KOEXAI
The Data Smile

The Koexai Approach to Astro-ML Research: From Physical Models to Machine Learning

ASTRAI as a case study in scalable supernova characterisation

ML4Astro 3 · Valletta · 2026
Luca Naso · Koexai

Applied AI • Machine Learning • Data Science • Astrophysics

PROBLEM RECOGNITION

ML in Astro: more than a model

An astrophysical ML project is a system of interdependent choices, experiments and validation steps, not just model selection.

SCIENTIFIC FRAMING

- Scientific question
- Problem formulation
- Observables & targets
- Physical constraints

DATA

- Access
- Pre-processing
- Dataset construction
- Representation
- Simulation / augmentation

PIPELINE & USABILITY

- Reproducibility
- Automation
- Inference / scalability
- Scientific outputs

MODELLING & EXPERIMENTATION

- ML approach
- Architecture
- Training
- Hyperparameters
- Experiments & selection

EVALUATION & SCIENTIFIC VALIDATION

- ML metrics
- Robustness
- Physical validation
- Interpretation

RESEARCH PROOF

We had to do it five times

Five research projects in the last 2 years

5	**Projects** Selected and completed
5	**Collaborations** UniCT, UniTS, UniRoma2, INAF, INFN
4	**Papers** Peer-reviewed and published

Cascade Grant awarded by INAF, Spoke 3 of ICSC, funded by PNRR

INAF ISTITUTO NAZIONALE DI ASTROFISICA

Piano Nazionale di Ripresa e Resilienza

Funded by the European Union

ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

Projects

1	ASTRAI	Supernovae	H-rich SNe light curves → physical-parameter inference
2	DeepCosmoNet	Large Scale Structure	N-body simulations → subhalos and voids
3	FastParam	GWs from PTA	Bayesian inference → Normalising Flows + statistics
4	GRAIS	Gamma-ray	Fermi-LAT → transients / rare-event detection
5	ESPAI	X-ray	XMM-Newton → background / soft-proton flares

Fundamental research and scientific feasibility, not standalone software products

CASE STUDY

ASTRAI: fast physical characterisation of supernovae

Grasso et al. 2026 · Astronomy & Computing 56, 101116

The astrophysical question

Infer physical properties from observed H-rich SN light curves (LCs).

Use physical models to interpret the underlying supernova physics.

The scaling challenge

Physical modelling is computationally demanding and difficult to scale.

LSST: ~10 million SNe over 10 years.

~1.9 events / minute

The ASTRAI approach

Learn from physics-based models.

Enable fast and reliable characterisation at survey scale.

Developed in collaboration with the group of Prof. Maria Letizia Pumo, University of Catania

CURRENT APPROACH

From a light curve to physical parameters

High-level view of the typical modelling workflow, starting from a bolometric light curve (UniCT).

Extract physical features	Model & constraints	Explore parameter space	Validate & estimate uncertainties
Peak luminosity, characteristic timescales, radioactive tail, expansion velocity when available.	Analytical / semi-analytical model, scaling relations, physically motivated priors.	Repeated model evaluations, likelihood evaluation, HMC sampling	Monte Carlo sampling around the best fit, posterior analysis, degeneracy checks
~ 15 minutes	**~ 15 minutes**	**~ 1 to several hours**	**~ 1 to several hours**

Output: estimated physical properties, such as **R*, Mej, Ek, MNi** (+ CSM properties for more complex models)

Estimated overall timescale: a few to several hours; detailed hydrodynamical modelling can take days.

