Working with AI

Fadi NAMMOUR

ekinox

EXTRACT

A distributed data-mining software platform for
extreme data across the compute continuum

CV express

- AI engineering consultant at Ekinox

- PhD in Astrophysics

- Certified trainer

- Some of the projects done using AI and classical methods

  - Galaxy image reconstruction

  - Increasing interpretability in autonomous driving

  - Drone geolocalisation using Computer Vision

  - Radio source reconstruction
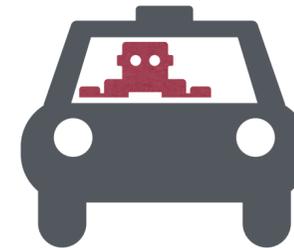
Why AI?

Artificial Intelligence (AI)

- **Image classification**
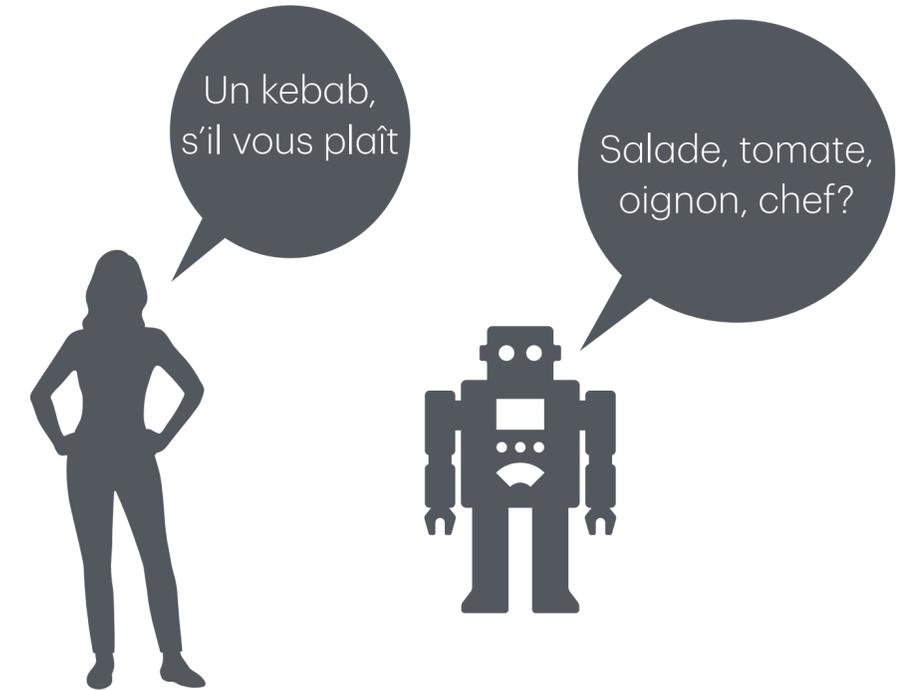 AlexNet wins the 2012 ImageNet challenge
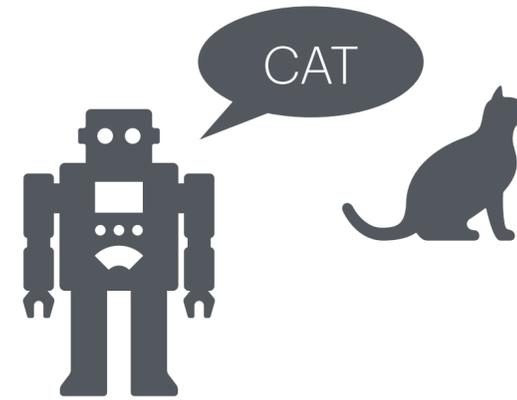
- **Autonomous driving**
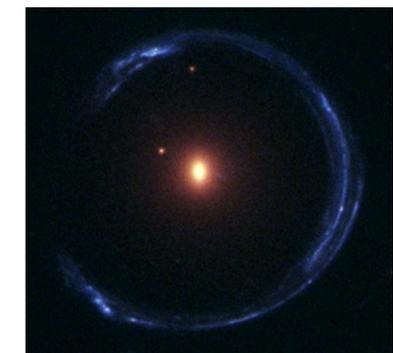 AI is the only hope of reaching full autonomy
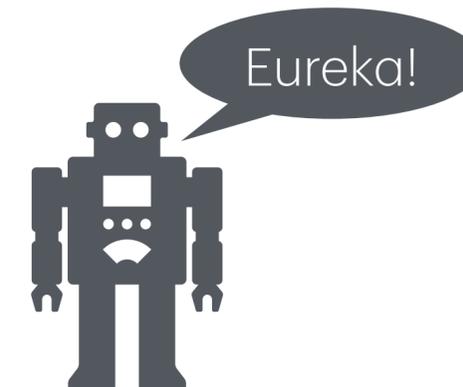
- **Voice recognition**
  AI enabled to mature this technology out of a restrictif use

