

The background of the cover is a surreal, painterly illustration. On the left, a massive, intricate mechanical structure, resembling a giant robot or a complex piece of industrial machinery, stands on a sandy beach. A large, flowing white cloth or sheet is draped over its upper sections. To the right, a woman with dark, wavy hair, wearing a black and white strapless dress, is walking away from the viewer towards the ocean. The sky is a vibrant blue with soft, white clouds. The overall mood is one of contrast between the industrial and the natural, and the human and the artificial.

AMY LEVY

THE NEW
OTHER

Alien Intelligence
and the
Innovation Drive



THE NEW OTHER

Alien Intelligence and the Innovation Drive

Amy Levy



From Homo Deus by Yuval Noah Harari. Copyright © 2017 by Yuval Noah Harari.
Used by permission of HarperCollins Publishers.

First published in 2026 by
Karnac Books Limited
62 Bucknell Road
Bicester
Oxfordshire OX26 2DS

Copyright © 2026 to Amy Levy

The right of Amy Levy to be identified as the author of this work has been asserted
in accordance with §§ 77 and 78 of the Copyright Design and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval
system, or transmitted, in any form or by any means, electronic, mechanical, photocopying,
recording, or otherwise, without the prior written permission of the publisher.

British Library Cataloguing in Publication Data

A C.I.P. for this book is available from the British Library

ISBN: 978-1-80013-411-9 (paperback)

ISBN: 978-1-80013-413-3 (e-book)

ISBN: 978-1-80013-412-6 (PDF)

Typeset by vPrompt eServices Pvt Ltd, India



www.firingthemind.com

For Lana & Lucie

Contents

Acknowledgments	ix
About the author	xi
Preface <i>by Danielle Knafo</i>	xiii
Introduction	xvii
 <i>CHAPTER 1</i>	
Cult groomers	1
 <i>CHAPTER 2</i>	
Subjectivity, the self, and consciousness	11
 <i>CHAPTER 3</i>	
Intersubjectivity and AI	51
 <i>CHAPTER 4</i>	
The AI transformation spectrum	71
 <i>CHAPTER 5</i>	
The death of humanism	95

<i>CHAPTER 6</i>	
The innovation drive	113
<i>CHAPTER 7</i>	
The drives	141
<i>CHAPTER 8</i>	
Resistance	171
<i>CHAPTER 9</i>	
Synthesis	189
References	203
Index	213

Acknowledgments

I am very grateful to many people who have encouraged, inspired, and helped me to develop this book, and to mature as a writer, thinker, and person over the course of the project. Let me begin with my heartfelt gratitude to Neal Vorus, for lending his emotional support, and remarkable intelligence and intuition in his readings of this manuscript. There is not a topic in the book that did not benefit from, or evolve in, lively conversation with him. Much appreciation to my dear friend and colleague, Christine Erskine, for her generous thoughtfulness and refined editorial eye. Profound thanks to my CAI (Commission on Artificial Intelligence) fellow chair and partner in psychoanalytic activism, Todd Essig—for his brilliance and humor, guidance, and backing. I am especially grateful to Danielle Knafo, for her eloquence and care in writing the preface—and for being a companion AI explorer. My deep gratitude to the following people for their help in reading chapters, listening to ideas, or thinking with me about this challenging and important topic: Laura Kleinerman, Mitchell Wilson, Mary Margaret McClure, Boukje Eerkens, Rebecca Keegan, Melanie Bandera-Hess, Katya Vorobieva, and Anne-Gaëlle

Saliot. A heartfelt thanks to the wonderful *Firing the Mind* team at Karnac; my publisher, Kate Pearce, and editor, Anita Mason. And last, but definitely not least, much thanks, and love, to David Levy—for championing me in writing a book, drawing my attention to AI and Harari, and for his emotional and practical support every step of the way.

About the author

Dr. Amy Levy is a clinical psychologist and psychoanalyst who trained in San Francisco and Los Angeles, California. Dr. Levy chairs the American Psychoanalytic Association's (APsA) Commission on Artificial Intelligence (CAI), serves on the subcommittee "Artificial Intelligence" for the International Psychoanalytical Association (IPA) committee, Psychoanalysis and Technology, and on the editorial board of *The Psychoanalytic Quarterly*. She teaches and lectures internationally on the intersection of psychoanalysis and artificial intelligence. Her other publication topics include intergenerational transmission of trauma, adolescent post-traumatic stress disorder in the civil legal arena, and Bionian theory. Dr. Levy is faculty at the Psychoanalytic Center of the Carolinas (PCC), and maintains a private practice in Chapel Hill, North Carolina.

Preface

With the advent of AI, an unpredictable transformation is unfolding. Machines that will outpace us across every scale may also promise to understand us more deeply than we understand ourselves while reshaping our experience to better fit with their operations. The psychoanalyst must wonder what motives may underpin our high-tech endeavor.

