

A Bridge over Deep Space: Astronomy and AI

ANNA JACYSZYN

FIZ Karlsruhe – Leibniz Institute for Information Infrastructure

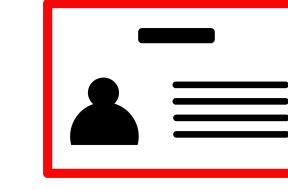
Anna.Jacyszyn@fiz-Karlsruhe.de



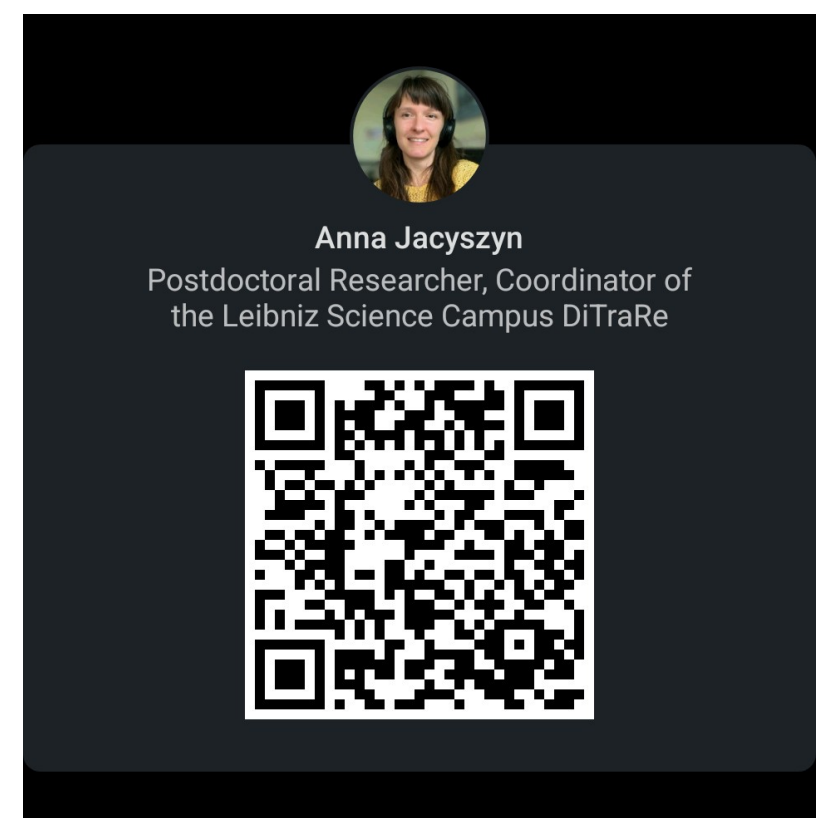
Leibniz ScienceCampus
Digital Transformation
of Research

About Ania

- ★ PhD in astrophysics (3d structure of the Magellanic System based on classical pulsators).
- ★ LSC DiTraRe coordinator.
- ★ Postdoctoral researcher at FIZ KA Knowledge-Driven AI (KDAI) group.



Add contact!



Connect on LinkedIn!

Abstract

- ★ Are you using AI at work often?
- ★ Do you consider yourself an AI expert?
- ★ Do you think that AI experts should collaborate with AI users more often?

Our DiTraRe experience (dimension *Exploration and Knowledge Organisation*) shows that **yes**: this collaboration is needed.

Below I connect my passion: astronomy, with my current work: applied AI. This literature overview: **to computer scientists**, presents the current topics in astronomy where AI is being applied, while **to astronomers**, gives a review of the state-of-the-art AI techniques.

Research Data Management in Astronomy



Highlights

- Rubin Observatory generates about 20 terabytes of data every 24 hours—it would take you over three years of watching Netflix, or over 50 years of listening to Spotify, to use that amount of data.
- It takes just 60 seconds to transfer an image from Chile to California in the US, compare the new image to older images to identify changes, and generate alerts based on those changes.
- At the end of the ten-year Legacy Survey of Space and Time, Rubin Observatory will have generated about 60 petabytes (60,000,000 gigabytes) of raw image data. Considering that the human brain can store about 2.5 petabytes of memory data, that's the equivalent of every single memory of 24 people, combined!