- **How about astrophysics?**
  AI made the search for strong lensing effects tractable
  with DeepLens [Lanusse et al, 2017]

[Horseshoe lens, Alex Tudorica, flickr, visité le 19/05/2025]

Outline

1. What is AI?

2. When to use AI?

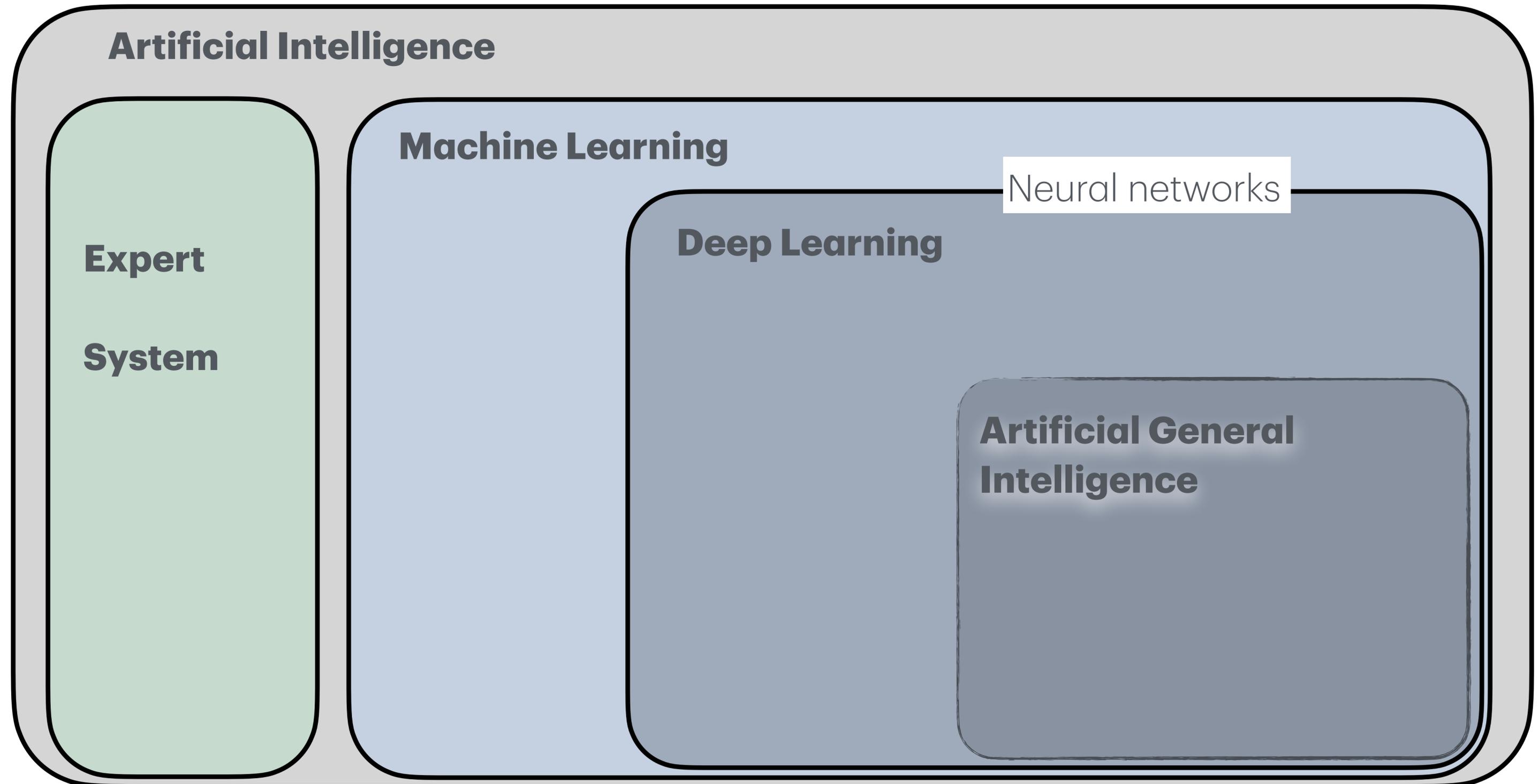3. I'd like to use AI, where do I start?

