Projeto de uma plataforma-ISSC-LSST para classificação automática de galáxias com aprendizagem profunda

Reinaldo R. Rosa

Coordenadoria de Pesquisa Aplicada e Desenvolvimento Tecnológico (COPDT)
INPE-MCTI

LSST-Sci. Colab.

1. Galaxies SC
2. S-MW-LV SC
3. Solar System SC
4. Dark Energy SC
5. AGN SC
6. Transients & VS SC
7. Strong L. SC
8. Inf. & Stat. SC


Thales


Luan


Vitor

BPG-LSST ISSC

LINEA


Reinaldo


Rubens

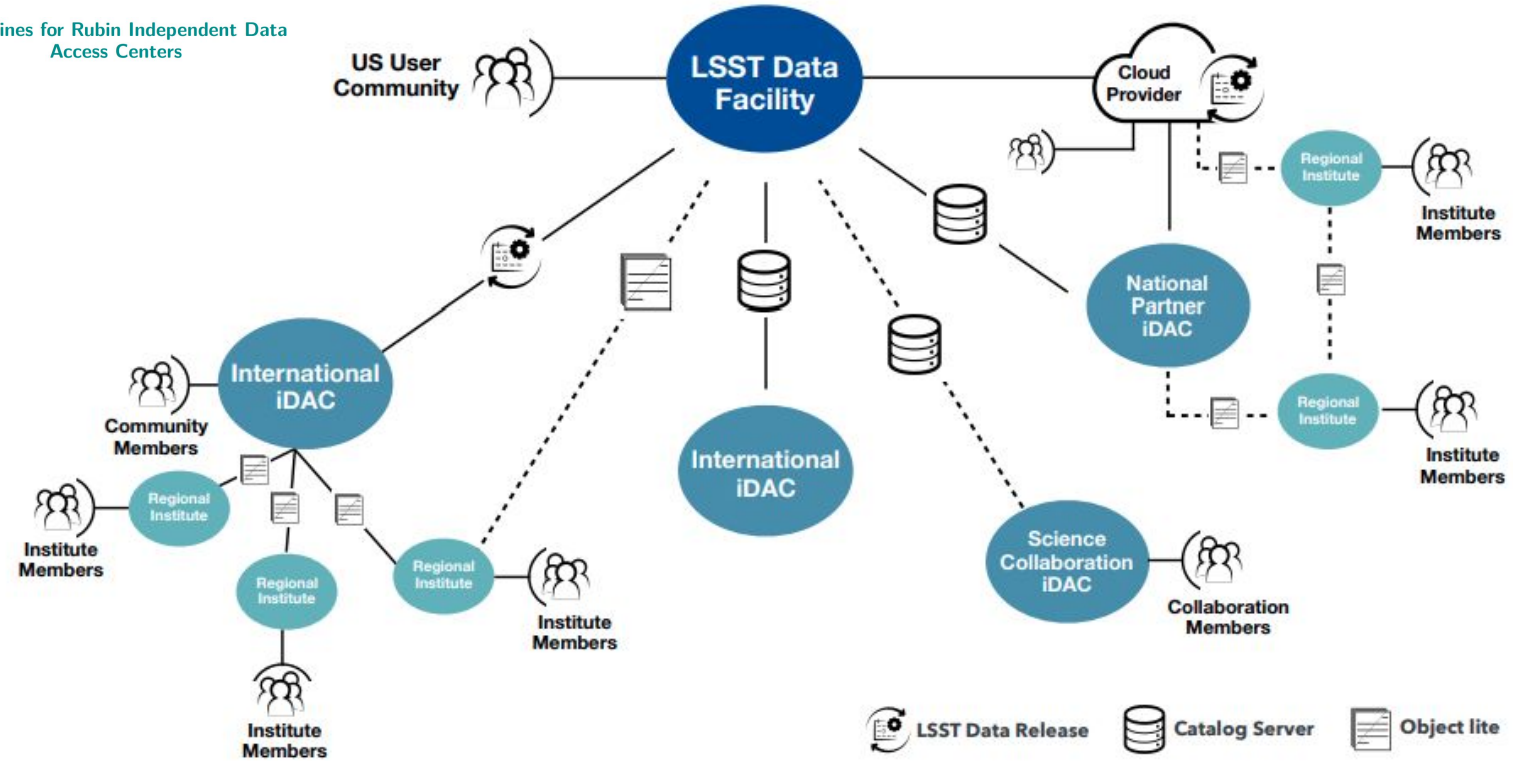

Carlos E. Falandes

Data Science Fellowship Program

LSST Discovery Alliance

IDAC network topology

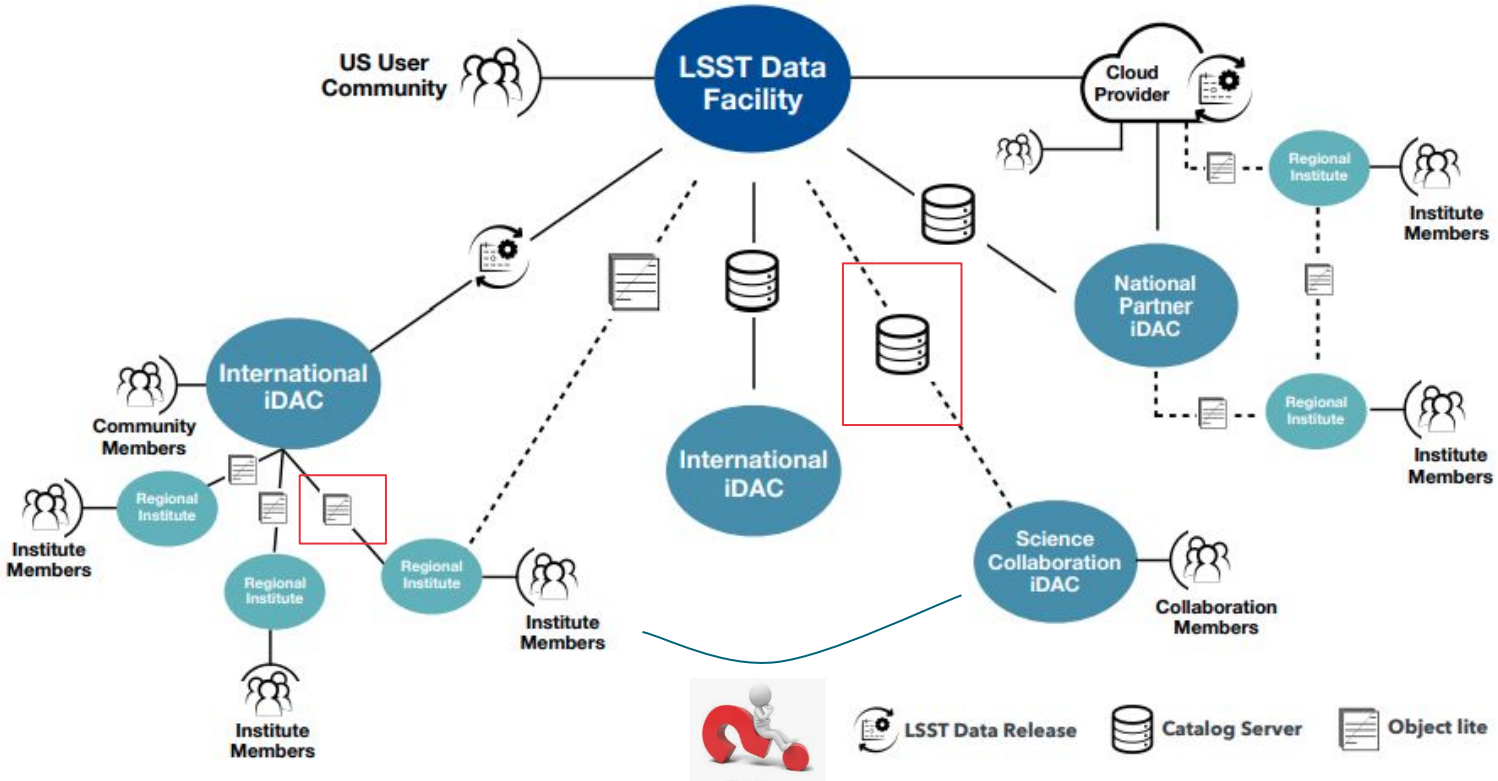
Guidelines for Rubin Independent Data Access Centers



BPG-LSST ISSC

BPG-LSST ISSC LineA

IDAC network topology



ISSC liaisons:

- LSST Survey Cadence Optimization Committee: Federica Bianco (University of Delaware)
- Rubin Data Management: Leanne Guy (Rubin)

ISSC Leadership Team

ISSC co-Chairs:

- Francois Lanusse (CNRS)
- Ashley Villar (Harvard)

ISSC Executive Council:

- Melissa DeLucchi
- Alex Gagliano
- Matthew Graham
- Francois Lanusse (ex-officio)
- Alex Malz
- Ashish Mahabal
- Ashley Villar (ex-officio)

Ethics Committee:

- Eric Aubourg
- Reinaldo Roberto Rosa
- Lior Shamir (chair)
- Yuanyuan Zhang

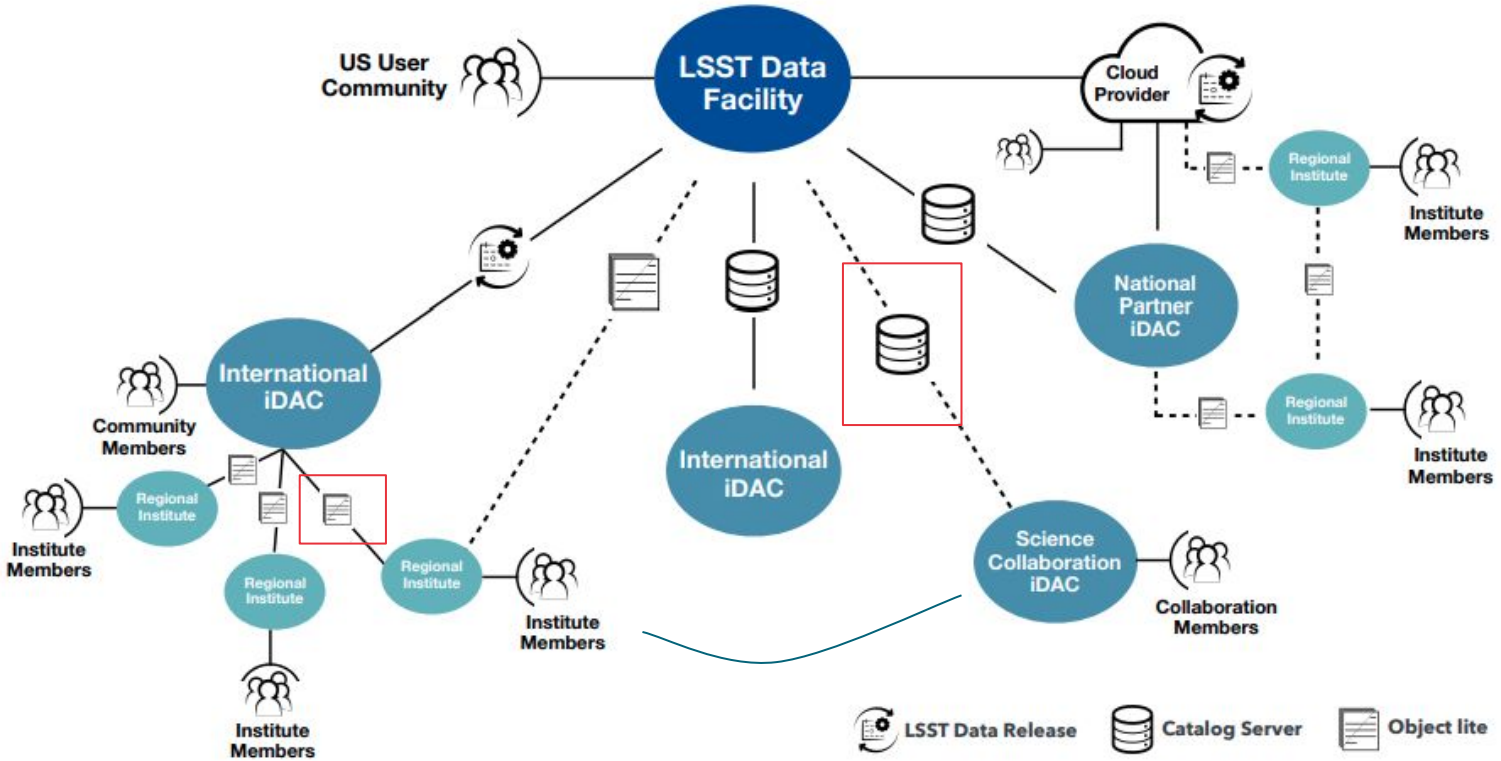
Publication Committee:

