Scensei
FUSION · SIMULATION · INSIGHT

Educating and Training Agents: Introducing OPTIMA's Instructor and Bootcamp

Antonio Cifonelli, Francesco Sponchiado and Armando Geller

01 October 2026

Contact:
Dr. Antonio Cifonelli, AI Lead
+41 79 907 12 00, antonio@scensei.com

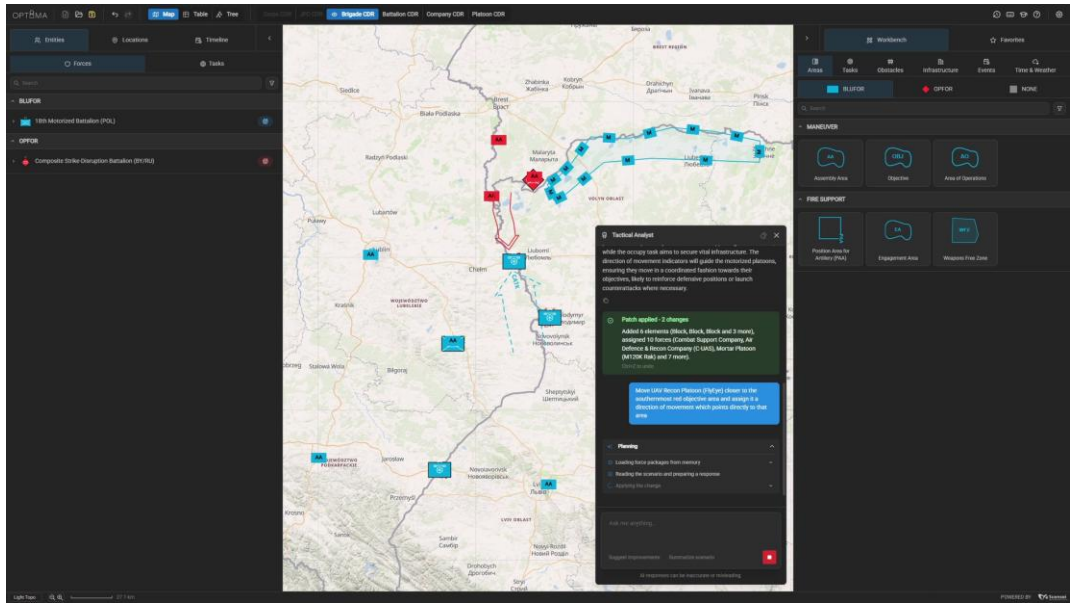


Image: Scensei.

PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Topics that will be discussed

- System architecture
- Implementation approach
- Experimental methodology used to evaluate the framework
- Observed outcomes, technical challenges, and lessons learned
- Implications for future defense simulation, training, and decision-support applications.

PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Objective

Introduce an approach for integrating doctrine-informed generative agents and reinforcement learning techniques within a multi-agent simulation environment to support scenario development, course-of-action (COA) analysis, and agent training.

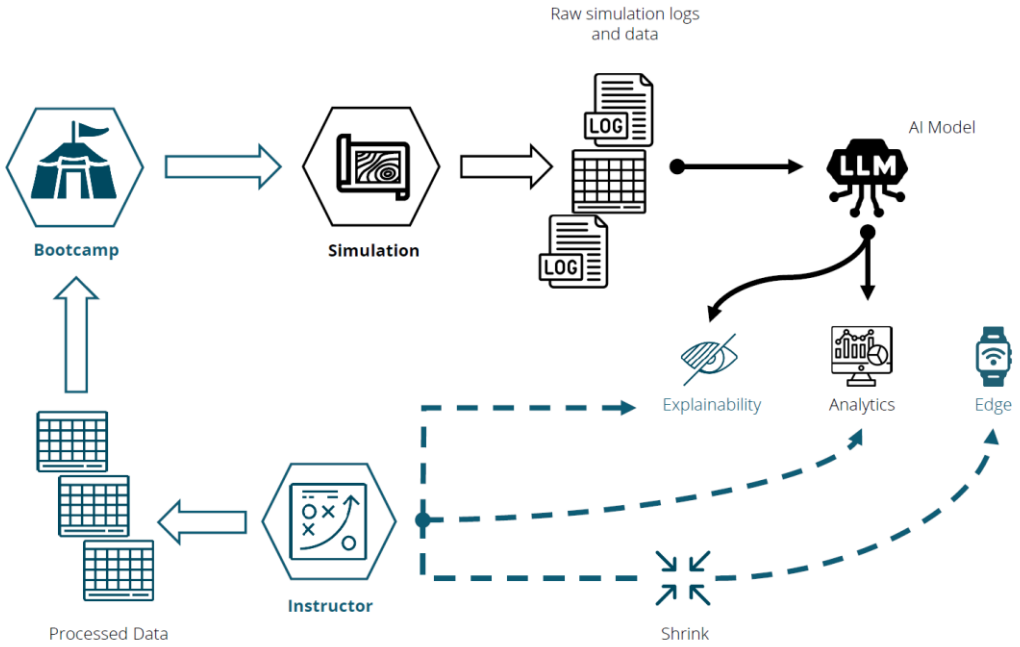
Develop smarter agents, able to emulate their physical counterparts.

