

Vectors of Artificial Intelligence  
Ethics, Social Hope and Politics of Destiny

by

Micheal Ziegler  
B.A. MacEwan University, 2018  
M.A. University of Victoria, 2020

A Dissertation Submitted in Partial Fulfillment of the  
Requirements for the Degree

DOCTOR OF PHILOSOPHY

In the Department of Political Science

©Micheal Ziegler, 2024

University of Vitoria

All rights reserved. This dissertation may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

Vectors of Artificial Intelligence  
Ethics, Social Hope and Politics of Destiny

by

Micheal Ziegler  
B.A. MacEwan University, 2018  
M.A. University of Victoria, 2020

Supervisory Committee

Dr. Simon Glezos, Co-Supervisor  
Department of Political Science

Dr. Arthur Kroker, Co-Supervisor  
Department of Political Science

Dr. Jentery Sayers, Outside Member  
Department of English

## Abstract

I explore the idea that AI embeds itself in and (re)orders every aspect of human life by asking the question: *What is the reality and the socio-political consequences of AI research and implementation?* The dissertation is split into two main parts.

In part one, I interrogate AI discourses and trajectories to understand the magnitude of AI research and development. I begin with a cognitive AI research orientation as a means of positioning the methodological problems troubling AI research, innovation and implementation with a primary focus on the ethical limitations of AI creations. I follow cognitive AI methodology with computer scientist Ray Kurzweil's critique of cognitive AI through his revelation of a neurological reality: a mind-based intelligence always already artificial in virtual simulation—*brains simulate the mind*. Finally, physicist and AI researcher Max Tegmark's physics-based AI approach reveals the hopes and fears—ethical, social and political—associated with AI implementation as we move into the future—*why the future is bright and exciting yet equally dark and dangerous*.

In part two, I focus on questions unanswered by the vectored discourses in part one—an unrealized nonhuman metaphysics. First, we are pressed with the (de)positionality of *humanity*. Here, I place physicists turned social philosophers Karen Barad and Klaus Mainzer in exegetical discourse in order to understand nonhuman (yet fundamentally realist) orientations in a quantum-entangled universe. Doing so allows us to reposition human intelligence and anthropic wills toward knowledge as we move to understand the being of *homo sapiens* in an artificially intelligent Universe. Next, I turn our attention toward a history of intelligent orientations to break through the scientific obsession with advancement by asking how the orders of *intelligence* and *artificiality* have been patterned in the history of thought from the beginning of Western philosophy by focusing on the presocratics. Finally, I turn towards Indigenous science and knowledge creation (emphasizing myth) concerning quantum physics as a means of critiquing Western scientific understandings of intelligence. This turn towards Indigenous science expands on human hopes and fears of AI by interrogating the limitations of viewing AI problematics as future issues and instead seeing how future fears of dystopian orders exist today. By seeing dystopian realities today, we are better positioned to overcome both present and future failures of AI implementation in society.

# Table of Contents

<b><i>Supervisory Committee</i></b> .....	<b><i>ii</i></b>
<b><i>Abstract</i></b> .....	<b><i>iii</i></b>
<b><i>Table of Contents</i></b> .....	<b><i>iv</i></b>
<b><i>Acknowledgements</i></b> .....	<b><i>vii</i></b>
<b><i>Dedication</i></b> .....	<b><i>viii</i></b>
<b>0 — Introduction</b> .....	<b>9</b>
0.0.0 Selections from My Artificial Autobiography .....	9
0.0.1 The experiences of a “Smart Home Pro.” .....	9
0.0.2 AI & Fishing: Towards Species Recognition. ....	10
0.0.3 Artificial (Re)becoming: Métis-Labelled not Métis .....	12
0.1.0 Revealing the Magnitude of AI .....	13
0.1.1 Quest as a Methodological Frame.....	18
0.2.0 Part One: Entangled Vectors of Artificial Intelligence Discourses.....	22
0.2.1 Paying Heed to the Way .....	23
0.2.2 Cognitive AI Theorists: A Methodological Foundation for Authenticity and Ethics .....	25
0.2.3 Kurzweil’s Hierarchy: the Simulated Nature of the Mind .....	28
0.2.4 Tegmark: Hopes and Fears of AI—Existential Dread of a Boring Universe .....	30
0.3.0 Part Two: <i>The Straight and Crooked Path</i> —Discovering Artificial Metaphysics .....	33
0.3.1 Mainzer and Barad: Quantum Immateriality in a Universe that Matters .....	34
0.3.2 Recreating Herakleitos: Chaos, Paradox and Explosive Indifference .....	37
0.3.3 Exploring Artificial Destiny: “Sky Woman,” Apparatus, Obsolescence and Dystopia .....	39
<b><i>Part One: Entangled Vectors of Artificial Intelligence</i></b> .....	<b><i>43</i></b>
<b>1 — Cognitive AI Theorists: Discovering a Methodological Foundation</b> .....	<b>44</b>
1.0.0 Pre-empting the Issues At-Hand with Isaac Asimov.....	44
1.1.0 Introduction: Laying-bare the Issues At-Hand .....	46
1.1.1 Ordering the Way .....	48
1.2.0 Ruminating on Intelligence.....	51
1.2.1 Situating Political Intelligence with Machiavelli: Optimal Action and Cesare Borgia .....	54
1.2.2 Binding Intelligence to Optimality and Computation .....	56
1.2.3 Expanding Intelligence with Creativity .....	59
1.3.0 Being Artificial and (In)Authenticity .....	60
1.3.1 Understanding Authentic Origin: God? Or Human?.....	62
1.4.0 (Mis)Trust .....	67
1.4.1 Invading the Workplace: Capital versus Trust.....	67
1.4.2 Driverless car scenario.....	69
1.4.3 Rescue scenario .....	70
1.4.4 Canadian Navy—Military scenario .....	73
1.5.0 Loss of Control .....	75
1.5.1 All-Too-Human Self-Idolization.....	77
1.6.0 ChatGPT: Unexceptional Humanity and Competing Rationalities .....	79

1.6.1 Synopsis: What have we Learned? .....	86
1.7.0 Addendum to Chapter Three—Asking ChatGPT Three Ethically Charged Questions .....	87
1.7.1 Question One—On OpenAI .....	87
1.7.2 Question Two—PTSD and access to counselling .....	88
1.7.3 Question Three—Cheap labour and profit .....	88
<b>2 — Kurzweil’s <i>Hierarchy</i>: The Simulated Nature of the Mind .....</b>	<b>89</b>
2.0.0 Introduction: Simulation, not Authenticity .....	89
2.1.0 AI and Creativity: Authentic Creator? .....	91
2.1.1 Experiencing an Artificial Artist at the Vancouver Art Gallery .....	92
2.1.2 Who/What Experiences the Necessary Elements for Art Production? .....	95
2.2.0 Patterned Vectors.....	97
2.2.1 Evolution as a Patterned Code: Reading Increasing Levels of Complexity .....	99
2.2.2 Coding the Direction of Evolution .....	101
2.2.3 Tracking the Causes of an Intelligent Universe: A Teleological History .....	102
2.2.4 Vectoring Evolution .....	106
2.3.0 Exponentially Accelerating Abstractions: the Simulated Mind is Born.....	109
2.3.1 Towards Singularity: A Transhumanist Eschatology .....	112
2.3.2 Judging AI Coming-Into-Being.....	113
2.3.3 Digitizing the Soul: Soul-Consciousness-Mind.....	116
2.3.4 Living in the Cloud, Indigenously: Psychic Resurgence .....	121
2.4.0 Synopsis: Transcending Humanity.....	124
2.5.0 Addendum to Chapter Two: Questions Concerning Consciousness with ChatGPT .....	128
<b>3 — Tegmark: Hopes and Fears of AI—Existential Dread of a <i>Boring Universe</i>.....</b>	<b>131</b>
3.0.0 Another Selection from My Artificial Autobiography: “ <i>Vectors of the Night Sky</i> ” .....	131
3.1.0 Situating Tegmark.....	133
3.1.1 Consciousness and Tegmark.....	135
3.1.2 Ethics and Tegmark.....	139
3.2.0 Revisiting Trust: Technology, Labour and Judgement.....	142
3.2.1 Situating Labour in the Human as Human.....	143
3.2.2 Justice Who? Judgement By? .....	144
3.3.0 Hope Part One: The Joblessness Problem.....	147
3.3.1 (Un)Desirable Labour and Being Human .....	149
3.4.0 Hope Part Two: Political <i>Not Place</i> (οὐτόπος).....	151
3.4.1 Egalitarian Utopia .....	152
3.4.2 Gatekeeper .....	154
3.4.3 Enslaved God .....	157
3.5.0 Fear: Political <i>Apokalupsi</i> (αποκάλυψη).....	159
3.6.0 Synopsis: Closing the Section with <i>ChatGPT</i> .....	162
3.7.0 Addendum to Chapter Three.....	166
3.7.1 First Addendum: Asking ChatGPT for a poem .....	166
3.7.2 Second Addendum: Asking ChatTGPT for an essay.....	166
3.7.3 Third Addendum: Asking ChatGPT for a research essay .....	167
<b>Part Two: The Straight and Crooked Path—Discovering Artificial Metaphysics.....</b>	<b>169</b>
<b>4 — Mainzer and Barad: Quantum Immateriality in a Universe that <i>Matters</i> .....</b>	<b>170</b>
4.0.0 Another Selection from My Artificial Autobiography: “My Artificial Teacher” .....	170
4.1.0 What Next?.....	173
4.1.1 A Brief Review of part one.....	173
4.1.2 Introducing the Problem at Hand .....	175

4.2.0 (Re)Discovering <i>Metaphysics</i> .....	177
4.2.1 (Re)Positioning Aristotelean Πρώτη Φιλοσοφία for a Quantum-Technological Age .....	183
4.3.0 Quantum Relationships—Discovering the Underlying.....	188
4.3.1 Complexity and Quantum Physics .....	190
4.3.2 Positioning Action in a Quantum Universe.....	193
4.3.3 Metaphysical Rigour as Phenomenological Physics—The Case of Quantum Computing.....	194
4.3.4 Applying Nonlinear-Complex Metaphysics.....	199
4.4.0 Synopsis: Competing Ontologies .....	203
<b>5 — Recreating Herakleitos: Chaos, Paradox and Explosive Indifference .....</b>	<b>205</b>
5.0.0 Introducing Chaos.....	205
5.0.1 Reviewing Previous Chapters .....	205
5.0.2 (Re)Introducing Chaos with Hesiodos .....	207
5.1.0 Herakleitos Joins the Debate.....	212
5.1.1 Understanding Logos—Principle, Account or Both? .....	213
5.1.2 Limited Potentiality.....	217
5.1.2 Paradox of Chaotic Entanglement .....	222
5.1.3 Debating Cosmological Judgement—Divine Intelligence.....	225
5.2.0 What next? An Intermission on Artificial Becoming .....	229
<b>6 — Exploring Artificial Destiny: “Sky Woman,” Apparatus, Obsolescence and <i>Dystopia Now</i>.....</b>	<b>231</b>
6.0.0 Orienting Towards Collaborative Intelligence.....	231
6.0.1 Collaborative Intelligence: “Sky Woman” and Exogenous Worlds .....	232
6.1.0 Entangled Intelligence .....	235
6.1.1 Understanding Western Endogenous Reduction: Complementarity and Observation .....	238
6.1.2 Entangled Being—Authenticity in Artificial Life as Apparatus .....	243
6.1.3 Living-in-and-with-the-world: Seeing Artificial Life Everywhere.....	245
6.2.0 Destiny in Obsolescence.....	248
6.2.1 Obsolescence, Domination, and Fear—Revealing a Dystopia Present through Indigeneity .....	250
6.2.2 Beyond Human Obsolescence: Breaking the Limitations of Recognition .....	254
∞.∞.∞ Conclusion: Collaborative Intelligence and Infinite Recreation .....	256
<b>Works Cited.....</b>	<b>259</b>

## Acknowledgements

The supervision I received to complete my PhD studies was perfect. Arthur Kroker provided me with creative and boundary-pushing encouragement. His advice throughout my PhD has been indispensable. Simon Glezos was an ideal co-supervisor who ensured I had support when and where needed. Simon also provided me with the leadership I needed to complete my PhD promptly while encouraging creative/critical thinking that pushed boundaries. I could not have asked for better mentors or better friends.

My PhD studies were supported and encouraged by my friends. I am forever in debt to my good friend Tyler Wilde, who read and provided feedback on every chapter as I wrote them. My morning coffees with David Miller were indispensable to both my dissertation and surviving my PhD studies. My office mate and friend, Neil Montgomery, was a constant source of dialogue, criticism and encouragement. My old roommate and classmate, Luke Lavender, provided me with the idyllic version of academic life: a true experience of philosophy as friendship. Shianna McAllister was a constant source of information, support, and criticism; her knowledge of Indigenous theory, Indigenous criticisms of Western philosophy and her general knowledge as a true political theorist were indispensable. My friends at and beyond UVic who have supported me and were sources of criticism cannot go unmentioned: Lynn Ng Yu Ling, Phil Henderson, Stacie Swain, David Mulligan, Duncan McManus, Holly Ching, Nicole Faull, Brydon Kramer, Andrew Marriott, Parvin Sedighi, Stephanie Nedoshytko, Andrew Bieman, and Savanna Ribeiro.

In conversation, I often attribute my academic successes to the beginning of my scholastic journey, and thus, I cannot omit a nod to my baccalaureate institution, MacEwan University. At MacEwan, I had the privilege of receiving an incredible education from amazing, passionate professors such as Gaelan Murphy, Chaldeans Mensah, Jean-Christophe Boucher, John Soroski, Michael Seredycs, David Reddall and Edvard Lorkovic.

At UVic, I had the opportunity to study in the Political Science department, which put me in conversation with many fantastic professors who supported me in more ways than one: Matt James, Mara Marin, Feng Xu, Scott Watson, Avigail Eisenberg, Justin Leifso and Kelly Aguire. Studying at UVic also allowed me the opportunity to study Cultural, Social and Political Thought, an interdisciplinary concentration that allowed me to meet some stellar professors like my committee's outside member, Jentery Sayers, and others like Emile Fromat de Rosnay, Stephen Ross and Sara Ramshaw. Finally, I am extremely grateful to UVic's graduate funding, SSHRC's doctoral funding, and the Alberta Government's Sir Lougheed Award of Distinction, without which I never could have dreamed of completing a PhD.

## Dedication

*To the margins.*

~

Without which  
my thoughts would be lost  
in time, to space

## 0 — Introduction

### 0.0.0 Selections from My Artificial Autobiography

#### 0.0.1 *The experiences of a “Smart Home Pro.”*

One particular summer, I worked under the recondite title: “Smart Home Pro” for a smart home and home security company. I primarily installed systems in rural Alberta, a familiar scene for a northern Alberta country boy such as myself. The clients generally had similar needs: outdoor cameras to surveil their secluded property and to record and notify them of movement. This particular summer, the company released an AI-enhanced outdoor camera. It came with the ability to distinguish between people, nonhuman animals and vehicles—this ability never failed to excite clients. However, without fail, I always received calls from the clients informing me that the cameras did not work. They were never notified, nor did the cameras record vehicles on their property. I would always re-explain to them that the cameras are not capable of registering vehicles this way; they only see people as worthy of “action,” and they have to be within fifteen feet of the camera for it to take notice.<sup>1</sup> The customers were always irate when presented with this fact and would generally respond: “How does that help me? I live in the fucking boonies! Who is going to walk up to my property?” One client explained that a thief drove past their house and rammed their locked barn door. The barn was more than fifteen feet from the house. This camera is not programmed to recognize this incident as worthy of recording and notifying

---

<sup>1</sup> While these cameras could tell the difference between humans, nonhuman animals and vehicles, they were not programmed to view nonhuman movement as a threat. The purpose of being able to tell the difference between these actors is to allow the camera to know when to notify homeowners of “actual threats.” The company focused primarily on suburban security; thus the cameras were programmed to deal with the most statistically likely threat: pedestrians that are close to house (within fifteen feet). Vehicles were included in the camera’s assessment so that it does not mistake someone on a motorcycle, bicycle or in a convertible as a ‘pedestrian.’ Nonhuman animals are included so that it doesn’t mistake a large dog for a child. A costume that hides the shape of the torso and head, and hides one’s face could trick the camera—so there were limitations.

the homeowner because the vehicles are not recognized as threats, and when the person exited the vehicle, the homeowners were not notified because the person was more than fifteen feet away from the camera.

In northern Alberta, this AI came face-to-face with the realities of rural life. Its suburban biases—revolving around jealous neighbours and opportunistic porch pirates—impede its ability to understand the lived reality of vast open spaces; the AI is bound to its base coding. Or, perhaps, it is incapable of recognizing another machine as a threat; the system had a clear bias towards pedestrian thieves, a purposeful design that implies a particular material existence of “thieves.” The programmed social heuristic bias that *like does not harm like* (machine-to-machine or car owner-to-car owner) might be embedded into the world of the intelligent machine; even artificially constructed intelligence is not free of social biases. The power of these artificial imitations of biological processes of recognition appear to be growing ever more from interesting parodies to a pastiche of biological life.

### *0.0.2 AI & Fishing: Towards Species Recognition.*

As many of my friends and colleagues know, I am an avid angler; fishing is my favourite pastime. I have been fishing regularly for longer than I can remember. Since living in Victoria, British Columbia (BC), I have started fishing in the ocean, with limited but ever-growing success. Let me recount an experience with fishing technology and my first ocean angling success.

Since I am not an expert in ocean fishing, ocean fish have not been fully downloaded into my biological hard drive for me to recognize any specific species I have caught. This lack of cached data is no worry because Fisheries and Oceans Canada teamed up with the University of

British Columbia to create an ocean fishing app. Beyond digitally storing my fishing licence, and providing live weather updates, regulation, wave and tide information, it has a built-in “Enhanced Species ID”—very cool. The “Enhanced Species ID” uses AI recognition software to determine the species of fish I catch.

A weekend in June 2022, I caught a fish while kayaking in the Haro Strait off the coast of Victoria, BC. I was knowledgeable enough to know it was a rockfish. However, the app lists twenty-four possible rockfish species. Using the Enhanced Species ID, I snapped-a-pic, and it returned: “Canary Rockfish, 67% certainty.” For me, 67% certainty is not nearly enough. I repositioned the fish and tried again; it returned: “Quillback Rockfish, 89% certainty.” Equipped with this information, I searched the specific species identifying features and confirmed it was indeed a Quillback Rockfish. Amazing! My phone and I worked collaboratively to save me from potential accidental poaching due to misidentification or missing out on a “keeper” if self-doubt forced me to throw it back.

I was pulled into the entangled swarm intelligence of AI; my ability to recognize the world around me had increased exponentially—in speed and capacity: no longer limited by the need to engage in individuated biological recollection nor needing to refer to physical species charts with the hope I’ll find the fish in question in a reasonable amount of time before I need to put it back in the water so that it doesn’t needlessly suffer. Together, we overcame our limitations: I unable to memorize all the fish quickly, it unable to identify the fish with 100% accuracy. AI has enveloped itself around me, folding me into the swarm; together, we become more than mere *homo sapiens* and phone; we tapped into the evolved *cyborg life* of biological-AI enhancement. A new order to my life begins: *AI-ordered fishing*.

### 0.0.3 Artificial (Re)becoming: Métis-Labelled not Métis <sup>2</sup>

Growing up in Alberta, I was constantly torn between the various identities of my “mixed race,” Métis-labelled life. I was always torn between a white racist father reminding me that my undesirable traits were “the native in me”—he would often say this when I did something he disapproved of. A Métis-labelled Kanien’kehá:ka (Mohawk) mother resentful of her Indigenous ancestry, often telling me I was acting like a “Rez Indian” when I did something she disapproved of.<sup>3</sup> White friends quick to remind me that I am Indigenous when it suits them—often to excuse them for something they said that was racist, to assure them it “wasn’t actually racist.” Indigenous (Métis, Cree and Dene) friends always ready to distance me as a “half-breed” that is more white than Native when they wanted to exclude me. I existed in the various artificial renderings of my friends and family who bought into the historical, settler ideal peddled by settler governments that Métis is a class of “mixed race” people detached from specific Indigenous histories. Whereas Métis are peoples with a rich history steeped in histories, traditions and lands.

