cortAIx

Artificial Intelligence by THALES

# AI & SIMULATION

THALES

Building a future we can all trust

# AI & Simulation

Frédéric BARBARESCO, Jérôme BERNIER, Bruno DELACOURTE, Jean-Yves DONNART, Jean-François GOUDOU,

Christophe LABREUCHE, Catherine LAMY-BERGOT, Pascal MAC, Emmanuel MICONNET, Jonathan SPRAUEL, Mathieu YTIER

## Summary

AI & Simulation (AI&S) integrate traditional simulation models (such as physics-based, agent-based, and system dynamics models) with advanced artificial intelligence technologies (including machine learning, reinforcement learning, and black-box optimization) to enhance their accuracy, efficiency, and adaptability.

This paper examines how the concept of “AI&S” enables significant advances in simulation through the application of artificial intelligence technologies. It demonstrates how AI can enhance simulation capabilities, enrich learning processes based on simulated environments, and support the development of sophisticated models commonly referred to as hybrid digital twins. Conversely, simulation technologies can improve AI by enabling the generation of relevant data or the simulation of realistic environments to build AI models.

The most immediate benefits of AI-driven simulation are observed in training applications, where trainees and instructors require greater autonomy to animate realistic simulated entities and to efficiently process the large volumes of data generated. Comparable advances are expected to become increasingly valuable in the near future for wargaming applications and for operational decision-support tools.

AI&S accelerates simulations and improves efficiency by using surrogate models that approximate complex computations in milliseconds, enabling rapid exploration of many scenarios for design, planning, and risk assessment. Through hybrid approaches such as Physics-Informed and Physics-Aware Machine Learning, AI combines physical laws with data-driven corrections, reducing modelling errors and continuously improving accuracy as new data becomes available.

AI&S also manages high system complexity by handling nonlinear, high-dimensional interactions and revealing emergent behaviors often missed by traditional models. It supports real-time decision-making through interactive, human-in-the-loop simulations and enables hybrid digital twins that adapt to specific contexts or users. Overall, AI&S lowers costs by reducing reliance on physical prototypes and minimizing computational and material waste in R&D.

# 1. Introduction and Context

AI-based augmented simulations are faster, more accurate, adaptable to new data, capable of managing complexity, and conducive to real-time decision-making—while simultaneously reducing costs and enabling the exploration of previously intractable problems.

The simulator has become an essential tool in the development of new systems, which can be confronted with simulated realities. Armored vehicles, drones, and multi-environment robots are now modelled by dynamic digital twins, which can be seamlessly immersed in tactical contexts of varying complexity.

By leveraging game engines or dedicated simulation tools, the developed virtual environments provide a realistic understanding and facilitate the design, development, and testing of equipment. These systems can, therefore, be exposed to simulator-generated scenarios, whose complexity or diversity varies depending on the specific objectives at hand. Furthermore, such simulations play a crucial role in the training of military personnel by immersing them in environments that mirror the operational situations they may face, while also assessing the acquisition of operational expertise.

In an era of rapid advancements in Artificial Intelligence (AI), it is therefore pertinent to question the contribution of this emerging technology to the simulation tools outlined above, and conversely, the role that simulation can play in the design of AI.

### *AI FOR SIMULATION*

This position paper develops 2 kinds of applications of AI for Simulation:

- AI FOR AUGMENTED TRAINING
- AI-BASED DESIGN BY SIMULATION

AI has the potential to enhance simulations, particularly through its contribution to the creation of realistic scenarios and the credibility of digital actors. AI enables the confrontation of complex systems during the design phase, before their actual deployment, within a reactive and relevant environment where their means of action can be simulated. This allows for the identification of the potential strengths and weaknesses of these systems, as well as an evaluation of their performance, both in terms of technology and operational doctrine. The same applies to systems of systems or multi-environment, multi-function swarms, whose employment doctrines can be tested, thereby saving valuable time that would otherwise be spent on operational field tests.

Furthermore, AI is also applicable to human training, enabling individuals to interact with artificial agents whose behavior is rendered intelligent. This empowers them to learn how to react to complex situations. An instructor utilizing such a tool will be able to assess a trainee's tactical training level, the coordination skills of a commander responsible for available resources, or the mastery of procedures by staff members. AI will facilitate the diversification of learning and role-playing scenarios, providing an expansive range of possibilities (non-compliant cases, attrition, mission changes during an operation), and offering automatic responsiveness to events. This represents the ultimate goal of any immersive training environment.

### *SIMULATION FOR AI*

This position paper develops 2 kinds of applications of Simulation for AI:

- SIMULATION FOR REINFORCEMENT LEARNING
- SIMULATION FOR DATA GENERATION

Simulation in reinforcement learning serves as a surrogate environment for agents to explore and learn optimal policies without the cost or risk of real-world deployment. It allows for the rapid generation of diverse scenarios, accelerating agent training by exposing it to a wide array of conditions. Simulated environments can be finely tuned, enabling the exploration of edge cases and rare events that would be costly to reproduce in reality. Ultimately, simulation aids in the development of robust RL models, facilitating faster deployment and better generalization to real-world tasks.

Simulation for data generation creates synthetic datasets that mimic real-world phenomena, especially when real data is scarce, expensive, or difficult to obtain. It enables the exploration of a vast range of conditions and rare events that might be underrepresented in real-world datasets. By adjusting parameters and randomizing inputs, simulators can produce large volumes of diverse data, enhancing the robustness of machine learning models. Synthetic data generated through simulation can be used to train models, test algorithms, and validate system behaviors under controlled yet variable conditions. Ultimately, it supports the development of predictive models by supplementing or augmenting real-world data with highly tailored, simulated instances.

Another area of investigation, perhaps less intuitive but equally pertinent, involves reversing the question and exploring the extent to which simulation can serve as a tool to assist in AI design.

Connectionist AI is data-intensive, and reinforcement learning relies heavily on vast quantities of simulations. In both instances, AI&S can generate the training data required, known as synthetic data. Furthermore, AI&S can be utilized to create test environments for an AI, allowing its behavior to be validated and certified within specific contexts.

### *AI-ENHANCED HYBRID DIGITAL TWIN*

A digital twin is a virtual representation of a physical object, system, or process. It uses real-time data from sensors and other sources to mirror the state of the physical counterpart, allowing for real-time monitoring and analysis.

Hybrid AI-based digital twins combine the interpretability of physics with the adaptability of AI, ushering in a new generation of self-improving, predictive, and real-time virtual representations of complex systems.

