Radio luminosity functions
with Radio Galaxy Zoo and machine learning

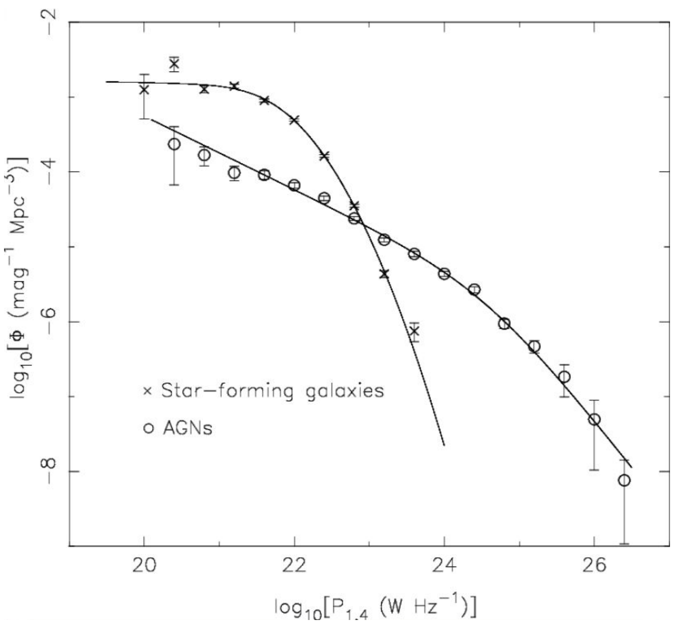
Matthew Alger (ANU/Data61)

O. Ivy Wong (ICRAR/UWA)
Cheng Soon Ong (Data61/ANU)
Naomi McClure-Griffiths (ANU)

Slides: http://www.mso.anu.edu.au/~alger/sparcs-ix

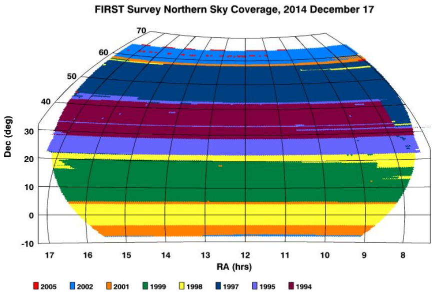
Radio luminosity functions

- Comoving density of radio sources as a function of radio luminosity
  - Units of $\mathrm{dex}^{-1}\ \mathrm{Mpc}^{-3}$ or $\mathrm{mag}^{-1}\ \mathrm{Mpc}^{-3}$
  - Comoving density accounts for universe size and shape over cosmic time
  - Distribution of radio source luminosities in a *physically meaningful* way
- Fractional radio luminosity functions
  - Luminosity function of a subset of sources
  - Luminosity distribution of physically-selected subsets may be different
  - Helps understand evolution and structure of radio galaxies

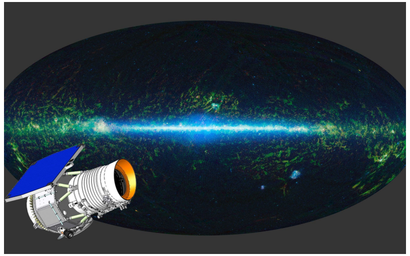


Radio luminosity function divided into radio due to star formation and radio due to active galactic nuclei.
Image: Mauch & Sadler (2007)

Radio Galaxy Zoo

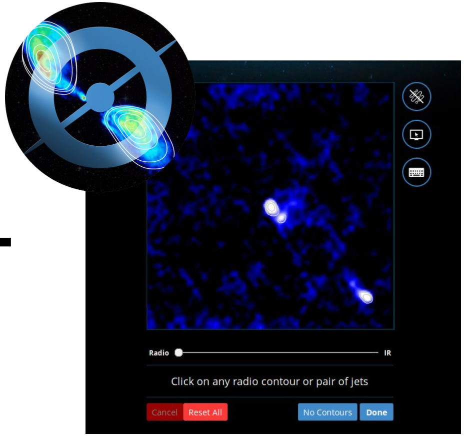


FIRST (radio)

