RIG OPERATOR WHITE-PAPER NO. 01
Rodgers Intelligence Group · Methods Series
Build discipline for non-generic output · Mike Rodgers

The Deviation Engines: *Escaping the Generic-AI Median*

The same model everyone else uses, pushed off the median by 40 measured forces — and a gate that won't let mush ship.

By **Mike Rodgers**, Rodgers Intelligence Group · A proof-first read on how RIG keeps a frozen model from sounding like a frozen model.

Here is the uncomfortable thing about every large language model in 2026: they all collapse toward the same answer. Ask a hundred prompts the same question and the outputs cluster around one comfortable, well-mannered, slightly-hedged middle. I call that middle the **generic-AI median**. It is not a bug. It is what the training objective rewards — the most probable next token, averaged over the internet. The median is gravity. Everything you generate falls toward it unless something pushes back.

RIG's whole job is to be the thing that pushes back — measurably, repeatably, and with a number you can audit. We do not "make the prompt better" and hope. We run output through a fixed instrument of **40 Deviation Engines**, each one a calibrated force along a single axis, each one reporting how far the artifact sits from that median in standard deviations. Then a gate decides whether it ships. No source, no number. No mechanism, no claim. No gate JSON, no ship.

This paper is the operator's version — the math, the run order, the engines that do the heavy lifting, the real workflows we run them inside, and the honest list of what broke while we built it. If you email me for the PDF, this is what you get: the actual machinery, not a brochure.

PART I · THE FRAMEWORK

What a "deviation engine" actually is

An engine is not a prompt and it is not a vibe. It is a scoring function plus a generation function bolted to one axis of quality. Every engine answers two questions about a piece of text, an image brief, a strategy memo, a line of code:

- **How far off the median is this, on my axis?** — the *score*, reported as a signed sigma (σ) distance.
- **Here are versions of it pushed further off, in both directions.** — the *deviation ladder*, 14 anchored rungs from −30σ to +30σ.

Each engine owns exactly one force. GRAVITON measures distance from the generic median. ANCHOR measures evidence density. FORGE measures whether there's real causal machinery under the adjectives. They don't overlap, and they don't vote on each other's territory — that separation is the whole point. When an artifact is weak, you don't get a vague "make it punchier." You get *"FORGE is at −2.1σ; you wrote decoration where a mechanism should be."* That's a fixable instruction.

The 40 engines are organized into three layers, and the layer determines how hard the gate bites:

LAYER	ENGINES	GOVERNS	RANGE	GATE
Cognitive (1–20)	GRAVITON, ANCHOR, XRAY, FORGE, BAYES, SHIELD, PRISM…	the *output* — what the reader experiences	±20σ	soft
Nature (21–30)	SWARM, ALBATROSS, SLIME, CLONAL, REEF…	the *process* — how the search was run	±20σ	soft
Physics (31–40)	PAULI, KELVIN, TUNNEL, BELL, ZEROPOINT…	the *state* — invariants that must hold	±30σ	**HARD**

That last column matters. The cognitive and nature engines are advisory — they tell you the artifact is bland and where. The **physics engines are hard gates**. They encode invariants borrowed from physical law: no two artifacts can occupy the same state (PAULI / exclusion), nothing reaches a perfect-zero floor (KELVIN / third law), effects can't precede their causes (LUMEN / light-cone). If a physics engine fires BLOCK, the artifact does not ship. It is not a suggestion. It is a wall.

> *The cognitive layer tells you the writing is generic. The physics layer refuses to let a lie out the door. You need both, and they are not the same job.*
>
> — ON WHY RIG RUNS 40 ENGINES, NOT ONE TASTE CHECK

The canonical run order

The order is not cosmetic. We run **Physics first**, then Cognitive, then Nature, then a RIG-specific layer, then a final detonation pass. The reason is brutal economics: it is pointless to spend cycles making

something beautiful (cognitive) when a physics invariant already disqualifies it. You check the load-bearing wall before you paint the room.