This position paper develops 3 kinds of applications of hybrid digital twins:

- DIGITAL TWIN FOR DESIGN OF SYSTEM OF SYSTEM
- DIGITAL TWIN FOR DESIGN OF EQUIPMENT
- DIGITAL TWIN FOR USE IN OPERATION

A hybrid digital twin, driven by AI, is an advanced iteration of digital twin technology that integrates physics-based models with data-driven AI/ML models, resulting in a more accurate, adaptive, and predictive representation of complex systems.

Hybrid AI fosters disruptive intelligent modelling, enabling the development of Hybrid Twins with minimal resources, as well as intelligent control, through optimizations-based techniques, algorithms, and prototypes that support optimal decision-making in complex infrastructures.

AI is incorporated to overcome the limitations inherent in purely physics-driven twins. Physics-Aware Machine Learning can bridge gaps in physics-based knowledge, learn system behavior from operational data, and enhance accuracy in the presence of uncertainty. Hybridization strategies couple Physics-Informed Neural Networks, which are constrained by physical laws, with residual modelling, where AI learns the "error" between physics-based predictions and actual outcomes. Surrogate modelling approximates computationally expensive simulations, facilitating faster real-time use, while data assimilation merges sensor data with simulation outputs to enable continuous calibration.

Digital Twin could be used for System of System Engineering, Equipment Design or Use in Operation.

We will describe in the following use-cases illustrating the following figure about Simulation for System Engineering including Digital Twin.

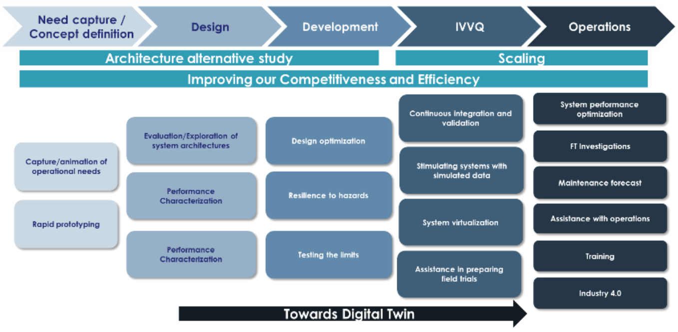



*Figure 1 – Simulation for Systems Engineering (SE)*

# 2. AI FOR SIMULATION

This chapter describes how the integration of AI technics can improve the performances and functionalities of simulation systems. This chapter will deal with:

- AI FOR AUGMENTED TRAINING
- AI-BASED DESIGN BY SIMULATION

## 2.1 AI For Augmented Training

The following paragraphs provide concrete examples of AI-augmented simulations across the three principal categories employed in training applications: live simulation, virtual simulation, and constructive simulation. While this chapter concentrates primarily on training, these three types of simulation, and the AI techniques applied to enhance them, may also be leveraged for other purposes, such as decision support or system design.

Artificial intelligence markedly enhances simulation-based training by increasing realism, adaptability, and pedagogical effectiveness. A principal contribution of AI lies in the autonomous animation of simulated entities. Through the use of intelligent agents, reinforcement learning, and behavior modelling, AI enables non-player entities to exhibit realistic, context-sensitive, and evolving behaviors, thereby reducing dependence on human instructors for scenario management.

AI further supports adaptive training environments. By continuously analyzing trainee actions, performance metrics, and cognitive load, AI-driven simulators can dynamically adjust scenario complexity, pacing, and difficulty. This facilitates personalized learning trajectories, ensuring that training remains appropriately challenging and that learning objectives are effectively reinforced.

Moreover, AI enables advanced assessment and debriefing. Machine learning techniques can automatically process the extensive data generated during simulation sessions to detect patterns, errors, and skill gaps. These insights permit objective, fine-grained performance evaluation and the provision of actionable feedback, substantially enhancing post-exercise debriefing and supporting long-term skill acquisition.

### 2.1.1 AI for Live Simulation (CERBERE Program)

Live Simulation, such as CERBERE designed by Thales, enables military forces, such as the French Army, to train in a real environment with real equipment, but with simulated firing devices and simulated effects.

These training systems generate large amounts of data that are too difficult for instructors to analyze during the exercise.

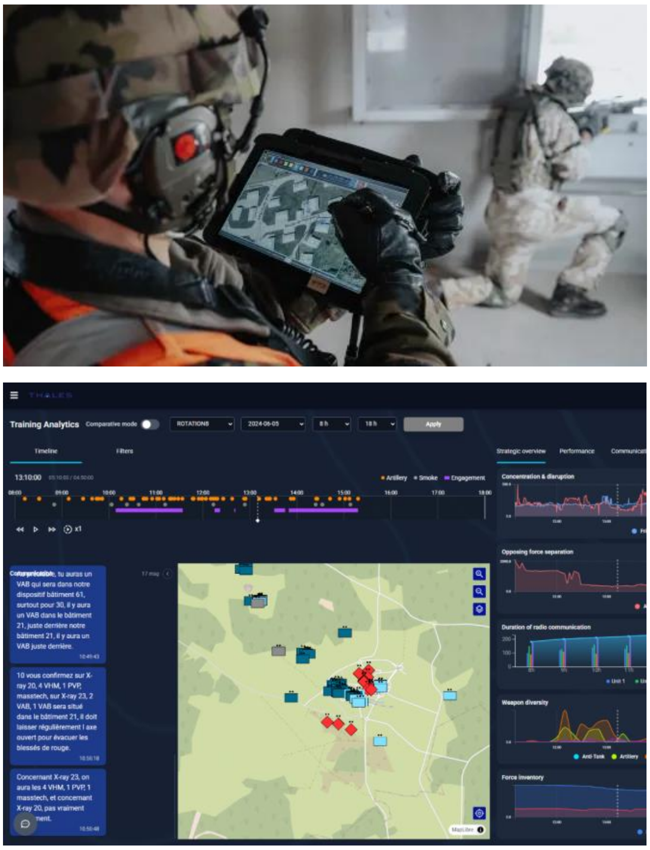



*Figure 2 – Example of AI-enhanced Live Simulation Training in CERBERE*

Thales has developed a debriefing tool capable of automatically analyzing this data to provide instructors with relevant indicators that they can select or specify when preparing for the exercise, enabling them to quickly identify the important facts of the exercise and visualize the evolution of these indicators during the exercise.

The tool will enable instructors to conduct their after-action analysis (AAA) based on objective facts in order to provide better feedback to trainees. It mainly uses data analysis techniques to identify key indicators enabling them to evaluate the forces' maneuvers during the exercise and track progress from one session to the next.

