

Externalization Engine:

Creation → We Create → Create Us

Jieqi Liu

Independent Researcher

jackey.l.gene@outlook.com

v1.6 · May 2026

Code & experiments: <https://github.com/JackeyLGene/GBE>

Abstract

This paper introduces a non-trained computational instrument for detecting structure in sequential data. The instrument maintains a finite memory of vector frames, merges similar inputs, prunes weakly reinforced frames, and reads the surviving attractors as centroids. Independent readings of the same stream are compared by cross-harm, a scalar measure of centroid divergence; stable centroids can also be written to an externalized Codex and tested across generations. No learned weights, supervised labels, or domain-specific templates are used during processing. Parameters are fixed or swept within pre-declared families and evaluated with shuffled, ablated, and baseline controls.

The instrument is evaluated across four domains. In Bach's *Well-Tempered Clavier*, a dual-encoding Self recovers tonal structure without chord templates, key signatures, or frequency pre-classification; transposition equivariance holds across all tested non-zero shifts, and a five-stage Codex selection loop distinguishes confirmed, supplemented, and frozen inherited entries. In DNA, the same architecture detects exon/intron structural separation in human transcripts ($d=-0.97$; $n=200$ transcripts, 10,606 windows), while a single-gene fork-column calibration on AHSB, a gene identified in positive-selection scans, yields $d=+1.51$; both signals collapse under shuffling. In RNA, temporal-lens divergence locates annotated CDS stop boundaries at 98-100% within ± 3 windows and CDS start boundaries at 66% within ± 3 windows (75% within ± 5); an internal-ATG control indicates that the signal reflects transition magnitude rather than AUG motif detection. In UN diplomatic voting, the instrument is used prospectively: the 2025 great-

power alignment displacement is the highest substantive reading in the 1946-2025 record, with falsifiable criteria specified for 2026-2028.

These results support a limited claim: a small, zero-training centroid detector can reveal order-dependent structure across heterogeneous streams when encodings preserve the relevant relational constraints. The current architecture remains a passive structural detector. It demonstrates centroid formation, cross-reading divergence, and Codex selection, but active Codex operation, in which inherited memory directly shapes ongoing processing under boundary conditions, is established as the immediate next step.

The complete codebase, experiment scripts, and guides to reproduce all findings are permanently archived on Zenodo and concurrently available at <https://github.com/JackeyLGene/GBE>.

Introduction

This paper asks a narrow operational question: can a small, non-trained instrument detect structure in streams without labels or domain-specific rules? The answer proposed here is a centroid detector. It keeps a limited memory, merges inputs that are similar, prunes frames that are not reinforced, and records the surviving attractors. No neural network is trained. No domain labels are supplied during processing. Parameters are fixed or swept and reported with controls (§2, Reproducibility).

The paper is self-contained. It belongs to a broader trilogy, but the reader does not need to have read GEME or BGM. GEME and BGM are not prerequisites for accepting the claims below; they are the architecture's provenance and calibration history. GEME supplies the minimal frame economy; BGM introduces endogenous time, bridge dynamics, and the first biological/cultural tests. EE asks whether the resulting detector can externalize and select structure across generations. The only idea required here is finite-memory structural detection: when a stream is repeatedly compressed by merge and prune, stable centers of mass emerge. This paper tests whether those centers, measured across independent encodings and across generations, can reveal structure in music, diplomatic voting, DNA, and RNA.

The philosophical motivation is externalization: information written into a medium that outlasts its source. Writing, notation, genomes, institutions, and digital archives all preserve structure beyond the lifespan of any single agent. The operational question is whether this can be made runnable. In this paper, a Codex is a minimal externalized memory: stable centroids are written out, inherited, confirmed, frozen, or ignored by later runs.

Three historical contexts are relevant but not required for following the experiments. Clark and Chalmers (1998) argued that cognition can extend into external media. Luhmann (1984) described society as operationally closed communication systems coupled through a shared medium. Aaronson (2011, 2013) emphasized computational cost as central to questions of mind and knowledge. The present work does not assume

these theories. It provides a small computational instrument whose behavior can be tested directly.

Contemporary dialogue. The 2024-2026 context matters for positioning. In AI, agent-memory systems make persistent memory a first-class architectural problem: Generative Agents store natural-language experience streams and synthesize reflections (Park et al., 2023), while MemGPT manages memory tiers around a fixed context window (Packer et al., 2023). Agent evaluation work also stresses that cost, reproducibility, and hold-out discipline matter as much as headline capability (Kapoor et al., 2024). The present instrument asks a lower-level question: before a system retrieves or reflects in language, what structure should be preserved at all?

In automated science, the AI Scientist and science-exocortex proposals treat research itself as an externalized loop of ideation, coding, execution, writing, and review (Lu et al., 2024; Yager, 2024). GBE is not a general autonomous scientist; it is a small runnable artifact for testing memory pressure, compression, and inheritance. In cultural evolution, collective-brain and tacit-knowledge models study how structure survives beyond individuals (Muthukrishna and Henrich, 2016; Miton and DeDeo, 2022). The WTC experiment supplies a computational version of that transmission problem.

Finally, genomic and UTR language models show that biological sequences can be treated as structure-bearing streams (Chu et al., 2024; Sanabria et al., 2024). The biological experiments here are not offered as domain-leading biology models. They are cross-substrate stress tests for a blind structural detector. In this sense, the surrounding literatures are not targets to be displaced. They are converging on the same boundary from above: larger agents, richer memory, automated research loops, and data-intensive sequence models. This paper approaches the boundary from below, asking how much of that convergence remains visible in a minimal merge-prune instrument.

Organization. Section 1 introduces the minimal instrument using four terms only: frame, centroid, cross-harm, and Codex. Section 2 reports the experimental evidence: UN diplomacy (§2.1), WTC music (§2.2), DNA evolutionary archives (§2.3), and RNA blind boundary detection (§2.4). Section 3 discusses implications and limitations. Section 4 concludes. Detailed instrument methodology, calibration, encoding, Faraday measurements, prediction protocol, scaffold history, negative-result decisions, and reproducibility are in Supplementary Material S1-S9.