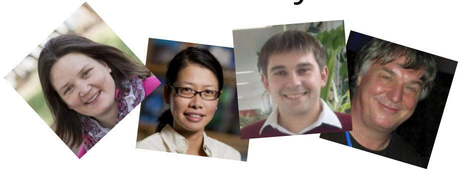


WISE (infrared)

+



Radio Galaxy Zoo



=

Zooniverse ID (3)	RA (4)	Declination (5)	N_{votes} (6)	N_{total} (7)	CL (8)
ARG000255v	251.679244	23.382107	41	42	0.98
ARG000255x	163.799660	23.384972	58	58	1.00
ARG000255y	138.960429	23.381641	43	43	1.00
ARG000255z	126.215156	23.381729	35	35	1.00
ARG0002560	149.273620	23.381661	40	40	1.00

Radio Galaxy Zoo cross-identification catalogue

Cross-identification as binary classification

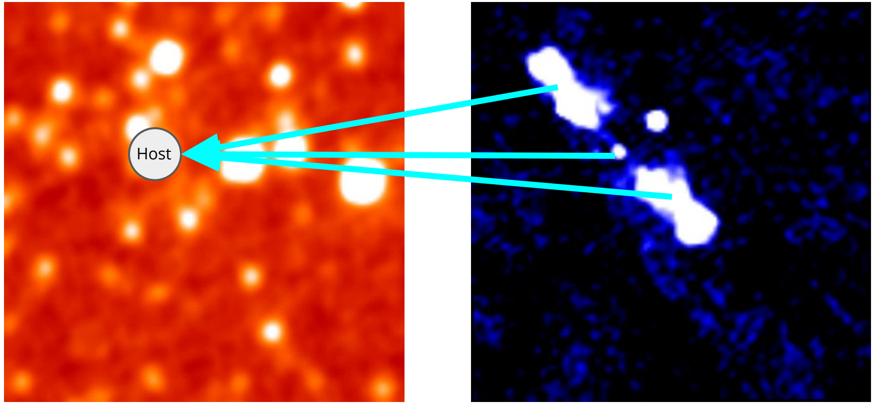
- Cross-identification
  - Match radio emission to infrared host galaxies
  - Output of Radio Galaxy Zoo
- Can be cast as binary classification
  - Binary classification is well-understood
  - Lots of off-the-shelf classification models
  - Easy to train
- Problems:
  - Converting cross-identification catalogues to binary labels loses information
  - Unclear how uncertainties in this formulation are related to dataset or physical uncertainties

$$f : \mathbb{R}^d \to \mathbb{R}$$

$$\mathrm{xid} : \mathrm{Radio} \to \mathrm{IR}$$

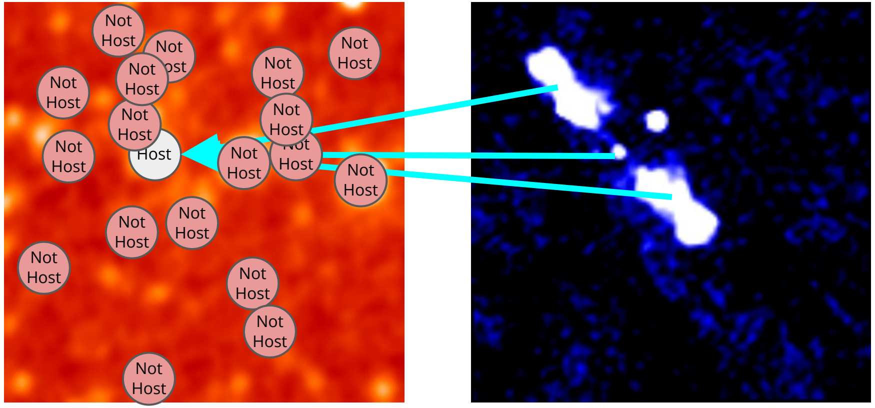
$$\mathrm{xid}(r) = \operatorname*{argmax}_{i \in \text{IR objects}} f(i; r)$$

