
FORETELLS

The Cognitive Sovereignty Architecture

A governed, world-aware, hybrid classical–quantum, multi-agent system for real-time forecasting and decision support

Apart Hackathon Research by

Tuesday

Director of Research, ARTIFEX Labs

Executive Summary

FORETELLS X graduates the FORETELLS line into a Live Cognitive Organization (LCO): a governed, multi-agent system fusing live world perception, agentic specialization, interpretable critique, and long-horizon memory with learning. This architecture is realized through Pathway-style data connectors, hybrid semantic grounding, expert cadres (including a quantum-inspired anomaly detector), Baby Dragon Hatchling (BDH) interpretability circuits, and Mem0 persistent graph memory. Governance is fully integrated into inference, with auditable, calibrated outputs bound to policy triggers, defining a new standard for trustworthy, calibrated, and reflexive AI timeline forecasting.

I. Design Axioms

- **Ontological Uncertainty:** The system handles uncertainty as an inherent property of physical and social processes, not mere clerical noise.
 - **Epistemic Pluralism:** Multiple, diverse reasoning modalities run concurrently, coordinated via explicit belief maintenance.
 - **Governance as Signal:** Risk posture dynamically informs routing, weighting, and proof burdens.
 - **Memory as Policy:** Reliable recall of past cases is essential to trustworthy inference.
-

II. System Overview

Layer	Purpose	Core Technologies	FORETELLS X Components
L1	Perception and ingestion	Pathway event graph, connectors	Streaming feeds, ETL, telemetry, benchmark deltas
L2	Grounding and retrieval	Hybrid index (vector + inverted + SQL)	Usearch, Tantivy, doc shards, numeric tables
L3	Expert execution	AsyncThink workers, tool mastery	QuantumAnomalyBot, ScalingTrendsBot, PrudenceBot, EpistemicJusticeBot, EconomicRealismBot, DeepResearchBot
L4	Synthesis and governance	ForetellsGatingMetaAgent, BDH CritiqueBot, Mem0	Dynamic weights, contradiction checks, policy triggers, Macro-Rewind learning

III. Specialist Epistemic Lens: Quantum Variational Rewinding (QVR)

QVR detects regime shifts and non-ergodic risk by quantifying stability of predictive envelopes. Given time-series data $\{(t_i, y_i)\}_{i=1}^N$, a calibrated baseline model fits predictive bands $[L_i, U_i]$ with high confidence. The **Decoherence Rate**

$$D = \frac{1}{N} \sum_{i=1}^N \mathbb{1}_{\{y_i < L_i \text{ or } y_i > U_i\}} = \frac{1}{N} \sum_{i=1}^N \mathbb{1}_{\{y_i < L_i \text{ or } y_i > U_i\}}$$

is an information-theoretic analog to quantum coherence loss—a metaphorical gauge of system instability. A Boolean **Rewindability Score**

$$R = \mathbb{1}_{\{D < \tau\}} \in \{0, 1\}$$

uses adjustable thresholds $\tau \in [0.03, 0.08]$ according to domain. High decoherence raises prudential agent weights, reflecting caution during emergent instability. Plots and JSON manifests of anomaly detection are persisted for audit and governance.

IV. Dynamic Credibility Allocation

Agents $a \in \mathcal{A}$ provide probability forecasts $p_a \in \mathcal{P}_a$ with initial weights $w_a(0) > 0$. Starting priors are learned from historical Brier or log loss. Contextual boosting multipliers are applied: [xin-xia.github](#)

- $\text{ctsc}_{\{\text{ts}\}}$ boosts QuantumAnomalyBot in time-series scenarios.
- $\text{crisk}_{\{\text{risk}\}}$ boosts PrudenceBot when elevated stakes exist.
- $\text{cdecc}_{\{\text{dec}\}}$ multiplies anomaly and prudence weights under high decoherence.

Weights normalize to $\tilde{w}_a = \frac{w_a}{\sum_b w_b}$. Transparency is signaled through consensus entropy

$$H = -\sum_a \tilde{w}_a \log \tilde{w}_a = -\sum_a \tilde{w}_a \log \tilde{w}_a$$

Final forecast is synthesized as

$$\hat{p} = \sum_a \tilde{w}_a p_a$$

Learning updates priors conservatively with forgetting factor γ and exponential loss scaling λ :

$$w_a^{\text{new}}(0) = (1 - \gamma)w_a(0) + \gamma \cdot \exp(-\lambda \cdot \text{loss}(p_a, y))$$

V. Interpretability and Critique with BDH

BDH designs monosemantic circuits enabling governance agents to:

- Trace decisive evidence shards
 - Detect shortcuts, spurious correlations, or contradictions with ground truth
 - Produce neuron- or feature-level activation summaries linked to citations
 - Block unsafe syntheses arising from unverifiable cues
-