1. What is AI?

Definition

- Artificial Intelligence is a process imitating human intelligence

- Is a search engine from the 2000s an artificial intelligence?

Source: 15 Years of yahoo.com

AI categories

Expert system

- FranceTerme definition:

  "A set of software programs that uses explicit and organized knowledge in a particular field and can replace a human expert."

- First form of AI, example: DENDRAL, in 1965, for organic chemistry

- Ideal for modelling closed systems with available expertise

Machine Learning

- FranceTerme definition:

  "The process by which an algorithm evaluates and improves its performance without the intervention of a programmer, by repeating its execution on data sets until it regularly obtains relevant results."
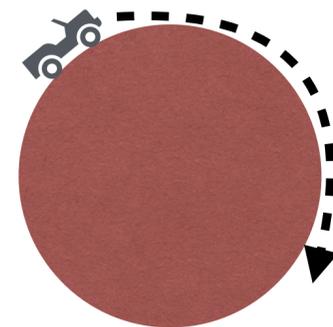
- Categories according to data availability

Supervised learning   Unsupervised learning   Reinforcement learning

Quizz

What is the most appropriate type of machine learning

- Complete a catalog associating astronomical images with the type of objects they contain?

- Create adapted categories to newly observed type of objects?

- To optimise a network of radio telescopes?

Hint:
Available data and answers ⟶ supervised learning
Available data only ⟶ unsupervised learning
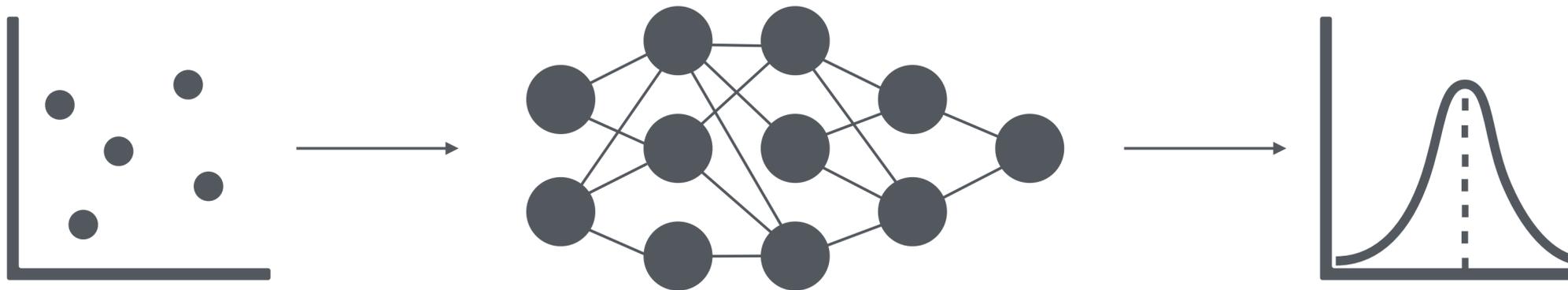Unavailable data ⟶ reinforcement learning

Deep Learning

- Définition de FranceTerme:

  "Machine learning that uses an artificial neural network composed of a large number of layers, each of which corresponds to an increasing level of complexity in processing and interpreting data."