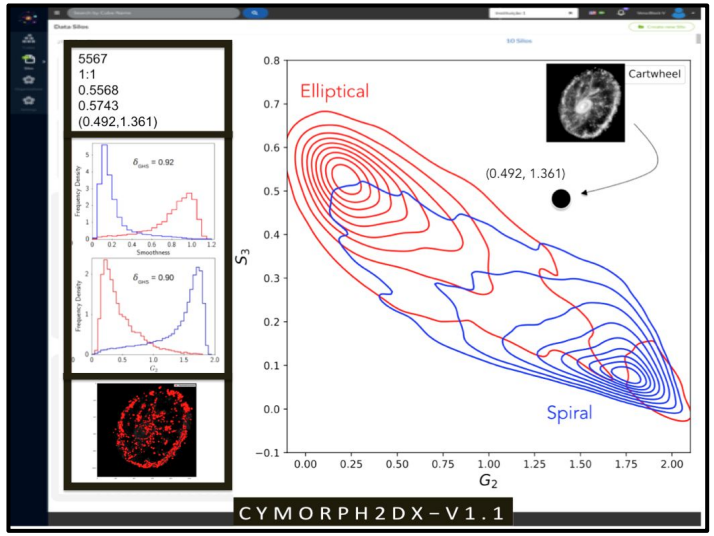
- Mariano Dominguez
- Hermine Landt (chair)
- Alex Malz
- Josh Speagle

Membership Committee:

- Tatiana Acero-Cuellar
- Alex Gagliano
- Bryce Kalmbach (chair)
- Konstantin Malanchev

IDAC network topology





ISSC liaisons:

- LSST Survey Cadence Optimization Committee: Federica Bianco (University of Delaware)
- Rubin Data Management: Leanne Guy (Rubin)

ISSC Leadership Team

ISSC co-Chairs:

- Francois Lanusse (CNRS)
- Ashley Villar (Harvard)

ISSC Executive Council:

- Melissa DeLucchi
- Alex Gagliano
- Matthew Graham
- Francois Lanusse (ex-officio)
- Alex Malz
- Ashish Mahabal
- Ashley Villar (ex-officio)

Ethics Committee:

- Eric Aubourg
- Reinaldo Roberto Rosa
- Lior Shamir (chair)
- Yuanyuan Zhang

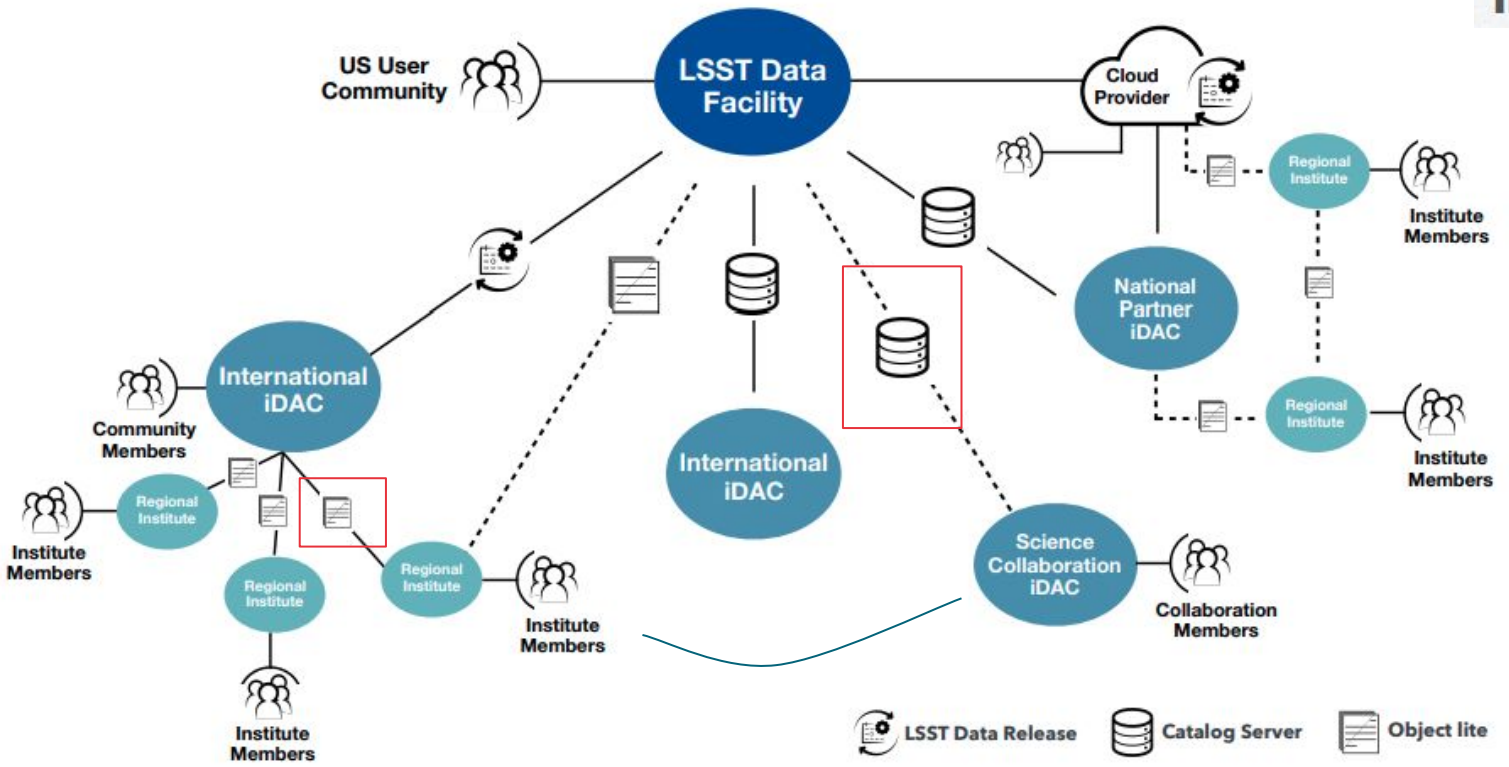
Publication Committee:

- Mariano Dominguez
- Hermine Landt (chair)
- Alex Malz
- Josh Speagle

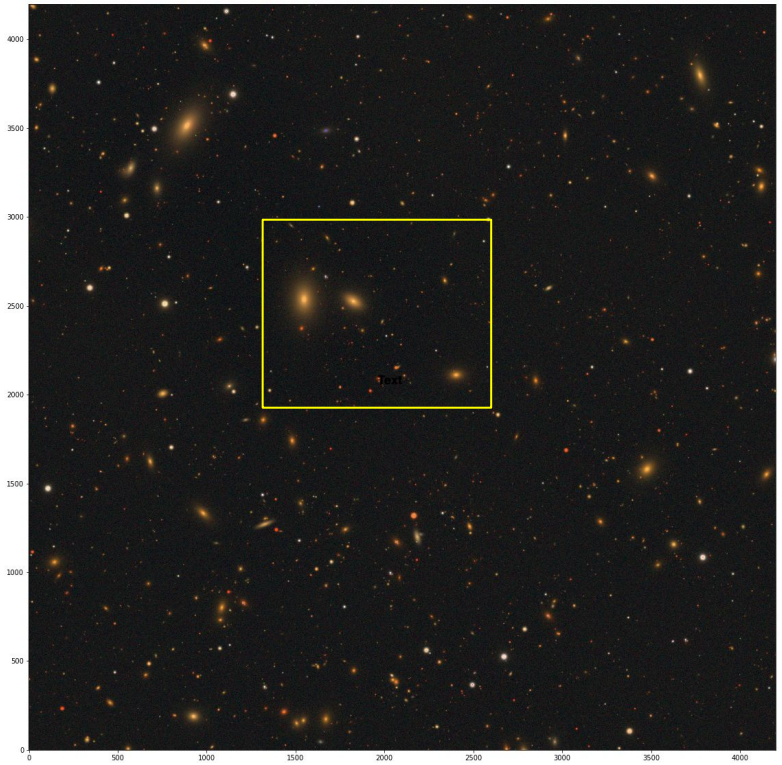
Membership Committee:

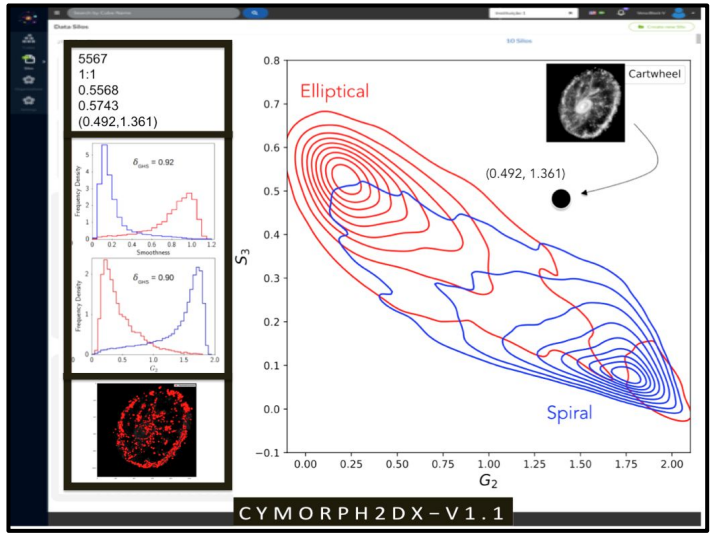
- Tatiana Acero-Cuellar
- Alex Gagliano
- Bryce Kalmbach (chair)
- Konstantin Malanchev

