

Lights-off Data Factory: Joint Modeling of Semantic Autonomy and Governance Resilience

Copula-Based Analysis of Phase Transitions in Autonomous Data Systems

Author: Sukant Pandey datamgmt2015@gmail.com

Keywords: Autonomous Governance, Copula Models, Epistemic Uncertainty, Semantic Reasoning Systems, AI Risk Management, Phase Transitions, Complex Systems, Data Governance Metrics, Human–AI Interaction, Enterprise AI Architecture, ADGMM (Autonomous Data Governance Maturity Model), AEGIS (Autonomous Epistemic Governance & Integrity System), CERES (Canonical Entity Reasoning & Epistemic Stewardship)

Abstract

Autonomous data governance emerges from the interaction of semantic reasoning, uncertainty management, and enforcement resilience. Prior work in complex systems and uncertainty modeling suggests that such interactions are nonlinear and cannot be captured through marginal metrics alone [1], [4], [5]. This paper introduces a copula-based framework for jointly modeling semantic autonomy (RAQ), governance resilience (REQ), and maturity phase transition (SST). Using large-scale simulation results, we demonstrate nonlinear collapse behavior in partially autonomous systems and show that true autonomy emerges only beyond a critical joint threshold.

I. Introduction

Autonomy in complex systems is rarely additive. Improvements in one subsystem often fail to compensate for weaknesses in others, a phenomenon well-documented in complexity science and risk modeling [4], [6]. In data governance systems, high semantic accuracy does not guarantee resilience, and low human escalation does not imply epistemic correctness.

Copula theory provides a principled method for modeling such dependencies without assuming linear correlation [1], [2], [3].

A. Relationship to the Governance-First Autonomous Data Framework

This paper is the third in a sequence examining governance-first autonomous data systems [13],[14]. The foundational framework [13] defines autonomous stewardship as an emergent property of three coupled architectural subsystems: semantic reasoning (CERES), governance enforcement (AEGIS), and maturity progression (ADGMM). The second study [14] introduces formal metrics—RAQ, REQ, and SST—to independently measure these dimensions and demonstrates their empirical separation between human-in-the-loop and Level-5 autonomous systems.

Building on these results, the present paper investigates the joint behavior of these metrics. Rather than proposing new architectural components, it analyzes how semantic autonomy, governance resilience, and maturity interact nonlinearly under scale and uncertainty. The copula-based modeling introduced here explains why partial autonomy regimes collapse and why full autonomy emerges only when the framework's components jointly exceed the Stewardship Singularity Threshold.

II. Motivation for Joint Modeling

Empirical studies in risk and governance demonstrate that systems fail catastrophically when dependency structures are ignored [6], [7]. Human-in-the-loop governance systems exhibit precisely such hidden coupling: uncertainty accumulates silently until human intervention becomes overwhelming.

III. Joint Metric Framework

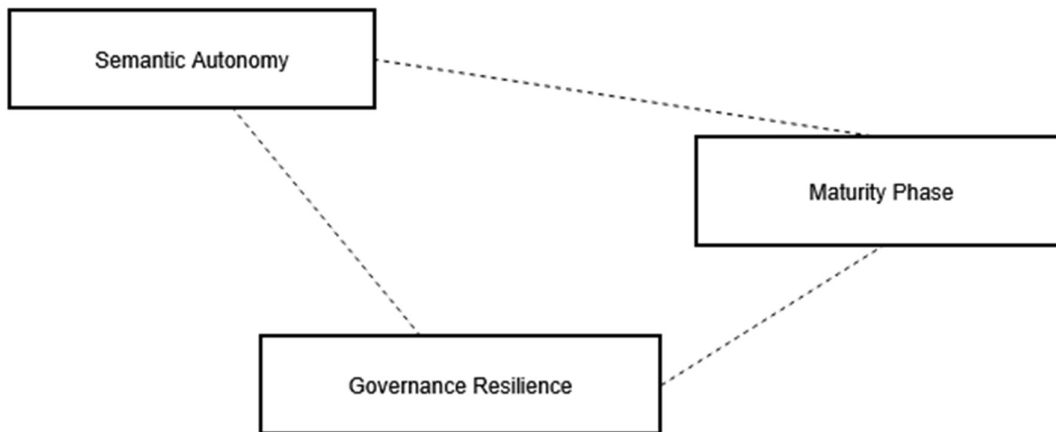


Fig. 1. Joint dependency structure between semantic autonomy (RAQ), governance resilience (REQ), and maturity phase transition (SST). Autonomy emerges only through the coupled behavior of all three metrics; no single dimension dominates independently.