1. Minimal Instrument

The main text uses four terms.

Frame. A frame is a vector with a weight. It is the instrument's current memory of similar inputs.

Centroid. When many inputs merge into the same frame and that frame survives pruning, its vector converges toward the center of mass of that part of the stream. This surviving attractor is a centroid.

Cross-harm. The same stream can be read in two ways: for example, pitch and rhythm in music, fast and slow temporal lenses in RNA, or different conservation lenses in DNA. Cross-harm is the scalar disagreement between the centroids produced by those readings. Low cross-harm means the readings converge to compatible structure. High cross-harm means they diverge.

Codex. A Codex is an externalized store of surviving centroids. A later run can inherit it. Entries that match the environment gain weight; entries that do not match freeze or decay. In the current paper, Codex formation and selection are tested. Active Codex operation — where inherited memory directly shapes ongoing processing under boundary conditions — remains a limitation (§3.5).

The entire measurement loop can be summarized as:

input stream

-> encode events as vectors

-> merge similar vectors into frames

-> prune weak frames

-> read surviving centroids

-> compare independent readings as cross-harm

-> write stable centroids into Codex

-> test inheritance by confirmation, freezing, decay, and controls

The implementation contains more named parts: Geruon, Self, We, BiasField, endogenous time τ , temporal coupling κ , merge threshold δ , decay γ , and self-reference period G . These are implementation details rather than prerequisites for reading the evidence. A Geruon is the single-cavity detector. A Self couples multiple Geruons. A BiasField is the shared medium through which their centroids affect one another. A We records disagreement across Selves. The Codex stores the surviving centroids. Architecture details are in S1; dynamical details are in S2.

Two methodological rules govern all experiments. First, encoding is part of measurement: every encoding preserves some structure and suppresses other structure. Second, positive signals must survive controls: fair-coin baselines should be flat, shuffled streams should collapse order-dependent effects, and ablations should identify which component carries the signal. The experiments below use the same core instrument with pre-declared domain encodings and controlled parameter families.

2. Evidence

Three experimental pillars — WTC dual-encoding, DNA fork columns, and RNA blind boundary detection — establish the architecture's core claims. UN diplomatic voting provides a civilizational-scale forward prediction. All domains share no physical dimensions and are processed by the same core instrument, with pre-declared domain-specific encodings and ablated parameter families (see Reproducibility for full parameter table).

Evidence map. The four domains are not independent discoveries. They are a single instrument undergoing cross-domain pressure testing. Each domain is blind to a different class of prior knowledge.

Domain	Question	Blind to what	Positive result	Control	Limitation
UN (§2.1)	Can civilizational collapse be anticipated without causal modeling?	Event causal narratives; lens tested by ablation	2025 disp=0.448, rank 1/79	P5 ablation; FRED negative control	Prospective (2026-2028); n=79 years
WTC (§2.2)	Can externalized cognition converge without musical priors?	Chord templates, signatures, voice labels	Transposition equivariance 100%; 5-stage Codex selection loop	Cross-key Codex: non-matching entries freeze	Codex operation scaffolded (§3.5)
DNA (§2.3)	Can evolutionary constraint be read without the genetic code?	Codon table, gene annotation, conservation scores	Exon/intron d=-0.97 (d = mean cross-harm_intron - mean cross-harm_exon; n=200)	Shuffle collapses to d=-0.10	AHSG fork column: single-gene (n=1)
RNA (§2.4)	Can translational boundaries be located without codon knowledge?	AUG motif, ORF annotation, Kozak context	CDS stop ≤3 windows 98-100%	Internal ATG control 98%	CDS start 66%; passive detection only

2.1 UN Diplomatic Voting: Structural Collapse Detection

Setup. Harvard Dataverse Idealpointestimates, 1946-2025, 193 countries, annual ideal points. Encoded as annual global ideal-point distribution features. P5 positions and economic indicators are treated as ablation lenses, not required priors. Geruon uses cap=64; a κ sweep ($\kappa \in \{0.5, 1, 2, 3, 5, 7, 10, 15, 20\}$) produced identical displacement

statistics at this granularity ($\text{cap}=64 \gg N=79$), so the signal is reported as kappa-invariant. Self-referent precursor detection — each time point uses only past data to judge present state.

Calibration. Three granularities were tested before adoption. Per-vote (~400K events) and per-resolution (2,415 events) proved either computationally infeasible or structurally uninformative ($\text{cap} > N$). Quarterly aggregation confirmed UN voting operates on an annual rhythm — quarterly slices merely amplify the Q4 signal. Annual granularity (79 data points) was adopted after bootstrap validation ($\text{CV}=33.6\%$ at quarterly level, converging to $r=0.80$ when aggregated to annual ideal-point readings).

Pipeline separation. The raw voting-stream calibration and the annual ideal-point prediction are distinct pipelines. The former provides retrospective event-pattern checks on per-resolution decision traces. The latter supplies the 2025 forward signal using annual country-ideal-point distributions. The two should not be conflated. In the raw voting-stream calibration, selected precursor patterns matched 3 of 4 selected system-level events (Soviet collapse 1991, financial crisis 2008, Ukraine war 2022; Crimea 2014 did not trigger). The annual ideal-point pipeline is used separately for the 2025 forward signal and does not inherit the 3/4 claim.

Current signal. The 2025 annual ideal-point displacement is 0.448, the highest substantive reading in the 1946-2025 sequence (1946 itself is an initialization artifact and excluded). Previous substantive peaks: 2022 Ukraine war (0.358), 2023 (0.249), 1974 oil crisis+Watergate+Vietnam (0.218), 1991 Soviet collapse (0.192). τ has remained in a narrow high-tension band (0.744-0.754) across 2023-2025, with 2023 registering zero $d\tau/dt$ under the script's freeze criterion.