IDAC network topology



ImSim https://github.com/LSSTDESC/Imsim





ISSC Leadership Team

ISSC co-Chairs:

- Francois Lanusse (CNRS)
- Ashley Villar (Harvard)

ISSC Executive Council:

- Melissa DeLucchi
- Alex Gagliano
- Matthew Graham
- Francois Lanusse (ex-officio)
- Alex Malz
- Ashish Mahabal
- Ashley Villar (ex-officio)

Ethics Committee:

- Eric Aubourg
- Reinaldo Roberto Rosa
- Lior Shamir (chair)
- Yuanyuan Zhang

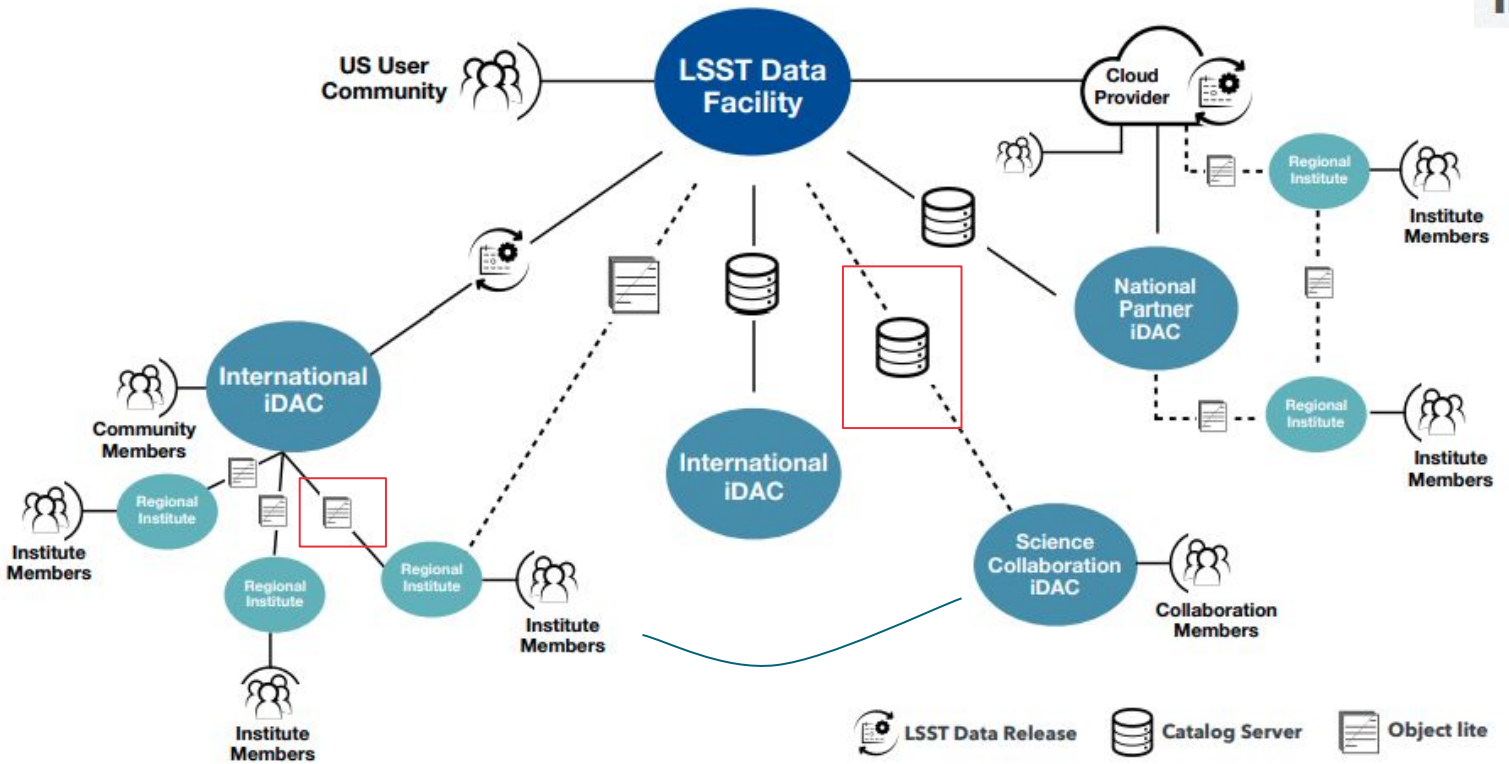
Publication Committee:

- Mariano Dominguez
- Hermine Landt (chair)
- Alex Malz
- Josh Speagle

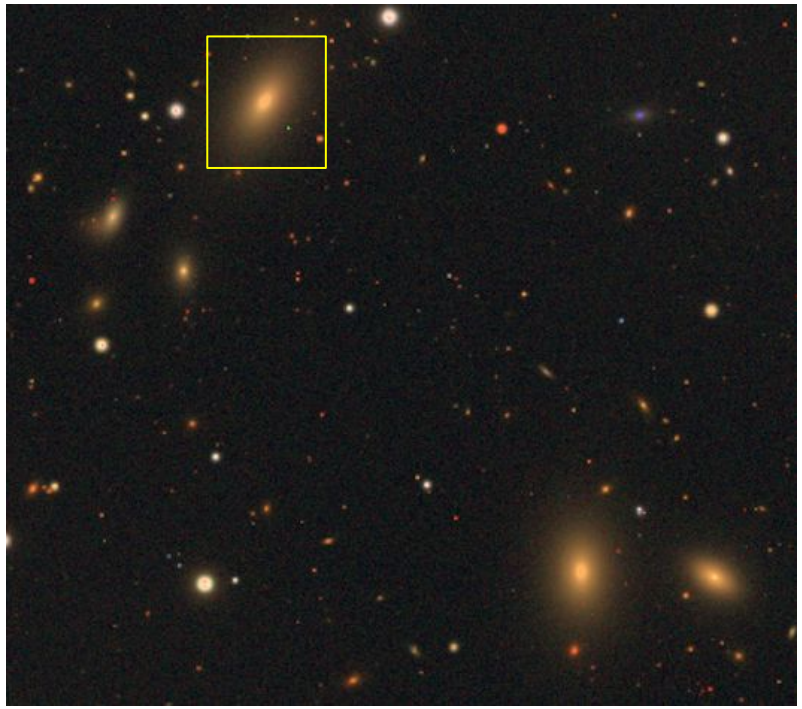
Membership Committee:

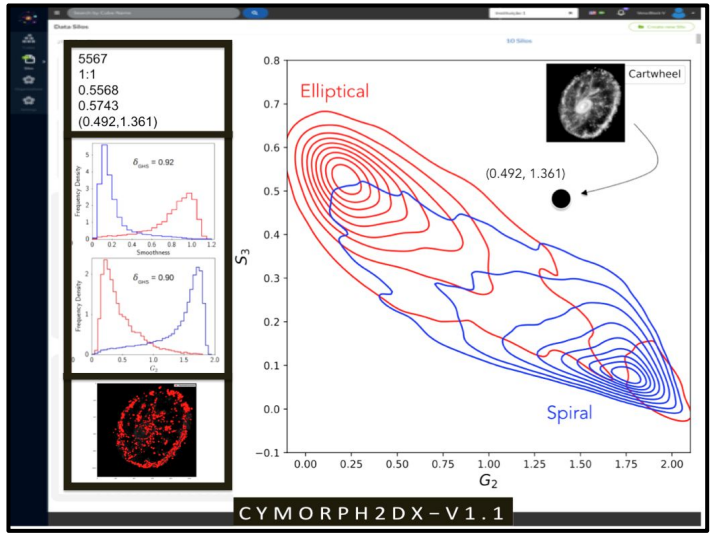
- Tatiana Acero-Cuellar
- Alex Gagliano
- Bryce Kalmbach (chair)
- Konstantin Malanchev

IDAC network topology



ImSim https://github.com/LSSTDESC/Imsim





ISSC Leadership Team

ISSC co-Chairs:

- Francois Lanusse (CNRS)
- Ashley Villar (Harvard)

ISSC Executive Council:

- Melissa DeLucchi
- Alex Gagliano
- Matthew Graham
- Francois Lanusse (ex-officio)
- Alex Malz
- Ashish Mahabal
- Ashley Villar (ex-officio)

Ethics Committee:

- Eric Aubourg
- Reinaldo Roberto Rosa
- Lior Shamir (chair)
- Yuanyuan Zhang

Publication Committee:

- Mariano Dominguez
- Hermine Landt (chair)
- Alex Malz
- Josh Speagle

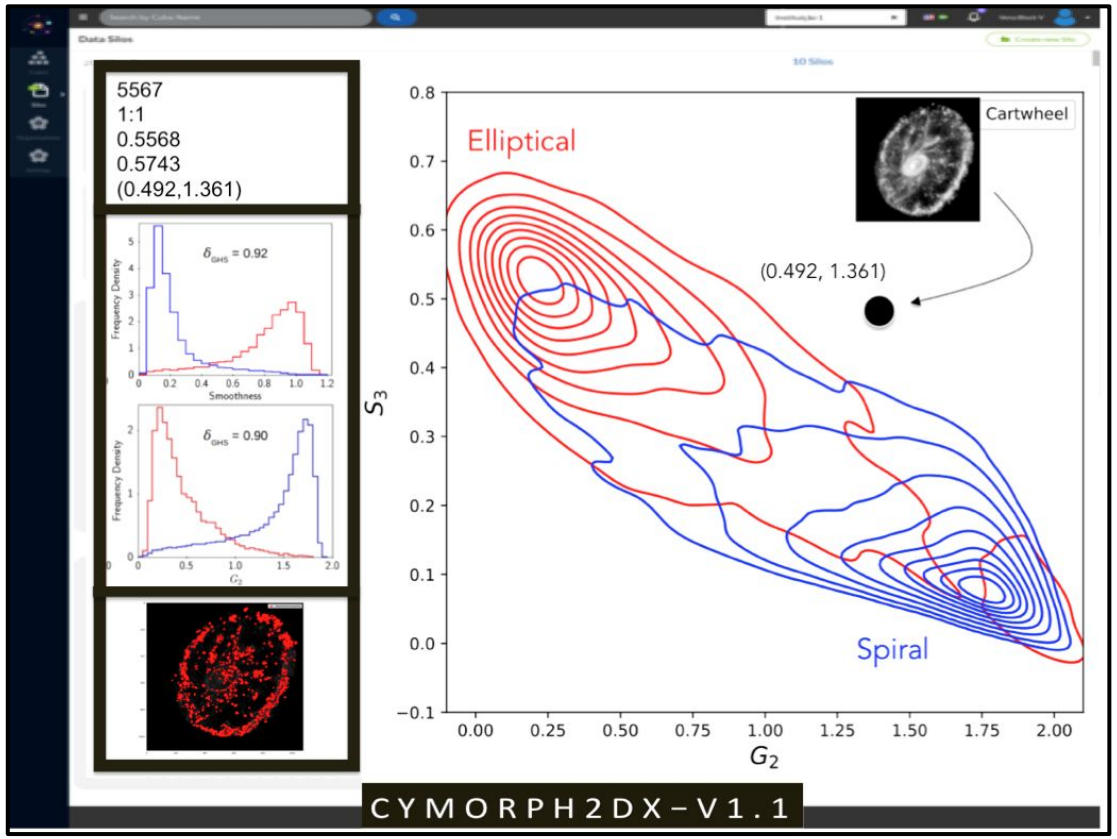
Membership Committee:

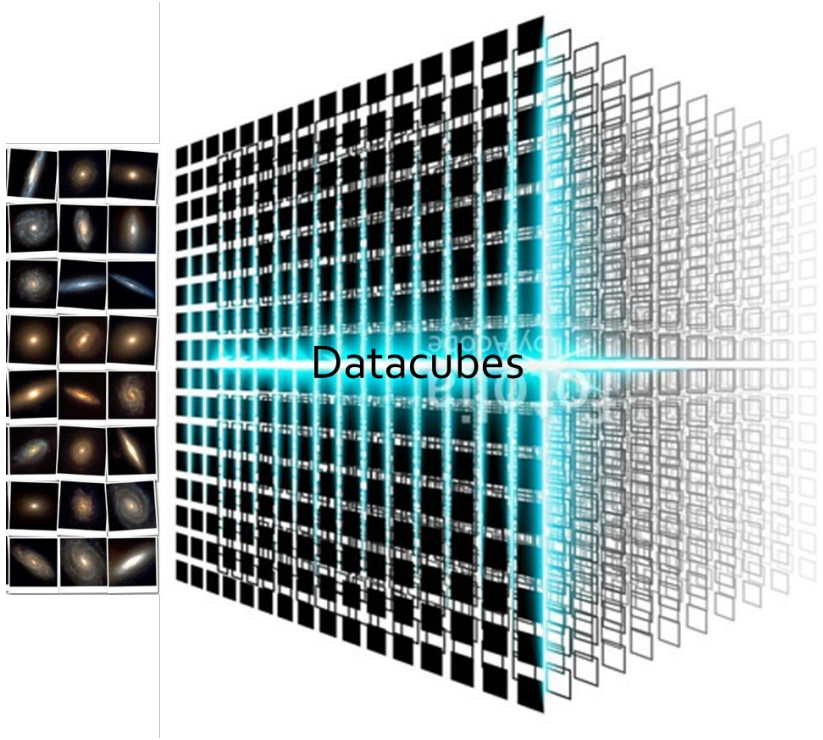
- Tatiana Acero-Cuellar
- Alex Gagliano
- Bryce Kalmbach (chair)
- Konstantin Malanchev

User Generated DP 100% baseado na nuvem?

RH/$: ISSC e/ou IDAC-BPG?

BPG-LSST ISSC





70-200 PB

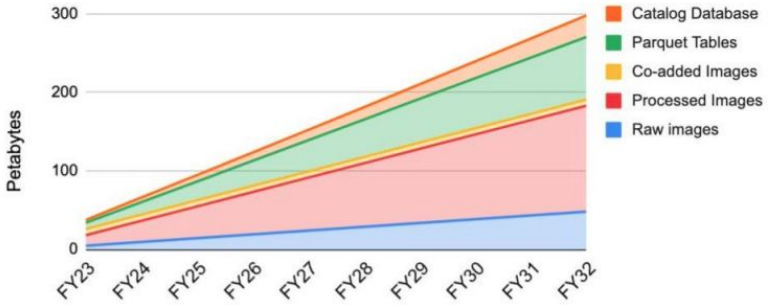


FIGURE 1: Volume of different datatypes over time from DMTN-135 data.

Guidelines for Rubin Independent Data Access Centers

Morphometrics

CAS

$$C = 5 \times \log\left(\frac{R_{80}}{R_{20}}\right)$$

$$A(\phi) = \frac{\sum |I_{\text{original}} - I_{\text{rotated}}|}{\sum I_{\text{original}}} - A_{\text{bkg}}$$

$$S = 10 \times \left(\frac{\sum |I_{\text{original}} - I_{\text{smoothed}}|}{\sum I_{\text{original}}}\right) - S_{\text{bkg}}$$

EGG

$$E(I) = -\sum_{k}^{K} p(I_k)\,\log[p(I_k)]$$

$$\tilde{E}(I) = \frac{E(I)}{E_{\max}}, \quad \text{where } 0 \le \tilde{E}(I) \le 1 \text{ and } E_{\max} = \log(K)$$

$$G = \frac{1}{2\bar{X}n(n-1)} \sum_{i=1}^{n}\sum_{j=1}^{n} |X_i - X_j|$$

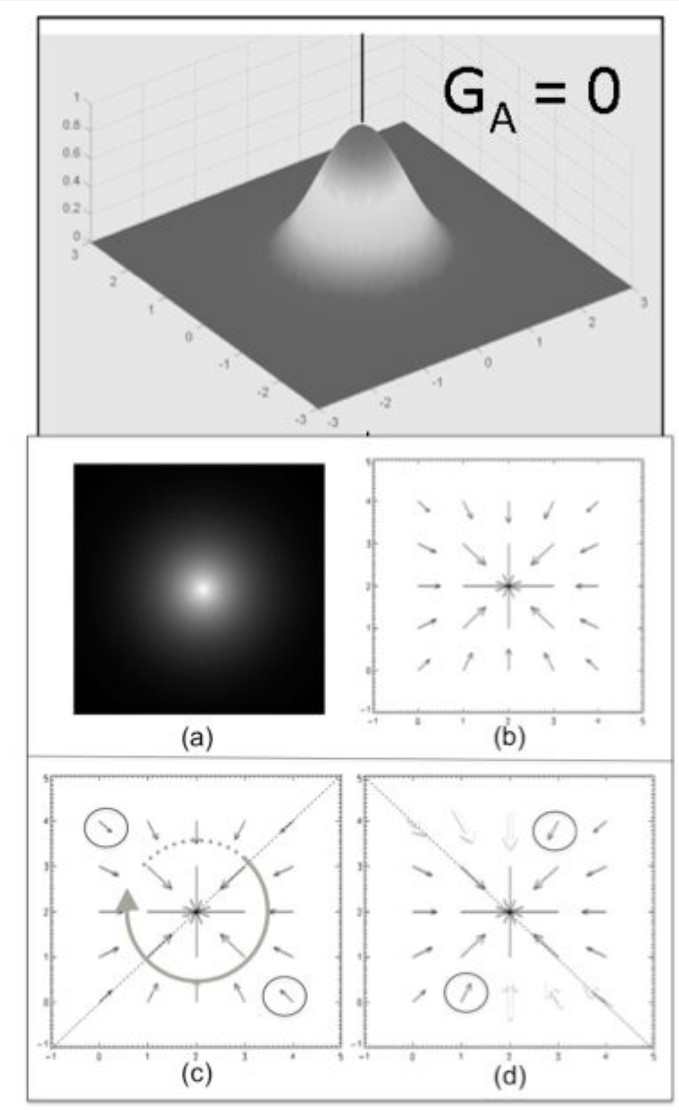
(total homoneity)0<G<1 (total inh.)

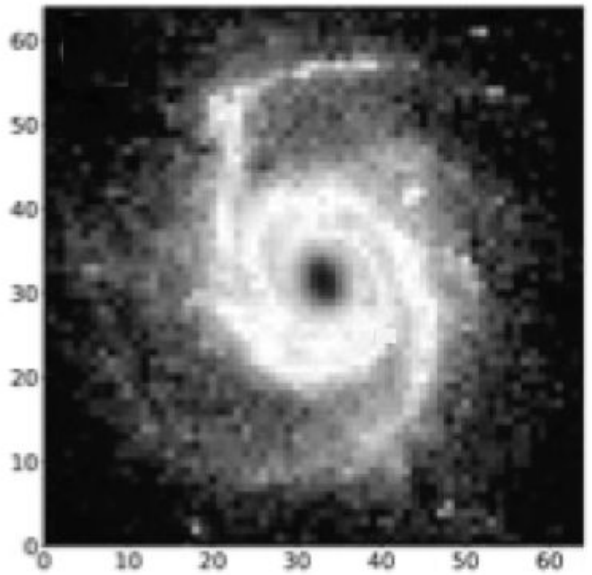
$$G_2 = \frac{V_A}{V}(1 - confluence)$$

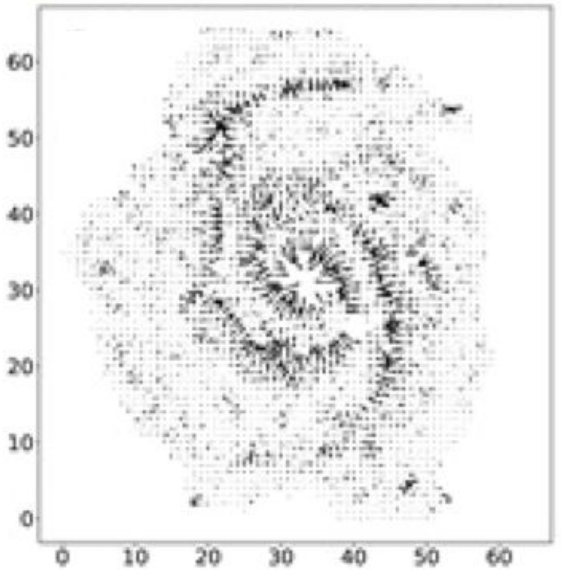
$$confluence = \left(\frac{\left|\sum_i^{V_A} v_a^i\right|}{\sum_i^{V_A} |v_a^i|}\right)$$

+ G_1, G_3, G_4

Morphometrics







$$G_2 = \frac{V_A}{N}\left(2 - \frac{\left|\sum_i^{V_A} r_i\right|}{\sum_i^{V_A} |r_i|}\right)$$

Morphometrics

https://doi.org/10.1093/mnrasl/sly054

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS 477, L101–L105 (2018) doi:10.1093/mnrasl/sly054
Advance Access publication 2018 April 04

Gradient pattern analysis applied to galaxy morphology

R. R. Rosa,[1★] R. R. de Carvalho,[2★] R. A. Sautter,[1] P. H. Barchi,[1★] D. H. Stalder,[1,3] T. C. Moura,[2] S. B. Rembold,[4] D. R. F. Morell[5] and N. C. Ferreira[1]

[1]Lab for Computing and Applied Mathematics, National Institute for Space Research (INPE), S. J. dos Campos, 12245-970, Brazil
[2]Astrophysics Division, National Institute for Space Research (INPE), S. J. dos Campos, 12245-970, Brazil
[3]NIDTEC/FP-UNA, San Lorenzo, 2169-000, Paraguay
[4]Departamento de Física, CCNE, Universidade Federal de Santa Maria, Santa Maria, 97105-900, Brazil
[5]DCET, Universidade Estadual de Santa Cruz, Ilheus, 45662-900, Brazil