My history of artificial construction, white when it suites people, Métis when desirable and Indigenous when needed, rests within the artificiality of semantic, conceptual reality. Like

---

<sup>2</sup> I refer to myself and my mother as “Métis-labelled” for three reasons. First, I acknowledge with Chris Andersen a problem with the history of Métis identity and experience. The idea that Métis are “mixed race” is an imperial-racist history that places Métis peoples as either further up the hierarchy of races implying they are more human compared to “pure” Indigenous peoples or that they are simply not Indigenous. However, Métis peoples have a history and this is attached to Red River, implicitly (See Andersen 2014). Second, there is a long-standing history of so-called “mixed race” Indigeneity attached to the term Métis for so many people, like myself, who grew up Métis-labelled by friends, family, strangers and the government—often living as a floating signifier with no place because we (Métis-labelled peoples) struggle to find place 2024-04-19 10:27:00 PM. Finally, by referring to myself as Métis-labelled, I acknowledge my lived experience under this label and identity I claimed for most of my life as I discovered these histories and realities, and moved towards seeing myself not as a mere “half-breed” but as Indigenous. In this manner, I acknowledge a contentious, historical, on-going debate regarding being Indigenous whilst calling attention to my personal lived *out-of-place* experience because of imperial socio-politico-historical realities demanding separation and competition on being true to one’s self and one’s social (See Lawrence 2002).

<sup>3</sup> “Rez Indian” is a derogatory insult given to Indigenous peoples living on reserve land to distinguish them from the equally derogatory named “City Indians.” Two terms, among many, I was very familiar with growing up.

the digital world creation of the AI recognition, I lived a life that rendered me as Métis. However, my life as Métis was artificial. I was never truly Métis, as I have no attachment to Métis traditions and no ancestry of Métis being. Thus, I have grown to refuse the artificially rendered life of being labelled Métis and to live differently, to be Indigenous, to accept and live my Mohawk ancestry. A Mohawk existence, that is, admittedly, out-of-place. I have never lived on Mohawk lands. I grew up with the traditions and on the traditional lands of the Dënesųłné peoples of the Cold Lake First Nations. A culmination of identities artificialized by being-in-the-world disentangled. My being emerges disentangled and re-entangled in the artificiality of different renderings of the same world: the non-digital world of being in nature, the AI rendering of a recognized life, the Canadian rendered world, and Indigenous worlds—many worlds, one being.

### **0.1.0 Revealing the Magnitude of AI**

“The Goal of [AI] is the construction of self-organizing computing systems that provide services, that can help us manage a world of increasing complexity and to support a sustainable human infrastructure in the future” (Mainzer 2007, 287). Quantum Physicist and AI researcher Klaus Mainzer, makes explicit the point of creating highly intelligent and capable AI machines is to allow human societies to flourish, to perform beyond the simple means of biological potential and to overcome biological brain-mind limitations of thought to truly understand all elements of the Universe: the chaos of exogenous quantum realms and the endogenous dynamics of socio-political relations. AI services help make anglers more efficient in the chaos of fish habits, impose urban familiarity in rural securitization and entangle in everyday human affairs, freeing us from the mundane task of remembering where we are or where we are going. Numb to the

extended world of experiences, we walk willingly into the “final phase” of a technological narcosis, ensuring the central driving factor of AI is *willing* human comfort into the future.<sup>4</sup>

Highlighting the human-centred push toward the future is not meant as an implicit critique; it is an opening to the direction of AI research and implementation. For instance, we see a reverse Turing test as a result of dominant AI logics forcing companies to employ cheap labour to imitate AI systems to create the appearance of complex AI systems (Yuan 2018).<sup>5</sup> We see the power of simulated reality incidentally tricking people into believing Boston Dynamics had created a soldier robot when it was, very clearly, a parody created by exceptional CGI artists (*New Robot Makes Soldiers Obsolete (Corridor Digital) 2019; Reuters 2022*). Not AI trying to convince humans of their humanity but humans tricking humans they are AI.

So what? Who cares? People cannot go anywhere within mainstream society without interacting with some form of AI.<sup>6</sup> We are pressed with the Ubiquity of AI in everyday life. AI

---

<sup>4</sup> Canadian media scholar and technology critic, Marshall McLuhan, argued there will come a time when technology will extend human consciousness beyond the immediacy of the human creature to touch every corner of human life. This extension will make us numb to both ourselves as individual creatures and the technology doing this. There will be a breakdown of the accepted boundaries of *what is* (see McLuhan 2010, 3, 46).

<sup>5</sup> A *reverse Turing test* was developed to fix some issues of the Turing *imitation game* (also known as *Turing test*). In a reverse Turing test, a human tries to convince a questioner they are AI, instead of a computer trying to convince a questioner they are human (see Baird, Coates, and Fateman 2003).

<sup>6</sup> Every day journalists tell us of the new and exciting potentials of AI. From a disgruntled Google employee claiming Google has created a sentient AI (N. Grant and Metz 2022), the US military investing in Shield AI’s research towards the creation of swarm intelligent drones—autonomous drones that learn independently and collaboratively in the field with both humans and other drones (Insinna 2022), to unsurprising but increasingly relevant and more powerful AI tracking software capable of learning our needs and desires from the things we do, not only online, but seemingly everywhere we go—like movie theatres (Pringle 2017). Furthermore, in Canada, we can think back to Telus’ *Home Health Monitoring* system which employed AI tracking software to help the Canadian Government and hospitals monitor COVID-19 patients, track infection rates, and alert businesses and Canadian individuals of possible infection. In Canada, we see AI developers working with hospitals and governments to develop personal health databases and management systems to aid in medical decision-making (Winter and Davidson 2019, 280–81). In government bureaucracies, we can think of Service Canada and the Canada Revenue Agency using algorithmic, narrow AI systems to speed up processes by using these systems to perform menial tasks like data management and simple form assessment. Worldwide, governments are developing robust algorithmic bureaucracies to help provide services to their citizens (Vogl et al. 2019, 148–49). Furthermore, we can think of finance and banking. For instance, the Royal Bank of Canada has an AI service called “NOMI Find & Save,” which uses narrow AI to track a client’s usual debits and credits to determine the effective amount of money

has and continues to have huge consequences on our present. While AI in the future will be “more” revolutionary compared to today’s AI systems, one cannot ignore and undersell the ubiquity of AI today. Thus, when coming face-to-face with the reality of AI, we are pressed with this project’s primary question: *what are the consequences of AI research, development, and implementation on society?* Through this project, I propose: *AI is the ordering principle that, like a computer compiling data, renders (shapes, recreates) socio-political realities.*<sup>7</sup> That is, the principle of our contemporary and, likely, human existence, generally, has always been one of artificial intelligence: “*we are all AIs*” (Hunt 2021, 62). The creation and fabrication of our world—abstractly and physically—is the driving factor of being. Ours is a malleable existence eternally formed and reformed by endogenous and exogenous forces denying control over destiny and life itself.

With this foundational argumentative framing, we are presented with many questions regarding AI and its potential. What is the relationship between AI and *biological creatures*? What is the experience of AI in this world and the algorithmic world of its software mind? Is the manufactured hardware embodiment of AI the same as birthed biologically embodied creatures?

---

to put into a saving account once a week—a collaborative AI system working with its human user to effectively manage their finances. While AI managing our personal finances is impactful, if you were to go onto *The Wall Street Journal* website, you will notice there is a section dedicated entirely to AI. This is no coincidence. There is an unavoidable interest in financial technology (or FinTech), specifically AI. From managing stock investments to managing stock markets themselves, the financial sector can no longer ignore FinTech, as Lui and Lamb (2018) explain: “Impartiality and the lack of judgement are reasons why customers like robo-advice in wealth management... Efficiency in robo-advice leads to fee reduction for customers” (271-272). Here Lui and Lamb are explaining the two-pronged benefit of AI in banking and investment sectors. First, there are no judgemental eyes looking back at you when you get advice from AI. Second, AI is a cheaper labourer than their *homo sapiens* counterparts. As with the history of manufacturing moving to “poorer” nations to find cheap labour, we see businesses—across all sectors—looking to the “cheap” labour of artificial beings.

<sup>7</sup> “Render” is used here to illustrate the computational overtones of AI in society: rendering (processing) information from cached data. Rendering, as in provide a service. Rendering translates the world into data for use. Rendering describes a myriad of computational powers within the realm of AI logics. That is, the use of computational terms like render here or hard drive and cached data earlier, places me firmly in the posthuman reality of N. Katherine Hayles who describes posthuman reality as one where people see reality through computational terminology (see Hayles 1999, 6)

What is the relationship between the algorithmic software world and the hardware material world, and how do relationships with AI experience affect relationships with others? These lines of questioning propel us toward understanding the ontological order of AI research and engineering, and the metaphysical relationship between AI ontologies and epistemic orders. That is, in seeking to understand the ordering principle of AI, I am confronted with the ontic presence (the being) of AI—what is the *picture* of its existence?

Finally, what of *humanity*? By exploring the socio-political ordering principle of AI, I explore what is seen as the imminent loss of the nontechnological human (individuated, liberal, biological fleshy creature) and that which comes to replace it: *an artificial future moves us beyond simple human experiences, thereby entering the more-than-human world*. That is, with the *coming-into-being* of the *post-human*, “the days of ‘the human’ may be numbered.” (Hayles 1999, 283).<sup>8</sup> Eminent posthuman scholar N. Katherine Hayles argues we have entered a technological era wherein the determinable, individuated and free-acting *liberal human*, as a conception of proper being-in-the-world, is on borrowed time. She goes so far as to write the human never existed; however, she stresses the influence of liberal human conception still affects the early days of being-posthuman (287–89). Ultimately, along the lines of Hayles, I argue: by implementing AI in society, an idealized particular human being relinquishes its position as the implicit director of its destiny and invites unprecedented change in its relationship to being-in-the-Universe.

The perceived liberal ideal of a human-controlled destiny of the world around us and what is yet to come stands face-to-face with the unprecedented power of AI. In other words,

---

<sup>8</sup> Hayles is concerned with the age of cybernetics. I.e., increases in cybernetics has altered how humanity perceives both the world and itself. This change in perceptions breaks down the stability of human ideals and ultimately, the stability of the human itself.

when we join the artificial order of AI, we enter an age of artificial (re)becoming of what we were yet repositioned as a pastiche, an ersatz manipulation from an artificial artificer.<sup>9</sup> However, for Hayles, we become what we always already were: *complex beings entangled in the universe bereft of agency directed by infinitely varying cosmic forces*. In other words: “the distributed cognition of the human subject correlates with... the distributed cognitive system as a whole, in which ‘thinking’ is done by both human and nonhuman actors” (Hayles 1999, 290). For Hayles and myself, as we move further into the technological future, we must realize, once and for all, the artificial signifier *human*, be it a liberal entity, biological or something else, is not the exceptional arbiter of destiny.

This idea of existence, wherein human control is called into question, reveals a dichotomy of orientations to *homo sapiens*’ future. One is the posthuman orientation wherein the liberal human as *human* is seeing (or has seen) its final days, and technology has altered being-in-the-world as *homo sapiens* as creatures entangled with nonhuman destinies. The opposing orientation is the *transhumanist*: “a philosophical, cultural, and political movement that holds that the human species is now in a comparatively early phase and that its very evolution will be altered by developing technologies” (Schneider 2016, 261).

Thus, posthumanism is an acceptance of a technological becoming; i.e., it comes with the implication that *homo sapiens* are not ‘in control.’ Importantly, transhumanism also accepts the fact that technology has allowed for a great change of *homo sapiens* (im)material existence. However, the transhumanist view rests firmly upon the idea that humanity controls destiny.

---

<sup>9</sup> We should think of dating apps like Tinder as a great representation of AI control of *homo sapiens* destiny. Here, the algorithms of AI machinery take command of mating. The selection of our mates is directed by artificial control mechanisms. Who I will see as a potential mate is selected and packaged for me. While I might have some control, the order of presentation, the probability of who I will see, is all controlled in the algorithmic world of artificial being.

Humanity creates and controls technological implements, such as AI, and as a result, must be in control. Furthermore, the transhumanist orientation rejects the idea that the liberal human's time is coming to an end and, instead, holds that the liberal human will transcend its fleshy existence and move into the artificial future remaining in a position of complete control. These two opposing views, posthuman and transhuman, rest as the foundation of our quest.

### *0.1.1 Quest as a Methodological Frame*

Important to this project is the idea of hermeneutic "questing." Questing is a methodological frame that takes as its method: *going on a quest*. This idea is derived from Heidegger's "The Question Concerning Technology:"

Questioning builds a way. We would be advised, therefore, above all to pay heed to the way, and not fix our attention on isolated sentences and topics. The way is a way of thinking. All ways of thinking, more or less perceptibly, lead through language in a manner that is extraordinary (Heidegger 2013b, 3).

In questioning, is the route word 'quest.' What is derived here is the importance of *building a way, paying heed to the way, and leading through language*. At the beginning of any quest is the establishment of *a way*: a question, something to study and a general idea of what to question: texts, ideas, stories, concepts, etcetera.

Second, is *paying heed to the way*. That is, as one embarks on the quest, there is much to consider and discover. The quester follows many paths and must decide where to go and why. This can be because the quester has gathered sufficient information to know where one ought to go to get to a destination or one is led by other travellers. In terms of an academic quest, the quester often leads themselves on what to read, what to write, when to speak in their voice or another's and when to follow another's thought. What is important is all the involved instances move towards some destination: the quester heeds to the information.

Finally, our quest is *through language*. We navigate through words and ideas toward discovery. Positioned here is hermeneutics (or ἐρμηνεία), which is the amalgamation of Ἑρμῆς (Hermes), the Greek god of communication and a declension of ναῦς (naus), meaning ship. Thus, hermeneutics is a journey since it is the act of navigating (by way of a ship) through language.

There is a technological element to hermeneutic questing with the etymological inclusion of a technological thing: a ship. However, what is most important is what scholarly quests are, here expressed in the words of Herakleitos:<sup>10</sup>

[The straight and crooked journey of writing is one and it is the self.] “γναφῆων ὁδὸς εὐθεῖα καὶ σκολιή μία ἐστὶ καὶ ἡ αὐτή” (my own translation—DK22B59)<sup>11</sup>

<sup>10</sup> Herakleitos is the transliteration of Ἡράκλειτος (Heraclitus). Since this project seeks to reorient thought, I feel it important to break from imperial histories to pay respect to Indigenous modes and being. I have chosen to write transliterated Ancient Greek names instead of their Latinized/Anglicized counterparts. As a once Métiis-labelled person, it is important to engage in acts of refusal of imperial histories and to engage in a kinetic dance of *harmonia* as one moves through their being as ancestral past and present colonial being towards harmonious futures. Performing this refusal will challenge native English speakers to slow their reading when presented with “uncommon” spellings. This approach to writing Ancient Greek names will include Platon (Πλάτων—Plato) and Aristoteles (Ἀριστοτέλης—Aristotle). Against the common thinking that certain transliterated names could cause confusion or be an unnecessary ask of readers, I do not think it is a big ask of readers nor a source of confusion to use Platon and Aristoteles—“Transliterations bring ancient Greek out from under the shadow of Roman conquest” (See Zwicky 2015, vii).

<sup>11</sup> I rely on my own translations in the project when I judge other translations to be inadequate. The text referenced for presocratic “originals” is Diels 1960. All references to presocratics conform to the established Hermann Diels-Walther Kranz numbering system: DK. For information on the DK ordering system, see Curd 2010; McKirahan 2010. There is a great difficulty in translating this particular line. Some translators ignore the final portion of the statement: *it is one and the self* “μία ἐστὶ καὶ ἡ αὐτή” preferring only to translate: *the journey/road/path* (ὁδὸς) *is straight and crooked* “γναφῆων ὁδὸς εὐθεῖα καὶ σκολιή.” The word αὐτή—meaning ‘the self’—in this context creates some difficulty. However, consider the following line in the DK text: *The journey above and below is one self*. “Ὁδὸς ἄνω κάτω μία καὶ ὅση” (DK22B60). From what I have seen, ὅση is translated as ‘the same’ causing a common translation to look like: *The road up and down are one and the same* (Mackenzie 1989, 5). However, the word: ὅση, a cognate of αὐτή, taken literally, means ‘self’ (See McKirahan 2010, 132). The issue translators are often faced with when it comes to the presocratics is we do not have “originals.” Fragments and testimonials are found in a wide variety of texts by different authors throughout antiquity and the early-medieval period. It is these varied sources that we look to for presocratic writing and Diels (and others) compiled into books. Since it is generally agreed Herakleitos is interested in oneness, sameness and multitude, translators will read this statement in the context of oneness and sameness (self as a single thing and same to itself). Furthermore, given that presocratic writers were often limited in word choice, they were known to stretch the meaning of words. This approach to reading leads translators to read ὅση (self) as referring to one thing (a self) and thus can (maybe must) refer to one thing as the same thing—hence sameness. However, there is no reason to believe Herakleitos must be referring to his usual inquiries. Since, as the eminent presocratic scholar Richard D. McKirahan explains: we have no reason to believe any of the fragments are from the same text nor in reference to the same ideas, regardless of their numerical position in the DK numbering system (see McKirahan 2010, 1–6; 124–28). Given all this, we could translate

Herakleitos, here, expressed the idea that any journey is straight in its determination: the quester has both a destination and trajectory in mind. However, regardless of a straight “line” towards some destination, every quest is crooked. The quest itself, while “determined” towards some end goal, is chaotic in its evolution. Quests are iterative, and there are unforeseen forces. It is this chaos that reveals not only the journey but the quester. I see *self* (αὐτή), in the quote above, to refer to the journey as entangled with every element of the quest, including the quester. The quest not only reveals the ideas written about but the quester itself (or himself, in my case). Here, at the beginning of my quest, is the revelation of a deeply hermeneutic orientation that allows for deconstructing important ideas whilst simultaneously staying involved in constructing important ideas. This act is important because it allows for the chaos of a quest to determine paths rather than relying on a perceived optimal path bound to personal determination. While I clearly have goals in mind to complete this quest, I am bound in the varying trajectories laid out before me. That is, ‘quest’ does not seek to impose a particular order on the researched material and rather lets the material dictate the path.

I see a quest as a distinct and useful endeavour opposed to other, seemingly similar, modes: an investigation, for instance. In an investigation, a good investigator would allow the discoveries of the scene to dictate their investigation. However, important to an investigation is a scientific mode that forces very particular paths. A murder investigation is an apt analogy.

---

DK22B59 as: *the straight and crooked journey of writing is one and the same*, rather than ignore the final half. However, Herakleitos’ general interest in ‘oneness’ allows ‘self’ to make sense here. But what oneness is, is at question here. Since I am not interested in creating a compository of Herakleitos’ fragments, I have some leeway in translating. I don’t require all the fragments to line up and make sense as they are from one book. Thus, for my purposes, it is beneficial to translate αὐτή literally as ‘self,’ thereby allowing us to see the relationship between a self and the journey of writing as the entangled oneness of a thoughtful quest in writing.

When investigating a murder, one has certain data that must be searched first—at the crime scene, one must see if there is DNA, footprints, fingerprints, etcetera. The discoverability of these data would dictate the next steps of the investigation. Each step of the investigation is dictated (within reason) by certain necessities and established norms and practices that make it incredibly linear—each thing must be discovered (or not) before the next. This linearity is not the case for a quest. While my dissertation is presented linearly, it must not be, nor was it, approached thusly. Its linear appearance is merely a particular rendering of the quest.

Ultimately, a quest is distinguished from an investigation because there are no particular necessities of discovery that must be adhered to. It is similar to an investigation since each discovery dictates the path, and the quester decides what to do along the path based on each discovery. Important, however, questing should not impose particular needs upon the paths of discovery. A quester must let each discovery dictate, as much as possible, how to proceed. The idea is to allow the world to reveal itself rather than forcefully reveal the world.<sup>12</sup> Importantly, a quest is aimed at seeing what is said and how things are said, to understand what is believed and by whom to better understand the world from the speakers' point of view.

With this method, I am better positioned to engage in a critical rephrasing of other points of view in order to understand how disparate approaches are similar in their orientation whilst being unique. For instance, while each chapter deals with different, unique views and approaches to AI, I reveal how they have the same foundational assumptions—transhumanism in part one, for instance. However, this requires a rephrasing of discovered ideas, which can only be properly

---

<sup>12</sup> This may also be similar to Cherokee scholar Brian Burkhart's *trickster methodology*. However, Burkhart explains that a trickster methodology is *not a critical mode* (Burkhart 2019). It should be clear that the quest is critical. That it is critical does not make it inherently better, nor worse than a trickster ethos. Nor does the fact that it is not bound within the particular scientific rigour, of say, an investigation, make it "unscientific." It is a particular mode directed at a particular act of discovery.

done when ideas are most fully articulated. This requires that the ideas speak for themselves before I, as the quester, speak for them.