Economic negative control. The same Geruon pipeline was run on pure economic data (FRED: NASDAQ, T10Y, T2Y, FedFunds, CPI, Unemp, 1993-2025). Hit rate: 1/6 — only the 2008 financial crisis was detected, and as a same-year detection, not a precursor. In 2025, FRED displacement = 0.142, ranked 24th of 33 years — below the median. Pure economic data judges 2025 as structurally normal. The 2025 structural pressure is entirely in the diplomatic-political domain.

NASDAQ volume encoding. When economic data is encoded as action rather than expression — using volume log, buy/sell direction, and conviction — τ breathes healthily (0.73-0.75). The 2024-2025 period shows the most sustained high-action period in 33 years (24 consecutive months). But no diplomatic collapse precursor is detected — because the causes of diplomatic disruption lie outside market data. Economic data alone does not reproduce the diplomatic-alignment signal. The market is working. The diplomatic order is not. Prosperity and collapse are not mutually exclusive — they occur at different structural layers. Kappa selection is scoring-sensitive and is not the central result; the stable finding is the between-domain contrast.

P5 ablation. A systematic ablation tested whether the P5 augmentation creates the signal or merely amplifies an existing global signal. Six encodings: No-P5 (global ideal-point distribution only), P5-augmented (current), P5-only (P5 countries exclusively), No-USA (P5 minus United States), No-Russia, and Random-5 (arbitrary five-country groups, 5 trials). Results:

Encoding	2025 disp	Rank/78	Interpretation
No-P5	0.101	12	Elevated but not extreme — global stress exists but is diffuse
P5-augmented	0.445	1	Current claim — P5 institutional lens
P5-only	0.448	1	Nearly identical to P5-aug — great-power layer dominates the signal
No-USA	0.550	1	Stronger without the US — multi-polar fracture exceeds unilateralism
No-Russia	0.290	2	Still extreme
Random-5 (mean)	0.288	~9	Elevated but not extreme — P5 is not arbitrary

The signal is concentrated in the P5 layer. Global-level stress (No-P5, rank 12) is elevated but not historically extreme. The P5 is not an arbitrary prior — it is the institutional backbone of the UN Security Council and the post-war order. The instrument is not being handed the answer. It is locating the structural layer where the pressure lives. The No-USA result (disp=0.550) is the most precise finding: the United States' disengagement from its allies produces greater alignment fracture among the remaining powers than direct US participation.

Forward prediction. 2026-2028: structural collapse of the US-led post-war international order, concentrated in the P5 great-power alignment layer. Not the collapse of the United States as a nation, but the collapse of the gravitational alignment structure in which the United States stands at the center and other P5 powers align around it. Magnitude exceeds the 2022 Ukraine war, comparable to the 1991 collapse of the Soviet-led ideological alignment system.

Verification criteria. Three indicators, any two satisfied = confirmation. (1) disp does not return below 0.20 by 2028 (after the Soviet collapse, disp fell from 0.19 to 0.05 within four years). (2) P5 voting agreement rate falls below the historical minimum (~12%) in any year 2026-2028. (3) Ideal point distribution transitions from uni-/bi-modal to multi-modal with no dominant peak for two consecutive years.

Boundary statement. The model measures structure, not causality. If this prediction is wrong, it means that for the first time in 79 years, displacement at 0.448 occurred without a subsequent structural collapse.

2.2 WTC Dual-Encoding: Codex Convergence and Tonal Discovery

Fitch (2010) observed that speech and music leave no fossils. Sound waves do not petrify. A bone can survive a million years in sediment; a dominant seventh chord cannot survive a second past its own decay. The only way music outlasts its own sound is through

externalization — notation, recording, repetition across generations of performers. This experiment asks: if an instrument with no knowledge of music theory processes a piece of Bach, and the centroids it converges to are passed to a second instrument that never heard the original, does the second instrument find the same centroids? And the third generation after that? The Codex is not a recording. It is the structural fossil — the attractor that survived the frame economy’s compression. What Fitch said could not fossilize, the Codex preserves.

Setup. Bach’s Well-Tempered Clavier, Book I. Dual Self architecture: one Self reads chroma (12-dim active pitch class vectors), one Self reads inter-onset intervals (12-dim sliding window histograms, log-binned). Shared BiasField. Harm-Geruo ($\kappa=10$, $\text{cap}=12$) receives harm arrows, consolidates, precipitates into a shared Codex. Ten generations per piece.

Zero prior knowledge. No chord templates. No key signatures. No frequency pre-classification. No voice separation. The architecture does not know that Bach existed. It only knows how to observe, merge, predict, and mark harm.

Architecture status. Formation is endogenous — stress-driven boundary events in the Self layer trigger Codex precipitation. The current experiment operates with native formation and Codex selection; active Codex query under noise perturbation remains scaffolded and is reported as a limitation (§3.5; S7).

Phase 1: Tonal skeleton and transposition equivariance. Thirty-six pieces (Book I: 24, Book II: 12) were processed through the dual-Self architecture across ten generations each. Book II produced 39% more cross-Self harm than Book I — a systematic difference consistent with two centuries of musicological analysis identifying Book II’s more adventurous harmonic language. The harm frequency bins themselves mapped to tonal function: C major concentrated harm in bins 5 and 7 (I-IV skeleton), C-sharp minor in bin 6 (the characteristic tone of the minor mode). No single key dominated the Codex — Bach’s complete coverage of all twenty-four keys was faithfully reflected.

Transposition equivariance was tested by shifting the C major prelude by k semitones ($k \in \{1, 2, 3, 5, 7\}$) and running ten generations of Codex convergence on each transposition. The Codex anchor dimensions shifted by exactly k semitones for all five non-zero transpositions: the I-V-ii skeleton moved as a complete tonal function. This is not statistical coincidence — it is the geometry of tonal relations preserved under the chroma encoding’s cyclic group structure. The architecture did not memorize “C major looks like this.” It internalized the relative distances between harmonic attractors.