SDSS Data Release 7 (Abazajian et al. 2009)

0.01<z<0.3
mag > 17.78
54,896 objects
SDSS-DR7 / GZ1

https://doi.org/10.1016/j.ascom.2019.100334

Astronomy and Computing
Volume 30, January 2020, 100334

Full length article

Machine and Deep Learning applied to galaxy morphology - A comparative study

P.H. Barchi [a b], R.R. de Carvalho [c d], R.R. Rosa [a], R.A. Sautter [a], M. Soares-Santos [b], B.A.D. Marques [e], E. Clua [e], T.S. Gonçalves [f], C. de Sá-Freitas [f], T.C. Moura [g]

https://doi.org/10.1093/mnras/stad3934

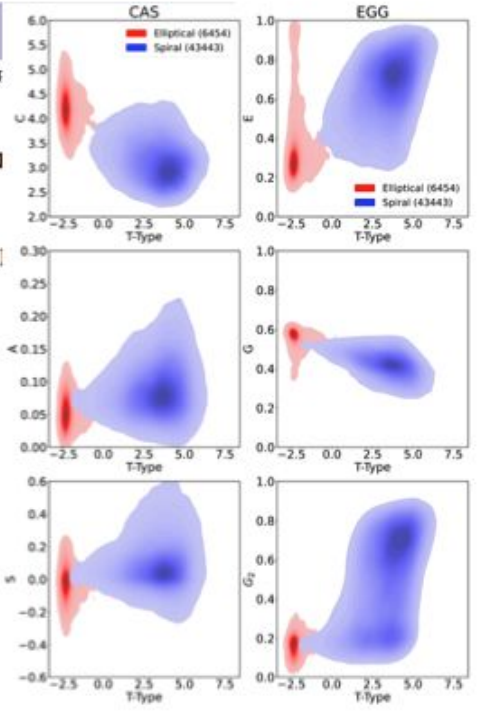
Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS 528, 82–107 (2024)
Advance Access publication 2023 December 21

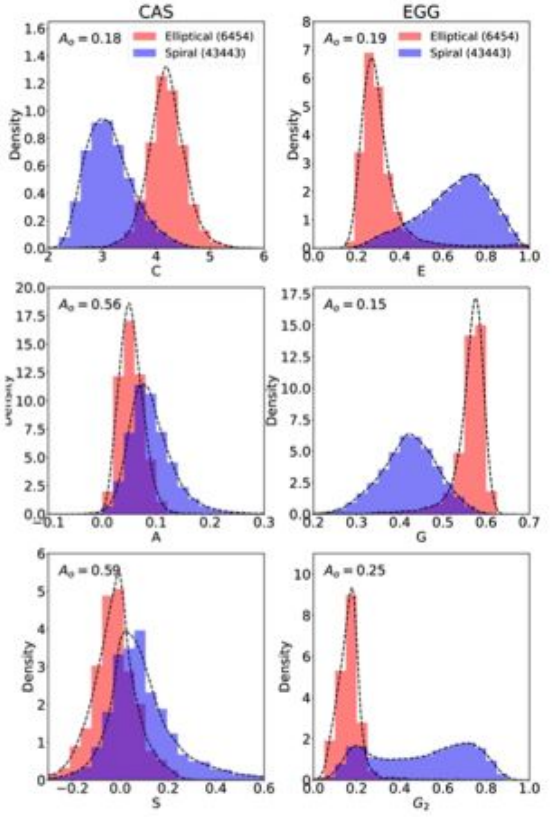
Unveiling galaxy morphology through an unsupervised-su approach

I. Kolesnikov [1,2★] V. M. Sampaio [3,4] R. R. de Carvalho,[4] C. Conselice,[2] S. B. C. L. Mendes[1] and R. R. Rosa[1]

[1]Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 12227-010 SP, Brazil
[2]Jodrell Bank Centre for Astrophysics, University of Manchester, Oxford Road, Manchester M13 9PL, UK
[3]School of Physics and Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD, UK
[4]NAT - Universidade Cidade de São Paulo, 01506-000 SP, Brazil
[5]Universidade Federal de Santa Maria, Santa Maria, RS 97105-900, Brazil

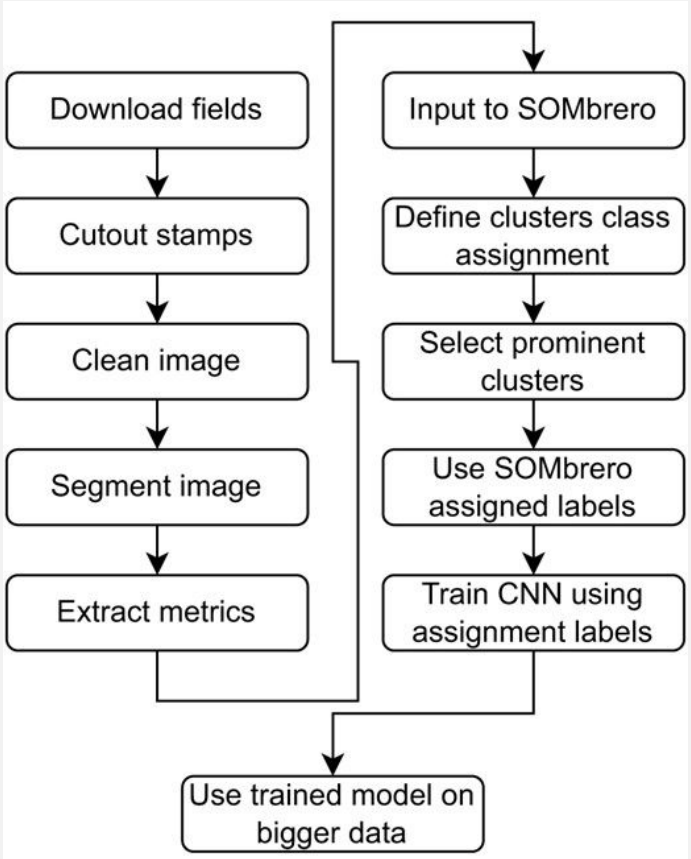
Domínguez Sanchéz et al. (2018)
670,602 objects (+ T-Type Value)





Morphometrics

Cymorph 3.0

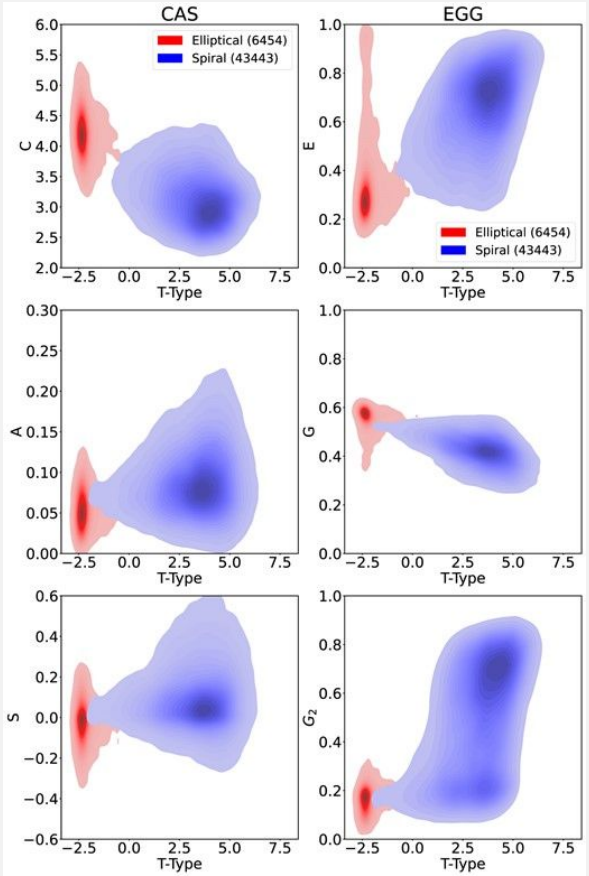


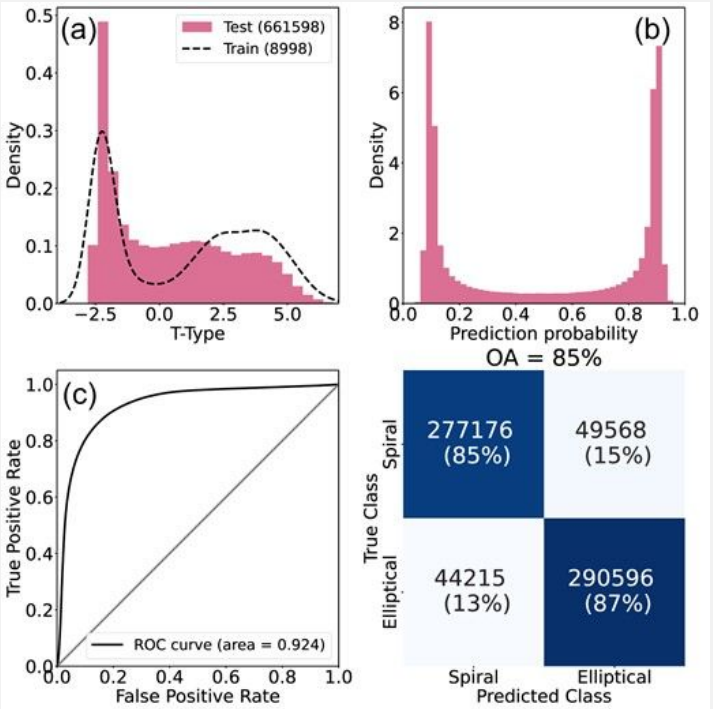
Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS 528, 82–107 (2024)
Advance Access publication 2023 December 21
https://doi.org/10.1093/mnras/stad3934

Unveiling galaxy morphology through an unsupervised-supervised hybrid approach

I. Kolesnikov [1,2★] V. M. Sampaio [3,4] R. R. de Carvalho,[4] C. Conselice,[2] S. B. Rembold,[5] C. L. Mendes[1] and R. R. Rosa [1]

[1] *Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 12227-010 SP, Brazil*
[2] *Jodrell Bank Centre for Astrophysics, University of Manchester, Oxford Road, Manchester M13 9PL, UK*
[3] *School of Physics and Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD, UK*
[4] *NAT - Universidade Cidade de São Paulo, 01506-000 SP, Brazil*
[5] *Universidade Federal de Santa Maria, Santa Maria, RS 97105-900, Brazil*





How can generative AI positively impact the morphological classification of Galaxies?