The joint framework reflects the coupled nature of semantic reasoning, governance enforcement, and maturity thresholds, consistent with complex adaptive systems theory [5].

IV. Copula Formulation

Copulas enable the decomposition of joint distributions into marginal behaviours and dependency structures [1], [3].

Let:

- $X = \text{RAQ}$
- $Y = \text{REQ}$
- $Z = \text{SST}$

Define joint autonomy index:

$$J = C(F_X(X), F_Y(Y), F_Z(Z))$$

Where C is a copula capturing dependency between semantic quality, resilience, and maturity.

We employ a Clayton copula to model asymmetric collapse behavior under low resilience conditions [2].

$$C(u, v, w) = \max((u^{-\theta} + v^{-\theta} + w^{-\theta} - 2)^{-1/\theta}, 0)$$

Collapse vs Autonomy Regions in Governance-First Data Systems

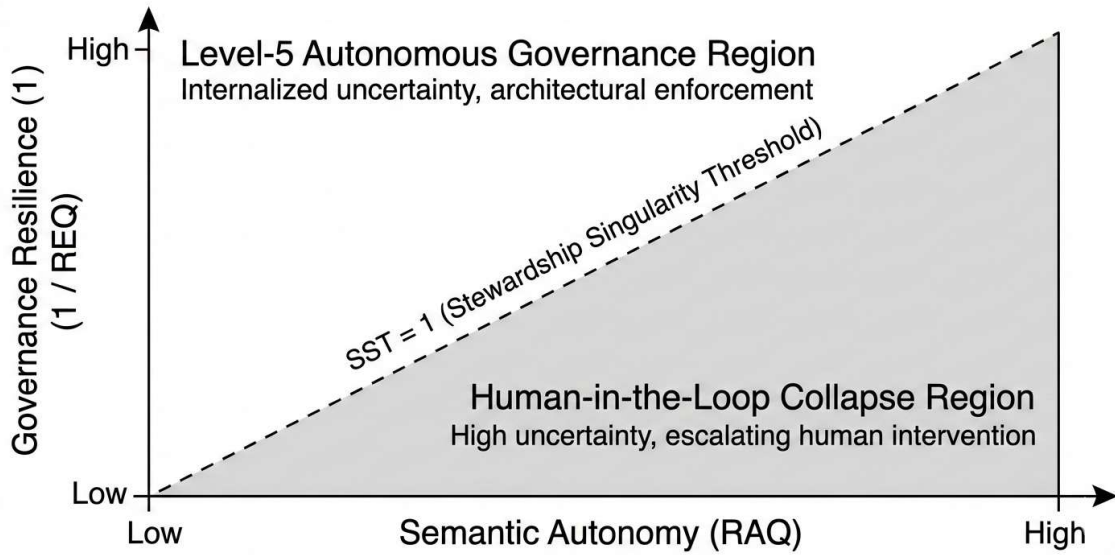


Fig. 2. Collapse versus autonomy regions as a function of semantic autonomy (RAQ) and governance resilience (inverse REQ). Systems below the SST boundary remain human-dependent, while systems above the threshold exhibit stable autonomous governance.

V. Simulation and Joint Behavior

Simulation results demonstrate that systems with high RAQ but elevated REQ fail to maintain autonomy under scale, consistent with fragility models in complex systems [7], [9]. Only systems exceeding SST across all dimensions exhibit stable joint autonomy.

VI. Phase Transition Analysis

Phase-transition behavior of this form has been observed in economic, computational, and governance systems [4], [11].

As record volume increases:

$$\lim_{N \rightarrow \infty} J_{\text{human}} \rightarrow 0$$

$$\lim_{N \rightarrow \infty} J_{\text{auto}} > \delta$$

This demonstrates structural dominance of autonomous governance.

Hence, as system scale increases, human-dependent governance collapses while autonomous systems maintain bounded uncertainty.

