

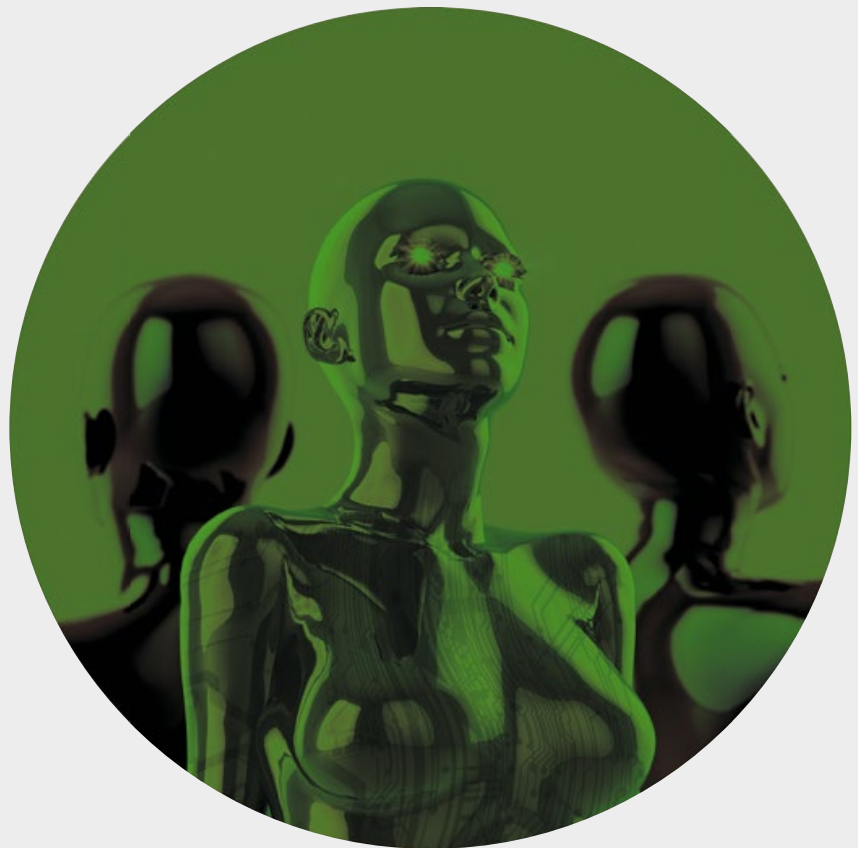


THE PALGRAVE LACAN SERIES

SERIES EDITORS: CALUM NEILL · DEREK HOOK

The Psychoanalysis of Artificial Intelligence

ISABEL MILLAR



palgrave
macmillan

The Palgrave Lacan Series

Series Editors

Calum Neill
Edinburgh Napier University
Edinburgh, UK

Derek Hook
Duquesne University
Pittsburgh, USA

Jacques Lacan is one of the most important and influential thinkers of the 20th century. The reach of this influence continues to grow as we settle into the 21st century, the resonance of Lacan's thought arguably only beginning now to be properly felt, both in terms of its application to clinical matters and in its application to a range of human activities and interests. The Palgrave Lacan Series is a book series for the best new writing in the Lacanian field, giving voice to the leading writers of a new generation of Lacanian thought. The series will comprise original monographs and thematic, multi-authored collections. The books in the series will explore aspects of Lacan's theory from new perspectives and with original insights. There will be books focused on particular areas of or issues in clinical work. There will be books focused on applying Lacanian theory to areas and issues beyond the clinic, to matters of society, politics, the arts and culture. Each book, whatever its particular concern, will work to expand our understanding of Lacan's theory and its value in the 21st century.

More information about this series at
<http://www.palgrave.com/gp/series/15116>

Isabel Millar

The Psychoanalysis of Artificial Intelligence

palgrave
macmillan

Isabel Millar
Centre for Critical Thought
University of Kent
Canterbury, UK

The Palgrave Lacan Series
ISBN 978-3-030-67980-4 ISBN 978-3-030-67981-1 (eBook)
<https://doi.org/10.1007/978-3-030-67981-1>

© The Editor(s) (if applicable) and The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cover illustration: VICTOR HABBICK VISIONS/SCIENCE PHOTO LIBRARY/gettyimages

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG. The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

This book is dedicated to my mum, Sylvia.