We are currently integrating new AI-based features, based on Natural Language Processing and Generative AI, in the debriefing tool for: (1) Automatic communication analysis, (2) Event prediction to estimate the outcome of engagements, (3) Force performance assessment, and (4) Doctrinal analysis based on the evaluation of tactics employed.

### 2.1.2 I for virtual Simulation (PERCEVAL & EPIIC projects)

Virtual Simulation enables individuals or crews to be trained in simulated 3D environments that are representative of real-world environments, such as aircraft cockpits, helicopter cockpits, or armored vehicle cockpits for example.

In such training environments, instructors need automated tools to monitor the cognitive state and to assess the behaviors and skills of the trainees. To this end, Thales has developed a Virtual Assistant for Training (VAT), based on the use of several AI techniques.

This assistant was first tested in several civil projects such as the DGAC's PERCEVAL project for training civil aviation crews, and is currently being tested as part of the European Project (EDF) EPIIC, which aims to design the advanced cockpit for future European fighter aircraft such as FCAS/SCAF or TEMPEST.

In such training systems, this assistant is able to estimate the pilot's cognitive state through the development of cognitive models (e.g., workload) using machine learning techniques.

This assistant is also capable of analyzing the data generated by the simulation in order to identify the trainees' observed behaviors and, ultimately, to assess their skills, using a combination of data analysis, rule based systems and generative AI.

Finally, this assistant uses the last advances in natural language processing and large language model to analyze aircraft communications and to interact with the instructor in natural language.

<u>Note</u>: Real-world platforms, such as aircrafts, already incorporate AI components (e.g., a planning system) that must be present in their simulation counterparts. One

solution to this requirement is to directly integrate these components into the simulation system and modify the simulation interfaces to generate the inputs required by the AI components and process their outputs.


*Figure 3 – Example of AI-enhanced Virtual Simulation Training in PERCEVAL*

### 2.1.3 AI for Constructive Simulation (Thales Combat Staff Trainer & OPOSIA)

Constructive simulation enables the training of officers at different levels of command. Such training requires the representation of large areas of terrain and the animation of numerous entities (allies, enemies and neutrals) thanks to what is called a "Computer Generated Forces" (CGF) component (see §5.3 for details).

In constructive training simulations, such as Thales Combat Staff Trainer (e.g., OPOSIA training system), the main application of AI is the automation of all entities that are not directly controlled by trainees (or exercise animators).

This AI technique, also called behavioral AI, aims to replicate the human or doctrinal behavior of real actors in the simulation. Modeling this behavior can be the result of parameterizing a behavioral engine with a specific doctrine, for example, or the result of a machine-learning algorithm.

Thales has successfully tested the latter for flight mission learning using Reinforcement Learning as part of a UK MoD study project and Genetic Algorithms with the AI toolkits originally developed by the former US Company Psibernetix, acquired by Thales in 2019 (see §3.1.5 for another use case of this technology).

Another interesting application of AI in constructive simulation is the use of natural language processing, not only to enable artificial entities to send voice feedbacks during their mission, but also to allow users to give tactical orders to their artificial subordinates by voice, as they would in the field.

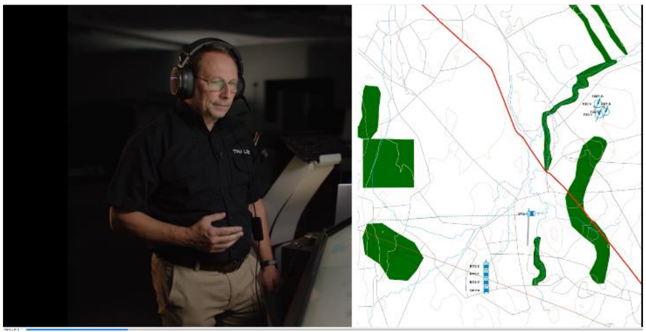


*Figure 4 – Example of AI-enhanced Constructive Simulation Training in T&S Combat Staff Trainer*

Finally, there are two other possible applications of AI for constructive simulation training:

- The analysis of the data generated during the simulation and the evaluation of the behaviors and skills of trainees, which would use the same type of AI technologies already described for identical purposes in Live Training and Virtual Training applications.
- Semi-automated, automated or automatic generation of exercise scenarios that will help instructors facilitate the preparation phase through a virtual assistant capable of selecting or generating all or parts of a scenario.

In addition to training, many other applications can benefit from the use of AI technologies in constructive simulations.

- Wargaming simulations allow the testing of new military platforms and new doctrines. This functionality may obviously benefit from the progress in behavioral AI, scenario generation and natural language processing.

*Figure 5 – AI for Wargaming*

- Constructive Simulation may also be used in operation, for instance to help an officer to make a decision or to prepare a mission. Once again, these features can benefit from advances in constructive simulation, such as the automatic generation and confrontation of courses of actions.
- Operations also generate large amount of data that may need to be analyzed by a tool similar to the virtual assistant experimented for training purposes.

*Figure 6 – Officer digital assistant in operation*

## 2.2 AI-Based Design by Simulation

We illustrate AI-based design supported by simulation through two principal use-cases: the next-generation cockpit for pilots and the Digital Air Traffic Controller for air traffic management.

Artificial intelligence fundamentally transforms design by simulation by accelerating exploration, optimization, and validation across complex design spaces. By learning rapid approximations of high-fidelity simulations, AI enables the evaluation of thousands of design variants at a fraction of the computational cost, thereby supporting early-stage design decisions and significantly shortening development cycles.

Beyond mere acceleration, AI facilitates intelligent exploration of the design space. Optimization algorithms, including evolutionary strategies and gradient-based learning, can automatically identify optimal or Pareto-efficient solutions under multiple, often conflicting constraints, such as performance, cost, robustness, and sustainability. This capability repositions simulation from a passive evaluative tool to an active instrument for design generation.

### 2.2.1 Design of New Cockpit & Evaluation for Pilot

We can improve many technical simulations using the same AI functionalities described above. For instance, Thales used the PERCEVAL platform to evaluate the design of new cockpit for aircrafts and helicopters.

*Figure 7 – AI-Based simulation of cockpit design*

### 2.2.2 Digital Atco Design & Evaluation for Air Traffic Controler

Digital ATCO would support ATCOs in managing increasingly complex and unpredictable air traffic flows while ensuring safety, punctuality, and environmental sustainability. The Digital ATCO can learn the flows (by observing traffic patterns), suggest optimal actions (best trajectory, avoid weather, solve conflicts, minimize clearances and communicate silently) and can help manage the tasks (identification, priority assignment, scan for changes, etc.).

