

ML-based, explorative web-service

Triangular

Inequality: with $\phi_a = \operatorname{argmin}_{\phi \in \Phi} (d(A, \phi(C)))$ and $\phi_b = \operatorname{argmin}_{\phi \in \Phi} (d(C, \phi(B)))$

how to deal with complex morphologies

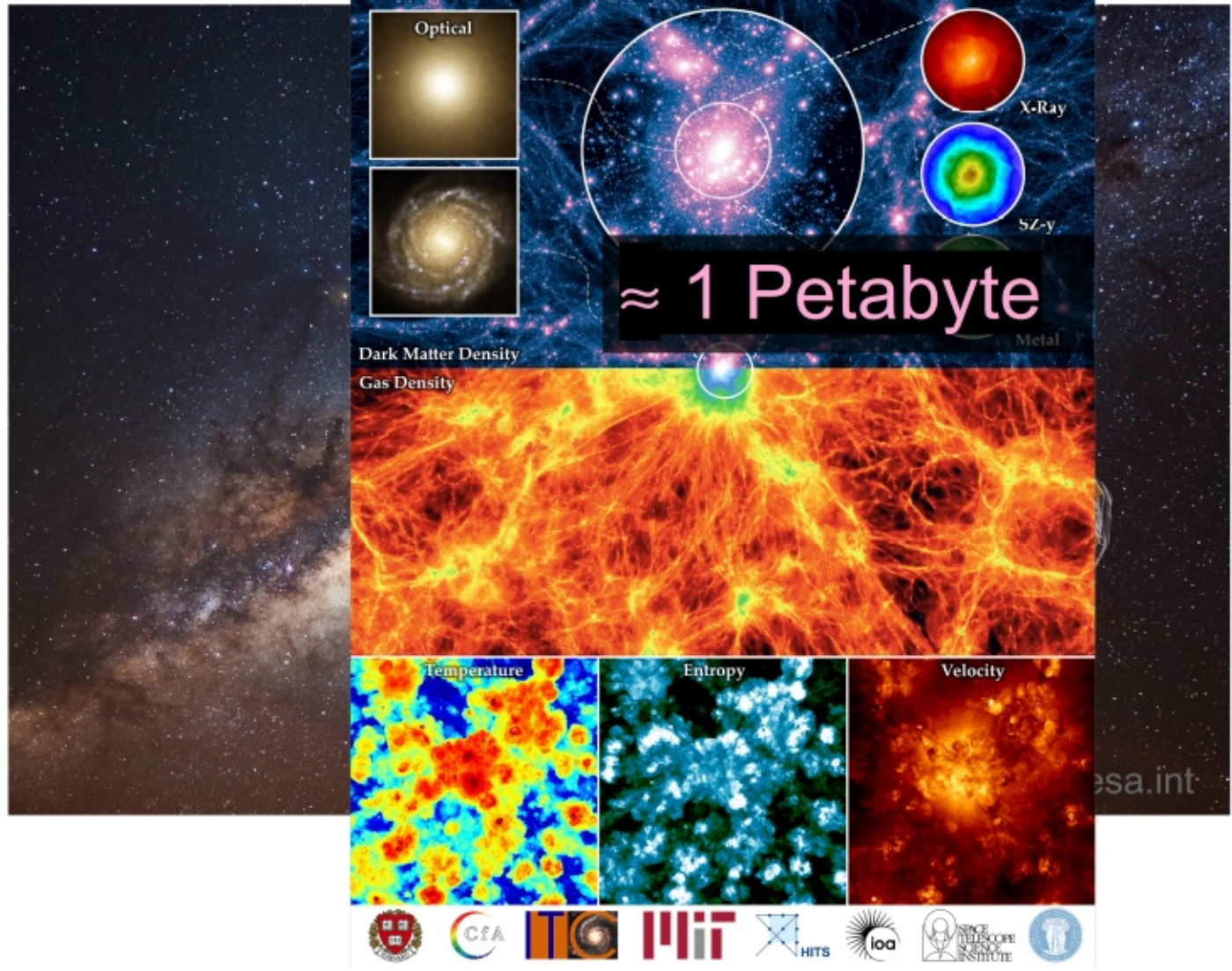
Complexity / Size of Data



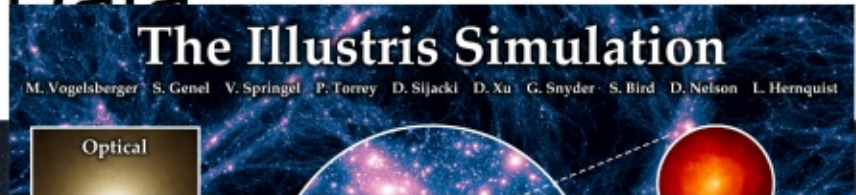
Gaia \approx 1 Petabyte



Complexity / Size of Data



Complexity / Size of Data



Manual Visual Inspection



200.000 stellar spectra

- “Pickering’s Computers” → Annie Jump Cannon



Smithsonian Institution -
Annie Jump Cannon (1863-1941), sitting at desk

Manual Visual Inspection

200.000 stellar spectra

- “Pickering’s Computers” → Annie Jump Cannon

50.000.000 images of galaxies

- Galaxy Zoo

Manual Visual Inspection



The screenshot shows the homepage of the Galaxy Zoo website. At the top, there is a navigation bar with the following links: CLASSIFY, STORY, SCIENCE, GALAXY ZOO (in a yellow banner), DISCUSS, PROFILE, and LANGUAGE. Below the navigation bar, there are social media icons for Facebook, Twitter, Google+, and RSS. The main heading reads "Few have witnessed what you're about to see" with a sub-heading "Experience a privileged glimpse of the distant universe as observed by the SDSS, the Hubble Space Telescope, and UKIRT". Below this, there is a large image of a spiral galaxy. To the left of the galaxy image, there is a section titled "Classify Galaxies" with a yellow button that says "Begin Classifying". Below the galaxy image, there are two columns of text. The left column is titled "How Do Galaxies Form?" and the right column is titled "History of Galaxy Zoo".

www.galaxyzoo.org/#/

hubble tuning fork

CLASSIFY STORY SCIENCE **GALAXY ZOO** DISCUSS PROFILE LANGUAGE

f t g+ rss

Few have witnessed what you're about to see

Experience a privileged glimpse of the distant universe as observed by the SDSS, the Hubble Space Telescope, and UKIRT

Classify Galaxies

To understand how galaxies formed we need your help to classify them according to their shapes. If you're quick, you may even be the first person to see the galaxies you're asked to classify.

[Begin Classifying](#)

How Do Galaxies Form?

Roughly one hundred billion galaxies are scattered throughout our observable Universe, each a glorious system that might contain billions of stars. Many are remarkably beautiful, and the aim of Galaxy Zoo is to study them, assisting astronomers in attempting to understand how the galaxies we see around us formed, and what their stories can tell us about the past, present and future of our Universe as a whole. [MORE](#)

History of Galaxy Zoo

The launch of this new version of Galaxy Zoo, the 4th, comes just a few weeks after the site's 5th birthday. It all started back in July 2007, with a data set made up of a million galaxies imaged by the Sloan Digital Sky Survey, who still provide some of the images in the site today. With so many galaxies, we'd assumed it would take years for visitors to the site to work through them all, but within 24 hours of launch we were stunned to be

Manual Visual Inspection



A screenshot of a web browser displaying the Zooniverse website. The browser's address bar shows 'https://www.zooniverse.org'. The page content is a grid of science-related articles. A large pink arrow points to the article titled 'Match growing black holes to their birthplaces'. The grid includes articles such as 'Wie entstehen Galaxien?', 'Entdecke die Oberfläche des Mondes', 'Untersuche Explosionen auf der Sonne', 'Finde Planeten um andere Sterne', 'Wie entstehen Sterne?', 'Explore the Red Planet', 'Match growing black holes to their birthplaces', 'the Birthplace of...', 'Sorting out Sunspots', and 'Help us discover near-Earth asteroids'. Each article features a small image and a brief description of the project.

