WHITE PAPER

Behavioral Data Entropy Collapse (BDEC)

How AI Systems Are Consuming the Minds That Built Them

Prepared by Jeremy Clawson / Zeno Global
Prepared for Executive Leadership / Board of Directors
Date March 2026
Classification Public

EXECUTIVE SUMMARY

AI systems are trained on human behavior. But human behavior is no longer a stable, generative signal. Decades of standardization, automation, and algorithmic prescription have systematically compressed the variance of human decision-making. As AI prescribes behavior in real time, the feedback loop between human agency and machine learning approaches a closed state. This paper calls that process Behavioral Data Entropy Collapse (BDEC), and it argues the implications reach well beyond the technical into the civilizational.

AI systems require a signal to learn from. That signal is human behavioral data: the accumulated record of decisions made, paths chosen, words composed, and questions asked. The richness and diversity of that signal is what enables machine learning to generalize. The problem is not volume. It is variance. The distribution of human behavior is narrowing as the same technologies that harvest behavioral data simultaneously shape it into increasingly predictable forms.

The existing model collapse literature, anchored by Shumailov et al.'s landmark 2024 Nature paper, establishes what happens when AI trains recursively on AI-generated data: the tails of the distribution disappear, diversity collapses, and generative capacity narrows. BDEC points to the problem that comes before this one. If the humans generating behavioral data have already been shaped by AI prescription into a narrower, more predictable distribution, then the training data is pre-compressed before a single synthetic token ever enters the pipeline.

KEY FINDINGS

- 53% of U.S. drivers say they would not feel confident driving anywhere at all without GPS, and only 12% feel completely comfortable without it (Murphy & Prachthauser, 2025, n=2,400+). Longitudinal neuroscience links heavier GPS use to a measurable decline in hippocampal-dependent spatial memory (Dahmani & Bohbot, 2020).
- 75% of global knowledge workers now use generative AI at work, with adoption nearly doubling in the six months before the survey, a rapid shift in how written work is produced (Microsoft/LinkedIn Work Trend Index, 2024).
- About 65% of U.S. Google searches now end without a click to the open web (SparkToro / Datos, 2024); when AI Overviews are present, the zero-click rate reaches 83% (Similarweb, 2025). Discovery-driven behavioral signal is undergoing structural displacement.
- Behavioral compression is not a future risk. It is a present condition. Because the harm is a function of duration, what the paper calls the Compression Duration Effect (CDE), every month without a structural response compounds the severity and reduces the reversibility of the outcome.

BACKGROUND & CONTEXT

BDEC draws on Shannon's foundational 1948 information theory to establish a precise language for the problem. Shannon entropy measures the unpredictability, and therefore the information content, of a signal. A high-entropy behavioral population generates rich, diverse training data. A low-entropy population generates a predictable, compressed signal that, used as training data, produces AI systems that gradually lose the ability to model the full range of human possibility.

Three successive waves of technological development have compressed behavioral entropy without policy recognition of the cumulative effect. Industrial standardization normalized behavior at scale: every McDonald's identical, every airport built on the same wayfinding logic. Digital automation structured the path of least resistance, so users follow GPS instructions rather than navigate, pick from curated sets rather than evaluate, and are served rather than left to discover. Algorithmic prescription now completes the transition, as AI systems draft communications, generate analysis, and propose strategies that humans simply ratify. The human has moved from the upstream of the decision to the downstream of it. The AI generates; the human confirms.

The Three Compression Waves

Wave	Compression Mechanism
W1 Industrial Standardization	Built environments, supply chains, and organizational procedures normalize behavior at scale. Conformity becomes the path of least resistance. The decision space narrows structurally before digital technology enters the picture.
W2 Digital Automation	Search engines, GPS, recommendation algorithms, and e-commerce platforms structure the path of least resistance. Users follow instructions rather than navigate; select from curated sets rather than evaluate. Agency is preserved as feeling even as its exercise is reduced.
W3 Algorithmic Prescription	AI systems draft communications, generate analysis, and propose strategies that humans ratify. The human moves from upstream originator to downstream approval mechanism. This is not augmentation. It is substitution with extra steps.