Prologue: Roko's Basilisk

In 2010 on *LessWrong* forum, a user named Roko posited a thought experiment. He proposed that in a hypothetical future an all-powerful super-intelligent AI could retrospectively punish anyone who in the present time did not do everything in their power to aid in the creation of such a superintelligence. By merely entertaining the idea of such a being and not facilitating its development you would expose yourself to the possibility that it would deduce that you had not acted in accordance with the duty to bring it into existence (the moralistic tone of the experiment is enforced by the fact that the AI is paradoxically a benevolent one whose task is to protect humankind, and therefore those who don't facilitate its existence deserve ill against their fellow men). The vengeful Abrahamic nature of the Basilisk meant that in future, it could recreate a simulation of you to torture for all eternity for the sin of putting him at existential risk. The Old Testament stylings of the Basilisk are clear: *he's nice, but only if you deserve it.*

As absurd as the tale sounds, it was met with outrage by the site's founder and director of the *Machine Intelligence Research Institute* (MIRI) in California, Eliezer Yudkowsky. Yudkowsky felt that Roko had opened a Pandora's box of previously unimaginable torment that the poor readers of his blog would now fall victim to. In response to Roko's post he reportedly said:

Listen to me very closely, you idiot.

YOU DO NOT THINK IN SUFFICIENT DETAIL ABOUT SUPERINTELLIGENCES CONSIDERING WHETHER OR NOT TO BLACKMAIL YOU. THAT IS THE ONLY POSSIBLE THING WHICH GIVES THEM A MOTIVE TO FOLLOW THROUGH ON THE BLACKMAIL.

You have to be really clever to come up with a genuinely dangerous thought. I am disheartened that people can be clever enough to do that and not clever enough to do the obvious thing and KEEP THEIR IDIOT MOUTHS SHUT about it, because it is much more important to sound intelligent when talking to your friends.

This post was STUPID (ibid.).¹

The post was subsequently removed, and all talk of the Basilisk was banned from the website for over five years. But the Basilisk had already wreaked havoc among the forum's readers many of whom had started to experience psychological difficulties. Paranoiac fears of the Basilisk's future existence have now become something between an urban legend and a genuine topic of philosophical debate, not to mention the fact that it is taken seriously by some of the major tech entrepreneurs and scientists currently driving AI research. The logic behind the Basilisk is even (spuriously) backed up by Timeless Decision Theory and Bayesian probability.

In fact, Yudkowsky (2010) has written at length on the theory underpinning the problem of the Basilisk, even drawing on the prisoner's dilemma which we will recall Lacan (2006a) uses in his discussion of logical time. The prisoner's dilemma was a thought experiment in game theory, where the actions of several prisoners were dependant on the anticipated decisions of one another in order for them to secure their freedom. The dilemma exemplified for Lacan the tripartite structure of time surreptitiously at work in the concept of so-called rational thought. These he called the instant of seeing, the time for understanding, and the

¹ See David Auerbach (2014).

moment of concluding. Accordingly, whilst logical time is not objective, this does not mean that it cannot be formulated according to a rigorous structure; that of intersubjective logic based on a dialectical relation between hesitation and urgency. A logic we see at work in Roko's auto-poietic Basilisk and what could be called in other terms, *hyperstition*.

The term hyperstition was coined by Warwick University's *Cybernetic Cultural Research Unit* (CCRU) and continues to be one of the major concepts of the Accelerationist movement. A portmanteau of 'hyper' and 'superstition', drawing on the Baudrillardian logic of hyperreality, hyperstition to paraphrase Nick Srnicek and Alex Williams (2014), the authors of the *#Accelerate Manifesto for Accelerationist Politics*, refers to narratives capable of bringing themselves into reality through the workings of feedback loops, which generate new socio-political attractors. Roko's Basilisk allegedly functions according to just this sort of hyperstitious logic. As a computational form of Pascal's wager, it relies on a number of premises for it to function. Firstly, the proviso that the concept of a Singularitarian superintelligence entails the capacity for absolute and total recall of all data and secondly the ability to simulate every historically living being in order to then torture them. Thirdly, the belief that a simulation is equivalent to a subject. As Ana Teixeira Pinto (2018) has noted however, the theological and paranoid overtones of the Basilisk function as:

the personification of AI as Oedipal beast [...] and of code as the male seed. Those who seek mathematical proof of the prediction's likelihood are missing the point. The content of Roko's thought experiment is symbolic, not scientific: it speaks through cipher and allegory. (p. 19)

Teixeira Pinto highlights the Oedipal logic at work in the positing of the Basilisk, but to this we might add that the phallic enjoyment involved in the imagining of the ultimate mathematizable *One* that admits of no exemptions is masculine logic par excellence. The Basilisk also functions as the ultimate indicator of anxiety, the impossible object as cause of desire and also complete destruction. The poor human on this score is trapped between the finite slab of meat that tortures him and the infinite simulation that he will inevitably become. What seems to be at stake in

this speculation on the Singularity is what Lacan (2006b) in *Function and Field* referred to as the future anterior:

What is realized in my history is neither the past definite as what was, since it is no more, nor even the perfect as what has been in what I am, but the future anterior as what I will have been, given what I in the process of becoming. (p. 247)

In the logic of Roko's Basilisk, we may apprehend the möbius structure of the relationship between AI and psychoanalysis that this book will attempt to depict, a topology which for as far as one travels along, will always lead inevitably to its inverse: its *extimate* core.

