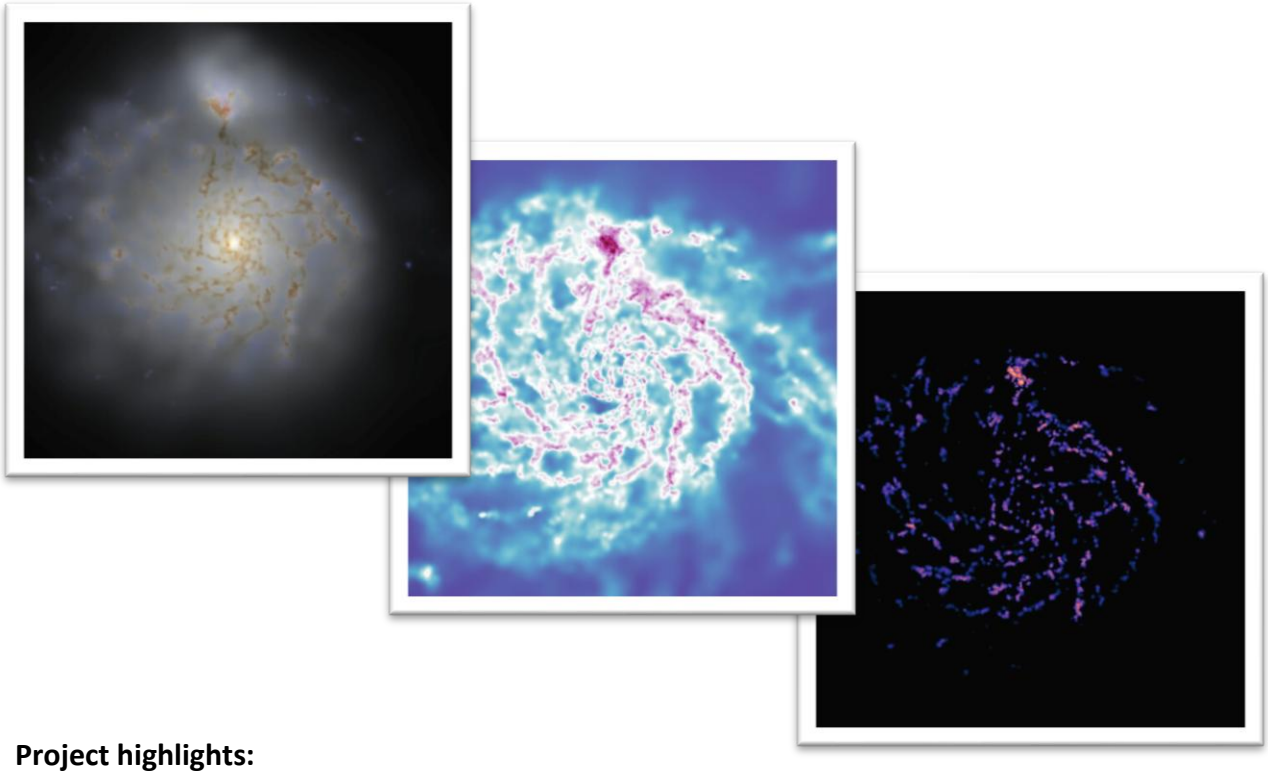
Investigating Giant Star-forming Clumps in simulations and the real Universe

Supervision team: Hugh Dickinson, Stephen Serjeant, Rob Crain (Liverpool John Moores University)

Lead supervisor: Hugh Dickinson, hugh.dickinson@open.ac.uk



Project highlights:

- This project investigates a mysterious class of objects called giant star-forming clumps (GSFCs). These are huge, kiloparsec-scale star forming regions that are very common in high redshift galaxies but much rarer in the local Universe.
- GSFCs were common when star formation in the Universe was at its most intense and they played a crucial but poorly understood role in building the most massive galaxies that exist today. Your work will help to understand how clumps form, how they evolve and how they affect their host galaxies' evolution. Answering these questions is an outstanding challenge for modern observational cosmology.
- You will analyse brand new data from the *Euclid* Space Telescope and use data from the cutting-edge COLIBRE simulation to compare the substructure of simulated and real GSFCs to unveil the physical processes operating within them.
- Your project will involve developing and applying sophisticated Deep Learning models to petabyte-scale datasets, using the Zooniverse citizen science platform to crowdsource labels for model training and using high-performance computing clusters to analyse the output of cutting-edge cosmological simulations.
- We are now at the start of a new era of Big Data in astronomy and cosmology. The work you do will pioneer the use of modern tools and techniques that will be essential to fully realise the scientific potential that new telescopes, surveys and simulations will provide.