Manual Visual Inspection



www.galaxyzoo.org/#/

hubble tuning fork

https://www.zooniverse.org

zooniverse

https://www.zooniverse.org

zooniverse

radio.galaxyzoo.org

radio galaxy zoo

polsterer SIGN OUT English

CLASSIFY SCIENCE TEAM PROFILE TALK BLOG

GALAXY ZOO RADIO

In Search of Erupting Black Holes

Help astronomers discover supermassive black holes observed by the KG Jansky Very Large Array (NRAO) and the Australia Telescope Compact Array (CSIRO)

Search for Black Holes

Black holes are found at the center of most, if not all, galaxies. The bigger the galaxy, the bigger the black hole and the more sensational the effect it can have on the host galaxy. These supermassive black holes drag in nearby material, growing to billions of times the mass of our sun and occasionally producing spectacular jets of material traveling nearly as fast as the speed of light. These jets often can't be detected in visible light, but are seen using radio telescopes. Astronomers need your help to find these jets and match them to the galaxy that hosts them.

[Begin Hunting](#)

NASA, ESA, S. Baum and C. O'Dea (RIT), R. Perley and W. Cotton (NRAO/AUI/NSF), and the Hubble Heritage Team (STScI/AURA)

Radio Galaxy Zoo



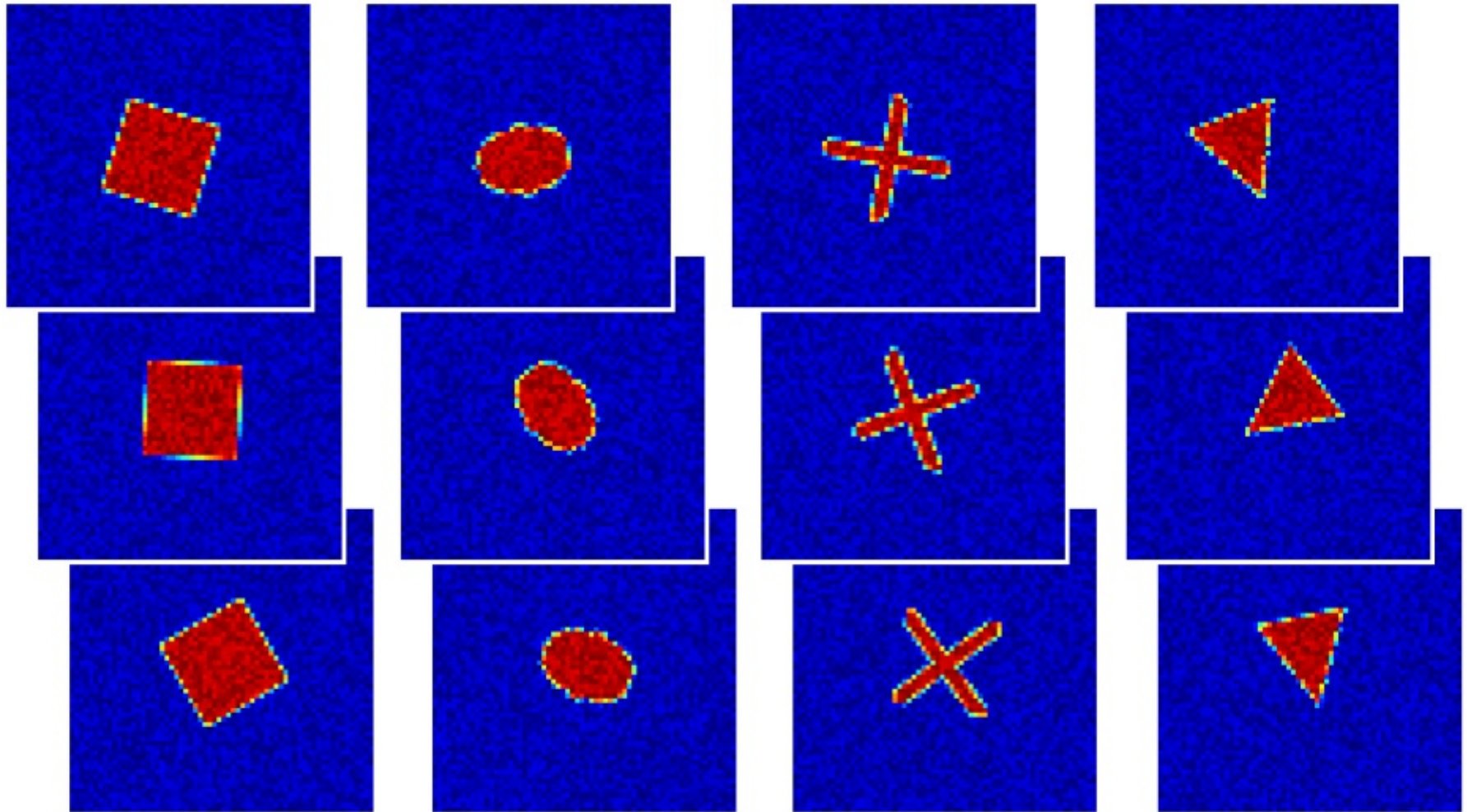
The screenshot shows the Radio Galaxy Zoo web interface. At the top, there is a navigation menu with links for CLASSIFY, SCIENCE, TEAM, PROFILE, TALK, and BLOG. The central logo reads "GALAXY ZOO" in blue and "RADIO" in yellow. In the top right corner, the user "polsterer" is logged out, and the language is set to "English".

The main content area displays a radio galaxy image with white contours overlaid on a blue background. To the right of the image are three circular icons: a crossed-out "ADD" button, a "VIEW" button, and a "GRID" button. Below the image is a slider control labeled "Radio" on the left and "IR" on the right. A text prompt below the slider says "Click on any radio contour or pair of jets". At the bottom, there are four buttons: "Cancel" (red), "Reset All" (red), "No Contours" (blue), and "Done" (blue).