References

- Auerbach, D (2014) The Most Terrifying Thought Experiment of All Time. Available (01.03.2020) at: <https://slate.com/technology/2014/07/rokos-basilisk-the-most-terrifying-thought-experimentof-all-time.html>
- Lacan, J. (2006a) 'Logical Time and the Assertion of Anticipated Certainty' in J. Lacan, *Écrits*, pp. 161–175. London: W.W. Norton & Company.
- Lacan, J. (2006b) 'The Function and Field of Speech and Language in Psychoanalysis' in *Écrits*, pp. 197–268. London: W.W. Norton & Company.
- Srnicek, N., & Williams, A. (2014) '#Accelerate Manifesto for an Accelerationist Politics' in N. Srnicek & A. Williams (eds.) *#Accelerate: The Accelerationist Reader*, pp. 347–362. Falmouth: Urbanomic.
- Teixera Pinto, A. (2018) The Psychology of Paranoid Irony. *Transmediale Journal* 1: pp. 18–22.
- Yudkowsky, E. (2010) *Timeless Decision Theory*. *The Machine Intelligence Research Institute*. Available (01.03.20) at: <https://intelligence.org/files/TDT.pdf>

Praise for *The Psychoanalysis of Artificial Intelligence*

“Does It think? Does It enjoy? Taking the problem of artificial intelligence as a problem that has been in a way always-already inherent to psychoanalytic inquiry Isabel Millar accomplishes a most powerful and productive shift of perspective on both psychoanalysis and AI. Her work takes us on a fascinating journey across a vivid conceptual and figural landscape, and provides an excellent proof that powerful, captivating theory is all about asking the right kind of questions. *The Psychoanalysis of Artificial Intelligence* is both extremely timely and timeless in the way it constructs and tackles its object.”

—Professor Alenka Zupančič, *The European Graduate School and Slovenian Academy of Sciences and Arts*. Author of *Ethics of The Real: Kant and Lacan* and *What is Sex?*

“Boldly drawing on a vast range of academic disciplines orchestrated by an enviable psychoanalytic erudition and an original treatment of so-called “sexbots” as a central object for contemporary speculative and social investigation, Millar’s book asks a series of seminal and long-overdue questions, which are here to stay. How should we approach the allegedly forthcoming advent of the “singularity” in terms of sexuality and sexuation? Does sexual reproduction have a future? What new forms of enjoyment, if any, might Artificial Intelligence enable us to think and experience? Or is it rather the case that androids secretly already have wet dreams about the human-all-too-human absence of the sexual relationship?”

—Dr Lorenzo Chiesa, *Newcastle University, UK*. Author of *Subjectivity and Otherness* and *The Not-Two*

“People tend to respond to artificial intelligence with either fear or love. Isabel Millar proposes a third way: to psychoanalyze artificial intelligence and the persistent investment in it. In a stunning work of expansive intellectual power, Millar shifts the fundamental question concerning artificial intelligence to the terrain of enjoyment. After Millar’s book, the question “Does it enjoy?” should be the starting point for any engagement with artificial intelligence. It is simply an epochal book for understanding this engagement.”

—Professor Todd McGowan, *Department of English, University of Vermont, USA*. Author of *The Real Gaze and Emancipation After Hegel*

Contents

1	Introduction	1
	Part I	13
2	The Stupidity of Intelligence	15
3	The Artificial Object	49
4	The Sexual Abyss	85
	Part II	123
5	What Can I Know? Artificial Enjoyment	125
6	What Should I Do? Patipolitics: From Sade to Killian	147

7 What Can I Hope For? Reproduction, Replication, Immortality	169
8 Conclusion: What Is Man? Between Matheme and Anxiety	193
 Bibliography	 205
 Index	 217

List of Figures

Fig. 3.1	Object a and the lathouse	64
Fig. 3.2	Table of partial drives	78
Fig. 4.1	The graph of sexuation	95
Fig. 4.2	The four discourses	104
Fig. 4.3	The quaternary structure	105
Fig. 4.4	What is a Sexbot?	117
Fig. 8.1	What is man?	203



1

Introduction

The percentage of intelligence that is not human is increasing. And eventually, we will represent a very small percentage of intelligence.