The rise of AI and advanced optimization algorithms solutions provides opportunities to advance ATC system to full automation of controller tasks. Aim is to harness such tech for developing a fully automated Digital ATCO that performs most tasks for the human controller. Automation of main controller actions will drop workload, and enable greater traffic capacity.

AI based or optimization algorithms for:

- Path proposals respecting RECAT
- Descent / climb proposals

- Automated clearances, coordination
- Speech recognition / text synthesis
- Task management
- Strategic flow / traffic management


*Figure 8 – AI for Digital ATCO*

# 3. SIMULATION FOR AI

Supervised AI training relies on datasets that must be representative of the real data that will be processed in the systems. Otherwise, detection biases may arise. This could involve the detection of an object or phenomenon that does not exist. This bias can lead to misinterpretations with various consequences. Another bias may result in the non-detection or erroneous detection of these same objects or phenomena. 'Missing' a detection can also have a significant impact and affect confidence in the systems.

In a context where certain training data is scarce or even non-existent, simulation is one of the methods used by Thales to produce the necessary training sets.

This dual assessment calls for different technological approaches:

- The value of simulation in supplementing insufficient quantities of real data sets.
- The importance of technologies that reduce the "gap to reality" of the data used to train AI systems.
- The key role of adapting algorithm training sets to real-world sensor and usage scenarios. This adaptation is based on professional knowledge of the sensor and its operating environment.

We will develop the following cases:

- ENVIRONMENT SIMULATION FOR MODEL, REINFORCEMENT & MACHINE LEARNING
- SIMULATION FOR DATA GENERATION

## 3.1 Environment Simulation for Model, Reinforcement & Machine Learning

The role of environment simulation in validating systems based on Reinforcement Learning and Evolutionary/Genetic Algorithms is crucial because these algorithms rely heavily on interactions with an environment to learn or evolve effective policies.

This provides Safe Testing and Validation because often Reinforcement Learning and Evolutionary/Genetic Algorithms often explore risky or unknown behaviors during learning or evolution where, by simulation, failures do not cause real-world damage.

Environment simulation accelerates iterations and helps Reinforcement Learning that requires thousands to millions of interactions to converge and Evolutionary/Genetic Algorithms that require evaluating many generations and populations to evolve good solutions. Real-world testing is too slow or expensive, whereas simulation allows fast-forwarding and parallel evaluations.

Simulations allow precise control of initial conditions, stochastic elements, and disturbances to benchmark different Reinforcement Learning policies, to measure robustness of evolved solutions, to ensure that performance metrics are reliable and reproducible.

Many real-world systems face rare or extreme conditions that are hard to test physically. Simulations can generate edge cases, dangerous scenarios, and diverse environmental conditions. Reinforcement Learning and Evolutionary/Genetic Algorithms can be trained and validated on these scenarios, improving robustness.

This approach is also cost-efficient, because testing in real systems can be expensive. Simulation reduces cost by allowing massive virtual experimentation.

Environment simulation is essential for validation, safety, speed, cost-efficiency, and robustness testing for Reinforcement Learning and Evolutionary/Genetic algorithms-based systems. It serves as a virtual testbed where algorithms can learn, adapt, and be evaluated under controlled and extreme conditions before deployment in the real world.

### 3.1.1 Complex behavior learning using constructive simulations (SE-STAR & SETHI)

Reinforcement learning has seen a recent growth and has been applied in many domains, from robotic to game playing. Most of the successes of Reinforcement Learning have been however in single agent domains, where modeling or predicting the behavior of other actors in the environment is largely unnecessary. However, there are many applications that implies interactions between agents that are co-evolving together, for example, robot swarms or autonomous cars. Those agents could also communicate with each other, creating complex behaviors. As such, successfully scaling RL to environments with multiple-agents is crucial to building artificially intelligent systems that can productively interact with humans and each other. Environment is simulated using the engine SE-Star, developed for simulating large crowd behavior, and is now use for drone's environment. In this simulation, n drones evolve in a 2D space of size 5000x5000 and have to prevent enemy drones from reaching the center red area. Enemy drones are controlled by the environment. Agents can see allies' position as well as enemy's position relative to themselves. Every drones have a small cone of vision that destroy enemy when they are inside.

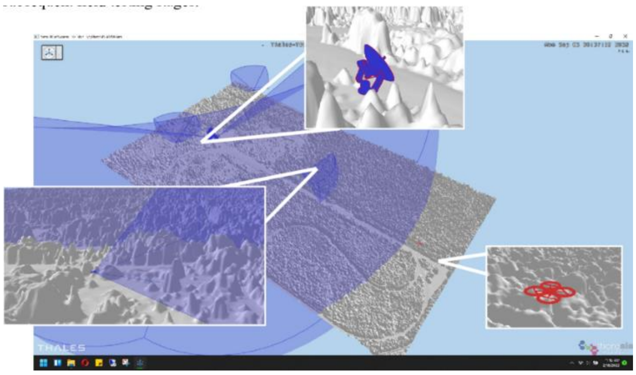


*Figure 9 – SE-STAR Drone simulation*

As part of the European project IV4XR, Thales has developed a use-case that is simple enough to verify that test agents can replace human testers in the most costly tasks, and close enough to a real-world example to allow the results to be extrapolated to more complex applications.

In this use-case, called "Nuclear Plant Intrusion Simulation", the behavior of the intruder was first learned using Reinforcement Learning in a simple simulation (SE-STAR), then this behavior was transferred to a more sophisticated simulation (SETHI) in order to finalize the learned policy.

