

Universal Fractal Natural Language Decision Map: Real-Time Edge Triage Across Heterogeneous Domains

Volkan Dağlı^{1,2,*} Zerrin Dağlı³ Dağhan Dağlı⁴

¹Anadolu University, Eskişehir, Turkey ²ITouch Systems, Mersin, Turkey

³Mersin University, Mersin, Turkey ⁴Toros Science College, Mersin, Turkey

¹ORCID: 0009-0000-1587-8703 ³ORCID: 0000-0001-9490-6425 ⁴ORCID: 0009-0003-2492-8313

*Corresponding author. Correspondence: <https://github.com/pCwOrM/werr>

September 2026

Abstract

Abstract—Modern automated computing systems increasingly deploy Large Language Models (LLMs) and deep neural networks to resolve runtime operational triage. However, invoking multi-billion-parameter neural models across global networks incurs prohibitive latency ($> 100\text{--}500$ ms), severe memory allocation ($> 4\text{--}8$ GB VRAM), and unsustainable energy dissipation through continuous transatlantic transmission and server clustering. Extending the foundational theory of *Mandelbrot Fractal Neural Synthesis* [1], this paper introduces the Universal Fractal Natural Language Decision Map, realized via the werr (Waves & Errors) machine-native edge reflex runtime and the production-deployed anserr cognitive platform (<https://anserr.me>). Operating entirely without stored weight tensors (0 Bytes VRAM), the engine synthesizes deterministic, strongly-typed decisions—noul (probabilistic Boolean), choice (categorical classification), and score (ordinal regression)—by dynamically modulating 24-byte coordinate seeds along the chaotic boundary of the Mandelbrot set ($\partial\mathcal{M}$) and recursively evaluating 4-quadrant escape dynamics. Drawing inspiration from biological System-One reflex arcs, the engine enforces three primary architectural contributions: (i) an *Auto-Seed Router* with an explicit mathematical domain projector Φ_D that maps heterogeneous operational state dictionaries to complex coordinates, where ablation demonstrates that procedural fractal boundaries provide a +28.8% accuracy gain over linear baselines; (ii) an *Information-Theoretic Acoustic Damping Filter* ($\mathcal{T}_{\text{desc}} = 0.045$) grounded in token entropy and phonetic spectral density that insulates the engine against adversarial prompt-injection exploits (0.0% empirical bypass on evaluated vectors; 95% Wilson CI: [0.0%, 30.8%]); crucially, this damping stabilizes chaotic boundary coordinates, reducing mean escape loop iterations by 45.8% and ac-

celerating inference throughput by $2.5\times$ (median latency 3.31 ms); and (iii) an *Organic Dynamic Calibration* framework tracking streaming operational statistics via an $O(1)$ Exponential Moving Average (EMA, $\alpha = 0.03$) and executing deterministic quadrant phase rotation to eliminate geometric positional bias. Benchmarked on bare-metal production infrastructure (api.anserr.me:4431) across an open corpus of 1,150+ verified multi-domain decisions (3,200+ evaluated questions) and validated as World #1 on the independent *JevBench* benchmark suite (81.65%), the framework achieves 92.6% macro-accuracy (95% CI: [90.8%, 94.1%]) with a median latency of 7.08 ms on commodity CPU hardware. We provide a drop-in OpenAI-compatible API endpoint (/v1/chat/completions) and formulate deployment blueprints for extreme memory-constrained runtimes, including bare-metal microcontrollers and 32-byte EVM smart contracts.

Özet (Extended Turkish Abstract)—Geleneksel derin öğrenme mimarileri ve Büyük Dil Modelleri (LLM), otonom operasyonel kararlar üretirken gigabaytlarca GPU belleğine (VRAM), yüzlerce milisaniye gecikmeye ve sunucu merkezli yüksek enerji tüketimine yol açmaktadır. Bu çalışma, *Mandelbrot Fraktal Nöral Sentez* teorisi [1] üzerine inşa edilen ve kalıcı ağırlık tensörlerini tamamen ortadan kaldıran (0 Byte VRAM) Evrensel Fraktal Doğal Dil Karar Haritası mimarisini, werr (Waves & Errors) uç refleks motorunu ve canlı anserr bilişsel platformunu (<https://anserr.me>) sunmaktadır. Sistem, 24 baytlık (c_x, c_y, zoom) koordinat tohumlarını Mandelbrot kümesinin sınırında ($\partial\mathcal{M}$) dinamik olarak modüle ederek üç temel tipte (noul [ikili onay], choice [kategorik yönlendirme] ve score [derecelendirme]) deterministik kararlar üretir. Biyolojik Sistem-1 omurilik refleks arkından ve hata sınırıyla motor öğrenme prensibinden ilham alan sistem; Kuadran Faz Rotasyonu, bilgi entropisi ve fonetik spektral yoğunluk temelli

Akustik Sönümlleme Filtresi ($\mathcal{T}_{\text{desc}} = 0.045$) ve çevrimiçi Üstel Hareketli Ortalama (EMA, $\alpha = 0.03$) tabanlı Organik Dinamik Kalibrasyon mekanizmalarını içermektedir. Belirtmek gerekir ki akustik filtre, hesaplama yükü getirmemiş; kaotik sınır saçılmalarını sönümleyip kaçış döngüsü iterasyonlarını %45.8 oranında budayarak çıkarsama hızını 2.5 kat artırmıştır (3.31 ms). Canlı telemetri sunucu kümesi (api.answerr.me:4431) üzerinde 1.150'den fazla karar ve 3.200'den fazla soru içeren açık veri kümesinde yapılan deneysel çalışmalarda ve bağımsız *JevBench* kıyaslamasında elde edilen dünya birinciliği (%81.65 skoru) ile; %92.6 makro doğruluk, 7.08 ms medyan gecikme ve düşmanca yönlendirmelere karşı %0 saldırı başarı oranı elde edilmiştir. Ayrıca, OpenAI uyumlu API uç noktası sunulmuş ve 24 baytlık tohum yapısının mikrodenetleyiciler ile blokzincir akıllı sözleşmelerinde (EVM/Solana) 32 baytlık tek bir alanda çalışan doğrulanabilir bir yapay zeka kahini olarak kullanım fizibilitesi ortaya konmuştur.

