

The shrinking island

RodMalpert.com

Executive summary

The paper studies documented human-AI collaborations, including philosophical, historical, forecasting, and creative work. Human contributions include framing, analogy, criticism, and research direction; AI contributions include synthesis, calculation, source retrieval, conceptual correction, and prose. Recorded contributions do not establish which functions an adequately instructed AI could not perform alone.

The Bostrom-Yampolskiy interview adds an external illustration of AI-assisted research and rapid intellectual communication. Research taste includes choosing valuable questions, recognizing a useful abstraction, and demanding precise formulations. Those are tasks to evaluate, not self-evidently permanent human monopolies.

The five-paper review demonstrates why a portfolio needs more than local arithmetic checks. It corrected outdated companion-survival references, distinguished differently conditioned probabilities, and separated generated-mind preservation from a residual local-production branch. The resulting sensitivity can cross 50% continuation without supplying empirical evidence for the chosen conditional.

Malpert selected the interview and directed the cross-paper goals; ChatGPT carried out the present comparison, formalization, calculations, and drafting. This is attribution within the collaboration, not a claim of historical priority or independent validation by the interviewees.

Timeliness affects research usefulness. Interviews, working papers, and short arguments can inform decisions sooner, provided their claims remain linked to inspectable support, uncertainty, and correction. Recentness and prestige are not likelihood multipliers.

A prospective comparison should hold task, source access, and resources constant across human-only, AI-only, and collaborative workflows. It should score question selection, omitted-mechanism discovery, source fidelity, error rates, uncertainty, and time to a useful result. Comparing open-ended AI assignments with explicitly instructed audits would help distinguish autonomous research direction from recoverable elicitation failures.

Source: Bostrom, Nick, interviewed by Roman Yampolskiy (2026). "Superintelligence Author: We May Already Have AGI | Nick Bostrom." The Roman Forum with Roman Yampolskiy, YouTube, October 16. <https://www.youtube.com/watch?v=m880ajt5sNM>. Interview discussion; conceptual claims and reported judgments, not an empirical calibration dataset.