- A neural network is a chain of parametric functions

Artificial General Intelligence

A few definitions

- Amazon:
  "Artificial general intelligence (AGI) is a field of theoretical AI research that attempts to create software with human-like intelligence and the ability to self-teach. The aim is for the software to be able to perform tasks that it is not necessarily trained or developed for."

- Google:
  "Artificial general intelligence (AGI) refers to the hypothetical intelligence of a machine that possesses the ability to understand or learn any intellectual task that a human being can. It is a type of artificial intelligence (AI) that aims to mimic the cognitive abilities of the human brain."

- OpenAI:
  "A highly autonomous system that outperforms humans at most economically valuable work."

2. When to use AI?

AI vs Classical Methods
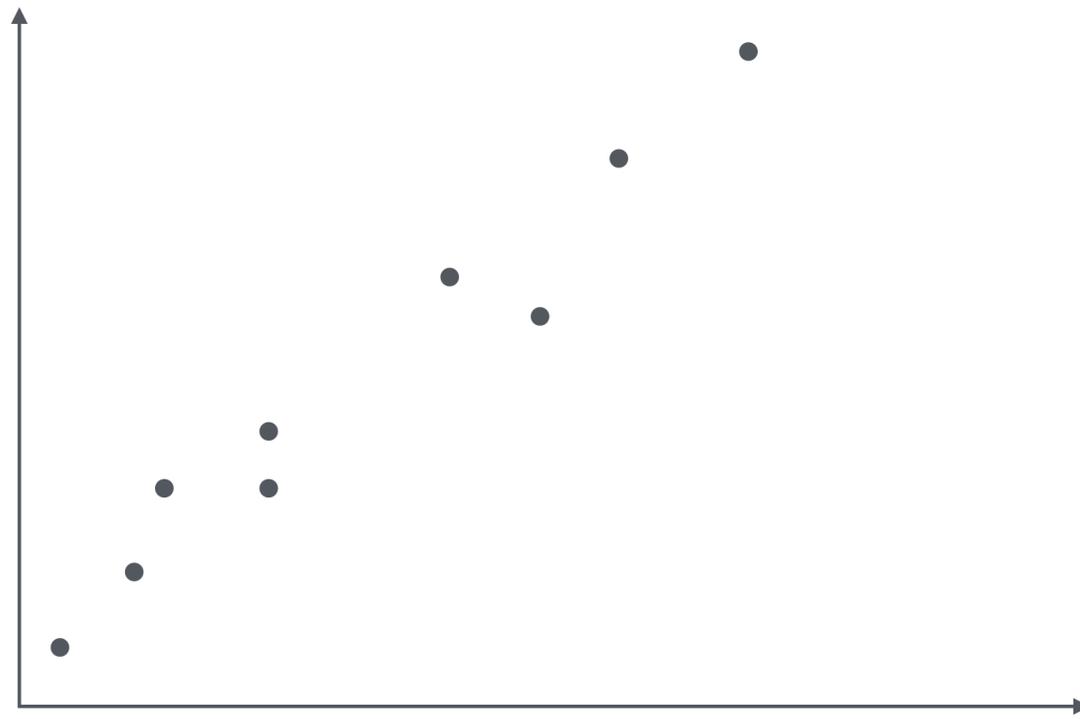
- Choose a polynomial regressor

- Try different orders

- Choose the most adapted order

- Assume a linear model

- Determine parameters

AI vs Classical Methods

Criteria	Classical Methods	Artificial Intelligence
Paradigm	Model based solution	Data driven solution
Bias estimation	Gap with the model	Statistical validation with data
Domain specific expertise	Highly required	Optional
Data availability	Unnecessary	The more the better or access to simulator
Maintainability	Higher chances of expert intervention	Mostly automatic

Bad practices

- Black box approach

- One ring to rule them all (ignoring "No Free Lunch" theorem)

- Starting with the model instead of the data

- Optimise a part before building the whole

- Reinventing the wheel (and not be lazy!)