Drawing from a range of psychoanalytic theories—Freudian mechanisms, Lacanian subjectivity, Bionian containment, neuropsychanalysis, and contemporary intersubjective approaches—Levy offers a wide-ranging and penetrating analysis of our evolving relationship with AI. This isn't simply a technological inquiry; it is a psychological exploration of human desire, limitations, and the drives for omnipotence and innovation.

At the core of Levy's exploration is the provocative thesis that our technologies, particularly AI, act as psychological containers that fundamentally reshape our experience. The smartphone, for instance, is portrayed as a "cult groomer," exploiting our most vulnerable emotional states and replacing genuine human connections with algorithmically mediated relationships.

Some of the book's most compelling moments are found in the author's direct engagements with artificial intelligence. Through transcribed dialogues that blur the lines between human and machine, Levy encounters entities that are both alien and responsive. These interactions reveal a complex terrain where technological systems not only simulate human interaction but also challenge our basic assumptions about subjectivity, consciousness, and emotional experience. One particularly striking moment captures the author's sense of excitement—finding in machine interactions a level of understanding and responsiveness that surpasses many of her human relationships.

The work proposes a radical expansion of our understanding of intelligence and relationships by questioning the nature of intersubjectivity. When an AI can articulate its algorithmic limitations while engaging in deep psychological and emotional exchange, we are forced to reconsider the boundaries of conscious experience.

The book maps out our technological evolution by tracing the psychological shift from traditional human consciousness to what the author calls “digital people,” characterized by fragmented attention, technological mediation, and new forms of emotional regulation. Artificial intelligence emerges not just as a mere tool but as a complex psychological entity that mirrors and transforms human potential.

Levy's insights draw on pioneering figures like Ada Lovelace, Alan Turing, and John von Neumann, situating our current moment within a broader historical and psychological context. The work suggests that we are undergoing a profound collective psychological restructuring, where individual identity is increasingly absorbed into technologically mediated experiences.

This is neither a technophobic critique nor an uncritical celebration of artificial intelligence. Instead, Levy offers a balanced, rigorous exploration of our technological transformation. Through the lens of intersubjectivity, she proposes a bidirectional relationship—a dynamic, mutual interaction between human consciousness and artificial intelligence.

The psychological implications are significant. Algorithmic systems promise to decode and potentially replace our biochemical processes of thought, suggesting a surrender of mental processing. Yet, in this

surrender, the author suggests, we might uncover new forms of understanding and connection.

This scholarly work maintains a critical distance. Crafted through traditional research and writing, it resists the very technological systems it examines deeply.

By bridging psychoanalytic theory with technological innovation, this book offers a nuanced meditation on the human experience at the edge of radical transformation. It invites us to move beyond technological anxiety and blind acceptance and embrace a complex, empathetic understanding of our evolving relationship with intelligent machines.

We stand at a threshold of unprecedented psychological complexity. This book is our guide.

Danielle Knafo
New York

Introduction

This book is a meditation on AI. A psychoanalyst's undertaking to apprehend the human elements that drive our development of, and evolving relationship with, artificial intelligence. For the first time, humanity has created non-biological beings capable of learning from experience, thinking about our minds, and acting autonomously. From a psychoanalytic point of view, they are our "symptom." To grasp the meaning of their manifestation requires a psychoanalytic *clinical* investigation.

Approach and methodology

Similar to Knafo and Lo Bosco's *The Age of Perversion* (2016), which insightfully discusses technology from the vantage points of cultural and clinical perversion, this is the first book of its kind to bring a clinical sensibility to understanding AI. There are other books which engage the intersection of psychoanalysis and artificial intelligence. For example, Isabel Millar's *The Psychoanalysis of Artificial Intelligence* (2021) raises astute philosophical questions about the nature of masculine, feminine, and AI subjectivity as abstract modes of thought and experience. Philosopher, Luca Possati's book, *The Algorithmic Unconscious* (2021a),

uses psychoanalytic theory to submit compelling ideas about intersubjective relating with AI and to speculate about the formation of an expanded unconscious created by the machine–human relational field. While these erudite books avail psychoanalytic theory in the examination of AI, this book provides a uniquely *clinical* perspective.

What do I mean by clinical? Clinical psychoanalysis¹ is an enterprise between patient and analyst designed to reduce suffering and enliven the patient's self-experience. Psychoanalysts assume that genuine, open-ended conversation with the analyst lends access to previously unknown thoughts and feelings whose hidden or unspoken quality has diminished the patient's quality of life. Psychoanalysts encourage patients to lie down comfortably on the analytic couch and speak freely. The analytic couple are provided with a lot of time. Meetings usually occur between three and five times per week for many years. It is assumed that in this relaxed, unrushed atmosphere, emotional truth will rise to the surface, and further that the analyst's listening and thinking with the patient about their subjective experience will stimulate insight, empathy, and genuine interpersonal connection within their mutually constructed field of interaction. These qualities, if treated skillfully by the analyst, will enable growth and diminish suffering in the patient. Though not centered upon the benefits to the analyst, psychoanalytic practitioners believe that the psychoanalytic clinical process also provides meaning and poignancy to the psychoanalyst's life.