Learning from Radio Galaxy Zoo



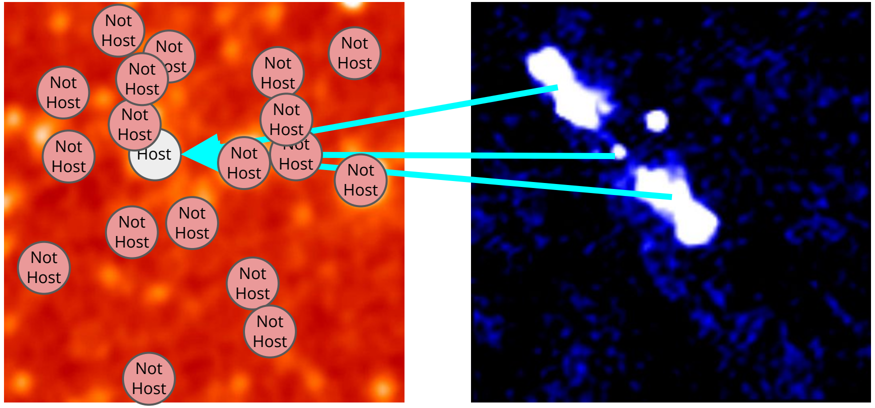
- Assign hosts positive labels

Learning from Radio Galaxy Zoo

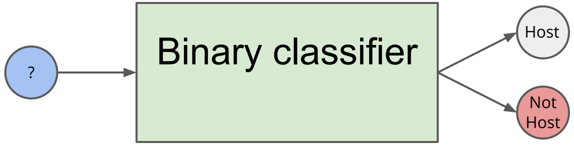


- Assign hosts positive labels
- Assign everything else negative labels

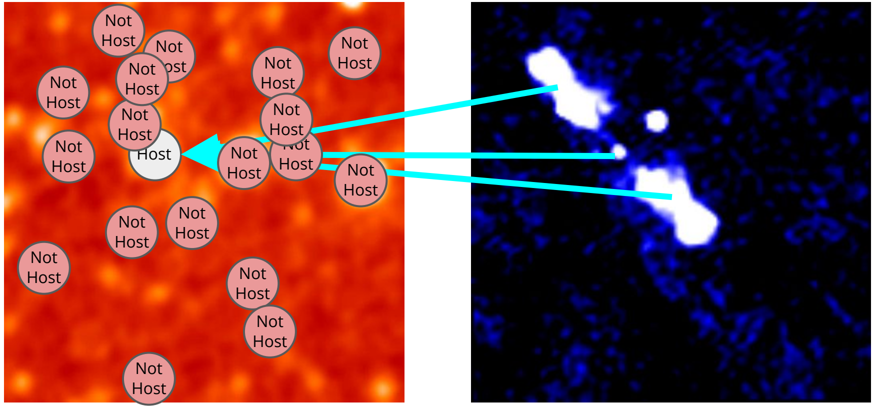
Learning from Radio Galaxy Zoo



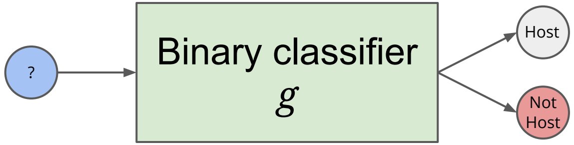
- Assign hosts positive labels
- Assign everything else negative labels
- Train classifier to identify *host* and *not host* classes



Learning from Radio Galaxy Zoo



- Assign hosts positive labels
- Assign everything else negative labels
- Train classifier to identify *host* and *not host* classes

