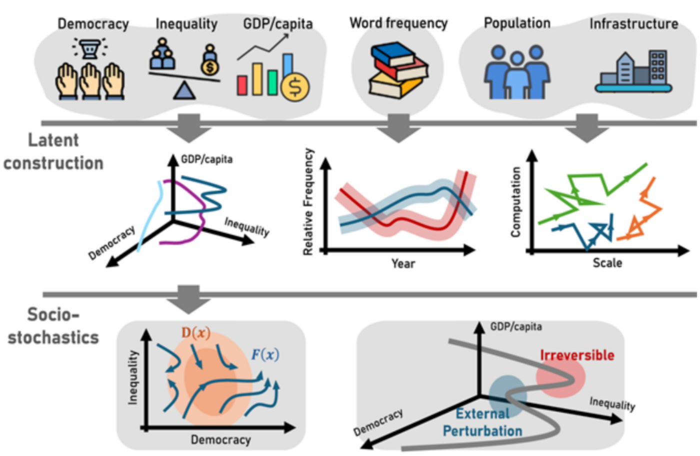
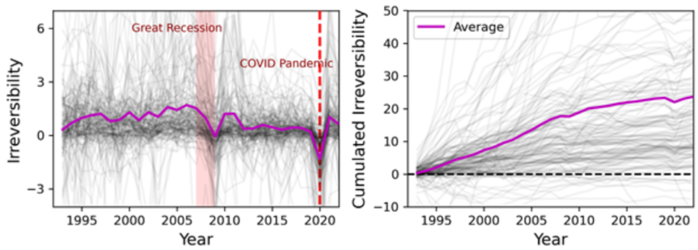
A Stochastic Dynamics Framework for Social Evolution from Observed Data

Youngkyoung Bae[1], Hajime Shimao[2], Seungwoong Ha[3], Luna Yang[4], David Wolpert[3]

[1]Seoul National University, Seoul, South Korea, [2]Penn State Great Valley, Malvern, USA, [3]Santa Fe Institute, Santa Fe, USA, [4]Pennsylvania State University Brandywine Campus, Media, USA





Are there hidden dynamical regularities underlying the evolution of social and cultural history? The rapid growth of digitized social data makes this question increasingly tractable, yet most quantitative approaches still rely on static comparisons, deterministic snapshots, or average effects. Such methods often miss the fact that historical change unfolds continuously over time and is shaped by both structural tendencies and intrinsic randomness. In this paper, we propose a framework for modeling social dynamics as stochastic processes governed by stochastic differential equations (SDEs). By treating societal evolution as a continuous-time process, our approach provides a principled framework for describing how long-term structural forces and endogenous fluctuations jointly shape historical trajectories, moving beyond simple accounts based solely on structural determinism.

Methodologically, our framework assumes Markovian dynamics and models observed historical trajectories as overdamped Langevin processes. Within this setting, we infer the underlying stochastic dynamics by estimating the structural trends (drift) and intrinsic fluctuations (diffusion) that govern their evolution, even under fragmentary and irregular observations. This allows us to reconstruct latent dynamics from incomplete data and to derive dynamical quantities that are inaccessible to standard static or deterministic approaches. Based on the inferred stochastic dynamics, we quantify the irreversibility of social change and estimate exogenous perturbations beyond those explained by endogenous dynamics alone.

We apply this framework to three distinct empirical settings: the modern Democracy-Inequality-GDP/capita dataset, the historical Seshat dataset, and the Google N-gram dataset. Across these cases, our approach enables us to examine how events affecting individual social agents—such as words, countries, or civilizations—as well as system-wide shocks shape the irreversibility of the overall dynamics. Here, irreversibility serves as a dynamical diagnostic of how strongly observed trajectories align with the intrinsic probability current of the inferred system. Specifically, we find that major disruptions, including the Great Recession and COVID-19, are associated with declines in average irreversibility across countries (see Fig. 2). More broadly, these results suggest that social evolution is better understood as a continuous stochastic process shaped by both structural forces and random fluctuations, rather than as a mere sequence of discrete states.

References:

[1] Youngkyoung Bae et al., Unveiling hidden features of social evolution by inferring Langevin dynamics from data. arXiv preprint arXiv:2601.17772 (2026).