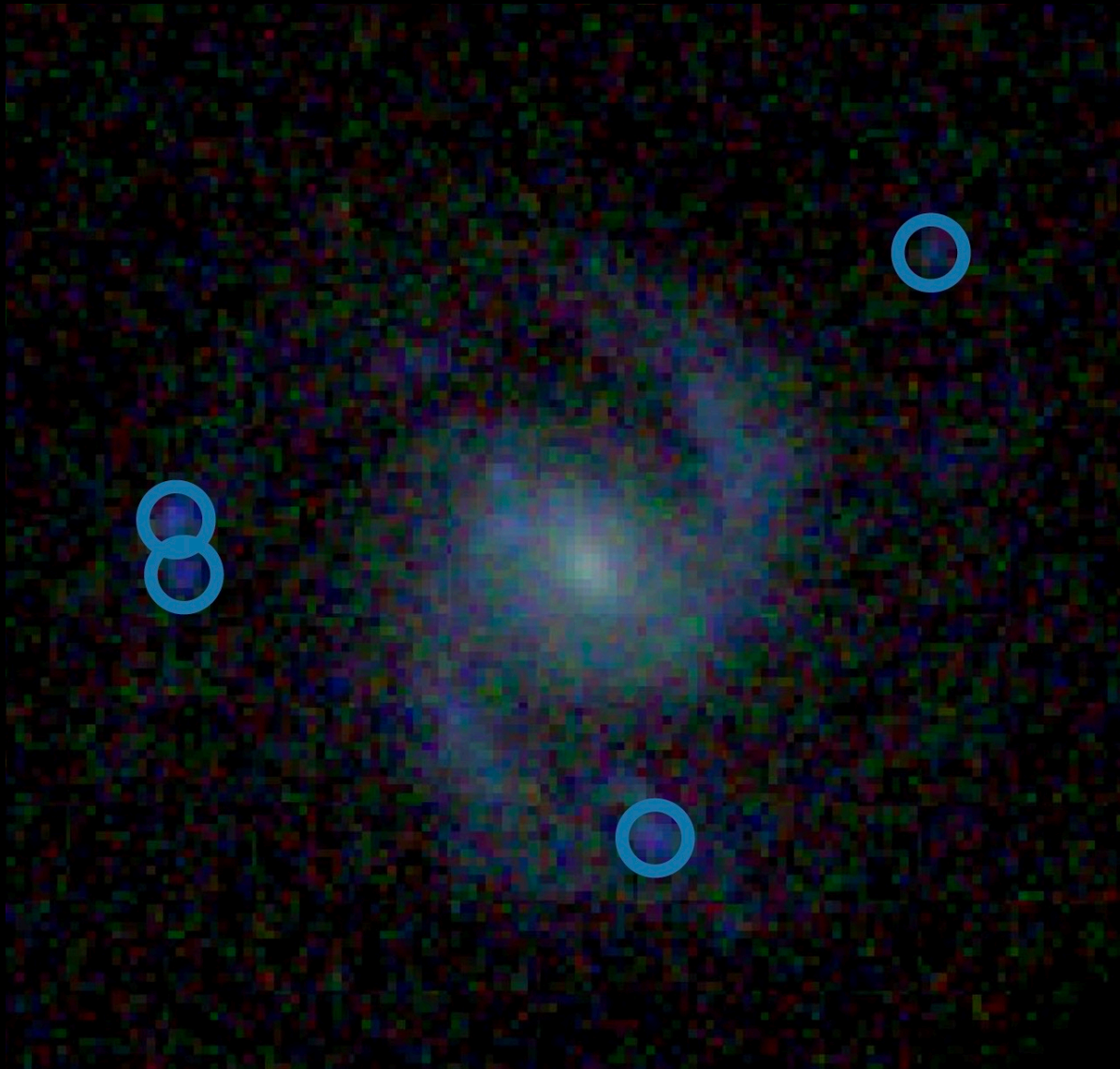


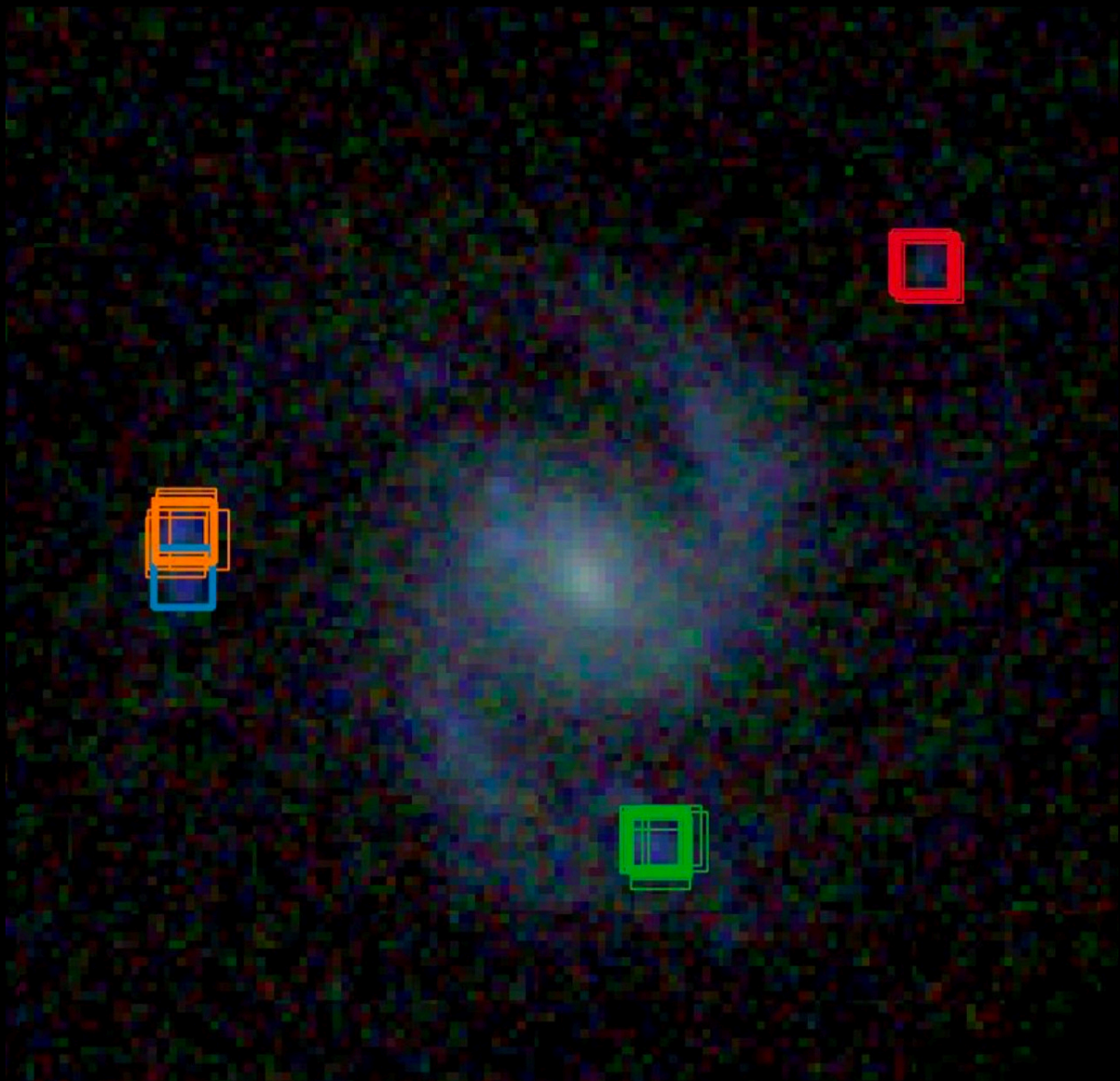
Virtual Observatory (VO)