Publication Culture

- ★ Astronomy: a shining example in the fields of **research data management** (RDM) and **open science** (OS).
- ★ Need for unified RDM recognised early and supported by large agencies (i.e. ESA, NASA).
- ★ For decades requirements for making the **data public** and **results open access**.
- ★ **Common standards** well developed and deeply established: catalogues, file types, archives, software packages, etc..

Graphics:
<https://rubinobservatory.org/explore/how-rubin-works/technology/data>.

Virtual Observatory

- ★ Globally recognised platform for RDM: **Virtual Observatory** (VO).
- ★ Gathers databases originating from different instruments and creates a **FAIR system**.
- ★ Workflow modelled after **W3C**, the World Wide Web Consortium.
- ★ **VO Semantics** Working Group: vocabularies and ontologies (symbolic AI).

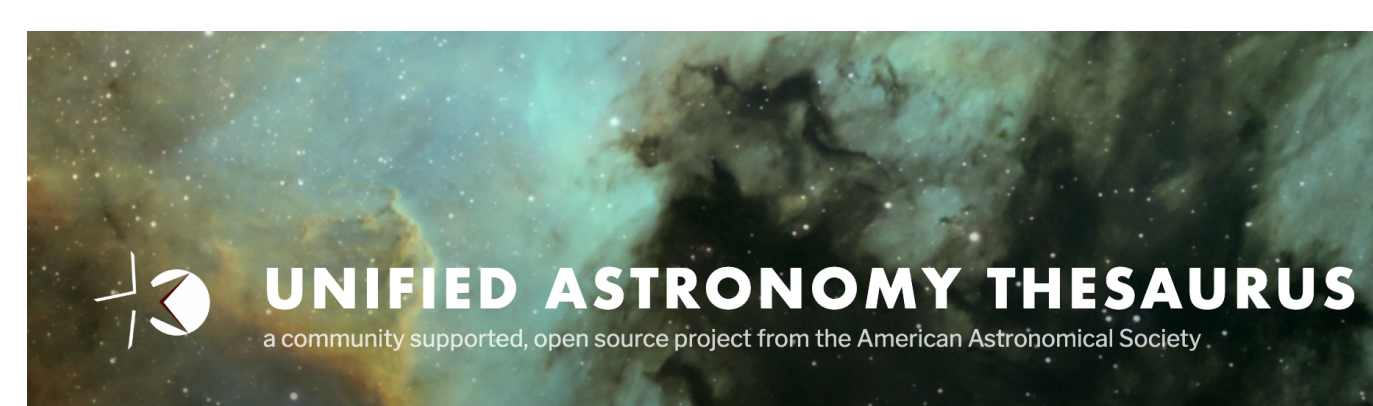
Graphics: IVOA members
(<https://ivoa.net/about/member-organizations.html>).