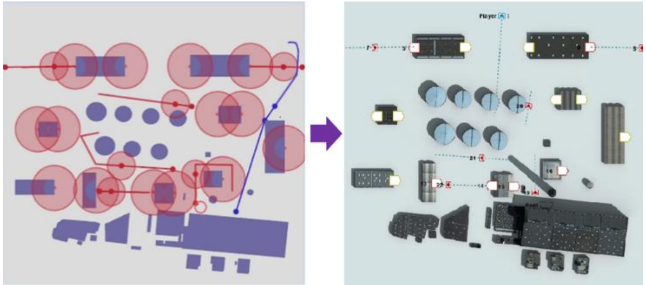
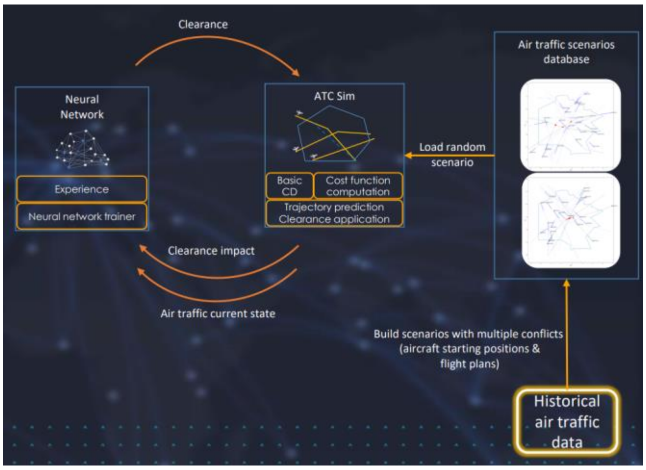


*Figure 10 – Transfer Learning between SE-STAR and SETHI of Reinforcement Learning Policies*

### 3.1.2 CDR (Conflict Detection & Resolution)

THALES CDR (Conflict Detection & Resolution), resolution function, is designed to help Air Traffic Control Operators (ATCO) to ensure the aircraft separation and safe transition of the air space, which also implies resolving the loss of separation (or conflicts) predicted by the CD function of the ATC system.

Thales's CDR is a cutting-edge, virtualized system that enhances air traffic control by providing earlier, more accurate conflict alerts combined with AI-driven decision support. Based on a hybrid AI approach (classical AI and deep reinforcement learning), CDR AI ultimately improves capacity, safety, efficiency, and controller workload management in complex airspace environments. PINN (Physics-Informed Neural Network) BADA are used to generate realistic flight trajectories with respect to airplane flight mechanics constraints.



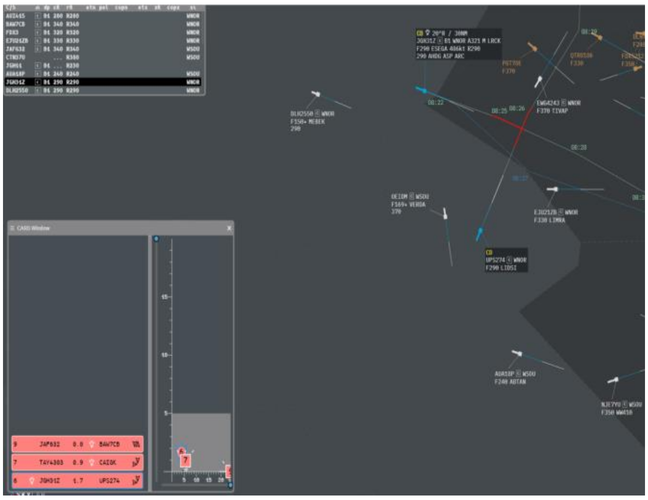


*Figure 11 –DIGITAL ATCO (Air Traffic Controller) for CDR (Conflict Detection & Resolution) based on Reinforcement Learning trained on AI-based simulation of Traffic*

### 3.1.3 TWA (weapon-to-target assignment) Simulation

Managing resource allocation dynamically across varying scenarios is a complex problem that necessitates robust and flexible approaches. Thales has developed a deep reinforcement learning approach based on an effector-target modeling framework to solve the dynamic weapon-to-target assignment problem (WTA). Using simulation environment, Thales created various WTA scenarios for an approach that includes a method utilizing MultiLayer Perceptron and another leveraging the attention mechanism. This method is input-size independent, enables knowledge generalization, and allows agents to perform well across diverse scenarios, while also providing an initial level of explainability for decision-making.

*Figure 12 –Simulation for Target-to-Weapon Assignment*

### 3.1.4 Mission Planning Satellites constellation

To maximize utilization of a constellation of satellites and to increase value creation by generating optimized mission plan, Thales has developed Reinforcement Learning algorithms to compute mission plans closer to the optimal, to be downloaded to each satellite in the entire constellation every hour as operational case. This combinatorial problem is NP-hard. Constellations with up to 5 satellites, 5,000 requests in 15 areas of interest are equivalent to ~$10^{169}$ combinations per satellite and area of interest, and so ~$10^{34,000}$ combinations in total. GPU-accelerated simulator used to train the agent and compute acquisition plans.

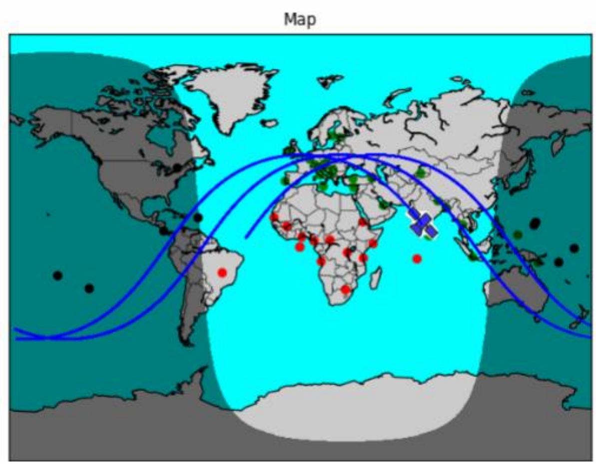



*Figure 13 – Satellites Constellation simulation for RL Mission Planning Algorithms*

### 3.1.5 PSIBERNETIX based on Evolutionary/Genetic Algorithms

THALES Psibernetix is based on a Genetic Fuzzy Tree (a combination of Genetic Algorithms and Fuzzy Logic), that breaks down complex problems hierarchically and in human-like non-binary fashion, makings sub-decisions whenever possible. The immense capabilities offered by Genetic Fuzzy Tree had notably been showcased in 2016 when they were put to the test in a US military project aimed at controlling the flights of Unmanned Combat Aerial Vehicles in air-to-air combat missions within a high-fidelity simulation environment. In the trials, the system defeated multiple renowned fighter pilots and expert tacticians. When the results went public, the story became a viral phenomenon that trended on social media. At Thales, the Genetic Fuzzy Tree methodology has been extensively taken on board, and has notably been applied to COMBI, a bidirectional operator/system translator concept currently in development which was awarded the European Defence Agency's 2021 Defence Innovation Prize. The COMBI concept is the first to integrate human 'intentions' in a system and provides the means of translating these high-level operator intentions between the operator and the multiple intelligent systems in a dynamic manner throughout a flight mission. Then, the pilot evaluates the information from those systems to make the best decision.