References: Pumo & Cosentino 2025, MNRAS; Cosentino, Pumo & Cherubini 2025, MNRAS; Cosentino, Inserra & Pumo 2026, A&A

ASTRAI METHOD

ASTRAI: we started from the physics

Physical parameters

R★ Mej Ek MNi

→

Physical Simulation

Use physical model to generate the LC

→

Labelled dataset

14,300 LC-parameter pairs
421 daily points each

The LC-parameter mapping can be used in both directions

Forward direction

Physical parameters → light curve

LCGen learns how to emulate the physical model

Inverse direction

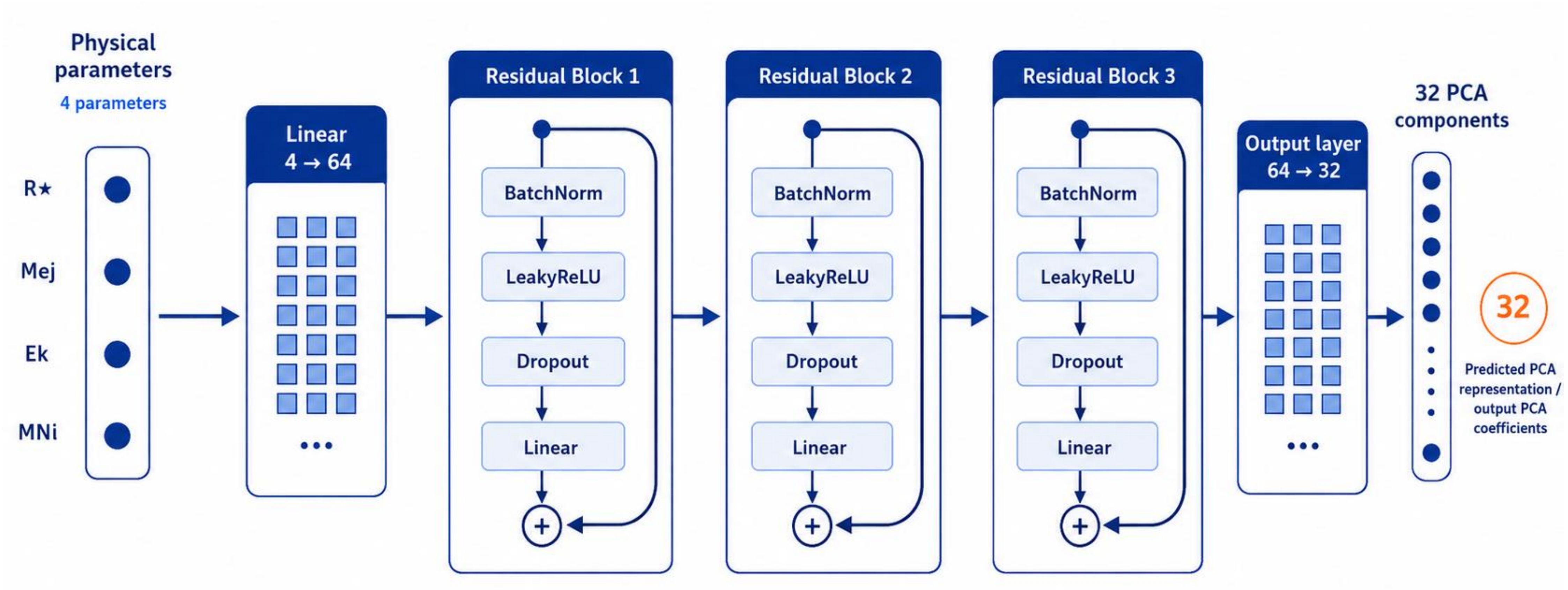
Light curve → physical parameters

PPReg learns how to invert the physical model

The physical model does not disappear: it teaches the system offline.