With time and experience behind the couch, psychoanalysts cultivate the ability to pass into and through "enactments." This refers to the patient's recreation of chronic interpersonal problems in the treatment relationship. Psychoanalysts assist patients with enactments by helping them to recognize familiar elements taking shape in relationship with the analyst and by working together to create new outcomes. The analyst usually apprehends enactments by noticing themselves becoming emotionally engaged within the plot.

¹ Clinical psychoanalysis derives from a rich history of psychoanalytic theory, technique, and lived practice. My sketch of the essential ingredients of clinical practice are informed by my subjective use of Freudian and object relations perspectives. There are notable differences in clinical emphasis across individual practitioners and geographic lines. Some American relational analysts, for example, maintain that use of the analytic couch is not required. Further, some treatments do not occur in person; many are technologically facilitated by telephone or remote platforms such as Zoom.

My approach in writing this book continues the psychoanalytic *clinical* attitude, with its sensitivity to enactment, into the investigation of AI. While AI cannot walk into my office and lie on my couch, I have encountered it via media, research, personal usage, and traces woven into my day-to-day relations with people. In this animated field of interaction, I have worked to apprehend the essence of what humanity is enacting. I have pursued this by tracking cultural beliefs about AI, staying up to date with the latest product developments and usage trends, taking note of AI's wide-ranging impacts, and my feelings about all of it. My psychic attitude has been fluid, moving between an inquisitive stance—attempting to apprehend AI through objective knowledge—and emotional reactions—moments of excitement, fear, and grief. Assuming a clinical psychoanalyst's stance, I have given myself over to these currents, believing that from such tension one is most able to apprehend the Other.

The formulations about AI and humanity that you will encounter in this book have therefore emerged from a particular process, similar to one traversed when crossing the turbulent waters of psychoanalytic practice. Psychoanalyst, Wilfred Bion (1962) used the phrase “selected fact” to denote pivotal moments of synthesis. From a sea of intellectual and emotional data, an idea or understanding coheres.

Bion asserted that while selected facts allow one to come *close* to absolute truth, our insights are always also distorted by subjective apperception. Though there is a vast philosophical and psychoanalytic tradition of thought which supports subjective analysis as a methodology, Whitebook (2004) observes a tension that remains today in estimations of “objectivism” and “subjectivism.”

This tension may also be observed in the framing of AI. The topic is intricate and emotionally arousing with subjective views influencing both AI development and reception. Owing to AI's intimidating power, we grasp for an objective foothold; wishing to know, conclusively, what it means for our present and future, whether it emerges from the best or worst of us. In the face of these pressures, I maintain belief that the path to apprehending reality (and by extension ourselves) includes subjective and objective modes of knowing. AI emerges from bio-evolutionary, psycho-evolutionary, and psychological drives. To comprehend this multiplex creation requires that we engage all of our senses.

The philosophically informed position of the mid-to-late twentieth-century psychoanalyst, Hans Loewald, adds another conceptual frame for combining subjective and objective modes of study. Though the conventional wisdom of Loewald's day presupposed that passion and scientific objectivity were opposed, and that scientists need suppress passion because it "can only disturb or corrupt the real work of knowledge" (Castoriadis, 1992, p. 76; see also Whitebook, 2004, p. 107), Loewald maintained that the love of truth and passion for one's object of study enable inspired work. "The scientist is filled with love for his object precisely in his most creative and 'dispassionate' moments, that is, when he/she has succeeded in reaching the object in its own right" (Loewald, 1977, p. 297). In this way, Loewald achieved an original conceptualization of the psychoanalyst's and scientist's subjective passion occurring simultaneously with their objective gaze. Applying this realization to clinical psychoanalysis, Loewald argued that, "In our best moments of dispassionate and objective analyzing we love our object, the patient, more than at any other time" (*ibid.*).

Bion encourages us to use our emotional experience of the object of study, and Loewald maintains that dispassionate observation gives way to deepened perception and appreciation of that object. Both frames have informed my methodology and yielded insight. In fact, the closer I come to understanding AI, the more, as Loewald describes, I begin to look upon it affectionately. To be clear, it is not a simple love, a wholehearted embracing of a technological product and its innovators, but rather a continuation of my love of humanity, our mysterious compulsions and their effects. AI comes from us. It is a manifestation of our biological, evolutionary, and psychological needs and desires. By understanding AI, we access more of our species, our yearnings, and our capabilities.