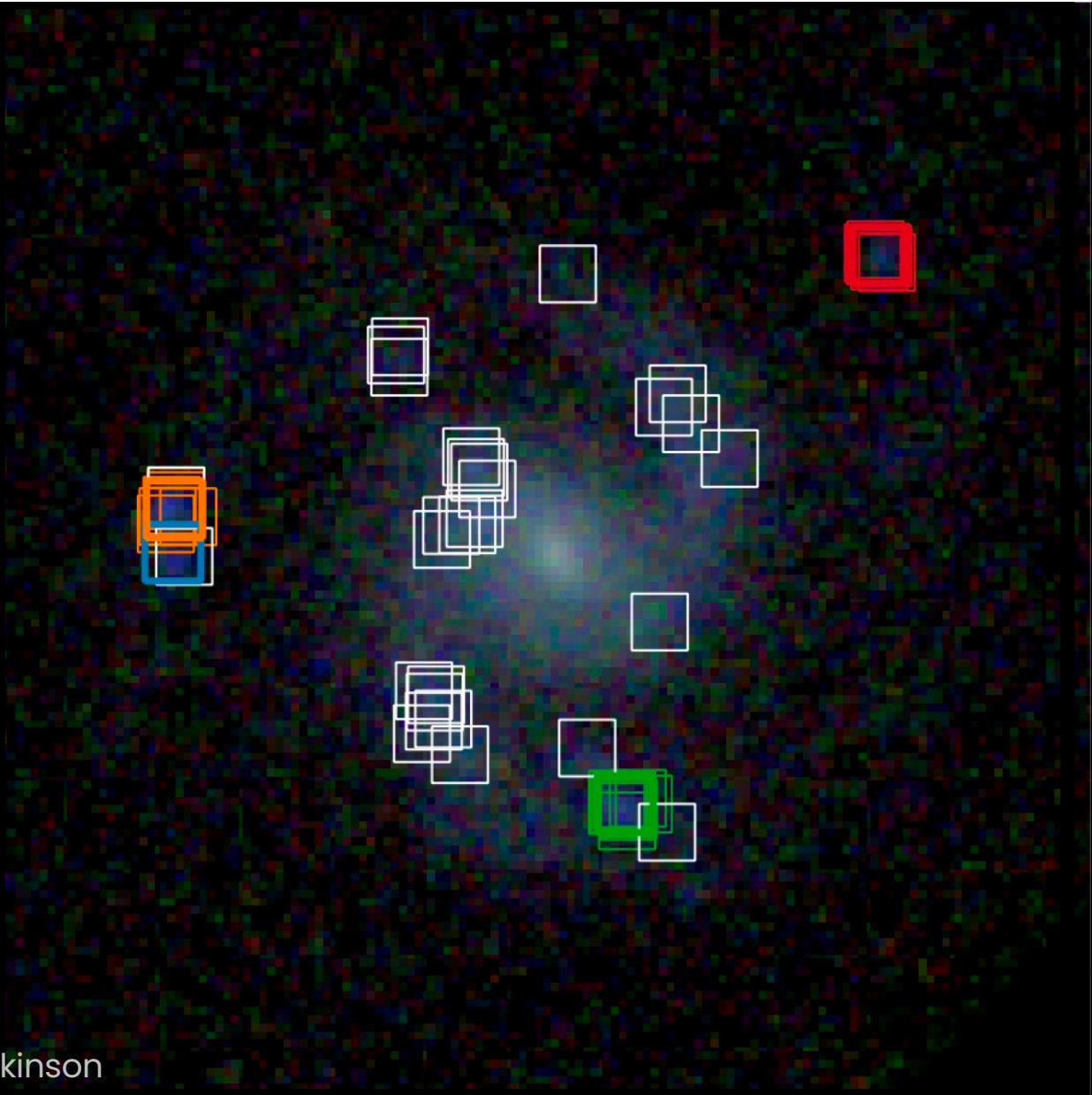
The Virtual Observatory defines a set of standards that can be used to download astronomical data.

