ANITI

Communauté d'universités et établissements de Toulouse

ANITI Days 2026

Reinforcement Learning & Robotics Foundation Models: the Revolution

Airbus Robotics & AI Research

Alexandre ARNOLD, Olivier COUDRAY

AIRBUS

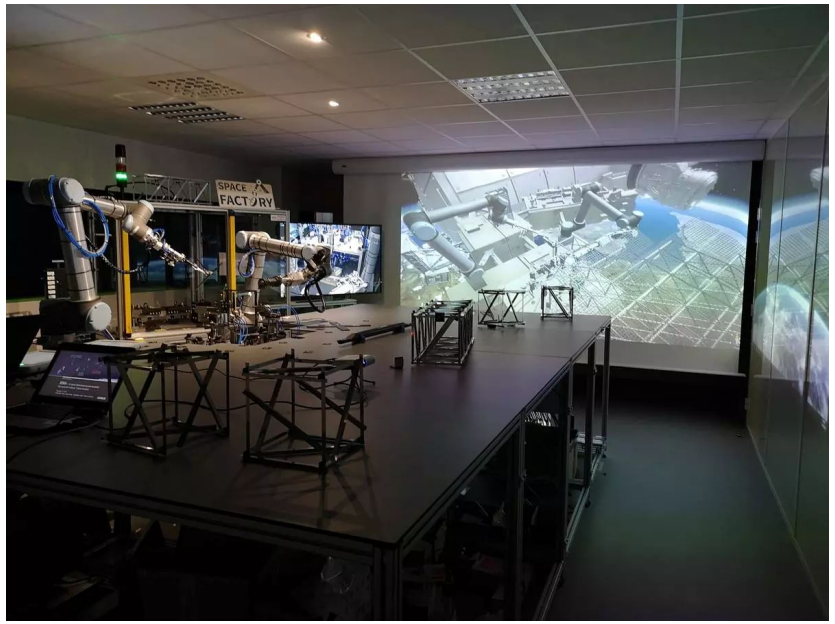
ANITI Robotic needs at Airbus

Airbus Commercial
Support ramp-up

> complement humans for repetitive tasks
(today only 7% of automated operations)

Airbus Defence & Space
Unlock new capabilities

> like in-Space manufacturing
(robotic reflector kit assembly in orbit)

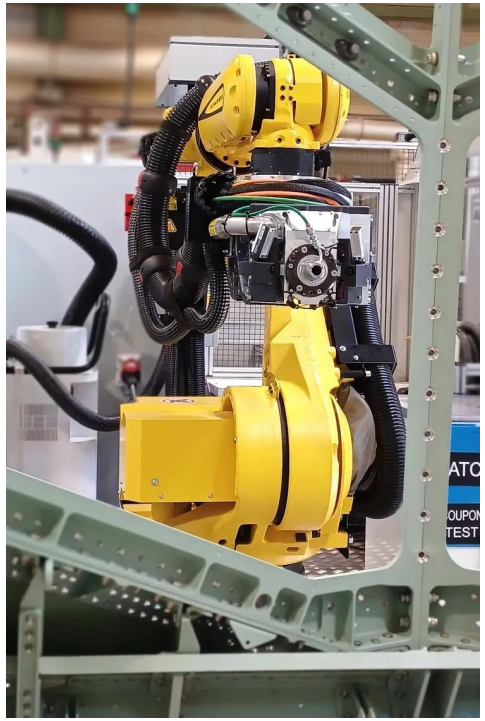


AIRBUS

ANITI Airbus Robotics: design industrial robots

Developing robotics expertise in-house with Airbus Robotics

→ internal development & integration of robotics
→ strategic acquisitions (e.g. MTM Robotics)



The Medium-Sized Drilling Robot

The Medium-Sized Drilling Robot (MSDR) is Airbus' first robot developed completely in-house. The MSDR was designed to be three times smaller than the industry standard in order to permit its integration into existing pre-assembly lines, and it is more accurate and more agile than non-customised robots available on the market.

The robot covers 87% of all pre-assembly-line drilling needs and is intended for use on fuselages, horizontal and vertical tail planes, and centre wing boxes.

Flextrack robotic systems

The Flextrack is a drilling robot primarily used on A320 Family fuselage pre-assembly lines. Flextrack robotic systems have either flexible or stiff rails that can be easily disassembled and reassembled around an aircraft, allowing lightweight and modular drilling robots to independently move alongside the fuselage.

Eventually, all existing and new single-aisle pre-assembly lines will be equipped with Flextracks, and their use on final assembly lines will be progressively increased.