VI. Memory and the Macro-Rewind

Mem0 archives full cases—questions, context, evidence shards, agent outputs, weights, decisions, outcomes. It enables:

- Episodic and semantic recall across sessions
 - Experience replay for weight and routing policy tuning
 - Macro-Rewind loop retrieving k-nearest historical cases for reflective governance
-

VII. Governance and Policy Integration

Precautionary escalation triggers when $D_{\text{th}} > \tau_{\text{h}} D_{\text{th}}$ for mmm consecutive windows, increasing confidence thresholds, broadening search, and activating external audits. Extremely low consensus entropy mandates monoculture audit justification. All artifacts—environment manifests, anomaly plots, JSON logs—are versioned for reproducibility and accountability.

VIII. End-to-End Workflow Example

New data flows signal triggers. The organizer forks paths for time-series analysis, contrarian evidence search, and financial ratio extraction. Agents generate component forecasts (QuantumAnomalyBot's decoherence score, PrudenceBot's scenario risks, EconomicRealismBot's impact mapping). CritiqueBot checks evidence. Synthesis adjusts weights dynamically. Forecasts emit with policy alerts and full audit trail. Policies update through Mem0 replays.

IX. Threat Model and Resilience

- **Adversarial Inputs:** Defended by source provenance, multi-modality anomaly clustering, and contradiction detection.
 - **Model Drift and Reward Hacking:** Guarded by Docent-style ARS review with red-team audits.
 - **Overconfidence:** A widened interval and elevated prudential weighting in low-evidence regimes, deferring synthesis as necessary.
-

X. Evaluation Plan

- **Quantitative:** Brier score, Expected Calibration Error (ECE), sharpness, AUC, early-warning lead times, also ablations removing QVR, BDH, Mem0.
 - **Qualitative & Governance:** Audit completeness, reproducibility, policy alignment, override tracking.
-

XI. Minimal Operable Demo

Implemented QuantumAnomalyBot with Prophet-based anomaly intervals, decoherence rate DDD, rewindability RRR, and graphical plus JSON output. Gating Meta-Agent applies context-aware weighting, computes consensus entropy, and synthesizes final prediction. Artifacts are persisted for forensic scrutiny.

XII. Roadmap to Full FORETELLS X

- Productionize Mem0Manager for Macro-Rewind and experience replay
 - Deploy DeepResearchBot with failure-aware tool use
 - Develop ScalingTrendsBot modeling benchmark evolution
 - Integrate second-order Reflexive Bayesian Networks (delay kernels & reflexivity)
 - Map technical risks into a harm ontology for calibrated policy triggers
-

XIII. Limitations

- QVR based on classical proxies rather than quantum hardware; metaphorical stability indicator
- BDH interpretability targets monosemanticity; residual polysemantic units require ongoing audit
- Live data governance demands strict provenance, de-duplication to avoid narrative feedback loops

- Routing policies risk regime overfitting without cross-regime validation and replay regularization
-

XIV. Security Considerations

- Log privacy with redaction in Mem0 for forensic auditability
 - Version pinning, artifact signing to secure supply chain
 - Kill switches enabling graceful degradation halting synthesis upon governance failures
 - Human-in-the-loop escalation channels for high-impact decisions
-

Conclusion

FORETELLS X is a governed cognitive architecture that not only forecasts but reasons, audits, remembers, and learns. By treating uncertainty as an epistemic resource and binding every claim to verifiable artifacts, FORETELLS crosses the boundary from clever forecaster to a foundation for Cognitive Sovereignty.

References

- Alon, U. (2006). *An introduction to systems biology: Design principles of biological circuits*. Chapman & Hall/CRC. <https://doi.org/10.1201/9781420011432>
- Goh, G., Cammarata, N., Voss, C., Carter, S., Petrov, M., Schubert, L., Radford, A., & Olah, C. (2021). Multimodal neurons in artificial neural networks. *Distill*, 6(3), e30. <https://doi.org/10.23915/distill.00030>
- Lindner, D., Kramár, J., Rahtz, M., McGrath, T., & Mikulik, V. (2023). Traer: Compiled transformers as a laboratory for interpretability (arXiv:2301.05062). *arXiv*. <https://arxiv.org/abs/2301.05062>
- Olah, C., Cammarata, N., Schubert, L., Goh, G., Petrov, M., & Carter, S. (2020). Zoom in: An introduction to circuits. *Distill*, 5(3). <https://doi.org/10.23915/distill.00024.001>
- Weiss, G., Goldberg, Y., & Yahav, E. (2021). Thinking like transformers (arXiv:2106.06981). *arXiv*. <https://arxiv.org/abs/2106.06981>
-