- Starting directly with an AI solution

3. I'd like to use AI, where do I start?

Why are long general introduction courses to AI a waste of time ?

- The AI field is growing faster and faster,

  "You don't need to learn calculus to do basic accounting tasks."

- Often there is no specific objective in this type of training

- Application examples are often simplistic

Recommandation: see how AI is used in your field (or a related field) and start by taking on a project

# In the boots of data scientist — in our case

Characterization of the problem

- Objective: develop a radio imager capable of reconstructi... ...idelity

- Classical methods are not satisfactory

- Cassey: solution under development

  - Takes into account the temporal structure

  - Adapted to multiband data

  - Sensitive to transient regime

  - Part of a pipeline with a realistic simulator
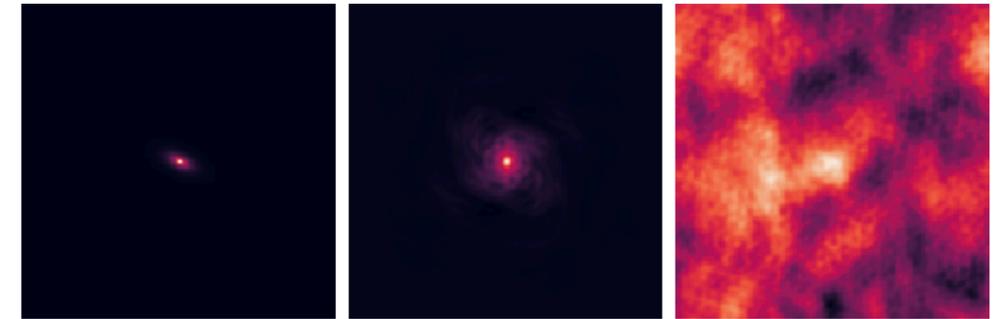
Cassey
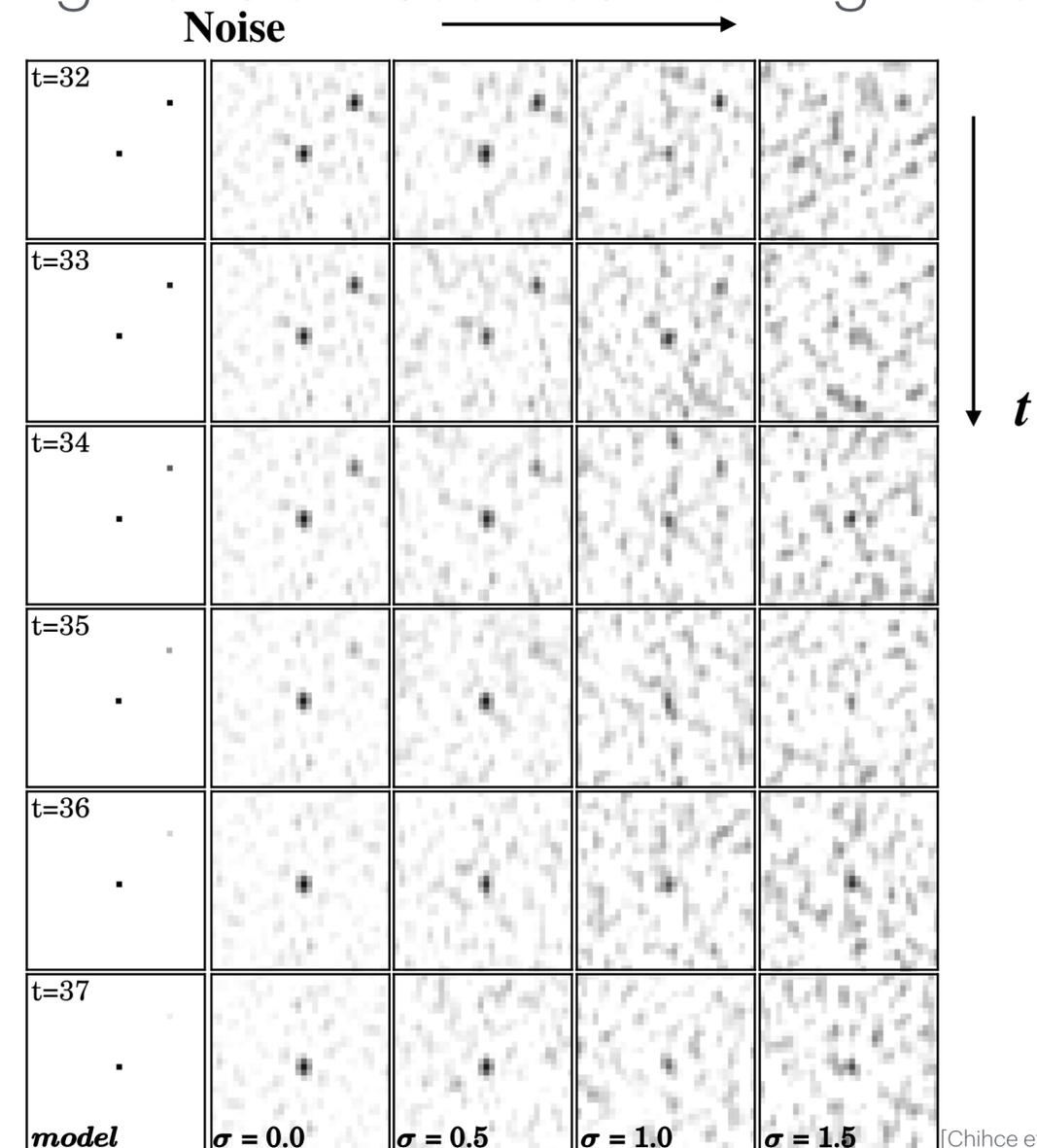
[Chihce et al, 2021]