AIRBUS

ANITI Robotic challenges in aerospace industry

- Huge product size
- High degree of customization per product
- Very high safety standards
- Fine-grained required accuracy…

→ Need for adaptive robots to address more use cases, e.g. kitting:
packing (few hundreds of) components from storage to briefcase

5-10 min

AIRBUS

ANITI Towards more flexible robots

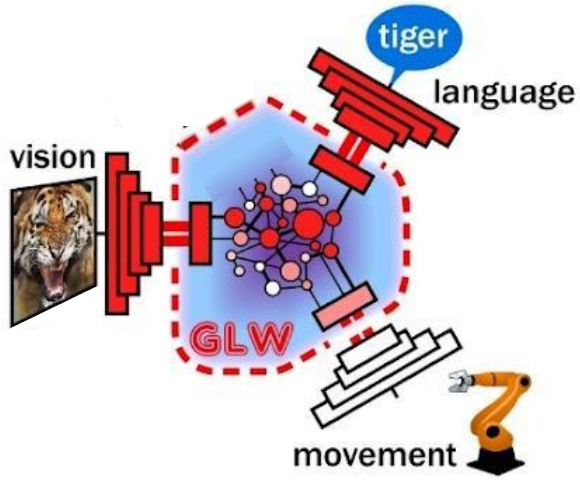
Airbus AI Research works with the C3PO* synergy chair in ANITI towards:

- **Adaptive robots**
  → to operate in environments & products as-built and not just as-designed
- **Communicative robots**
  → to dynamically adapt their behavior according to human instructions

C3PO

*** Cobots with Conversation, Cognition & Perception**

- Chairs: R.VanRullen (CerCo), N. Asher (IRIT), T. Serre (Brown), O. Stasse (LAAS)
- Frugal multimodal robotic systems with grounded perception, language and action



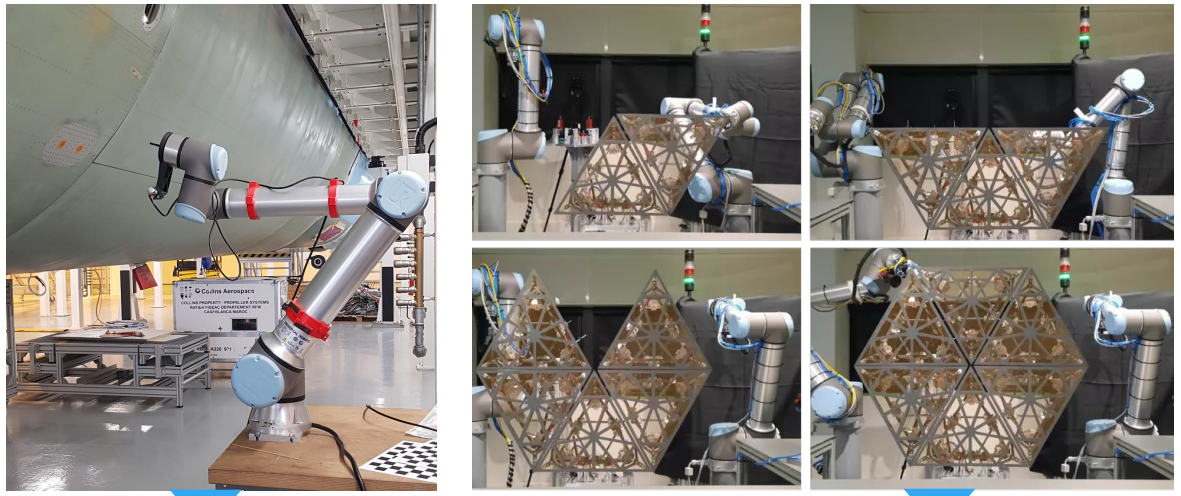
AIRBUS

ANITI Digital twins for safe AI training of robots

Requirements to simulate industrially-inspired robotics scene with camera in the loop:

- Robotics-oriented simulator
- Realistic physics engine
- Realistic visual rendering
- Highly parallel physics & rendering
- 3D import from usual CAD software
- Ability to connect to real robots

→ Use of NVIDIA Isaac Sim/Lab (built on Omniverse)



Aircraft manufacturing & inspection on ground

In-space assembly

AIRBUS

ANITI Research axis: RL with Perception in the Loop

Deep Reinforcement Learning with a focus on multimodality & data-efficiency (DrQ v2, Dreamer v3, GW-Dreamer…)

→ end-to-end approach (no complex pipelines)
→ high-dimensional inputs (camera + proprioception)
→ learning from scratch to find/reach various targets (with random pose)

OBSERVATIONS:

Mounted camera view

+

Joints pos/vel

AIRBUS

ANITI Research axis: Foundation Models for robotics

Pretrained Foundation Models for robotics inspired by LLMs

→ avoid retraining from scratch (lots of knowledge backed-in)
→ hope of zero/few-shot generalisation (towards generalist robots)

LLM

Large Language Models (often pretrained on web-scale text data, then finetuned with dialog data & RLHF) paved the way for *Foundation Models*: reusable on many domain-specific tasks via modest finetuning.