ANALYSIS

The Evidence We Already Have

BDEC is currently untestable at precise scale because no institution has recognized behavioral entropy as something worth measuring. But proxies exist, and they are not ambiguous. The paper proposes a Behavioral Entropy Index (BEI) built from four domain-specific proxies: Navigational Autonomy Rate (NAR), Original Composition Rate (OCR), Decision Variance Score (DVS), and Discovery Independence Index (DII).

What We Know, What We Infer, What We Hypothesize

BDEC spans a causal chain whose links are not equally evidenced, and intellectual honesty requires marking where measurement ends and argument begins. The strongest evidence supports the individual-capability claims. The population-scale and civilizational claims are, respectively, a reasonable inference and a hypothesis that this paper argues is serious enough to act on before it can be precisely measured.

Epistemic Status	The Claim	What Backs It
OBSERVED	Individuals are offloading specific cognitive tasks to AI and automation, and measured capability declines in at least one domain. Adoption is widespread and still rising.	Dahmani & Bohbot (2020), spatial-memory decline; Murphy & Prachthauser (2025), GPS dependence; Microsoft/LinkedIn (2024), AI adoption; SparkToro/Datos (2024) and Similarweb (2025), zero-click search.
SUPPORTED INFERENCE	Aggregated across domains, this offloading is narrowing the variance (the entropy) of human behavior at population scale.	Strong evidence within individual domains plus a coherent information-theoretic mechanism (Shannon, 1948). No direct population-level entropy measurement yet exists: the BEI is proposed, not built, and the DVS has no instrument.
HYPOTHESIS	Pre-compressed human behavioral data degrades the training signal for frontier AI, and sustained over time (the CDE) compounds into reduced innovation, epistemic, and democratic capacity. These are the civilizational stakes.	The recursive-synthetic form of model collapse is established (Shumailov et al., 2024); the human-upstream form is an extension by analogy and remains untested. This is the claim BDEC exists to put on the research agenda.

53%
would not feel confident driving at all without GPS (Murphy & Prachthauser, 2025)

65%
of U.S. Google searches end without a click to the open web (SparkToro / Datos, 2024)

75%
of knowledge workers use GenAI at work (Microsoft/LinkedIn, 2024)

Domain	Proxy Metric	Current Reading	Trajectory	Research Basis
Navigation (NAR)	GPS Dependency Rate	53% not confident driving at all without GPS; only 12% fully comfortable without it	Degrading; neuroscience confirmed	Murphy & Prachthauser (2025); Dahmani & Bohbot (2020)
Composition (OCR)	AI Writing Usage Rate	75% of global knowledge workers use GenAI at work; adoption ~doubled in 6 months	Rapid adoption; nearly doubled in 6 months (2024)	Microsoft/LinkedIn Work Trend Index (2024)
Discovery (DII)	Zero-Click Search Rate	65% overall; 83% when AI Overviews triggered (2025)	Rising; zero-click up from ~50% (2019) to ~65% (2024)	SparkToro/Datos (2024); Similarweb (2025)
Decision Variance (DVS)	No direct measure yet	Instrument does not yet exist	Urgent research gap	BEI proposal; see full paper

SPOTLIGHT: Google Maps and the Disappearing Sense of Direction

Consider what has happened to human navigation since Google Maps became ubiquitous. Before turn-by-turn GPS, getting somewhere unfamiliar required active spatial reasoning: building a mental map, tracking landmarks, estimating distance and orientation. That cognitive effort, often frustrating, was also generative. It built and maintained the neural architecture for spatial memory.

A 2020 study by Dahmani & Bohbot in *Scientific Reports* found that habitual GPS users had worse hippocampal-dependent spatial memory than non-GPS users, and a longitudinal follow-up linked heavier GPS use to a steeper decline in that memory over time. The capacity was not just unused; it was measurably weaker. A 2025 national survey of 2,400+ U.S. drivers found that only 12% feel completely comfortable driving without GPS, and 53% would not feel confident driving anywhere at all without it.

This is BDEC made tangible. The convenience was real. The cost was invisible until the capacity was gone. The generation growing up with GPS as the default has never developed the spatial baseline that older users are losing. For them there is nothing to recover, only something that was never built. Multiply this pattern across composition, discovery, and professional judgment, and the scale of the compression becomes clear.