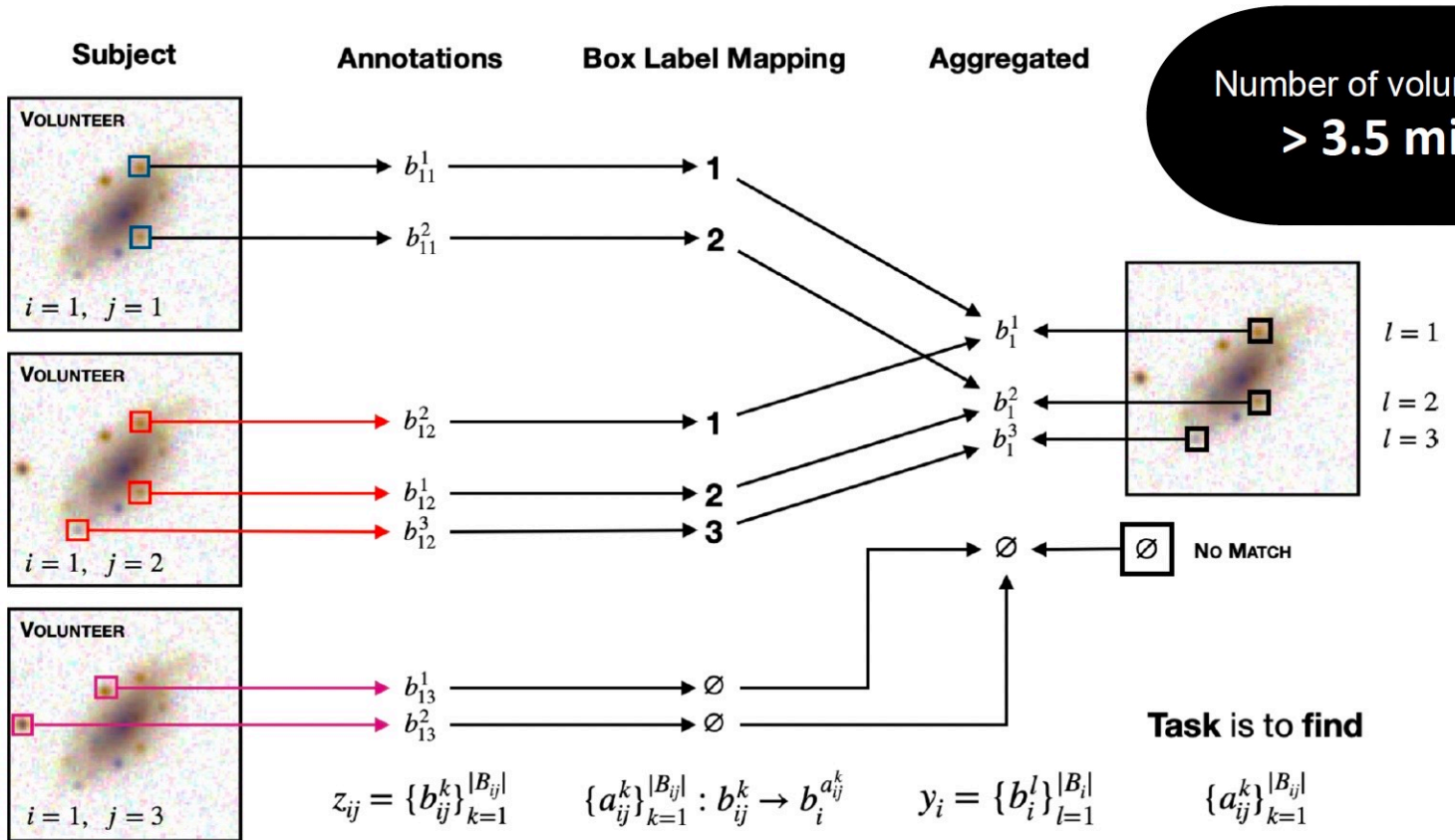
[Visit Virtual Observatory \(VO\) Archives](#)