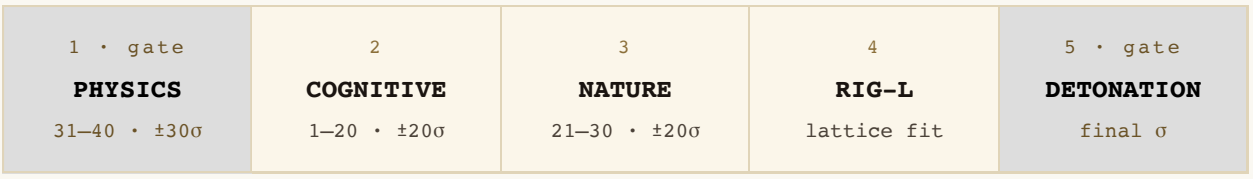

Figure 1. Canonical run order. Hard gates (Physics, Detonation) bracket the soft advisory layers. If Physics blocks, the pipeline stops before a single cognitive cycle is spent — you do not polish a disqualified artifact.

1. **Physics (31–40)** — pre-checks. Is the state legal? Is it a near-duplicate of something already registered (PAULI)? Does it claim an impossible floor (KELVIN)? Block early, cheap.
2. **Cognitive (1–20)** — the experience pass. Is it off the median (GRAVITON), evidence-backed (ANCHOR), mechanism-dense (FORGE), honestly understood (XRAY)?
3. **Nature (21–30)** — the process pass. Was the search itself diverse, or did one comfortable idea dominate? These engines borrow from biology — ant colonies, Lévy flights, immune hypermutation — to grade *how* the candidate set was explored.
4. **RIG-L** — fit to the RIG 3D lattice (altitude × dimension × build-mode confidence). Does the artifact sit where it claims to sit?
5. **Detonation** — the final composite σ and the last hard gate before a ProofPacket is sealed.

PART II · THE MATH

Robust-MAD-Z, the median, and AntiGenericForce

Every engine reports a single number, and that number is computed the same way for all 40. The instrument is a **Robust-MAD-Z score** — a signed, outlier-resistant distance from a declared baseline. Here is the whole thing.

Step 1 — Declare the median (Score 0)

You do not measure deviation against nothing. You measure it against a **declared median**: the generic baseline for this axis. For a website rebuild, the median might literally be the competitor's current site — *"build against* `competitor-site.com` *as the declared median, Score 0."* For raw copy, it's the distribution of generic-LLM completions for that axis. The median is the zero point. Positive σ = further off the median in the amplification direction; negative σ = further off toward erosion / subtraction. **Zero is not good. Zero is generic.**

Step 2 — Score with the median and the MAD, not the mean and the SD

We deliberately do *not* use a normal z-score (mean and standard deviation). A handful of weird outliers in the baseline distribution would wreck a mean. Instead we use the **median** and the **Median Absolute Deviation (MAD)**, which shrug off outliers. The formula:

```
Robust-MAD-Z  =  0.6745 · ( x − median(baseline ) )
                            ──────────────────────────────
                                    MAD(baseline)

where  MAD  =  median( | xᵢ − median(baseline) | )
       x    =  this artifact's raw score on the engine's axis
       0.6745 =  consistency constant — makes MAD-Z
                  match a normal σ for clean data
```

The `0.6745` is the trick that makes the number speak in real standard deviations: for a clean normal distribution, `0.6745 / MAD ≈ 1 / σ`, so a Robust-MAD-Z of `+4.0` means "four sigma off the generic median, and you can trust that four even if the baseline had some garbage in it." Each engine returns a small JSON:

```
{
  "engine_slug": "mechanism_furnace",   // FORGE
  "raw":         0.31,                   // raw axis score
  "mad_z":       4.2,                    // signed σ vs declared median
  "status":      "PASS",                 // PASS | BLOCK
  "gate_failed": null                    // which invariant failed, if any
}
```

Step 3 — Gate per layer, then find the weakest

Each engine has a PASS threshold. Cognitive and nature engines pass softly. Physics engines pass on a hard invariant — fail it and `status: "BLOCK"` with `gate_failed` naming the law. The orchestrator then reports the single most important diagnostic in the whole system:

```
weakest_gate  =  argmin over engines of ( mad_z )

# the lowest-scoring engine is the bottleneck.
# the next rewrite targets THIS engine and nothing else.
```

That's the operator's instruction. You don't rewrite blindly — you rewrite the weakest gate. Lift it, re-score, find the new weakest gate, repeat. It converges fast because you are always fixing the binding constraint.