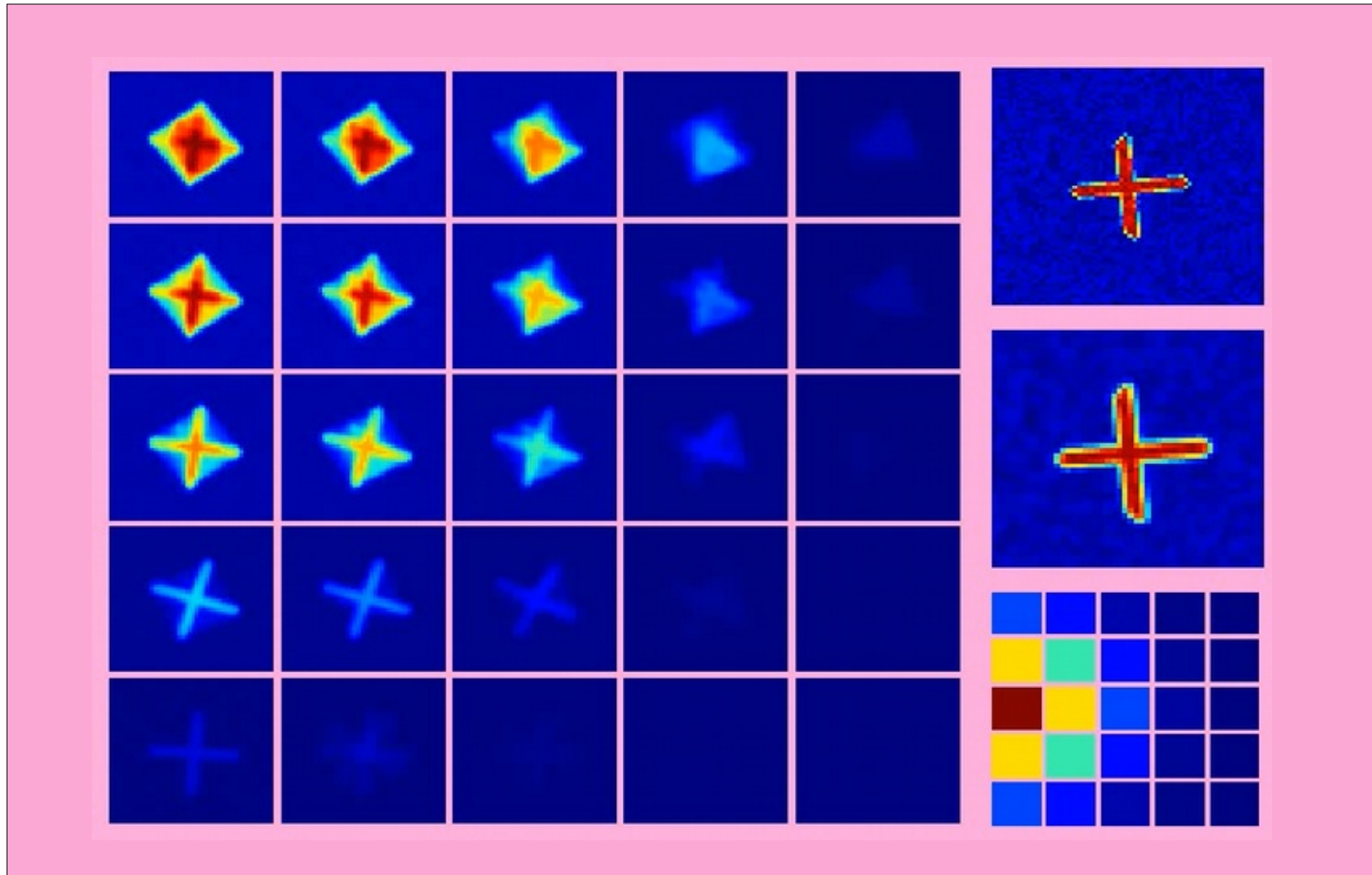
Human Pattern Recognition



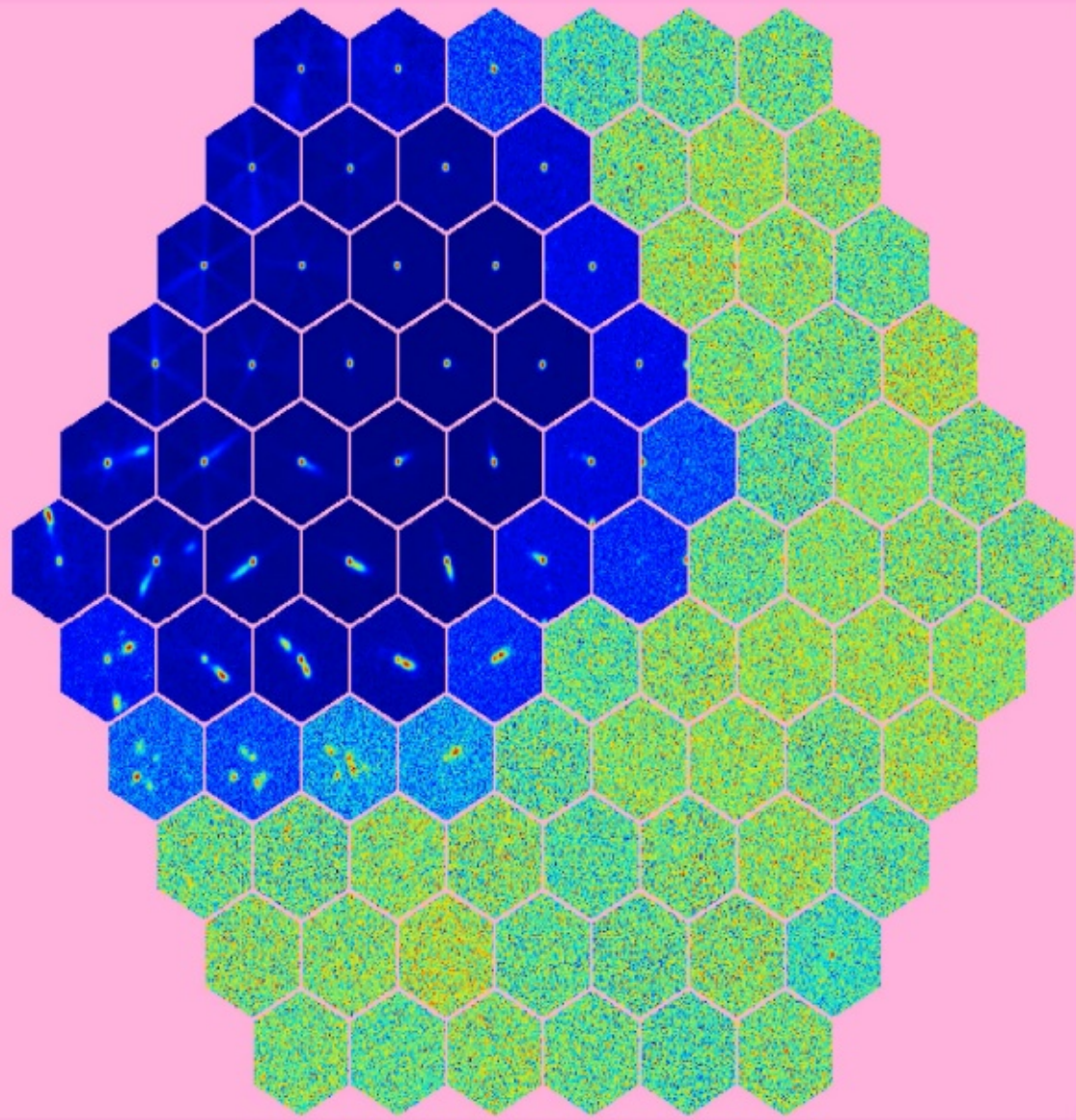
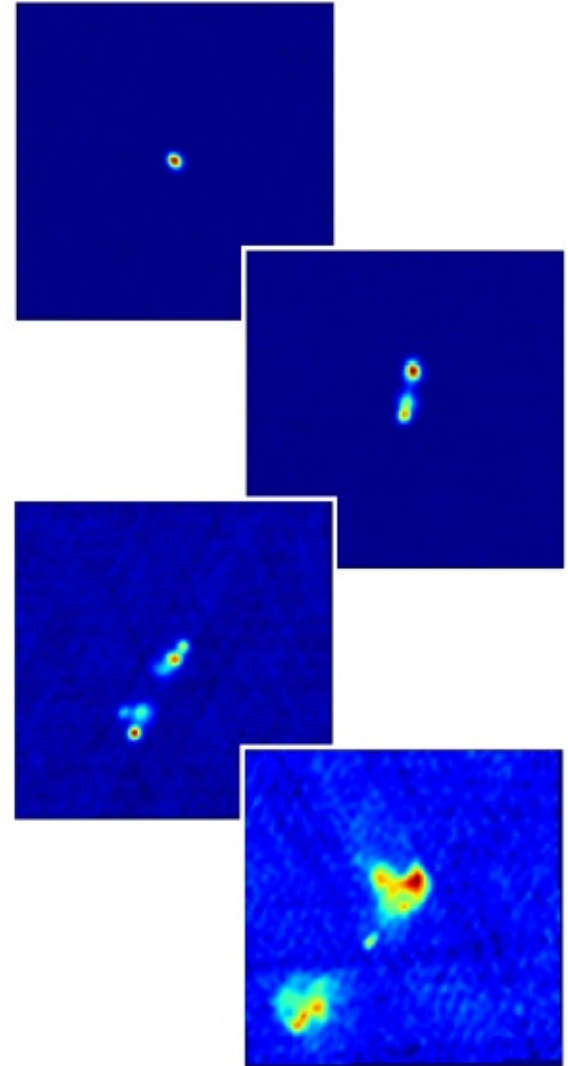
Dimensionality Reduction