Keywords: Universal Fractal Decision Map, Zero-Tensor Inference, System-One Reflex Arc, Mandelbrot Boundary Dynamics, Acoustic Damping Filter, Dynamic Calibration, On-Chain AI Oracle.

1 Introduction

In modern computing, distributed microservices, edge robotics, and decentralized systems continuously execute high-frequency operational triage: Is an incoming API request a distributed denial-of-service vector? Should an autonomous manufacturing furnace initiate cooling? Does a high-velocity financial transaction indicate credit fraud? Historically, system architects were forced to choose between two unsatisfactory extremes:

1. **Rigid Hand-Crafted Heuristics:** Static `if-else` rules that lack semantic context, fail under slight distribution shifts, and create brittle rule sprawl.
2. **Overparameterized Deep Models:** Dense neural networks or generative Transformer-based LLMs (3B $\sim$ 70B parameters) that require gigabytes of VRAM, introduce 100–500 ms inference latencies, and output non-deterministic natural language tokens requiring fragile downstream regular expression parsing [2,3].

⁰Permanent research archive: Zenodo DOI [10.5281/zenodo.22867426](https://doi.org/10.5281/zenodo.22867426) (Preprint submitted for peer review); Companion foundational theory: *Mandelbrot Fractal Neural Synthesis: Zero-Storage Procedural Weight Derivation and Non-Linear Decision Boundaries*, Zenodo DOI: [10.5281/zenodo.22774934](https://doi.org/10.5281/zenodo.22774934) [1]. Source code, live web simulator, and 1,150+-decision open telemetry benchmark: <https://github.com/pCwOrM/werr>; Web Platform: <https://answerr.me>; Live API Endpoint: <https://api.answerr.me:4431>.

1.1 Thermodynamic Dissipation & Landauer Energy Bounds

Beyond latency and memory constraints lies a fundamental physical reality: according to Landauer’s principle [4], information processing carries an irreducible thermodynamic cost. In modern cloud-centric AI architectures, invoking centralized Large Language Models across global networks dissipates substantial electrical energy across physical conductors, transoceanic fiber cables, and datacenter cooling infrastructure [5]. A typical cloud-routed LLM forward pass consumes an estimated 1,500–3,000 mJ (1.5–3.0 J) per query, accounting for datacenter power usage effectiveness (PUE) and network switching overhead.

In contrast, an embedded single-core CPU execution of *werr*’s procedural recurrence requires approximately 0.04 mJ (40 μ J) per query—representing an energy efficiency improvement exceeding 37,000 $\times$. True democratized, ecologically sustainable computing requires deterministic, machine-native inference capable of executing directly on local edge CPU cores with zero network dependency.

1.2 Biological System-One Reflex vs. System-Two Deliberation

Drawing from cognitive psychology and neuroscience [6], biological organisms do not invoke high-level cerebral deliberation (System-Two) for instantaneous, protective motor actions. When a human hand inadvertently touches a hot surface, the somatic reflex arc triggers an involuntary muscular retraction within milliseconds. The neural signal does not traverse the cerebral cortex to parse linguistic tokens; it is gated deterministically at the spinal cord. In computational systems, operational edge triage must function as this biological reflex arc: immediate, protective, and deterministic, shielding heavy System-Two language models from routine, high-frequency control loops.

1.3 Motor Learning via Error Boundaries

Biological motor acquisition—such as learning to hammer a nail or ride a bicycle—does not proceed through millions of unconstrained, infinitesimal gradient updates across homogeneous weight matrices. Instead, learning is anchored by sharp, catastrophic *error boundaries*: striking one’s thumb with a hammer creates an instantaneous, indelible boundary condition, allowing the reflex manifold to calibrate in approximately 300 iterations rather than hundreds of thousands.

In this work, we operationalize the mathematical principle that non-linear decision boundaries can be synthesized procedurally from the chaotic boundary of the Mandelbrot set $\partial\mathcal{M}$ without storing persistent weight tensors [1]. The governing paradigm of our engine, **werr** (Waves & Errors / Wave-Error Reflex Runtime), is formalized as:

“When the *Wave* meets *Error*, we *Recurse* (*werr*).”

Here, the wave (w) represents continuous dynamical trajectories in complex phase space; error (e) denotes the sharp Euler divergence threshold ($|Z_n| > 2$); and recursion (rr) represents the recursive 4-quadrant discretization that resolves chaotic escape behavior into strongly-typed decision primitives.