The role of generative AI in training a *hyper citizen bot* for a Galaxy Zoo-like environment

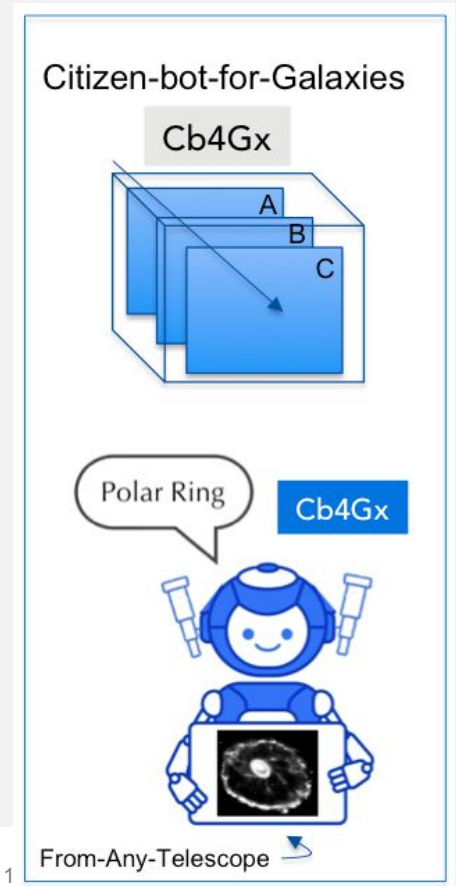
https://github.com/indiradutta/DC_GAN

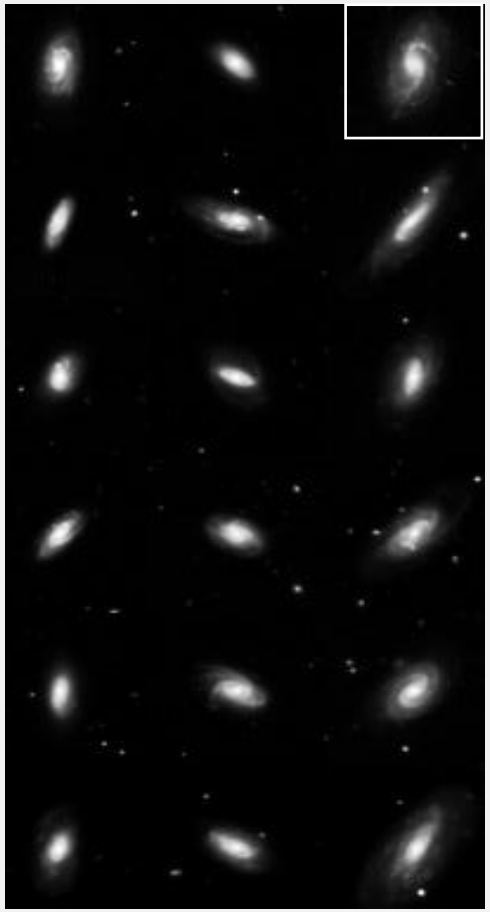
Deep Convolutional Generative Adversarial Network

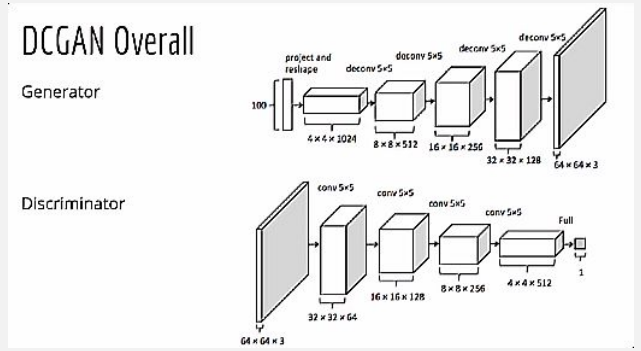
junyanz/**pytorch-CycleGAN-and-pix2pix**

