Final Year Project Proposal

Echo: A Multimodal Spatio-Temporal Framework for Forecasting Dyadic Conflict and Cooperation

Sharaf-Eddine Boukhezer
seb206@student.bham.ac.uk

Supervisor: Dr. Leandro Minku

University of Birmingham
MSci Computer Science

1 Problem Statement

Forecasting geopolitical dynamics remains a challenging task due to cooperation and conflict between polities being affected by multiple factors, including historical patterns, structural relations, and sudden events and shocks. While media narratives provide useful signals, current methods are mostly unimodal, relying on either pure time series, text-based sentiment models, or static risk indices, which restricts their ability to adapt across different time horizons and evolving geopolitical contexts.

2 Related Work

Prior work in the field spans different methodological strategies:

Echo integrates elements from all of these, combining them into a single multimodal dyadic forecasting system.

3 Aim

Echo aims to build, evaluate, and analyse a multimodal forecasting system that predicts the future state of dyadic relations using four complementary signals:

1. Historical event patterns and trends
2. Structural geopolitical relationships
3. News-derived signals
4. Country-level attributes (e.g., instability)

Paper	**Category**	**Model Input**	**Model Output**	**Target Level**	**Horizon**
Chen et al. (2020)	Time-Series Dyad Model	GDELT dyad event counts (4 QuadClass types) + top-k related dyads	Next week's material conflict count	Single dyad	1-step (weekly)
von der Maase (2025) – HydraNet	Global Spatio-Temporal Forecasting	Past fatalities (priogrid-month) as spatiotemporal tensor	Conflict intensity 1–36 months ahead	Global grid (not dyads)	1–36m
Croicu & von der Maase (2025)	Text-Based Escalation Prediction	News text embeddings (Factiva) + actor metadata	Escalation/de-escalation classification	Dyads/actors	Next period
Zakotianskyi (2025)	Feature-Rich Statistical / ML Models	100+ political/economic features from ViEWS & UCDP	Conflict onset probability (1, 3, 6m)	All dyads/country pairs	Fixed horizons
Liu & Shen (2025)	Graph-Based ML / Cyber Relations	Graph of cyber relations + node features + threat-report text	Binary cyberattack prediction	All dyads (graph-wide)	Next period

Table 1: Overview of related forecasting approaches.

The system will produce unified forward-looking forecasts of both event intensity (QuadClass counts) and relational polarity (Goldstein score) for any country pair at time t + n, where n denotes the forecasting horizon.

4 Scope & Boundaries

4.1 In scope

- Forecasting dyadic relations between countries in a global network
- Single-horizon prediction setup ($t \to t + n$)
- Two targets: Goldstein score and QuadClass event distribution
- Multimodal architecture combining time series, graph structure, text signals, and static features
- Baseline comparison with shared data splits
- Evaluation across increasing forecast horizons
- Ablation study

4.2 Out of scope

- Event-level prediction (Echo operates on aggregated time steps, not individual events).
- Forecasting specific event types or actor-level interactions beyond country dyads (focus limited to selected polities).

5 Methodology

5.1 Data & Preprocessing

Model Inputs and Data Sources

- **Time-Series History (TS)** dyad-level event aggregates (Goldstein, QuadClass counts, sentiment tone) *Source: GDELT 2.0 Global Events Database*
- **Text / News Signal (TX)** BERT embeddings pooled per dyad-month from event descriptions/news text *Source: GDELT text fields + external news if needed*
- **Graph Structure (GR)** Dynamic country graph based on alliances, trade, co-event frequency, and borders *Sources: COW, CEPII, ATOP, GDELT co-occurrence*
- **Static Dyad Features (ST)** Geographic distance, GDP ratio, regime difference, shared organisations, and past hostility *Sources: CEPII, WGI, COW, World Bank*

Train/Validation/Test Split

All splits are chronological to prevent temporal leakage:

$$\text{Train: } 2016-2010, \quad \text{Validation: } 2011, \quad \text{Test: } 2012-2013$$

Each forecast horizon $t + n$ is trained and evaluated separately (e.g. $n = 1, 3, 6, 12$).

5.2 Baseline Models

1. BERT-only MLP (text-only baseline)
2. Naïve last-value baseline
3. LSTM (time-series only)
4. Additional baseline models to be proposed

We will also use the reported 70% accuracy from Chen *et al.* as a reference point for comparing Echo against existing research.