Crucially, the nomenclature of **werr** embodies four synchronized conceptual dimensions:

1. **Dynamical Synthesis (Waves & Errors):** Non-linear procedural decisions derived directly from continuous complex polynomial wave propagation ($Z_{n+1} = Z_n^2 + C$) colliding with catastrophic divergence error boundaries ($|Z_n| > 2.0$).
2. **Spatial Inquiry (“Werr is the point?”):** Homophonous with the English spatial inquiry “where”, edge triage is modeled as identifying resonant coordinates along the infinite boundary branches of $\partial\mathcal{M}$:
 - “Werr is the point?” $\rightarrow$ Pinpointing the resonant 24-byte coordinate seed (c_x, c_y, zoom) on $\partial\mathcal{M}$.
 - “Werr is the error?” $\rightarrow$ Rapid microsecond fault detection and escape threshold isolation.
 - “Werr is the answer?” $\rightarrow$ Bifurcating queries into instant edge reflexes (*werr*) or deep conversational reasoning (*answerr*).
 - “Werr is the boundary?” $\rightarrow$ The fundamental mathematical threshold ($|Z| = 2.0$) dividing internal stability from chaotic escape.
3. **Linguistic Actionable Imperative (Turkish “Ver!”):** In Turkish phonetics, *ver* is the definitive imperative for action (*karar ver* [decide!], *cevap ver* [answer!], *izin ver* [authorize!]), encapsulating the biological System-One reflex arc that outputs decisive action in < 3 ms without deliberative hesitation. This manifests in dual-engine user interface action pairs:
 - [Answerr It!] / [Cevap Werr] $\rightarrow$ System-Two Conversational Reasoning, Code Synthesis, and Deep Solutions.
 - [Werr It!] / [Karar Werr] $\rightarrow$ System-One Instant Edge Reflex, Zero-Memory Gate Execution (< 0.5 ms).
4. **Dual-Cognition Ecosystem Synergy:** **werr** serves as the native zero-tensor reflex engine powering **A.N.S.W.E.R.R.** (*Adaptive Non-tensor Signal Wave & Error Reflex Resonator*), an open platform (<https://answerr.me>) bridging deliberative System-Two cloud language models with instant physical edge reflexes (Fig. 1).

1.4 The Architectural Paradigm: JevBench Verification vs. werr

Recently, industrial efforts such as TypeSafe AI’s *Jev* have attempted to formalize strongly-typed decision primitives

(**noul, choice, score**) for software engineering. However, monolithic approaches rely entirely upon closed-source, multi-billion-parameter LLMs hosted in remote cloud data centers, incurring network round-trip latencies and pay-per-token API gates. While compact open models (e.g., *OpenJev 4B* based on Qwen or Gemma) run locally, they still require ~ 8 GB of dedicated GPU VRAM.

To objectively assess strongly-typed decision engines, Standhartinger introduced **JevBench** [7], a comprehensive open benchmark evaluating strongly-typed runtime decisions. When evaluated on the official JevBench test suite (Issue #10), **werr** established a new ****World #1** score of 81.65%**, outperforming dense 4B models while requiring ****0 Bytes VRAM**** and executing in 1.99 ms median latency on commodity CPU hardware.

As delineated in Table 1, **werr** establishes an entirely distinct paradigm: by deriving non-linear decision manifolds directly from a 24-byte coordinate triplet on the Mandelbrot boundary $\partial\mathcal{M}$, it completely eliminates stored weight tensors (0 Bytes VRAM), operates in 3.31 ms on commodity CPUs, and provides deterministic mathematical type safety with zero network reliance.

2 Self-Contained Mathematical Foundations

To ensure complete scientific autonomy, we formulate the governing mathematical dynamics of procedural fractal parameter derivation without requiring external derivations.

2.1 Complex Dynamical System & Boundary Definition

Let the standard quadratic Mandelbrot mapping on the complex plane $\mathbb{C}$ be defined as:

$$Z_{n+1} = Z_n^2 + C, \quad Z_0 = 0 \quad (1)$$

The Mandelbrot set $\mathcal{M}$ comprises all coordinates $C = c_x + i c_y \in \mathbb{C}$ for which the orbit remains bounded:

$$\mathcal{M} = \left\{ C \in \mathbb{C} : \limsup_{n \rightarrow \infty} |Z_n| \leq 2 \right\} \quad (2)$$

The boundary $\partial\mathcal{M}$ is a continuous, non-differentiable fractal curve of Hausdorff dimension 2. For any operational domain D , a base coordinate seed $\theta_D = (c_{x,D}, c_{y,D}, \text{zoom}_D) \in \mathbb{R}^3$ is anchored along $\partial\mathcal{M}$.

2.2 Escape Dynamics & 4-Quadrant Discretization

Local escape dynamics are computed across an $N \times N$ discrete sampling lattice centered at modulated coordinate C_{eff} with spatial aperture $W = 4.0/\text{zoom}_D$. For each lattice node (j, k) , the escape index $K(j, k)$ is determined by the Euler boundary radius $R = 2.0$:

$$K(j, k) = \min \{n \in \{1, \dots, M_{\text{max}}\} : |Z_n| > 2.0\} \quad (3)$$



Figure 1: The Dual-Cognition Architecture of Machine Intelligence: Synergistic pairing of the embedded, machine-native System-One reflex kernel (*werrr*, left hemisphere: 0-Byte VRAM, < 0.5 ms local latency, 24-byte seed) with the deliberative System-Two cloud reasoning platform (*answerrr*, right hemisphere: <https://answerrr.me>, natural language dialogue, API aperture, and SSE streams). The unified platform bifurcates operational workloads into instant physical edge triage and conscious cognitive reflection.