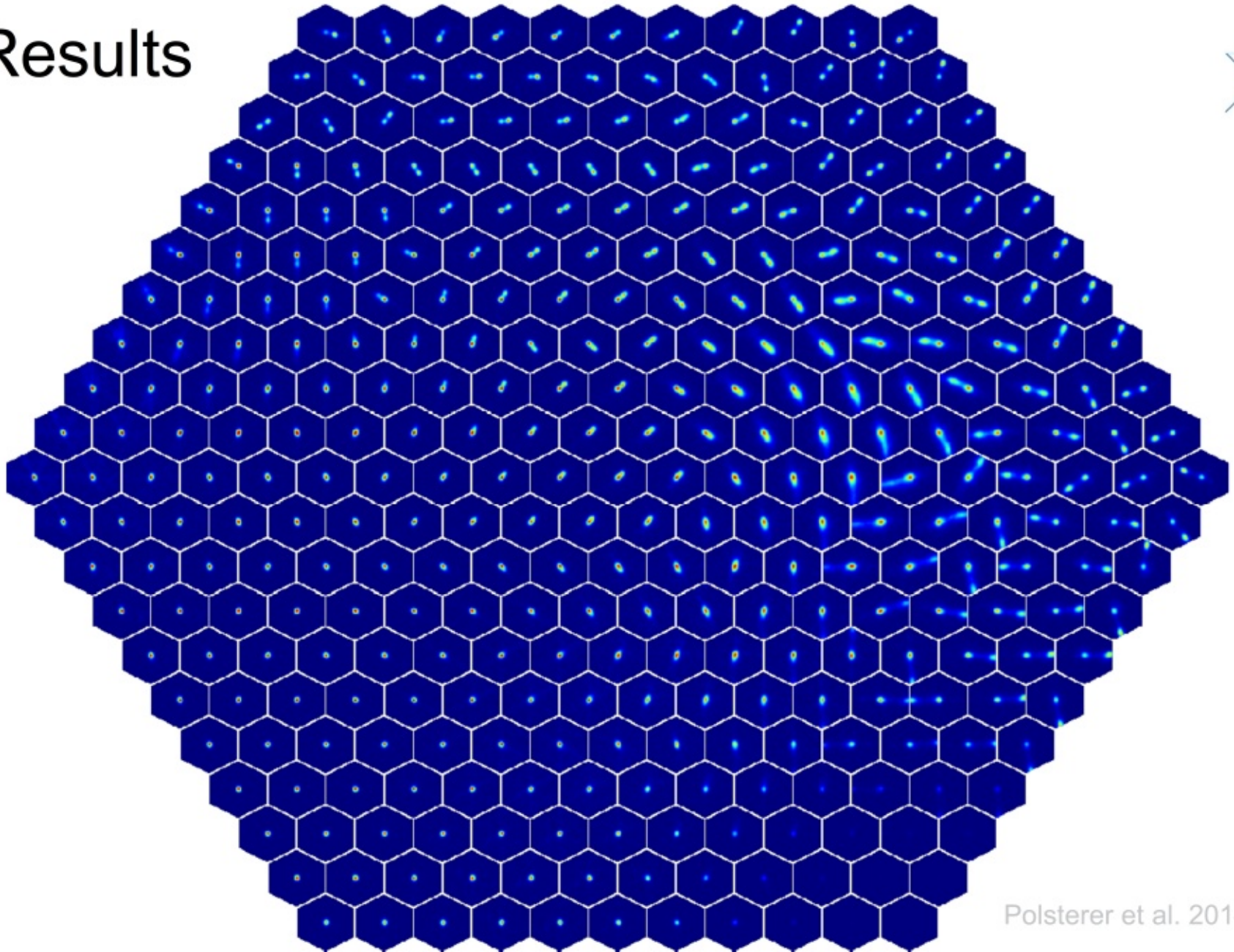
Self-organizing Maps / Kohonen Maps



PINK

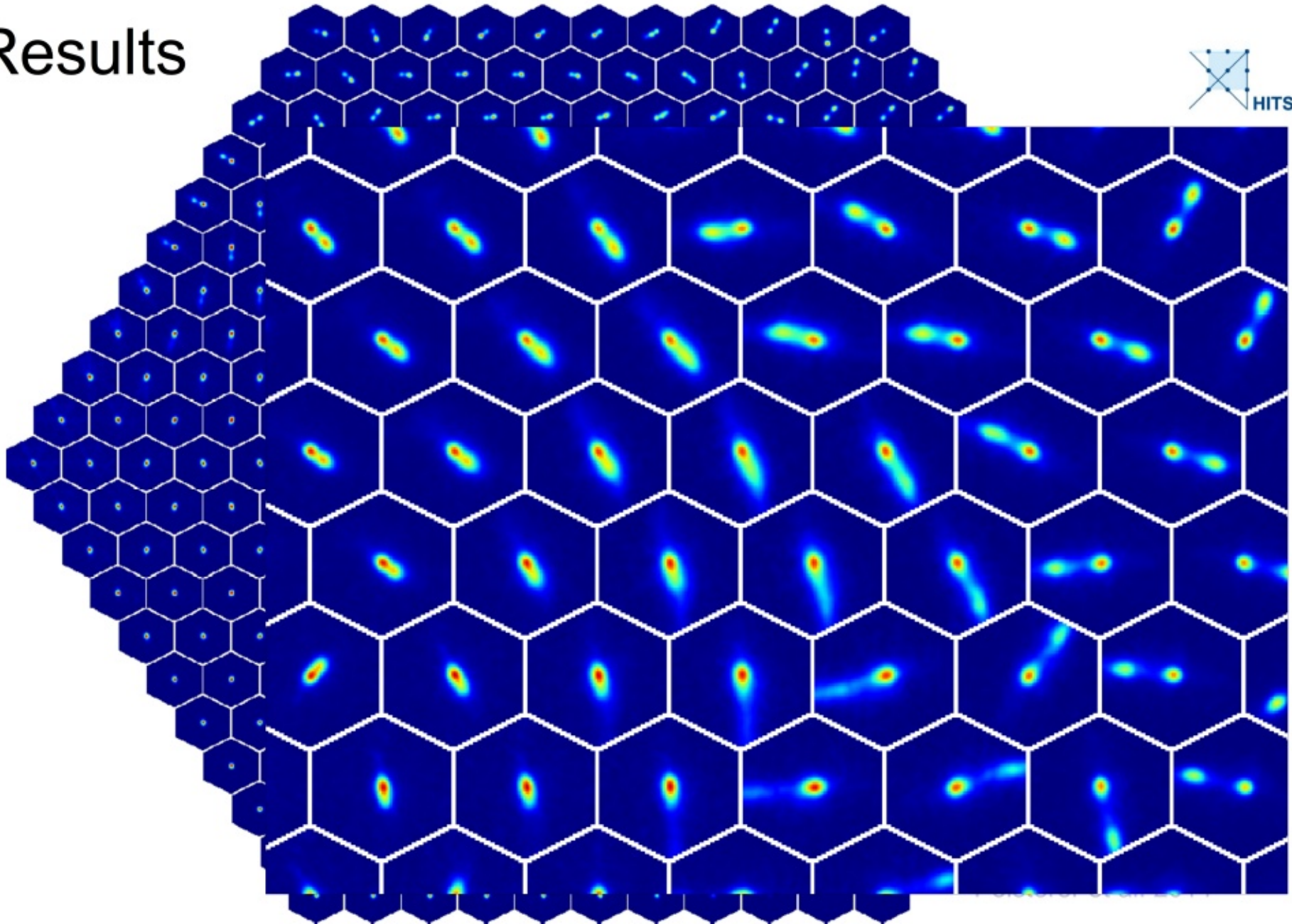


Results

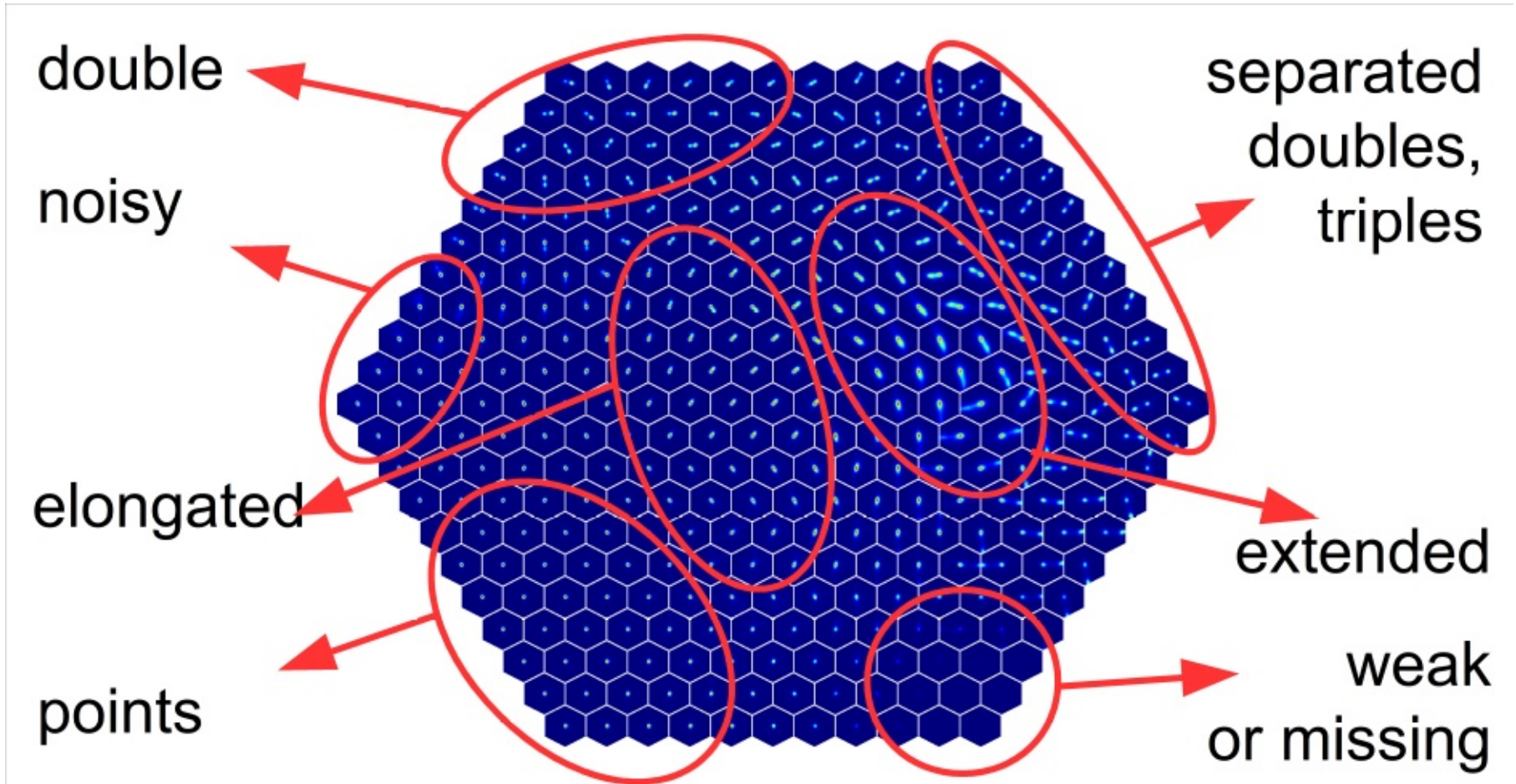


Polsterer et al. 2014