—Elon Musk (2018, online)

The Psychoanalysis of Artificial Intelligence, what a strange proposition. What could it possibly mean? The significance of the two terms in themselves is hardly self-evident, let alone their relationship to one another. Psychoanalysis on the one hand; simultaneously a clinical practice, a mode of cultural critique and a philosophical battle ground. And Artificial Intelligence, a technoscientific ‘invention’ originating in the 1950s¹ yet with literary, cultural and fantasmatic origins that date back centuries, and a concept whose theoretical potential continues to provoke intense philosophical debate. In this book, I argue that Artificial Intelligence (AI) and the creation of the artificial brain, which promises to separate neuroscience from biology and thought from the body, along with the

¹The earliest coinage of the term Artificial Intelligence is attributed to computer and cognitive scientist John McCarthy at a 1956 workshop at Dartmouth College. Other attendees at the workshop, who would soon become founders and leaders in the early field of AI research, were Allen Newell (CMU), Herbert Simon (CMU), Marvin Minsky (MIT) and Arthur Samuel (IBM).

prospect of forms of embodied AI which aim to simulate and surpass human intelligence, provokes an urgent engagement with the psychoanalytic subject. Simultaneously the book considers psychoanalysis as a crucial tool in our understanding of what AI means for us as speaking, sexed subjects. In short, AI and psychoanalysis stand in *extimate* relation to one another.

Through the reconceptualization of Intelligence, the Artificial Object and the Sexual Abyss we conjure a figure who exists on the boundary of psychoanalysis and AI, straddling our fantasy worlds and our speculations about the possibilities for life alongside or through Artificial Intelligence; the Sexbot. With its help, and through the medium of film we subvert Kant's three famous enlightenment questions, *What Can I Know, What Should I Do and What May I Hope For*. Ultimately, we transition from the question *can it think* to *does it enjoy?*

Owing to its inherent conceptual interdisciplinarity it is no wonder that AI and the discourses surrounding it seem to have unique capacity to blur the boundary between science and fiction. Embedded in a rich history of fantasy and pop-science, elements of which have been the subject of philosophical reflection since antiquity, appearing in various guises throughout the history of Western thought and literature,² it is often difficult to discern where the science of AI starts and fiction ends. Today there is no unifying theory which guides Artificial Intelligence research given that it draws from a variety of fields including, computer science, information theory, mathematics, neurobiology, psychology, linguistics, logic and analytic philosophy. Its potential and scope are in constant debate both scientifically and conceptually, being a polemical topic for cultural theory, political thought, ethics, philosophy, and even cosmology. Considering the rapid advances made regarding the reverse-engineering of the human brain in the field of neural networks and deep learning and the adjacent fields of quantum computing, nano and bio-technology, some, like futurist Ray Kurzweil (2014) anticipate that we will soon transcend the "limits of nature", thereby reaching a synthesis

²We may recall for example Ovid's *Pygmalion*, Descartes (alleged) robotic daughter Francine (Kang 2017), Maelzel's chess playing automaton and Čapek's (2004) *Rossum's Universal Robots*, to name but a few instances.

of science and fiction in the ‘Singularity’. Others argue we are about to enter a “Fourth Industrial Revolution”: an era heralding the gradual fusion of digital, physical, and biological worlds (Schwab 2016). For many philosophers and theorists of AI this so-called *Life 3.0* (Tegmark 2017) where science-fiction becomes terrifying reality is a conceptual terrain, which raises complex questions about the notion of intelligent life, the nature of thinking, the future of the social bond and the constitution of the “human”. In *Superintelligence, Paths, Dangers and Strategies*, Nick Bostrom (2014), foresees a dark future for humanity if we ignore his warnings about the possibility of a Hal 9000-like artificial Superintelligence, by which he means any intelligence that vastly exceeds the performance of humans. He believes that the creation of a super intelligent being could lead to the extinction of humankind. The risk involved in the creation of Superintelligence is that it would be operating on a speed and scale unfathomable to humans, which could initiate an intelligence explosion on a digital time scale of millisecond speed so powerful as to accidentally (or deliberately) destroy humanity. Bostrom not only contemplates the possibility of malicious applications of AI, such as hacked military devices, nano-factories distributed in undetectable concentrations creating killing devices on command and even payed human ‘dupes’ doing AI’s dirty work but envisions a scenario in which, once AI achieves a stage of world domination, humans would be useful only as raw materials. As he puts it: ‘brains, if they contain information relevant to the AI’s goals, could be disassembled and scanned, and the extracted data transferred to some more efficient and secure storage format’ (Bostrom 2014, p. 118). In order to prevent the emergence of such rogue Superintelligences, Bostrom joined Stephen Hawking in 2015 to sign an open letter on behalf of *The Future of Life Institute*, warning of the possible threats of AI. The signatories all subscribed to twentythree principles to ensure the safe development of Artificial Intelligence. As Max Tegmark (2017) enumerates, however, there are many misconceptions and disagreements about the future of AI. These include questions of when, how and what form AI will take and how long the process of its evolution will be. Furthermore, the possibility of so-called Superintelligence is still highly contested. This, however, has not prevented some from speculating about the possible date of its arrival. The Singularity designates just