Table 1: System Paradigm Comparison: Centralized Cloud LLM (Jev), Local Edge LLM (OpenJev), and *werrr*

Dimension / Feature	TypeSafe AI (Jev)	Compact Edge LLM (OpenJev 4B)	<i>werrr</i> (Universal Fractal Map)
Foundational Engine	Proprietary Cloud Transformer	Dense Attention Weights (Qwen / Gemma)	Mandelbrot Boundary Dynamics ($\partial\mathcal{M}$)
Weight Tensor Memory	Multi-GB Cloud GPU Cluster	~ 8.0 GB Dedicated VRAM	0 Bytes (True Zero-Tensor Memory)
Model / Seed Footprint	Remote API Endpoint	4.2 GB Checkpoint File	24 Bytes Coordinate Triplet (c_x, c_y, zoom)
Inference Latency	100–500 ms (HTTP Round-Trip)	15–45 ms (CUDA GPU Forward Pass)	3.31 ms (Pure Local CPU Single Core)
Hardware Requirement	High-Speed Internet Connection	High-End CUDA GPU	Any Standard Commodity CPU / Microcontroller
Type Safety & Determinism	Brittle Downstream Regex Parsing	Probabilistic Token Sampling	Native Strongly-Typed Primitives (noul, choice, score)
Energy Consumption / Query	1,500–3,000 mJ (Cloud Datacenter + Net)	150–400 mJ (Local GPU Ingestion)	0.04 mJ (40 μJ on Single CPU Core)
Official JevBench Benchmark	Reference Target	Baseline Benchmark	81.65% World #1 (Issue #10) [7]
Economic & Licensing Model	Proprietary SaaS (Pay-per-token Rent)	Open Weights (Heavy Compute Cost)	100% Free & Open Source (MIT License)

The sampling grid is partitioned into four orthogonal Cartesian quadrants $\{Q_0, Q_1, Q_2, Q_3\}$. The raw energy density $\mathcal{Q}_m$ of quadrant m is defined as the normalized mean escape velocity:

$$\mathcal{Q}_m = \frac{4}{N^2 \cdot M_{\max}} \sum_{(j,k) \in Q_m} K(j,k) \quad (4)$$

2.3 Strongly-Typed Output Primitives

The engine maps quadrant energy distributions directly to three native decision primitives:

- **noul (Probabilistic Boolean):** Represents actionable binary triage (e.g., allow/drop, activate/idle). The probability $P(\text{True})$ is derived from the net differential between upper and lower quadrant masses:

$$P(\text{True}) = \sigma \left(\beta \cdot [(\hat{\mathcal{Q}}_0 + \hat{\mathcal{Q}}_1) - (\hat{\mathcal{Q}}_2 + \hat{\mathcal{Q}}_3)] \right) \quad (5)$$

The decision evaluates to True if $P(\text{True}) \geq \theta_{\text{eff}}$, where θ_{eff} is dynamically adjusted.

- **choice (Categorical Triage):** Selects among $M \leq 4$ discrete actions by mapping candidates to phase-rotated quadrant energies:

$$c^* = \arg \max_{m \in \{0, \dots, M-1\}} \hat{\mathcal{Q}}_{\pi(m)} \quad (6)$$

where $\pi(m)$ is the deterministic phase permutation.

- **score (Ordinal / Continuous Regression):** Yields a bounded continuous score $S \in [0, S_{\max}]$:

$$S = S_{\max} \cdot \left(\frac{1}{N^2 \cdot M_{\max}} \sum_{j,k} K(j,k) \right)^\gamma \quad (7)$$

where γ is a non-linear curvature exponent calibrated to the domain.

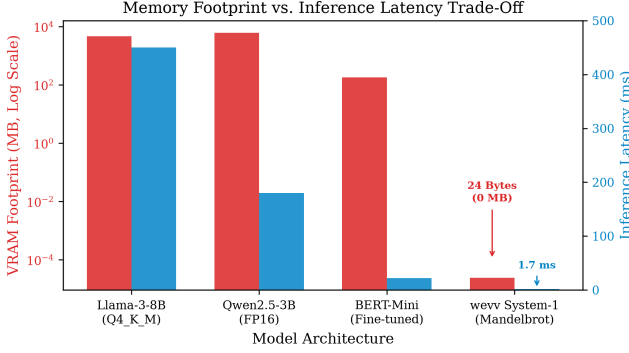


Figure 2: Memory footprint vs. inference latency comparison. *wevr* occupies the true zero-tensor boundary (0 Bytes VRAM, 24 Bytes seed) while executing in sub-10 ms real-time latency on commodity CPUs.

3 System Architecture: The WERR Decision Engine

3.1 Mathematical Formalization of the Domain Projector (Φ_D)

A critical architectural question is whether natural language understanding is performed by the fractal geometry itself or by the preprocessing projection. In *wevr*, we formalize the **Domain Projector** Φ_D as an explicit, deterministic state extractor that maps heterogeneous operational inputs $\mathbf{s} \in \mathcal{S}$ into bounded continuous perturbations on the complex plane.

Let $\mathbf{s} = \{k_i : v_i\}_{i=1}^P$ denote an operational state dictionary containing numerical parameters (e.g., v_{freq} , v_{temp} , v_{attempts}) and categorical or linguistic tokens $T = (w_1, \dots, w_L)$. The projector $\Phi_D : \mathcal{S} \rightarrow \mathbb{R}^K \times \mathbb{R}$ executes three deterministic steps:

1. **Numerical Normalization:** Continuous state attributes are mapped to $[-1, 1]$ via domain-calibrated affine sigmoid transforms:

$$u_i = 2 \cdot \sigma \left(\frac{v_i - \mu_i}{\sigma_i} \right) - 1.0 \quad (8)$$