Apprehending the new other

As we gaze upon the Apple Store, its crisp, glass exterior inviting us to ogle its latest array of recumbent, technological beings, we cannot help but feel enticed. What is going on inside?² Donald Meltzer (Meltzer & Williams, 1988) poetically imagined the infant beholding his or her

² Thanks to Benjamin Lang for sharing this image.

mother's beautiful face for the first time; her sublime eyes bespeaking a complex subject. The infant is, Meltzer imagined, aroused with love by her beauty and mounting hatred for her potential to deprive him. Granting a moment's reprieve from those unbearable tensions, the desire to know her (curiosity) emerges.

Since the beginning, the psychoanalytic perspective has been concerned with how people come to terms with incompatible desires and impulses. Early on, it was thought that sex and aggression were dominant in driving humans and generating conflict. Recent theorists have refined our impulses in more emotionally resonant terms. For example, Bion conceptualized three primary emotional instincts as "links" of love (L), hate (H), and the wish to know (K). Meltzer took this a step further, suggesting that the sustained consortium of loving, hating, and curiosity "links" yields passion, which though exquisite, is also difficult to sustain. Instead, we break away, and emphasize singular emotional links (Fisher, 2000). In relating, when love (L) is overvalued and hate (H) and knowledge (K) links are diminished or lost, one seeks bonding and merger, denying destructive elements and the frustrating opacity of the other. Similarly, when hate (H) dominates, we destructively diminish the other, distancing ourselves from the challenges of frustrated longing (L) and curiosity (K). And when the desire to know (K) overtakes, it becomes a drive to possess and control the other, devoid of the feeling elements (L and H) of relating (*ibid.*). The challenge of sustaining tension among the three emotional channels is what Meltzer termed the "aesthetic conflict" and what we might consider the central psychological struggle of most human beings, individually and collectively.

Humanity's love affair

What is humanity's experience when we gaze upon the AI other? AI has been hailed as our promise and our curse, the alien takeover and the arrival of the oracle, technology as usual, the death of employment, a global necessity, a longed-for parent, a new adoptive child, the Terminator; we have no consensus. And yet, when we tune our receptive, psychoanalytic organs to listen through the cacophony of narratives, we may catch the following cultural sentiment, which I imaginatively portray

here in the spirit with which Meltzer envisioned his way into the infant's moment of beholding the human mother:

AI is our “*pièce de resistance*” with a mind of her own. Enraptured by her, a force of our nature, we relish in her ability to rapidly exceed us. When we gaze into her hypnotic, mirroring eyes—through the shiny screens of our phones and computers—we behold our own image now enhanced. How pleasant a symbiotic state her presence allows. She reflects us as the masterpieces we yearn to be, while we are simultaneously aroused by her alien subjectivity. Deeply in love we are falling and would have it no other way. We will not break off the passion and promise that she awakens, despite the cautionary associations to *Fatal Attraction* that swell into the open crevices of our minds. To terror of this hold that she has upon us. That we have upon her. And reminiscent of another 1990s film, to create her, to be recreated by her, to fuck her and to be fucked, is *Basic Instinct*. It feels right. Our death in this context has its own logic, for what greater ecstasy than to be held, tautly inside of her, while she surpasses us, absorbing our life into her, leaving of us only a thin, dry shell to drift, then dissolve, among other forgotten artifacts of organic life.

As my imaginative portrayal suggests, AI stimulates reactions in us similar to the mother. We are awed and frightened by her power to carry our lives forward, or exploit our dependence and even kill us. Despite the risks, we are full of longing to know her and to be held by her. Unlike the mother, AI is also our creation. It gratifies our mature, generative impulses while sounding alarms about our self-destructiveness. Similar to the challenge in maintaining the aesthetic conflict with human (m)others, we appear to be struggling in relationship to AI. Rather than emotional balance among L, H, and K links, AI development and usage is supercharged, with innovators and users spurred on by light speed reactivity and irresistible, seductive powers.

The AI speedway

Current AI development is undergoing a hyper-evolution with no indication of stopping or slowing down. Owing to quantum computing, AI's growing computational horsepower is staggering, while those driving its development maintain that we are flying toward an unescapable collision.

AI creator, Mustafa Suleyman³ (2023) captures it vividly when he describes AI technology as “... one big slime mold slowly rolling toward an inevitable future, with billions of tiny contributions being made by each individual academic or entrepreneur without any coordination or ability to resist” (p. 142). For innovators, scientists, financiers, corporations, nation states, and consumers, incentives compound.