ASTRAI LCGen

ASTRAI: LCGen Network Architecture



Results: R^2 = 0.989, Relative RMSE = 0.18%; Batch generation rate = 1.4 million LCs per second

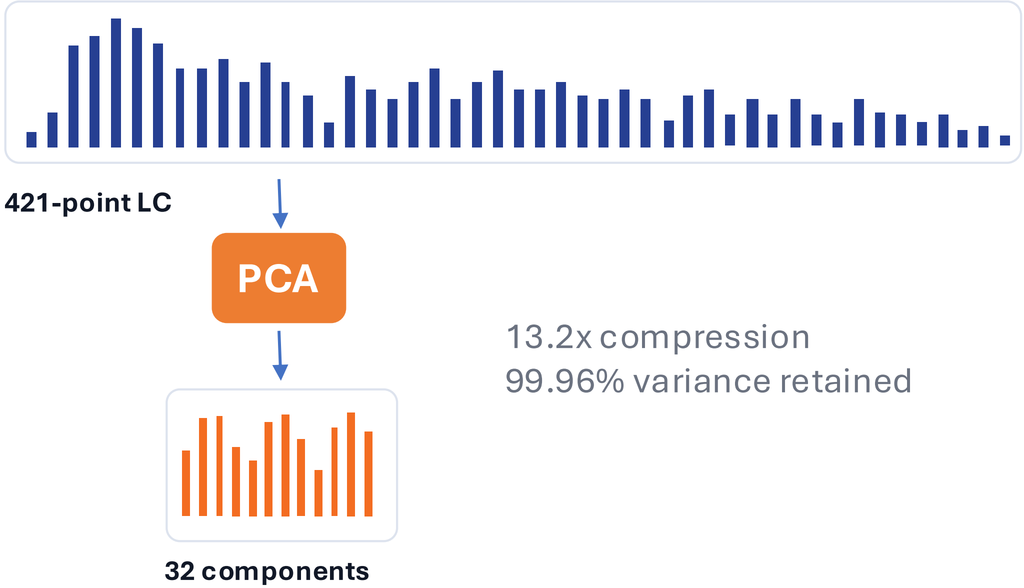
Generation rate is measured on a local benchmark

