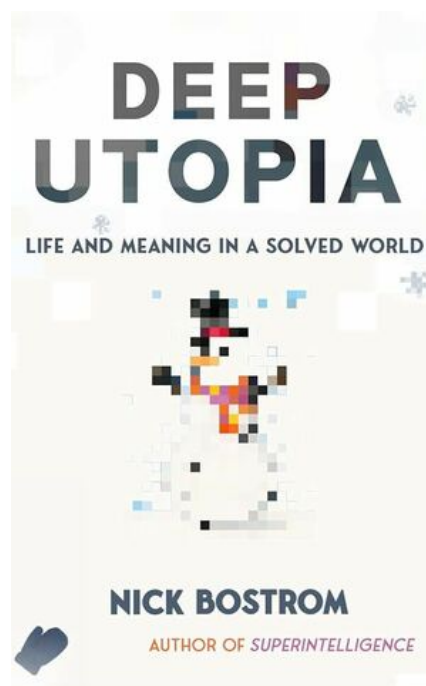


Deep Utopia: Life and Meaning in a Solved World

Nick Bostrom, *Deep Utopia: Life and Meaning in a Solved World*, Ideapress Publishing, 2024, 525pp., \$32.95 (hbk), ISBN 9781646871643.

Reviewed by Matthew Hammerton, Singapore Management University

2024.11.4



Nick Bostrom is perhaps best known for his 2014 book *Superintelligence*, which explores the existential threat posed to humans by superintelligent AI. In his new book, *Deep Utopia*, he shifts gears to explore the *existentialist* threat that superintelligent AI poses. Suppose that humans successfully navigate the technological, moral, and political challenges that come with advanced AI. The result of such success, according to Bostrom, would be a kind of deep utopia. We would not only live in a post-work, and post-scarcity society (things already covered by shallower conceptions of utopia); we would also live in a “post-instrumental” society where almost anything we might consider doing, including pursuing our own personal enrichment, could be better accomplished by AI. Bostrom’s central concern is how we might find meaning and purpose in such a world. If everything you might consider doing can be done better by machine intelligence, what is there to live for?

Before we delve into Bostrom’s arguments, prospective readers should be aware (and beware) of two ways in which this is an atypical philosophy monograph. First, Bostrom’s book is not tightly organized around a central thesis that it systematically defends. Instead, it explores its subject matter in an open-ended way, trying out various ideas, pursuing tangents, and hedging on several key questions.

Second, the book employs an experimental style and structure. It consists of a series of fictional lectures that Bostrom imagines an older version of himself delivering over the course of a week (each chapter takes a day of the week as its title).

Thankfully they do not read like stale lecture notes—students regularly interject with questions and various events (including a blackout and a pesky dean cutting short a lecture) add drama and inject humour into the proceedings. The text also includes handouts distributed during the lectures and assigned readings—the latter being standalone literary pieces, often allegorical, that Bostrom wrote specifically for this book. The framing narrative within which all of this occurs is the interactions of three philosophically astute friends who are gatecrashing Bostrom’s lectures and like to wind down afterwards at the public baths. Their sharp dialogue, including reactions to the lectures and readings, bookends many chapters.

These two features—the open-ended approach and experimental structure—will likely delight some readers and infuriate others. Detractors will call the book “bloated”, arguing that a more focused version of it, without the literary pretensions and tangential discussions, could be one-third the length. Others will find Bostrom’s non-standard approach refreshing and will relish the playfulness and wit that pervade the book. My reaction was primarily the latter, though sometimes veering toward the former. Overall, I would describe the book’s non-standard approach as a worthwhile but not wholly successful experiment.

Now, let us turn to the philosophical content in *Deep Utopia*. The book is packed with arguments, thought experiments, case studies, and empirical data. In this short review I cannot cover everything of interest and will limit myself to a few points that merit particular attention.

In the book’s third chapter, Bostrom discusses how we might find meaning and purpose in a post-work world where virtually all occupational labour has been automated. His answer is that we need to develop a leisure culture that empowers and educates individuals to thrive in the absence of traditional employment. This culture would “encourage rewarding interests and hobbies, and promote spirituality and the appreciation of the arts, literature, sports, nature, games, food, and conversation, and other domains which can serve as playgrounds for our souls that let us express our creativity, learn about each other and about ourselves and about the environment, while enjoying ourselves and developing our virtues and potentialities”. While not the first to advocate such an idea, Bostrom’s discussion is

especially poignant and persuasive and includes a set of interesting case studies of historical precedents for this leisure culture. Thus, it makes a valuable contribution to contemporary debates about the promises and pitfalls of a post-work society.