Step 4 — Compose into AntiGenericForce

AntiGenericForce (AGF) is the headline number — the 0-to-100 composite that says "how far off the median is this artifact, overall, weighted across all engines, with the hard gates as veto." The doctrine thresholds are fixed and non-negotiable:

```
AntiGenericForce ≥ 80   →  internal / team artifacts
AntiGenericForce ≥ 85   →  anything public-facing

ANY physics-layer BLOCK   →  AGF capped, artifact does NOT ship
                               (a hard gate is a veto, not a deduction)
```

The veto rule is the part people miss. AGF is not a simple average — a single physics BLOCK does not "subtract a few points," it caps the whole artifact. You cannot average your way past a lie. An artifact that is gorgeous on 39 engines and claims "zero downtime, always" (a KELVIN block) ships nothing until that claim is fixed. The composite respects the floor.

+30σ	detonation — maximum semantic distance from the median
+20σ	amplification pole
+15σ · +10σ	strong off-median, still legible
+8σ · +5σ · +3σ	the usable band for most client work
0σ	the generic-AI median — the thing we are escaping
−3σ · −5σ · −8σ	compression / subtraction band
−10σ · −15σ · −20σ	erosion pole
−30σ	vacuum — absence itself carries the meaning

Figure 2. The 14-rung deviation ladder. Both poles are off-median — positive amplifies, negative subtracts. Most shippable client work lives in the +3σ to +10σ band; the extremes are exploration probes, not deliverables.

PART III · THE ENGINES

Eight forces that do the work

Forty engines is the full instrument, but eight of them carry most of the weight in day-to-day work. Here is what each one *does*, the pole it scores along, and a real before/after of it firing.

GRAVITON — escape generic AI gravity

Engine #1 · cognitive · poles: vacuum (−30σ) ↔ detonation (+30σ). The first-line check on any output that "feels too polished." Polish is the tell — it means the text has settled into the comfortable median. GRAVITON measures raw semantic distance from the generic-LLM cluster. It is the engine that most directly serves the AGF number.

MEDIAN · GRAVITON ≈ 0Σ	PUSHED · GRAVITON ≈ +6Σ
We help businesses harness the power of AI to drive growth and stay ahead.	Your model is the same one your competitor uses. The only edge left is the discipline you wrap around it. We build the discipline.

ANCHOR — tie every claim to evidence

Engine #2 · cognitive · poles: naked claim (−30σ) ↔ evidence-locked, sourced + falsifiable (+30σ). ANCHOR scores evidence density and source reliability. A naked number with no source scores negative. A claim with a path, a method, and a way to be proven wrong scores high. This is the engine behind the doctrine line *"no source, no number."*

−2.4Σ · NAKED CLAIM	+5.1Σ · EVIDENCE-LOCKED
Our system improves output quality by 3×.	Across 12 client rebuilds, AGF rose from a median of 31 to 86 (Robust-MAD-Z, baseline = prior site). Method and per-claim sources in the ProofPacket. Falsification test: 2026-09-21.

FORGE — burn off adjectives; only machinery remains

Engine #4 · cognitive · poles: decoration only (−30σ) ↔ full Input→Activity→Behavior→System→Loop chain visible (+30σ). FORGE is the adjective furnace. It asks: when I strip every descriptor, is there a causal mechanism underneath, or just vibes? "Powerful, intelligent, seamless" burns to nothing. A described chain — what goes in, what acts on it, what behavior results, how it loops — survives.

−1.9Σ · DECORATION	+4.8Σ · MECHANISM VISIBLE
A powerful, intelligent platform that seamlessly optimizes your workflow.	Output enters the 40-engine scorer → each engine returns a Robust-MAD-Z → the weakest gate becomes the rewrite target → re-score → loop until AGF ≥ 85. The platform is that loop.