AI in Astronomy

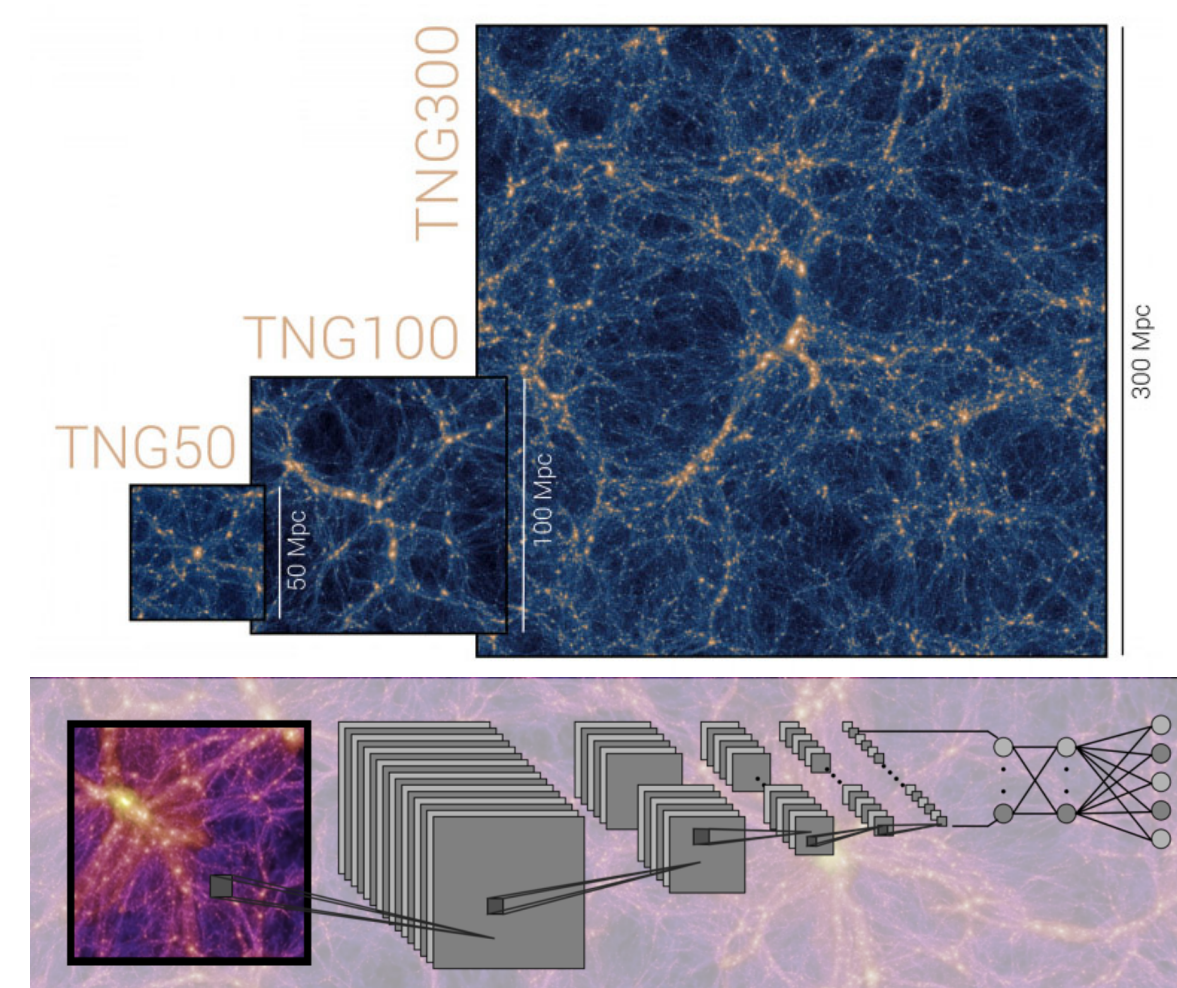
Semantics in Astronomy

- ★ VO vocabularies and thesauri are **limited**.
- ★ Advancements needed in the semantics environment.
- ★ Recently started: OPAL – **Ontology Portal for Astronomy Linked-data**.
- ★ OPAL will combine existing vocabulaires (like UAT) and FAIRify them, it will also expand them into neighbouring fields.



Machine Learning in Astronomy

- ★ Different ML methods are deeply incorporated into the field.
- ★ ML led to multiple **milestone discoveries** in astronomy.
- ★ Example: large volume, cosmological, gravomagneto-hydrodynamical **simulations** of the Universe – Illustris TNG.
- ★ A collection: CAMELS. Cosmology and Astrophysics with Machine Learning Simulations.



Graphics: <https://www.tng-project.org/>,
<https://sites.google.com/site/aicosmo2019/>.



GenAI in Astronomy

- ★ Astronomy-specific **LLMs** have been trained, i.e. AstroLLaMa and astroBERT.
- ★ AstroMLab – a group dedicated to advancing the applications of LLMs.
- ★ Latest models are capable of winning a **gold medal** at the International Olympiad on Astronomy and Astrophysics.

Graphics: <https://huggingface.co/UniverseTBD/astrollama>.

Towards Neurosymbolic AI in Astronomy

- ★ GenAI may provide incorrect results (i.e. hallucinations).
- ★ It is profitable to combine symbolic AI + genAI → neurosymbolic AI.
- ★ Collaborations between astronomers and computer scientists resulting in neurosymbolic methods are soon to be expected to start influencing the field.

Graphics: <https://www.nao.ac.jp/en/news/science/2020/20200811-subaru.html>.