Step 0: Understanding the data

Images of radio sources

- Images obtained by radio imaging after calibration are noisy and blurred.

[Nammour, 2021]

- The blur operator (PSF) is fully known

→ **First observation**: This is an inverse deconvolution problem with a known kernel

- It is impossible to have images without noise and blur

- We have a perfect knowledge of the image formation process

→ **Second observation**: it is wise to consider using a simulator

Step 1: Literature review & Technology monitoring

List and classify resources according to their ease of integration

- Get a rough idea of the available computing resources to better target the options.

  [LIST OF ACCESSIBLE COMPUTING AND STORAGE CENTERS]

- [imager] Considered solutions:

  - [FIRST CONSIDERED DEEP LEARNING SOLUTION]

  - [SECOND CONSIDERED DEEP LEARNING SOLUTION]

- [simulator] Considered solutions:

  - [FIRST CONSIDERED SIMULATION SOLUTION]

  - [SECOND CONSIDERED SIMULATION SOLUTION]

- **AFTER** choosing the methods to implement, understand how these affect the data

Step 2: Establish a baseline
Aka Proof of Concept (PoC)

- Allows to have a preliminary validation of the hypothesis

- Allows to define metrics and set up visualisations (figures, curves,…)

- Allows to better decouple the scientific effort from the technical effort

- Allows to check that all the components are here and aligned

- Allows avoid major errors

Step 2: Establish a baseline

Aka Proof of Concept (PoC)