 github.com/ou-astrophysics/BoxAggregator


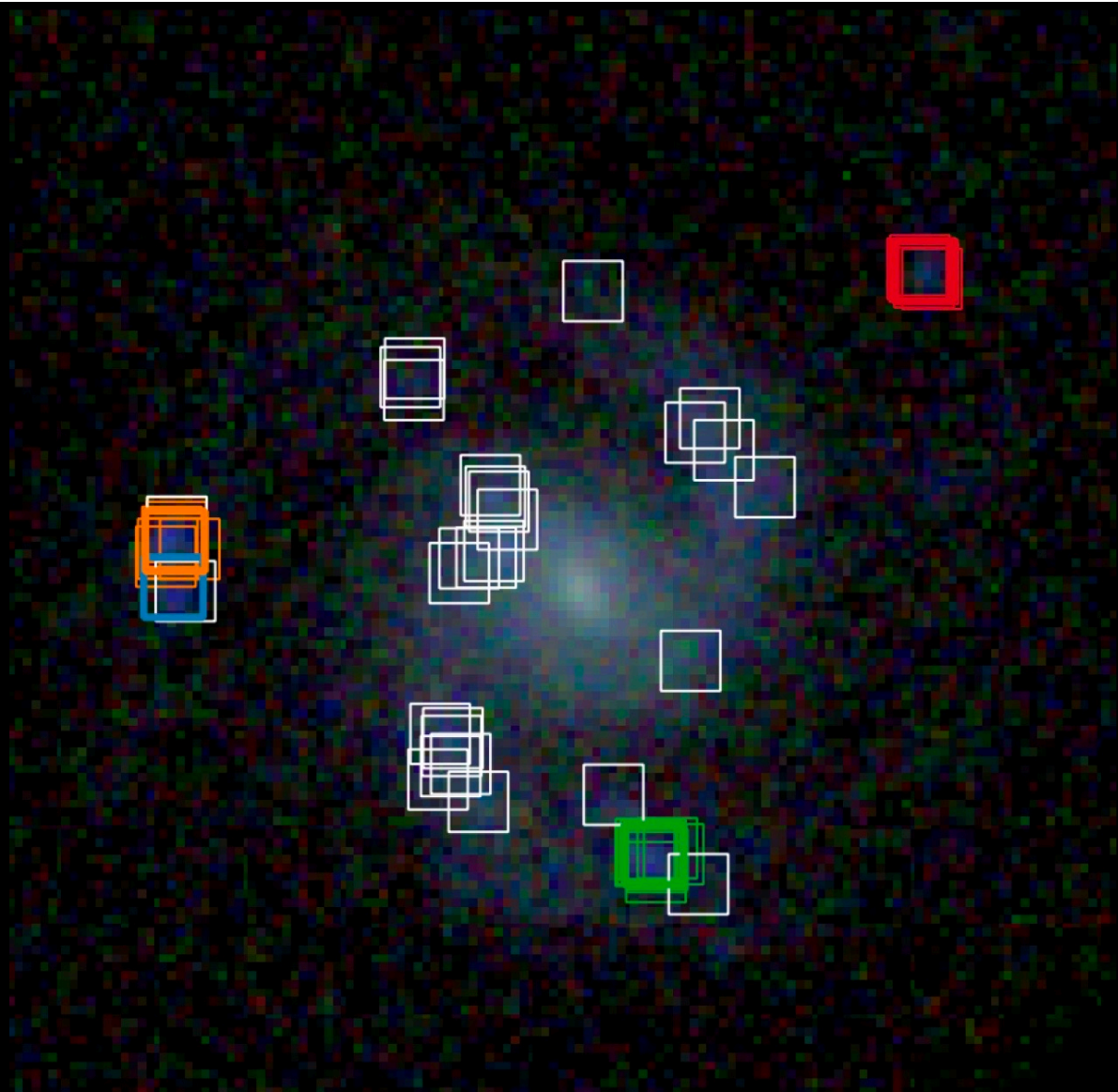
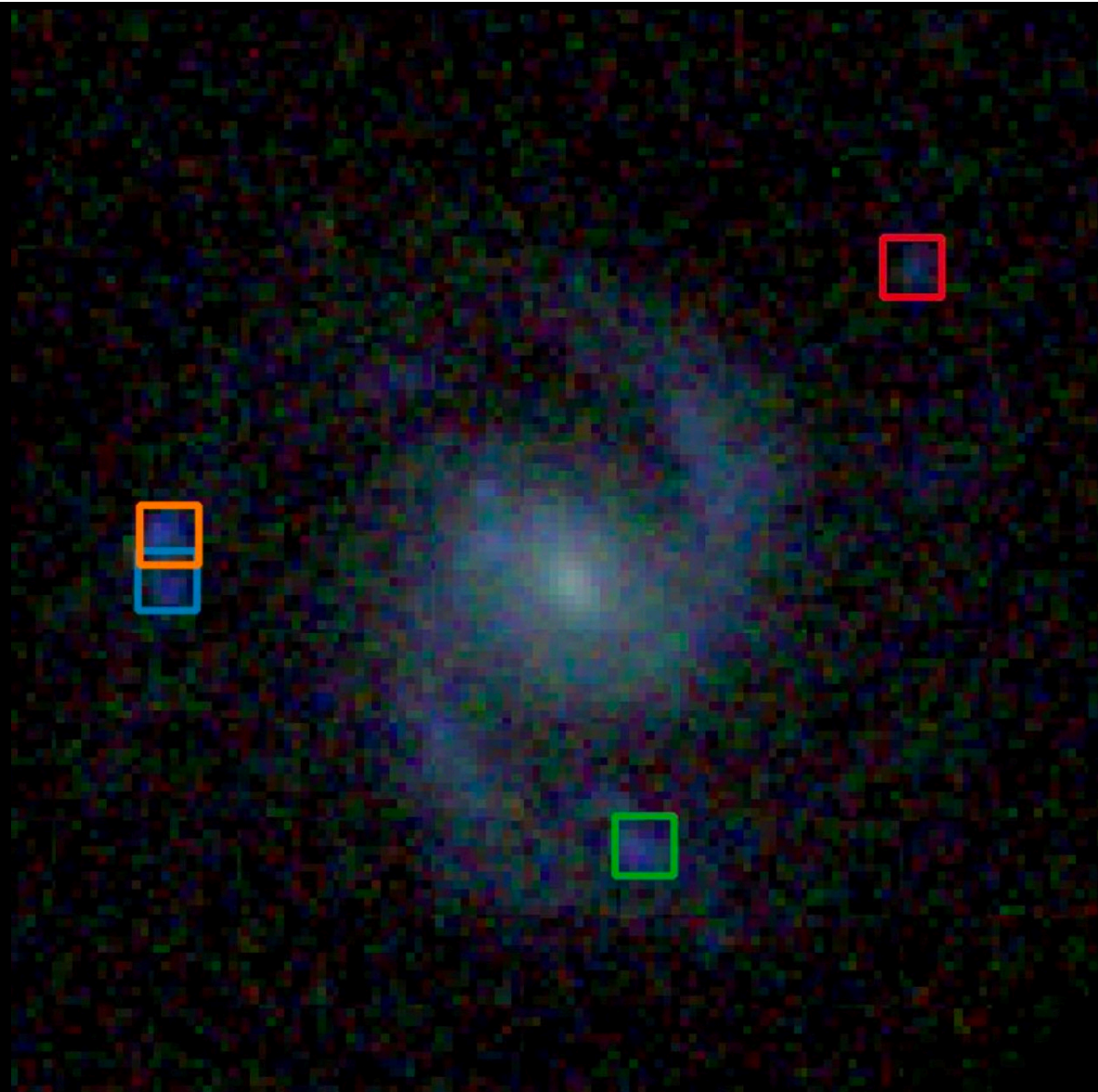
 Dickinson et al (2022) - arxiv.org/abs/2210.03684