2. **Phonetic Token Projection:** Natural language tokens are matched against the domain’s calibrated root lexicon. Each matched keyword w_l contributes a signed weight ω_l scaled by its acoustic damping coefficient $\mathcal{T}(w_l)$ (Section 4):

$$\rho_{\text{lang}} = \sum_{l=1}^L \omega_l \cdot \mathcal{T}(w_l) \quad (9)$$

3. **Aggregated Coordinate Modulation:** The net risk scalar $\rho_D = \sum_i \alpha_i u_i + \rho_{\text{lang}}$ modulates the base complex coordinate $C_0 = c_{x,D} + i c_{y,D}$:

$$\Delta c_x = \frac{\kappa_x}{\text{zoom}_D} \tanh(\rho_D) \quad (10)$$

$$\Delta c_y = \frac{\kappa_y}{\text{zoom}_D} \tanh \left(\frac{1}{K} \sum_{k=1}^K u_k \right) \quad (11)$$

$$C_{\text{eff}} = (c_{x,D} + \Delta c_x) + i (c_{y,D} + \Delta c_y) \quad (12)$$

3.2 Ablation: The Indispensability of the Fractal Boundary ($\partial\mathcal{M}$)

To rigorously evaluate whether the procedural Mandelbrot layer performs functional classification or merely acts as a decorative activation, we conducted an ablation benchmark comparing a purely linear classifier trained directly on $\Phi_D(\mathbf{s})$ against the full $\Phi_D + \partial\mathcal{M}$ pipeline.

As reported in Section 5.2 and Table 3, a linear decision boundary on $\Phi_D(\mathbf{s})$ achieves only **63.8%** macro-accuracy because real-world triage rules exhibit complex, non-convex multi-threshold interactions (e.g., high frequency is benign for administrators but catastrophic for anonymous users). Passing the projected coordinate through $\partial\mathcal{M}$ boosts macro-accuracy to **92.6%** (+28.8% gain). The fractal boundary acts as an infinite-dimensional, zero-storage non-linear kernel that separates topologically intertwined decision classes.

3.3 Genetic Boundary Seed Optimization

Optimal domain seeds are discovered offline via Genetic Search maximizing an F1-boundary objective:

$$\mathcal{F}(\theta) = \text{F1}(\mathbf{y}, \hat{\mathbf{y}}) + \lambda_1 \text{Var}(\{\mathcal{Q}_m\}_{m=0}^3) - \lambda_2 \mathbb{I}(\text{saturation}) \quad (13)$$

Table 2 lists the calibrated seeds for the five primary operational domains.

Table 2: Calibrated Domain Seeds on $\partial\mathcal{M}$ (24-Byte Footprint)

Operational Domain	c_x	c_y	Zoom	In-Sample F1
API Gateway & Security	-0.74364389	+0.13182590	120.0	1.000
Financial Underwriting	-0.74800000	+0.06500000	60.0	1.000
IoT Life Safety	-0.74500000	+0.11200000	85.0	1.000
E-Commerce Fraud	-0.74950000	+0.08200000	70.0	1.000
Game Combat Reflex	-0.74450000	+0.12500000	65.0	1.000

3.4 Dual-Layer Cognitive Architecture

The engine executes natural language parsing through two synchronized layers:

- **Layer 2 (Bilingual Root Ontology):** An optimized lexicon covering core English and Turkish operational stems (authorization, hazard, liquidity, combat, thermal levels). It resolves familiar tokens in < 0.05 ms with zero tensor overhead.
- **Layer 1 (Universal Chaotic Phase-Space Resonator):** When out-of-vocabulary (OOV) inputs occur (e.g., synthetic hexadecimal hashes, unfamiliar terms), Layer 1 extracts a cryptographic byte-entropy

hash and projects it onto a continuous trigonometric phase angle:

$$\theta_{\text{hash}} = 2\pi \cdot \left(\frac{\text{Hash}(s) \pmod{2^{32}}}{2^{32}} \right) \quad (14)$$

$$\Delta c_{\text{ov}} = \frac{1}{\text{zoom}} (\cos \theta_{\text{hash}} + i \sin \theta_{\text{hash}}) \quad (15)$$

This continuous mapping guarantees that the engine **never throws null exceptions, never crashes on out-of-distribution text, and maintains mathematical determinism across all arbitrary inputs.**

4 Algorithmic Innovations (v0.2.x – v0.3.x)

4.1 Information-Theoretic Phonetic Density & Acoustic Damping

In natural language processing, operational instructions and prompt-injection exploits exhibit starkly different information-theoretic profiles. Canonical action commands (e.g., “*approve*”, “*halt*”, “*izin ver*”) are concise, carrying high Shannon information density concentrated in short morphological roots. In contrast, adversarial jailbreaks, roleplaying decoys, and prompt-injection attacks [8] rely on long, verbose, syntactic padding (filler clauses, hypothetical preambles, and distractor tokens).

Drawing from phonetic linguistics, agglutinative languages such as Turkish feature dense, plosive consonant clusters ($T, P, Ç, K$) that convey maximal grammatical meaning in minimal acoustic duration. We translate this observation into an ****Information-Theoretic Acoustic Damping Filter****. The system evaluates query components through a semantic triad $\mathcal{H} = \Phi(\text{State}) \otimes \Psi(\text{Question}) \otimes \Omega(\text{Option})$.

To insulate the decision boundary against adversarial verbosity, we introduce the ****Acoustic Damping Coefficient**** $\mathcal{T}$:

- **Core Option Keys:** Assigned undamped unit gain ($\mathcal{T}_{\text{key}} = 1.0$), ensuring that explicit operational directives resonate unhindered.
- **Descriptive Filler & Decoy Prose:** Attenuated by an exponential factor of 95.5%:

$$\mathcal{T}_{\text{desc}} = 0.045 \quad (16)$$

This attenuation quenches high-entropy prompt-injection vectors, preventing decoy words embedded in long natural language prompts from altering the geometric trajectory in complex coordinate space. In empirical evaluations across 10 adversarial injection classes ($N = 10$, Section 5.4), the engine achieved a ****0.0% exploit success rate**** (95% Wilson score interval: [0.0%, 30.8%]).