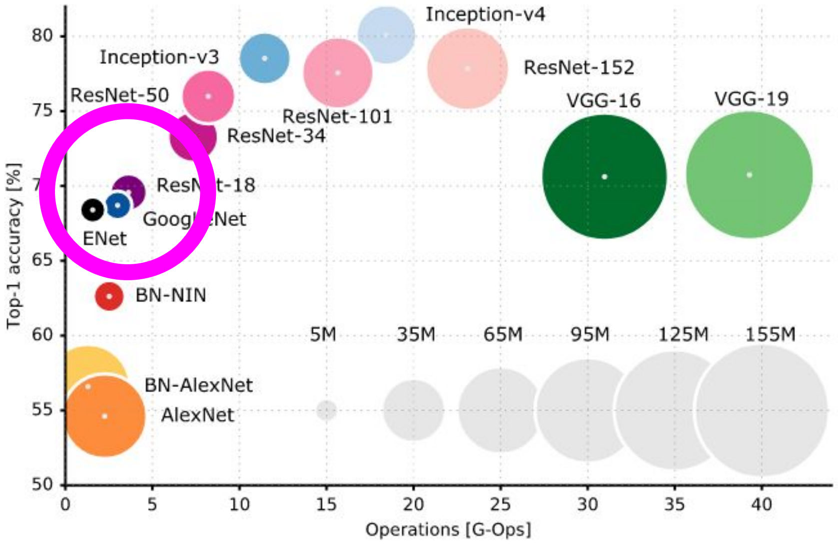


$$\mathrm{xid} : \mathrm{Radio} \rightarrow \mathrm{IR}$$

$$\mathrm{xid}(r) = \underset{i \in \text{IR objects}}{\operatorname{argmax}}\ g(i)\ \mathcal{N}(r, i)$$

