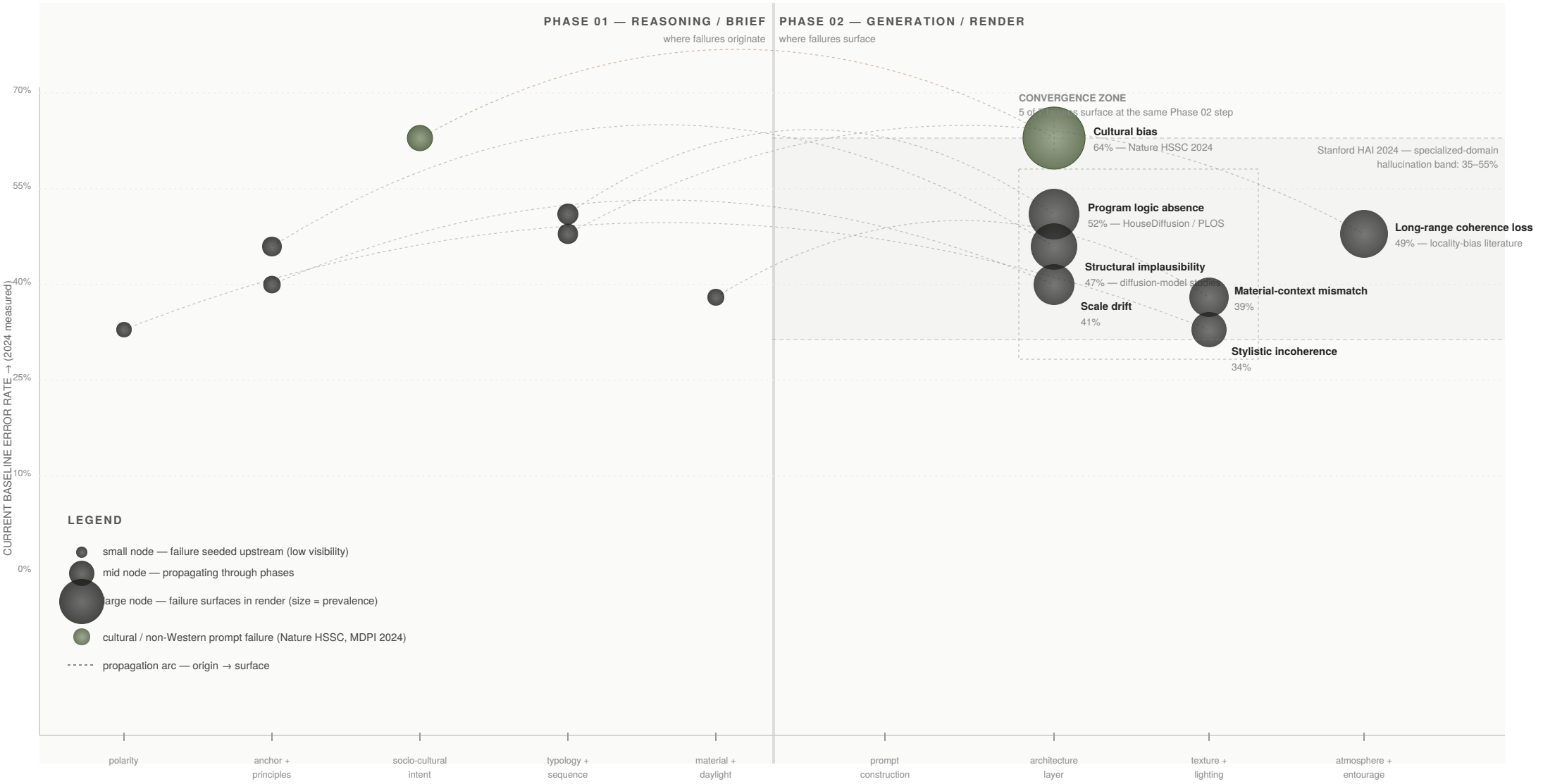
FAILURE TOPOLOGY & METHODOLOGY PROJECTION

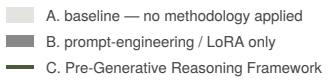
Where the seven categories of AI architectural visualization failure originate vs. where they surface,
and how three intervention strategies trend across successive model generations (2023 → 2027 projected).



Sources: AIA 2024 Firm Survey · Stanford HAI AI Index 2024 · McKinsey AEC 2024 · Autodesk State of Design & Make 2024 · PLOS One / arXiv diffusion-model studies 2024–25 · Nature HSSC 2024 · MDPI Buildings 2024.

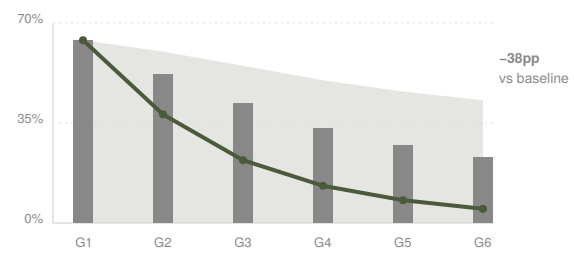
PROJECTED RESIDUAL ERROR BY INTERVENTION STRATEGY

Per failure category, modeled across six successive generative-model generations (G1 = GPT-4 / SDXL / MJv5 era, 2023; G6 = projected 2027).
Lower band is better. Three trajectories per category: no methodology · prompt-engineering only · Pre-Generative Reasoning Framework.