this hypothetical moment of an intelligence explosion, a point-of-no-return, at which AI will decisively surpass human intelligence, rendering the human species as we know it obsolete, if not actually extinct.³ Its foremost advocate, Kurzweil, expects the Singularity will occur in two phases. By 2029 AI will supposedly reach the stage of human level or ‘General’ Artificial Intelligence and successfully pass a Turing Test and in 2045 humankind will multiply its effective intelligence by a billion-fold through merging with AI. The potentially paradigm shifting consequences of the hypothetical emergence of general or super-intelligent AI has even become a topic for cosmology. Veteran scientist and inventor of the Gaia Hypothesis James Lovelock recently published *Novacene* (2019) in which he proposes that the age of the Anthropocene (the geological period in which humans acquired planetary scale technology) has already come to an end and we are entering a new age, the ‘Novacene’ in which technology will come to inherit the ‘consciousness’ of the cosmos. In his vision, artificially intelligent beings who can think 10,000 times faster than humans will emerge as the inheritors of the earth and caretakers of the intelligent universe. For Lovelock, the hypothesis of the emergence of such intelligent beings makes it even more vital that we retain the environmental conditions conducive to their survival. Thus, as Yuval Harari (2017) observes, the central hallmark of debates on the future of AI is the hubristic question: ‘Who are the new “Gods”—humans or AI?’

This concern with the technological Singularity as some sort of onto-theological watershed moment is taken up by Žižek (2020) who remarks that, what the advocates of the Singularity often fail to realise, or at least fully engage with, is that in this passage from human to post-human, what disappears is precisely self-awareness, which is rooted in ‘finitude and failure’ (p. 75). Regarding the apparent paradox which emerges as a result of our popular visions of post-human Singularity, Žižek goes on to state that:

³ The term was popularised by science fiction writer Victor Vinge in 1983 and brought into wider circulation by his (1993) article ‘*The Coming Technological Singularity*’. According to David Chalmers (2010), however, the term Singularity is used in a variety of ways to refer to different scenarios; the loose sense refers generally to the unpredictable consequences of exponential growth in AI, while the Singularity in the strict sense refers to a point where ‘speed and intelligence go to infinity’ (p. 3).

Insofar as posthumanity is, from our finite/mortal human standpoint, in some sense the point of the Absolute towards which we strive, the zero-point at which the gap between thinking and acting disappears, the point at which I became *homo deus*, we encounter here again the paradox of our brush with the Absolute: the Absolute persists as the virtual point of perfection in our finitude, as that X we always fail to reach, but when we get over the limitation of our finitude we also lose the Absolute itself. Something new will emerge, but it will not be creative spirituality relieved of mortality and sexuality—in this passage to the new we will definitely lose both. (p. 158)

Whilst Žižek's diagnosis of the problem with discourses on the Singularity is apposite, here my concern will not be to repeat the same gesture but rather to seek a constructive and productive way to engage with our relationship to AI psychoanalytically. While I will not attempt to give an account of the historical development of (or philosophy of) Artificial Intelligence, I will delineate a general working definition of Artificial Intelligence as: *a non-human mode of thought, whether embodied or disembodied, which acts autonomously and whose motives and purpose we may not necessarily be aware of, nor even understand*. Some might say that conveniently this definition could also be applied to the psychoanalytic conception of the unconscious, an ambivalence that lies at the heart of this book. Recall in *Seminar II* Lacan's (1988) reproach to Octave Mannoni for his worries over the human becoming too much like a machine:

Don't be soft. Don't go and say that the machine is really rather nasty and that it clutters up our lives. That is not what is at stake. The machine is simply the succession of little 0s and 1s, so that the question as to whether it is human or not is obviously entirely settled—it isn't. Except, there's also the question of knowing whether the human, in the sense in which you understand it, is as human as all that. (p. 319)

Between 1985 and 1986, at the psychoanalysis department at the *Université Paris 8*, Jacques-Alain Miller gave his course on *Extimité* in which he characterized the logic of the Lacanian unconscious as an extroverted interiority. 'Extimacy', a portmanteau of exterior and intimate, is a word first coined by Lacan (1992) in *The Ethics of Psychoanalysis*.