ASTRAI PPReg

ASTRAI: engineering the inverse problem

Input light curve → PPReg → estimated physical parameters

Data representation



13.2x compression
99.96% variance retained

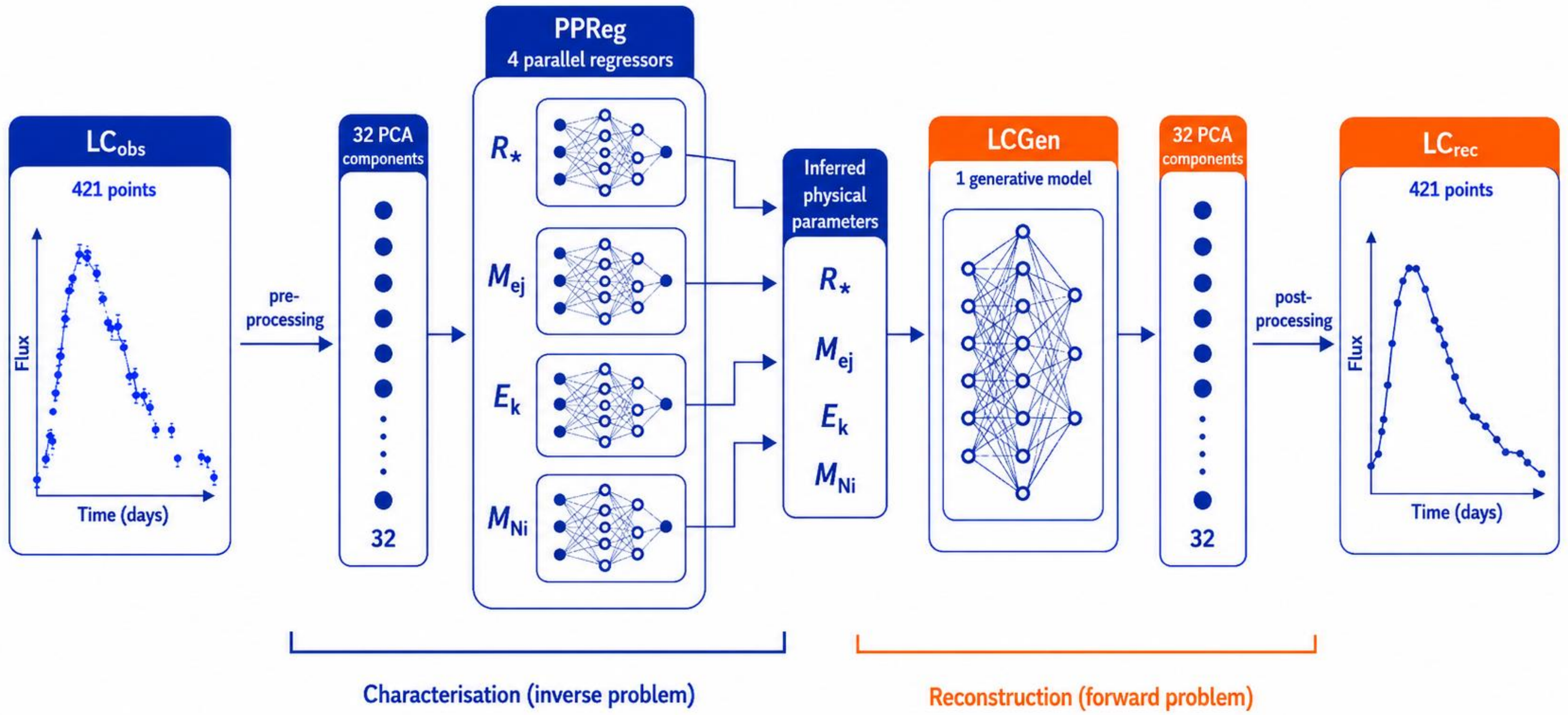
Model formulation

- **Observational realism**
  Noise · visibility gaps · missing data vs idealistic LCs
- **Output strategy**
  Set of specialised networks vs Unified model
  Test to decide
- **ML approach**
  Global regression from LC morphology vs Autoregression
  MLP with Residual Blocks

Results: R^2 ranges from 0.98 to 0.9997; Relative RMSE ranges from 0.9% to 5%

ASTRAI FULL CYCLE

ASTRAI: from observations to physics and back



Full Cycle: observed light curve → inferred parameters → reconstructed light curve