*Figure 14 – Psibernetix AI won out during simulated aerial combat against U.S. expert tacticians.*

## 3.2 Simulation for data generation

Simulation provides unlimited and scalable data. Real-world data collection can be time-consuming, expensive, or physically constrained. Simulations can generate massive amounts of labeled data quickly, which is particularly useful for deep learning that requires large datasets. This avoids manual annotation, which is costly and prone to error. This ensures that the Machine Learning system learns robust features across a wide range of conditions (e.g. weather variations). By training on slightly varied simulated data, Machine Learning models are more robust when deployed in real-world conditions. Simulation provides cheap, safe, scalable, and controllable data generation, often with perfect labels and coverage of rare conditions, which makes it extremely valuable for Machine Learning training.

### 3.2.1 SIMULATION OF SCENARIO FOR COLLABORATIVE SITUATION & THREAT ASSESSMENT

Thales has worked on improving the anticipation capacity of a Joint Tactical SubGroup commander through the estimation of enemy intent in COSITA (Collaborative Situation and Threat Assessment). Current tactical situation is evolving with a much larger number of units on the battlefield than in the past, due to the rise of the use of autonomous vehicles (UAVs, UGVs). Understanding this situation is more and more complex for the commanders. COSITA's prediction AI module gives commanders some prediction of enemy intent. This prediction is based on the federated observation of enemy units' positions, movements and activities through the deployment of several sensors and observation vehicles as well as the use of sensing, decision and fusion algorithms at different levels. The COSITA predictive tool is integrated in several Thales products, including CDP - Thales tactical C2 and Atlas - Thales artillery C2.

In order to train the predictor, as combat situations with actual observations of red troops remain quite difficult to gather apart from training exercises, Thales has worked with operational referents and simulated several operational scenarios and a large amount of runs of these scenarios with data augmentation on the red units' parameters and on the units' observations by friendly units and dedicated sensors. This vast dataset has completed available operational data to create the training and validation dataset for the predictor.

### 3.2.2 Application to Space Observation Images

Thales has developed space observation systems that are now used in military and civilian applications. Very high-resolution systems provide access to capabilities for identifying equipment of military interest. The algorithms used to process these images are trained using combinations of real and synthetic images.

Synthetic imagery is generated from simulations that take into account knowledge of observation payloads. These simulations incorporate the environment, such as lighting conditions and objects of interest to be included in the images.

*Figure 15 – Combination of real and synthetic images*

In addition, technologies that reduce the 'gap to reality' are used to blend in objects of interest to enhance training datasets. Recent progress in the domain of Generative Adversarial Networks (GAN) help us better bridge this gap between the real and simulated worlds. Models trained separately perform well on their respective domain but lack when cross-trained to fit both worlds. Thanks to GANs, we can create datasets representing images at the cross of both domains.

### 3.2.2 Application to SAR & SAS (Synthetic Aperture Radar & Sonar) Imagery

Synthetic Aperture Radar (SAR) imaging supports applications from environmental monitoring to defense. For Automatic Target Recognition (ATR), Deep Learning (DL) delivers strong results but needs large datasets, and SAR data is scarce due to cost and confidentiality. A common workaround is training on synthetic data generated by simulators and Computer-Aided Design (CAD) models, but these simplify complex electromagnetic effects, creating a domain shift between training (synthetic) and test (measured) domains. Although Data Augmentation (DA) is used to improve representativeness and robustness, many lack semantic, physics-informed changes to improve recognition performance. Thales has proposed a physics-based DA to address the Synthetic-to-Measured (S2M) gap, first validating physical parameter extraction from measured images and then leveraging this knowledge to improve ATR.

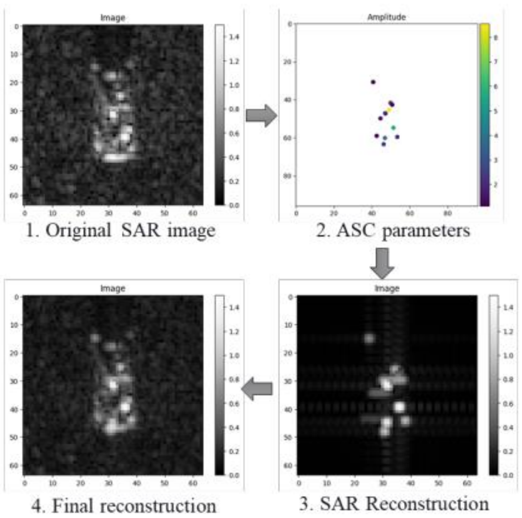



*Figure 16 – Synthetic Image Reconstruction*

# 4. ENHANCED SIMULATION WITH HYBRID DIGITAL TWIN

Hybrid digital twins combine physics-based models with data-driven models, capturing unmodeled dynamics, noise, and complex correlations. This results in more accurate and realistic simulations, even for complex systems that are hard to fully model analytically. Enhanced simulation with Hybrid Digital Twins brings together the best of physics-based modelling and data-driven AI models, providing advantages far beyond traditional simulation or purely data-driven approaches. Hybrid Digital Twin can simulate complex interactions between multiple subsystems, agents, or environments, useful for applications where interactions are critical.

We will illustrate Hybrid Digital Twin for design of SoS (System of System) and equipment (e.g. Radar) and for use in operations.

## 4.1 Digital Twin for design of system of system

### 4.1.1 Digital Twin for collaborative system architecture design

Modern combats rely on connectivity between users, and as such communications systems are a key part to be finely defined with the overall system that will use them. For this THALES has developed a Communication Network Management System (CMNMS) that is a network component dedicated to define long-term network policies to be applied on the network in the near to long term future (according to the mission timing). Whereas Predictive Communications reacts only to the network variations, and adapt the network short term status, the Communication Network Management System is longer-term oriented. Located in the management plan, it can reconfigure the network for future mission needs or in anticipation of future network reconfiguration.

This system allows to provide answers to a wide panel of operational scenarios. A large number of variations of these scenarios have been simulated, to ensure representativity and were exploited by a first version of CMNMS based on heuristics. We are currently developing an AI version in order to provide a much more adapted answer to variations of the situation. Again, the limited availability of real data is compensated by the use of extensive simulation with use case validated by operationals.

### 4.1.2 Collaborative Combat LAB (CCLab)

The CCLab is dedicated to technico-operational simulation of a multi-domain operations tactical system-of-systems, in order to confirm operational added value of collaborative combat modules.