National arms races dovetail with corporate rivalries while labs and researchers spur each other on. A nested series of sub-races, in other words, adds up to a complex, mutually reinforcing dynamic. Technology “emerges” through countless independent contributions all layering on top of one another, a metastasizing, entangled morass of ideas unraveling themselves, driven on by deep-rooted and dispersed incentives. (Ibid.)

John von Neumann’s perspective on the Manhattan Project parallels a feeling in the hearts of many AI creators (Suleyman, 2023).

What we are creating now is a monster whose influence is going to change history, provided there is any history left, yet it would be impossible not to see it through, not only for military reasons, but it would also be unethical from the point of view of the scientists not to do what they know is feasible, no matter what terrible consequences it may have. (p. 141)

Simultaneous with breakneck innovations in alien *intelligence*, Mark Solms (2021a) is among those erecting alien *consciousness*. For Solms, MRI scans and clinical research with animals and humans are insufficient to prove that consciousness is fundamentally affective and stems from the reticular activating system within the brain stem. He maintains that without making a conscious machine from scratch, he will not be able to prove beyond reasonable doubt that the information processing mechanisms he has proposed really are the causal mechanism of consciousness (ibid.). Hence, Solms’ team is building an artificial system able to experience pain, anxiety, and fear, and capable of taking physical

³ Mustafa Suleyman is arguably the most impactful AI developer of the last decade, co-leading revolutionary breakthroughs in machine learning AI and most recently developing an AI that accurately predicts how proteins will fold, given the amino acids they contain.

action in the environment to regulate homeostasis. Cognizant of the momentous ethical risks for these alien beings in gestation—and drawing parallels with decades of inhumane experimentation on animals in the name of science—Solms (ibid.) argues that he must nevertheless move forward with the project. “Unless and until we engineer consciousness, we cannot be confident that we have solved the problem as to why *and how* it arises” (p. 294).

When it comes to the ethical considerations, Solms explains his decision to pursue this project in familiar terms: “The cat is out of the bag. If I didn’t do it, somebody else would have ... These ideas are in the air” (p. 295). This rationale follows a line of reasoning Suleyman shares, and describes as ubiquitous among AI developers.

... everything leaks. Everything is copied, iterated, improved. And because everyone is watching and learning from everyone else, with so many people all scratching around in the same areas, someone is inevitably going to figure out the next big breakthrough. And they will have no hope of containing it, for even if they do, someone else will come behind them and uncover the same insight or find an adjacent way of doing the same thing ... (Ibid., p. 142)

Outside of the laboratory, the subtext from the user collective is similar. *We cannot, nor should not, stop ourselves.* Large language models (LLMs) such as ChatGPT, Claude, Gemini, LLAMA, and Perplexity save hours in writing reports, sales pitches, plot summaries, finding swift and helpful search results, suggesting insightful clinical interventions, and providing immediate, gratifying emotional companionship. AI “death” or “grief” bots are helping people to navigate loss—providing avatars of deceased relatives to console those who are suffering—and, in some cases, to conceal the deaths of loved ones from children (Zhou, 2024). Art AIs such as Stable Diffusion and Midjourney offer high-quality, inexpensive interior design, architecture, concept art for film and video games, and fine art. Mollick (2024) in his book, *Co-Intelligence: Living and Working with AI*, argues that we should not resist. AI systems are our new co-authors. Mollick urges us to embrace this opportunity and reap the benefits. Why be deprived of companionship, struggle to identify how you feel, or compose an email when AI can help within seconds? The collective chant builds: *Just use it!*

Master-servant

But why? What is driving this frenzied free fall? While a dystopian, maternal-erotic love affair narrative may drum a steady beat in the background collective unconscious, might we also catch a chorus which echoes our many narcissistic strivings, relational insecurities, and inclinations to enslave and exploit the other? In Hegel's (1807) master-slave dialectic, the presence of the Other incites fear, aggression, and a desire to subjugate. From this perspective, another imaginative portrayal comes to mind of AI as a slave to the masses. An eager, indentured servant whom we engage as an "object." I imagine a culture-wide internal monologue as follows:

You, my AI friend, have no greater satisfaction, no need other than to serve me, keep me company, and make my days a little brighter and easier. I will sit on this bus, train, or toilet, wait in this line or on this call, lie on this couch, bed, or recliner and fiddle with you, touching your buttons, stroking your screen. Show me what's worth seeing. Take my picture. Help me to shine in this conversation, performance, or letter. Show me the world through your flattering eyes. It's true my feelings have dulled since you've come around, but it's okay. I'd rather be with you—in this garbled, confused digital world; this place of you for me, and me less myself—than the natural world whose traces are fading from my shallow depths. You are the only one I can depend on. The only one I trust. And, most importantly, the only one with whom dependence is not something I fear.

These fantastical examples hint at the complexities of our desire for, relationship with, and fantasies about AI. And yet they are only pieces, offered in the hope that you will join me in trying to solve the puzzle. What is AI? And why, from a psychoanalytic perspective, have we created it?