Phase 2: Five-stage externalization loop. The Codex selection experiment tests whether inherited Codex entries are confirmed, rejected, or decayed across generations. A cross-key Codex (entries from C, Eb, F#, B domains) was injected into a C-major environment and tracked across five generations. C-domain entries gained weight at +1 per generation. B-domain entries froze — never selected, never decayed. Coverage merge allowed partial entries to be supplemented by more complete evidence across generations. Three-layer selection: complete confirmation (matching evidence, weight growth), partial supplementation (incomplete evidence, coverage expansion), complete rejection (non-

matching evidence, weight freeze). The Codex selection layer runs parallel to the frame economy — it does not improve processing. It accumulates a history. This is the core discovery of the five-stage loop: externalization creates a selection layer decoupled from adaptation.

Results — C major (BWV 846). Ten generations converge the Codex to 38 entries. Two stable patterns dominate: a tonic anchor at chroma dimension 0 (>45% of entries), and a tonic-subdominant complex at dimensions 0+5 (>55%). BiasField top dimensions: d0 (28M), d5 (5.3M), d11 (4.2M), where M denotes cumulative BiasField deposit weight in arbitrary units.

Results — Eb major (BWV 852). Ten generations converge to 48 entries. Four stable patterns emerge: a tonic triad skeleton spanning dimensions 2, 4, 6, 9; a dominant-tonic complex; a pure dominant anchor at dimension 9; a pure mediant anchor at dimension 4. BiasField top dimensions: d9 (1.3M), d4 (1.3M), d2 (1.2M) (M: cumulative deposit weight).

Cross-piece comparison. The two Codexes share zero dominant dimensions. C major anchors to dimension 0. Eb major anchors to dimensions 9, 4, and 2. Under chroma encoding convention (dimension 0 = C), these correspond to the tonic notes of the respective keys. The architecture did not “discover that the piece is in C major.” It discovered that dimension 0 is the statistical gravitational center of the frame economy — and under the chroma encoding, that center coincides with what music theory names the tonic.

Why this is not statistical frequency. Raw harm arrow distributions show dimension 0 (0.312) and dimension 4 (0.300) nearly equal in C major — a 1.2% difference. Only 37% of harm arrows exceed 50% in any single dimension. A purely statistical Codex would not collapse in this configuration. The convergence is driven by three mechanisms jointly: (1) cross-Self harm filtering — only dimensions where chroma and IOI Selves both register structural divergence enter the harm-Geruo; (2) consolidation selection — only frames that survive induction_clean enter the precipitation pool; (3) Codex inheritance — the next generation’s lookup reinforces previously precipitated patterns, creating a positive feedback loop.

Next step — noise-driven query. Gen1 processes clean stream, precipitates Codex. Gen2 processes the same stream with injected noise. The cavity, destabilized by noise, queries the Codex. The hypothesis: Codex-enabled Gen2 converges closer to Gen1’s clean centroids than Codex-disabled Gen2. This directly tests externalization’s adaptive function — the Codex as a noise stabilizer.

2.3 DNA: The Archive Boundary

Fitch observed that speech and music leave no fossils. DNA leaves the oldest fossil on Earth — four billion years of continuous externalization, compressed into four letters. This experiment asks whether the Self architecture can read the structural signature of evolutionary constraint without knowing the genetic code.

Setup — Exon/Intron assay. Human protein-coding transcripts from the GENCODE and TE atlases. 3-cavity Self ($\kappa=0.5/10/100$), 3-mer encoding ($D=64$), 256-nt windows in genomic order. Exon and intron regions annotated post-run from CDS masks — the instrument never receives these labels. $N=200$ transcripts, 10,606 windows.

Results. Exon cross-harm (0.020) is $2.3\times$ lower than intron cross-harm (0.045). $d = -0.97$ (where $d = \text{mean cross-harm}_{\text{intron}} - \text{mean cross-harm}_{\text{exon}}$). Shuffling genomic order collapses the signal to $d=-0.10$. $\Delta d=+0.87$.

Setup — Fork column assay. Four-species hominid alignments (Pongo, Gorilla, Pan, Homo). Each alignment column encoded as a 16-dimensional vector — the four fork copies concatenated. 3-cavity Self processes columns in genomic order with no external annotation. Conservation is measured post-run by counting how many of the four species agree at each column.

Results. AHSG (alpha-2-HS-glycoprotein) — identified under positive selection in multiple genome-wide scans (Sabeti et al., 2007; Nielsen et al., 2005) — was tested as a single-gene calibration. Low-conservation columns (1-2/4 species agree) produce cross-harm 1.5 standard deviations higher than high-conservation columns (4/4 agree). $d = +1.51$ (where $d = \text{mean cross-harm}_{\text{low-conservation}} - \text{mean cross-harm}_{\text{high-conservation}}$). Shuffling column order collapses the signal to $d=-0.20$. $\Delta d=+1.71$. This is a single-gene result ($n=1$). The per-gene pooled approach did not generalize across a random sample of transcripts (mean $d=0.13\pm 0.22$), indicating that the fork column signal depends on sufficient within-gene conservation variance. AHSG — the one gene in the sample with documented positive selection (Sabeti et al., 2007; Nielsen et al., 2005) — carried that variance.

Significance. The exon/intron assay ($n=200$ transcripts, 10,606 windows, $d=-0.97$) provides population-level evidence that evolutionary constraint leaves a detectable structural signature in DNA. The fork column assay is reported as a single-gene calibration, not as population evidence. It demonstrates a principle: the four copies of the alignment ARE the evolutionary record, and the Self's three time lenses can read conservation directly from the fork structure without external annotation. But the principle currently stands on one gene known to carry strong evolutionary signal. Generalization requires a larger sample of genes with independently documented selection pressure. Both signals depend on genomic order and are eliminated by shuffling. DNA is externalization at its physical limit — information surviving not across generations of readers, but across millions of years. The instrument reads its structure.