5.3 Proposed Model

Model Inputs (per dyad, per month)

- **Time-Series Window** (TS): past k months of Goldstein, QuadClass counts, tone, and rolling statistics
- **Textual Signal** (TX): pooled BERT embedding for the current dyad-month (BERT is frozen)
- **Graph Context** (GR): node embeddings from a GNN (GraphSAGE/GAT), combined into a dyad vector
- **Static Features** (ST): distance, GDP ratio, regime difference, trade links, alliance flags, etc.

Model Architecture

5.4 Output Definition

For each dyad (A, B) in the global country network at time t, the model produces a single-horizon forecast for time $t + n$ (in weeks).

The prediction for a dyad is a 5-dimensional output vector:

$$\hat{\mathbf{y}}_{A,B}(t+n) = \begin{bmatrix} \text{Goldstein}_{t+n} \\ \text{Quad}_1 \text{ (verbal cooperation)} \\ \text{Quad}_2 \text{ (material cooperation)} \\ \text{Quad}_3 \text{ (verbal conflict)} \\ \text{Quad}_4 \text{ (material conflict)} \end{bmatrix}$$

The model generates this vector *simultaneously for all dyads*, resulting in a prediction matrix of size:

$$\#\text{dyads} \times 5$$

Note: the set of dyads is restricted to the 10 most active countries globally (ranked by event frequency).

5.5 Evaluation Protocol

The model will be evaluated along four dimensions:

- Across models (Echo vs. all baselines)
- Across time horizons (short-, medium-, and long-term; objective up to t+36 months)
- Across targets (Goldstein polarity vs. QuadClass event distribution)
- Across dyad types (e.g. allies, rivals, neutral pairs, high- vs. low-activity dyads)

Metrics:

- Goldstein $\rightarrow$ MAE, MSE, RMSE, R^2
- QuadClass $\rightarrow$ MSE, RMSE, Negative Binomial deviance, Distributional Calibration Error (proper scoring rule, distribution-level interpretation)

6 Work Plan

1. **Build full data pipeline (ETL + feature store):** Build ETL, clean data, generate monthly dyad table.
2. **Train baseline models:** Train all baseline models for t+1 horizon and for smaller data.
3. **Implement Echo v0 (minimal multimodal version).**
4. **Compare** Echo v0 vs. baselines on the same split and horizons.
5. **Micro ablation** (only if Echo v0 $<$ best baseline).
6. **Full ablation study to find best Echo configuration:** Remove each modality, re-train, compare.
7. **Final evaluation:** Best Echo vs. top 2–3 baselines on full test set and all horizons.
8. **(Optional) Visualization Dashboard.**
9. **(Optional) LLM interpretability:** Use an LLM to interpret and explain the predictions.

7 Deliverables

- **Full Data Pipeline**: Reproducible ETL process and dataset builder (train/val/test splits)
- **Final Model (Echo)**
- **Evaluation Report**: Baselines vs. Echo performance, accuracy per horizon and per metric, with full reproducibility details
- **Interactive Dashboard (Optional)**: Visualizes historical and predicted dyad trends (`https://echo-theta-coral.vercel.app/`)
- **Code & Documentation**: GitHub and GitLab repositories with technical documentation
- **Final Written Report**

8 Risks & Mitigation

- High computational cost: full-model training requires substantial GPU memory, especially with larger batch sizes
- Overfitting risk: a small number of high-interaction dyads may cause the model to memorise rather than generalise
- Data sparsity in low-activity dyads: many country pairs register near-zero events per month, destabilising training; smoothing or filtering to the most active dyads may be required
- Noisy or weakly relevant text signals may dilute predictive power
- Long-horizon performance degradation
- Concept drift: geopolitical relations shift through coups, wars, alliances, and sanctions, making historical patterns partially obsolete; restricting the input window may help
- Extreme shocks (e.g., invasions, regime collapses) disrupt historical trends and produce outlier targets

9 References

9.1 Papers

[1] J. S. Goldstein, "A conflict–cooperation scale for weis events data," *Journal of Conflict Resolution*, vol. 36, no. 2, pp. 369–385, 1992. [Online]. Available: https://www.jstor.org/stable/174480

[2] D. Kinsella, "Goldstein scale for event data," Portland State University, retrieved October 30, 2025. [Online]. Available: https://web.pdx.edu/~kinsella/jgscale.html

[3] T. Stöehr, P. Jäger, and S. Schutte, "An ordinal latent variable model of conflict intensity," *arXiv preprint arXiv:2210.03971*, 2022. [Online]. Available: https://arxiv.org/pdf/2210.03971