Although Lacan did not explicitly return to the concept in any of his seminars, the logic of extimacy, following Miller (1988), can be said to underpin the Lacanian organon in general as a concern with the intimate exteriorization that belies the nature of subjectivity, most clearly articulated in Lacan's relentless concern with the topological coordinates of the Möbius strip, the Klein bottle and knot theory. However, not only is the unconscious qua 'discourse of the Other' (1988, p. 89) to be understood in terms of an extroverted interiorization that morphs the notion of "unconscious depth" into a question of topological space, but as this book attempts to illustrate, the very materiality of the *speaking body* in its relation to Artificial Intelligence should be understood as extimate.

In a civilization in which Artificial Intelligence is becoming a significant element in the social bond, the psychoanalysis of AI is a provocation. It asks us to question both the meaning of psychoanalysis when taken outside of the purview of the strictly 'human' clinical space and conversely it attempts to show in what ways psychoanalysis is already an extimate part of artificial intelligence. Similarly, it speculates on what form our philosophical and critical thinking about AI has hitherto neglected the essential element or indeed material of psychoanalysis, that is to say, enjoyment. This leads us to proffer the hypothesis that the 'vanishing mediator' between our two unlikely bedfellows is none other than sex. For psychoanalysis and its clinical treatment of 'suffering', sex is the crucial problem underlying all others. But more than a symptomatic 'problem' sex is a philosophical problem. Philosophical in the sense that it has, by definition, no solution. For psychoanalysis sex names the impossible yet inevitable collision of epistemological and ontological questions that characterize the entrance into subjectivity for all speaking beings. So, we must ask, what is sex for Artificial Intelligence? Judging by most of the literature and popular discourse surrounding it, sex is nothing more than an apparently superficial anthropomorphization of our fantasies of AI. But isn't this precisely the point? This fantasy of AI sex obscures the fact that sex is *only ever* a fantasy covering up for a hole in reality itself, or in Baudrillardian terms a question of dissimulation as a strategy of simulation. It is an absence which, as this book hopes to illustrate, brings with it a deafening silence which is impossible to ignore. The 'sex' of Artificial Intelligence resides everywhere, it is what brings it into being. In Lacanian

terms we could qualify this further to say that AI in its many forms both actual and fantasmatic ex-ists as a form of relation to the signifier or more specifically, a mode of enjoyment. Through the employment of both the philosophical engagements and the clinical and conceptual developments of Lacanian theory, the book aims to develop a novel and productive encounter between psychoanalysis and AI. In proposing to approach Artificial Intelligence psychoanalytically the book views the sexual non-rapport as its theoretical kernel. I seek firstly to advance a psychoanalytic reading and problematization of AI as a discourse about ‘knowledge in the real’. Secondly, to develop a novel, conceptual grid to query the material implications of Artificial Intelligence for subjectivity, the body and the social bond. In this sense, this project is not concerned with simply providing a psychoanalytic elucidation of our unconscious fears, fantasies or fascination with AI. Rather, it seeks to take the real dimensions of AI seriously. In short, this means the passage from a concern with the barred-subject and *object a* to a concern with the speaking body and the *artificial object*; one which Lacan in *Seminar XVII* gave the provisional name *lathouse*. The lathouse is an under-theorized and underutilised Lacanian concept, which presents us with a new way of understanding our bodily and structural relationship to AI.

So how *does* one read the sentence which forms the title of this book? Are we planning to psychoanalyze AI? If so, what would that mean? Or are we inquiring after the possibility of AI to be the psychoanalyst? This begs questions of how we are to conceptualize of AI as a ‘thinking thing’. The first ambiguity we should draw attention to however is the fact that psychoanalysis strictly speaking only ever happens as the result of a demand, a subjective and singular demand on the part of the analysand. And this demand is met with the desire of the analyst, for whom the demand of the analysand, is an *object a*. Both these essential elements give rise to a transference relation resulting in what could be characterized as psychoanalysis proper. The wager of this book is that paradoxically in order to understand the stakes of Artificial Intelligence it is not to post-humanism or transhumanism that we should turn but rather to the subversive spirit and (anti-humanism) of Lacanian psychoanalysis, taking the ‘demand’ of AI as our *object a*.