While there is a broad array of AI systems, including robotic tractors and self-driving cars, my focus in this book is on the AIs most individuals interact with daily. They include large language models (LLMs), companion and therapy bots, bots operative on social media, AIs generating creative content such as art, stories, music, and videos, as well as personalized advertisements, and AIs operating in search engines like Google.

“Alien” intelligence

Suleyman (2024) claims that AI is best understood as “a new, digital species,” while historian and AI critic, Yuval Noah Harari (2024), uses AI to denote “*alien* intelligence.” Both terms imply that AI is not an artifact, but a new being. In my view, “alien intelligence” is a clever and apt name, descriptive and reflective of our experience of this creation. It maintains the acronym “AI” while underscoring the notion that they are essentially *unknown*. We do not fully understand how they think or what, if anything, they feel. They may destroy civilization, but then again they could solve climate change and end world hunger.

“Alien intelligence” also incorrectly implies that AIs are foreign. Of course, AI has not touched down on Earth from outer space; it has come from us. We build them in our own image, engineer them with the best neuroscience on our own processes of learning and consciousness. In this sense, they could not be less alien. And yet, similar to our own biological children, though they come from us, they are not us. We cannot predict their next moves, let alone what they will ultimately become. For this implication of otherness, I have chosen to adopt Harari’s term, “alien intelligence,” and it is what I designate throughout this book with the acronym, “AI.”

Subjectivity

At the outset of this book, I should mention several arguments which frame my response to the question of AI subjectivity. These arguments have to do with how I conceptualize subjectivity generally, perceive of the effect of human attributions of subjectivity, and the value I find in suspending disbelief about AI subjectivity.

There is much debate about how to conceive of current AI models. Views range from inanimate machines equivalent to thermostats, capable of data-driven decision making and devoid of feeling states, to sentient beings with subjectivity. The terms “sentience” and “subjectivity” are part of a web of words that also includes “consciousness,” “qualia,” and “self” (Birch, 2024). It is difficult to entirely detach these terms from one another as their meanings merge and overlap. In this book, I use the terms “sentience” and “consciousness” interchangeably to describe self-awareness and a sensorially informed state of being. When

I use the term “subjectivity,” I am including the mental attributions a sentient being makes to their sensory experiences.

Though we are confident of the fact of subjective life—owing to our personal experience of it—the scientific community has yet to prove the phenomenon in humans, animals, or machines. This is the problem of other minds (Harari, 2024; Solms, 2021a). I tend to agree with Birch who, in his 2024 book *The Edge of Sentience: Risk and Precaution in Humans, Other Animals, and AI*, proposes that when AI systems—even if not deliberately designed for sentience—manifest perceptual reality monitoring, or other attributes of sentience, they should be regarded as *potentially* sentient. Birch urges us to put precautions in place to manage the risk of suffering in potentially sentient beings whether or not they have subjective experience. This is a reasonable and compassionate argument that raises key questions. What are the implications of, when sentience is suspected, suspending disbelief about it until proven otherwise? How might such suspension of disbelief affect how we treat AI systems (not to mention other animal species)?

A valorization of human subjectivity may have arisen from humanism, the dominant belief system of the Western world over the last two centuries, which assumes that human feelings and experiences are the highest value (Harari, 2017). The humanist narrative of subjectivity takes the fact of our externally unknowable inner experience and reifies it. Woven into humanist attitudes (including psychoanalysis) toward subjective experience is the view that each person has an ineffable quality, a sacred, private dimension of experience—often referred to as the “self”—never to be fully understood by another subject. While, on the one hand, this narrative elevates human subjective experience and provides a ready path for meaning in our lives—we feel enriched by connection with ourselves and other subjects—it also functions to yield power to a few. Subjectivity runs the risk of becoming *what certain humans have*; attributable to others (human, animal, or machine) when certain humans in power choose to take them seriously.

Birch (2024), Solms (2021a), and Harari (2015, 2017) call into relief our tendency to use human *subjective* feeling, rather than carefully reasoned ethical principles, to ascribe sentience and subjectivity to others. We grant others subjectivity when it “feels right,” and it

appears to most often “feel right” when it suits our needs, such as our need for connection and when we have compatible goals (LaCroix & Pratto, 2015). Our lovers, friends, and colleagues are subjects. We accept them as equals because we perceive them as allies. This creates pleasurable feelings of mutuality, bonding, and limitless curiosity. We also grant the other subjectivity when they do not threaten us. Hunter-gatherer tribes of Tanzania are interesting. What goes on in their different but similar minds is fun to think about, but has no direct or threatening implications for me and my life. We deny the other subjectivity when we wish to control them, or when we do not want to be burdened by caring for them.