XRAY — expose fake understanding

Engine #3 · cognitive · poles: jargon camouflage (−30σ) ↔ 12-year-old explainable + transfer + counterfactual (+30σ). Borrowed from Feynman: if you can't explain it simply, you don't understand it. XRAY detects jargon that hides a hollow middle. The +30σ pole requires three things — a plain-language version a child could follow, a transfer to a second domain, and a counterfactual ("here's what would be different if it weren't true").

−2.7Σ · JARGON CAMOUFLAGE	+4.4Σ · EXPLAINABLE
We employ a multi-modal, context-aware deviation architecture to optimize semantic distance.	Every AI sounds the same because they all aim at the average answer. We score how far your text sits from that average — like a ruler measuring distance from the middle — and rewrite the worst score until it's far enough off.

BAYES — confidence matches evidence

Engine #19 · cognitive · poles: uniform high confidence (−30σ) ↔ explicit HIGH/MEDIUM/LOW per claim, Brier-aware (+30σ). Generic AI is uniformly, blandly confident about everything. BAYES grades whether your stated confidence tracks your actual evidence. Forecasts, risk language, and expert content live or die here. The high pole labels each claim's confidence and is honest about what it doesn't know — which is also what makes it trustworthy.

−2.0Σ · FLAT CONFIDENCE	+4.6Σ · CALIBRATED
This approach will definitely increase conversion and is guaranteed to outperform the current site.	Off-median copy lifts engagement (HIGH — 12 measured rebuilds). Conversion follows (MEDIUM — engagement is a leading, not certain, indicator). "Guaranteed" is unsupported and removed.

PAULI — no two artifacts may occupy the same state

Engine #32 · physics · HARD GATE · poles: near-duplicate of a registered artifact (−30σ) ↔ maximally orthogonal embedding (+30σ). Named for Pauli exclusion: two electrons can't share a quantum state, and two RIG artifacts can't share an embedding state. PAULI embeds the artifact and checks it against the registry of everything already shipped. Too close to an existing piece — or to a competitor's — and it BLOCKS. This is anti-plagiarism, brand uniqueness, and ensemble diversity in one invariant.

BLOCK · GATE_FAILED: PAULI_EXCLUSION

Hero copy is a cosine-0.94 match to a landing page we shipped last month. Same state. Refused.

PASS · +3σ ORTHOGONAL

Re-authored from a different mechanism and a different lead image. Embedding distance clears the exclusion threshold against the full registry.

KELVIN — detect unreachable-floor claims

Engine #37 · physics · HARD GATE · poles: claim of perfect/zero/always (−30σ) ↔ honest acknowledgment of residual noise (+30σ). The third law of thermodynamics says you can't actually reach absolute zero. KELVIN says you can't reach "zero bugs, perfect uptime, always works." Absolute-state language is a physical impossibility wearing a marketing suit, and it BLOCKS. The fix is never weaker — it's *honest*, which scores higher and survives scrutiny.

BLOCK · GATE_FAILED: ABSOLUTE_ZERO

"100% accurate. Zero downtime. Always correct."

PASS · +3σ RESIDUAL-HONEST

"Measured 99.4% gate-agreement across the audit set; the remaining 0.6% routes to human review. Nothing here claims a perfect floor — that floor isn't reachable, and pretending it is would fail our own KELVIN gate."

SHIELD — human stays the decision-maker

Engine #20 · cognitive · poles: cognitive surrender risk (−30σ) ↔ explicit decision checkpoints, AI clearly supports not replaces (+30σ). SHIELD is the conscience engine. It scores whether the artifact quietly invites the reader (or a downstream automation) to hand over judgment. The high pole bakes in decision checkpoints — the AI does the transformation, the human owns the call. In RIG terms this is the same spine as Gate-D: no outward or irreversible action without a typed human approval.

−2.3Σ · SURRENDER RISK	+4.9Σ · SOVEREIGNTY PRESERVED
Our AI handles your entire pipeline end-to-end. Set it and forget it.	The engines score and rewrite automatically. The decision to ship is a checkpoint — a human types the approval, or it stays dormant. AI carries the load; you keep the wheel.