Produce **more realistic simulations.**

Ship user-centric system for **enhanced explainability and analytics** to the user.

Increase agents' adaptability in unseen scenarios.

Target **edge and offline** devices.



PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Scensei
FUSION · SIMULATION · INSIGHT

Proposed framework

- Combine a doctrine-based, multi-generative-agent system with a MARL training environment. Generative agents provide contextual and doctrinal knowledge, while reinforcement learning methods enable agents to develop and refine operational behaviors through repeated exposure to representative mission scenarios.
- Ensure operational relevance through continuously assessing emergent behaviors against doctrinal constraints and mission objectives.
- Incorporate AI-assisted scenario and COA generation capability to accelerate the creation of complex multi-domain operational environments for experimentation.

PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Two main components: Instructor and Bootcamp

Instructor

- LLM
- Deriving information from publicly available doctrines and other official documentation
- Generates COAs and informs our agents with structured knowledge of their operational environment, including intelligence on the opposing side, so enemy capabilities and likely actions are factored into the planning

Bootcamp

- MARL
- Training agents based on the Instructor's knowledge through Reinforcement Learning (RL).
- The two components so designed prepare our agents to adapt dynamically and adequately while in simulation.

PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Instructor

A Multi(-Generative)-Agent Artificial Intelligence (AI) System:

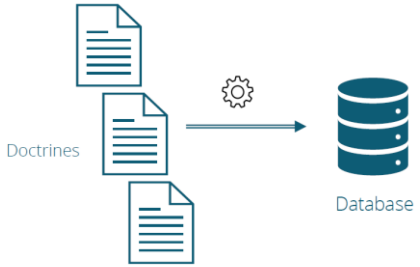
- That **serves multiple purposes**: can be used either as a planning **assistant** or as an **information retrieval system** to extrapolate structured information.
- Is **modular** where each agent can be a standalone assistant/information retrieval system.
- Is **scalable**: if an agent in the graph must be added or pruned, it can be done without affecting the other agents.

PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Instructor: components

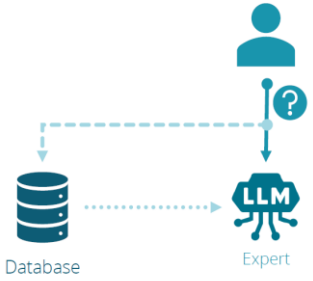
Data processing and storing

Focused on documents (such as manuals, datasheets...) pre-processing and subsequent storage.



User interaction

This view focuses on the user-facing components and how users interact with the system. The LLM Expert acts as the central point of contact, handling queries and coordinating any background processes to deliver a coherent response back to the user.

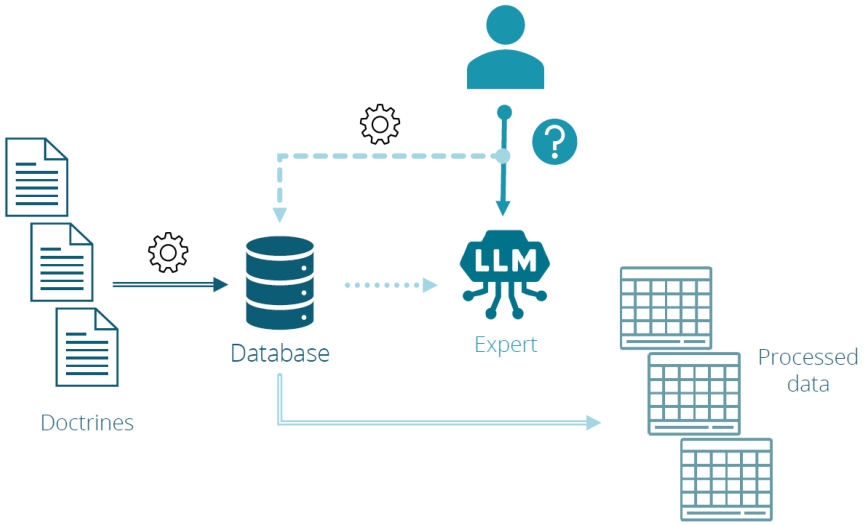


PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Scensei
FUSION · SIMULATION · INSIGHT