Image credit: Hugh Dickinson





Number of clumpy galaxies:

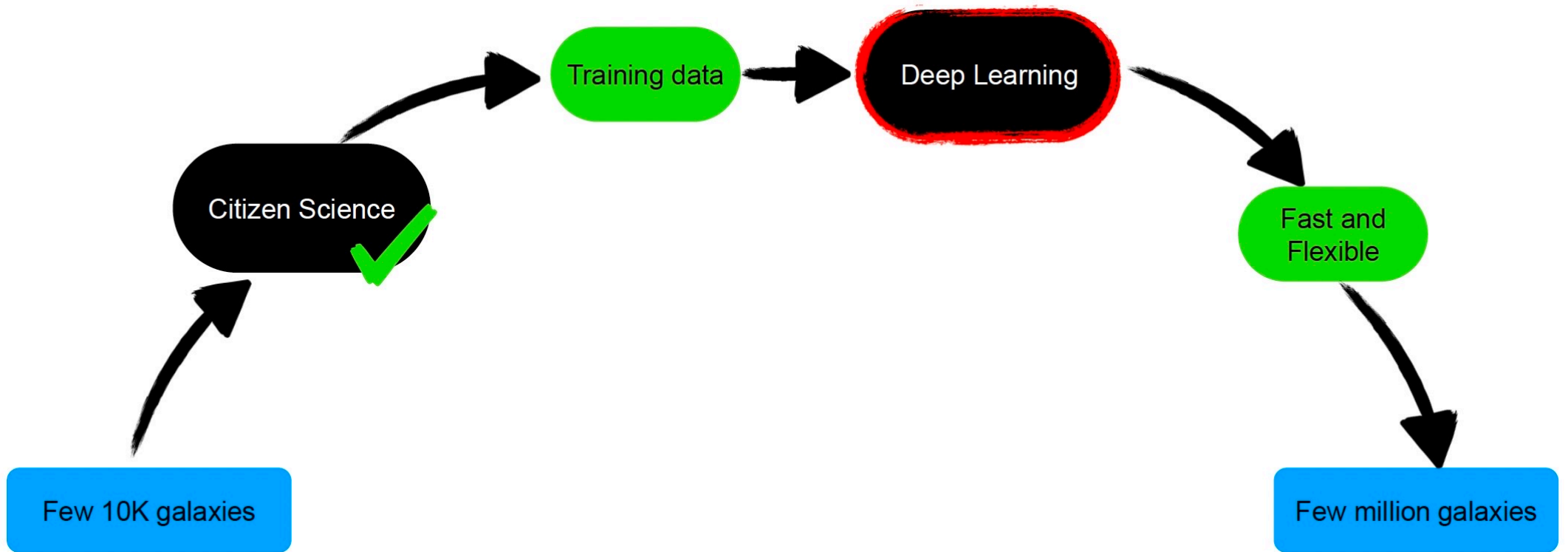
~35,000

Number of potential clumps:

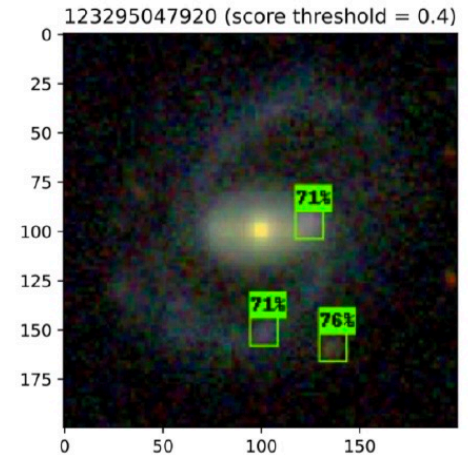
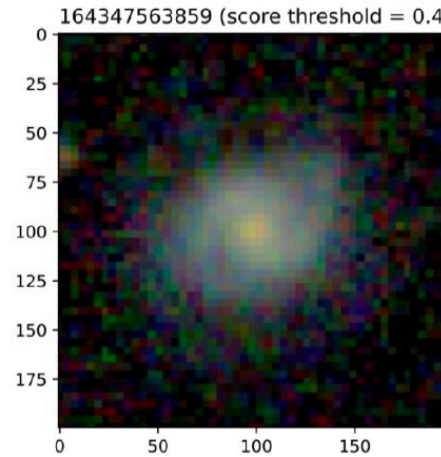
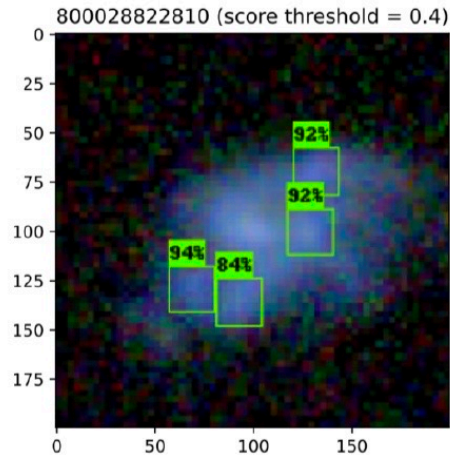
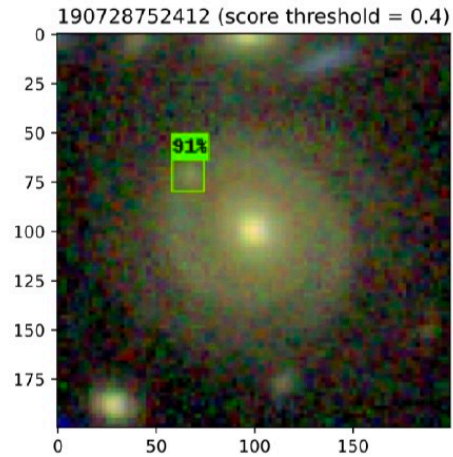
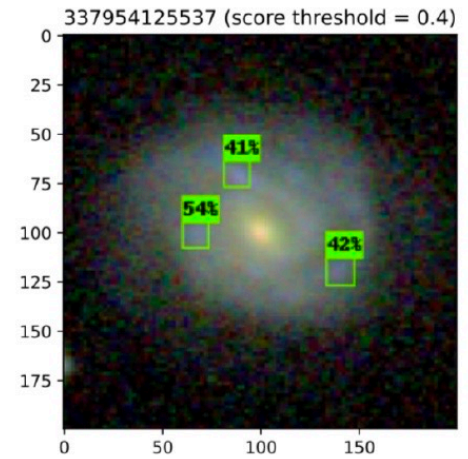
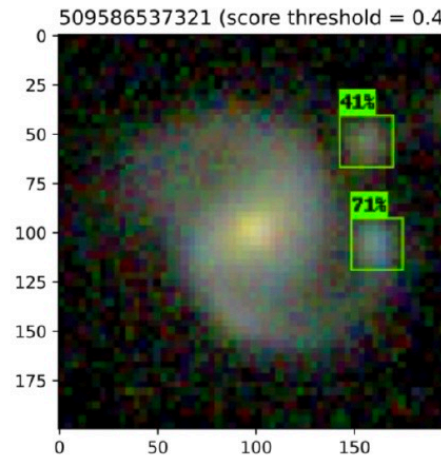
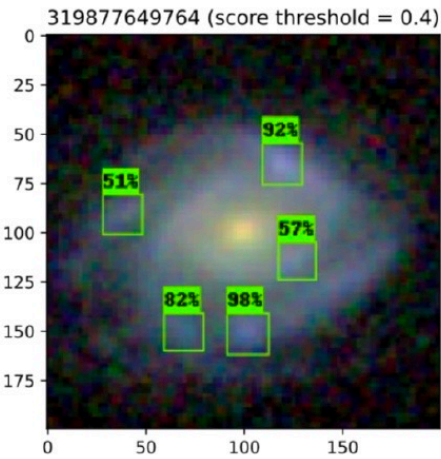
~100,000

**First catalogue
released!**

~~ar~~xiv Adams et al (2022) - arxiv.org/abs/2201.06581



Generic Deep Learning model finds **four times** as many clumps as Clump Scout volunteers!