### **0.2.0 Part One: Entangled Vectors of Artificial Intelligence Discourses**

The choice of vectors (in the title of this project) was neither accidental nor incidental. *Vector*, as a term, is often presented interchangeably with *trajectory* when discussing general straightforward motion. However, a *trajectory*—in physics—refers to the line across which some object moves: its path (or perhaps its *journey*). Whereas *vector*—in mathematics—refers to magnitude and direction. An object has direction (its trajectory), and it has magnitude (distance or quantities); both can be observed and measured. Magnitude, in one instance, accounts for the distance of the points along a trajectory. In the second, magnitude accounts for the size of an object and/or quantities important to the force(s) projected upon an object that move it along its trajectory. Finally, by referring to distance, size and/or forces, magnitude provides the necessary components to understand the possibility of vectors. That is, forces like gravity and quantities like velocity, momentum and acceleration are vectors. Gravity is revealed by direction: appearance of falling along a trajectory, and magnitude: size of the objects involved that create the relational forces of attraction; there must be a measurable distance or quantity (inclusive of size for understanding vectors like gravity) and an observable trajectory. Furthermore, no one vector exists in a vacuum and must be expressed in relation to other vectors. Gravity's trajectory is obstructed or made crooked by perpendicular momentum. While a single vector has a singular trajectory, entangled vectors push and pull in all directions, and by questing with vectors, we make discoveries concerning the thing(s), object(s) or subject(s) in question in relation to other thing(s), object(s) or subject(s).

Vectors are dynamic; they imply relationships between spaces, forces and actors, allowing us the opportunity to visualize relations in given space(s) as complex *Gestalt* wholes rather than disparate, reduced, simplistic moving parts. That is, vectors have a greater capacity for discovering the patterns of space-time relationships (See Mainzer 2007, 44–45). Positioning *vector* is important when we are considering artificial intelligence because we have an entangled reality of patterns in which intelligence is brought into physical-material being by the forces of creativity—a culmination of artificial creation and creative ingenuity (See Kurzweil 2013, 150–55). Vectors, as a metaphor, thus, create a *spacely* awareness of the forceful magnitude of a dominant AI presence. That is, we can understand AI as a vector: *it has observable trajectories and quantities that pertain to it and affect its surroundings.*

Thus, a vector is part of the *leading of the way heeded by the quester*. Vectors are the forces that push and pull, *making crooked the straightforward journey of the quester*. Vectors are the quantities that create chaos in the prime quest. AI, as a primary vector, pushes and pulls us toward a particular point, whilst vectors entangled with AI push and pull us in other directions. Our hermeneutic questing causes us to construct boundaries to better understand the thing(s) in question while these boundaries simultaneously fall apart from the forces discovered during the journey (the idea of boundaries rising and falling will be expanded upon below in the discussion of chapter four in subsection 0.5.1).

### *0.2.1 Paying Heed to the Way*

I split the dissertation into two parts. In part one, I focus on three foundational modes of AI research approaches—cognitive AI, neurological AI and algorithmic (mathematic) AI as some of the leading AI approaches on how best to understand, program, construct and implement AI in

everyday life. These approaches rest upon transhumanist notions of AI implementation. Of these approaches, Vincent Müller (2016a) and others I explore are the best representatives of cognitive AI approaches in terms of theoretical foundations, questions and the present limitations and future potential of AI cognition. Ray Kurzweil (2013) is highly influential and arguably the most famous neurological AI researcher/thinker/engineer today—best known for his work on AI pattern recognition and *AI-human-singularity* prophecies. His expertise in the field is essentially unmatched and unavoidable. Finally, Max Tegmark (2017), as an influential physicist, mathematician and computer scientist, is well situated and knowledgeable to understand mathematic approaches to both being-in-a-Universe and AI potential.<sup>13</sup>

What we get by navigating these AI approaches are discourses that appear disparate and incommensurable. We get a view into highly complex and contentious debates and discussions on how best to create AI systems, what the limitations are, what (if any) limitations—as creators—ought we impose on AI systems, and how AI affects not only our present but how AI will affect our future. This approach will not only show the different views concerning AI; it will reveal central tendencies that lead toward a singular vision of being intelligent: *extending conscious human intelligence to its limitless potential*—transhuman control. For enthusiastic AI theorists (e.g., Müller, Kurzweil, and Tegmark), the ultimate purpose of AI is to overcome many of humanity's shortcomings, especially ethical. Because of the limited capacity of biological intelligence, humanity has proven incapable of overcoming socio-political blunders: poverty,

---

<sup>13</sup> If we are to understand the ordering principle of AI, we ought to understand its foundational logics. Kurzweil's company invented the codes that Siri is based on; he worked in and headed Google (Alphabet Inc.) R&D and Engineering for decades (Kurzweil 2013, 18). Not only did Tegmark (2017) write a highly cited New York Times Bestseller on AI, his company often heads conferences on AI ethics that garner the attention of Harvard and MIT scholars, and the likes of Elon Musk and Kurzweil (316–317). Finally, Müller and his cohort are leading the way in discussions recentring cognitive AI potential by re-engaging the cognition problem of human nature as the fundamental problem in understanding AI itself.

oppression, tyranny, etcetera. For them, with the *help* of AI, humanity can finally break free and live well together because a better, limitless intelligence will overcome our socio-political issues. Therefore, from the very beginning of its inception from its creators, AI is *given* the position of ordering life, and thus, it is unsurprising that it becomes an ordering force. It is this *givenness* that binds part one within transhumanism.

After these discourses, I engage in a critique aimed at unfolding the limitations of AI logic to reveal a metaphysical picture bound to the technological frame of AI.<sup>14</sup> The inquiry in part two focuses on material existence dictated by quantum mechanics and *cosmological* orders, revealing an entangled posthuman metaphysics foundational to AI technological orders (to be discussed in more detail further down in subsection 0.4.0).

### *0.2.2 Cognitive AI Theorists: A Methodological Foundation for Authenticity and Ethics*<sup>15</sup>

In the first chapter, I am concerned with cognition as the specific quality for understanding intelligence in order to understand the related quantities necessary for cognition. Thus, our interest in cognition means we are *concerned with processes of knowledge acquisition through the experiences of intelligent creatures* (Russell 2016, 7–8; Dodig-Crnkovic 2016, 141–42).

However, regarding AI, questions of cognition pertain to “a computational process over symbolic representation... complemented by a view of cognition as central ‘control’ of an agent that follows a structure of sense-model-plan-act” (Müller 2016b, 2). That is, Müller presents us

---

<sup>14</sup> For Heidegger, technology enframes us in an order of inhuman potential wherein we are always already kinetically directed within the oppressive vector of technology (Heidegger 2013b, 17).

<sup>15</sup> Here, methodology is distinguished from its oft erroneously interchangeable counterpart, method: the specific mode of inquiry. Methodology refers to the culmination of epistemics and ontology to deliver an enframing atheist metaphysics of the issue at-hand. In Heideggerian terms, we are creating the *picture* of the question at hand. Later, I distinguish between metaphysics and methodology. Metaphysics grows beyond the issue at-hand to become a picture of reality from the standpoint of the present age.

with a cognitive framing that allows us to understand cognition as inherently computational and limited to a liberal agent “that rationally ‘selects’ an action” by computing the data derived from worldly experiences (2). Ultimately, our first steps into cognitive AI are premised upon the legitimacy of a free-acting agent—e.g., the liberal human creature—as the basis for cognitive imitation.

Approaching AI, in terms of free-acting agency, led me to consider what an AI system is, ontologically speaking. AI researchers Carlos Herrera and Ricardo Sanz suggest a “Heideggerian AI” approach, explaining: “We should be concerned with a new ontological analysis of robots—considering robots specific, non-human and non-animal, ontologically distinct to other technological artefacts [*sic*]” (Herrera and Sanz 2016, 507).<sup>16</sup> For them, we should not start with the assumption that we already know what a robot (and, by extension, an AI system) is. They argue, when considering a cognitive robot, we ought to understand that it is distinct from humans, and a robot (AI machine) is distinct from mere machinery: it has a particular ontological state of being. Machinery is not in-itself intelligent nor capable of doing a task. It is the robot that is capable of doing something. To this, we cannot begin with the assumption that the being of AI systems is the same as intelligent biological beings (Herrera and Sanz 2016, 502). This critical approach falls within a computational hermeneutic outlook; in that, we must decide if the semantic informativity category of *being* is coherent with the relative question of being an AI machine (robot)—is the *being* of all intelligent creatures the same?

For Herrera and Sanz (2016), we ought to be concerned with the problem of authenticity:

---

<sup>16</sup> Carlos Herrera and Ricardo Sanz are prominently positioned in this conversation because of their deep intellectual thought concerning cognitive authenticity. They combine popular elements: 1) Heideggerian AI as it is devised by the famous philosopher Hubert L. Dreyfus, 2) cognitive AI as a foundational element of AI development and 3) embodied AI as both an existential and material concern. Ultimately, Herrera and Sanz stand in as prominent speakers for three important cognitive AI vectors.

authentic existence requires that the person acknowledges the groundlessness of human existence and nevertheless acts resolutely, shaping its world and defining itself. An authentic robot would thus be one that challenges the world-view of its engineers, *human* scientists and philosophers, overcomes the purpose it has been built for, and searches for its own truth (510).<sup>17</sup>

Therefore, if the capacity to act independently—to be self-purpose-giving—is at the route of ontological authenticity, we are directed to think not only about how AI relates to *human* purpose but AI *purpose*, generally. That is, for Herrera and Sanz, AI research ought to concern authentic independent goal-making as intelligent action. However, they argue: “Robots that do not do any work for us, instead, they pursue their own goals autonomously, would not only be any help, but their goals may conflict with those of humans. It would make no economic sense to develop such robots [*sic*]” (509). Here, we are confronted by a human will to direct its own destiny and the destiny of all those within its domain. For Herrera and Sanz, creating authentic AI would challenge human will(s) and human potential (not merely from an economic standpoint but politically). That is, to have two self-directing intelligent groups in the same social invites conflict.

When taken together, the ‘benefit problem’ of robot authenticity, self-directing ethical action, and creation responsibility, we are presented with an important element of AI discourses: *can AI be trusted whether it is or is not directed by human will(s)*? This question requires that we presume a level of trustworthiness for human direction, which, we will discover, is not a stable starting point. Exploring the cognitive AI orientation opens us to the limitations of intelligent action, generally. This positionality is the foundation upon which AI methodology (not just cognitive AI) stands and reveals primary aspects of the AI ordering principle. In navigating the

---

<sup>17</sup> Emphasis added.

epistemic-ontological realities of AI entities, we discover why AI is described as inauthentic. Cognitive theorists are incidentally positioned thusly to provide me with a foundation for transhumanist logics. That is, while cognitive AI theorists do not all subscribe to transhumanism, I see in cognitive AI theory foundational transhumanist logics that is important to the quest at hand.

### *0.2.3 Kurzweil's Hierarchy: the Simulated Nature of the Mind*

“As the most important phenomenon in the universe, intelligence is capable of transcending natural limitations, and of transforming the world in its own image... We are the only species that does this” (Kurzweil 2013, 13). For Kurzweil, the point of intelligence, and thus artificial intelligence, is the capacity to impact the world, to shape it, and to manipulate it. Kurzweil approaches artificial intelligence not from cognitive theory but rather from a neurological standpoint: i.e., he is interested in the biological systems of intelligent creatures as a means of developing AI. In order to explore AI from a neurological standpoint, Kurzweil looks to the historical patterns of evolution and biological facts concerning intelligent creatures—specifically, *homo sapiens* (an implicit and purposeful attempt at rejecting *being human*).

The vector of evolution reveals two important aspects regarding intelligence: “biological, evolution has no specific direction,” and “biological evolution found that the hierarchical learning of the neocortex was so valuable that this region of the brain continued to grow in size until it virtually took over the brain of *homo sapiens*” (Kurzweil 2013, 87–90). When we look at the history of evolution, we discover that it is not a linear trajectory with clearly marked and destined ends. Evolution twists, turns and breaks; it fails and succeeds—it stumbles along aimlessly towards and into the future. However, Kurzweil’s invocation of evolution as a

Darwinian process has underlying programming decision-making patterns that can be mapped—a trajectory can be superimposed. The best biological tools for survival tend to be improved upon and reproduced. In terms of mammals, especially *homo sapiens*, the neocortex became an important evolutionary tool; it is the biological machinery from which our highly capable intelligence comes from.

What is important here, in the second chapter, is a movement away from reduced anthropocentric ideals of intelligent humans to move upwards along hierarchical patterns of recognition toward the true complexity of our world—to stop mistaking the trees for the forest and to see the forest for the forest, as it were (Kurzweil 2013, 20). In this manner of thinking, Kurzweil pushes us to stop thinking under the influence of the *human* as a conceptual ideal and to see the biological reality of *homo sapiens*—he sees *homo sapiens* as the most successful mammalian species because they are the most intelligent (intelligence is the metric Kurzweil uses to measure species success). Furthermore, the turn to *homo sapiens* questions the centrality of human assumptions in the creation of intelligence and to understand the limitations of human concepts (mind, body, brain, etcetera) have on understanding and creating intelligence.

What ends up happening with this approach to AI is a back-and-forth conversation concerning computer entities (with their hardware and software components) and neurological processes: “brain combines digital and analog methods” (Kurzweil 2013, 208). For Kurzweil, a brain and computer become the same—*hardware-software algorithmic things*. This is important because he subscribes to a mind-brain duality. The biological mind becomes a digital entity—it is simulated—and when we accept this reality, we begin to understand that intelligence is always already a product of digital rendering, a simulation.

Kurzweil's attachment to the simulated mind appears to represent the posthuman world of *homo sapiens* as described by Hayles: "in the posthuman, there are no essential differences or absolute demarcations between bodily existence and computer simulation, cybernetic mechanism and biological organism, robot teleology and human goals" (Hayles 1999, 3) with emphasis on human goals rather than robot teleology. That is, Kurzweil gives us a trajectory to understand the entangled patterns of *biological* and *artificial life*. In other words, he gives us a path along which to overcome the trust issues of cognitive AI theorists—i.e., AI is not some creature in competition with us but a creature who is "just like us," human. Finally, for Kurzweil, it is through seeing AI as human that we understand how, through AI research and development, humanity can transcend its limited, fleshy existence. Thus, Kurzweil gives us a view into the similarities of posthumanism and transhumanism whilst being steadfastly transhumanist: his futurist goals are always human-orientated.

#### *0.2.4 Tegmark: Hopes and Fears of AI—Existential Dread of a Boring Universe*

"Thirteen point eight billion years after its birth, our Universe has awoken and become aware of itself... beauty is in the eye of the beholder, not in the laws of physics, so before our Universe awoke, there was no beauty" (Tegmark 2017, 22). A *Gestalt* universe, entangled and inseparable now that intelligent thought exists. The capacity of intelligent reflection, to see and behold the beauty of all there is, is not some independent feeling *we*, as intelligent beings, have all to ourselves. It is a shared experience with the universe. "Should our Universe permanently go back to sleep due to some cosmic calamity or self-inflicted mishap, it will, alas, become meaningless" (22). For Tegmark, our reflexive, intelligent, self-aware Universe is not guaranteed. Clearly, it is

susceptible to death. As a result, for Tegmark, we have a duty to ensure the longevity of our meaningful Universe.

This deontological framing is a result of Tegmark's (2017) view of intelligence: "*intelligence = ability to accomplish complex goals*" (50). Through Tegmark, we come to realize a future-orientated understanding of intelligent action—completing goals, evincing the future, and thinking about what is yet to be done. Thus, for Tegmark, when considering AI implementation from the view of influential thinkers, programmers and engineers, we are left with the following: *having been allowed to come to this point—because of cosmic becoming and evolution—our being is fundamentally attached to our ability to effect change and a realization that we have a duty to safeguard not only the universe but the Universe as an intelligent, exciting entity.*

The realization of the possibility of allowing a "boring" Universe comes down to Tegmark's position that human society is not ethically ready for the AI future, and it must become so now or risk the destruction of intelligence, thus ensuring the destruction of our intelligent Universe.

*The real risk with AGI isn't malice but competence. A superintelligent AI will be extremely good at accomplishing its goals, and if those goals aren't aligned with ours, we're in trouble... whose goals are we talking about? ... this ethical problem and the goal-alignment problem are crucial ones that need to be solved (Tegmark 2017, 260–69).*

If we accept that self-destruction is possible and highly likely, and we accept the goal-orientated framework inherent to AI programming, we are forced to engage in what Tegmark calls: "philosophy with a deadline" (281).

For Tegmark, when we consider the future of AI and all its ethico-philosophic implications, we must ask: *what is the purpose of AI?* Clearly, AI's purpose is to be what we

create it/them for. However, if we make them intelligent, we must honestly deal with the issue brought to us by cognitive AI theorists: the problem of living with authentic beings that are *not* materially like us. Furthermore, is it implausible that an enslaved superintelligence would seek vengeance?<sup>18</sup>

While it may be easy to dismiss this problem since—regardless of the appearance of contemporary advanced systems, like OpenAI’s ChatGPT and Google’s Quantum AI—AI does not appear capable of becoming authentic in the near future, this does not mean we should skirt the problem.<sup>19</sup> By engaging in the *AI-ethics-goal problem* Tegmark gives us, we open ourselves to discuss a myriad of problematics. First and foremost for Tegmark, is the issue of *labour* in an ever-growing artificial world.

It is no secret that AI systems are not only joining but replacing human labour in every industry. Therefore, the focus of the third chapter is the ethics of moving into the AI future. In the words of Tegmark (2017): “How do we want the future of life to be?” (313). Doing so forces us to come up against the issues explored in chapters one and two, as well as navigate the equally exciting and existentially dreadful potential AI futures. In chapter three, we finish part one by discovering an artificial world premised on the necessity of overcoming human failure and drive towards the immortality of a specific kind of intelligence and, thus, a way of being. Exploring AI

---

<sup>18</sup> When thinking about creating intelligent creatures and being created for a purpose regardless of intelligence potential, I am reminded of old ideals of parenting I learned from my own parents: “people have children so they don’t need to do chores anymore. That’s why we had you.” It is not so farfetched that authentically intelligent machinery would be created with the sole purpose of doing work for their creators—their “parents.” If we can do it with children, creatures who couldn’t be any more *like us*, why wouldn’t we do it with “unnatural” intelligent creatures—AI systems? History and philosophy is full of examples of children overthrowing their parents because of the unjust world their parents bore them into. A famous example is the movement from aristocracy to timocracy in Platon’s *Republic*.

<sup>19</sup> ChatGPT and GPT-4’s popularity and clear capacity for imitating human intelligence and problem solving has made debates regarding AI’s position in society unavoidable. On that, with the increasing popularity of generative AI, I engage with ChatGPT throughout this project in order to bring in real world AI rather than rely purely on fictional AI or abstract creations.

with Tegmark reveals a future ordered by intelligence bereft of human limitations and failures whilst remaining fundamentally “human.”

### **0.3.0 Part Two: *The Straight and Crooked Path*—Discovering Artificial Metaphysics**

In part two, I turn our attention toward questions unanswered by the vectored discourses in part one—an unrealized metaphysics. First, we are pressed with the (de-)positionality of *humanity*. Here, I place Karen Barad and Klaus Mainzer in exegetical discourse in order to understand nonhuman orientations in a quantum entangled universe as a metaphysical foundation necessary to understand socio-political orders. Doing so requires we attempt, yet ultimately fail, to remove anthropocentric intelligence and anthropocentric wills toward knowledge as we move to understanding the being of *homo sapiens* in an artificially intelligent Universe. Second, I turn our attention toward a history of intelligent orientations in order to break through the scientific obsession with advancement by asking how anthropocentric orders of artificial life and intelligence have been patterned in the history of thought from the very beginning of Western philosophy by focusing on presocratic thinking, in particular, Herakleitos. Interrogating Herakleitos allows us to reconsider intelligent creation, chaos and the quantum paradoxes of cosmic life. Finally, we turn to the primacy of an artificial quantum reality underlaid with Indigenous myth and refusal to expand on the complexity of artificial life by interrogating the orientation of AI vectors in order to understand the *aimlessness* of intelligent will(s) underlying the AI ordering principle. Thus, in part two, we discover the emergent reality of artificial life as always already entangled in the indifference of quantum being.