For that purpose, the activity in the CCLab consists in modelling a *blue* vs *red* scenario that is manageable for the blue side, and progressively augment the red forces threat. Both forces have accurate behavioral models for platforms, sensors and effectors, but also for connectivity, C4I and sensing collaboration.

This aggressive baseline can be played with standard policies or embed AI collaborative modules. The two versions of the blue order of battle, with or without collaborative modules, are then played against the same red opposition to gather data on how AI collaborative modules improve current policies.

The simulation provides technical parameters as well as operational KPIs that make explicit, tangible and measurable the operational added value of the AI collaborative modules.

Several simulators work in parallel within a High-level Architecture (HLA) federation:

- A Computer-Generated Forces – SETHI from Thales AVS (see §3.1.1 and §5.3)
- A modular digital platform for collaborative services prototyping as well as external models integration – KAST from Thales Services (based on Kubernetes)
- A library of sensor models – Virtual Sense from Thales DMS
- A Communications simulation tool – COMSaaS from Thales SIX (Communications Simulation as a Service)

The simulation of autonomous swarm is an illustration of this way of working (cf. fig.16): A blue force is tasked with stopping a red platoon of armored vehicles plus anti-drone warfare, using artillery. Legacy solution leads to firing quite a lot of shells on the enemy presence area but the possible changes in speed and direction of red vehicles make this option highly inefficient. Other currently proposed solutions rely on loitering munitions, which each requires an operator, a dedicated datalink, and non-synchronized time on target. On the contrary, the CCLab AI module, deployed on top of current modules, turns these munitions into a synchronized swarm, reducing the number of operators, the electromagnetic footprint and the efficiency of anti-drone warfare. Metrics such as the number of operators, the number of loitering munitions, as well as the mission efficiency are therefore quite easy to evaluate.

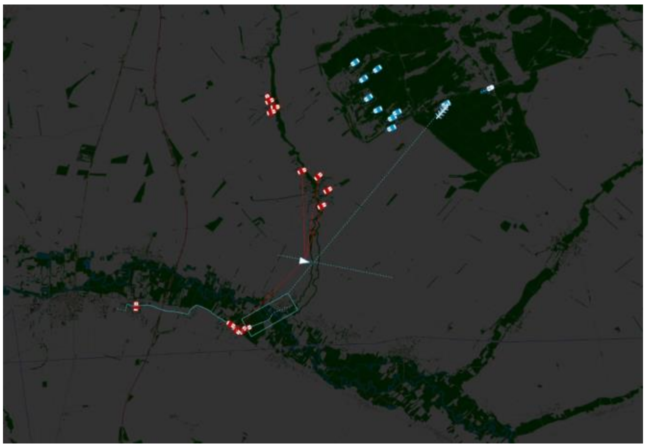


*Figure 17 – Simulation of autonomous swarm*

## 4.2 Digital Twin for design of equipment

### 4.2.1 TWINFIRE (SF500 Naval Radar Digital TWIN)

A digital twin TwinFire is a digital replication of a Naval SF500 radar dedicated to test and/or exploratory trials. The Multi-Function Radar (MFR) digital twin consists in two parts: the real SW and a simulation of the environment (target, propagation, clutter) and of the Front End. This simulation is handled autonomously or through a distributed simulation through DIS interfaces. For Thales, TwinFire is used to optimize radar processing validation schedule (TwinFire being included into a virtuous loop of comparison between the results obtained on the digital twin and those obtained during field tests), to perform efficient and cost-effective non-regression testing and to speed up implementation of new functions. For system integrator, TwinFire is also used to secure system integration schedule without requiring the radar itself. For customer, TwinFire is used to assess performances in specific scenarios and or threats and for new radar modes if needed. Main advantages of AI-based Digital Twin TwinFire is for faster Radar definition and validation, performances assessments when trials are not possible or are complex/expensive to set up (saturation scenarios, missiles…), to prepare Integration with Combat Management System, and for verification of evolutions and non-regression testing.

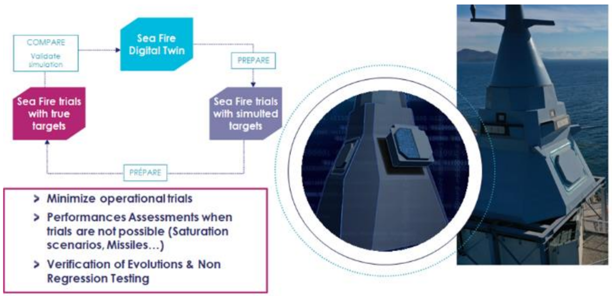



*Figure 18 – TWINFRIRE: SEA FIRE 500 Naval Radar Digital Twin for Frigates*

## 4.3 Digital Twin for use in operation

### 4.3.1 NO-FLY ZONE simulation & Glocalization Accuracy Assessment for UTM (Uav traffic management) in urban environment

The knowledge of the wind velocity at the street level affects optimal drone trajectory planning. The main challenge is that wind forecasts have kilometer-level resolution, while urban applications require meter-level resoluion. By using coarse wind conditions as boundary conditions for fine-grain urban analysis, street-level wind in a city or district can be evaluated using state-of-the-art CFD software. However, this approach is costly and impractical for real-time applications. Instead, by generating fine-scale wind data for different macroscopic wind conditions and training a physics-informed regression model, wind maps can be evaluated almost in real-time. These wind maps can then be used to assess other related fields (temperature, air quality, etc.) in real-time.

The main challenges include optimal trajectory planning to ensure minimal energy consumption given environmental conditions (wind and turbulence maps). Additionally, managing collisions and facilitating collaborative drone operations are crucial. Hybrid artificial intelligence addresses these technical objectives, while social and human considerations are also taken into account.