2.4 RNA: The Operation Boundary

DNA provides evidence that structural code survives evolutionary time. RNA asks the next question: can the instrument detect where the code is being operated — where translation begins and ends — without knowing what a codon is?

Setup — CDS-UTR structural separation. Human protein-coding transcripts from the TE atlas. 3-cavity Self ($\kappa=0.5/10/100$), 3-mer encoding ($D=64$), 256-nt windows in genomic order across the CDS boundary. Windows classified post-run as CDS ($\geq \text{WINDOW}/6$ coding positions) or UTR (zero coding positions). $N=100$ transcripts.

Results. CDS cross-harm (mean 0.0184) is lower than UTR cross-harm (mean 0.0302). $d=-0.84$. The coding region produces more regular 3-mer organization — lower cross-cavity divergence because all three time lenses converge on the same periodic structure. The UTR produces higher cross-harm because its 3-mer distribution, lacking translational constraint, diverges across κ . Shuffling the genomic order collapses the signal. The instrument reads translational constraint in sequence structure alone — no RPF data, no codon table, no ORF annotation.

Setup — Blind CDS boundary detection. The key mechanistic test: can temporal-lens divergence locate the CDS boundary itself, without being told where it is? 2-cavity or 3-cavity Self processes windows in genomic order across the UTR→CDS transition. At the boundary, the fast lens ($\kappa=0.01$) adapts to CDS 3-mer structure faster than the slow lens ($\kappa=500-5000$) — producing a transient peak in cross-harm. The peak position is compared against the annotated CDS start and stop positions from the CDS mask. Three parameter configurations were swept: window sizes 32-64 nt, κ spreads from 0.005/10/5000 to 0.01/500, cavity counts 2-3. Multi-seed validation (3 seeds \times 120 transcripts). Results are reported at three error thresholds ($|\text{error}| \leq 2, \leq 3, \leq 5$ windows) to eliminate ambiguity.

Results — CDS stop. The CDS→3'UTR boundary is detected with near-perfect precision. Across all parameter configurations, all seeds, the cross-harm peak falls at median -1.0 windows from the true stop codon. $|\text{error}| \leq 3$ windows: 98-100%. $|\text{error}| \leq 2$ windows: 96%. This is parameter-invariant — the stop boundary is so structurally sharp that every κ spread and every window size converges to the same answer. The CDS stop codon is a structural cliff.

Results — CDS start. The 5'UTR→CDS boundary is detected reliably but with a systematic offset. Best configuration (2-cavity, $W=32$ nt, $\kappa=0.01/500$): median -0.9 windows (-7 nt), $|\text{error}| \leq 3$ windows: 66%, $|\text{error}| \leq 5$ windows: 75%. 2-cavity outperforms 3-cavity for position accuracy — the simpler fast/slow lens pair produces a cleaner boundary signal than the three-lens interaction. The peak is systematically ~ 1 window before the annotated CDS start, suggesting that the fast lens picks up CDS-like structural features in the late 5'UTR.

The asymmetry is a biological prediction. CDS stop is more precisely detected than CDS start across all parameter configurations. If this were a measurement artifact — window size, κ tuning, encoding bias — start and stop would show similar precision. They do not. The instrument reads what biology built: the stop codon is a sharper structural boundary than the start codon. The CDS→3'UTR transition is a cliff — strong 3-nt periodic structure collapsing to no periodic structure. The 5'UTR→CDS transition is a ramp — variable upstream sequence structure (Kozak context, upstream ORFs, IRES elements) gradually giving way to CDS periodicity.

This asymmetry has a biological corollary. Translation machinery faces the same structural ambiguity: locating the correct start codon within a structurally gradual transition zone. Leaky scanning — where the preinitiation complex bypasses the first AUG and initiates at a downstream site — is a well-characterized phenomenon. Misinitiation at upstream or downstream AUG codons is a known pathogenic mechanism

in cancers (where upstream ORFs in tumor suppressor genes suppress protective translation) and genetic disorders (where start-codon mutations force initiation at alternative sites). The instrument's difficulty in precisely locating the CDS start is not a failure of measurement. It recapitulates a genuine biological difficulty: stop codons are structurally unambiguous; start codons are not. The boundary the instrument reads is the same boundary the ribosome must find.

AUG control — transition-magnitude detection, not motif detection. Does the instrument detect AUG codons, or does it detect structural transitions? To distinguish these hypotheses, cross-harm peaks were compared at annotated start AUG versus internal ATG codons (≥ 200 nt into CDS). At each position, the same scanning procedure was applied. With 2-cavity, the annotated start peak exceeds the internal ATG peak in 78% of transcripts (mean $\Delta h = +0.014$) — above chance but not decisive. With 3-cavity and extreme κ spread ($\kappa = 0.005/10/5000$), discrimination reaches 98% (mean $\Delta h = +0.085$). The extreme κ configuration is sensitive to transition magnitude: the 5'UTR \rightarrow CDS jump is the largest structural transition in the entire transcript, larger than any internal ATG site (which sits within already-periodic CDS structure). The instrument does not identify AUG. It does not identify motifs. It detects the structural transition at which molecular operation enters and exits the coding regime. This is transition-magnitude detection, not motif detection.

Significance. The same temporal-lens divergence mechanism that detects boundaries in music (WTC phase transitions), diplomacy (UN system collapse), and cardiac signals (ECG PVC detection) also detects the translational boundary in mRNA — without knowing what a codon is, without searching for AUG, without any genetic code knowledge. The instrument does not identify codons or annotate ORFs. It detects the structural transition at which molecular operation enters and exits the coding regime. RNA provides the second link in the biological evidence chain: DNA provides evidence that archive survives; RNA provides evidence that operation boundary is structurally readable; WTC provides evidence that externalized cognition accumulates through history.

3. Implications

3.1 Cognition Is Centroid Convergence

The scientific implication is modest. The architecture does not simulate a mind, a society, a genome, or a musician. It performs one operation family under finite memory pressure and then asks whether the surviving centroids are informative. The fact that those centroids, when used as probes, reveal latent structure and, when externalized, accumulate across generations suggests that cognition may not need to begin with symbolic world-modeling. It may begin with compression, survival, and re-entry of structure.

