

PROPOSAL FOR A “SECOND PROGRAM” (SHORT FORM)

We argue, with Isabel Millar (*The Psychoanalysis of Artificial Intelligence*, 2021) and Giambattista Vico (*New Science*, 1725), that AI is ancient and even foundational for human thinking. We envision a Second Program that focuses on interface. ChatGPT should provoke new *kinds* of behaviors. Research should aim to test interface subjectivity to spin off new special projects in all current user interfaces.

We envision a new ambitious era of AI as a whole. Our Second Program should be scalable and verifiable, drawing from ethnological evidence that AI has been present, as a “foundational prosthesis,” from the beginning of human thought (<https://sites.psu.edu/boundarylanguage/files/2023/06/OpenAI-proposal-long-form.pdf>). We employ topology to define *symmetrical difference*, the essential logic behind this prosthetic relation to ritual practices, religious beliefs, social formations, and the production of arts and cultural institutions.

All of these employ prosthesis to effect conditions of instrumental convergence, but until now theory has not considered how the human subject, topology, AI, and prosthesis are combined. The Second Program aims to construct “polythetic” interfaces — blurs, “tolerance spaces,” wedges. It proposes workshops/symposia, where AI theorists and Lacanian psychoanalysts meet to collaborate on an alternative AI interface. This is not about tweaking current practices, but rather realizing a new paradigm involving ethnology, AI, psychoanalysis, and topology as if they had been conditioning each other since the beginnings of humanity. Out of this collaboration, existing operations benefit.

Nick Bostrom (*Minds and Machines*, 2012) articulated five points as the foundation of AI. (1) Self-preservation; (2) Goal-content integrity; (3) Cognitive enhancement; (4) Technological perfection; (5) Resource acquisition (controlling systems and processes to reach primary goals). We must revise these foundational assumptions. We revise the goal self-preservation in relation to resource acquisition through the idea of “instrumental convergence,” but point out that Bostrom does not allow for IA failure (~1), either by crashing (accidental) or suicide (intentional). To realize goal-content integrity (3), however, AI is capable of guiding more interesting user-behaviors, where convergence creates resonance (4) and considers anticipation as a resource (5).

We revise Bostrom’s thesis to recover a *foundational* structure. Our “back to the future” proposal shows how technological perfection can be achieved *if and only if* self-sacrifice (oblation), implicit in the process of “instrumental convergence,” is implicit; and *if and only if* “archaic” models of oblation in ethnology, art, and literature are reconsidered.

What if HAL, Kubrick/Clarke's spaceship computer, had discovered a Higher Mission than the one his programmers had set, but knew that this more ambitious program had to proceed without him? This conjecture plays out the predictions of Norbert Wiener (*Cybernetics*, 1948), who argued that *both* human and machine thinking trended to breakdown, adding complexities exponentially increasing the likelihood of crashing. Using the human-machine as a dual, we use Lacanian psychoanalysis (the only form of psychology to explicitly address AI) to consider the question of self-termination, inadvertent or designed, as the decisive issue of instrumental convergence. It will take three seminars and five speakers to explain this thesis. [499 words]