Project description:

Giant star-forming clumps (GSFCs) are discrete regions of enhanced star formation within galaxies with apparent radii ~ 1 kpc and stellar masses $M \gtrsim 10^5$ $M_\odot$ [1]. They are a very common feature of massive star-forming galaxies at $z \sim 2$ but very rare in the local Universe, where discrete star-forming structures within galaxies are typically much smaller and less massive [2]. Understanding how, when and why clumps became so rare is an outstanding challenge for modern observational cosmology. The epoch around $z = 2$ is often referred to as the cosmic "noon" because it corresponds to the most intense period of star forming activity in the history of the Universe. This means that the GSFC formation represents a crucial stage in the evolution of the most massive galaxies that exist in the present-day Universe, including our own Milky Way.

The overarching goal of this project is to compare the detailed substructure of GSFCs that are observed by the *Euclid* Space Telescope with the predictions made by the next-generation COLIBRE cosmological simulation (https://colibre.strw.leidenuniv.nl). Doing so will allow the physical processes driving the formation and evolution of huge populations of GSFCs to be inferred while also revealing the effects of GSFCs on their host galaxies and the surrounding intergalactic medium. This project applies to GSFCs in general but will focus primarily on understanding the characteristics of low-redshift ($z \lesssim 0.05$) GSFCs because studying these nearby GSFCs allows their substructure to be studied on small spatial scales. Until very recently the rarity of GSFCs at low redshift has made studying them very challenging. Fortunately, the arrival of next-generation survey instruments like *Euclid* means we now have high-quality imaging for ~250 million galaxies that we can search for GSFCs.

To detect and analyse these GSFCs, the student will adapt a deep learning model that we have developed at the Open University [3]. The existing model can detect GSFCs in galaxies using lower resolution images from the Subaru Hyper Suprime-Cam, but it must be fine-tuned so that it can be applied to *Euclid* images. Labels for this fine-tuning step will be collected using a citizen science project building on our successful *Galaxy Zoo: Clump Scout* project (https://tinyurl.com/clump-scout). By applying their fine-tuned model to the full *Euclid* imaging dataset, the student will assemble and analyze the largest ever catalogue of GSFCs at low redshift with associated high-resolution imaging.

This catalogue will allow the student to undertake the first detailed comparison between the observed substructure of low-redshift GSFCs and the predictions of a high-resolution hydrodynamical simulation of GSFCs in a cosmological context. Comparing observations and simulations in this way will allow the realism of the simulations to be evaluated while also providing crucial insights into the physical processes driving the formation and evolution of GSFCs in the real Universe.

The project will be supported by Professor Rob Crain at Liverpool John Moores University (LJMU) who is a world-renowned expert in the processing and analysis of cosmological simulation data. Computational support needed for the analysis of COLIBRE data will be available via the *Prospero* high-performance computing facility at LJMU.

The work undertaken during this project will continue the Open University's pioneering use of deep learning to analyse one of the first genuinely Big astrophysical datasets. The tools developed during this project will be essential to fully realise the scientific potential of the petabyte-scale datasets that *Euclid* and other next-generation facilities will provide.

References:

1. Elmegreen 2007, IAU Symposium, Vol. 235, pp. 376–380.
2. Guo et al. 2015, ApJ, Vol. 800, p. 39
3. Popp et al. 2024, RASTI, Vol. 3, p174

Qualifications required: Minimum BSc 2:1 in a relevant discipline; recommended MPhys 2:1 or above.