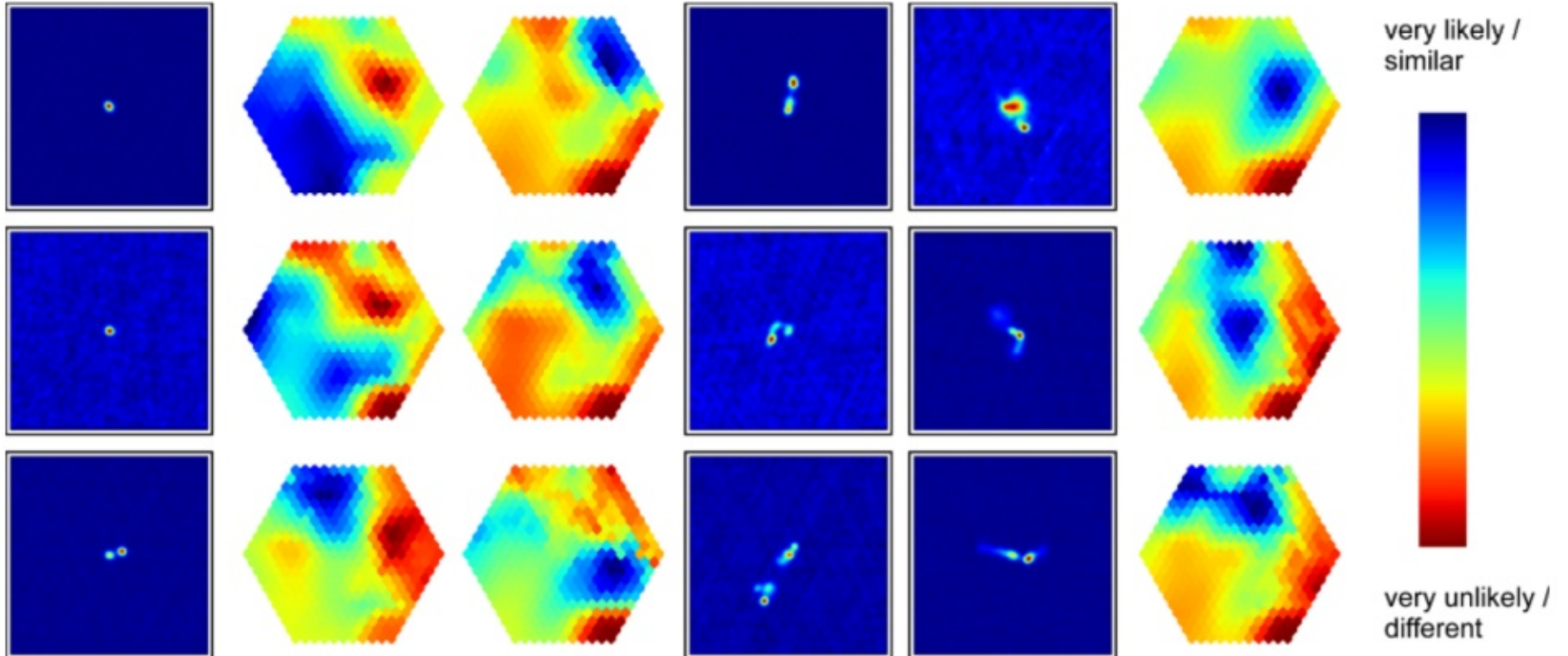
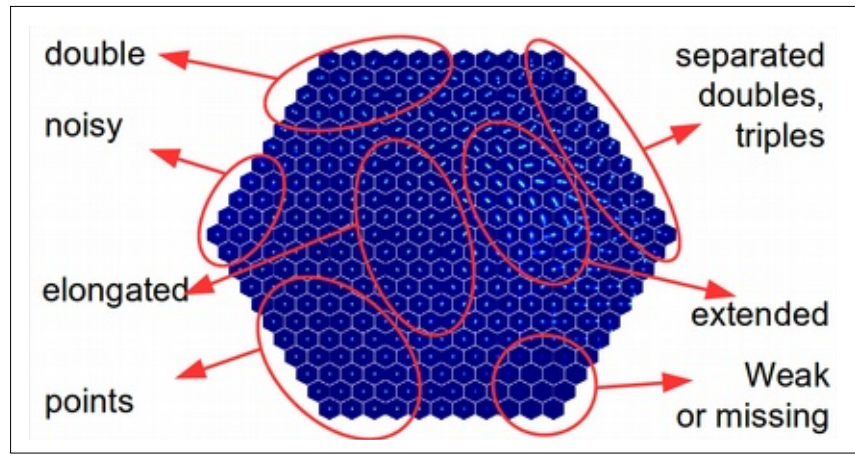
Results



Results / Analysis / Annotation



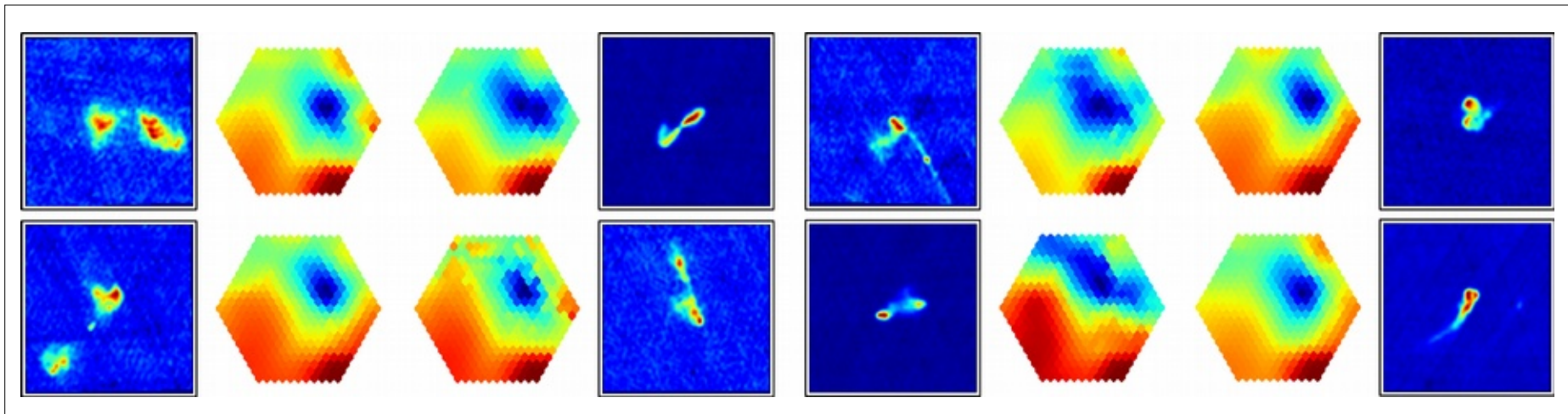
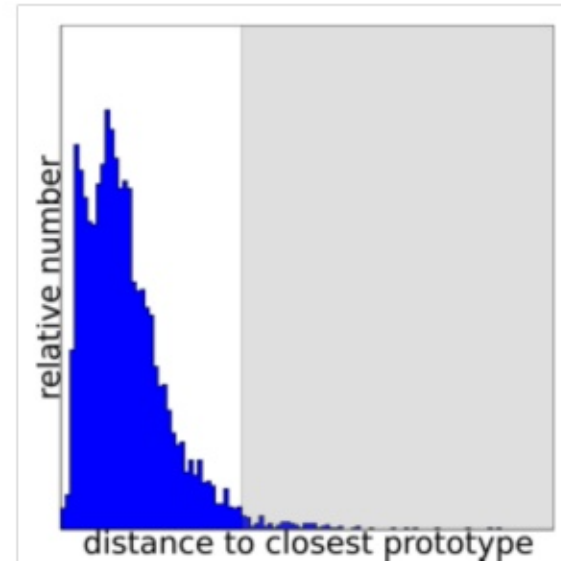
Transfer

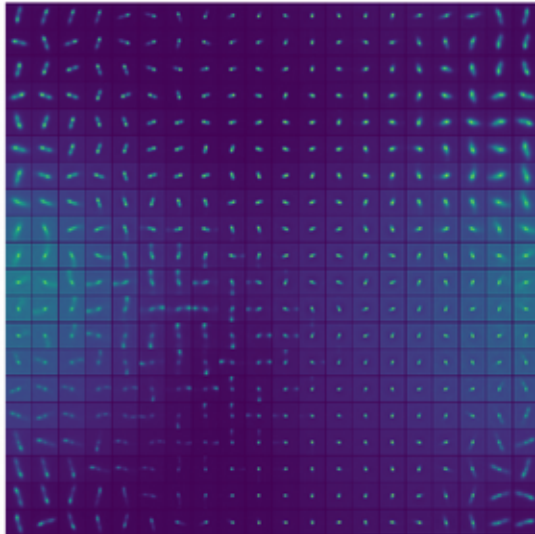


Rare Objects



find outliers based on similarity measures





Show heatmap SOM properties

This is a Self-Organizing Map, trained on sources from the LOFAR survey. Click on one of these prototypes.