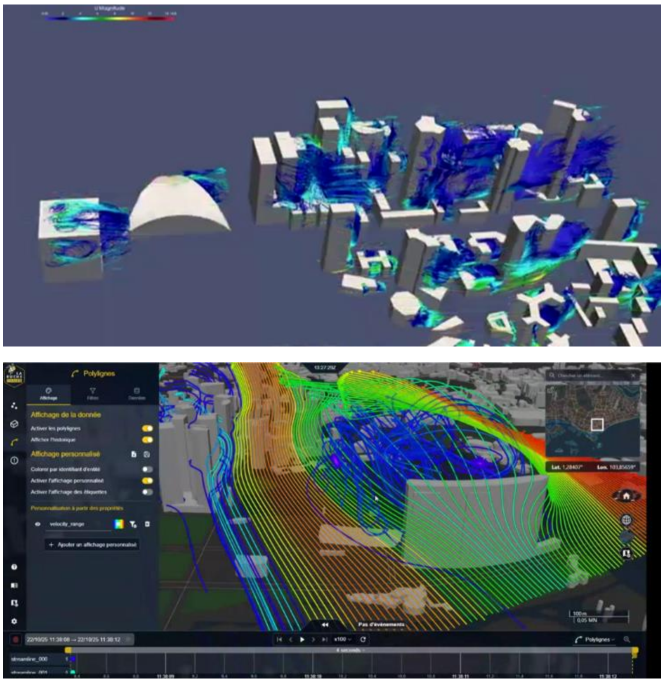


*Figure 19 – UTM (UAV Traffic Management) in Urban Environment in DesCartes CNRS@CREATE project*

# 5. TECHNOLOGIES OF AI BRICKS FOR SIMULATION

## 5.1 PINN (Physics-Informed Neural network)

PINNs are neural networks trained not only on data but also to respect physical laws, typically expressed as partial differential equations (PDEs). This is achieved by incorporating terms into the loss function that penalize violations of these equations. To render PINNs trustworthy, Thales has proposed an extension of the Hamiltonian Neural Network (HNN) by integrating symmetry constraints and Noether invariants within a Symplectic Foliation-Informed Neural Network. This analytically grounded model-informed machine learning framework is further extended to learn dissipative physics with Thermodynamics-Informed Neural Network (TINN), based on metriplectic equations and Souriau's Lie Group Thermodynamics, constrained by transverse Riemannian foliations.

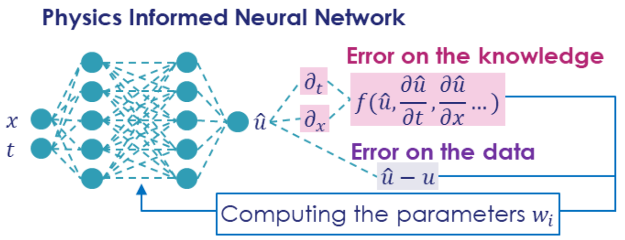



*Figure 20 –Physics-Informed Neural Networks*

## 5.2 MYRIAD: Multicriteria DecIsion AiD

The evaluation of simulation results is a critical step in the process. It involves identifying objective measures, known as Key Performance Indicators (KPIs). These KPIs may pertain, for example, to the assessment of technical-operational simulations - such as testing a blue force architecture in a scenario against a red force for border surveillance — or to evaluating the performance of a trainee in a tank simulator.

MYRIAD has been utilized to aggregate KPIs and assess simulation outcomes. The first step consists of building the aggregation model by eliciting stakeholder preferences. The second step involves applying the previously developed multi-criteria model to the simulation results. The main advantages of MYRIAD are its ability to model complex decision-making strategies (including the representation of interacting criteria), its user-friendly algorithms for learning from preference data, and its capability to generate explanations and provide recommendations for improvement.


*Figure 21 –MYRIAD interface to explain aggregation*

To ensure representative evaluations, Monte Carlo simulations are frequently used. This involves executing each solution across multiple runs. MYRIAD can be integrated with statistical tools to, for example, measure the confidence that an evaluation surpasses a given threshold, based on the statistical distribution of the Monte Carlo simulation results.

### 5.3 CGF: Computer Generated Forces

The “Computer-Generated Forces” (CGF) module is the main component of a constructive simulation. This component is responsible for animating all entities in the simulation using more or less sophisticated behavior models, taking into account interactions with the terrain model and with all entities or objects in the simulation.

As quickly described in §2.1.3, CGFs may use different categories of AI techniques to model the behavior of simulation entities.

Basic CGFs use rule-based systems or behavior trees to animate simulation entities with stereotypical behavior, while more sophisticated CGFs can use human behavior models to simulate crowd behavior or to increase the level of autonomy and realism of military entities in the simulation for example.

*Figure 22 –Thales CGF SETHI used for doctrine studies in FCAS/SCAF project*

In recent years, it has been demonstrated that some or all of the behavior of entities can be learned from their experiences in accelerated simulations using reinforcement learning techniques or genetic algorithms (as illustrated respectively in §3.1.1 and §3.1.5). In the future, this behavior could also be the result of a Machine Learning algorithm based on data, as could the generation of simulation scenarios.

## 6. References

[1] Mathieu Klein, Thomas Carpentier, Eric Jeanclaude, Rami Kassab, Konstantinos Varelas, et al.. AI-Augmented Multi Function Radar Engineering with Digital Twin: Towards Proactivity. 2020 IEEE International Radar Conference, Sep 2020, Florence, Italy

[2] Frédéric Barbaresco, Lie Groups Machine Learning & Foliation Structures of Thermodynamics-Informed Neural Networks, 2nd International Workshop on Artificial Intelligence and Augmented Engineering (AIAE’22), 2022

[3] Margareta Fretas. Bringing Facts and Metrics to Competency-Based Training and Assessment – A Civil Approach for the Military Domain. International Training Technology Exhibition and Conference (ITEC) 2025, Oslo, Norway.

[4] Jean-Yves Donnart. De l’aide à l’instruction à l'embarquement d’un instructeur virtuel dans un aéronef à l’horizon 2035. Séminaire Combat Aéro-Terrestre 2035, Nov. 2025, Ecole Militaire, Paris, France.

## 7. ABOUT THALES

Thales is directly engaged with the challenges and opportunities outlined in this document through its diverse business activities. For several years, artificial intelligence has been embedded across its product portfolio—particularly in critical systems and defense solutions deployed or intended for deployment in multiple countries worldwide. This operational experience has enabled Thales’ research and engineering teams to develop AI components, tools, and methodologies tailored to both sovereign and international contexts.

Today, Thales brings together more than 800 AI experts within its dedicated internal organization, cortAIx, established in 2024. These experts are distributed across three core teams:

- Labs, focused on foundational research and model development;
- Factory, responsible for toolchains, cybersecurity, and critical infrastructure;

### Sensors, integrating AI into embedded and edge systems.

Together, these teams ensure that AI capabilities are not only cutting-edge but also operationally viable, secure, and aligned with sovereignty requirements. Thales’ AI workforce is strategically located across France, the United Kingdom, Canada, Singapore, Germany and the United Arab Emirates, reflecting the company’s global footprint and commitment to supporting regional sovereignty through local expertise.

----

Building a future we can all trust

4, rue de la Verrerie 92190
Meudon FRANCE

Tél. + 33(0)1 57 77 80 00

www.thalesgroup.com