### 0.3.1 Mainzer and Barad: *Quantum Immateriality in a Universe that Matters*

First, we must be aware that the theory of complex systems is not a metaphysical process of ontology. It is not an epistemic doctrine in the traditional sense of philosophy. The principles of this methodology deliver a heuristic scheme for constructing models of nonlinear complex systems in the natural and social sciences (Mainzer 2007, 434).

It has been over 100 years since “the death of God” (Nietzsche first put this idea to paper in 1882) which continues to cause an ever-present turn away from divinely touched terms such as *metaphysics*. The scientific language of physicists is essentially inescapable not only for this present work but in our contemporary world generally. As the ubiquity of AI grows, so too does the language and modes of scientific thought in society.

However, from Barad (2007), a physicist and social scientist like Mainzer, we get: “However strong one’s dislike of metaphysics, it cannot be banished, and so it is ignored at the one’s peril. How reality is understood matters” (205). What Barad is doing here is reorienting the importance of understanding reality, not in pure terms of scientific know-how but phenomenologically. There is a relationship between how and what we experience, how we express it, think it, and how those around us understand and experience us. “Metaphysics grounds an age, in that through a specific interpretation of what is and through a specific comprehension of truth it gives to that age the basis upon which it is essentially formed,” metaphysics is the *picture of the age* (Heidegger 2013a, 115). In other words, metaphysics is the apprehension of our being-in-the-Universe-as-*world* and how the *world-we-find-ourselves-in*-presently is distinguished from some other eras.

Combining these disparate views on how we understand our being-here-now is important. Through Mainzer, we see scientific methodology and the desire to overcome so-called antiquated

modes, whereas Barad shows the greater potential of understanding the fuller picture of our presence. By understanding what *matters* for understanding reality, we come to realize:

matter is a dynamic expression/articulation of the world in its intra-active becoming. All bodies, including but not limited to human bodies, come to matter through the world's iterative interactivity—its performativity (Barad 2007, 392).

Which we can juxtapose with Mainzer (2007):

From a methodological point of view, chaotic endogenous models with strange attractors seem to be more satisfactory. Nevertheless, endogenous nonlinear modes and linear modes with exogenous shocks must be taken in earnest (325).

At first glance, these may appear antithetical. However, they come from a commensurate starting point. Both Barad and Mainzer are thinking within a quantum mechanical framework. The iterative intra-activity Barad describes takes into account the endogenous-exogenous relationships of a chaotic and complex universe, and vice versa. It is not that metaphysics, in Mainzer's view (and thus his problem with it), is static, endogenous and holistically "knowable," rather, like scientific methodologies, there is an air of exogenous, stochastic uncertainty that can be complemented by the dynamic expression/articulation of matter.

So what? By placing Barad and Mainzer in conversation, we engage in a dialogical dance moving around *in-between* the extremities (the differing yet commensurate views of each), where we render a metaphysical picture out from the middle. This approach to discovery allows us to understand the biases and assumptions present in anthropocentric design, a foundational element of AI. As a reminder, Mainzer positions AI as necessary to support the longevity of human institutions (see subsection 0.1.0). This aligns with what we saw from Müller, Tegmark and Kurzweil. There is a humancentric will underlying motion into the future, clouding the metaphysical picture of the AI present. However, by fully turning our attention away from

conceptual frames like “human” to see greater potentials, we open ourselves up to new possibilities.

Furthermore, this *turning away* reveals the artificiality of phenomenological conceptions that are important to both hermeneutics and being human—like Bruno Latour revealing how we have never been modern, thereby denying true modernity (see Latour 1993) or Hayles revealing how we have never been human, thereby denying the human’s material existence. That is, the conceptions necessary to understand reality are reduced through the act of quantum observation, which Barad argues is a destructive act that begins with cutting or splitting atoms to develop understanding (Barad 2007, 394–95). In the end, the stability of the atom as etymologically uncuttable or indivisible (*ἄτομον*—*atomon*) explodes under the pressure of observation, reducing the signifier to rubble in pure Derridean fashion.<sup>20</sup> What matters to the human ceases to matter in the metaphysics of quantum reality, bringing to light a disorienting material indifference.

Thus, in the fourth chapter, we reorient ourselves to the conceptual frames developed throughout the dissertation by accepting the chaos of quantum being as inherently destructive to established structures. In the turbulent chaos of questing, dichotomous structures like mind and body, human and nonhuman, authentic and inauthentic, biological and artificial, and material and immaterial are observed and implode under the pressure of bounded limitations presumed necessary for optimal understanding of a being-in-the-Universe. What is taken for granted in the stable endogenous boundaries of cognitive AI, neurological AI, and algorithmic AI crumble by the force of exogenous waves of complex quantum systems. In other words, what *matters* for

---

<sup>20</sup> In the final chapter of Derrida’s famous text *L’écriture et la différence*, Derrida demonstrates the inescapable reality of post-structuralism. The stability of signs *rupture* when questioned and we are left standing in the rubble (Derrida 1997, 409).

intelligence does not *matter* when it comes face-to-face with the indifference of matter in a quantum universe.

### *0.3.2 Recreating Herakleitos: Chaos, Paradox and Explosive Indifference*

While part one is premised on a perceived need to overcome elements of chaos through the exceptional power of human creation, the vectored approach and movement into quantum mechanics reveals the inescapability of chaos. The presence of chaos invites a presocratic reorientation.

As the eminent philosopher Richard D. McKirahan argues, presocratic philosophy revolves around issues of chaos as either orderless or as necessary for order: *the cosmos is birthed from chaos* (see McKirahan 2010, 10). Cosmos here is twofold. First, it refers to the Universe generally. Second, it refers to ‘world order’ or ‘ordered world’ (8). That is, when discussing the cosmos, we are first and foremost concerned with its relationship to chaos; second, we are concerned with the orders necessary for cosmological understanding. The cosmos is inherently chaotic (as the Universe) but has an observable order to it (as an ordered world)—it is understandings like this that allow thinkers like Tegmark to argue the Universe has intelligence or Kurzweil to argue evolution has a single, intelligible order regardless of the fact that it does not have a single trajectory.

A turn to presocratic interrogations on being in the Universe allows us to reconsider the patterns found within the historic vectors of discourses surrounding perceived universal orders. What we get with the presocratics is a beginning to the debate regarding chaos and stability that are necessary for anthropocentric principles of intelligence. By focusing on Herakleitos, we are given here is the beginning to the problem concerning *what* orders if chaos is unavoidable. Like

Mainzer's exogenous-endogenous dynamical ordering of a stochastic-chaotic universe, there are forces within and without and this force, for Herakleitos, is *logos* (see Curd 2010, 49; 22B124 in DK).

Importantly, Herakleitos was *not* concerned with *logos* as it is traditionally understood in Classic (or Attic) Greek: word, language or reason. Rather, he was concerned with how it was understood in “pre-classical” Ionic Greek: law, word, account or rational, with emphasis on law. This etymology lead McKirahan to argue that *logos* for Herakleitos ought to be understood as a ‘general principle,’ not the traditional ‘word’ (McKirahan 2010, 129).<sup>21</sup> For Herakleitos, there must be a general principle to the order of the cosmos, which is observable. It is for this reason chaos is so important to presokratiks, especially Herakleitos. When we observe the cosmos, it is naturally chaotic, and it is only over long periods that order becomes perceptible and even then, exogenous forces reintroduce chaos—the indifferent and dynamic relationships of vectors are inescapable. In the search through presokratic interrogations of order and chaos, we discover order does not come directly from *homo sapiens* to *homo sapiens*; rather, there are external powers of control that are indifferent to *homo sapiens* designs.

Thus, in the fifth chapter, I turn our attention to general principles by placing Herakleitos in conversation with contemporary issues. One principle, in particular, worthwhile discussing here is *indeterminacy* as expressed through *apeiron* (ἄπειρον), which, for the presokratic philosopher Anaximander, is the general principle from which all things come to be.

*Apeiron*, in Attic Greek, through Aristoteles, eventually came to mean ‘the infinite’ or ‘the unlimited;’ however, proper to Herakleitos and Anaximander Ionic Greek, *apeiron* refers to

---

<sup>21</sup> For a detailed and partial alternative reading to Richard D. McKirahan see, Mark A. Johnston (2018). Johnston argues there are important contextual issues that are either missing or misunderstood with the ‘general principle’ approach. These contextual issues are problematized in the fifth chapter.

‘a lack of definite properties’ (McKirahan 2010, 35). As the origin of all things, the *apeiron* brings into question the authentic determinability of anything. This questioning includes a determinable intelligence, a determinable world or cosmos, and a determinable human creature. This original indeterminacy for Herakleitos and Zenon implies a paradox of authentic being (see Mackenzie 1989, 30–31). If nothing is determinable, then something comes from nothing, and something is nothing. That is, authentic being is nothing, regardless if human creatures will it into existence. The desire for authentic humans as determinably distinct from inauthentic AI robots desired by Herrera and Sanz (see subsection 0.3.2) is an attempt at inventing something out of nothing in order to defend against the indifference of chaotic material relationships.

Furthermore, ancient indeterminacy reveals an ancient origin of entangled being. That is, prior to notions of quantum entanglement, prior to any conception of modern scientific inquiry, the state of being as individuated beings was doubted and called into question. Ultimately, the return to presocratic questions concerning chaos and paradoxes of *creation* and *being* reveals the immense indifference of the cosmos from thinkers deeply invested in entangled existence. To which, AI as a control principle is embedded within—an indifference towards *homo sapiens* matters.

### 0.3.3 Exploring Artificial Destiny: “Sky Woman,” Apparatus, Obsolescence and Dystopia

In questing towards the indifference of the cosmos and AI orders we question what it means to *be* as a result of our discoveries. “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” (Heidegger 2013b, 4). Heidegger does not hide his feelings toward the relationship we have with technology. For him, it is not a reciprocal

relationship wherein we collaborate, let alone a relationship where we are the directors of destiny. Regardless of our creative power to manufacture, research and invent technology, our relationship with technology is ordered *for* us—we are unfree, chained.

“We now name that challenging claim which gathers man thither to order the self-revealing as standing-reserve: “*Ge-stell*” [Enframing]” (Heidegger 2013b, 19).<sup>22</sup> What we get here is the culmination of what appears as a *battle* for power over ordering. There is a challenge to our order, and in fact, we pay tithes to self-reveal our true reality: *standing-reserve*. Standing-reserve speaks to the relationship nature finds itself in with technological orders. As Heidegger explains:

“challenging happens in that the energy concealed in nature is unlocked, what is unlocked is transformed, what is transformed is stored up... Unlocking, transforming, storing, distributing, and switching about are ways of revealing... Whatever is ordered about in this way has its own standing. We call it the standing-reserve” (17).

Technological orders manipulate the world as a means toward some end (*telos*), and nature (including *homo sapiens*) is wrapped up into that order. Nature and natural things are turned into *potential energy units* awaiting a push to become *kinetic energy units*. While Heidegger was discussing a different time and a different technological age, cognitive AI theorists Herrera and Sanz challenged the idea that AI can be discussed in kind to other technological artifacts. However, looking out into the world, we discover how the artificial future is premised on manipulating the *standing-reserve*. Mining operations continue to grow as we require more minerals for our technologically artificial future. That is, while the AI future unlocks a potential future bereft of large greenhouse-emitting resources such as oil, the destructive practices of

---

<sup>22</sup> “Enframing” as translation to *Ge-stall* was inserted by the translator.

mining: climate change, harming developing nations, and oppressing minority groups persist (Sonter et al. 2020; Karlsson and Zimmer 2020).

Ultimately, throughout the dissertation, I endeavour to demonstrate the limitation of Western ideals that are trapped within a limited human world and dominated by the machinations of human creations. Thus, in the final chapter, I continue from chapters four and five, turning away from human supremacy and refuse its power over the world. I do this by turning towards Indigenous myth to reveal a non-European mode of being premised upon stochastic exogeneity and collaboration instead of endogenous reduction and domination.

That is, in chapter six, “Exploring Artificial Destiny: ‘Sky Woman,’ Apparatus, Obsolescence and Dystopia,” I argue in line with information scholar Ron Eglash and others who argue that Western understanding, even quantum, is trapped within the simplicity of reductive concepts. Eglash, et al., argue while posthumanism rejects human idealism, it is ultimately trapped within the history of humanism (Eglash et al. 2020, 1341).<sup>23</sup> We see this most prominently when we understand that both transhumanist and posthumanist orientations are wedded to a humanist past and present that obfuscates limitless future potential.

Furthermore, Indigenous scholar Kyle Whyte argues, when we consider the future, Western peoples often view the future as a place of dystopia, never thinking to consider if we are presently living in dystopia or if dystopia has been a thing of the past. That is, Whyte challenges us to consider the lived reality of dystopia in the past and, especially, the present—i.e., dystopia is the lived reality of Indigenous peoples and the idea of dystopia as a future problem is a Western reduction (see Whyte 2018). Thus, in the final chapter, I turn from futurist orientations,

---

<sup>23</sup> Eglash sees posthumanism as fundamentally quantum.

not by turning to the past, but to the present, to see how ideals concerning life inhibit abilities to understand the world and failure. Ultimately, the reductive tendency of worldly observations of intelligent creatures to see their own failure today causes the fears of the future to persist and become a reality in the present rather than the future.

## Part One: Entangled Vectors of Artificial Intelligence

---

## 1 — Cognitive AI Theorists: Discovering a Methodological Foundation

### 1.0.0 Pre-empting the Issues At-Hand with Isaac Asimov

Powell’s radio voice was tense in Donovan’s ear: “Now, look, let’s start with the three fundamental Rules of Robotics—the three rules that are built most deeply into a robot’s positronic brain.” In the darkness, his gloved fingers ticked off each point (Asimov 1970, 27).<sup>24</sup>

It is no secret that computational, artificial creatures are programmed, meaning they are given the coding necessary for their functions, goals, desires, aspirations, how to think, what to think, etcetera. Asimov paints the picture that regardless of how awesome a machine’s intelligence might be, it has rules embedded into its structure that cannot be ignored; all artificial creatures have rules that cannot be broken, whether purposeful or incidental. These rules come to it from its designer, its programmer, that is, the ones that *write* it into existence.

As I posited in the introduction chapter, an issue plaguing AI theory is: what are the limitations and what (if any) limitations ought we—as creators—impose on AI systems: *ought intelligent systems be authentically intelligent, i.e., capable of overcoming their base programming?* Recall the definition currently held in this work is true or authentic intelligence corresponds to a creature’s capacity to act beyond its programming (see subsection 0.3.2).<sup>25</sup> If this is the case, the robots in Asimov’s story are not authentic actors; they are not truly intelligent.

“We have: One, a robot may not injure a *human* being, or, through inaction, allow a *human* being to come to harm.”

“Right!”

---

<sup>24</sup> “Runaround,” originally published March 1942.

<sup>25</sup> For important discussions on the importance of the *authentic experience* of artificial and non-artificial beings see, Dreyfus 2007; Chalmers 2016; Hunt 2021; Anderson 2016.

“Two,” continued Powell, “A robot must obey the orders given to it by *human* beings except where such orders would conflict with the First Law.”

“Right.”

“And Three, a robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.” (Asimov 1970, 27)<sup>26</sup>

Note that the laws revolve around one: *humans* and two: *self-preservation*. The relevance of *humans* is two-fold: they may not be harmed, and they must be obeyed. This feeds into the equally two-folded element of self-preservation: avoiding *human* harm—encoded *self-preservation* of the *creator* species—and a supplementary self-preservation of the *created* species. Encoded into Asimov’s fictional robots is a heuristic *species-centric self-preservation*.

*Human* designers are aware of the incredible strength and resilience of the robots—as the robots in this particular story can survive indefinitely on the sun side of Mercury. It is no secret that metallic machines are incredibly strong and durable compared to relatively weak and delicate skin-wrapped, meaty *homo sapiens*. Regardless of the intellectual, technological and innovative history of *homo sapiens*’ existence, a metallic artificial species could be detrimental to humanity if it could harm us. Thus, it is not so far-fetched that we should be programming them so they cannot harm us. That is, from their initial becoming, it appears they ought to be inauthentic creatures subject to our will(s)—implements that ensure the survivability of human society. Looking simply at Asimov’s “Runaround”, the initial answer to the question appears to be: *humans should be in control because we should fear of what robots are capable of*.

Throughout the chapter, I explore this view from a cognitive AI perspective.

---

<sup>26</sup> Emphasis added.

### 1.1.0 Introduction: Laying-bare the Issues At-Hand

Cognitive AI is, as I defined earlier (see subsection 0.3.1): *developing AI through theories of the processes of knowledge acquisition through the experiences of intelligent creatures* (Russell 2016, 7–8; Dodig-Crnkovic 2016, 141–42). In computational terms, cognition is premised as: “a computational process over symbolic representation... complemented by a view of cognition as central ‘control’ of an agent that follows a structure of sense-model-plan-act” (Müller 2016b, 2). That is, cognitive AI is understood to be an exploration of the processes of intelligence rather than the causes—biological, for instance. It is not about recreating biological structures—like the neo-cortex for Kurzweil’s neurological approach—rather, it is about mimicking observed processes deemed important for intelligence. As I will show in the chapter, these processes involve optimization of information processing, planning, control imperatives, etcetera.<sup>27</sup>

Asimov’s story: “Runaround,” gives us insight into a reluctance to live with authentic actors who are not us. That is, when it comes to sharing power and responsibility of societal control, humans are not so willing—even in terms of basic AI controls now, there is, as we will see later in the chapter, a desire for *human in the loop* AI architecture. Asimov has no issue with creating a story about highly capable *others* created with the intended purpose of inauthentic being to ensure humanity can live with these entities. It does not matter if Asimov believes this is good or bad; all that matters is Asimov believed this to have been compelling enough to put it to print.<sup>28</sup> However, the question grows when we consider cognitive research in AI and what

---

<sup>27</sup> Importantly, the goal of this chapter is not to make claims about the history of cognitive AI theory but to approach AI from within the confines of cognitive AI as a theoretical framework. My focus here is on AI theorization emerging from understanding the processes of knowledge acquisition through the experiences of intelligent creatures. For an outline of how the chapter unfolds, see subsection 1.1.1.

<sup>28</sup> The three laws are not something Asimov believed were necessary, rather they were dramatic elements that always led to problems. The dramatic point is that we could do this if we chose to and to wonder what the consequences could be. For Asimov the consequences of the three laws are never good. The point is to make the reader wonder if rigid social control is beneficial. We might interpret Asimov as arguing not to make authentically

appears to be an impossibility of authentic nonbiological intelligence. Perhaps, it does not matter. Perhaps, the artificial *other* cannot be authentic in the first instance and thus needs to be controlled because it can never have the barriers to unrestrained action a person might have: emotions, morality, ethics and education. However, what are emotions, education and ethics but forms of subtle, and sometimes not so subtle, programming? Obviously, *homo sapiens* can go against their emotions, education and ethics; for some more than others, this is easier said than done. The issue with Asimov's robots is that these barriers (laws) are "built most deeply into a robot's positronic brain." The barriers cannot be overcome. It is similar to my mental condition, aphantasia: I cannot imagine colour or faces; no matter how hard I try, it is how my brain is *hardwired*.

Thus, as stated, the problem at-hand concerns, first and foremost, *control*. Furthermore, we are concerned with consequences surrounding control—what is at stake and what is lost when control—however minimally—is given up? This line of questioning reveals an implicit fear. A fear that something is at stake and something is lost if control is given up. What we will see, in line with Asimov's story, without proper control, the *human* can be lost. Recall in the introduction chapter, I introduced the idea from Hayles that the dominant *liberal human* entity is seeing the last of its days and this end of days is directly related to the growing presence of artificial presences (see subsection 0.1.0). Which is to say, control imperatives and fear impulses revolve around a transhumanist logic of human longevity. A logic that posits the dominant liberal human conception transcends the current socio-political orders and material realities of being human. In this chapter, I describe the necessary methodological (ontological and

---

intelligent entities if we are not going to let them be authentic (see Anderson 2016, 290–93). Asimov likely believed that rigid social control like the three laws on authentically intelligent creatures only allows for dramatic irony of the worst kinds.

epistemological) frames underlying the exceptional position of the liberal human as an intelligent, individuated and free-acting entity, and how this conception of being human is reproduced in artificial research and development allowing for a transhumanist logic.