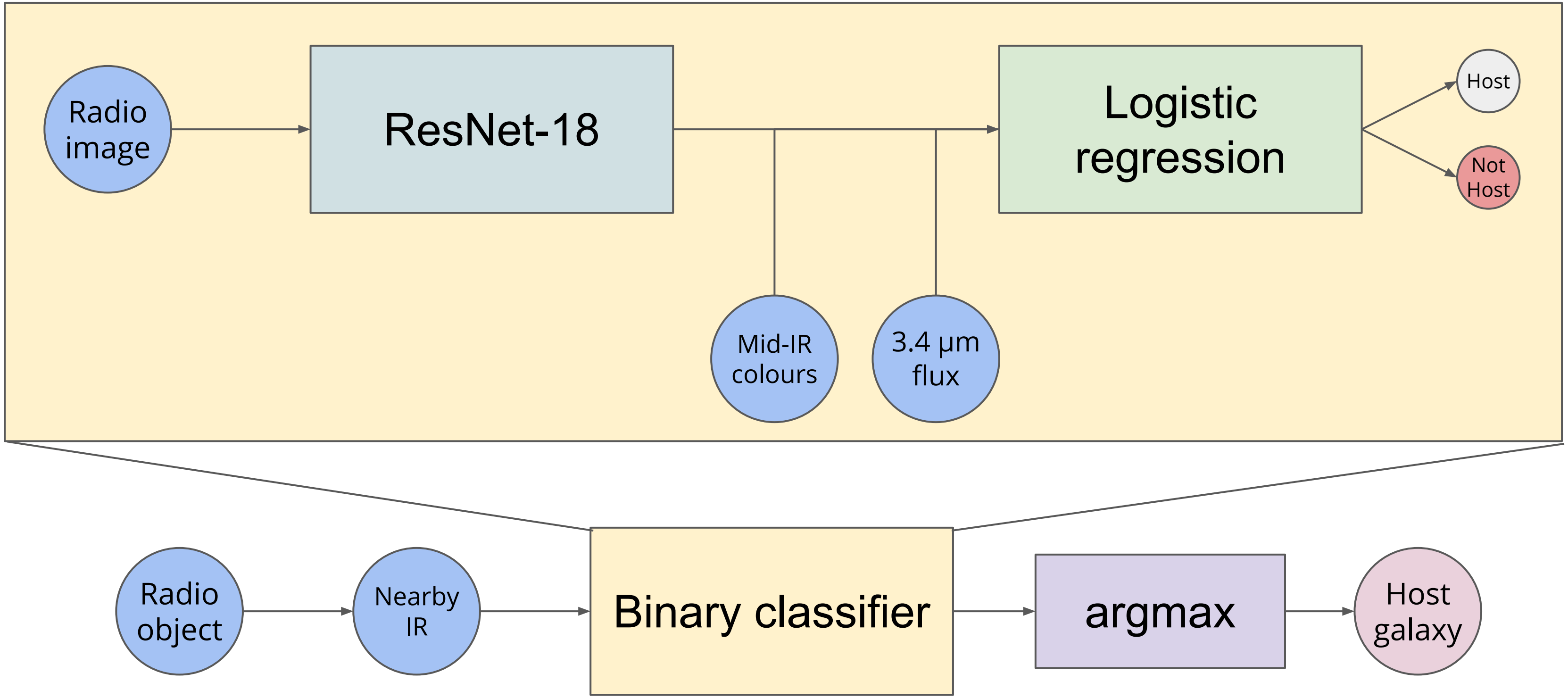
Binary classification model

- ResNet-18 (multiclass)
  - Good accuracy
  - Low complexity
  - Very fast to train and use
- Remove last layer and replace with a binary classifier
- Add non-image features
  - Mid-infrared colours
  - 3.4 μm flux
  - Room for improvement — e.g. add redshifts
- Accurate for ~96% of sources



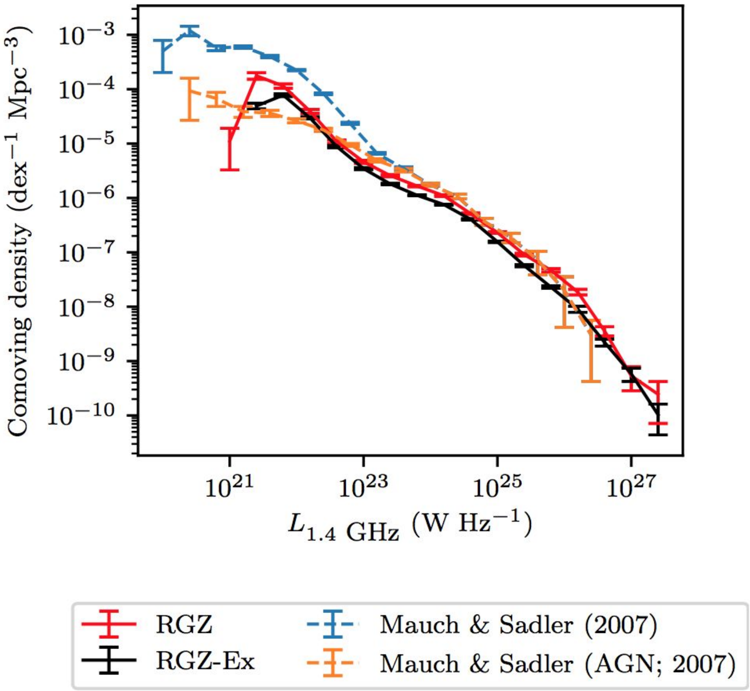
Trade-offs between network complexity and accuracy on ImageNet.
Image: Canziani et al. (2016)