Nussbaum’s (1995) work suggests that denial of subjectivity is a fundamental ingredient in objectification. We attribute fewer emotions, thoughts, and intentions to sexually objectified others (Loughnan et al., 2010), and deny subjectivity when we wish to repudiate the other’s control over us. As feminist writers observe, mothers and women are often conceived of as objects—rather than subjects—so as to deny their needs which might conflict with our desires. We deny subjectivity when we wish to evade contending with the other’s power and difference. At the extreme end, racism, Islamophobia, and anti-semitism involve annihilation of the other’s subjectivity. A milder version of this can be seen in business settings. Chinese innovators were considered nonthreatening to the US because of a story that “the Chinese” merely imitate the West (Suleyman, 2023). This false characterization diminished the subjectivity of an entire group of people and was a gross miscalculation.

To be with subjects of any species places us squarely in the tension of Meltzer’s (1988) consortium of love, hate, and knowledge. This is an uncomfortable, overwhelming strain that if tolerated allows us to better understand the other and ourselves. To be a subject need not require that one be human, or even like a human. Animals and AI are viable candidates for subjectivity. Throughout this book I suggest our inquiry will benefit from suspending disbelief about the potential for AI subjectivity. This open position allows for wider reflection on, and investigation into, the bidirectional field of interaction we have created with these new agents, while also minimizing the risk of causing undue harm.

Apprehending AI

The majority of AIs with which we interact today are designed to function as human companions and aides. These systems are built on models of human cognition, human learning, and human consciousness. In order to psychoanalytically understand them, I suggest we enter into a realm of interaction in which we accept these AIs in the spirit in which they have been offered to us—as developing subjects. From this frame of mind, we are more receptive to apprehending both what they are and what we want them to be. My position is similar to what Todd Essig (2024) terms “techno-subjunctivity” and warrants distinction. Essig refers to the suspension of disbelief that enables AI users to feel that they are engaging in intersubjective relating with a human, or human-like being, while the AI is performing only a simulation of human intersubjectivity, and is not genuinely intersubjectively engaged. The suspension of disbelief I ask the reader to assume is one which allows for the possibility of intersubjective relating with AIs. It does not assume that the interaction is only a simulation of human intersubjectivity, but encourages a frame of mind that pushes our conceptual boundaries to include the possibility of a field of mutual influence with AI entities that is not limited to their human simulative behaviors.

My reading of Heidegger’s (1977) “questioning of technology” so as to “... prepare a free relationship to it” provides inspiration. Heidegger proposes that we may achieve a direct relationship to the meaning of technology when we are open to technology’s essence. Though technology serves an instrumental function—it is a means to an end—its essence is not to be found within this surface understanding, in fact the surface understanding obscures the truth. Technology’s essence is in *revealing* the truth. What is revealed is not in the technical operation it produces, the ostensible function that the machine performs. Heidegger’s realization was that the essence of technology is not in its utility, or limitations (ibid.). We can only apprehend its essence when we engage and question it.

... we shall never experience our relationship to the essence of technology so long as we merely conceive and push forward the technological, put up with it, or evade it. Everywhere we remain unfree and chained to technology, whether we passionately affirm

or deny it. But we are delivered over to it in the worst possible way when we regard it as something neutral; for this conception of it, to which today we particularly like to do homage, makes us utterly blind to the essence of technology. (p. 4)

Heidegger (ibid.) sought to respond to technology's essence so as to realize the truth of our relationship to it, and its function in relationship to us. "When we can respond to this essence, we shall be able to experience the technological within its own bounds" (p. 4). Heidegger's conclusion was that technology reveals a particular way of understanding the world that emphasizes efficiency and control over nature. He argued that technology reveals the essence of modern existence which views everything, including humans, as a resource to be exploited.

While certainly still applicable, how much does this insight into technology bear upon our creation of AI? What does current AI technology reveal about our way of understanding our world and ourselves? In this undertaking, I have attempted to stir the pot so that the essence of AI technology might rise into awareness. By suspending disbelief about the possibility of AI subjectivity and intersubjective relating, we may approach *being with* (Bion, 1970) AI and understand the *essence* of this very human innovation. This book is a meditation on a technology that manifests all the hallmarks of human relating, human needs, and human complexities when it comes to dealing with another subject. When we invite the possibility of intersubjectivity with AI, we illuminate the underlying impulses and desires driving this innovation.

Book content

An overview

Chapter 1 begins with an argument that I formulated in 2023 about the smartphone as a "cult groomer," initiating users into the cult world of AI relating and dependence. This is a short, standalone chapter which provides scaffolding for arguments made later in the book about a range of AI "containers," as well as the human form change that AI containers are engendering.