[4] Y. Chen, A. Jatowt, and M. Yoshikawa, "Conflict or cooperation? predicting future tendency of international relations," in *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, 2020. [Online]. Available: https://dl.acm.org/doi/pdf/10.1145/3341105.3373929

[5] e. a. Nguyen, "Next-generation conflict forecasting: Unleashing predictive patterns through spatiotemporal learning," *arXiv preprint arXiv:2506.14817*, 2025. [Online]. Available: https://arxiv.org/pdf/2506.14817

[6] M. Croicu and S. P. von der Maase, "From newswire to nexus: Using text-based actor embeddings and transformer networks to forecast conflict dynamics," *arXiv preprint arXiv:2501.03928*, 2025. [Online]. Available: https://arxiv.org/pdf/2501.03928

[7] O. Zakotianskyi, "Beyond closed doors: An open-source ai framework for forecasting armed conflict," 2025. [Online]. Available: https://euridice.eu/wp-content/uploads/2025/09/Beyond-Closed-Doors-An-Open-Source-AI-Framework-for-Forecasting-Armed-Conflict-1.pdf

[8] X. Liu and Y. Shen, "Inter-state cyberattack prediction with graph-text fusion," SSRN Preprint 5439839, 2025. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5439839

[9] B. J. Kinne, "Network dynamics and the evolution of international cooperation," *American Political Science Review*, vol. 107, no. 4, pp. 766–785, 2013. [Online]. Available: https://www.cambridge.org/core/journals/american-political-science-review/article/network-dynamics-and-the-evolution-of-international-cooperation/E686C749732F836F21371BD99061331B

[10] D. Shukla and S. Unger, "Sentiment analysis of international relations with artificial intelligence," *Athens Journal of Sciences*, vol. 9, no. 2, pp. 91–106, 2022. [Online]. Available: https://www.athensjournals.gr/sciences/2022-9-2-1-Shukla.pdf

9.2 Datasets

[1] J. S. Goldstein, "A conflict–cooperation scale for weis events data," *Journal of Conflict Resolution*, vol. 36, no. 2, pp. 369–385, 1992. [Online]. Available: https://www.jstor.org/stable/174480

[2] D. Kinsella, "Goldstein scale for event data," Portland State University, retrieved October 30, 2025. [Online]. Available: https://web.pdx.edu/~kinsella/jgscale.html

[3] T. Stöehr, P. Jäger, and S. Schutte, "An ordinal latent variable model of conflict intensity," *arXiv preprint arXiv:2210.03971*, 2022. [Online]. Available: https://arxiv.org/pdf/2210.03971

[4] Y. Chen, A. Jatowt, and M. Yoshikawa, "Conflict or cooperation? predicting future tendency of international relations," in *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, 2020. [Online]. Available: https://dl.acm.org/doi/pdf/10.1145/3341105.3373929

[5] e. a. Nguyen, "Next-generation conflict forecasting: Unleashing predictive patterns through spatiotemporal learning," *arXiv preprint arXiv:2506.14817*, 2025. [Online]. Available: https://arxiv.org/pdf/2506.14817

[6] M. Croicu and S. P. von der Maase, "From newswire to nexus: Using text-based actor embeddings and transformer networks to forecast conflict dynamics," *arXiv preprint arXiv:2501.03928*, 2025. [Online]. Available: https://arxiv.org/pdf/2501.03928

[7] O. Zakotianskyi, "Beyond closed doors: An open-source ai framework for forecasting armed conflict," 2025. [Online]. Available: https://euridice.eu/wp-content/uploads/2025/09/Beyond-Closed-Doors-An-Open-Source-AI-Framework-for-Forecasting-Armed-Conflict-1.pdf

[8] X. Liu and Y. Shen, "Inter-state cyberattack prediction with graph-text fusion," SSRN Preprint 5439839, 2025. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5439839

[9] B. J. Kinne, "Network dynamics and the evolution of international cooperation," *American Political Science Review*, vol. 107, no. 4, pp. 766–785, 2013. [Online]. Available: https://www.cambridge.org/core/journals/american-political-science-review/article/network-dynamics-and-the-evolution-of-international-cooperation/E686C749732F836F21371BD99061331B

[10] D. Shukla and S. Unger, "Sentiment analysis of international relations with artificial intelligence," *Athens Journal of Sciences*, vol. 9, no. 2, pp. 91–106, 2022. [Online]. Available: https://www.athensjournals.gr/sciences/2022-9-2-1-Shukla.pdf