About this project

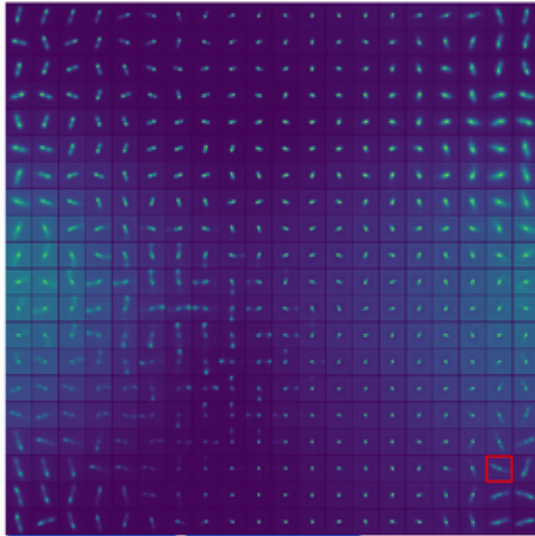
From the shape or morphology of a radio source we can infer physical properties of the source and its environment.

To find out what different morphologies are present in the LOFAR survey, we use a dimensionality reduction technique known as a *Self-Organizing Map*.

This is an unsupervised neural network that projects a high-dimensional dataset to a discrete 2-dimensional representation.

The map contains 20x20 neurons or prototypes, each represents a group of sources.

Demo

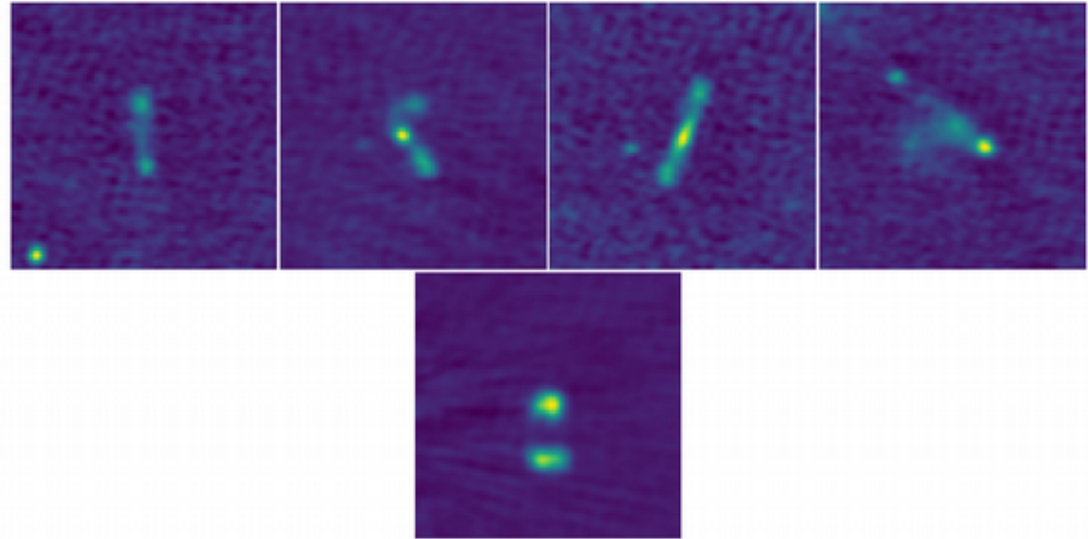


Show heatmap SOM properties



This is a Self-Organizing Map, trained on sources from the LOFAR survey. Click on one of these prototypes.

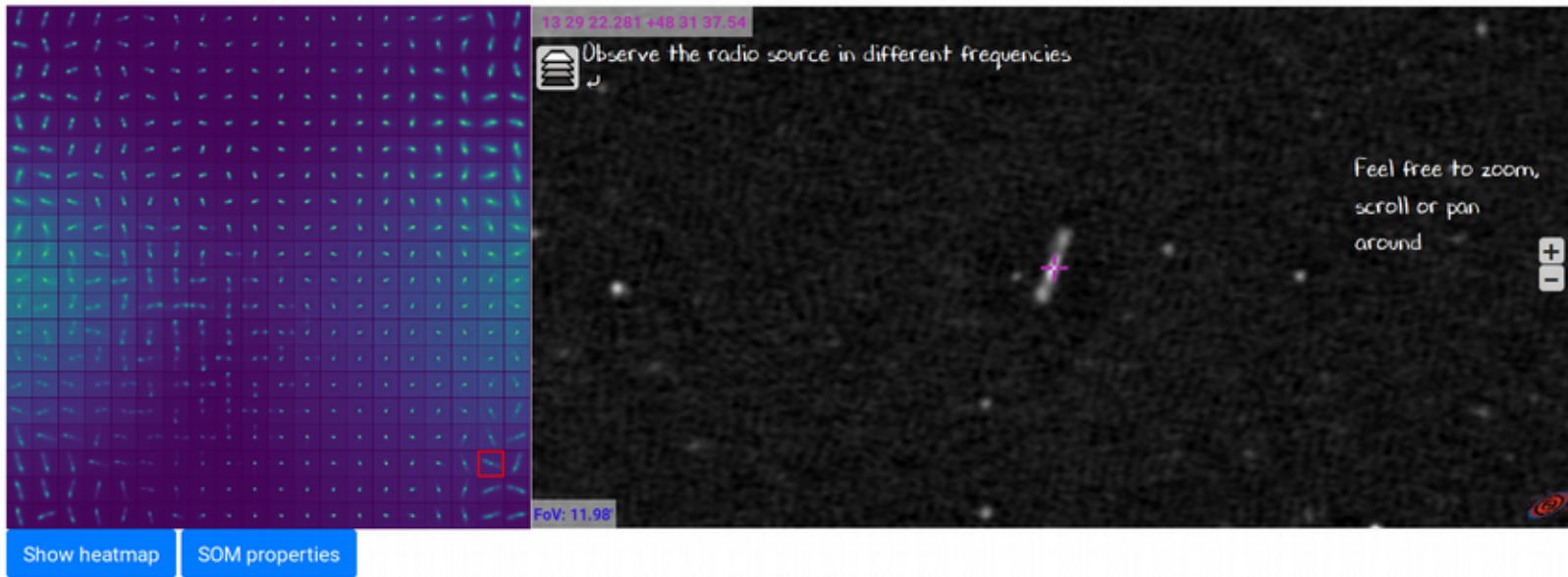
Radio sources from LOFAR survey that resemble the selected prototype (18,17):



Here are 5 of the radio sources that best resemble the prototype you just selected.

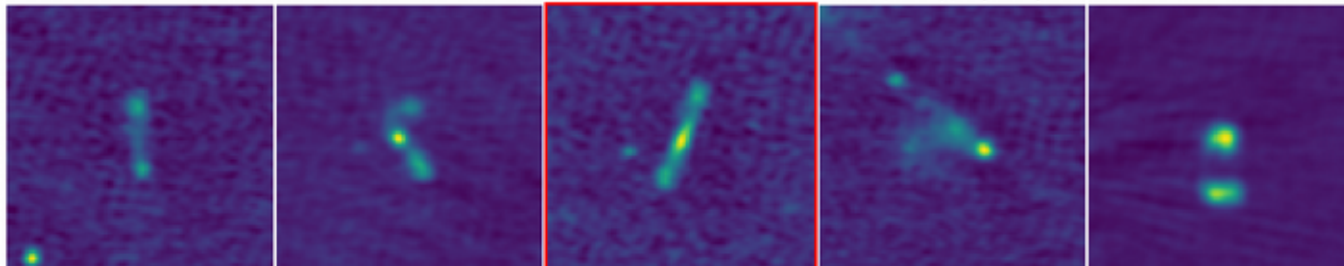
Click on a source to view it in the sky.