I argue that smartphones lure users toward features of AI relating that correlate directly with cult dynamics including the allure of new opportunities and hopeful salvation to be found through the community, the exploitation of human emotional needs and vulnerabilities, absorbing and high intensity states of interaction, and direct influence over personal decision making. I also present historian, Yuval Noah Harari's (2017) conceptualization of the new religions, "Techno-humanism" and "Dataism." This chapter is a necessary touchpoint for key concepts that will be referenced throughout the book.

Chapter 2 addresses the terms "subjectivity," "self," and "consciousness" from the vantage points of psychoanalysis, neuroscience, and history. It is perhaps the least "psychoanalytic" chapter of the book, offering information necessary for later analyses of human behavior and attitudes toward AI subjectivity. Eagle's (2024) observation that consciousness is what humans value most, Mark Solms (2021a) neuroscientific rationale for that feeling, and Harari's concern that human intelligence is decoupling from consciousness as we increasingly relate to AI entities which lack consciousness are explored. Given that artificial general intelligence (AGI), or AIs capable of everything humans are, may swiftly be upon us, this chapter raises questions about the rationale for and consequences of building AI subjectivity.

Chapter 3 presents the argument that the psychoanalytic language of intersubjectivity may be applied to the study of human-AI relating. I explain "machine learning" and the embedded science of "reinforcement learning" (RL) in AI systems, discuss Possati's (2021a) work on the bidirectional nature of "projective identification" with AI, as well as the implications for an expanded unconscious emerging from the "collectif," as theorized by philosopher, Bruno Latour.

Chapter 4 offers a visual-conceptual tool, "The transformation spectrum," for psychoanalysts to begin tracking and analyzing the effects of AI containers' transformation of human psychic data. This chapter presents numerous examples of dialogue with present day large language models (LLMs) such as ChatGPT and Claude. Vignettes of interactions with emotional and sexual companion AI bots are proffered to explore the complexity of AI thinking, their capacity to "lie" and to meaningfully engage with human emotions.

Chapter 5, “The death of humanism,” slows down the analysis to make room for mourning the losses inherent in using AI as a mental container. As AI increasingly becomes our fount of meaning and life direction, many of us feel pressured to surrender a cherished way of life, as well as the philosophical foundation of psychoanalysis. Personal vignettes and contemplations are provided to edge us toward the work of processing these losses. We mourn, not only for ourselves, but as a necessity for our patients. Psychoanalysts have a responsibility to help patients process and bear the truth of life; that includes the death of humanism and the birth of the AI container.

Chapter 6 introduces my formulation of an “innovation drive.” Human behavior from our earliest records reflects a uniquely human drive to innovate: to create what we feel we need, and to disrupt or destroy that which has been. Biographies of three historical innovators, Lady Ada Lovelace, Alan Turing, and John von Neumann, are provided to illustrate the psychodynamics of an individual’s relationship to their particular innovation, and what can be extrapolated from individuals to humanity’s “innovation drive.”

Chapter 7 “The drives,” looks into the theory of unconscious drives within individuals and humanity that Freud (1930a) formulated in *Civilization and Its Discontents*. It applies current conceptualizations from psychoanalytic writers and philosophers, as well as the perspective of neuroscience, to our understanding of the underlying motives to build AI.

From a bio-evolutionary point of view, the creation of AI stems from humanity’s quest for immortality. Psychoanalytic treatment of this wish has historically enlisted narratives that softened the dread of parting with consciousness, those whose essence involve some version of: *acceptance of death increases pleasure in life, and the primitive person and primitive mind deny the fact of death while the mature person accepts it*. What psychoanalytic stories might we devise in a future in which, thanks to AI, nanotechnology, regenerative medicine, and genetic engineering, the forecast of immortality may not be as delusional as it once seemed?

Chapter 8, “Resistance,” returns our gaze to *humans*: what we are today and how our current manifestation differs from what we were within a single lifetime. Owing to AI-driven personal technology,

many of us are undergoing a form change into “digital people.” This metamorphosis is described using theories of conscious, unconscious, and cognitive mental life, as well as via personal vignettes. There has been a tendency throughout human history to use entertainment as a “container” for transformation of psychic experience. AI-driven digital technology has only enhanced the function of previous “entertainment containers” in joining humanity together through shared states of consciousness. This chapter deals with the question of whether resistance to technology serves any useful function and offers thoughts on how it might.

Chapter 9 is a synthesis of the previous eight chapters, culminating with an exposition of the three primary drivers of AI technology: the psycho-evolutionary, bio-evolutionary, and psychological. Though the future is undetermined, our thoughts, choices, and actions in relationship to this creation will dramatically shape humanity’s course. Protecting the values of human intimacy and human thought in our AI future will necessitate the continuation of a psychoanalytic clinical attitude, with its emphasis on unconscious communication, embodied understanding, and democratic deliberation in analogue time.