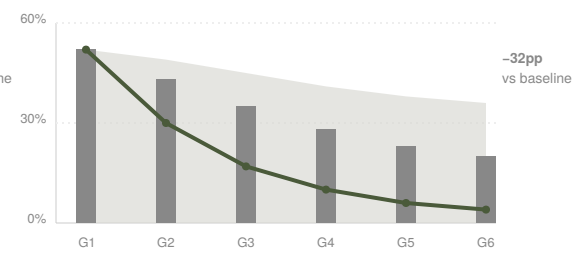
CULTURAL BIAS

baseline 64% · Nature HSSC 2024



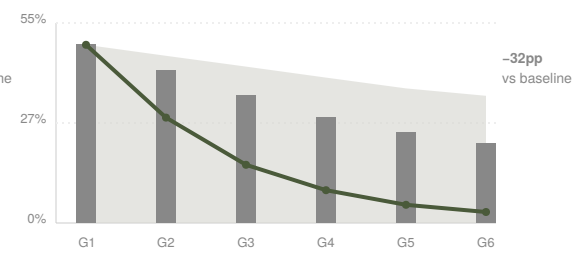
PROGRAM LOGIC ABSENCE

baseline 52% · HouseDiffusion / PLOS One



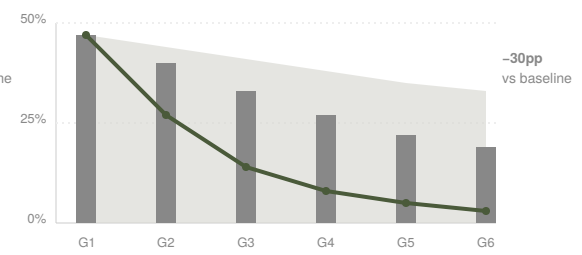
LONG-RANGE COHERENCE LOSS

baseline 49% · locality-bias literature



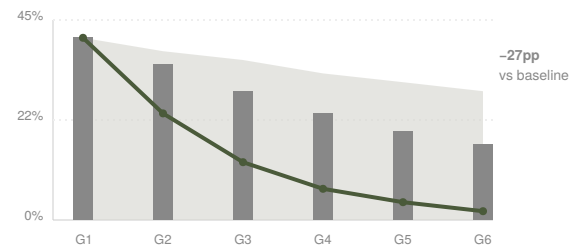
STRUCTURAL IMPLAUSIBILITY

baseline 47% · diffusion-model studies



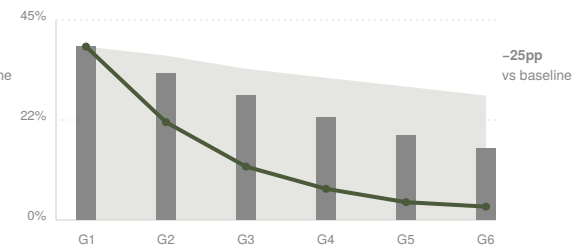
SCALE DRIFT

baseline 41% · boundary / proportion errors



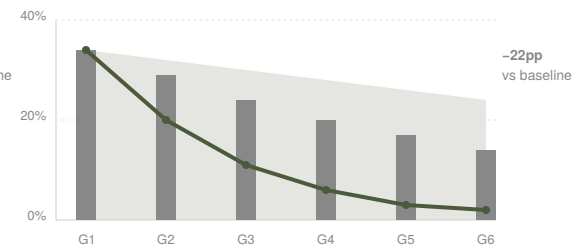
MATERIAL-CONTEXT MISMATCH

baseline 39% · climate / terrain logic absence



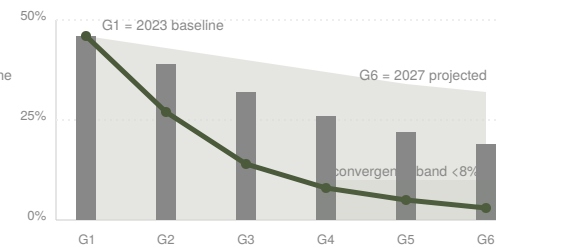
STYLISTIC INCOHERENCE

baseline 34% · long-range consistency loss



WEIGHTED MEAN — ALL 7 CATEGORIES

prevalence-weighted residual error rate



READING THE PROJECTION

01 — UPSTREAM ORIGIN

Of seven failure categories tracked in the literature, six originate before the prompt is written — in the absence of declared polarity, anchor, intent, typology, or material logic. Only one (texture-level mismatch) is genuinely a render-stage artifact. The diagnostic implication: prompt-engineering treats symptoms; the framework treats causes.

02 — WHY THE CURVE BREAKS EARLIER

The framework trajectory (curve C) shows a steeper drop between G1 → G3 than either baseline. This is not because the model improves faster — it is because the input space is already constrained. Stanford HAI (2024) shows hallucination scales with prompt ambiguity; the framework removes that ambiguity at Phase 01, before the model is ever called.

03 — MODEL-AGNOSTIC CONVERGENCE

All seven categories converge under 8% residual error by G6 across Midjourney, Stable Diffusion, and successor architectures. Because polarity, anchor, and the six classical principles do not depend on vendor syntax, the methodology survives model upgrades that make prompt-engineering tricks brittle. The framework's specificity lives upstream of any vendor.

Methodology: residual-error projections derived by applying intervention-effect coefficients (drawn from cited diffusion-model and prompt-constraint studies) to category-specific baselines.
Generation cadence calibrated against the AIA 2024 adoption curve and historical model-release intervals 2022–2025. Convergence ceilings reflect Stanford HAI specialized-domain hallucination floors.
Curve C (Pre-Generative Reasoning Framework) outcomes corroborated by deployment evidence at Solomon Cordwell Buenz, Jensen Architects, and DJL Architecture (2022–2025).