Bostrom ultimately views the problem of post-work purpose as “relatively superficial” when compared to the deeper issues raised by a post-instrumental society. He labels the former problem “shallow redundancy”, contrasting it with the “deep redundancy” of the latter. Most of *Deep Utopia* is devoted to exploring the problem of deep redundancy. However, this leads to some highly speculative and futuristic places. Among the ways that human effort is redundant in a post-instrumental world, Bostrom mentions exercise being rendered unnecessary by nanobots that can physically condition our bodies while we sleep, and learning becoming pointless due to AI-directed brain editing that can incorporate new information and skills into our minds without the need for study (there is a five page handout with footnotes that analyses the prospects of this).

Much of this discussion is fascinating, but the question of how we will find meaning and purpose in a post-instrumental world of futuristic technologies that make us post-human will strike many as a low priority in our present moment. This will lead some readers to wonder why they should really care about the book’s central concern. Bostrom addresses this head-on in the opening pages before the reader has a full sense of what the book will cover. He tells us that we should care because “Our civilization looks to be approaching a critical juncture” where we will face consequential choices “about what kind of future we want”. Rather than allowing this juncture to catch us unprepared, we ought to engage with these questions proactively. Additionally, he suggests that exploring these issues “can serve as a kind of philosophical particle accelerator in which extreme conditions are created that allow us to study the elementary constituents of our values”.

These are excellent points to make. However, Bostrom unfortunately never circles back to them. The book would have benefited from several reminders of these motivating concerns as well as concrete details about how specific points of discussion contribute to them. Especially jarring is the lack of a conventional conclusion that crisply articulates the point and payoff of the preceding discussion (another sacrifice of the book’s experimental approach). As a result, many readers may remain unconvinced of deep redundancy’s importance as a philosophical problem worthy of such extensive attention.

Setting these concerns aside, let's consider Bostrom's response to the post-instrumental purpose problem. Overall, he is optimistic about the prospects of life in deep utopia. Although the open discussion style often leaves definite conclusions elusive, he emphasizes two key reasons for optimism. First, he argues that, even if lives in deep utopia have a meaning deficit, they would be incredibly rich in other ways that might make up for this. Such lives could be filled with immense pleasure stimulated through intense experiences of "heartrending beauty". With cognitive enhancements and sophisticated AI content programming, our intellectual, emotional, and aesthetic capacities could be fully engaged. Furthermore, such lives need not be passive—even if human effort is redundant, AI could design engaging tasks and challenges for us to tackle (Bostrom calls this "artificial purpose") that utilize our enhanced capabilities.

No doubt some would be fully satisfied with such a life. But others will still feel the sting of the purpose problem—if all human effort is redundant, what is the point of it all? This leads Bostrom to engage with the philosophical literature on meaning in life, particularly the theory of Thaddeus Metz, a prominent scholar in the field. Bostrom acknowledges that, according to Metz's theory, life in deep utopia would lack one of the crucial ingredients required for a meaningful life—namely being orientated towards the good in a way that actually produces value in the world. If human effort is entirely redundant, then we would no longer be in a position to produce value and thereby fail to meet this standard for meaning. However, Bostrom notes that we could still orient ourselves towards the good simply by being a virtuous person who loves and appreciates the good around us. Indeed, deep utopia would offer us better opportunities to do this than our current world. This leads Bostrom to ask: why isn't this enough to make our lives meaningful?

Bostrom pushes this thought further by developing his own theory of meaning in life, which he terms "meaning as encompassing transcendental purpose". This theory includes familiar elements from other theories of meaning in life put together in a somewhat original way. While not groundbreaking, it merits consideration alongside other contenders in the literature on meaning in life. Crucially, according to this theory, a meaningful life does not require actively producing good—it can emerge from other pathways, including virtuously loving and appreciating the good.

As an exploration of these questions, Bostrom's discussion is highly effective. However, some readers may be disappointed that he doesn't push towards more definite answers. It is one thing to suggest that the rich experiential goods found in deep utopia may compensate for a deficit in meaning; it is another to defend an account of how different sources of value in human lives should be weighed against one another to produce an overall assessment of a life's value. But the latter is what we need to determine whether the compensations of deep utopia are indeed worth it. Likewise, although Bostrom helpfully shows us that there is at least one plausible theory of meaning in life that allows for a highly meaningful life in utopia, he does not systematically defend it against competing frameworks. In response, Bostrom might note that delving deeply into the technical aspects of value theory was not the intention behind this book.

In conclusion, *Deep Utopia* is an ambitious, unconventional book that offers fascinating discussions and original arguments about humanity's possible post-instrumental future. Readers who are expecting something with the power, focus, and impact of *Superintelligence* will be disappointed. Nonetheless, the book is a timely contribution to ongoing debates about how future advances in artificial intelligence will impact us. It will especially be of interest to those working on AI, the future of work, and meaning in life. However, philosophers outside of these areas will likely find aspects of the discussion both provocative and intellectually rewarding.