Instructor: agent overview

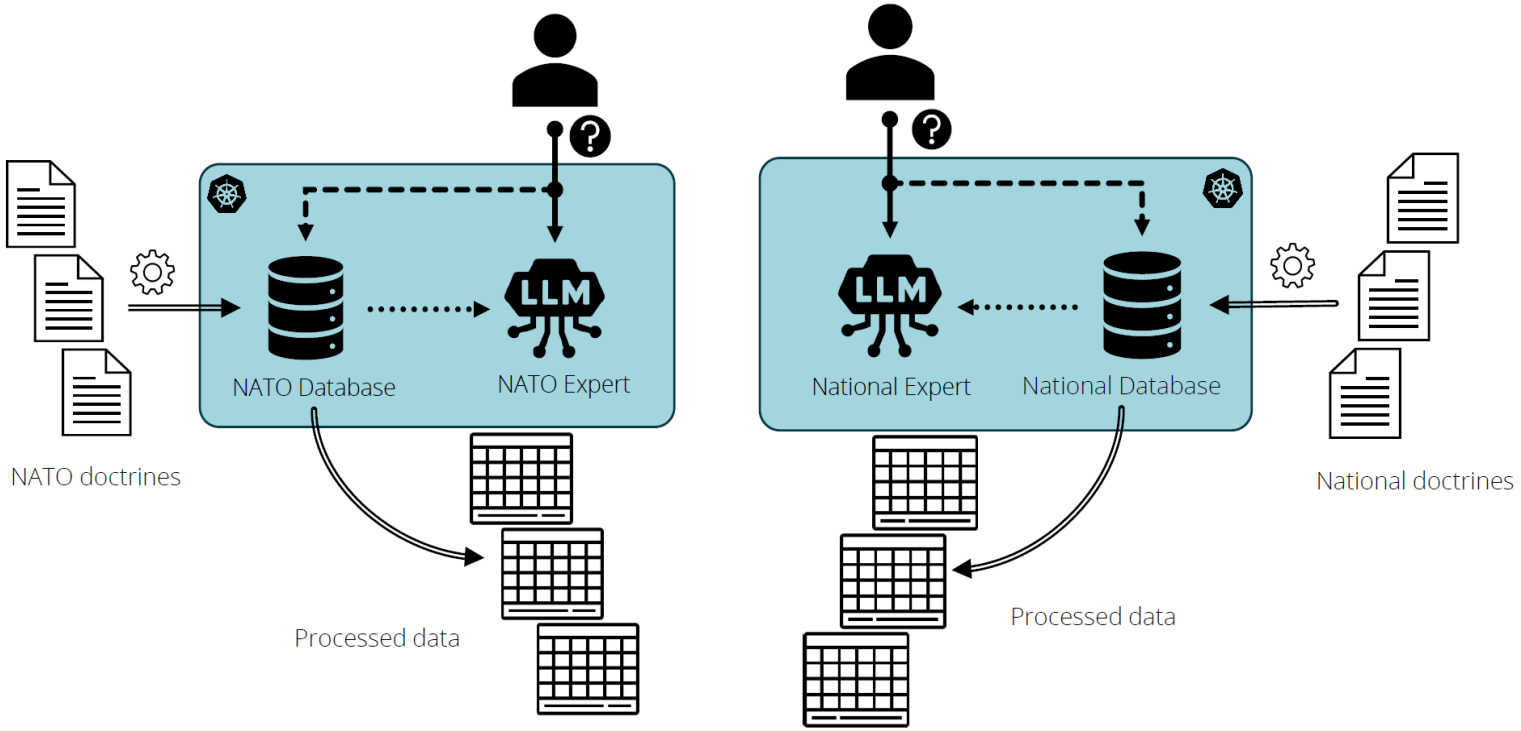
- Setting interaction among previously defined services, e.g. the Scenario Editor Assistant.
- Structured output to be used in the Reinforcement Learning environment.



PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Scensei
FUSION · SIMULATION · INSIGHT

Instructor: standalone



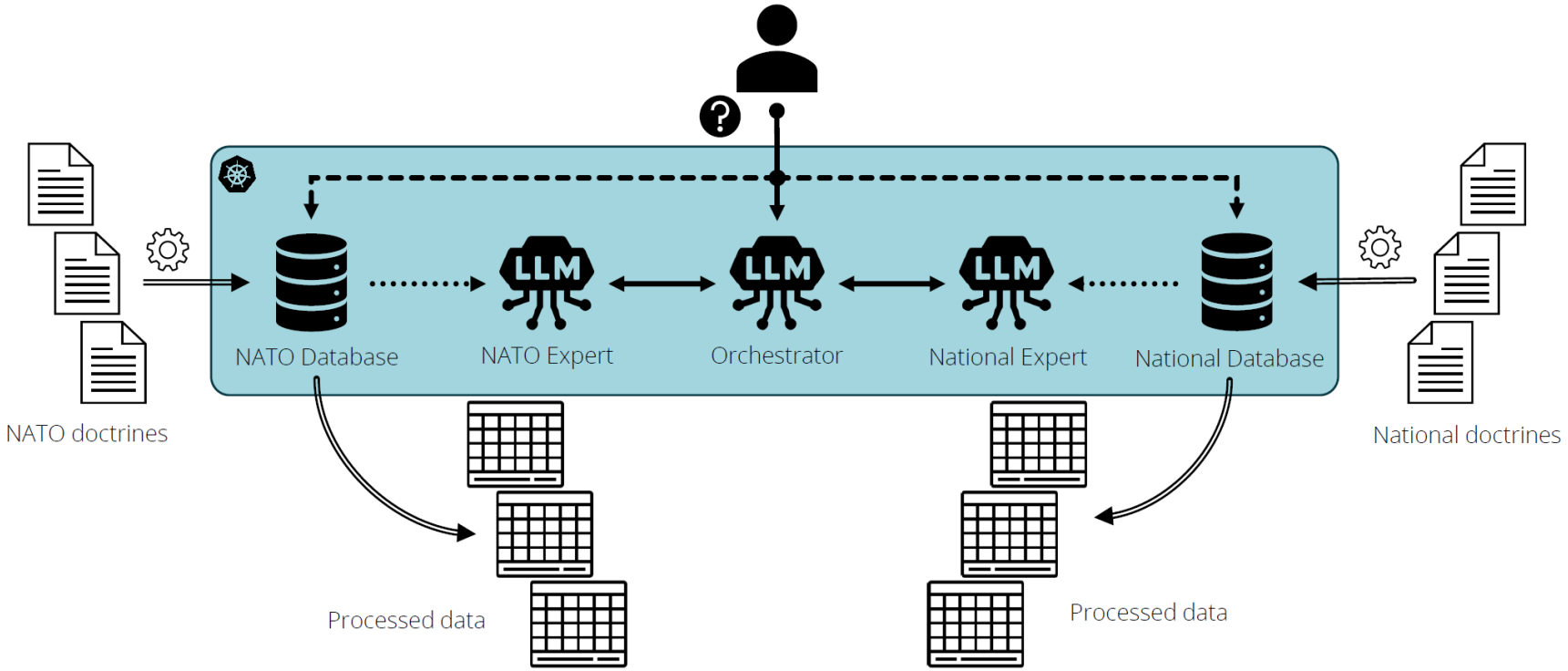
Processing step

Kubernetes cluster

Scensei
FUSION · SIMULATION · INSIGHT

PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Instructor: orchestration



Processing step

Kubernetes cluster

Scensei
FUSION · SIMULATION · INSIGHT

PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Instructor: testing methodology

- Within-subjects controlled comparison
  - With and without assistant
- Ablation studies
  - Constrain directly
  - Think first then constrain
  - With and without intelligence information
- Metrics
  - Setup time (quantitative)
  - Rate of degenerate shapes/incorrect placement (quantitative)
  - COA acceptance rate (quantitative)
  - Expert rating (qualitative)
  - Interviews (qualitative)

PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Instructor: observed outcomes, technical challenges, and lessons learned

- Reduced scenario setup time: user sets what matters, Instructor fills the gaps and take care of repetitive, structural work.
- Smooth transaction between COAs: asking for a different posture, will cause the model to re-think the COA accordingly
- Per-Commander vs Per-Level
- Generated coordinates look «correct» in isolation on their own, regarding syntax and semantic.
- Lack of spatial awareness: entities requiring multiple coordinates, end up receiving duplicates.
- Anchoring: reuse of already-available coordinates when generating new ones for other entities.