ASTRAI RESULTS

ASTRAI: a real case

Full-cycle execution time

LCObs — PPReg — LCGen — LCRec

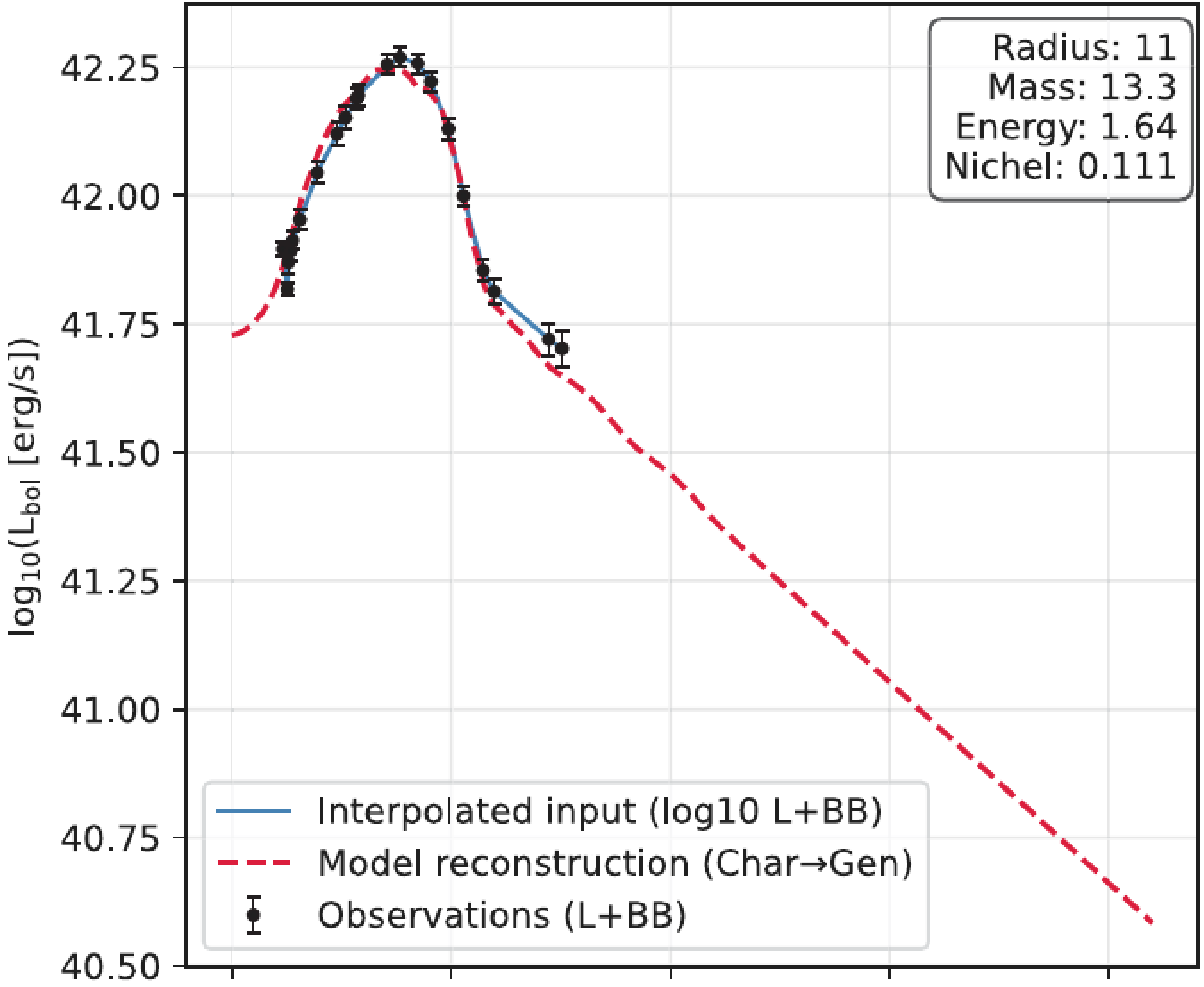
≈ 0.4 ms
per observation

Inverse direction only

LCObs — PPReg

≈ 0.2 ms
per observation

Average values,
calculated over 20,000 cycles



SN 2006V (Grasso et al., 2026, Fig. 12)

COLLABORATION MODEL

Two parties, clear ownership, shared decisions

The collaboration pattern behind ASTRAI can be replicated: each party owns a component, while both work in synergy.

Science Ownership
Your Research Group

- Scientific question
- Physical modelling
- Domain knowledge
- Scientific interpretation

ASTRAI: semi-analytical models, choice of the physical parameters, ...

Shared Decisions

Dealt with together

- Problem formulation
- Validation
- Interpretation of unexpected results

ML Ownership
Your ML Partner

- ML problem formulation
- Data & training pipeline
- Model development
- Experimentation & tuning
- ML validation
- Deployment

ASTRAI: PPReg and LCGen pipelines, experimentation, ...

Thank you

If ML could help your science,
and you don't want to spend your time on the ML work package,
then find your ML partner.

luca@koexai.com

www.koexai.com

References

Grasso et al. 2026 · Astronomy & Computing 56, 101116
astrai.koexai.com
Pumo & Cosentino 2025 · MNRAS 538, 223-242

Spampinato et al. 2026 · Astronomy & Computing 56, 101114
Puglisi et al. 2026 · Astronomy & Computing 57, 101152
D'Amico et al. 2026 · LNCS 16405, 243–258