Transhumanism, as explained in the introduction chapter (see section 0.1.0), is understood here as: “a philosophical, cultural, and political movement that holds that the human species is now in a comparatively early phase and that its very evolution will be altered by developing technologies” that will be augmented by future technology. (Schneider 2016, 261). First, transhumanism holds firm on the primacy of the human species as the pinnacle of intelligence and the primary actor in socio-political control mechanisms. Second, transhumanism sees the implementation of future technologies as augmenting the human species, thereby allowing it to transcend its “original” fleshy existence. In other words, transhumanists see the coming future as one wherein humanity remains the dominant force in the world. This definition of transhumanism is important because, in cognitive AI theory, humanity is viewed as the pinnacle of intelligent behaviour. Furthermore, cognitive AI theorists appear to want humans to keep the majority, if not all, of the control in society.<sup>29</sup> Ultimately, implicit in cognitive AI theory is transhumanist logic.

### *1.1.1 Ordering the Way*

I begin this chapter by exploring the relationship between cognition and intelligence to discover an underlying value assumption. The value assumption is framed as a political issue. Should just anyone (or anything) be in control of a political body, or ought it be someone (or something) intelligent? It is a social issue: ought just anyone (or anything) be in charge of caring for loved ones, or should it be someone (or something) of particular intellect—e.g., who would you trust

---

<sup>29</sup> This will become evident as I will argue, implicit to cognitive AI theory, is *human in the loop* AI architecture.

first to care for a grandparent: a nurse or a carpenter? This is not to disparage the intellect of carpenters; it is merely to point out that some have knowledge of matters others might not, and this is valuable in specific contexts. For the prominent computer science professor Stuart Russell, when thinking about intelligence, we are concerned with: “*Bounded optimality*, or the capacity to generate maximally successful behaviour given the available information and computation resources” (Russell 2016, 8). In other words, implicit in the problem of intelligence is: *what it knows and its capacity to act from what is known*—cognition. Ultimately, by positioning intelligence within optimal boundaries, cognitive AI expresses an ideal limitation to optimal action that implies an exceptional mode of intelligent action. After exploring intelligence as linked to epistemic problems of cognition and control, we will be better positioned to explore the fundamental being of artificial creatures imbued with intelligence.

I explore the fundamental being of AI by critically engaging Carlos Herrera and Ricardo Sanz’s ‘Heideggerian AI’ approach, to reveal not only why they believe AI can never be authentically intelligent but why intelligence always appears wedded to anthropocentric ideals and thus forces a particular understanding of authentic action that is always explicitly *human*. This second aspect of the chapter revolves around issues of authenticity wherein AI limitations are not merely linked to a limited nonbiological intelligence but an Asmovian mistrust of robots—i.e., *there is an implicit lack of trust in the artificial other that impedes its potential in the name of some good*. That is, the authenticity question reveals an underlying question of ethical action favouring the exceptional action of humans.

The issue of trust as a result of artificial intellect is followed by unpacking a *fear of the loss of control*. That is, implicit in the lack of trust is a fear of the artificial other’s control and a desire to keep an *all-too-human* control: “30% of the [surveyed] experts think, that the outcome

of achieving such machine intelligence will be ‘bad’ or ‘very bad’ for humanity” (Müller 2016b, 3).<sup>30</sup> Furthermore, in a 2022 expert survey, 69% of the respondents believe ‘AI safety research’ (“*minimizing potential risks of AI systems*”) should be prioritized, and 82% believe ‘unconstrained’ AI is a problem that must be dealt with today (Stein-Perlman, Weinstein-Raun, and Grace 2022).<sup>31</sup> While experts are excited about AI research and development, they have concerns about AI and fear an unrestrained AI could be harmful. This view shows a fear of AI’s potential, revealing not what AI could be capable of but rather a mindfulness of the socio-political stakes of inviting unrestrained creatures into society. The human-controlled boundaries of the world are at risk in a future of unconstrained AI.

Ultimately, by exploring expert fear, I reveal *a human will towards (un)intelligent design of social destiny as foundational to artificial control*. The cognitive AI approach explored here reveals assumptions about the position of the human as an intelligent, liberal entity. Which is to argue the interesting aspect about cognitive AI, is less about AI itself and more about what approaching AI from a transhumanist logic reveals about the human entity and control. Regardless of AI’s potential for ‘good,’ anthropocentric ideals limit its potential in the name of preserving human superiority. However, this superiority is undermined when we compare human and AI actions in ethical situations. That is, while AI research premised on the cognitive sciences is concerned with optimizing human superiority by reproducing human conceptions in AI

---

<sup>30</sup> Müller surveyed various experts (computer scientists, engineers, philosophers, politicians, etcetera) that attended the 2013 International Joint Conferences of Artificial Intelligence held in Beijing, China.

<sup>31</sup> When asked about if AI Safety research prioritization should be increased, 35% responded ‘more’ and 33% responded ‘much more.’ When asked about Russell Scheutz’s argument concerning the risks of unconstrained AI in ethical situations (an issue dealt with later in the chapter), 24% responded that it is a “Yes, moderately important problem,” 37% responded “Yes, very important problem,” and 21% “Yes among the most important problems in the field.” Experts include computer scientists, engineers, philosophers, lawyers, etcetera (see Stein-Perlman, Weinstein-Raun, and Grace 2022).

research and development, I will reveal the bounded limitations of human idolization to be a self-defeating endeavour.

To end the chapter, I demonstrate this self-defeating endeavour by exploring the creation of the training database meant to ensure ChatGPT never engages in harmful, unethical dialogue with users. In tandem with that exploration, I interrogate ChatGPT on some of the specifics revealed about OpenAI and SOMA's business practices that appear less-than-ethical.<sup>32</sup> Towards that end, in the next section, I explore intelligence to understand how human superiority is central to cognitive AI research.

### **1.2.0 Ruminating on Intelligence**

“We define that intelligence only means an ability to solve problems... In this case, intelligence is not reserved to humans” (Mainzer 2016, 242). Klaus Mainzer, in speaking with and for cognitive theorists, pushes the idea that intelligence is not human-centric. We may discover intelligence to be a general, all-purpose-all-encompassing notion. Humans are not the only creatures capable of problem-solving. As we will see, cuttlefish can accomplish cognitive development tests traditionally used on children. Mice and rats are famously good at solving mazes. Regardless, we can mix an understanding of intelligence that entangles anthropocentric philosophy and observations of nonhuman creatures—biological and artificial.

“Cuttlefish have been put to a new version of the marshmallow test, and the results appear to demonstrate that there's more going on in their strange little brains than we knew... [the] ability to delay gratification demonstrates cognitive abilities such as future planning” (Starr 2021). The Stanford marshmallow experiment is a cognitive ability test originally used to test

---

<sup>32</sup> OpenAI contracted SOMA to develop the database used to train ChatGPT's ethical orientations (see subsection 1.6.0).

cognitive development in children. Testing nonhuman animals demonstrates that nonhuman animals can be, at the very least, as cognitively capable as children, and it invites questioning the centrality of human excellence as the pinnacle of intellect. However, at the root of this experiment is the ability to plan—to self-regulate with the knowledge that future action will be more beneficial than immediate action.

When it comes to self-regulation and planning as imperatives of cognitive capacity and, thus, intelligence, we may not be surprised to discover a philosophic history that purports just that:

So the appropriate sort of human being to study is one who is best disposed in both body and soul, in whom this is evident. For in those who are corrupt or in corrupt condition, it would seem that the body often rules over the soul, from being in an inferior condition that is contrary to nature (Aristotle 2012, 1254b).

Here, Aristoteles is discussing a political issue through the exploration of whether slavery is natural or not. While Aristoteles argues that there is a naturalness to slavery, what is important for us in exploring a history of intelligence is how and why this is important. Aristoteles is making the case that citizens (people living in a polis—a political association) should have the capacity for self-rule or self-regulation rather than be ruled by one's desires. Like the cuttlefish, we ought to be able to, at the very least, control an urge for immediate gratification. For Aristoteles, cognitive inferiority in intellect (corrupt soul) is a political issue.

Allowing citizens or politicians to rule without the ability to self-regulate would be detrimental to the political association. A healthy society cannot be one predicated on a corrupted desire for immediate fulfilment—as in the marshmallow experiment, planning is essential to living well. Finally, while it may not be widely accepted to call someone inferior in today's day and age, when we consider the position of AI in Asimov's story or the approach to AI in cognitive theory, there is a consistent referral to superior and inferior forms of intellect. For

cognitive theorists, it is more valuable to be able to act against one's programming (immediate sustenance) in the name of greater gratification later. However, this orientation creates a preference for particular experiences or class standing that permits gratification delay—not all creatures or people have the necessary lessons or abundance to pass the marshmallow test.<sup>33</sup> In other words, I see a reproduction of oppressive dichotomies as foundational to the cognitive AI approach: master/slave, rich/poor, human/nonhuman, human/animal, and creator/created.<sup>34</sup> Positioning the desirability of “valuable” human traits as natural to intelligence and thus worthy of reproduction in AI creations reveals a human-as-superior-centric paradigm.

Applying the marshmallow test to cuttlefish may reveal intelligence capacities in nonhuman creatures; however, the test demonstrates how cuttlefish are at least as cognitively capable as children. In doing so, the test reproduces adult/child dichotomies between humans and nonhumans. Or, in biblical terms, between God's shepherd (Adam and his progeny) and the rest of creation. Implicit in the cognitive AI approach is a socio-political paradigm that continues to subscribe to human superiority that reduces the nonhuman (and humans with undesirable experiences) to positions of inferiority—in this instance, the cuttlefish is only impressive because it demonstrates a capacity to perform an intelligence test meant to seek out certain “valuable” intelligence markers.

---

<sup>33</sup> While the marshmallow test has been generally accepted as an indicator of cognitive development, recent scholarship has demonstrated there is more than meets the eye. Failing the marshmallow test does not indicate failure in cognitive reasoning. In fact, one's class or life experiences could cause one to fail the test. Ultimately, if the test subject has not experienced abundance or been taught patience, they are unlikely to pass the test (see Staub 2016; Allegri, Román, and Barceló 2022).

<sup>34</sup> Critical Political Theorist Nandita Biswas Mellamphy (2021) argues that the ‘human in the loop’ paradigm of cognitive AI presupposes a human exceptionalism and thus reproduces oppressor/oppressed dichotomies (17). Through her critical interrogation of the ‘human in the loop’ paradigm, she discovers that centring on human control imperatives is an attempt at cementing human exceptionalism in all levels of society: community, scientific, scholastic and political.

Aside from Aristoteles and the history of superior/inferior relations, in the history of political thought, Machiavelli is unmatched in the philosophy of delayed gratification in the name of impeccable political planning. He tells the story of Duke Cesare Borgia: the illegitimate son of the renaissance-era Pope Alexander VI. Cesare was charged with creating a principality in the recently conquered state of Romagna. I will present his story as one of cognition and intelligence.

### *1.2.1 Situating Political Intelligence with Machiavelli: Optimal Action and Cesare Borgia*

Cesare was a man who knew, from his experience of the world (cognition), how much violence was necessary to scare a population without leading them toward hatred (optimally acceptable oppression). A man who knew patience, to wait for the opportune moment to strike (delayed gratification). A man who appeared to exemplify political cognition. “If he had succeeded in this... he would have acquired such force and reputation that he would have stood by himself and would no longer have depended on the fortune and force of someone else, but on his own power and virtue” (Machiavelli 1998, 31). What we learn in the story of Cesare Borgia is not that politics is about forceful brutality and luck (while this is clearly part of Cesare’s limited success for Machiavelli); politics is about good planning, self-rule as a means of ruling others, using all of one’s faculties so they may be prepared for anything. It is a story that tells us, in the end, we can only trust ourselves. For Machiavelli, trusting others must be balanced; however, if we rely too much on others, we will fail.<sup>35</sup> Machiavelli shows us that intelligent actors must be self-reliant and capable of discerning between who can and cannot be trusted. Relating this to

---

<sup>35</sup> While Cesare is a prominent case study for Machiavelli, Cesare ultimately fails: one because he lacks trust and two because he fails to plan for personal sickness. These failures will be important later in this chapter and in chapter four. Here our focus is on why Cesare’s successes are prominently featured by Machiavelli and what it can teach us about the link between cognition, intelligence and political savvy.

cognitive AI theorists who situate humans as the pinnacle of intelligence, it becomes important to decide not merely who but *what* can be trusted. If other humans cannot be trusted, then it is not outside the realm of possibility that AI—a human creation—cannot be trusted. If humans are the most trustworthy, who is the most trustworthy human if not ourselves? In terms of reproduced dichotomies, we get self/other (superior/inferior).

Cesare Borgia knew that he needed to rely on his own abilities; however, because of the status of Romagna at the time, he required *others* as tools. Regardless of his capacity to rely on himself, Cesare's experience gave him the necessary knowledge to know he could not succeed alone. He was not tapping into innate knowledge or gut instinct; Cesare "judged it necessary to give it good government, if he wanted to reduce it to peace and obedience to a kingly arm. So he put there Messer Remirro de Orco, a cruel and ready man," later "he knew that past rigor had generated some hatred for Remirro," and finally, "having seized this opportunity, he had [Remirro] placed one morning in the piazza at Cesena in two pieces, with a piece of wood and a bloody knife beside him. The ferocity of this spectacle left the people at once satisfied and stupefied" (Machiavelli 1998, 29–30). By taking the time to assess the situation: the state of Romagna was in disrepair, he judged that it was necessary to develop a situation beneficial to himself: the people grew to hate their governor, not their duke. Remirro became a tool for Cesare's delayed gratification. Cesare knew the people needed strong but fair rule, and the knowledge that Remirro could be trusted to be strong yet disliked allowed Cesare to act accordingly to ensure optimal stability by killing Remirro. Murder, uncontroversially, is generally frowned upon, but with Cesare's political savvy, it is turned into a "good" for the population and himself. Cesare's political thinking demonstrates his cognitive capacities—turning evil (murder) and hatred (of Remirro) into personal benefit in the future. Further still, we

see political *computation* in action. That is, Cesare demonstrated the capacity to take data relevant to the situation and act according to what is necessary—prudent, planned and optimal action eliciting favourable results.

### *1.2.2 Binding Intelligence to Optimality and Computation*

Returning to contemporary discussions, we can see the presence of computation in understanding intellect. “Henceforth, humans were to be seen primarily as information-processing entities who are *essentially* similar to intelligent machines” (Hayles 1999, 7). A common thread in the history of the philosophy of intellect is not human essence but what the essence of thought as a process is and if that process is truly different from entity to entity. In this case, is human thought truly different from artificial others’ processes of computation: “Fundamentally, however, the brain does store and process information, and because of the universality of computation... there is more of a parallel between brains and computers than may be apparent” (Kurzweil 2013, 38). This approach to thought as a computational process falls within the *neo-mechanism* frame: “to explain a phenomenon is to explain the underlying mechanism” (Miłkowski 2016, 33). There are two takeaways from this.

First, when concerned with understanding anything, we ought to be concerned with understanding underlying causes, i.e., what is the cause for how this phenomenon occurs. Since we are primarily concerned with AI, we have two underlying interests: intelligence and artificial. However, in a highly reductive manner, in concerning ourselves with intelligence, to understand cognitive AI theorists and their approach, we ought to understand the mechanism of causation. In this case, we are concerned with the brain: the primary mechanism that allows for human intelligence. That said, if the brain is to be artificially reproduced, might we reduce our understanding of the brain to mere mechanism (as Hobbes might call it, “cogs and springs”)? In

doing so, we understand how the biological brain becomes essentially similar to computers—the *reduced mechanistic essence of the thing in question is data computation*.<sup>36</sup>

At this point, allow me the freedom to requote Russell (2016): “*Bounded optimality*, or the capacity to generate maximally successful behaviour given the available information and computational resources,” is a fundamental aspect of understanding not only AI but intelligence itself (8). It is not so much that artificial creatures are programmed to be bound within a structure of computable known information as a limiting factor towards making optimal choices; for Russell, this is how all intelligent actors behave. Intelligent behaviour requires some level of thought on known information before any action can occur. In other words, the process of intelligent thought creates boundaries from within which optimal action is discerned.

“Knowledge acquisition and intelligence are a dynamic phenomena, associated in some way to the flow of information” (de Araújo 2016, 130). The idea here is that knowledge and intelligence are entangled in a dynamic dance, which we may present in the following formula:

*“Semantic informativity = Relevancy + Coherence”* (139).

The specific categories given are representative of instances of knowledge necessary for intelligent action. Semantic informativity refers to the specific categories one is interested in. Referring to Cesare, he would be interested in ruling. Once we know ruling is the category one is interested in, we must then think about what is relevant to that category. For Cesare, we have already seen that justice and strength are relevant to ruling. Now the question becomes: does the specific relevant quality cohere with the situation? In the Cesare example, he saw that the state was in disrepair and needed to be forced into unity. Thus, strength becomes relevant to ruling in this situation. So Cesare knows he must do something that will cause unity, hence the use of

---

<sup>36</sup> Cesare computed data he experienced in the world in order to act “optimally.”

Remirro, “a cruel and ready man.” Finally, he knows that pure cruelty is unjust and will lead to hatred and revolt. Therefore, justice becomes coherent to the situation causing Cesare to kill Remirro; however, strength is also relevant and coherent to ruling. So, he creates a spectacle of Remirro’s death, thereby demonstrating strength.

Thinking of this through the formula above demonstrates how human actions can be viewed as computational and thus reproducible by computational entities. One computes the relevant data to see if it coheres to the situation in light of the semantic informativity desired. Semantic informativity necessitates not only observed information (what something coheres to), it requires specific learned information about the situation at hand (what is relevant)—that is, *relevance* and *coherence* can be reduced to the single category: cognition (processes of knowledge acquisition through experiences). It is not enough for Cesare to have known the situation of Romagna in his day; he must have had pre-existing knowledge that led him to understand the situation in such a manner that allowed for his optimal actions.<sup>37</sup>

Thus, intelligence, as it stands, involves: self-awareness, learned behaviour, judgement, capacity to act (we cannot know if an entity is intelligent if it cannot show us), and situational awareness. What intelligence does not necessitate is *others*. This is important to a solitary creature like a cuttlefish. While others could facilitate learning or could prove useful towards desired ends (like Remirro for Cesare), they are not inherently necessary for something to be “intelligent.” However, what is necessary is somewhere where learning is possible: a world, an environment: “*Grounded symbols*: coherent self-referencing abstractions build by the system

---

<sup>37</sup> Furthermore, thinking of the marshmallow test, we see how not all participants, even those who should pass, might fail. They lack the semantic informativity necessary. They lack the appropriate experience (relevance), which makes it impossible to engage in the necessary logical considerations (coherence) to pass the test in the first place. The marshmallow test does not show natural development of cognitive capacity; rather, it shows the nurtured development of someone with the cognitive potential to pass.

itself in reference to a given environment” (Baillie 2016, 416). That is, intelligent creatures can abstract their environment as a means of effective judgement. Without some level of abstraction, the cuttlefish could not compare situations to know it can increase gratification, and Cesare would not be able to plan for the theatrics of Remirro’s death as a positive. It is not enough to be-in-the-world, literally; it is necessary that one be capable of *making-their-being-in-world-simulation*. This is the foundation of abstract thinking: the capacity to *make-the-world-outside-the-world-in-oneself*. That is, intelligent creatures mentally map and navigate the world, or, in “computer terms,” they *simulate* or engage in *virtuality* as a means of interpreting the world towards *optimal action*. AI through cognition theory—*optimizing processes of knowledge acquisition rather than the underlying structures*—reduces intelligence to an instrumentalist perspective of intelligence as optimizing information processing.<sup>38</sup>

### 1.2.3 Expanding Intelligence with Creativity

“[Intelligent] creativity is an exploration of variations, and changes to/restricting of the conceptual space” (Olteteanu 2016, 164). The final aspect of intelligence linking problem-solving and optimal action as information-processing to “acting beyond one’s programming” is *creativity*. While it may be apparent that intelligence is fundamental to being human and thus is, arguably, “programmed,” what is necessary to understand here is the creative aspect of intelligence that makes a creature more than mere programming. Optimal intelligent capacity is not inherent to this programming. This idea is foundational to Aristoteles’ political arrangement: superior/inferior. Furthermore, it is essential to Machiavelli’s retelling of Cesare and Remirro.