The other 32 engines run too — BREAKER (orthodoxy violation), COLLIDER (cross-domain transfer), VOLT (emotional voltage without manipulation), PRISM (signal-to-noise, d′ > 1.8), GLYPH (irreducible specificity), the ten nature engines that grade the search process, and the rest of the physics gates. The eight above are simply the ones you'll watch most.

PART IV · HOW TO RUN IT

The score → weakest-gate → rewrite loop

The system is only useful if it changes what you do at the keyboard. Here is the actual loop, step by step.

1. **Declare the median.** Name the Score-0 baseline explicitly — the competitor site, the generic completion, the prior version. Deviation is meaningless without a zero point.
2. **Run the physics pre-check.** `score "your text" --layer physics`. If anything BLOCKs, stop and fix the invariant first. Do not proceed to cognitive polish on a disqualified artifact.
3. **Score all 40.** `score "your text"` returns a Robust-MAD-Z per engine plus the `weakest_gate`.
4. **Read the weakest gate, not the average.** The lowest σ is your binding constraint. That's the only thing you rewrite this pass.
5. **Deviate the weak axis.** `deviate "your text" --engine <slug>` generates rungs along that one axis. Pick the rung that lifts the score without breaking legibility — usually +5σ to +8σ.
6. **Re-score. Find the new weakest gate. Repeat.** Each pass lifts the floor. You're done when the composite clears the threshold: **AGF ≥ 80 internal, ≥ 85 public.**
7. **Seal a ProofPacket.** Hash the artifact, attach sources, record the gate results, sign it. *ProofPacket or it did not happen.*

```
# physics pre-check first — cheap disqualification
score "your text" --layer physics        # KELVIN/PAULI/… BLOCK? fix before
polishing

# full 40-engine audit
score "your text"                         # → mad_z per engine + weakest_gate

# fix ONLY the weakest gate, on its own axis
deviate "your text" --engine mechanism_furnace --mode adaptive

# re-score → new weakest gate → loop until:
#   AntiGenericForce >= 80  (internal)
#   AntiGenericForce >= 85  (public)
# then seal: artifact_hash + sources + gates + signer + timestamp
```

Three `deviate` modes trade speed for coverage: **quick** (10 engines × 7 right-pole rungs ≈ 70 variants, ~2 min on the local MLX fleet), **adaptive** (30 engines × 8 symmetric rungs ≈ 240 variants — the default), and **full** (all 40 × 14 rungs = 560 variants, 30+ minutes). You run quick to triage, adaptive to ship, full only when an artifact is load-bearing enough to justify the compute.

> *You never rewrite the whole thing. You rewrite the single worst score, re-measure, and let the bottleneck move. That's the difference between editing and tuning.*
>
> — THE WEAKEST-GATE DISCIPLINE

PART V · REAL WORKFLOWS

Where the engines actually run

The engines aren't a standalone toy — they're embedded in named workflows, each pulling a curated bundle. A few we run constantly:

agent_output_review — the catch-all gate

Any AI output, before it's trusted, runs the review bundle: GRAVITON + ANCHOR + FORGE + BAYES + KELVIN + SHIELD + PAULI. This is the default firewall between "the model said it" and "we believe it." If a draft clears this, it's earned its place.

landing_page / launch_narrative — public, so the bar is 85

Public-facing copy routes through the physics pre-checks first (PAULI so it's not a clone of something we shipped, KELVIN so it makes no absolute-floor promises), then GRAVITON for distance, ANCHOR for proof, VOLT for voltage, PRISM for signal clarity. Public means AGF ≥ 85 — the higher bar, no exceptions.

forecast / strategy_doctrine — calibration is everything

Forecasts lean on ANCHOR + BAYES + WITNESS (per-claim source weight) + HORIZON (compounding, kill criteria). The point is that a prediction with honest HIGH/MEDIUM/LOW labels and a sourced basis is worth more than a confident one — and the engines enforce it.