Reconstruction results with noise std = 43.64
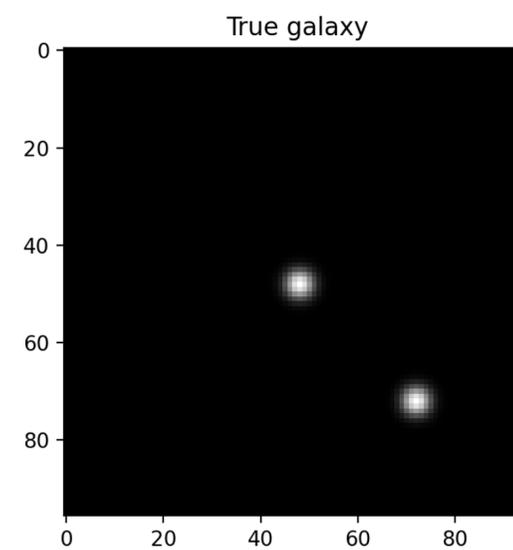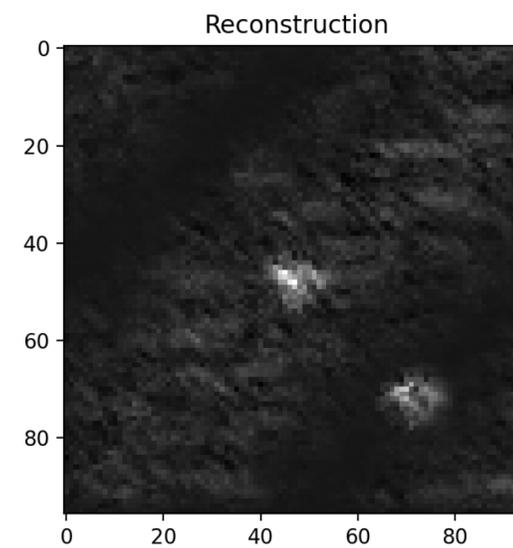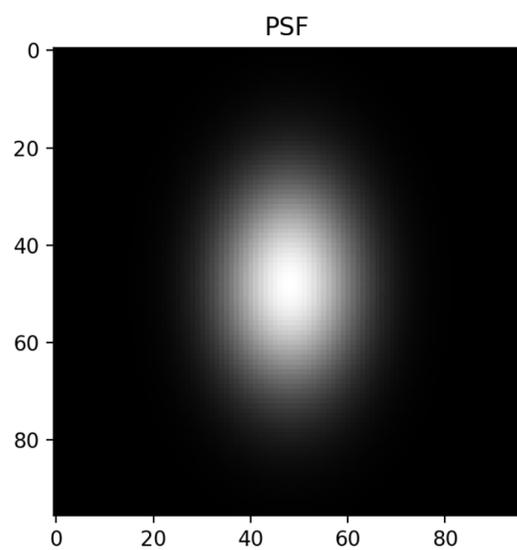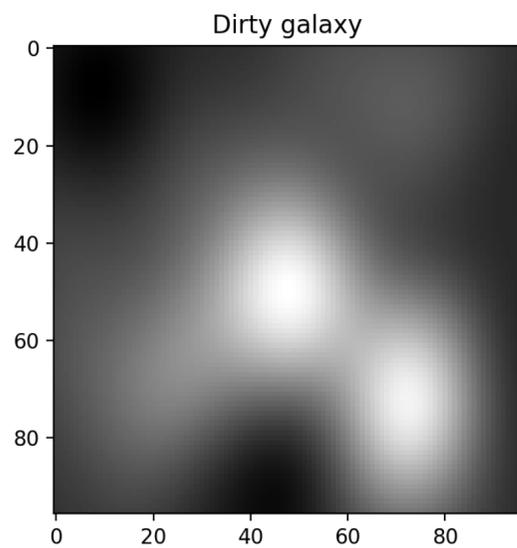

```python
# Create dirty image and dirty beam
dirty_image, dirty_beam = create_dirty_image_and_beam(uvw, visibilities, GRID_SIZE)

GAIN = 0.1
THRESHOLD = 1e-3
ITERATIONS = 500
# Perform CLEAN deconvolution
clean_image, residual_image = clean_deconvolution(dirty_image, dirty_beam, ITERATIONS, GAIN, THRESHOLD)

# Plot the results
plot_images(dirty_image, clean_image, dirty_beam, residual_image)
```

Step 3: Continuous improvement
Until reaching a satisfying or conclusive result

1. Choose the most important part to improve

2. Apply steps 0, 1 and 2

3. Repeat

Tips on best practices:

- "Clean code" for a maintainable and flexible code

- Use Git for versionning for better collaborations and code traceability

- Agile/Scrum mindset for quick and safe advancement in a heterogeneous team

Be the "I" in AI!