The Compression Duration Effect (CDE)

The most alarming dimension of BDEC is not the compression level but the duration. The CDE holds that the societal consequences of behavioral compression are not binary but cumulative: they build up over time, so the total accumulated compression matters far more than any single snapshot. The dose-response analogy is a good one. A brief period of high GPS dependency and AI writing assistance may leave cognitive infrastructure largely intact. A decade of the same compression does more than deepen those effects. It restructures the institutions, educational systems, and professional norms built during that period, encoding the compressed state as the new baseline.

Adults who experience compression have a prior high-entropy baseline, so the loss registers as loss and recovery means getting something back. A generation raised entirely within a high-compression environment has no such baseline. Compression is simply their experience of normal. For them, recovery is not a matter of getting something back but of building capacities that were never developed in the first place. This is the CDE's most severe and least reversible implication.

Domain	What Prolonged Compression Produces
Epistemic Systems	Scientific discovery depends on questioning the prevailing paradigm, which is a generative act. Sustained compression produces researchers who optimize within existing frameworks ever more efficiently while growing steadily less capable of the frame-breaking that produces breakthrough science.
Democratic Systems	Meaningful deliberation requires citizens capable of forming independent political judgments. The CDE degrades the reasoning process itself, producing an electorate that ratifies

	algorithmically curated political positions rather than constructing independent ones. This is qualitatively different from misinformation. Misinformation corrupts the inputs; the CDE degrades the process itself.
Economic Innovation	Transformative innovation emerges disproportionately from the long tail of the behavioral distribution. AI-driven efficiency gains compound quickly; the innovation deficit is structural and long-term. An economy optimized within current paradigms for a decade will encounter their limits without the generative capacity to transcend them.
Intergenerational Transmission	The most severe CDE implication. A generation raised in a high-compression environment has no prior high-entropy baseline. Recovery is not restoration but construction of capacities that were never developed. This is the civilizational stakes of letting the clock run.

RECOMMENDATIONS

Every serious diagnosis invites a serious response. The six interventions below are offered in that spirit, not as a fixed prescription but as provocations meant to spark thought and surface creative solutions. They constitute a multi-level response addressing individual behavior, economic incentive, data infrastructure, and epistemic transparency. No single intervention is sufficient, and none is meant to be the final word; the point is to show how a combination might work, and to invite better ideas in the process.

Intervention	What It Does	Key Risk
Behavioral Autonomy Score (BAS)	Personal entropy index measuring degree of independent vs. AI-ratified behavior. Friction nudges below threshold. Professional licensing integration in high-consequence domains.	Surveillance paradox: measuring freedom requires observing behavior. Must be user-owned data.
Cognitive Convenience Levy (CCL)	Tax on high-income, high-AI-dependency individuals. Revenue funds Generative Data Commons and generative education programs.	Measurement requires same surveillance infrastructure. Perverse incentives toward usage concealment.
Entropy Sabbaticals	Mandated AI-tool-free recertification periods for licensed professionals in medicine, law, engineering, and governance. Tests unassisted competency.	Competitive disadvantage without industry-wide regulation. Classic collective action problem.
Generative Data Commons (GDC)	Publicly funded repository of high-entropy, unassisted human behavioral data. Seed bank against BDEC-compressed commercial training pipelines.	Quality and representativeness risk. Skews toward educated, tech-skeptical populations without active demographic outreach.
Corporate Entropy Liability	Platforms above scale threshold conduct annual Behavioral Entropy Impact Assessments. Systematic compression triggers graduated regulatory exposure.	Regulatory capture and measurement gaming without standardized, auditable entropy metrics.
AI Disclosure at Point of Use	Standardized AI generation index on every output (document, decision, recommendation). Markets price cognitive provenance transparently.	Disclosure fatigue. Requires contextual weighting and gaming-resistant methodology to remain salient.