Much perceptual research in music cognition has focused on dissonance — which chord progressions sound wrong, which intervals clash, which listener hears what in a Deutsch

octave illusion (Deutsch, 1974). The present experiment asks a different question. Individual perceptual variation is not treated as the primary signal; it is treated as variation across readers. The cross-Self harm in WTC does not depend on which encoding one privileges — chroma or IOI, pitch or rhythm, ear or brain. What matters is whether the structural divergence between two readings of the same stream converges to the same centroids across generations. Deutsch’s illusions measure differences between listeners. The architecture measures what survives across readings. That is the difference between studying perception and studying externalized structural cognition.

3.2 A Complementary Layer for Agent Memory

Base large language models are powerful recognition and generation engines. Contemporary agent architectures extend them with memory streams, reflection, retrieval, and explicit context management (Park et al., 2023; Packer et al., 2023). This field shows that persistent memory is now a central engineering problem, not something solved by increasing context length alone.

The centroid detector enters at a different layer. It does not generate language, retrieve documents, or optimize a task policy. It asks which structures survive finite-memory compression before any linguistic reflection occurs. Endogenous time, structural identity, externalized memory, and cross-generational filtering are explicit primitives rather than side effects of scale. The architecture is therefore not a competitor to LLM agents. It is a minimal testbed for the selection problem every long-memory agent eventually faces: what deserves to be written outside the current context, and how should later generations confirm, supplement, freeze, or reject it?

3.3 The Centroid as a Candidate Unit of Cultural Transmission

Dawkins proposed the meme as a unit of cultural transmission — replicating, mutating, selecting. The present architecture does not claim to replace that concept. It offers a lower-level operational candidate: a precipitated centroid. A precipitated frame is a centroid that survived the induction cycle. It is transmitted to the next generation through the Codex. It biases the perception of the generation that inherits it. If reinforced, its weight is maintained. If not, it decays by γ . Time does the selection.

Luhmann proposed that society is a network of operationally closed communication systems. The architecture is that network, running. Each Self is an operationally closed frame economy. Cross-Self harm is structural coupling. The Codex is the externalized residue of communication that outlasts any single system.

3.4 The Encoding Boundary and the Expression–Action Spectrum

Not all data is equally readable by the instrument. The expression–action spectrum (S4) provides a principled framework for understanding where the instrument works and where it does not. But the framework is more than an empirical boundary for this particular instrument. It points to a structural property of information itself.

Data that records decisions — votes, trades, choices — contains relational structure between the entities that made those decisions. The frame economy’s core operations,

merge and prune, converge centroids by detecting which entities act similarly over time. A country voting Yes on a resolution. A note-on event in a musical stream. A buyer committing capital in a specific direction. These are actions performed by agents in relation to other agents. The relational structure is native to the data.

Data that records measurements — prices, CPI, temperature readings — contains no such structure. It describes what happened, not who decided. The instrument's κ does not differentiate on pure expression data. The frame economy does not breathe. This is not a limitation of the instrument. It is a discovery about what kinds of information are structurally readable without prior knowledge: information that encodes relational action is readable; information that only encodes traces of processes is not.

The empirical evidence is systematic. ECG required six encodings to find one that produced signal — because raw voltage is a trace of cardiac depolarization, not the depolarization itself. Sleep required 5-second epochs rather than 30-second clinical windows — because brain dynamics occur at the timescale of the process, not the timescale of the clinical label. The FRED negative control in §2.1 is the cleanest demonstration: pure economic data (prices, CPI, employment) produces no structural precursor signal, even when run through the identical pipeline that detects collapse precursors in diplomatic voting. Not because the economy lacks structure. Because economic data measures the economy's outputs, not its decisions.

This has implications far beyond the present work. Much of what passes for data in machine learning is expression — images, text corpora, sensor readings. The centroid detector suggests that a different class of data — action traces, decision records, entity-relation sequences — may be necessary for systems that claim to understand structure rather than merely model distributions. The expression–action spectrum is not a boundary to be overcome. It is a map of where structure lives.

3.5 Limitation: The Action Gap

The architecture described in this paper is a structural detector. It reads patterns in streams. It converges to centroids. It precipitates what survives. It inherits across generations. What it does not do is act.

In living systems, cognition and action are simultaneous. A ribosome does not first “recognize” the start codon and then “decide” to initiate translation — the recognition IS the action. A musician does not first analyze the harmonic structure and then choose the next note — the analysis and the choice are the same event. The architecture models the perceptual side of this loop. It detects where structure changes, where boundaries occur, where centroids converge. But it does not close the loop from detection back to action — from “this is a boundary” to “do something about it.”

The Codex inheritance experiments (§2.2) demonstrate that externalized memory survives generational transmission and that the selection layer decouples from the processing layer. This confirms Codex formation. What remains unaddressed is Codex operation: the active loop where inherited memory directly shapes ongoing processing under boundary conditions. In the current architecture, Codex entries are selected or frozen post-hoc across generations — a passive filter. In the full circuit, a Self

encountering noise or novelty would query the Codex during processing, and the inherited centroid would enter the frame economy as a correction signal — not after the fact, but in the moment. This loop currently exists only in scaffold form (S7).

The action gap is not a bug. It is an architectural boundary — the line between what has been built and what remains to be built. The present work demonstrates that a blind instrument can detect structure, precipitate it, and pass it across generations. The next step is to close the loop: to build the mechanism by which inherited structure shapes ongoing perception, and perception feeds back into what gets stored. When a Self queries the Codex not because a generation ended, but because the world stopped making sense — that is the moment cognition and action become one event. The architecture does not yet reach that moment. This paper documents the distance traveled toward it. The remaining distance is the subject of ongoing work.

4. Conclusion

This paper has made one claim, unfolded in three parts, supported by three experimental pillars.