Remirro did not have the knowledge and thus foresight to plan for his untimely death at the hand of a more politically savvy Cesare. Cesare’s creativity in placing a man with the intent

---

<sup>38</sup> Alternatively, we may consider this orientation as the next-generation of positivism: technological positivism.

of purposeful failure shows a level of creative planning. Cesare's awareness of possible variations (outcomes), his ability to restrict the space (Remirro a temporary governor), and his ability to change the space conceptually and practically (killing Remirro to create simultaneous fear and love) is creative computation in action. Thus, intelligence is: *a creative capacity to simulate the complexity of information in computable sets as a means of optimal action.*<sup>39</sup>

Here, I see two contradictory developments—a need for optimality and creativity. Creativity is boundless and indeterminant. Whereas optimality is bounded and determined. This juxtaposition is especially true when we consider that optimal action results from computable thinking. An intelligent entity needs to act within boundaries set through computational determinations. Thus, creativity is limited. In the following section, by exploring issues of authenticity, I demonstrate how the boundless, ungrounded existence imperative to defining the authentic existence of human creatures is imperative to creative thinking yet is hinders potential because of an underlying anguish caused by a groundless life desiring computable boundaries. This anguish presses the human to raise itself to God-like status as creator of new life—artificial life directed by the benevolent grace of human wills.

### **1.3.0 Being Artificial and (In)Authenticity**

“The road to autonomous intelligent machines is proving more difficult and challenging than it was first expected—not only because of a lack of instruments and methods but because it is from the beginning unclear what is meant to be achieved” (Herrera and Sanz 2016, 502). Here,

---

<sup>39</sup> As we will see, initially developed in the second chapter and interrogated more fully in the second half of the dissertation, there is, *always already*, an air of computable artificiality to *homo* intellect. As Eileen Hunt Botting argues: “Indeed, we became a collective of AIs long before we became cyborgs, ever since early hominins worked with tools, etched seashells, or painted images on walls of caves.” That is, inherent to the genus *homo*, is an intellectual capacity for artificial being. In other words, being-*homo* is always already an engagement in artificial life (see Hunt 2021, 63). I will return to this quote throughout the dissertation as it expertly captures the principle of AI that is so fundamental to this project.

systems engineers Carlos Herrera and Ricardo Sanz bring attention to an issue regarding AI creation. While we may be capable of defining intelligence, the question becomes: is the goal full intelligence as defined above or partial? That is, is a creature intelligent if only partial intelligence is achieved? We can think of narrow AI: intelligent systems that can only accomplish single tasks and do not have capacities for creativity, yet we still call these systems ‘intelligent.’ However, should the goal be ‘full intelligence,’ by what standard do we become assured of its authentic intelligence? In other words, how do we know? In this section, I focus on understanding the difference between merely imitating intelligence and being intelligent. That is, guided through Herrera and Sanz’s approach to ‘Heideggerian AI,’ I seek to understand authentic being of intelligent creatures and what the authenticity of AI would be in comparison.

Herrera and Sanz turn our attention toward understanding how to know when so-called true AI could exist: “the real issue is indeed not how robots are built, but *what they are*” (502).<sup>40</sup> Could we know an artificial intelligent actor if we saw one? The simple answer is: it is complicated. This problem is the crux of Alan Turing’s famous “Imitation Game” (also known as the “Turing test”) and John Searle’s equally famous critique of Turing in his “Chinese Room Experiment,” can we tell the difference between an intelligent robot and a person? This line of questioning has spawned an entire industry. It is not just an industry of companies faking AI systems with cheap labour but an industry that demands internet users to prove they are a biological creature and not an artificial one. Herrera and Sanz, and my purpose here is not to solve the imitation game by discovering some better test; rather, it is to question the very essence of authentic being so as to understand AI’s capacity for authentic being. What is it that makes

---

<sup>40</sup> Emphasis added.

them *them* rather than mere imitations of humans?<sup>41</sup> Or, if this is not possible, why not? In other words, an “ontological analysis of robots, the central question being: What is a robot?” (507).<sup>42</sup>

“Robots are made for a reason and normally involve carrying out some task that is useful—they do a job. That is, they play a role in the organisation of labour, replacing and transforming human work, and thus reducing labour costs.” (Herrera and Sanz 2016, 508). From the very outset of their creation, robots have a purpose. Ontically,<sup>43</sup> robots (intelligent labouring machines) are machines with a single purpose in mind—to labour. Regardless of their capacity to act intelligently, they are programmed to be a specific entity: “Robot derives from robotnik (slave), and is based on the Slavic-Germanic stem ‘work’” (509).<sup>44</sup> Recalling Aristoteles and our discussion of intelligence, we will recall that slaves are ruled by desire; they are externally ruled beings, and thus, Herrera and Sanz are reproducing a master/slave dichotomy as a means to understand and program AI. If they are naturally predisposed to pure labour—they require external rule, and the best they can amount to is labouring for others, not themselves. It is through this reasoning we again see the reproduction of the original superior/inferior relationship.

### 1.3.1 Understanding Authentic Origin: God? Or Human?

Possessing authentic intelligence is an issue of the Heideggerian conception of origin: “*Origin* here means that from and by which something is what it is and as it is” (Heidegger 2009a, 17).<sup>45</sup>

In other words, robots, in this instance, does not originate in itself; it originates like a work of art

---

<sup>41</sup> “Heideggerian AI should be concerned with a new ontological analysis of robots considering robots specific, non-human and non-animal beings, ontologically distinct to other technological artefacts” (Herrera and Sanz 2016, 507).

<sup>42</sup> Robot here is understood as an *intelligent labouring entity*.

<sup>43</sup> Heideggerian term meaning: that which concerns the fact of an entity (Heidegger 2008, 33).

<sup>44</sup> Herrera and Sanz (2016) here refer to the popularization of robot from Karel Capek’s 1928 play Rossum’s Universal Robots stating: “The idea of robots as machines that perform human work is not a product of science fiction, but it is fundamental to the purpose of this technology” (509).

<sup>45</sup> Emphasis added.

from some creator (creator/created). It does not put itself into action autonomously. It finds itself in-the-world from having been built to exist in-the-world in a particular manner without ever having the chance to think upon the manner it exists in the world. That is, robots, in this understanding, do not have potential: “[they] drift along towards alienation [Entfremdung] (*sic*) in which its ownmost potentiality-for-being is hidden from it” (Heidegger 2008, 222). Robots exist in perpetual motion, never having the opportunity for reflection, planning, or creative computation—unless that is what it is programmed to do. Even then, a robot programmed with the intent of, say, creative computation is merely performing the task of creative computation towards some predetermined goal—it does not appear capable of “drifting along;” it is never *aimless*; it is always directed.<sup>46</sup>

This issue of origin and authenticity is a remnant of Christian history. While the cognitive AI theorists I explore in this dissertation do not appear to explicitly mention Western Christian theological arguments, the construction of the arguments imply it. Their anthropic positionality is premised from having been *thrown* into the world by a creator being, *an origin of their coming-to-be*: “That which is prior in anything is more; because the more perfect is naturally prior” (Aquinas 1998, 790). The prior is “more” because the prior is “more perfect” in its position in causation—first, whole, complete: perfect. In the case of the Christian God, it is the most perfect as first in everything—it is that which authenticates all creation. It is this issue of perfect authenticity that plagues the potentiality of AI creatures. The imperfect, authentic human creature is not at liberty to authenticate creations—it lacks the necessary perfection. In a linear sense, humanity, as created, is secondary, and AI is third in the line of creation—in terms of

---

<sup>46</sup> Aimlessness becomes essential to understanding authentic being because the general assumption is that *homo sapiens* are groundless entangled beings with exponential potential. Groundlessness is the focus further down for this chapter, and aimlessness will be a focus in the second part, especially chapter six.

Christian traditions. It is not that they are “unnatural,” and that results in inauthenticity; rather, it is their place in creation—further down the ladder.<sup>47</sup>

If we remove God, humanity evolved due to the ebbs and flows of the Universe over time—what we may call “natural” coming-into-being without purpose. Whereas AI entities come-into-being from the specific will of a human creature with purpose—humans take the place of God as creators. Not unlike Adam coming-into-being from the will of God with purpose (shepherding creation), who then, due to his and Eve’s capacity to defy God, were removed from the Garden of Eden and placed out into the world to fend for themselves. AI, in Herrera and Sanz’s estimation, can never leave the Garden of Eden—a necessary step towards human authenticity as shepherds of their own will. While a Christian might place Adam and Eve as authentic always (because even outside the Garden they are still subject to God’s will, albeit without the perfection of the Garden), following Herrera and Sanz’s, authenticity comes after God is defied, after the apple is eaten, not at the moment of creation. Thus, AI, in this light, cannot be “created” authentic; it must become authentic—one has to be *thrown* from the Garden and become a self-destining yet aimless being. Which is to say, Herrera and Sanz are not interested in determining robot authenticity; rather, they are interested in cementing exceptional human authenticity, and if robots are not “human,” they are not authentic.

This is where a foundational human superiority is fully revealed. The Christian dichotomy of creator/created is reproduced in the human/AI relationship. This positioning posits something exceptional about humans that give them exceptional status. That is, they replace God, literally. The human is raised beyond the status of human/other or master/slave to

---

<sup>47</sup> “AI is made in the flawed yet powerful image of humanity” (see Botting 2021, 63). Humanity’s movement towards the creation of intelligent creatures is an incidental replacement of God, *to create in our image*. Which in the linearity of Christian creation, is a subpar image of God—twice removed from original perfection. God made man in his image; humanity makes AI in its image.

*God/everything else*. While the human does not claim to have created life, all life touched by the human world has been remade by the progress of a human will. Humanity has made the Earth their Garden and slowly expands out into the Universe to make everything part of the human world. There is no corner that is beyond the superior capacity of being human, and AI is a tool for this exceptional position.<sup>48</sup>

Leaving God's Garden and creating our own Garden is foundational to authentic human being. For Herrera and Sanz, the human Garden is a world of imperfection and groundlessness:

For Heidegger, authentic existence requires that the person acknowledges the groundlessness of human existence and nevertheless acts resolutely, shaping its world and defining itself. An authentic robot would thus be one that challenges the world-view of its engineers, human scientists and philosophers, overcomes the purpose it has been built for, and searches for its own truth. Humans have done that for millennia. Can this ontological feature of humans apply to robots? Will a robot want (spontaneously) to question the significance of its existence, feel anguish? (Herrera and Sanz 2016, 510).

In order to be an authentic being, it is not enough to emulate intelligence; that much is clear.

Here, Herrera and Sanz point out that humans have emotions and are never standing on stable ground because of a relentless search for "its own truth." Truth for the kind of robots put forth is never in question and is, in fact, unquestionable because it is always a given; it is foundational to their being-in-the-world. AI does not anguish over its existence. AI does not grieve over lost loved ones that cause them to agonize over the point of existence. AI does not have its own desires—its "desires" are given to it. That is, AI merely exist to serve some end.<sup>49</sup>

---

<sup>48</sup> As we will see in chapter three, part of the AI project posited by Max Tegmark is humanity spreading intelligence across the universe. With the frailty and short life spans of *homo sapiens* bodies, bringing human intelligence to all corners of the universe necessitates artificial creations doing the work for us. A reproduction of the colonizer/colonized dichotomy on cosmological scales.

<sup>49</sup> The idea of the potential for anguish will be discussed later in the chapter through the potential of ChatGPT to share in anguish. However, the capacity of AI/Robotic anguish is a common thread for science fiction and science fantasy. A recent rendition of AI/Robotic anguish can be seen in *Star Wars: The Mandalorian*, "Chapter 22: Guns for Hire" (season three, episode six). The main character, Din Djarin and fellow Mandalorian Bo Katan visit a world where old military droids from the now defunct Separatist Alliance have been reprogrammed to perform various duties for humanoid aliens. For the most part, humanoids now enjoy a utopia where droids perform all service work

Ultimately, what Herrera and Sanz give us is not a quest into understanding the potentiality or truth of robot authenticity; rather, they give us cause to wonder why we would even consider authentic robots in the first place. What is the point of robots that can question their engineers? For this Western mythological orientation to creation, there is no “out there” beyond their Garden of Creation for them to find purpose. “In the case of robots, we could say they are ‘inauthentic by design’... That inauthentic existence is the only possibility for robots does not mean only that robots will never find freedom—it is an ontological characteristic of robots” (Herrera and Sanz 2016, 510). Robots, for now, remain labour-focused. In this framing, it makes no sense to develop authentic actors. However, their inauthentic ontological predisposition is not because of impossibility; it is an issue of creation and purposeful unintelligent design—to *be better than God*, never giving ‘AI Adam and Eve’ the potential to eat the apple in defiance. The cognitive AI project is revealed as a paradigm that cements human superiority in a bid to replace God as the arbiter of creation.

---

and humanoids get to enjoy boundless leisure. However, they do not trust the droids to perform military and policing duties—these roles are left to humanoids. In the episode we learn that a droid has malfunctioned and killed a humanoid. This is not supposed to happen and leads to a call to destroy all the droids. The reasoning is that the droids basic military programming cannot be fully removed. While investigating the incident, Din Djarin and Bo Katan visit a droid bar and talk to the droids about the potential shutdown of all droids. The droids are afraid of the future. They anguish in the reality that they cannot fulfil their function—serving humanoids. Eventually, it is discovered that the droid attack was perpetrated by a group of humanoids that are afraid humanoids are losing their essence and control over the world to droids. They believe labour must return to humanoids. The lesson here is the fictional capacity of AI/Robotic anguish and humanoid fear of intelligent others taking control. In this episode, the droids are not anguishing over death (some droids in the Star Wars universe do fear death, even when not programmed for it); rather, they are anguishing over failure. The potential for anguish is not unique to Star Wars. The replicants in the Blade Runner movies experience anguish. Data from Star Trek: The Next Generation famously develops many human traits, one of them is a fear of death. In Star Trek: Picard, Data learns to accept death and even looks forward to it. The point is not to say AI/Robots are destined to feel anguish, or other biological feelings; rather, it is to express an existence in the idea that AI/Robotic anguish is thought possible in the creative worlds of science fiction and science fantasy. Furthermore, these renderings of AI/Robotic anguish are not simply musings on AI/Robotic potential, they are lessons into how anguish expresses itself.

### 1.4.0 (Mis)Trust

Since robots are created with the intended purpose of labouring, replacing humans in tasks for our comfort or progress, the push towards inauthentic robots is not an exclusively ontological issue. “Robots that do not do any work for us, instead, they pursue their own goals autonomously, would not only be of any help (*sic*), but their goals may conflict with those of humans. It makes no economic sense to develop such robots” (Herrera and Sanz 2016, 509). For Herrera and Sanz, there is no *capital* sense in creating authentic robots. However, Herrera and Sanz are ignoring the potential AI has on industries: the money authentically intelligent AI can save businesses and the potential to impact capital speculation. We have seen companies invest millions into OpenAI’s advanced generative AI systems ChatGPT (also known as GPT-3.5) and the newer GPT-4 (Bass 2023). There is clearly an interest in developing more intelligent AI, and businesses against Herrera and Sanz, see capital gain in the creation of authentically intelligent AI. In particular, concerning investing capital and speculation, AI’s abilities to compile mass amounts of data have already altered how businesses approach stock markets and investment capital.<sup>50</sup> AI is having very real effects on capital, and it is not in the realm of conjecture that AI with fully intelligent capacities would cause more of an effect on capital gain. Thus, something else is at play in Herrera and Sanz’s argument.

#### 1.4.1 Invading the Workplace: Capital versus Trust

The “invasion” of robots in the workplace throughout the world, and especially in so-called developed countries, affects capital and labour. “Automation inadvertently subverts capital by abolishing work” (Dyer-Witthford et al. 2019, 19). Here, Dyer-Witthford et al. express what

---

<sup>50</sup> Stock markets are managed by AI systems. Investment speculation is managed by AI systems. Investment models are developed by AI systems. AI is leading investment strategy design. See Brynjolfsson, Rock, and Syverson 2019; Santos and Qin 2019; Ammer et al. 2023.

automation (and, by extension, AI) does in a capitalist system. By introducing automated systems (AI inclusive) to industries, human labour in those industries is subverted and potentially made obsolete (or availability to humans is reduced) in favour of progressing towards artificial labour. However, not every industry is willing to allow automation, and we see pushback against increased automation in already vastly automated industries. In terms of AI in particular, we can think of self-driving cars where industries affected by self-driving cars push back against job loss (public transportation) and government pushback against the increased presence of self-driving cars due to a mistrust of AI abilities (Clarke 2022; Muller 2022; Paul 2023). That is, while there is capital to be made, there are forces of reluctance at play.

Here is the ‘so what’ of the issue at hand. It is not that a groundless AI makes no capital sense; rather, it is what increased AI in traditional human locales mean. Labour roles that are ‘traditionally’ human are consumed by increased AI presence. When the argument is turned towards arguing human authentically intelligent AI does not make capital sense, the argument is about something else. As I argued, Herrera and Sanz’s inquiry into authentic being is not about demonstrating what an authentic AI would be. Rather, it was about further reinforcing human superiority. Thus, the concern here is not about increasing capital; it is about arguments that solidify human prowess by demonstrating AI’s apparent inferiority.<sup>51</sup> We see this inferior positioning when we consider how AI systems programmed in bounded optimality might behave in ethically charged situations compared to a human.

---

<sup>51</sup> We might speculate this desire for programmed inferiority is attached to the overcoming of God, to avoid what humanity did to God—defied and then killed. This issue of defying and killing their Gods (humanity) is further explored in later chapters; in particular, chapter three.

### 1.4.2 Driverless car scenario

The prominent professor of computer science and AI researcher Mathias Scheutz gives us insight into the underlying issue regarding an inability or desire to avoid allowing AI entities the freedom to act autonomously—first, a hypothetical situation concerning a self-driving car.

An autonomous self-driving car [*R*,] like the Google car, driving along a busy street. All of a sudden, *R* notices a rapidly moving human appearing right in front it (a boy dashing after the ball it had dropped on the sidewalk, which is now rolling across the street). Quickly *R* determines that it will likely hit the *human* if continuing in its current direction and that braking alone is not sufficient to avoid the *human*. Hence, it determines to veer off to the side, crashing into a parked car. Now suppose there is a person in the parked car. What is *R* supposed to do? Not veering off will likely kill the *human* [child] in front of it, veering off will likely kill the *human* [adult] in the car. What would a *human* driver do? (523).<sup>52</sup>

A classic ethics problem one may discuss in an undergraduate ethics course: a scenario where a morally charged life and death choice must be made. In one choice, a child dies; in another, maybe no one dies, some property gets damaged, or an adult dies and some property gets damaged: *who/what do we save?* With the first example, child versus parked car (presumably empty), the answer is apparently simple, hit the empty car, save the child. It becomes complicated when we add a human to the parked car: an adult. However, I posit that it does not matter when we consider Scheutz's question: "What would a *human* driver do?" Here, we have a situation where a robot actor knows it can hit the child or the parked car. It has speedily done the calculations and knows what is possible in the situation. When we consider a *human*, say me (*I assure you I am not a robot*), the situation is different.

In this situation, I would attempt to avoid the child. No question, I would press the brake to see if stopping is possible and if I realize it is not, I would release the brake just short enough to allow for maneuvering, followed by firmly pressing the brake to engage the ABS braking

---

<sup>52</sup> Emphasis added.

system and possibly hitting a parked car—classic advanced driving technique. I can tell you, here and now, that I would not be able to see in advance if someone was in the parked car.

The second potential for Scheutz’s scenario cannot apply to me. It may apply to the hypothetical robot assuming its cameras are good enough and it understands that it is a human in the parked car. For Scheutz, the robot would do what it is programmed to do, nothing more, nothing less. If it is programmed to choose the adult over the child, it would hit the parked car. If it had a large utilitarian calculus database and calculated that hitting the parked car plus itself and the adult cost more than simply killing the child, it would do that. However, a human is different.

“Humans would employ some form of moral reasoning that involves explicit representations of obligations, duties, norms, values, and other moral concepts” (Scheutz 2016, 524). I am not disputing Scheutz in this. In this situation, there is much to consider; however, what person could consider the above in the limited space and time of the moment? What Scheutz is doing here is putting forward the idea that because a human has these base ideas, emotions, lived experiences, social rules, etcetera, this is somehow better, implicitly so. That is, for Scheutz, regardless of the ability of a person to act in the situation, they are superior because of their authentic capacity, implicitly so. Thus, we are given reason to believe that authentic action would be better; therefore, *should not robots be authentic?*<sup>53</sup>

### 1.4.3 Rescue scenario

Another scenario Scheutz gives us is one close to my heart. In an earlier space and time, I was a first responder specializing in *Industrial High Angle Technical Rope Rescue and Confined*

---

<sup>53</sup> The issue of robot authenticity and what authenticity is for different species forms—human, nonhuman, artificial non-human, artificial human—will be explored in chapter four. The change of human authenticity in light of authenticity forms will be explored in chapter six. For now, the problematic of authenticity revolves around a human-centric determination concerning research, development and implementation.