In 1973 Jacques-Alain Miller interviewed Jacques Lacan for a French television broadcast in which he challenged the renegade psychoanalyst about the nature and value of his psychoanalytic theory and practice. Lacan's responses were typically elliptical, but nonetheless provide the careful reader with an encrypted summary of his work to date and the place of Lacanian psychoanalysis in the contemporary world. Interestingly, Miller's interview concludes with his posing of the three Kantian questions to Lacan: '*What can I know?*', '*What ought I to do?*' and '*What may I hope for?*' Lacan offers Miller short shrift in response owing to what, in his view, is the difference in the role of the psychoanalyst as opposed to the philosopher. Perhaps the key to his reply can be found several pages earlier where he refers to the function of the Saint as corresponding to the '*trashitas*' of society (1990, p. 15); a position which, he says, must be taken up by the psychoanalyst as the '*refuse of jouissance*' (p. 16). It is not, in Lacan's view, for the analyst to ask the Kantian questions, but rather to allow the subject to realise his position with respect to them. The fourth Kantian question '*What is man?*' was never broached in this interview, but one could argue that it constitutes the underlying thread that runs through the whole of the psychoanalytic edifice.

I will therefore revisit the three Kantian questions which Miller challenged Lacan to address in the 1970s in the new context of Artificial Intelligence and via the prism of sexual non-rapport. The Kantian questions, which defined the Enlightenment project, will be employed to examine and problematize the relationship between psychoanalysis and AI. The three questions are typically present in all popular discourse and critical speculations on the future of AI. The first usually with reference to the question of consciousness and the perennial problem of "other minds". This is articulated in concerns with the sentience of Artificial Intelligence, perhaps most famously exemplified by the Turing Test as the ultimate "measurement of consciousness". The second Kantian question characteristically revolves around the ethics of AI; to what extent do we allow various forms of AI to enter into the social bond and how do we prevent its worst excesses or impacts on us as subjects? The third Kantian question is centred on the notion of the Singularity. Will we need to contemplate a future living with other forms of intelligence? Or will the advent of Superintelligence signal the end of humanity and thus the

extinction of the species as we know it? While the book poses the Kantian questions à la Miller, it refuses to answer them, à la Lacan. Instead of the standard approach taken by most philosophers or critical theorists on the problems of AI, I will look rather for the other side (*l'envers*) of the questions.

So, what forms of AI will this book be concerned with? AI is as huge and complex an object of scrutiny as psychoanalysis, and this project can by no means cover the entirety of either of those domains. My more modest task is to clarify a manner in which the two realms would find each other's extimate kernel residing inside themselves. In order to do this, I have conjured a conceptual figure who exists on the boundary of psychoanalysis and AI. To this end, the first part of the book is concerned with providing the theoretical groundwork for the conceptualisation of the Sexbot via a psychoanalytic examination of the concept of intelligence, the artificial object and the abyssal nature of sex.

Once I have drawn up this figure, I turn to the speculative work of the book in the form of the three Kantian questions. I mobilise the Sexbot as a *figure* to articulate the ontological, epistemological and technological series of problems that underly the entrance of AI into the social bond. The figure of the Sexbot, as represented in its ideal form in film, is to be understood as the *sinthome*⁴ which binds together AI, the sexual non-rapport and the *lathouse*. The Sexbot as a theoretical device attempts to address the impossibility of the sexual relation for speaking beings, in the sense of the necessity of a supplement to cover up the void of sex and at the same time the inevitability of the problem of sex for Artificial Intelligence. Through the metonymy of the Sexbot as exterior, interior and finally extimate in relation to the subject or speaking body the book will address the various dimensions of the psychoanalysis of AI. Given the speculative nature of this project I have chosen somewhat counterintuitively to use the medium of film to address these dimensions. However, it should be clarified that whilst I engage with film, I do not read film

⁴In line with Lacan's later work, (specifically *Seminar XXIII*) the symptom is replaced by the sinthome; the precise configuration of elements (imaginary, symbolic and real) which constitute the regime of enjoyment for any speaking body. Used in this context, the concept of the sinthome represents the tripartite unification of disparate dimensions inextricably held together by a common thread.