Demo



On the left you can see where the radio source you clicked on is located on the sky. The source might be accompanied or interacting with other sources or

Radio sources from LOFAR survey that resemble the selected prototype (18,17):



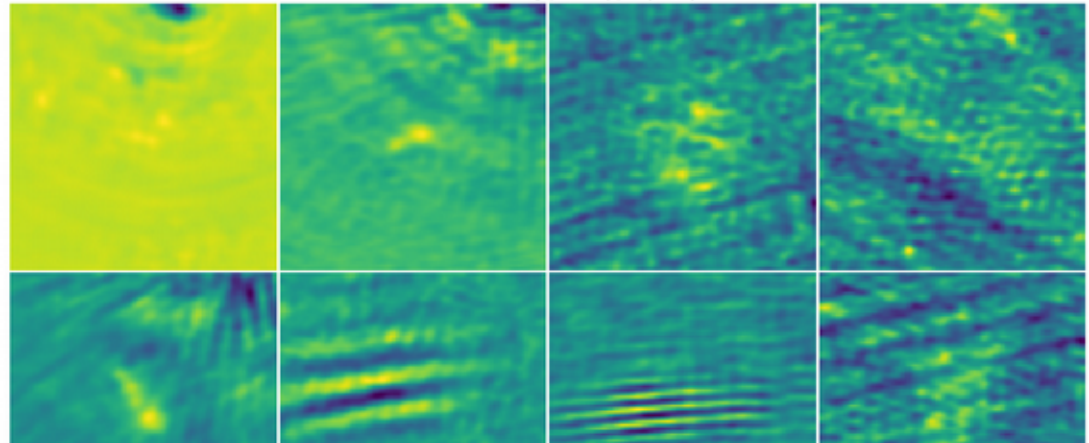
100 morphologically rarest sources

The Self-Organizing Map is a condensed representation of the most occurrent morphologies present in our dataset.

If a source barely resembles any of the prototypes in the Self-Organizing Map, it is thus a morphological outlier.

Using this heuristic, we show the 100 most morphologically unique radio sources below:

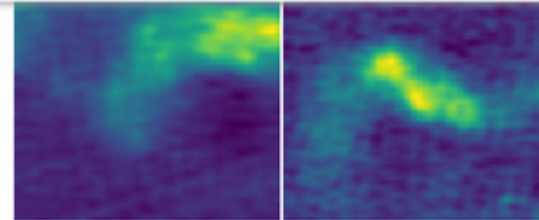
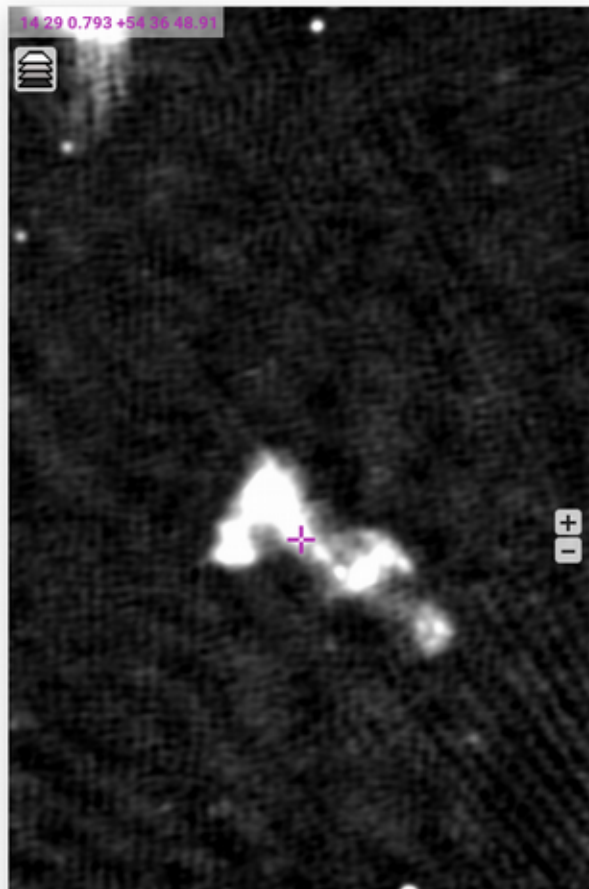
Best resembling prototype (10,0):



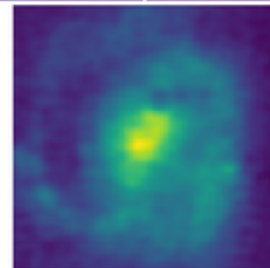
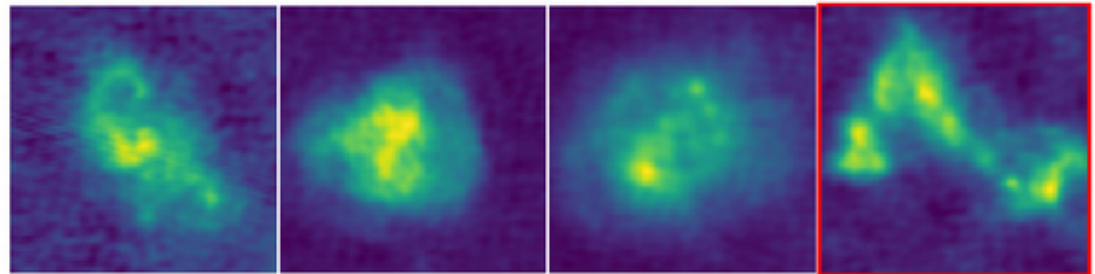
Click on one of the outliers to show them in context.



Demo



Best resembling prototype (8,0):

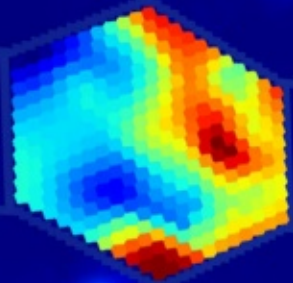


Best resembling prototype (8,19):



Open Questions

- How to store annotations?
- How to describe projections?
- How to transfer/preserve/publish annotations?
- What about semantics?



Symmetry: $\Delta(A, B) = \min\{d(A, \phi(B)) \mid \phi \in \Phi\}$

$$= \min\{d(\phi(B), A) \mid \phi \in \Phi\}$$

$$= \min\{d(B, \phi^{-1}(A)) \mid \phi \in \Phi\}$$

$$\geq \Delta(B, A)$$

The same way we get $\Delta(B, A) \geq \Delta(A, B)$



Thanks for your attention!



@Astroinformatix / @HITStudies



HITStudies



TheHITSters



the_hitsters

Inequality:

$$\begin{aligned} \Delta(A, C) + \Delta(C, B) &= d(A, \phi_a(C)) + d(C, \phi_b(B)) \\ &= d(A, \phi_a(C)) + d(\phi_a(C), \phi_b(B)) \\ &> d(A, \phi_a \circ \phi_b(B)) \end{aligned}$$