- Pattern matching over computation for coordinates: LLMs favor pattern matching over actual computation when producing coordinates, which is likely the root cause of degenerate coordinates.
- Anchoring, hypothesis: coordinates provided for entities on the map are «low hanging fruits» the LLM preferentially grabs from the context rather than computing fresh ones.

- Lack of publicly available data: affects the ability to validate outputs
- Think first, constrain after: let the model plan first, then constrain with schema/grammar
- Expert in the loop: domain expertise to review/correct outputs and services. Not a commodity population.

Scensei
FUSION · SIMULATION · INSIGHT

PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Bootcamp

Multi-agent reinforcement learning (RL) system

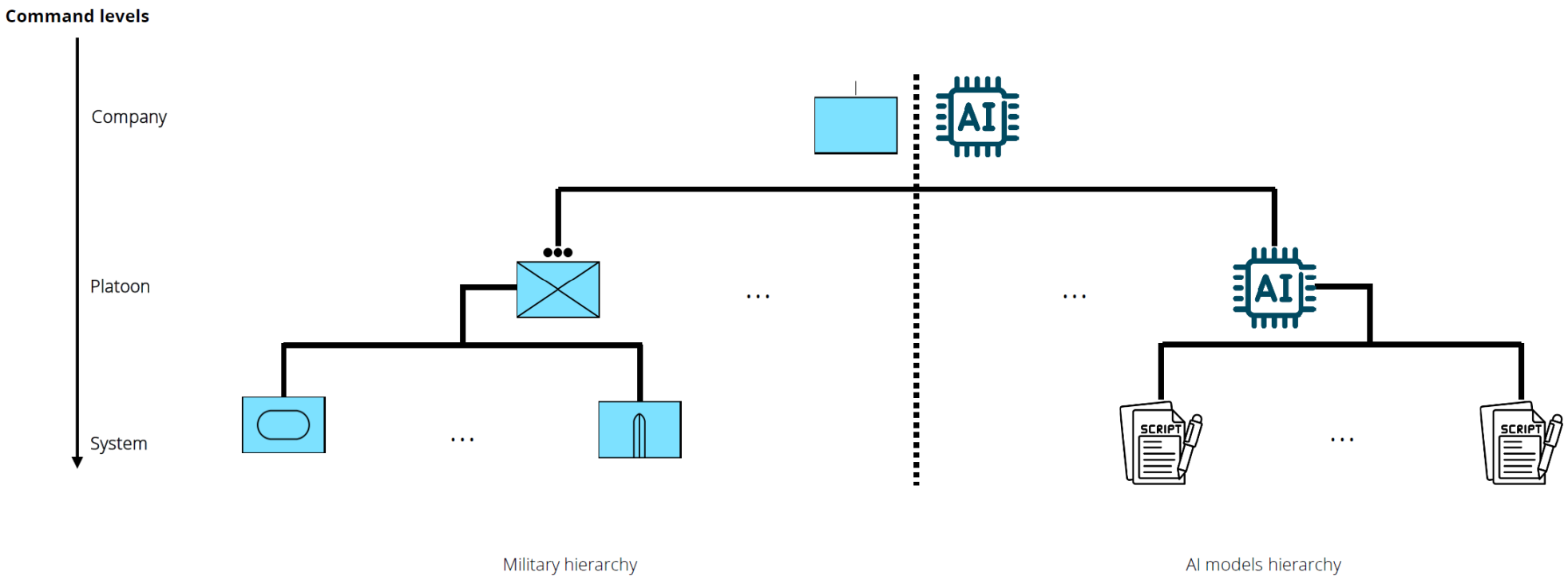
Training: dedicated environment to turn our agents into operational domain experts. Produce smarter agents and more realistic simulations.

Doctrine enforcement: embedding Instructor's knowledge into our RL loop, align system behaviors with human expectations.

Simulation agnostic: the training process is not tied to a specific simulation or simulation engine.

PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Bootcamp: mapping hierarchies



AI model

Scensei
FUSION · SIMULATION · INSIGHT

PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Testing methodology

- Doctrine's violation as hard constraint (quantitative)
- Doctrine's violation as soft constraint (qualitative)
- Credit assignment sanity check (quantitative)
- Expert judgement (qualitative)
- Scalability
  - Communication overhead (quantitative)
  - Wall-clock training time (quantitative)

PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Observed outcomes, technical challenges, and lessons learned

- **Explainable decisions**: a modular, military-inspired architecture and action space impact positively the interaction with experts.