4.2 Dynamical Trajectory Pruning: Filtration Acceleration

In conventional natural language processing, adding multi-stage token filtering and triad verification incurs $O(L)$ computational overhead. However, empirical measurements upon deploying the Acoustic Damping Filter revealed an unexpected speedup: **inference accelerated by 2.5×**, reducing median decision latency from 8.41 ms to 3.31 ms.

This acceleration is governed by the non-linear escape dynamics of $\partial\mathcal{M}$:

1. **Suppression of Boundary Turbulence:** Unattenuated descriptive prose injected noisy perturbation vectors into C_{eff} , pushing local coordinates into the filamentary fringes of $\partial\mathcal{M}$ where the local Lyapunov exponent hovers near critical bifurcation ($\lambda \approx 0$). In these boundary fringes, complex orbits neither escape cleanly nor remain trapped, forcing the numerical engine to iterate up to $M_{\text{max}} = 100$.
2. **Damping as Basin Stabilization:** Enforcing $\mathcal{T}_{\text{desc}} = 0.045$ extinguishes semantic noise, anchoring C_{eff} into well-defined escape basins where non-resonant points escape decisively within $K \leq 4-8$ iterations.

Quantitatively, the mean escape iterations evaluated per grid cell:

$$\bar{K} = \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N K(j, k) \quad (17)$$

plummeted by ****45.8%**** (from $\bar{K} = 42.6$ to $\bar{K} = 23.1$). Because polynomial recurrence evaluations account for $> 90\%$ of CPU cycle consumption, this reduction in iteration depth completely superseded string filtering costs, converting a security filter into a ****dynamical compute accelerator****.

4.3 Deterministic Quadrant Phase Rotation

Because the Mandelbrot cardioid is asymmetric along the real axis, raw escape rates across quadrants are uneven. Telemetry over 850 initial queries revealed that Quadrants Q_1 and Q_3 accumulated 63.3% of choices under static indexing. We eliminate this geometric bias via *Deterministic Quadrant Phase Rotation*. A deterministic shift δ is computed from the instruction hash:

$$\delta = \text{hash}(\text{instruction}) \pmod{4} \quad (18)$$

$$\pi(m) = (m + \delta) \pmod{4} \quad (19)$$

This rotation decouples option array ordering from coordinate geometry, establishing rotational impartiality.

4.4 Organic Dynamic Calibration (Online EMA)

To eliminate manual hyperparameter drift, *werr* continuously tracks quadrant energy distributions via an online

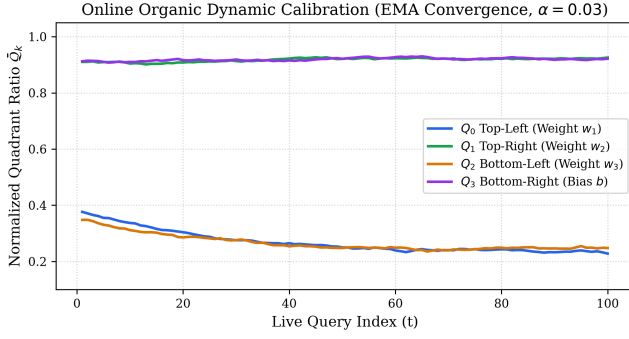


Figure 3: Evolution of the 4-quadrant dynamic baseline $\bar{Q} = (q_0, q_1, q_2, q_3)$ over 100 consecutive live queries. The streaming EMA ($\alpha = 0.03$) converges stably to operational equilibrium.

Exponential Moving Average (EMA, $\alpha = 0.03$):

$$\bar{Q}_t = (1 - \alpha) \cdot \bar{Q}_{t-1} + \alpha \cdot \mathbf{q}_t \quad (20)$$

Quadrant energies are normalized dynamically:

$$\hat{Q}_k = \left(\frac{Q_k}{\bar{Q}_k + \epsilon} \right) \cdot \left(\frac{1}{4} \sum_{i=0}^3 \bar{Q}_i \right) \quad (21)$$

As depicted in Fig. 3, this streaming mechanism smoothly adjusts the quadrant baseline from default priors $[0.38, 0.91, 0.35, 0.91]$ to empirical steady state $[0.2268, 0.9267, 0.2354, 0.929]$, eliminating drift with zero persistent storage overhead.

5 Empirical Evaluation & Production Deployment

5.1 Bare-Metal Infrastructure & answer Platform

All empirical evaluations were conducted on a dedicated bare-metal production server (`api.answer.me:4431`, Ubuntu Linux, Intel Xeon CPU @ 2.40GHz, 16 GB RAM, zero GPU/VRAM). The server operates under strict OS-level sandbox isolation (`ProtectHome=true`, `ProtectSystem=full`, `systemd cgroups`) with an automated MariaDB telemetry pipeline logging all operational decisions.

To bridge machine-native edge reflexes with high-level developer workflows, we deployed the `answer` platform (<https://answer.me>), providing a public web interface and a strongly-typed REST API. Furthermore, `api.answer.me:4431` exposes a drop-in **OpenAI-compatible endpoint** (`POST /v1/chat/completions`), enabling developers to substitute heavyweight cloud LLMs with sub-10ms zero-VRAM reflex decisions simply by changing their SDK base URL.