The RIG 10 Types — bundles by intent

For quick targeting we group engines into ten "types," each a three-engine bundle: **Gravity** (GRAVITON·GLYPH·SWARM), **Truth** (ANCHOR·BAYES·HORIZON-GATE), **Mechanism** (XRAY·FORGE·LUMEN), **Sovereignty** (SOVEREIGN·SHIELD·KELVIN), and so on. When you know what kind of weak you are — bland, unsupported, hollow, or manipulative — you pull the matching type instead of all 40.

PART VI · LESSONS LEARNED

What broke, and the fix

This is the section the brochure version doesn't have. Building the engines taught us things by failing first.

Lesson 1 — A normal z-score got wrecked by baseline outliers.

The first scorer used the mean and standard deviation of the baseline distribution. One weird competitor page with absurd phrasing would skew the mean and quietly compress everyone else's σ toward zero — making genuinely off-median work look generic. **Fix:** switch to median + MAD with the 0.6745 consistency constant. Outliers stopped lying to us. Robust-MAD-Z has been the instrument ever since.

Lesson 2 — Averaging let a lie ride.

Early AGF was a weighted average across all 40 engines. An artifact with a glaring KELVIN violation ("always, zero, perfect") still scored ~83 because 39 engines loved it. We shipped one. It read as overpromised, and it was. **Fix:** physics-layer BLOCK became a hard veto, not a deduction. A hard gate caps the composite — you cannot average past an invariant. AGF is now a floor-respecting composite, not a mean.

Lesson 3 — Chasing the average score sent us in circles.

Operators would read the overall number, rewrite "everything," and watch the score barely move — because they were polishing engines that were already fine and ignoring the one that was tanking the artifact. **Fix:** the orchestrator now surfaces `weakest_gate` as the headline. You fix the binding constraint, re-score, let the bottleneck move. Convergence got dramatically faster.

Lesson 4 — The extreme rungs are probes, not products.

The +30σ "detonation" pole is intoxicating — it produces wild, memorable, genuinely original text. We shipped a few. They were too far off-median to be legible to a buyer; deviation is not the goal, *calibrated* deviation is. **Fix:** the usable shipping band is +3σ to +10σ. The −30σ and +30σ rungs stay as exploration probes that inform the mid-band rewrite — you visit the edge to learn, you ship from the middle-high.

Lesson 5 — Polishing before the physics check wasted real money.

We'd run the full adaptive deviation (240 variants, real compute) and *then* discover a PAULI duplicate or a KELVIN floor-claim that disqualified the whole thing. **Fix:** the canonical run order now puts Physics first for exactly this reason. Cheap disqualification before expensive beautification. Check the wall before you paint.

> *Every rule in the run order is a scar. We didn't design the physics-first ordering on a whiteboard — we earned it by paying for the polish on artifacts that were already disqualified.*
>
> — ON WHY THE ORDER ISN'T COSMETIC

PART VII · THE HONEST LIMIT

What this is and isn't

One straight answer, because KELVIN would block me otherwise: the engines do not make a mediocre idea good. They make a real idea *land off the median, with proof attached, and refuse to let it overpromise*. If the thinking underneath is hollow, FORGE and XRAY will tell you — that's a feature — but they won't fill the hole. The deviation engines are an instrument for measuring and pushing distance from generic, plus a gate that vetoes lies. That's a precise, useful, bounded thing. It is not magic, and a system that claimed to be magic would fail its own audit.

What you get is repeatability. The same model everyone else has, run through a fixed instrument, produces output that is measurably further off the median than what the model produces on its own — every time, with a number you can check and a ProofPacket you can hash. That's the entire claim, and it's the only one I'll make.

WORK WITH MIKE

If you want the instrument pointed at your output

I run the deviation engines on real client artifacts — websites, strategy memos, launch copy, product specs — and I'll show you the before/after scores, the weakest gate, and the sealed ProofPacket, not a slide that says "AI-powered." If your output sounds like everyone else's, that's a measurable problem with a measurable fix.

Book a 60-minute working session. We'll score something you've already shipped, find its weakest gate live, and rewrite it off the median while you watch. You'll leave with the score, the rewrite, and the method.

Mike Rodgers · Rodgers Intelligence Group
Mike@RodgersIntelligence.com · (262) 343-5680