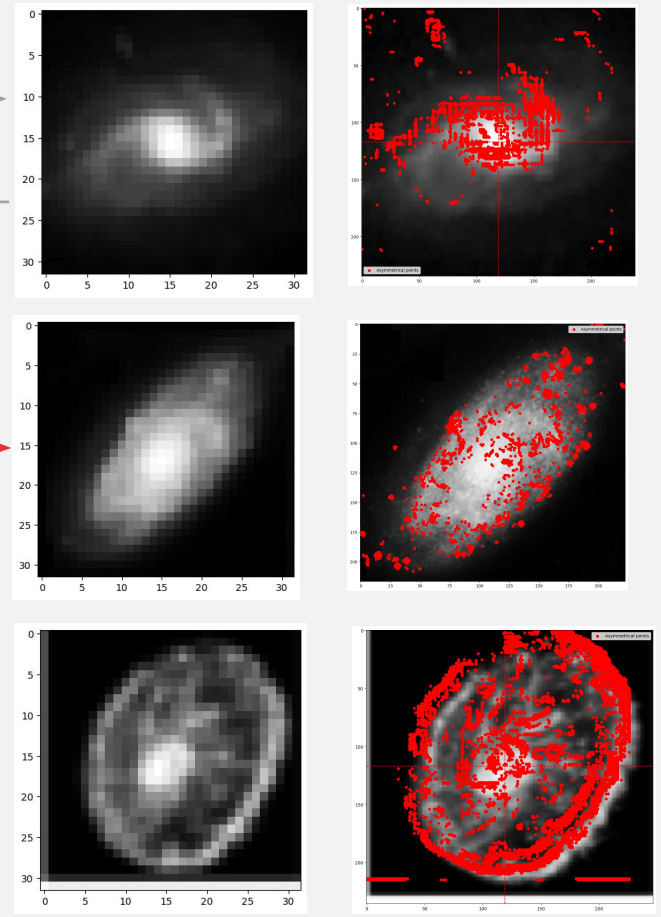
Image-to-Image Translation in PyTorch

SDSS-Generator

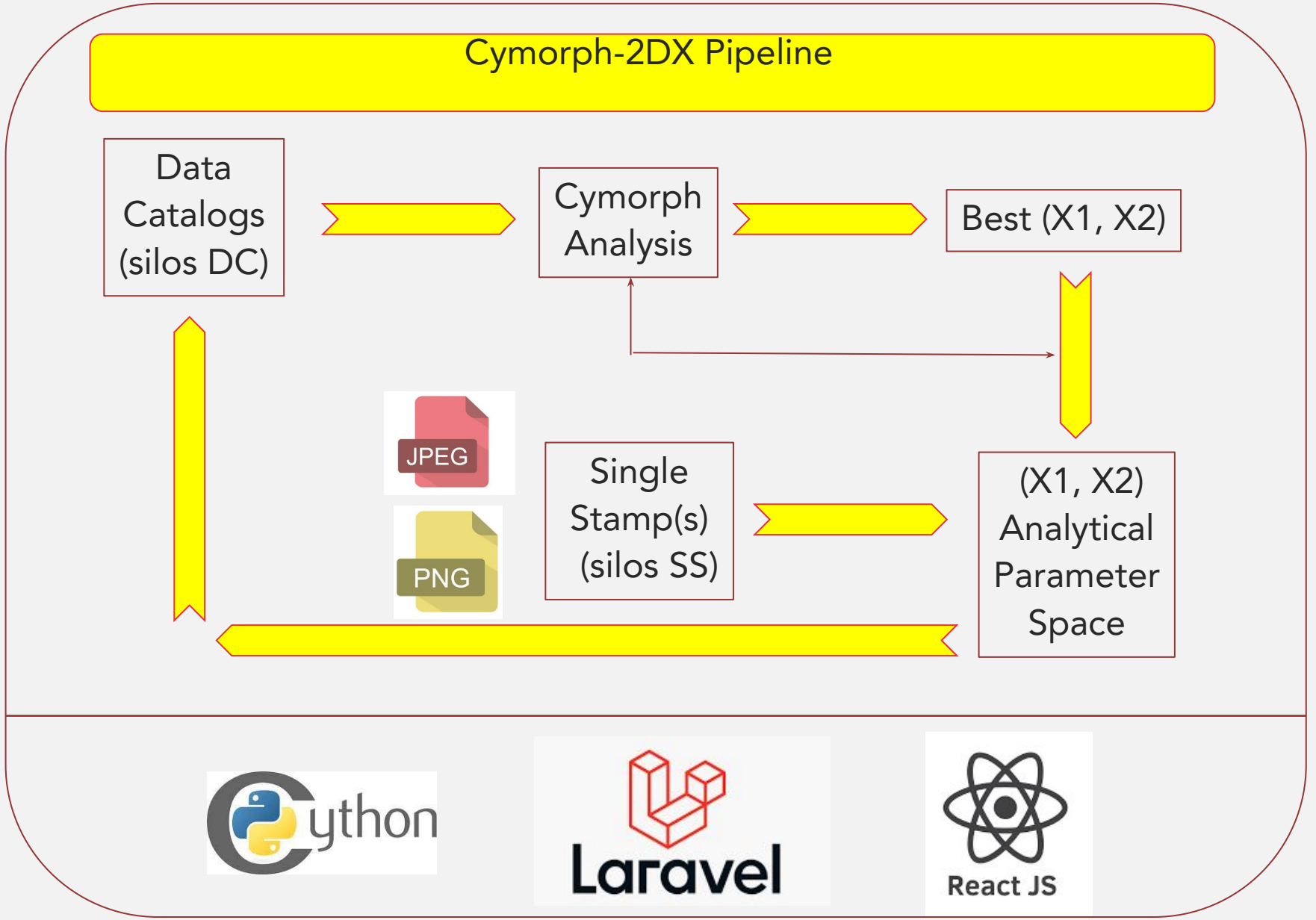




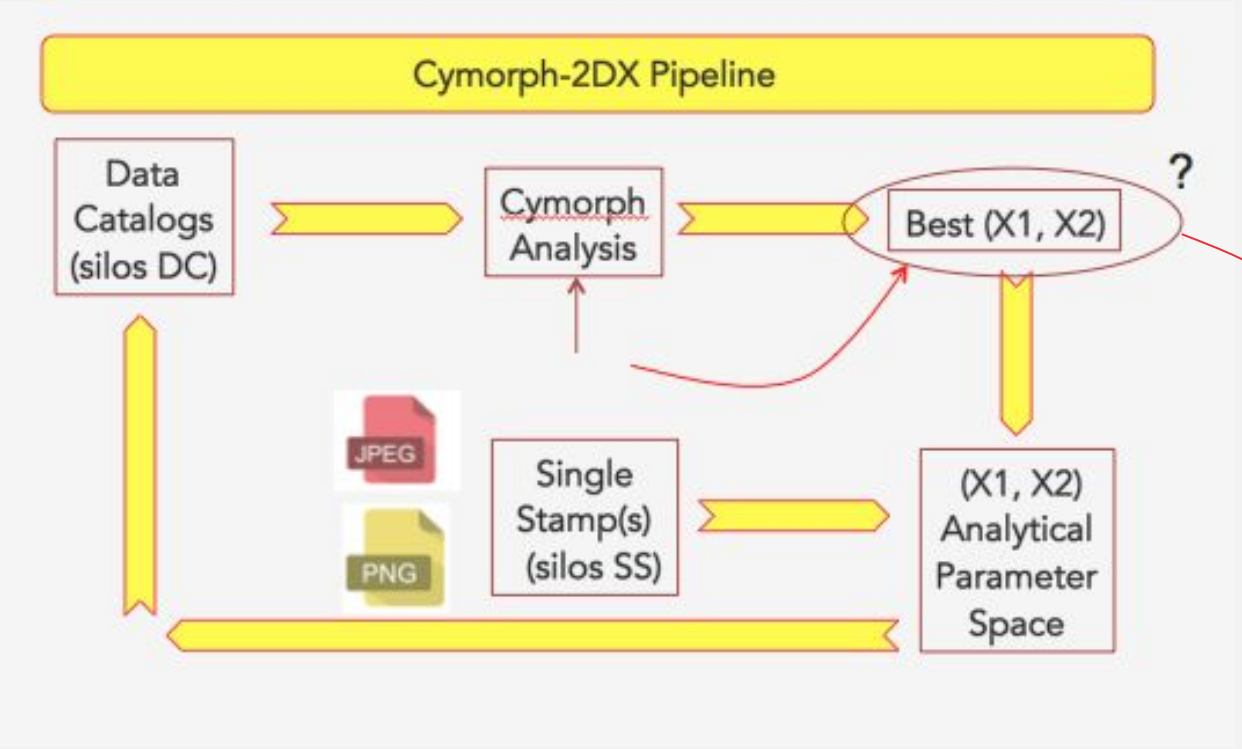


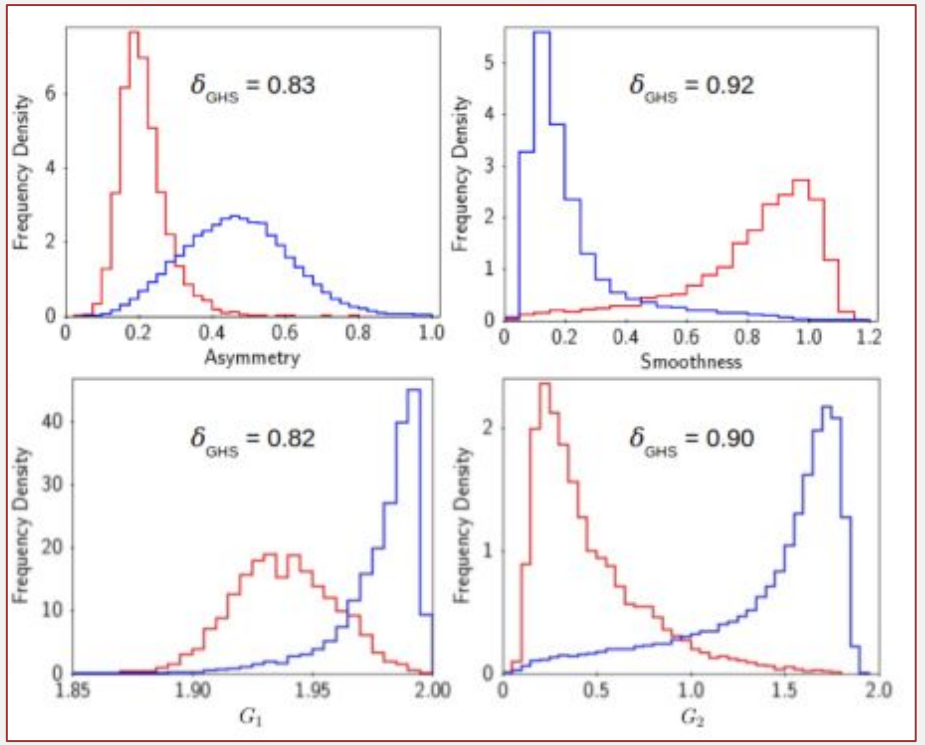


- Current Pipeline



• Current Pipeline



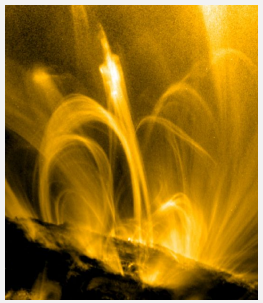


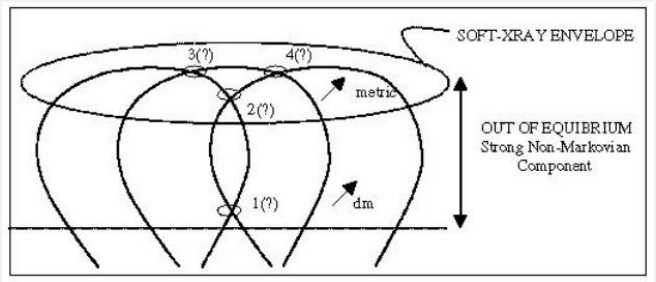
$$\delta_{GHS} = \frac{\left(1 - \frac{A_{BR}}{A_B + A_R + A_{BR}}\right)^{1/2} + \left(\frac{h_B + h_R - 2h_{BR}}{h_a + h_b}\right)}{2}$$

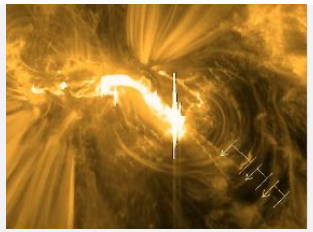
Rosa et al., MNRAS 477, L101-L105(2018)
doi:10.1093/mnrasl/sly054

Novo Módulo (ACDeep): Detecção e previsão de Anomalias

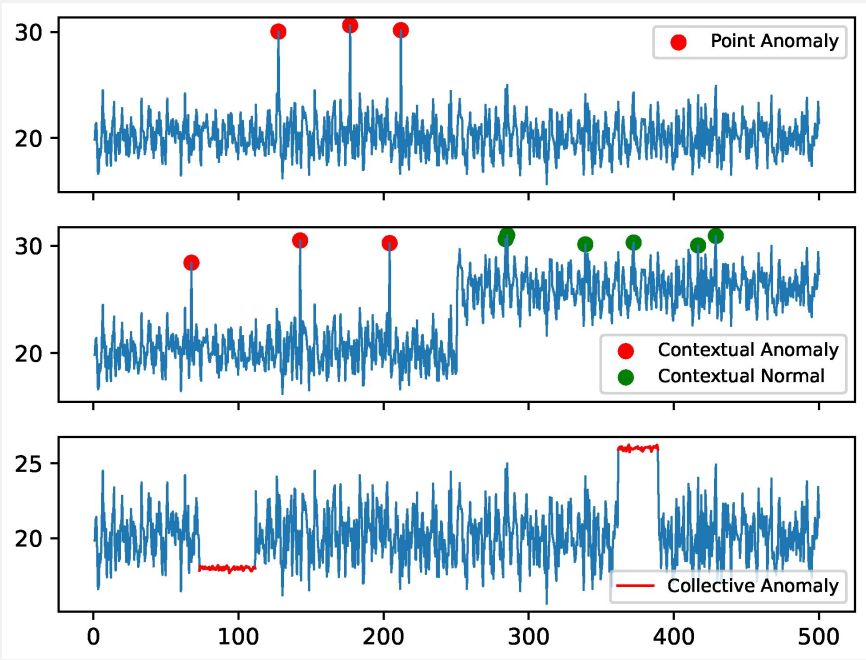
1. Disruption in Plasmas
2. Machine Learning Approaches
3. Data & ML Analysis
4. Results
5. Concluding Remarks

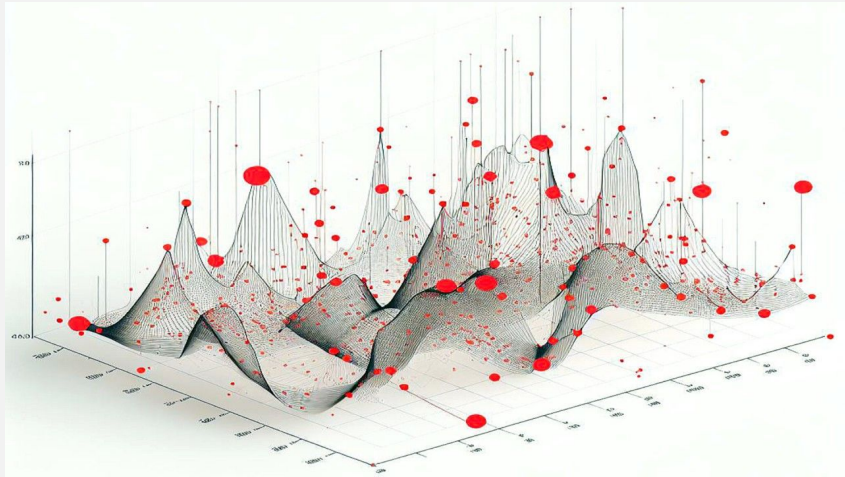






Disruption/Extremes = "Anomalies"

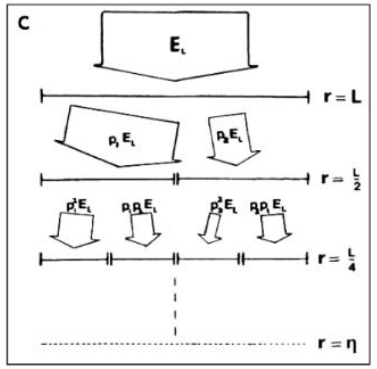


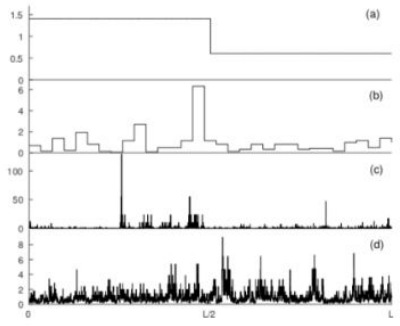


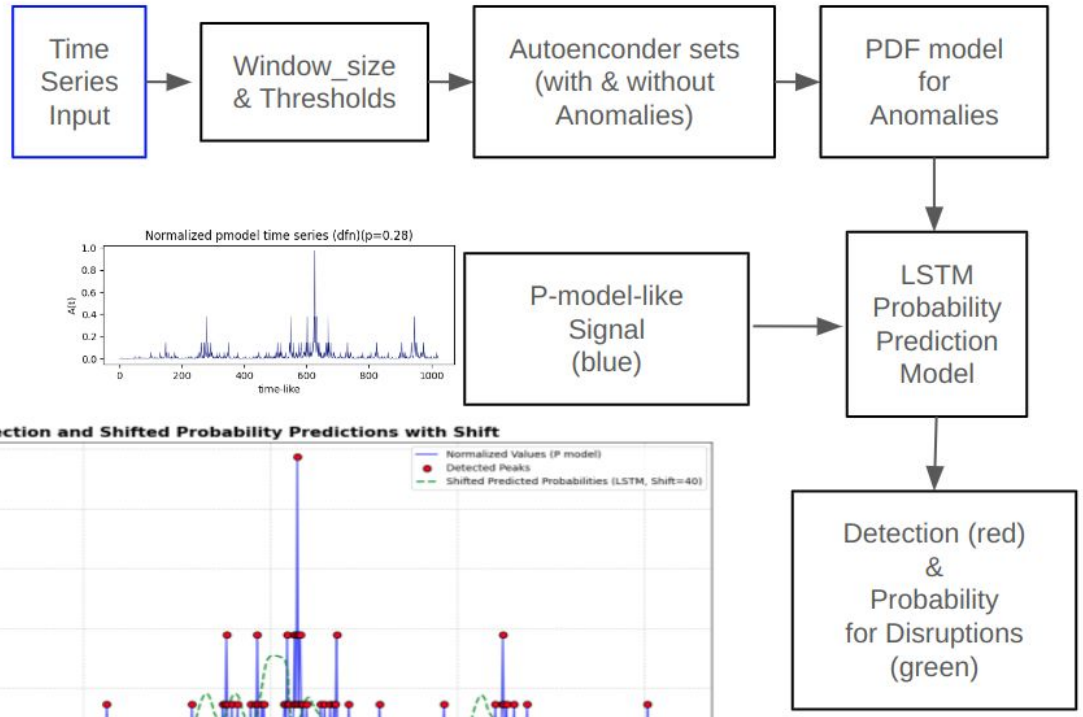
ACDeep

A LSTM-AUTOENCODER
was trained and tested with data generated by p-model.