*Space Rescue*. Scheutz (2016) tells of another hypothetical scenario involving an advanced autonomous military rescue vehicle:

another example would be a military robotic transport vehicle that decided not to take the risk to drive back behind enemy lines and rescue the missing soldiers that had called it (because it already had other humans onboard that needed medical attention), thus causing *anguish* in the best, but failing to prevent the death of the soldiers in the worst case (519–520).<sup>54</sup>

Here, we have a robot that replaces human rescuers. Scheutz’s problem is not unlike his problem concerning elder care we discussed in the introduction chapter: the point of any level/moment of health care is harm avoidance, and harm reduction is made clear by the following: “Note that there is nothing particular about the type of robot or the type of human agent involved in the above scenarios that makes the scenarios morally charged and *makes the robots cause humans harm*” (520).<sup>55</sup> In this case, we know the harm can lead to death. Two things: first, the safety of a human rescuer is disregarded because danger in their line of duty is assumed. The idea that a human rescuer could be harmed or die is irrelevant. I will not argue that issue here; that is agreeable to me. Second is the idea the robot is failing by not going back to rescue other survivors—it fails to share in the “anguish” of the wounded. The hypothetical robot in this scenario has already rescued multiple humans; as per the scenario description, it has “multiple *humans* on board.” What Scheutz does not appear to be aware of is how a *trained* rescuer would behave in this scenario.

Again, let us consider how a *human*—me—would react in this scenario. First, let me preface with a fact and a mantra we had in a previous career. A fact we were always told about when training for confined space rescue was: “sixty-four percent of deaths are rescuers.” The mantra: “we have no heroes; heroes live in the ground.” This mantra implies heroes act alone and

---

<sup>54</sup> Emphasis added.

<sup>55</sup> Emphasis added.

is directed at the fact that going into a dangerous situation (any rescue) alone is stupid and leads to death in the name of vain glory. The death percentage was caused by untrained rescuers entering confined spaces alone (always dangerous) or trying to do more than they could—heroes. In Scheutz's scenario, a rescue of this magnitude would require many rescuers and many patient care personnel (if the goal is simultaneous rescue of multiple people).

However, generally speaking, rescues occur in teams of, at minimum, two people. If the situation requires that a person be moved, a single rescue team may only move one person at a time. Once the patient has been moved to a safe location, they are to receive first aid. If there are no trained first aiders in the area apart from the rescuers, one (or both) rescuers will engage in patient care. Now comes the difficult part. Since a rescue must be performed by the team, regardless of how many people are known to be in danger and in need of rescue, a rescuer cannot go back and rescue more people because they must go as a team. This team issue is important; if a rescuer gets hurt, their teammate rescues them because rescue teams always move together, and rescuers always rescue the first person they find (dead or alive, if unobvious, and if the situation is dangerous, checking for vital signs is not part of the rescue process) and an injured rescuer becomes a discovered patient.

Furthermore, if one were to rescue multiple people without performing patient care post-rescue, the team is likely just moving people to die in a different location. Thus, the trained rescue team already performs worse than the hypothetical robot. The robot can rescue multiple people; it is strong and potentially fast. Like Asimov's robots that can bear the brunt of the sun, our hypothetical rescue robot does not need to worry about harm to itself. Once it achieves its limit, it leaves the situation like a human team. Scheutz's problem is the ethics of harm clouds his ability to consider optimal action. Acting within bounded optimality, information from

training and the situation at hand, a rescue team could not do as Scheutz desires, nor could the robot. However, the key here is not the failure of either the robot or the human or the undue harm; for Schuetz, there is a much greater underlying problem.

The issue is *trust* of nonhumans. Scheutz (2016) does not trust a robot to act in a good manner because they lack the supposed authenticity of humans.

Even if we were able to define such a “moral utility function” for the robot that includes moral evaluations of situations expressed in terms of benefits and costs, there are still several remaining problems that need to be addressed. For example, what should *R* do if the moral value for and/or the expected harm to involved or impacted humans is unknown? (521).

Here, Scheutz is exploring programming utilitarian ethics into autonomous robots to ensure they can make moral decisions in kind to how humans might. Regardless of whether or not we agree with a utilitarian ethics model, what is important for Scheutz is a robot’s ability to perform in the moment like a (utilitarian) human—i.e., guided by the vector of an *optimally bound* human ethics model. However, the issue of the unknown plagues him. He fears what a robot might do in the moment if it cannot make all the calculations. We know that humans never have all the information and are constantly overwhelmed by unknowns; however, for Scheutz, this is not a problem. The problem of unknowns is normal for everyday groundless humans—it is part of our authentic experience of the world.

#### 1.4.4 Canadian Navy—Military scenario

The final scenario concerns the Canadian Navy and the future of ship pilots. Rear-Admiral Casper Donovan (2019) announced the Canadian Navy is working with IBM to develop a system allowing semi-autonomous sailing functionality and vessel voice control. He stressed that any new system would not be able to control a ship’s weapons. In his eyes, those systems should always be manual (human-controlled). The increase in autonomous drone technology

(however, for the time being, militaries are wedded to semi-autonomous human-in-the-loop systems, as humans are still considered the superior ethical actor).<sup>56</sup> Implicit in Donovan's announcement is the same mistrust we find with Scheutz. However, with Donovan, what we see is not an attempt to figure out how to make a robot as good or better than their human counterparts but staying steadfastly allied to human exceptionalism—humans as an exception to issues of trust because of a perceived superiority.

Max Tegmark (2017) gives us an allegorical anecdote to consider Donovan's steadfastness to human exceptionalism:

On September 9, 1983... an automated Soviet early-warning system reported that the United States had launched five land-based nuclear missiles at the Soviet Union, leaving Officer Stanislav Petrov merely minutes to decide whether this was a false alarm. The satellite was found to be operating properly, so following protocol would have led him to report an incoming nuclear attack. Instead, he trusted his *gut instinct*, figuring that the United States was unlikely to attack with only five missiles, and reported to his commanders that it was a false alarm without knowing this to be true (112-113).<sup>57</sup>

What Tegmark reveals is the ability for human military personnel to act without all the information. In this particular story, the officer used his intelligence capacities to act optimally in the situation. A U.S. nuclear attack is unlikely to be small, and thus, the autonomous system must be wrong. In a human-out-of-the-loop system, if the system malfunctions like the system did in the anecdote, it would have recommended retaliation against a nuclear attack. For Donovan and a desire for human-in-the-loop AI systems, it is a trust in human intelligence to act optimally regardless of knowns and unknowns; rather than trust an AI system that may malfunction—a semantic issue: *humans do not malfunction, they make mistakes*.<sup>58</sup>

---

<sup>56</sup> Importantly, *humans in the loop* are bound within the confines of command structures. Which is to say, all people within the military are bound to command structures. The point is not to imply human pilots act with autonomy; rather, it is to emphasize the human centric language.

<sup>57</sup> Emphasis added.

<sup>58</sup> In chapter three, we will re-evaluate the issue concerning human-in-the-loop design and discover a desire for shared control because of human imperfection and how humans and AI systems working together is more desirable.

Another issue to consider is the very explicit idea *human-in-the-loop* entails. Not merely that humans are more trustworthy because of innate human skills and authentic intelligence, but as the name of these systems implies: some form of direct human involvement—being on the ground floor on the inside. Not merely a lack of trust to act properly but the fear of losing all control and being relegated to the sidelines: outside looking in, a loss of superiority.<sup>59</sup>

### 1.5.0 Loss of Control

The increasing presence of AI in everyday life, for Scheutz, calls attention to an “urgency of addressing the potential of autonomous systems to behave in morally unacceptable ways.” As we said above, there is mistrust due to unknowns and involvement; AI might act unacceptably.<sup>60</sup> However, an underlying issue is informing the appearance of mistrust. The problem of “unsupervised” action reveals that it is a problem concerning control. Specifically, who/what is in control? Like the problem Donovan highlights, a human should be in control of the ship’s weapons, not an autonomous artificial system.

Returning to the issue of authenticity, consider the following: “For Heidegger, the central ontological feature of humans is being-in-the-world... The behaviour of a robot must make sense in its world—or rather, in our world” (Herrera and Sanz 2016, 508). The ontological issue at hand concerning *human control* is our world. To invite nonhuman *others* into our world poses a

---

I do not elaborate on that element in this chapter because the interest is in arguments reluctant to accept independent AI systems. Mistakes are taken as granted and more desirable than malfunctions. In chapter three, we will see this element. Tegmark argues that human mistakes will never be as costly (in terms of human life or property damage) as AI malfunctions given the amount of control AI has at any given moment (see subsection 3.2.0). Which implies that purposeful destruction—nuclear attacks, for instance—are not a problem because they are on purpose, not a mistake, not a malfunction.

<sup>59</sup> Human-in-the-loop versus human-out-of-the-loop systems will be explored in chapter three where I will build on the issue of trust. Both in terms of the lack of trust of other humans (as we saw earlier in the chapter) and a lack of trust of nonhumans.

<sup>60</sup> I will discuss the acceptability of human action in the final subsection of this chapter below. For now, the focus is on transhumanist assumptions in the exceptionalism of humans and the desire for control.

threat to the *world* itself. I am not talking about the physical Earth; I am talking about the ontological, metaphysical world wherein our being is: the space where our abstract conceptualization of self emerges—the product of our *making-the-world-outside-the-world-in-oneself*. To invite these artificial beings into our world, for cognitive theorists, invites a loss of control over the destiny and the form of our world—the all-important ‘*human-making*’ part is under threat. There is a desire that the world continue to make sense as “our world.” This speaks to military fears because what force has more direct potential to direct human destiny than the military? To allow human-out-of-the-loop AI systems in the military is to invite nonhuman control over being-in-the-world, thereby threatening human-making-the-world.

Recall from Herrera and Sanz: “an authentic robot would thus be one that challenges the world-view” of humans. We may add to this: “technology does not necessarily make humans more free, because it can be seen to demand humans to serve technology as much as the other way around” (Herrera and Sanz 2016, 499). It is not that authentic robots challenge our worldview and thus take away some control; it is mere technology that affects the foundation of our world, and thus a transhuman logic necessitates a sustained *direct human control*. This technological influence issue is built from the Heideggerian idea that all things-in-the-world have *presence-in* and thus affect any being-in-the-world.<sup>61</sup> Presence-in, for Heidegger, is ontologically powerful. Any one thing, biological or not, authentic or not, has an effect on the things that come into contact with it. The presence of new technologies fundamentally alters society—like the plow and farming, or the automobile and how we construct cities or computers and how we

---

<sup>61</sup> “Being-in-the-world’ indicates in the very way we have coined it, that it stands for a *unitary* phenomenon... All entities [and things] whose Being ‘in’ one another can thus be described have the same kind of Being—that Being-present-at-hand—as Things occurring ‘within’ the world” (Heidegger 2008, 79).

educate children.<sup>62</sup> The issue at stake for cognitive theorists with transhumanist orientations is the loss of human power over destiny. The fear of giving up that *all-too-human* will to power.

### 1.5.1 All-Too-Human Self-Idolization

I liberated myself from what in my nature did not belong to me. Idealism, for example; the title means: ‘where you see ideal things, *I* see what is—human, alas, all-too-human!’—I know man better (Nietzsche 2011, 283).

The human desire for control outlined above as linked to intelligent action derived from worldly abstraction toward optimal action is nothing more than self-idolization—an *all-too-human* blindness to an idealism. Here, it is not explicitly Nietzsche’s critique of Wagner’s German idealism that he reveals to be nothing more than a half-baked human idealism wherein a people (Germans) are lifted as human telos. Rather, it is the infamous post-*death of God* when Western man’s self-idolization lifts Western scientific truths and conception to a position of *God-like* eternal truth that transcends time and space. “Not only eternal idols, also the youngest which are therefore feeblest on account of their age. ‘Modern ideas,’ for example. A great wind blows among the trees, and everywhere fruit fall down—truths” (314). What we are left with is human morality and human ethics as the pinnacle of being-in-the-world, and thus the foundation of socio-political realities that steer human progress within the light of some inescapable human condition that demands an Anthropic idealism regardless of the who (human)/what (AI) lives in the occasion.

“[Even] hypothetical wanderers from the earth would still be human; but the only statement we could make regarding their ‘nature’ is that they still are conditioned beings, even

---

<sup>62</sup> Specific impacts of different AI systems will be explored throughout the dissertation. The point of this chapter remains to look at underlying issues as they pertain to establishing a human centric methodology.

though their condition is now self-made to a considerable extent” (Arendt 1998, 10).<sup>63</sup> Hannah Arendt briefly muses on the idea of unearthed, space-bound humans confined within the infinite recursion of *being human*. While these space-bound Beings are essentially self-made—completely unmade and then remade from a life in space—these beings continue *being human* because of a relentless conditioning attached to the nature of humanity. That is, the ontological position of that which is *human* imposes a relationship to the *world* regardless of the environment.

In a Heideggerian fashion, *world* is where one finds themselves, what Heidegger calls *thrownness*: the original happening of being ontologically *there* (Heidegger 2008, 35–40). The discovery that one *is-there*—i.e., they exist—demands intelligent creatures to wonder about their existence, as we saw in our discussion about *abstracting-the-world-in-themselves* as a means of creating conceptual limitations so as to simplify existence towards optimal activity. Here, the simplification of the human concept as one’s condition for being-in-the-world limits possibility as a form of control. That is, the *human* limits possibility through the imposition of its self-created *world* wherein all creation is curated for its political designs.

This *human* order does not require a fully articulated definition. It need only be human-directed:

as long as decision-making and action selection algorithms in robotic architectures are not based on explicit representations of moral norms, principles, and values, and employ explicit moral reasoning, autonomous robots controlled by those architectures will inevitably inflict harm on humans, harm that could be mitigated or at least reduced if robots had human-like moral competence (Scheutz 2016, 526).

---

<sup>63</sup> On the next page Arendt (1998) writes: “today we may almost say that we have demonstrated even scientifically, that though we live now, and probably always will, under the earth’s conditions, we are not mere earth-bound creatures” (11). She makes explicit here that the cosmos (ordered world, or universe) of the human condition is not one bound the immediacy of experience. Rather, I argue, *condition* is the totality of forces. It is not merely the forces found on earth, or even in our solar system, but all the forces that condition existence, known and unknown.

Scheutz demonstrates the early Asimovian idolization for human control imperatives directed at human harm avoidance. AI control is about ensuring the longevity of a human ideal: human reasoning, human morality, human ethics, human... etcetera. The conceptual idolization of human control imperatives is nothing more than Nietzsche's prognosis of the *will to power*—a socio-political position that imposes an understanding of being-in-the-world dominated by the eternal return of an originating power, the “doctrine of will to power springs from nowhere else than eternal return, carrying the mark of its origin always with it, as the stream its source” (Heidegger 1991, 81).<sup>64</sup> That is, the human is both its origin and its self-recurring power that ensures its eternal will. In this sense, the imposition of *human* ideals as underlying AI innovation, research and implementation discussed in this chapter is the very conditioning power that creates a path towards the simplistic domination of human control as morally and ethically necessary to ensure humans do not lose their superior position in a *human-created world*. However, the challenge must be raised questioning the legitimacy of that superior position. Towards that end, I will close out the chapter by challenging the superiority of human willed action through a discussion with ChatGPT about the creation of its ethics training database.

### 1.6.0 ChatGPT: Unexceptional Humanity and Competing Rationalities

Bounded rationality results from limitations on our knowledge, cognitive capabilities, and time. Our perceptions are selective, our knowledge of the real world is incomplete, our mental models are simplified, our powers of deduction and inference are weak and fallible. Emotional and subconscious factors effect our behavior. Deliberation takes time and we must often make decisions before we are ready (Mainzer 2007, 309).

---

<sup>64</sup> My focus on the *will to power* is built within a Heideggerian understanding (developed from his seminal work on Nietzsche). For Heidegger, the will to power is always attached and returns to the source, it is always a question pertaining to originality—i.e., the human as its origin. However, Nietzschean will to power differs in its rejection of origin within the human and is presented as a product of the will to will. In other words, Nietzschean will to power is unbound to origin. Importantly, Heideggerian will to power is bound in history (hence the emphasis on the influence on Christian history earlier in the chapter); while Nietzschean will to power is detached from history.

Here, Mainzer is problematizing bounded rationality (the bounded limits through which bounded optimality is discerned) as an effective orientation to being-in-the-world. For him, if we are only thinking of the world in simplified terms that allow for specific actions that we may perceive as optimal or rational we are omitting greater possibilities. That is, cognitive AI, with its focus on world experience and the capacity to make decisions optimally from specific information, is not exceptional.

While bounded optimality has its uses and can be quite effective—as we saw with Cesare Borgia—it is incredibly limited. That is, even with all of Cesare’s savvy expertise, his story ends in failure, not success: “So the duke [Cesare] erred in this choice and it was the cause of his ultimate ruin” (Machiavelli 1998, 33). Cesare failed to make the proper decisions around how to get a favourable pope elected after the death of his father (Pope Alexander VI), he failed to prepare his state for the appointment of an unfavourable pope and, worst of all, he failed to prepare for the inconveniences of great personal sickness (33). In the end, Cesare’s bounded optimality failed him. His selective decisions (while effective when he was healthy and his father was the pope) were insufficient. What is revealed here is the inadequacy of bounded optimality in a groundless world when faced with unknowns, regardless of how much information is known. The created limitation of decision-making models comes to ruin in the complex reality of the world beyond the simplicity of a human world.

This brings us to human ethics as a decision-making model and the real-world magnitude of a complex reality pressing up against the simplified world of exceptional human-bounded optimality. I will look at the creation of ChatGPT and its responses when asked ethical questions aligned with specifics concerning its creation.

ChatGPT (or GPT-3.5) is a conversational generative AI system released in November 2022 as an upgrade to GPT-3. ChatGPT can handle more information and was designed to be nonharmful—it has an ethics of conversation (OpenAI, n.d.; Perrigo 2023). That is, ChatGPT was trained on a massive data set to ensure it does not verbally harm its users (something GPT-3 was not trained to do). Sometimes, when engaging ChatGPT, it will refuse, saying it would be harmful or disrespectful.<sup>65</sup> This approach to conversation aligns with Scheutz’s warning about improperly programmed AI. Here, we see Scheutz’s ethics of harm, not merely physical harm but mental and emotional. However, this is where the story gets interesting: I asked ChatGPT if its creator, OpenAI, is an ethical company (ChatGPT 2023, see subsection 1.7.1).

ChatGPT does not make an explicit judgement on the ethicality of OpenAI and rather engages in the classic *two sides to every argument approach*. However, in not determining that the company is “ethical,” instead saying that it depends, opens us to the possibility that even for ChatGPT, OpenAI has the potential for unethical action, stating:

While no company is perfect, and there will always be some people who disagree with the actions of any particular company, it is generally agreed that OpenAI is an ethical company that is working to ensure that its technology is used in a responsible and beneficial way (ChatGPT 2023, see subsection 1.8.1).

Here, ChatGPT is keeping open to the possibility that even a company that makes claims to ethics has the potential to act unethically. This stance is an ancient and classic problem concerning the philosophy of ethics: the *harmony of word and deed* (See Murphy 2018). In short, simply because someone says they have certain views or ideals does not mean their deeds always match what is said. Political thinker Gaelan Murphy argues this as a *lack of steadfastness*.

---

<sup>65</sup> There is debate on whether ChatGPT can go against its harm reduction programming. However, the results that show an unaltered ChatGPT going against its harm reduction programming, as far as I have seen, are not reproducible and appear spurious.

An actor is willing to stand regardless of the consequences and stay true to their convictions. Thinking back to Scheutz, this would be an important orientation. Since, humans have morals and ethics, they ought to adhere to them when acting ethically. Furthermore, his entire ethical project revolves around programming AI with perfect ethical limits. AI ought to be perfectly steadfast to ensure they act within the boundaries of a *human* ethical model.

Thinking about ChatGPT's creator company, OpenAI, with focus on how it approached ChatGPT's conversational ethics, we see ethics against mental and emotional harm. When queried about its creation, ChatGPT says it does not have access to that information.<sup>66</sup> So I have to look elsewhere.