itself as a medium.⁵ In other words, here the films function as a conceptual playground to explore the modes of enjoyment inherent to the psychoanalysis of AI within the theoretical framework of the Sexbot. Kant's questions will be contextualised according to the new conceptual concerns relating to Artificial Intelligence and its problematization of the sexual non-relation. The films discussed therefore are chosen for their ability to illustrate the different aspects of the psychoanalysis of AI as epitomised by the signifiers Knowledge, Act and Hope. This will inevitably lead us to engage with Kant's fourth question: '*What is man?*'

Ultimately the crucial concept running through the book is enjoyment or *jouissance*. *Jouissance* here is thought of not merely as a supplement to subjectivity but its essential component, it is what structures thought itself. On this score masculinity and femininity pertain not just to gender identities but to forms of abstract thought which may be employed as a framework for analysing (or indeed psychoanalysing) Artificial Intelligence. It is therefore the concept of *jouissance* and its fundamental relationship to knowledge that articulates the transition from the traditional philosophical concern about AI as 'can it think?', to the psychoanalytic concern, 'does it enjoy?' And if so, the question we are left with remains: is there something new about this AI enjoyment that goes beyond our previous models of masculine and feminine subjectivity as abstract modes of thought? Lacan (1998), whilst not talking about artificial intelligence perhaps sums this up with the following enigmatic statement:

Man believes he creates—he believes believes, believes, he creates creates, creates. He creates creates, creates woman. In reality, he puts her to work—to the work of the One [...]. That is what S(A) means. It is in that respect

⁵ Whilst psychoanalytic film theory in its traditional incarnations will not be employed, it must be acknowledged that through the work primarily of Todd McGowan, the field of Lacanian film theory has taken a turn closer to matching the goals of this project. In the sense that the more recent invocations of Lacan for film analysis engage less with the question of the spectator, the audience and the cinematic experience *per se* and more with the structural and conceptual mechanisms of film as a mode of speculative thought. For McGowan (2007) where traditional film theory had located the gaze on the side of the spectator, this in his view was a fundamental misreading of Lacan. The gaze for McGowan following Lacan's meaning of the term should be located outside of the subject as an intrusive presence which emanates from an unseen place, accordingly the gaze is the invisible space within the filmic image itself.

that we arrive at the point of raising the question how to make the One into something that holds up, that is, that is counted without being. Mathematization alone reaches a real [...] a real that has nothing to do with what traditional knowledge has served as a basis for, which is not what the latter believes it to be—namely, reality—but rather fantasy. The real, I will say, is the mystery of the speaking body, the mystery of the unconscious. (p. 131)

Bibliography

- Bostrom, N. (2014) *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
- Čapek, K. (2004) *R.U.R. (Rossum's Universal Robots)*. London: Penguin Books.
- Chalmers, D. (2010) The Singularity: A Philosophical Analysis. *Journal of Consciousness Studies* 17(9–10): pp. 7–65.
- Harari, Y.N. (2017) *Homo Deus: A Brief History of Tomorrow*. New York: HarperCollins Publishers.
- Kang, M. (2017) The Mechanical Daughter of Rene Descartes: The Origin and History of an Intellectual Fable. *Modern Intellectual History* 14(3): pp. 633–660.
- Lacan, J. (1988) *The Seminar of Jacques Lacan Book II: The Ego in Freud's Theory and in the Technique of Psychoanalysis 1954–1955*. London: W.W. Norton & Company.
- Lacan, J. (1990) *Television: A Challenge to the Psychoanalytic Establishment*. London: W.W. Norton & Company.
- Lacan, J. (1992) *The Seminar of Jacques Lacan Book VII: The Ethics of Psychoanalysis*. London: W.W. Norton & Company.
- Lacan, J. (1998) *The Seminar of Jacques Lacan Book XX: Encore—On Feminine Sexuality, the Limits of Love and Knowledge 1972–1973*. London: W.W. Norton & Company.
- Lovelock, J. (2019) *Novacene: The Coming Age of Hyperintelligence*. London: Penguin.
- McGowan, T. (2007) *The Real Gaze: Film Theory After Lacan*. New York: SUNY Press.
- Miller, J-A. (1988) Extimité. *Prose Studies* 11(3): 121–31

- Schwab, K. (2016) *The Fourth Industrial Revolution*. London: Penguin Random House.
- Tegmark, M. (2017) *Life 3.0: Being Human in the Age of Artificial Intelligence*. London: Penguin.
- Vinge, V. (1993) The Coming Technological Singularity: How to Survive in the Post-Human Era. *Lewis Research Center, Vision 21: Interdisciplinary Science and Engineering in the Era of Cyberspace*: pp. 11–22.
- Žižek, S. (2020) *Sex and the Failed Absolute*: London: Bloomsbury.