

AI Simulation for Materials Procurement Negotiations



BUSINESS PROBLEM

Materials procurement negotiations recur at a fixed cadence, affect a large supplier portfolio, and have meaningful cost and supply implications. In the setting studied here, buyers have only partial visibility into supplier constraints and limited time to compare alternative strategies before live negotiations. At the same time, complete negotiation transcripts are not consistently retained in the historical record. The business need was therefore a safer offline environment for planning, strategy comparison, and rehearsal grounded in the strongest available evidence.

DATA SOURCES

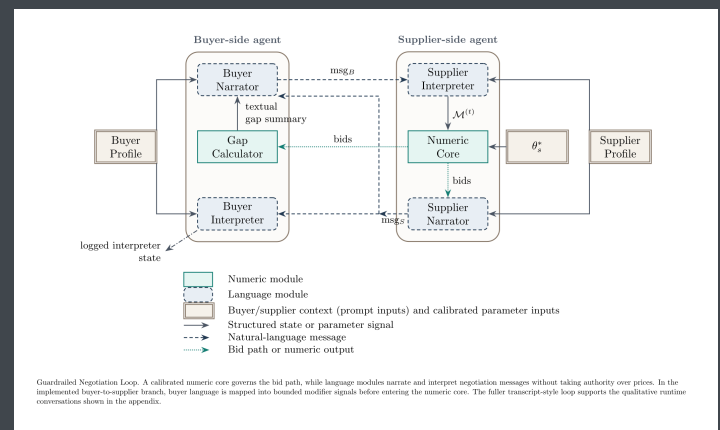
1. Retained operational negotiation workbook snapshots for 2025-H2, 2026-H1, and 2026-H2 with prices, targets, bids, and volumes.
2. A canonicalized read-only research dataset derived from those snapshots for calibration and evaluation.
3. SME interviews and limited de-identified communication artifacts used for workflow interpretation and language-interface design, not quantitative fitting.

Data Types and Format

Boolean, Categorical (nominal), Count (integer), Numeric, Relational Data, Structured, Text, Time series (Dynamic), Unstructured

APPROACH

The project builds a guardrailed multi-agent negotiation simulator anchored on a supplier-aware numeric core calibrated from canonicalized historical negotiation traces. The numeric model uses prior price, disclosed target, per-round bids, stopping depth, and forecast volume to generate bounded bid paths and is evaluated on fixed held-out supplier/cycle splits against heuristic baselines. LLM modules are added only as narrator and interpreter components that generate or interpret language through bounded modifier channels rather than proposing prices directly.



IMPACT

Validated thesis results show that the Numeric Core Model achieved the best held-out Masked MSE in 2 of 4 feasible supplier/cycle comparisons. On Supplier A, calibrated on 2025-H2 and evaluated on 2026-H1, it reduced Masked MSE by about 20% versus a heuristic benchmark and about 35% versus the persistence baseline. On Supplier B, calibrated on 2026-H1 and evaluated on 2026-H2, it improved Masked MSE by about 6% versus heuristic benchmark. The guarded structured-language extension improved APS error versus its paired numeric-only reference on both overlapping held-out tests, but improved Masked MSE on 0 of 2. Directionally, the artifact enables safer offline planning, more explicit strategy comparison, and a reusable training or rehearsal environment. No live deployment dollar savings, cost reduction, or hours-saved measurement is reported in the thesis.

 DRIVERS	The project was motivated by repeated, commercially important negotiations where buyers needed better support for preparation and strategy comparison. The strongest retained evidence was numeric rather than conversational, creating a need for a simulator that could learn from sparse observation without overstating what the data could justify.
 BARRIERS	The main barriers were limited transcript retention, uneven target coverage, noisy operational workbook traces, and sparse late-round support on some supplier-cycle pairs. These constraints made it difficult to validate richer behavioral models and required careful canonicalization, benchmark pressure, and conservative claim boundaries.
 ENABLERS	Key enablers were recurring negotiation structure, retained round-level numeric traces, cross- functional SME input from sourcing and analytics stakeholders, and a reproducible pipeline that linked canonical data, calibration, diagnostics, and reporting artifacts under one contract. LGO alumni network at company was very supportive.
 ACTIONS	We mapped the negotiation process, canonicalized retained workbook snapshots into a read-only dataset, defined a supplier-aware numeric model and heuristic benchmarks, calibrated and tested them on fixed cross-cycle splits, and then added a bounded language layer in which narrator and interpreter modules could affect the simulator only through clipped modifier signals.
 INNOVATION	The technical novelty is the combination of a calibrated supplier-aware numeric bidding core with guardrailed LLM communication modules that are not allowed to set prices directly. This creates a modular negotiation simulator that is both more realistic than a language-free replay tool and more auditable than an end-to-end autonomous LLM negotiator.
 IMPROVEMENT	The clearest validated improvement was predictive: on one held-out Supplier A transfer, the Numeric Core Model lowered Masked MSE by about 20% versus heuristic benchmark and about 35% versus persistence, and on the one feasible Supplier B transfer it improved Masked MSE by about 6% versus heuristic benchmark. Broader business improvement remains directional rather than live-measured.
 BEST PRACTICES	In sparse enterprise settings, keep numerical authority in an interpretable calibrated backbone, benchmark against strong simple alternatives, and let language influence the simulator only through bounded, inspectable interfaces. Clear provenance handling and reproducible artifacts matter as much as model complexity.
 OTHER APPLICATIONS	The same framework could be adapted to other repeated buyer-supplier negotiations where structured offer traces are retained more reliably than full dialogue, including other materials categories, sourcing negotiations in adjacent industries, and training or rehearsal environments for commercial teams under human oversight.