A Data-Centric Simulation Ecosystem for Accelerated Operational Outcomes

Modern defense organizations face an expanding gap between the pace of technological innovation and the speed at which new capabilities are fielded. This paper proposes a data-centric simulation ecosystem beyond those currently envisioned—enabled through a data fabric and rapidly adaptable to new technologies—to accelerate the transition of ideas into operational outcomes across the U.S. Department of War and NATO. Rather than treating simulation as a collection of isolated models and platforms, the approach positions data as the persistent, governed, and shareable foundation that connects modeling and simulation, experimentation, analysis, training, and operational planning.

Further, by embedding governance, provenance, and automated validation within the data layer, the framework can ensure traceability and credibility for verification, validation, and accreditation activities. This creates a persistent evidence base that allows experimentation results to inform planning, acquisition, and mission rehearsal in near-real time. The outcome is a shift from episodic experimentation to a continuously learning defense enterprise where simulation is integrated into daily operational workflows.

At its core, the proposed capability contains a military M&S domain data model that provides a common representation of units, entities, platforms, relationships, and behaviors spanning tactical, operational, and strategic levels. This shared information model enables consistent interpretation of data, models, and analytical results across diverse simulation systems, operational planning tools, and synthetic environments. By exposing the data model through the underlying data fabric, users and systems can discover, access, and exchange authoritative data products using a common semantic framework rather than relying on point-to-point data translations.

To support scalable data exchange and persistence, the proposed capability implements data distribution using Google Protocol Buffers (Protobuf) and Apache Kafka (event-driven data streaming capability). Protobuf provides an efficient, platform-independent mechanism for transporting simulation entities, events, and state information. Kafka enables the continuous publication, distribution, and storage of simulation data streams across distributed environments. Together, these technologies allow simulation outputs, operational data, analytical products, and synthetic environment updates to be captured as persistent, re-playable data products that can be shared across applications.

The capability requires a Parametric Data Service (PDS) which is the single source of truth for model data. Accessed through a UI, operators can view, modify, create, and curate model data. The PDS uses Attribute-based Access Control (ABAC) to restrict access to parametric data to authorised users. The PDS service maintains metadata about each parameter that is observable and modifiable by authorised users. Every parametric data modification creates a revision history associated with parametric data. The revision history maintains the lineage of a parameter as it is modified so that traceability to the original value is always maintained, thereby supporting credibility, traceability, and auditability. All modifications can be reversed to the original/previous state.

Ultimately, a data-centric simulation fabric transforms synthetic environments into a shared, coalition-ready capability that supports multi-domain operations, strengthens interoperability, and delivers measurable improvements in decision speed and mission effectiveness. This approach provides a practical pathway for ensuring that innovation moves rapidly from concept to operational impact.