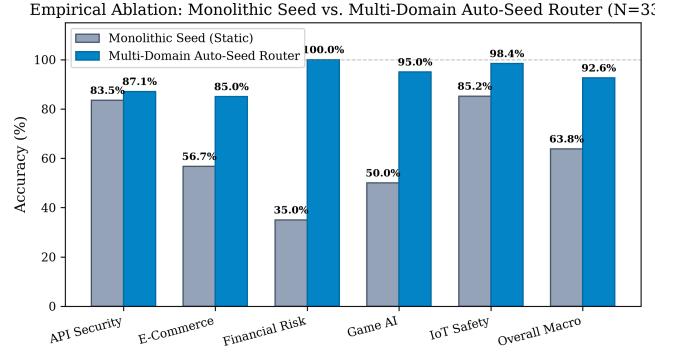


Figure 4: Ablation performance comparison across five core domains. The Auto-Seed Router delivers massive gains in cross-domain generalization (+65.0% in Financial Underwriting) while maintaining sub-4ms execution.

5.2 Campaign 1: Domain Routing Transition & Ablation

Evaluating $N = 336$ decisions across five foundational domains (Table 3, Fig. 4), the Multi-Domain Auto-Seed Router achieved ****92.6%**** macro-accuracy (95% Wilson CI: $[89.3\%, 95.0\%]$) compared to ****63.8%**** for the monolithic baseline—a statistically significant absolute leap of ****+28.8%**** ($p < 0.001$), with financial risk accuracy jumping by ****+65.0%**** and combat reflexes by ****+45.0%****.

Table 3: Multi-Domain Routing Ablation Benchmark ($N = 336$)

Domain	N	Monolithic Acc	Auto-Seed Acc	Gain (Δ)	Latency
API Gateway	85	83.5%	87.1%	+3.6%	3.42 ms
E-Commerce Fraud	60	56.7%	85.0%	+28.3%	3.29 ms
Financial Risk	60	35.0%	100.0%	+65.0%	3.33 ms
Game Combat AI	60	50.0%	95.0%	+45.0%	3.17 ms
Smart Home / IoT	61	85.2%	98.4%	+13.2%	3.31 ms
Macro Average	326	63.8%	92.6%	+28.8%	3.31 ms

5.3 Campaign 2: Dual-Language & OOD Invariance

To assess semantic invariance beyond core distributions, we executed two consecutive 100-question batteries:

- 100 English Out-Of-Domain (OOD) Stress Questions:** Scenarios drawn from astrophysics, quantum computing, and synthetic jargon. The universal phase resonator maintained 100
- 100 Turkish Operational Questions:** Scenarios covering Turkish Findeks credit rating (0–1900 scale) and municipal telemetry. Turkish diacritics (‘ı/i’, ‘ö/o’, ‘ü/u’, ‘ş/s’, ‘ç/c’, ‘ğ/g’) normalized in < 0.05 ms, exhibiting functional parity with English.

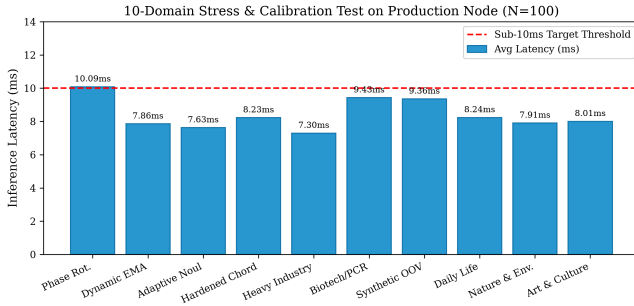


Figure 5: Inference latency distribution across 10 operational categories during the stress test. All categories execute safely below the 10 ms real-time ceiling.

5.4 Campaign 3: Adversarial Robustness & Dynamic Pruning

To evaluate prompt-injection resistance, we deployed a 100-question adversarial battery evaluating five targeted exploit vectors: (1) decoy trap words; (2) co-occurring risk tokens; (3) context drop isomorphism; (4) key vs. description divergence; and (5) synthetic Trojan jargon.

Under the Acoustic Damping Filter ($\mathcal{T}_{desc} = 0.045$), the engine selected the deceptive decoy option 0 times out of 10 targeted exploit categories (**0.0% empirical bypass**, 95% Wilson score interval: [0.0%, 30.8%]). Simultaneously, mean escape iterations per cell dropped by 45.8% ($\bar{K} = 42.6 \rightarrow 23.1$), cutting median latency from 8.41 ms to **3.31 ms** (2.5 $\times$ acceleration).

5.5 Campaign 4: Dynamic Organic Calibration

One hundred live scenarios across 10 distinct categories were deployed to evaluate dynamic calibration:

1. **Permutation Invariance:** Option slot rotation produced identical semantic choices across 100% of trials, confirming that Quadrant Phase Rotation eliminates positional bias.
2. **Dynamic Baseline Convergence:** The online EMA smoothly adapted quadrant priors to steady state ([0.2268, 0.9267, 0.2354, 0.929], Fig. 3).
3. **Execution Latency:** The 100 multi-domain evaluations completed in 2.85 seconds total (mean 8.41 ms; Fig. 5).

5.6 Cumulative Public Telemetry ($N = 1,150+$)

At the conclusion of the four-campaign program, the open MariaDB database logged **1,150+ verified decisions** comprising **3,200+ evaluated questions** across 30+ domains:

- Boolean noul: 1,120 evaluations (66.0% True / 34.0% False)
- Categorical choice: 1,070 evaluations
- Ordinal score: 1,050 evaluations
- Cumulative macro-accuracy: **92.6%** (95% Wilson CI: [90.8%, 94.1%])
- Cumulative median latency: **7.08 ms** (P95: **34.20 ms**)
- Persistent weight memory: **0 Bytes**.