Number of Volunteers engaged:

~15,000

Number of Galaxies inspected:

~80,000

Number of clumpy galaxies:

~35,000

Number of potential clumps:

~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



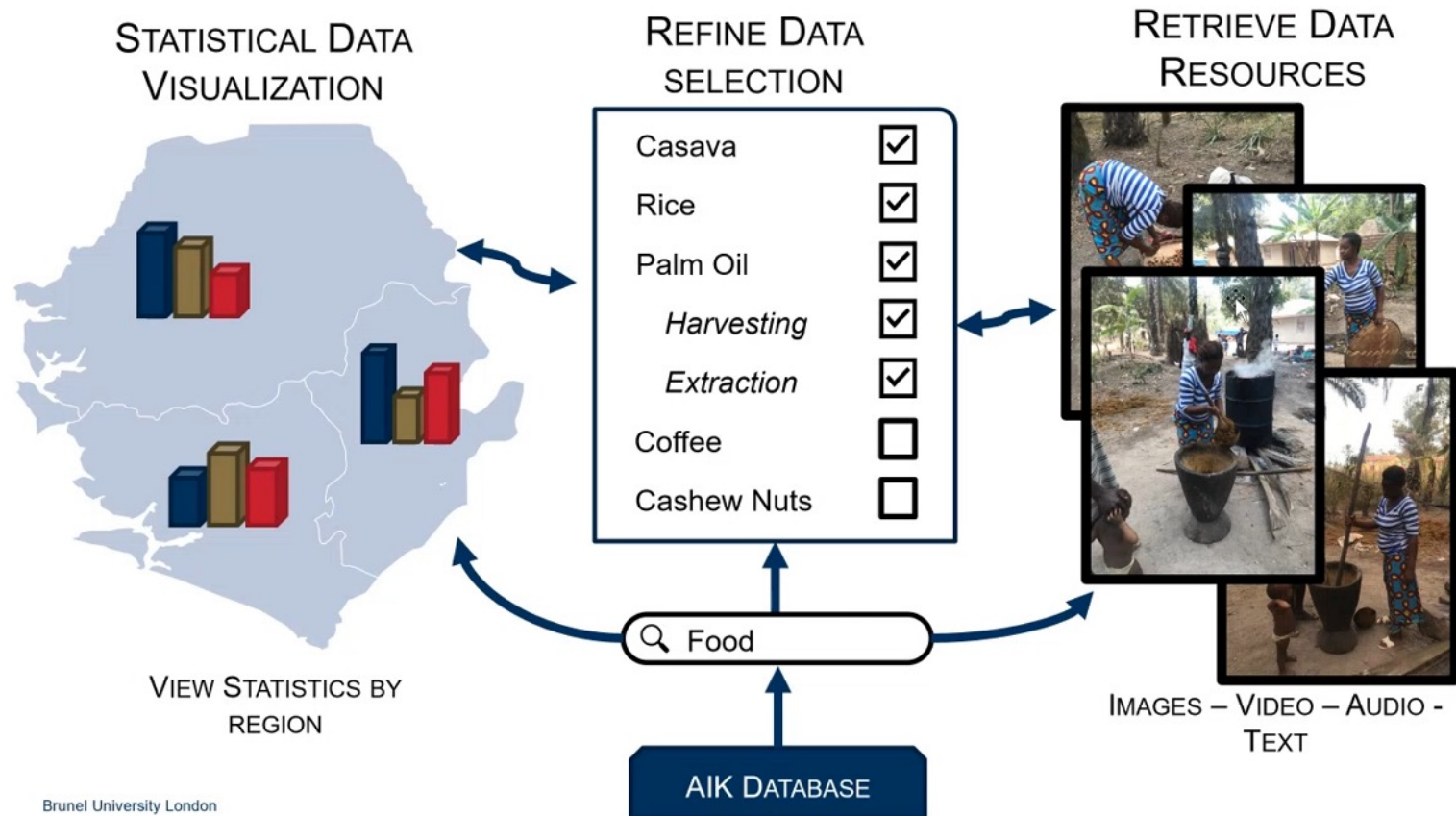
Knitting Patterns
in Domestic
Magazines

Image credits:
James Pearson



African Indigenous
Knowledge (AIK)
for Resilient Food
Systems

African Indigenous Knowledge



Knitting patterns

Completed item



Four Colour Favourite

Title

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

It is by passing the working wool over it, at start of each right-side row. 1st pattern row: K in W. 2nd row: P in W. 3rd row: K, * 2 R, 2 W; rep. from * to end of row. 4th row: P, * 2 W, 2 R; rep. from * to end of row. 5th row: K, in B. 6th row: K, in B. 7th row: K, in G. 8th row: K, in G. 9th row: K, in W. 10th row: P in W. 11th row: K, * 2 W, 2 R; rep. from * to end of row. 12th row: P, * 2 R, 2 W; rep. from * to end of row. 13th row: K, in B. 14th row: K, in B. 15th row: K, in G. 16th row: K, in G. Rep. these 16 rows twice more, then rep. 1st to 8th rows again. To shape armholes, cast off 4 sts. at start of the next 2 rows, then dec. 1 st. at each end of the following 4 rows, leaving 68 sts. Keeping continuity of pattern, work another 42 rows, thus completing 8th row of 7th pattern from start. To shape shoulders, cast off 7 sts. at start of next 6 rows. Cast off remainder for neck.