NEXT STEPS

Recommended Action	Why It Matters
Commission a Behavioral Entropy Baseline Study *Partner with academic institutions to calibrate the BEI framework across domains using existing proxy data*	The BEI instrument does not yet exist at the precision the problem requires. Without it, BDEC governance operates without a speedometer. The baseline study converts a structural risk into a measurable, governable variable, and it is the foundation for every intervention that follows.
Pilot Entropy Sabbaticals in One Licensed Profession *Begin with medicine or law, professions where unassisted competency is already a licensing concern and regulatory frameworks exist*	The collective action problem that makes Entropy Sabbaticals difficult is solvable at profession scale before it is solvable at economy scale. A successful pilot generates the evidence base for broader adoption and demonstrates that competency and AI augmentation can coexist with deliberate structural friction.
Establish a Generative Data Commons Seed Fund *Public-private partnership to begin collecting demographically representative, unassisted behavioral data at scale before compression further degrades signal quality*	The compressed state of commercial behavioral training data cannot be reversed after the fact; it has to be archived before it is lost. The GDC is the institutional equivalent of a seed bank, and its value is proportional to how early it is established relative to the compression event it is meant to survive.
Introduce Corporate Entropy Liability Reporting *Require platforms above defined scale thresholds to publish annual Behavioral Entropy Impact Assessments as a condition of operating license*	The chief beneficiaries of BDEC, namely the surveillance economy, industry, and governments, have every incentive to continue compression with no countervailing obligation. Mandatory reporting is the minimum structural response. It does not stop compression; it makes compression visible. And visibility is the precondition for accountability.

The most dangerous thing about BDEC is not that it ends in collapse. It is that every stage of the process looks, from the inside, like progress.

APPENDIX A: SELECTED REFERENCES

[1] Shumailov, I., et al. (2024). AI models collapse when trained on recursively generated data. Nature, 631, 755–759. ● Peer-reviewed. Foundational downstream collapse literature.

[2] Dahmani, L., & Bohbot, V.D. (2020). Habitual use of GPS negatively impacts spatial memory during self-guided navigation. Scientific Reports, 10, 6310. ● Peer-reviewed. Primary neuroscience source for navigational autonomy degradation.

[3] Murphy & Prachthauser. (2025). America's Reliance on GPS While Driving: National Survey (n=2,400+). ○ Commercial (law-firm) survey. Source for GPS dependency figures: 53% not confident driving at all without GPS, and only 12% fully comfortable without it.

[4] Microsoft & LinkedIn. (2024). 2024 Work Trend Index: AI at Work Is Here. Now Comes the Hard Part. ○ Primary source for 75% knowledge worker GenAI adoption figure.

[5] Gallup. (2025). AI Use at Work Rises. Q3 2025 Workforce Survey, n=23,068 U.S. adults. ○ Practitioner research.

[6] Similarweb. (2025). Zero-Click Searches and AI Overviews. ○ Source for the 83% zero-click rate on AI-Overview-triggered searches; zero-click news searches rose to ~69%.

[7] SparkToro / Datos. (2024). 2024 Zero-Click Search Study. ○ Practitioner research. Source for the overall U.S. zero-click baseline (~58.5% strict; ~65% widely cited).

[8] Zuboff, S. (2019). The Age of Surveillance Capitalism. PublicAffairs. ● Scholarly text. Source for surveillance economy / behavioral futures market framing.

[9] Scott, J.C. (1998). Seeing Like a State. Yale University Press. ● Scholarly text. Source for government legibility and population-standardization analysis.

[10] Shannon, C.E. (1948). A Mathematical Theory of Communication. Bell System Technical Journal, 27(3), 379–423. ● Peer-reviewed. Foundational information theory; entropy as a measure of signal diversity.

[11] Truong Xuan Khanh & Truong Quynh Hoa. (2025). Entropy Collapse: A Universal Failure Mode of Intelligent Systems. arXiv:2512.12381. ● Preprint.

[12] Dillion, D., et al. (2025). AI assessment changes human behavior. PNAS, 122(25). ● Peer-reviewed.

● = peer-reviewed or established scholarly text. ○ = practitioner or commercial research.

Methodology & AI disclosure: The research and analysis in this white paper are grounded in the peer-reviewed and practitioner sources listed in Appendix A. The following AI tools supported source discovery, drafting, structural editing, and multi-model critical review: Anthropic's Claude Opus 4.7, Google's Gemini 3.1 Pro, and Perplexity Pro (Sonar Pro / Pro Search). All sources, claims, and conclusions were verified by and remain the responsibility of the author.