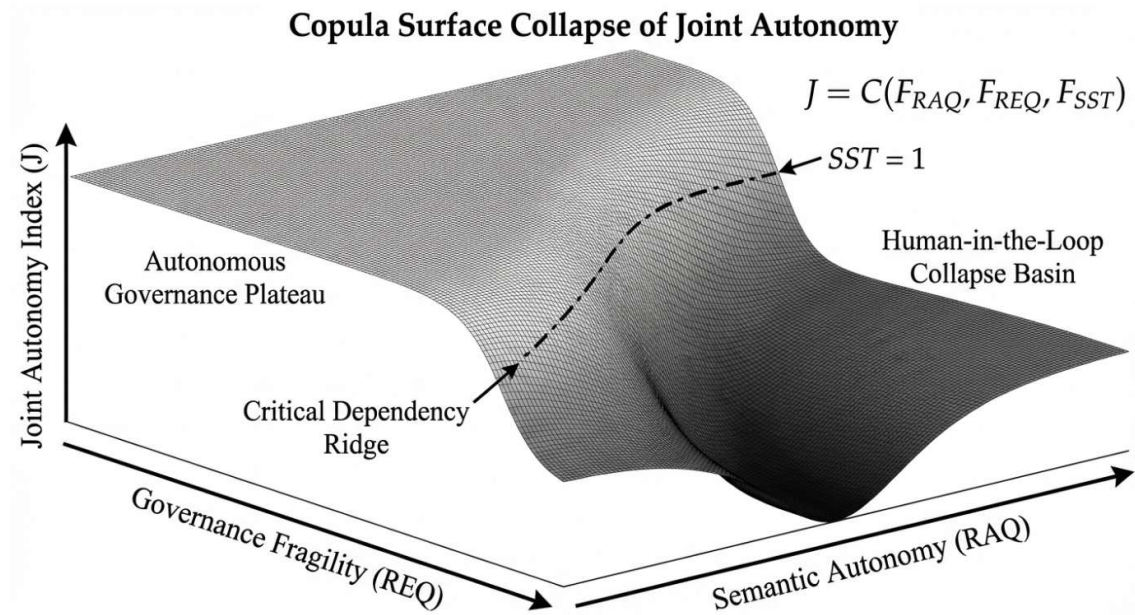


Fig. 3. Conceptual copula surface illustrating joint dependency between semantic autonomy (RAQ), governance resilience (REQ), and maturity phase transition (SST). Autonomous governance emerges as a stable high-value plateau only when all three metrics jointly exceed critical thresholds. Systems below the SST boundary experience nonlinear collapse despite improvements in individual dimensions.

Fig. 3 illustrates why marginal improvements in semantic accuracy or reduced human escalation are insufficient to achieve autonomy. The copula surface demonstrates a nonlinear collapse phenomenon in which partial autonomy regimes fall into a low-value basin unless all three dimensions—semantic correctness, governance resilience, and maturity—jointly exceed the Stewardship Singularity Threshold.

VII. Implications

These findings explain why AI-assisted governance systems stall despite improvements in accuracy and throughput, echoing critiques of superficial automation in AI ethics literature [8], [10], [12].

VIII. Conclusion

Joint modeling reveals autonomy as an emergent system property that cannot be achieved through incremental improvements alone. Copula-based analysis provides a formal foundation for evaluating and designing governance-first autonomous data systems.

References

1. R. B. Nelsen, *An Introduction to Copulas*, 2nd ed. New York, NY, USA: Springer, 2006.
2. H. Joe, *Dependence Modeling with Copulas*. Boca Raton, FL, USA: CRC Press, 2014.
3. M. Sklar, "Fonctions de répartition à n dimensions et leurs marges," *Publ. Inst. Stat. Univ. Paris*, vol. 8, pp. 229–231, 1959.
4. S. Wolfram, *A New Kind of Science*. Champaign, IL, USA: Wolfram Media, 2002.
5. M. Mitchell, *Complexity: A Guided Tour*. Oxford, U.K.: Oxford Univ. Press, 2009.
6. D. Helbing, "Globally networked risks and how to respond," *Nature*, vol. 497, pp. 51–59, 2013.
7. N. Taleb, *Antifragile: Things That Gain from Disorder*. New York, NY, USA: Random House, 2012.
8. S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Upper Saddle River, NJ, USA: Pearson, 2010.
9. B. Schneier, *Liars and Outliers*. Hoboken, NJ, USA: Wiley, 2012.
10. L. Floridi, "The ethics of information," *Oxford Handbook of Information and Computer Ethics*, Oxford Univ. Press, 2010.
11. K. Arrow, "The economic implications of learning by doing," *Review of Economic Studies*, vol. 29, no. 3, pp. 155–173, 1962.
12. ISO/IEC TR 24028, *AI Trustworthiness*, ISO, 2020.
13. P. Sukant, "Lights-off Data Factory: A Governance-First Architecture for Level-5 Autonomous Data Systems", Zenodo, Preprint, 2025.
14. P. Sukant, "Lights-off Data Factory: Measuring Epistemic Autonomy in Governance-First Data Systems", Zenodo, Preprint, 2025.