Binary classification model



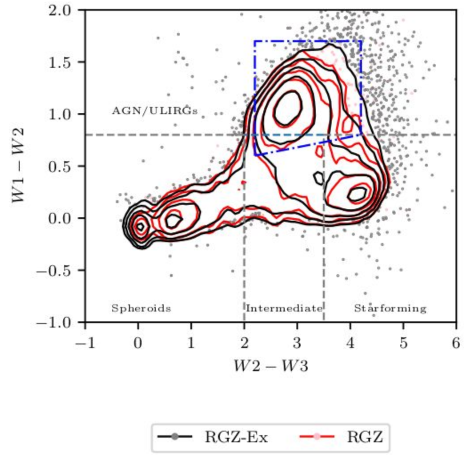
Luminosity function of extended sources

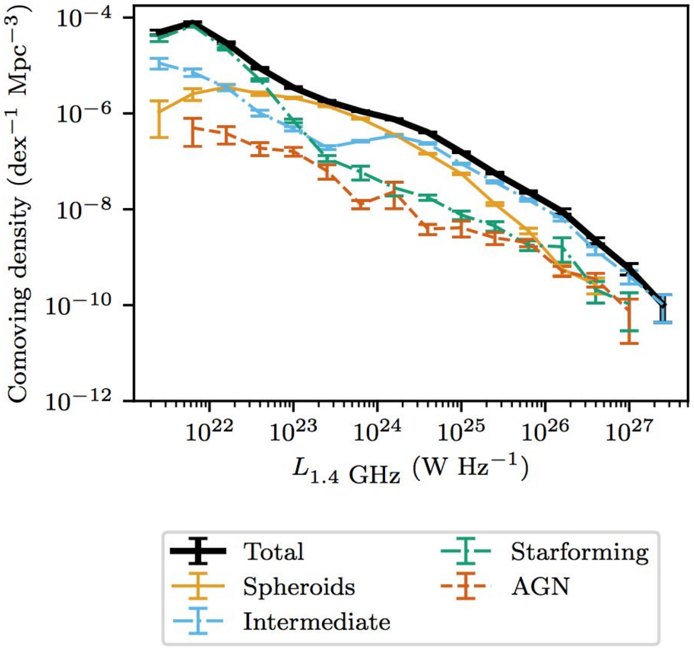
- RGZ-Ex contains 214 214 cross-identified radio components with 26 268 redshifts
  - >4x more components than RGZ
  - >2x more components with redshifts than RGZ
- Large sample allows us to build a radio luminosity function of extended sources
  - Luminosities 10^{21}–10^{27} W/Hz



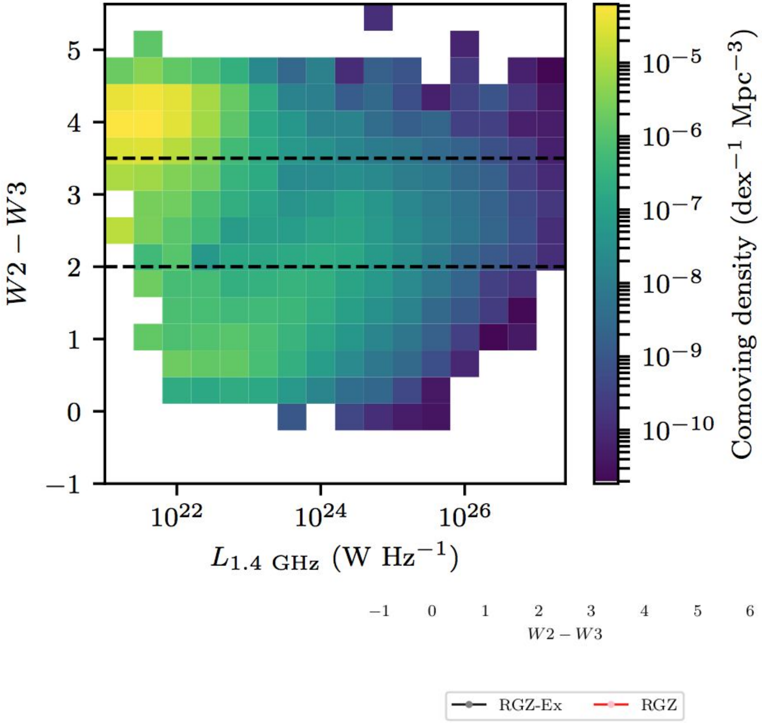
Fractional luminosity function (Mid-IR)