IN: text
OUT: text

VLM

Vision Language Models (often based on a LLM + vision adapter) enable taking into account images along with text. Sometimes their text output can be interpreted as a (robotic) action, effectively becoming a VLA.

IN: vision + text
OUT: text (action)

VLA

Vision Language Action models (often based on a VLM) enable directly predicting robotic actions or action chunks, given current observations (vision & sometimes proprioception) and a task textual description.

IN: vision + text
OUT: robotic action

WAM

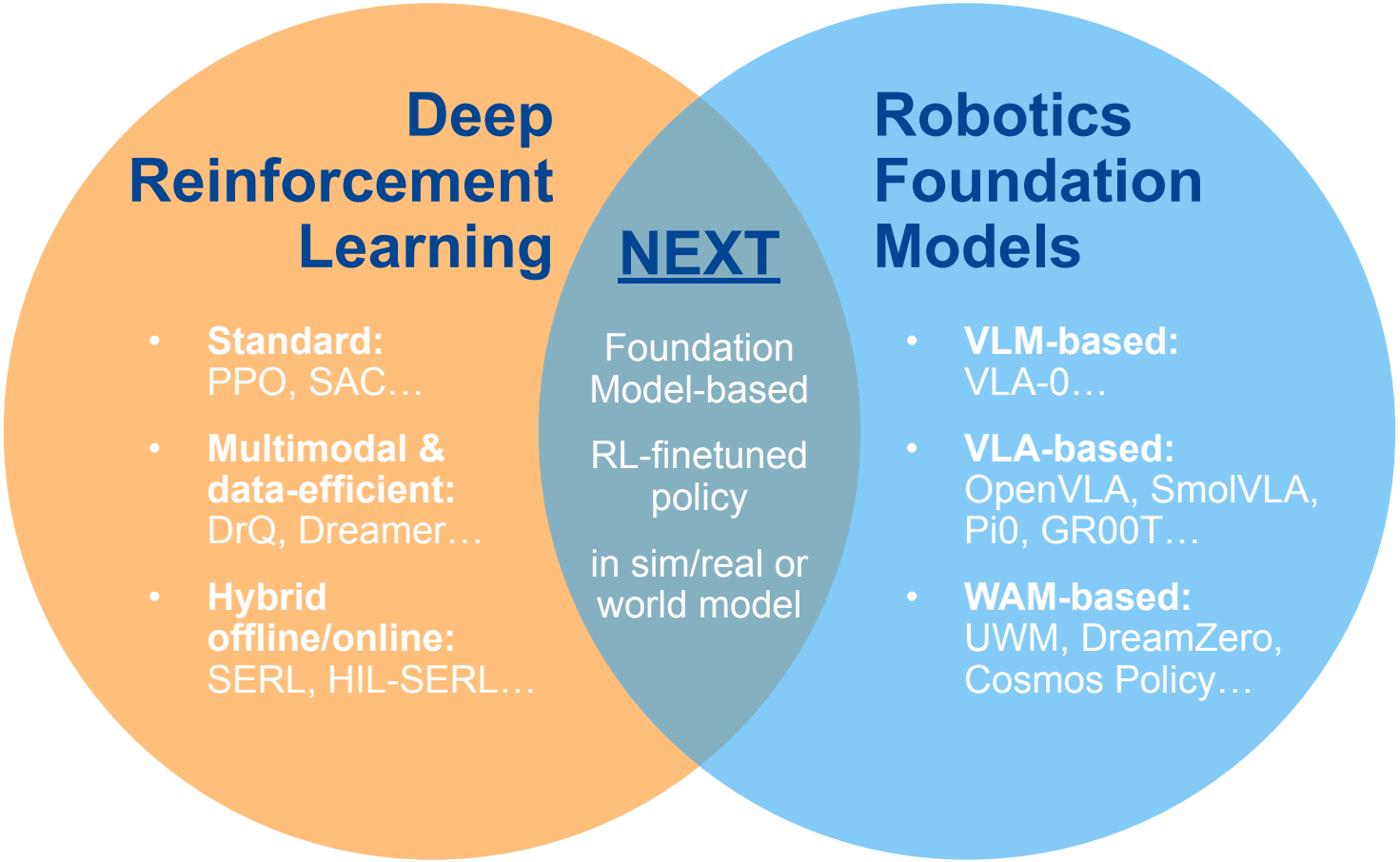
World Action Models (often based on a video generation model or a world model) also predict next state in addition to the next action. This enables better action planning (or even visualization before operation).

IN: vision + text
OUT: next state + action

Inspecting objects on-the-fly without retraining

AIRBUS

ANITI Research perspective: best of both worlds



Start from Foundation Model (pretrained)

Add RL finetuning (limit teleoperations)

AIRBUS

[Backup slide] Another challenging use case

Vacuum bagging: position a deformable plastic foil on the skin & stringers to generate an isolated atmosphere