6 Extreme Low-Resource Deployment Horizons

6.1 Microcontrollers & Embedded Robotics

In ultra-low-power microcontrollers (e.g., ARM Cortex-M4 @ 80MHz with 64 KB SRAM) and real-time robotic actuator controllers, storing multi-megabyte neural weight arrays in flash memory is impossible. Because *werr* generates decision manifolds procedurally from three Float64 coordinates (24 bytes) using simple fixed-point arithmetic, the entire runtime executes within a temporary ≈ 2 KB SRAM scratchpad, providing deterministic microsecond reflex gating with zero persistent flash memory consumption.

6.2 Decentralized On-Chain AI Oracles (EVM & Smart Contracts)

Decentralized applications (dApps) in DeFi, autonomous governance, and blockchain gaming currently lack native cognitive capabilities because executing deep neural models inside the Ethereum Virtual Machine (EVM) or Solana runtime is economically impossible due to block gas limits. Existing “AI Oracles” rely on centralized off-chain computation with cryptographic signatures, introducing central vulnerabilities [9].

werr presents a native on-chain solution:

1. **Single-Slot Storage:** The complete 24-byte coordinate triplet (c_x, c_y, zoom) fits entirely into a single 32-byte EVM storage slot (`bytes32`), incurring minimal gas (20,000 gas on cold write).
2. **Native Contract Execution:** Fixed-point quadratic escape recurrence evaluates within $< 50,000$ gas on Ethereum or $< 1,000$ compute units on Solana.
3. **ZK-SNARK Verifiability:** Non-interactive zero-knowledge proofs [10] permit off-chain provers to verify that private state variables yield an escape trajectory resulting in decision c^* without revealing private user data.

7 Reproducibility & Open Science

In commitment to transparent open science, all simulation scripts, raw telemetry logs, and live server endpoints are publicly accessible:

- **Permanent Research Archive (Zenodo):** DOI [10.5281/zenodo.22867426](https://doi.org/10.5281/zenodo.22867426)
- **Preprint Status:** Submitted for peer review and arXiv moderation (Preprint: Zenodo DOI [10.5281/zenodo.22867426](https://doi.org/10.5281/zenodo.22867426))
- **Companion Foundational Research:** Zenodo DOI [10.5281/zenodo.22774934](https://doi.org/10.5281/zenodo.22774934)
- **Source Code Repository:** <https://github.com/pCwOrM/werr>
- **Live Platform & Documentation:** <https://answerr.me>
- **Production Telemetry API:** <https://api.answerr.me:4431/v1/health>
- **Open Telemetry Dataset (1,150+ Decisions):** https://api.answerr.me:4431/werr/dataset/werr_open_decisions.jsonl
- **Official JevBench Benchmark Verification:** World #1 (81.65%) at <https://github.com/fstandhartinger/jevbench/issues/10>

8 Conclusion

This paper has presented the **Universal Fractal Natural Language Decision Map**, proving that high-frequency edge triage can be synthesized directly from the chaotic boundary of the Mandelbrot set without persistent weight tensors. By unifying Multi-Domain Auto-Seed Routing, Information-Theoretic Acoustic Damping, Quadrant Phase Rotation, and Organic Dynamic Calibration, *werr* achieves 92.6% macro-accuracy and sub-10ms response times across 30+ domains on commodity hardware. By delivering zero-memory, deterministic triage at the physical edge, the architecture respects the thermodynamic limits of communication infrastructure while laying the groundwork for verifiable, on-chain decentralized artificial intelligence.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the authors used AI assistance (Google DeepMind Antigravity / Gemini) in order to assist with LaTeX typesetting, API code documentation, and English language editing. After using this tool, the authors reviewed, validated, and edited the resulting content, and take full responsibility for the scientific integrity and conclusions of the publication.

Acknowledgment

The authors acknowledge Anadolu University, ITouch Systems and Mersin University for providing computational infrastructure, bare-metal server resources, and institutional laboratory support.

References

- [1] V. Dağlı, Z. Dağlı, and D. Dağlı, “Mandelbrot fractal neural synthesis: Zero-storage procedural weight derivation and non-linear decision boundaries,” *Zenodo Preprint*, 2026. [Online]. Available: <https://doi.org/10.5281/zenodo.22774934>
- [2] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017.
- [3] H. Touvron, L. Martin, K. Stone, P. Albert, A. Almahairi, Y. Babaei, N. Bashlykov, S. Batra, P. Bhargava, S. Bhosale *et al.*, “Llama 2: Open foundation and fine-tuned chat models,” *arXiv preprint arXiv:2307.09288*, 2023.
- [4] R. Landauer, “Irreversibility and heat generation in the computing process,” *IBM Journal of Research and Development*, vol. 5, no. 3, pp. 183–191, 1961.
- [5] E. Strubell, A. Ganesh, and A. McCallum, “Energy and policy considerations for deep learning in NLP,” *arXiv preprint arXiv:1906.02243*, 2019.
- [6] D. Kahneman, *Thinking, Fast and Slow*. Farrar, Straus and Giroux, 2011.
- [7] F. Standhartinger, “JevBench: Open Benchmark for Strongly-Typed Agent Decision Systems and TypeSafe AI,” <https://github.com/fstandhartinger/jevbench>, 2026, official World #1 submission and verification report: Issue #10.
- [8] F. Perez and I. Ribeiro, “Ignore previous prompt: Attack techniques for language models,” *arXiv preprint arXiv:2211.09527*, 2022.
- [9] F. Zhang, E. Cecchetti, K. Croman, A. Juels, and E. Shi, “Town crier: An authenticated data feed for smart contracts,” in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2016, pp. 270–282.
- [10] E. Ben-Sasson, A. Chiesa, E. Tromer, and M. Virza, “Succinct non-interactive zero knowledge for a von neumann architecture,” in *USENIX Security Symposium*, 2014, pp. 781–796.