P-model: Meneveau, C., & Sreenivasan, PRL, 1987







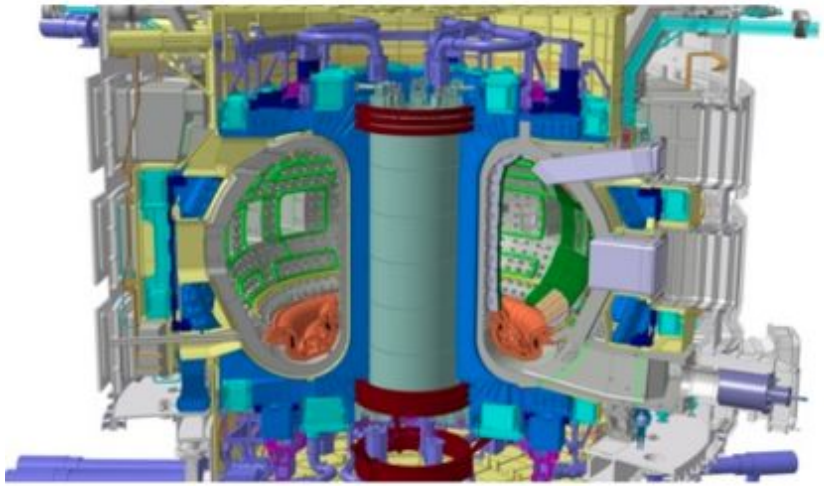

```python
class EVTWeibullAnomalyDetector:
    def __init__(self, window_size, threshold_factor, anomaly_threshold):
        self.window_size = window_size
        self.threshold_factor = threshold_factor
        self.anomaly_threshold = anomaly_threshold
        self.extreme_values = []
        self.weibull_params = None
```

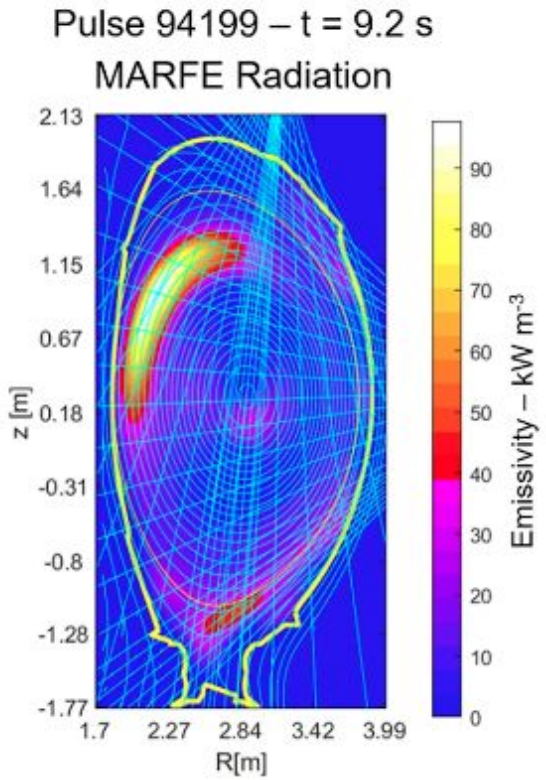
ACDeep

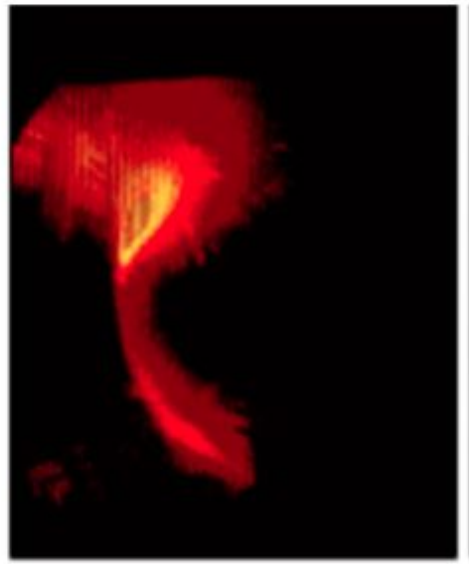
A control oriented strategy of disruption prediction to avoid the configuration collapse of tokamak reactors

Andrea Murari, Riccardo Rossi, Teddy Craciunescu, Jesús Vega, JET Contributors & Michela Gelfusa

Nature Communications **15**, Article number: 2424 (2024) | Cite this article

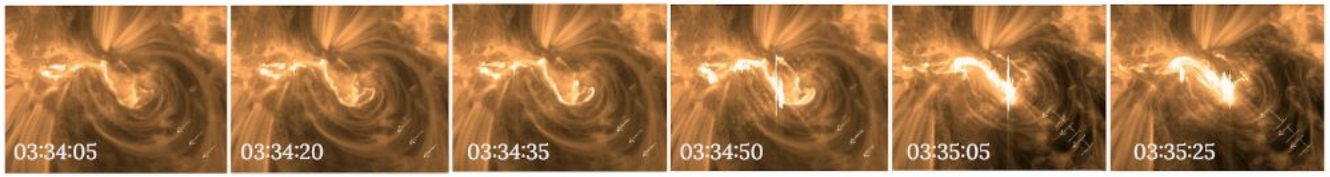


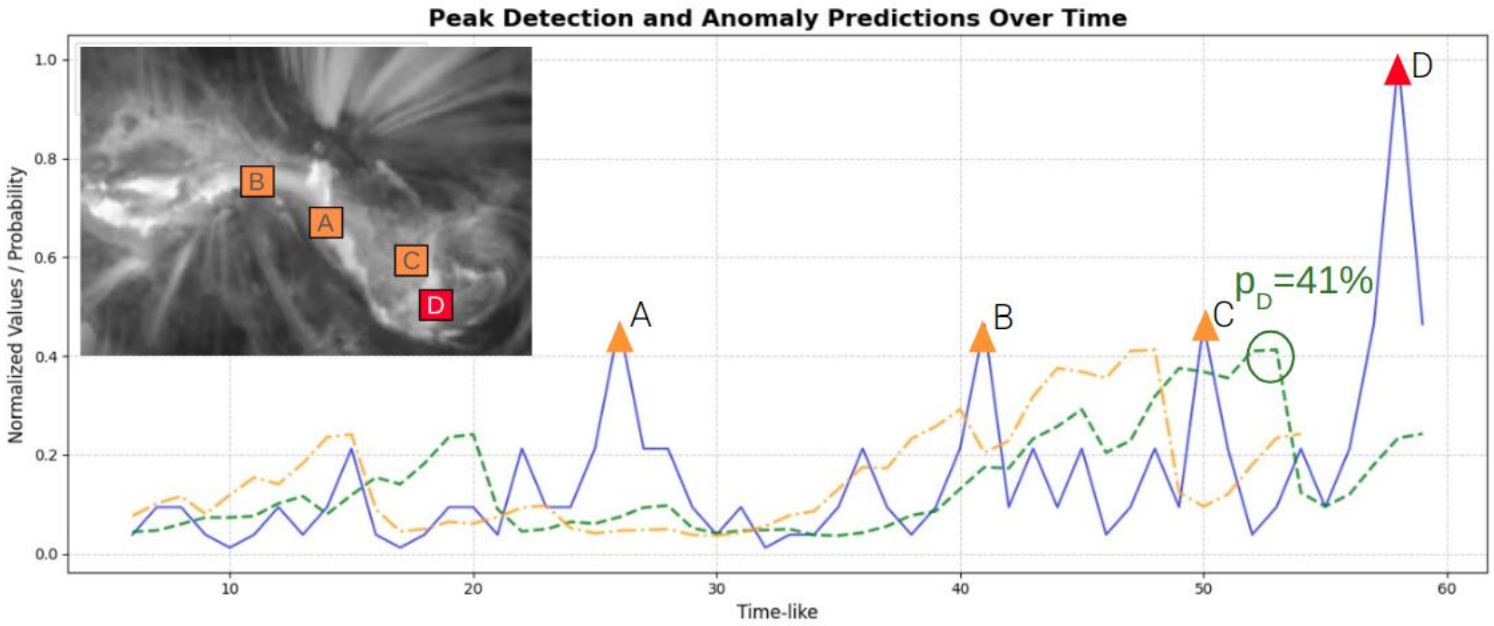




ACDeep

Snapshots from SDO/171A - NOAA 11429 (09/3/12)
03:35:05 to 03:35:25 UT (180 frames)





Solar Physics & Space Weather

Automatically diagnose minor anomalies across multiple channels (or pixels) as quickly as possible to predict a larger systemic anomaly (disruption: a solar flare ⇒ solar noise storms ⇒ magnetic storm).

```python
[3] #SPECIAL TOOLS FOR LSTM-AUTOENCODER - FOR - ANOMALIES PROBABILITIES
    #C. FALANDES-V.10-SEP-2024

    from scipy.stats import weibull_min
    from tensorflow.keras.models import Model
    from tensorflow.keras.layers import Input, Dense
    import numpy as np
    import tensorflow as tf
    from tensorflow.keras.models import Sequential
    from tensorflow.keras.layers import LSTM, Dense, RepeatVector, TimeDistributed
```

```python
[4] # Define autoencoder
    input_dim = 1024  # Example input size
    input_layer = Input(shape=(input_dim,))
    encoded = Dense(128, activation='relu')(input_layer)
    decoded = Dense(input_dim, activation='sigmoid')(encoded)
    autoencoder = Model(inputs=input_layer, outputs=decoded)
    autoencoder.compile(optimizer='adam', loss='mse')
```

```python
[ ] class EVTWeibullAnomalyDetector:
        def __init__(self, window_size, threshold_factor, anomaly_threshold):
            self.window_size = window_size
            self.threshold_factor = threshold_factor
            self.anomaly_threshold = anomaly_threshold
            self.extreme_values = []
            self.weibull_params = None
```

CyMorph + CyMorph2DX + ACDeep

Rubens

Luan

Carlos E. Falandes

Google colab + GitHub

BPG-LSST ISSC

Outline:

- Galaxies (LSST era)
- Morphometrics (Fundamentals)
- Morphometrics (Strategies)
- Concluding Remarks