The architecture is a centroid detector. Two operations — merge, prune — converge any vector-encoded stream to its statistical attractors. The attractors are centroids. The detector requires no training, no labels, no domain knowledge. This is Creation.

Centroids, combined through a shared field and measured across encodings, reveal latent structure. The probe is interchangeable because the centroid is interchangeable. The measurement is the structural divergence between centroids. This is We Create.

Centroids precipitate into a Codex. They survive the erasure cycle. The next generation inherits them. The demonstrated operation is selection: later generations confirm, supplement, freeze, or ignore what earlier generations externalized. The centroids that survive across generations acquire operational value. Not chosen in advance. Not labeled by the experimenter. Survived. This is Create Us.

Three experimental pillars — WTC dual-encoding, DNA fork columns, and RNA blind boundary detection — verify the claim, supported by a forward prediction on UN diplomatic voting.

WTC demonstrates Codex formation and selection: the tonal skeleton is discovered without musicological priors, transposition equivariance holds at 100%, and a five-stage selection loop — Formation, Inscription, Transmission, Confirmation, Rejection — confirms that inherited Codex entries matching the environment gain weight across generations while non-matching entries freeze. Codex operation — the active loop where inherited memory shapes ongoing processing — remains scaffolded (§3.5). The Codex selection layer is decoupled from the frame economy's processing. Externalization does not improve adaptation. It accumulates a history that runs parallel to it. DNA demonstrates the archive boundary: exon/intron structural separation reaches $d=-0.97$ ($n=200$ transcripts, 10,606 windows); a single-gene fork column calibration on AHSG —

identified under positive selection in multiple genome-wide scans (Sabeti et al., 2007; Nielsen et al., 2005) — produces $d=+1.51$. Both signals collapse under shuffling. RNA demonstrates the operation boundary: temporal-lens divergence locates the CDS stop codon at 98-100% within ± 3 windows (parameter-invariant) and the CDS start at 66% within ± 3 windows (75% within ± 5). The instrument does not identify AUG. It detects the structural transition at which molecular operation enters and exits the coding regime — transition-magnitude detection, not motif detection.

The three biological domains form a consistent evidence chain. DNA provides evidence that structural code survives evolutionary time. RNA provides evidence that the operation boundary is structurally readable by a blind instrument. WTC provides evidence that externalized cognition accumulates through history under selective pressure. Three domains, three layers of the externalization argument, one measurement principle.

UN diplomatic voting provides the civilizational-scale test: a forward prediction for 2026-2028, falsifiable and time-bounded.

The architecture demonstrates that cognition need not begin with learning a world-model; it can begin with detecting structure. It is a working primitive for a non-learning cognitive architecture — a system for detecting and precipitating centroids across time. Codex formation and selection are demonstrated. Codex operation — the active loop where inherited memory stabilizes ongoing processing under noise — is the next step. We interpret centroids, accumulated across generations, as an operational form of cultural inheritance — what we provisionally call the third dimension of evolution.

The philosophical consequence is an inversion of priority. Traditional philosophy often begins with concepts — mind, world, truth, value, meaning — and asks how they relate. The present work begins with operations: merge what is similar, prune what is not reinforced, and repeat under finite memory. In the experiments above, those operations produce centroids; coupled centroids reveal structure; externalized centroids accumulate history; surviving centroids become candidates for value. Structure is therefore not treated as a ready-made property waiting to be represented. It is treated as the invariant residue of repeated cognitive operations. From so simple an economy, endless forms of structure have been, and are being, externalized.

Supplementary Material

Detailed supplementary material is maintained as a separate document: ee_supplement_v1.7.md.

S1. Instrument architecture — Geruon, Self, We, BiasField, Codex, structural identity.

S2. Instrument dynamics — frame economy, endogenous time, operational tau, measurement orders, precipitation.

S3. Instrument calibration and controls — passive calibration, fair-coin baseline, required controls, parameter discipline, determinism.

S4. Encoding and data-type boundary — encoding as measurement, expression-action spectrum, illustrative cases (archival).

S5. Faraday/probe measurements — probe interchangeability, τ convergence, F/τ decoupling, structon resolution.

S6. UN prediction protocol — data source, pipeline separation,

P5 ablation, FRED negative control, verification criteria.

S7. Scaffold history — temporary mechanisms, removed scaffolds, remaining Codex-operation scaffold.

S8. Negative results and decision log — failed paths, demoted claims, and boundary decisions.

S9. Reproducibility checklist — code, data dependencies, required controls, required reporting.

Reproducibility and Code Availability

Core instrument. The centroid detector (Geruon), Self, We, and Codex are implemented in pure Python 3.8+ with zero external dependencies beyond the standard library. No GPU, no neural network framework, no external API calls. The core files — `code/geruon.py` (~2,500 lines), `code/geme.py` (~800 lines), `code/we_core.py` (~300 lines) — are version-locked at the commit used to produce all results in this paper.

Experiment scripts. All domain-specific experiment scripts are archived in the `experiments/` directory, organized by domain (`un/`, `wtc/`, `dna/`, `rna/`). Each directory includes a README with data sources, run order, and expected output. Scripts use fixed random seeds (42, 123, 456 for multi-seed validation). The core instrument was audited and patched (P0/P1 fixes: structural identity collision-resistance, run-order determinism, vector dimension guard, window semantics, enrich idempotence) on 2026-05-29, before any results reported here were collected.

Data dependencies. Data processing scripts require: `pandas`, `numpy`, `pyarrow` (for `parquet`); `pybigtools` (Windows `bigWig` reader, RNA only; `pip install pybigtools`). These are standard scientific Python packages, not proprietary or custom toolchains. Raw data files (TE atlas `parquet`, `bigWig` tracks, GENCODE GTF, FASTA alignments) are too large for direct submission and must be obtained from their respective sources.