RIGHT FRONT

** With No. 11 needles and W wool, cast on 44 sts. and work in k. 1, p. 1 ribs for 1½ inches. Changing to No. 10 needles, work the 16 pattern rows of back twice, then work 1st to 4th rows again.** To commence front shaping, dec. 1 st. at start of next row and every 4th row after, until 40 sts. remain. Work 2 rows straight. To begin armhole shaping, cast off 4 sts. at start of next row. Dec. 1 st. at armhole edge on next 4 rows and dec. 1 st. at front edge on the 3rd of these rows. Work 2 rows straight, then dec. 1 st. at front edge on next row.*** Continue to dec. 1 st. at front edge on every 4th row until 21 sts. remain. Work 4 rows straight (or 3 rows when working left front). To shape shoulder, cast off 7 sts. at start of next row and the following alternate row, then work 1 row, after which cast off remainder.***

LEFT FRONT

Work as right front from ** to **. To start from shaping, dec. 1 st. at end of next row and every 4th row after until 40 sts. remain. Work 1 row straight. To begin armhole, cast off 4 sts. at start of next row. Dec. 1 st. at armhole edge on next 4 rows and dec. 1 st. at front edge on the 4th of these rows. Now proceed as right front from *** to ***.

THE SLEEVES (both alike)

With No. 11 needles and W wool, cast on 44 sts. and work in k. 1, p. 1 ribs for 1½ inches. Changing to No. 10 needles, work the 16 stripe pattern rows of back once, then keeping continuity of pattern, inc. 1 st. at each end of next row and every 4th row after until there are 58 sts. Work 11 rows straight. To shape top, dec. 1 st. at each end of the next 2 rows, then dec. 1 st. at start of every row until 18 sts. remain. Cast off.

TO COMPLETE

With No. 11 needles and W wool, cast on 10 sts. and work 4 rows in k. 1, p. 1 ribs. Place a buttonhole on next 2 rows thus: 5th row: Rib 4, cast off 2, rib 4. 4th row: Rib 4, cast on 2, rib 4. Work another 14 rows in ribs. Rep. the last 16 rows twice, then rep. 5th and 6th rows again. Then continue in ribs until strapping is long enough to extend along both from edges of cardigan and across back neck, noting that it should be slightly stretched on; then cast off. Press work. Join seams. Sew on strapping; add buttons.

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, 1½ 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

With No. 11 needles and W wool, cast on 84 sts. and work in k. 1, p. 1 ribs for 1½ inches. Changing to No. 10 needles, proceed in the following stripe pattern, joining on B wool at start of 3rd row, B wool at start of 5th row, and G wool at start of 7th row, and seeing that as soon as you discard one colour for another (at the end of a wrong-side row), you need not break off the discarded colour; simply carry it up side of work to the next position required, and to avoid it thus forming loose loops at side, catch

pink smudged of damp earth, of lime blossoms and crushed rose leaves. The whole world was washed in the pure rain-cooled air, the deep golden light. Noddin's Claire was so happy she didn't know what to do about it. "I like my women fat," Alan was saying cheerily. She turned, laughing. "I'm not 'your women,'" she said. He put his hands on her shoulders, drawing her close. "You are. You're all my women, every one of them. All I'll ever want." When his hat came down on her nose she stayed very still for a moment, then lifting her arms she bowed them about her neck.

In the small intimate restaurant where they ate, they talked of Niewekid; it was Claire who introduced the subject, telling Marjorie how, after all, it had been odd news. "I give the servants very kept on," she said. "I was awfully glad about that. I'd been worrying about them." "Like to run out and have a look at the old place some time?" Alan offered. And when Claire shook her head, he gave her a quick, speculative glance. "Is it still, then, most important to you than anything else in the world?" he asked. "Has it meant everything to you, being it?" "Niewekid? Oh, no." Her grey eyes were very honest. "I hated giving it up, of

course. It hurt dreadfully. But there were other things that hurt me more." The last words were very softly spoken and a quick colour ran up into her thin cheeks. "Do-construct I'd made about myself; the feeling that I'd muddled my life, misunderstood so much. And then..." She covered her eyes with her hand. "For a time I went through a horrid period of bitterness against Zena. But in the end, even before the crash came, living there alone at Niewekid I began to find out how empty and unsatisfying it can be to give one's heart to a house." She looked at him across the table and the

Please turn to page 20.

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

Materials

› **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?**



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



60 SECOND
ADVENTURES IN
ARTIFICIAL INTELLIGENCE