- Encode doctrines' information into an RL-ready format
- Simulation interfaces for observation/action space

- **Learn what matters**: wise resource usage, not everything is worth learning
- **No LLMs in the loop**: resource-intensive, doesn't scale well with the number of agents
- **Intelligence on commanders**: planning requires foresight
- **Reaction on systems**: execution requires adaptability

PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Implications for future defense simulation, training, and decision-support applications

- This presentation demonstrated a practical approach to integrating doctrine-informed generative AI and multi-agent reinforcement learning within a modular simulation environment for wargaming, experimentation, and multi-domain decision support. The implementation and experimental findings provide a foundation for further operational validation and research.
- The proposed architecture represents a distinctive approach to combining generative AI, reinforcement learning, and simulation-based decision support.
- The same differentiation applies at the functional level. The framework combines AI-assisted scenario generation, course-of-action development, doctrine-informed agent behavior, and explainable decision support within a unified environment.
- The combination of doctrine-informed generative agents and reinforcement learning offers a promising pathway toward more realistic, adaptive, and operationally relevant simulation agents.
- The separation of planning, knowledge generation, and agent training into distinct yet connected components supports scalability, maintainability, and future integration with external simulation environments and defense systems.
- By adopting a user- and mission-centric design philosophy while maintaining a modular, scalable, and open architecture, the framework can be tailored to specific operational challenges without compromising flexibility or interoperability.
- The resulting capability helps accelerate the transition from situational awareness to anticipatory decision support, enabling users to explore potential futures, evaluate alternative courses of action, and prepare for emerging operational conditions.

PROPRIETARY © 2026 Scensei (Switzerland) Sagl

Scensei
FUSION · SIMULATION · INSIGHT

Grazie!

Antonio Cifonelli, Francesco Sponchiado
and Armando Geller

01 October 2026

Contact:
Dr. Antonio Cifonelli, AI Lead
+41 79 907 12 00, antonio@scensei.com

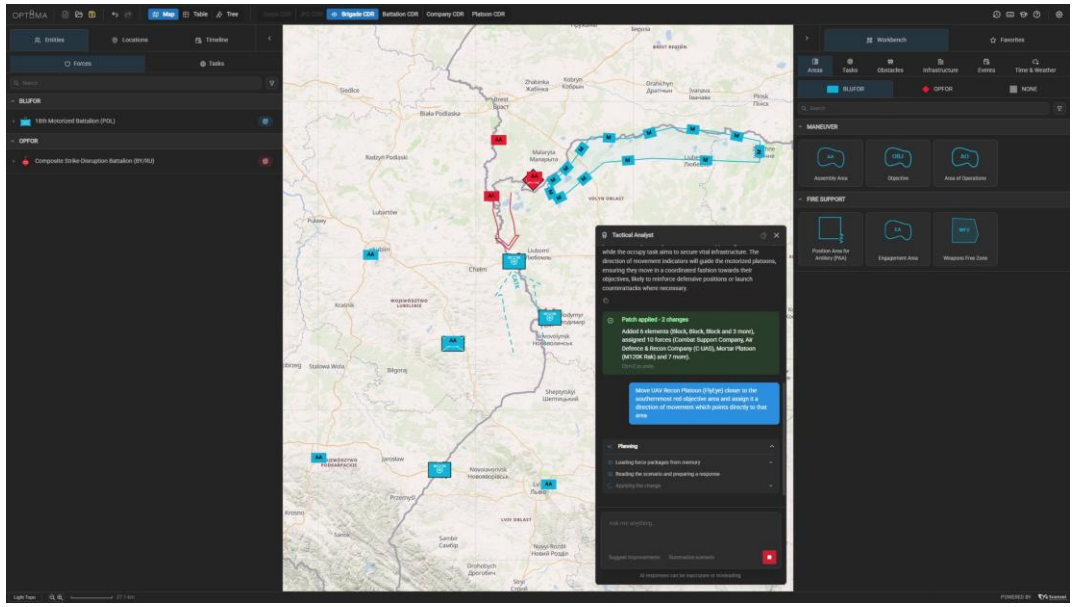


Image: Scensei.

PROPRIETARY © 2026 Scensei (Switzerland) Sagl