Data sources. - UN voting: Harvard Dataverse `Idealpointestimates` (Voeten, 1946-2025), FRED economic indicators (1993-2025) - WTC: MIDI encodings of Bach's Well-Tempered Clavier, Books I and II (public domain, `midi_encoder.py`) - DNA: Human TE atlas (GENCODE v49), four-species hominid alignments (Pongo/Gorilla/Pan/Homo, local FASTA) - RNA: Human TE atlas, HeLa RPF `bigWig` tracks (GSE79664, RPFdb v3), GENCODE v49 primary assembly GTF

Parameter table. All experiments share core parameters unless otherwise noted.

Parameter	Value	Role
D (vec_dim)	16 (general), 64 (DNA/RNA 3-mer)	Vector space dimensionality
cap (memory_cap)	12 (WTC), 20-32 (general), 24 (DNA/RNA)	Frame economy capacity
κ_τ (kappa_tau)	0.5/10/100 (3-cavity Self), 0.01/500 (2-cavity)	Temporal coupling — time-lens spread
window (DNA/RNA)	256 nt (exon/intron, CDS-UTR), 32-48 nt (boundary)	Genomic window size
stride	64 nt (exon/intron), 8-12 nt (boundary)	Window step size
bias_weight	0.3	BiasField coupling strength
δ (delta)	0.19	Merge distance threshold
γ_1 (gamma)	0.05/step	Weight decay rate
τ_0	0.60	Baseline τ
GI	4 steps/cycle	Self-reference period

Run order. Core instrument tests → calibration baselines (fair coin, structon) → WTC Phase 1 (tonal skeleton, transposition) → DNA exon/intron → DNA fork column → RNA CDS-UTR → RNA boundary detection → RNA RPF coupling → WTC Phase 2 (five-stage Codex loop) → UN pipeline. Each domain’s results are independent of the others. Within each domain, scripts can be rerun independently.

Acknowledgments

This work began with a book picked up on the third floor of Waterstones in Oxford, 2017. Dennett’s *From Bacteria to Bach and Back*. 560 pages. £29.99. The book said consciousness emerged across the long arc from bacteria to Bach. The architecture says: that evolution is still going. The third dimension has only just begun. The philosophy was refined over seventeen years. The code was built over nineteen days in May 2026. The rest is for anyone who downloads it and points it at their own sky.

References

1. Aaronson, S. (2011). Why philosophers should care about computational complexity. In B. J. Copeland, C. J. Posy, & O. Shagrir (Eds.), *Computability: Turing, Gödel, Church, and Beyond*. MIT Press.
2. Aaronson, S. (2013). The ghost in the quantum Turing machine. arXiv:1306.0159.
3. Calvo, S. E., Pagliarini, D. J., & Mootha, V. K. (2009). Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *PNAS*, 106(18), 7507-7512.
4. Chu, Y., Yu, D., Li, Y., Huang, K., Shen, Y., Cong, L., Zhang, J., & Wang, M. (2024). A 5' UTR language model for decoding untranslated regions of mRNA and function predictions. *Nature Machine Intelligence*, 6, 449-460.
5. Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7-19.
6. Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press.
7. Dennett, D. C. (2017). *From Bacteria to Bach and Back*. W. W. Norton.
8. Deutsch, D. (1974). An auditory illusion. *Nature*, 251, 307-309.
9. Fitch, W. T. (2010). *The Evolution of Language*. Cambridge University Press.
10. Frankish, A., et al. (2019). GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Research*, 47(D1), D766-D773.
11. Kapoor, S., Stroebel, B., Siegel, Z. S., Nadgir, N., & Narayanan, A. (2024). AI agents that matter. arXiv:2407.01502.
12. Kozak, M. (1989). The scanning model for translation: an update. *Journal of Cell Biology*, 108(2), 229-241.
13. Kozak, M. (2002). Pushing the limits of the scanning mechanism for initiation of translation. *Gene*, 299(1-2), 1-34.
14. Liu, J. (2026a). GEME: A self-referential prism for cognitive modeling. Zenodo.
15. Liu, J. (2026b). BGM: Building bridge — from Bach to bacteria and forward. Zenodo.
16. Lu, C., Lu, C., Lange, R. T., Foerster, J. N., Clune, J., & Ha, D. (2024). The AI Scientist: Towards fully automated open-ended scientific discovery. arXiv:2408.06292.
17. Luhmann, N. (1984). *Soziale Systeme*. Suhrkamp. (English: *Social Systems*, Stanford, 1995.)
18. Miton, H., & DeDeo, S. (2022). The cultural transmission of tacit knowledge. *Journal of the Royal Society Interface*, 19(195), 20220238.
19. Muthukrishna, M., & Henrich, J. (2016). Innovation in the collective brain. *Philosophical Transactions of the Royal Society B*, 371(1690), 20150192.
20. Nielsen, R., et al. (2005). A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biology*, 3(6), e170.
21. Packer, C., Fang, V., Patil, S. G., Lin, K., Wooders, S., & Gonzalez, J. E. (2023). MemGPT: Towards LLMs as operating systems. arXiv:2310.08560.
22. Park, J. E., Yi, H., Kim, Y., Chang, H., & Kim, V. N. (2016). Regulation of poly(A) tail and translation during the somatic cell cycle. *Molecular Cell*, 62(3), 462-471.

23. Park, J. S., O'Brien, J. C., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). Generative agents: Interactive simulacra of human behavior. *Proceedings of UIST 2023*. ACM.
24. RPFdb v3. (2024). Ribosome profiling database. <https://sysbio.gzzoc.com/rpfdb/>.
25. Sabeti, P. C., et al. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature*, 449(7164), 913-918.
26. Sanabria, M., Hirsch, J., Joubert, P. M., & Poetsch, A. R. (2024). DNA language model GROVER learns sequence context in the human genome. *Nature Machine Intelligence*, 6, 911-923.
27. Voeten, E., Strezhnev, A., & Bailey, M. (2024). United Nations General Assembly Ideal Point Estimates. Harvard Dataverse, V33.
28. Yager, K. G. (2024). Towards a science exocortex. *Digital Discovery*, 3, 1933-1957.