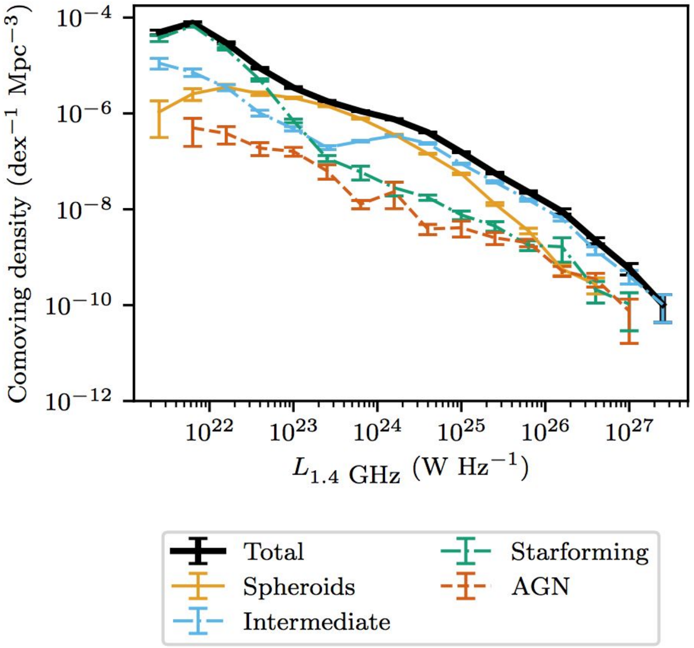
- Divide radio luminosity function based on mid-infrared host colours
  - “Extended” star-forming sources below 10^{23} W/Hz
  - Radio-loud sources dominated by “intermediate” galaxies





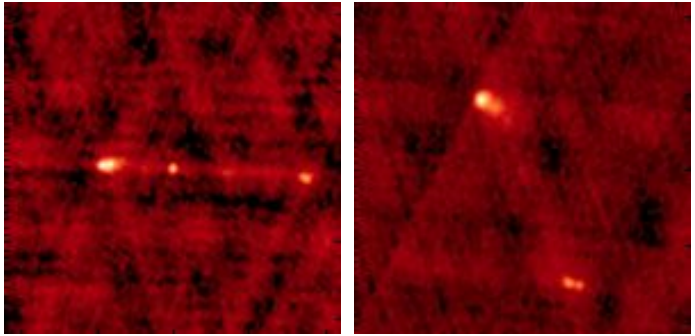
Fractional luminosity function (Mid-IR)





Source finding with cross-identification

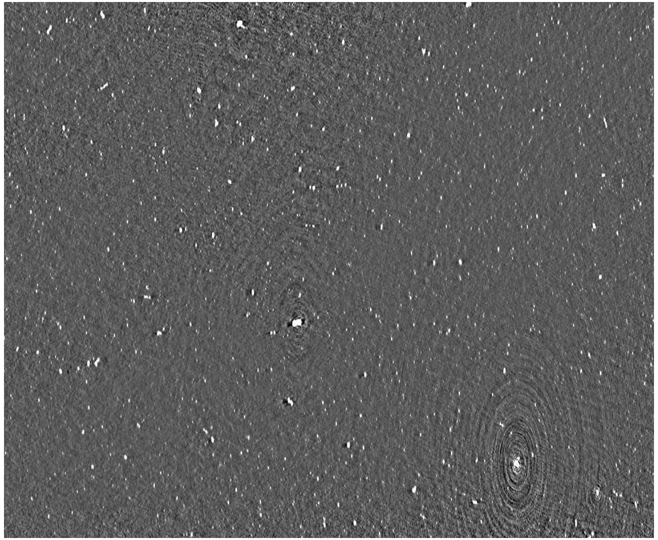
- We matched 214 214 radio components to 157 007 host galaxies
- If two components have the same host they are part of the same source
  - Many false positives
- RGZ-Ex candidate source catalogue
  - 157 007 candidate extended sources
  - All data required to reproduce our results
  - At least 10 new giants



Two giant radio galaxies identified in our catalogue, each 1.1 Mpc across.
Images: FIRST

Looking forward: EMU

- Very different results to previous work on EMU-ATLAS
  - Much lower source density on sky
  - Higher angular resolution — more extended sources
  - More training and prediction data
  - Alger+18, doi:10.1093/mnras/sty1308
- Generalisation to EMU will be non-trivial
- Radio Galaxy Zoo for EMU?



ATLAS observations of CDFS.
Image: ATLAS DR3, Franzen+2015