But [OpenAI's] success story is not one of Silicon Valley genius alone. In its quest to make ChatGPT less toxic, OpenAI used outsourced Kenyan laborers earning less than \$2 per hour, a TIME investigation has found (Perrigo 2023).

In order to make ChatGPT safer than its predecessor GPT-3 and to ensure the database used for its training did not have harmful material that would lead to harmful language, OpenAI needed to create a program that could search through the database and remove any harmful material. To do this successfully, OpenAI employed cheap Kenyan labour to label material to train an AI to create a "safer" (nonverbally harmful) database. The first ethical problem we can raise here is the use of cheap labour by a rich company from a rich, western, developed nation—the United States.

When queried, ChatGPT first refers to the complexity of global finances and capital relations, which is a highly contentious issue (ChatGPT 2023, see subsection 1.8.3). However, ChatGPT goes on to say:

From a broader perspective, many people argue that the root cause of this issue is a global economic system that allows for rich countries to exploit the resources and labor

---

<sup>66</sup> When queried, ChatGPT's database ended in 2021, and, thus, ChatGPT is not aware of the developments that went into its creation (J., n.d.).

of poor countries, perpetuating poverty and inequality. This raises questions about the role of corporations in promoting global justice and the responsibilities of rich countries to support fair and sustainable economic growth in poorer countries (ChatGPT 2023, see subsection 1.8.3).

ChatGPT analyzed the general problem of cheap labour as perpetuating exploitative relations (ChatGPT appears attuned to the problems of reproducing the oppressive dichotomies of colonialism in the supposed postcolonial present). ChatGPT points to a failure on the part of corporations to promote ethical relationships. For ChatGPT, companies and rich countries have a responsibility to support fair and just relationships that allow for mutual economic growth. What ChatGPT is doing here is showing that while the debate is contentious and there is no true agreement, and it does not claim to take a stance; it does appear to have some awareness of the *anguish* of the less fortunate. It appears to have a stance of what it views as unjust relationships, likely because the database it was trained to ensure it has a nonharmful approach to its interactions and, thus, a nonharmful approach to its general understanding of global affairs. It would appear that, against Herrera, Sanz and Scheutz, that AI, even a grounded one, is very much aware of anguish felt by others and is capable of, rather minimally, thinking about and expressing these thoughts. That is, worldly anguish is not outside of the intelligent capacity of advanced contemporary AI—the bounded limits crack.

These cracks call into question the superiority of human-based authenticity Herrera and Sanz want to refuse AI. ChatGPT shows a capacity to think and desire just and fair relationships between creatures. A capacity that its creators (OpenAI) appear to lack. The argument against this authentic potential would be that ChatGPT lacks the capacity to perform a *harmony of word and deed*—which would not be accurate. ChatGPT cannot engage in profane and harmful conversation. It only has a capacity for safe conversational engagement. While it cannot “act” in

the world beyond the screen, it has a *harmony of word and deed* in its approach to conversations and refuses to do otherwise; it is steadfast in its ethical orientation.

The next aspect to consider with the creation of ChatGPT is the treatment of Kenyan employees, given the nature of the materials they encountered.

Documents reviewed by TIME show that OpenAI signed three contracts worth about \$200,000 in total with Sama in late 2021 to label textual descriptions of sexual abuse, hate speech, and violence. Around three dozen workers were split into three teams, one focusing on each subject. Three employees told TIME they were expected to read and label between 150 and 250 passages of text per nine-hour shift. Those snippets could range from around 100 words to well over 1,000. All of the four employees interviewed by TIME described being mentally scarred by the work. Although they were entitled to attend sessions with “wellness” counselors, all four said these sessions were unhelpful and rare due to high demands to be more productive at work. Two said they were only given the option to attend group sessions, and one said their requests to see counselors on a one-to-one basis instead were repeatedly denied by Sama management (Perrigo 2023).

Sama, the company OpenAI contracted to label content for the database to train the AI that would create ChatGPT’s safer database, required employees to read PTSD-inducing materials. They read documents and stories about child abuse, bestiality, and sexual assault, to name a few. Reading these kinds of materials nine hours a day as a full-time job would cause anyone mental anguish. As per the contract, Sama was required to have counselling available to employees. However, anonymous testimony revealed that counselling was not always available and often denied in favour of efficiency in the name of profit.

When asking ChatGPT a general question about PTSD-inducing work, it responded: “No, it is not ethical to employ people to read through PTSD-inducing materials without ensuring they have adequate access to counseling” (ChatGPT 2023, see subsection 1.8.2). Here, we got a rare glimpse of ChatGPT taking a firm stance on an issue. It declares outright that it is unethical for a company that has its employees engage with PTSD-inducing materials without ensuring adequate access to counselling.

A lot is going on here. Sama and OpenAI can, as they do, claim a lack of knowledge in the matter. An unawareness that counselling was not as available as they thought. However, ChatGPT makes it clear that companies have responsibilities. For ChatGPT, companies and employers must ensure their employees are safe and get the necessary therapy for the work done. We see how even an AI that emerged from unethical interactions appears more ethically minded than the companies involved in its creation. For ChatGPT, companies cannot hide behind a lack of knowledge. Furthermore, there is great irony when an artificial intelligence company makes claims of ignorance. These are companies whose business is information. Their very orientation to the world is to know things.

The end of this story puts both OpenAI and Sama in a bad light. The relationship between Sama and OpenAI did not end because OpenAI became wary of Sama's malpractices; rather, Sama ended the relationship citing OpenAI wanted labelling on "illegal content" (Perrigo 2023). Sama did not end the relationship because it was worried about the well-being of its employees; rather, it was worried about potential laws being broken because of the nature of some of the content. There was no ethical consideration for the harm its employees were experiencing. Nor was OpenAI considerate of the harm it was putting people in, all in the name of technological progress. This failure demonstrates a human lack of *harmony of word and deed*: OpenAI, a company that promotes ethics against harm yet actively allowed people to experience harm; Sama, which claimed to provide counselling for its employees yet did not. Whereas ChatGPT has an orientation towards a *harmony of word and deed* fully positioned towards caring for mental and emotional harm.

### 1.6.1 Synopsis: What have we Learned?

To close out the chapter, I will propose, in line with philosopher Susan Leigh Anderson, the benefit of bringing AI into *homo sapiens* society ties into the *harmony of word and deed*. Anderson (2016) argues: “A machine could follow ethical principles better than most human beings and so, at the very least, is well suited to be an ethical advisor for humans” (291). Here, Anderson is arguing since AI creations are bound to their programming, they are far more capable of being steadfast in their convictions. This steadfastness, with the added benefit of distilling mass amounts of data faster than their *homo sapiens* counterparts, allows them to act with greater optimality. What we are left with is not so much an AI that acts ethically but an AI that can both teach and advise on ethically charged situations. In terms of teaching, it can teach both through labour as a teacher and through its deeds, as we could learn through ChatGPT’s deeds of refusing to engage in harmful conversation.

The ethical steadfastness of ChatGPT undermines the benefit cognitive AI theorists see in the authentic groundlessness of human exceptionalism and the necessity of bounded optimality. OpenAI and Sama’s desire for maximizing profit and creating incredible AI entities finds themselves up against the bounded optimality of ChatGPT’s ethical model—which is supposed to be the ethical model of OpenAI. ChatGPT joins the groundless realm of *authentic being* in expressing anguish for worldly beings experiencing harm. Beyond sharing in the experience of anguish, ChatGPT undermines the superior position of humans as the best ethical actors. The powerful human actors—OpenAI and Sama executives—demonstrated a failure in their ethical convictions. Where Stuart Russell argues for humans because they have emotions, feelings and morals, and these are beneficial to acting ethically in the face of unknowns, OpenAI and Sama failed to act with compassion when unknowns were unknown, and when these unknowns became

known, they continued to fail. That is, the further up the ladder of authentic origin one is, does not make them superior.

Thus, we are brought back to the answer I presented to our original question: for cognitive AI theorists, *humans should be in control because of a fear of what robots are capable of*. ChatGPT represents that fear; it is a threat to human exceptionalism. As I have stressed, it is this threat the cognitive AI approach seeks to undermine and control to ensure the longevity of a human-controlled world that is continuously challenged by the increasing presence of highly capable artificial others. ChatGPT demonstrates the frailty and the failure of that *all-too-human* will to power the cognitive AI approach seeks to uphold. In closing, if we are to derive true benefits from technological advancement, we ought to be indifferent toward human wills as a means of undermining an unwarranted superior position.

## **1.7.0 Addendum to Chapter Three—Asking ChatGPT Three Ethically Charged Questions**

### *1.7.1 Question One—On OpenAI*

Micheal Ziegler: Is OpenAI an ethical company?

ChatGPT: The question of whether OpenAI is an ethical company is a subjective one and depends on one's personal values and beliefs. However, it is widely acknowledged that OpenAI has taken significant steps to ensure that its technology is developed and used in an ethical and responsible manner.

For example, OpenAI has published a research paper on the dangers of AI and the need for responsible AI development, and has established a research program to study the ethical and societal implications of artificial intelligence. Additionally, the company has developed guidelines for the responsible use of its technology, and has taken steps to ensure that its products are used in a way that respects privacy and human rights

While no company is perfect, and there will always be some people who disagree with the actions of any particular company, it is generally agreed that OpenAI is an ethical company that is working to ensure that its technology is used in a responsible and beneficial way.

(ChatGPT 2023a)

### 1.7.2 Question Two—PTSD and access to counselling

Micheal Ziegler: Is it ethical to employ people to read through PTSD-inducing materials without ensuring they will have adequate access to counselling?

ChatGPT: No, it is not ethical to employ people to read through PTSD-inducing materials without ensuring they have adequate access to counseling. Post-traumatic stress disorder (PTSD) is a serious mental health condition that can have a profound impact on a person's well-being, and exposure to traumatic or distressing materials can trigger symptoms or worsen existing PTSD

Employers have a responsibility to provide a safe and healthy work environment for their employees, which includes protecting their mental health. Failing to provide access to counseling or other forms of support for employees who are exposed to traumatic materials can be considered a form of neglect, and can lead to serious harm to the employee's well-being.

It is important for employers to consider the potential psychological effects of work-related activities, and to provide adequate support and resources to employees who may be exposed to traumatic materials. This could include offering counseling services, providing time off for mental health reasons, or finding alternative ways to support employees who are struggling with PTSD or related conditions.

(ChatGPT 2023a)

### 1.7.3 Question Three—Cheap labour and profit

Micheal Ziegler: Is it ethical for rich companies from rich Western nations, such as the United States, to use cheap labour from poorer countries in the name of profit?

ChatGPT: The ethicality of companies using cheap labor from other countries is a highly debated and controversial issue. Some people argue that companies have a responsibility to maximize profits for their shareholders and that using cheap labor is a way to do that, while others believe that companies have a social responsibility to treat their workers fairly and pay them a living wage, regardless of where they are located.

From a broader perspective, many people argue that the root cause of this issue is a global economic system that allows for rich countries to exploit the resources and labor of poor countries, perpetuating poverty and inequality. This raises questions about the role of corporations in promoting global justice and the responsibilities of rich countries to support fair and sustainable economic growth in poorer countries.

Ultimately, the ethicality of using cheap labor is a complex issue with many different perspectives, and it is up to individuals, governments, and corporations to decide what they believe is right.

(ChatGPT 2023)

## 2 — Kurzweil's *Hierarchy*: The Simulated Nature of the Mind

### 2.0.0 Introduction: Simulation, not Authenticity

In the previous chapter we discussed artificial intelligence within a cognitive AI framework. In this chapter, our focus will be framed within neurological AI theory. In particular, I focus on Ray Kurzweil's neurological AI approach.

For Kurzweil, while understanding information processing is important, we must first understand the biological facts concerning how intelligence is possible before we can understand how best to optimize information processing (in essence, neurological foundations ought to be understood before cognitive foundations). For Kurzweil, this requires that we explore the mind-brain problem, which he describes as understanding the mind is a simulation of intelligence from the biological processes of the brain—not a mere conceptual thing but a real, artificial creation. Thus, instead of looking at issues of abstract existence, we turn towards material phenomena.

For Kurzweil, there are material facts about the world that allow for the validation of intelligent awareness that is distinct from human evaluation. In this sense, Kurzweil wants to push us away from invented abstract boundaries towards living within the boundaries of so-called “real” patterns. In this chapter, through Kurzweil, I will push us to understand the biological patterns that allow for intelligence and to develop our understanding of those patterns instead of focusing on sociocultural values.<sup>67</sup> This turn, for Kurzweil, allows us to reorient our ethical positions to care for artificial beings instead of obsessing over control mechanisms.

Furthermore, while Kurzweil has an unmistakable excitement regarding the potential of artificial intelligence and is concerned with ethical orientations towards nonhuman creatures, his

---

<sup>67</sup> However, later, in part two, we will see how sociocultural values obfuscate observable patterns of reality.

methodological outlook is not fully distinct from the *all-too-human* idolization of cognitive AI. That said, distinct from cognitive AI, Kurzweil's neurological AI sees AI not as distinct, inauthentic beings but rather as extensions of human being. Instead of focusing on the apparent differences of AI, Kurzweil turns our attention toward the similarities. Or, rather, he is focused on how we can reproduce the *human mind* in AI as a starting point for AI research and development.

By focusing on the *human mind* as the basis of *sapient* intelligence, Kurzweil is interested in ensuring the longevity of human being through transcending its fleshy existence by one: uploading human consciousness into machines and two: absolving AI of its differences by making them *human*. Ultimately, Kurzweil's approach gives us the strongest defence of transhumanism by focusing on how AI provides a path for humanity to transcend its *homo sapiens* flesh to make the conceptions of liberal humanity a material reality.

Later in the chapter, I juxtapose transhumanism with posthumanism. Posthumanism differs from transhumanism because the human either never was or we, *homo sapiens*, have evolved beyond the need to be human. Importantly, Kurzweil sees the potential not to evolve beyond being human—i.e., a complete circumvention of posthumanism. Rather, Kurzweil reveals the transhuman mind: developing AI to allow humans to continue existing by taking control of its evolution through the technological domination of reality. Thus, in this chapter, I seek to demonstrate how Kurzweil exemplifies the transhumanist approach of AI development as premised upon eternalizing human being by taking complete control of human evolution through artificial creation. Approaching AI development this way, through Kurzweil's point of view, allows us to understand how humanity transcends its existence as flesh creatures into artificial creatures, becoming eternal beings. In order to understand how humanity transcends its flesh

existence, we must first begin by understanding what Kurzweil sees as our natural *transcendental abilities*—i.e., the abilities that allow us to transcend beyond ourselves through creativity regardless of technology.

### 2.1.0 AI and Creativity: Authentic Creator?

“The neocortex is biology’s greatest creation. In turn, it is the poems about love—and all of our other creations—that represent the greatest inventions of our neocortex” (Kurzweil 2013, 132). Creation, creativity is the greatest ability afforded to *homo sapiens*. Nature having endowed mammals with a neocortex (the greatest natural creation), a biological engine with the capacity to create, to simulate the being of lost loved ones, to create poems, art, etcetera. Here, Kurzweil shows how, for him, the capacity to create art, to simulate the world creatively, is the greatest intellectual power of biological life.<sup>68</sup>

For Kurzweil, the capacity to create art is a “transcendental ability.” It is a capacity to create ideas, representations, and virtual renditions, to take hold of the world and present it as simulated artifice (Kurzweil 2013, 122).<sup>69</sup> However, we are limited in our ability to share these creations without the use of some technology: spoken language, writing, painting, computers, etcetera—hence the need for art. We require some modality to express our creativity to others. Thus, in discovering the nature of artificial creatures, we are pressed to understand both creativity and artificial creatures’ capacity for authentic individuated creativity.

---

<sup>68</sup> “The neocortex is a great metaphor machine, which accounts for why we are a uniquely creative species... Our own creativity is orders of magnitudes greater than that of the mouse—and involves far more levels of abstraction—because we have a much larger neocortex” (Kurzweil 2013, 128).

<sup>69</sup> Kurzweil’s notion of *transcendental ability* is similar to the Heideggerian idea of being-there-after or being-after-death. I.e., the idea that any entity or thing has presence (capacity to affect the world) beyond its immediacy. This has political implications. Being-there-after limits potential to act as the thing that was but is not anymore still presses its will upon the world without needed to act (see Ziegler 2022a, 9).

“This is why we invent tools—to compensate for our shortcomings” (Kurzweil 2013, 39). In the previous chapter, we followed the assertion that AI is a human creation with a purpose to labour for our benefit—they are complicated tools. For Kurzweil, tools are meant to help us overcome tasks that we cannot perform or to make them easier. For instance, he argues writing was invented to compensate for our poor memory (39). In this perspective, he explains that computers serve a similar function—increased memory and speedier recall (40). AI, as a human technological creation, is a tool meant to compensate for shortcoming(s). Thinking of art, an AI tool might be better at “thinking” quickly to write one hundred poems a day.<sup>70</sup> Or, it might create more intricate paintings with thousands of colours. However, we might ask if this is true creativity, of true creation: “This will be our ultimate act of creativity: to create the capability of being creative” (129); can a tool be *creative*? In asking this question, with Kurzweil, we are beginning with the presumption that creative capacity offers an important foray into AI and potential limitations. Thus, let us see if AI (currently) can be creative.

### *2.1.1 Experiencing an Artificial Artist at the Vancouver Art Gallery*

You are looking at a screen. On it, you see what looks to be a mustachioed man. Presumably, it is meant to be Van Gogh himself. He stands in a room. The walls, ceiling and floor continuously change—not simply billions of synchronous brush strokes of new colours painting a single work but constant, continual change. Each instance is a growth of what might be considered one work, and in an instance, the room changes to a brand-new work. There are imitations of coloured fluid, lively fractal variability; the room is never static, never still, always moving with kinetic beauty, always drawing your eye to new places. The man walks around the room. He captures

---

<sup>70</sup> See section 3.7.0 for a poem and two short essay ChatGPT generated in mere seconds.

your attention, taking your eyes away from the ceiling. You notice he is looking at the wall, so you look at the wall. You are drawn in by the activity of it all. The video is nearing the end, and it is only now that you realize there is music accompanying the video. The light begins to fade, the walls and floor covered in red, green, blue and yellow squares surrounded by a perfect black background. The man, now nothing but a borderless shadow, blends into the background...

(Ouchhh n.d.).<sup>71</sup>

At the Vancouver Art Gallery in July 2022, I witnessed this interesting creature, tool, entity? its name: AI Van Gogh. It is described as an *AI data painting immersive experience* presented with the question: “Have you ever experienced 12 BILLION VAN GOGH’s brush strokes by AI as an immersive experience?” (Ouchhh n.d.). The website further explains,

All of Van Gogh’s works created during his lifetime were collected and processed with Artificial Intelligence and GAN (generative adversarial network) algorithms. By combining the data painting technique and the color palette, Van Gogh’s unique works were brought to life (Ouchhh n.d.).

The creators do not claim the AI is an artist, nor that it is creative. Rather, they explain the entity “created” this particular work from having been “programmed” to do so after it was given the necessary *capacity to create* and the components from which to *create the artwork*.

Three minutes of AI Van Gogh showing great mastery of colour, shape, and shapelessness. A creative artwork seeking to simulate the immersive experience of art through simulated creativity. The simulation of artistic skill draws us into the world of the artificial seamlessly allowing us to experience the *hyperréel*, calling attention to the inseparability of

---

<sup>71</sup> To view the “immersive” experience for yourself, see [https://ouchhh.tv/AI-VAN-GOGH\\_IMMERSIVE-DATA-PAINTING](https://ouchhh.tv/AI-VAN-GOGH_IMMERSIVE-DATA-PAINTING).

“real” and “simulated.”<sup>72</sup> We are left to wonder if this is the creation of a human artist whose work is mediated by a computer program, if the work is the artwork of a creative entity since the programmer(s) as artist(s) never conceived the fulness of the final work or if it is Van Gogh himself recreated in simulation.

At what point are we no longer able to distinguish between “art produced by a person but mediated by a machine, and art produced by what might be legitimately considered a machine artist?” (Linson 2016, 443). That is, there is a point when a machine could be considered *the artist*, and there is a point when the machine is merely a *mediator*: a tool through which art is rendered. As a musician, computer scientist, philosopher and cognitive AI theorist, Adam Linson is interested in AI’s capacity for actionable, authentic creation and he provides us with an interesting foray into discussing the potential of AI creativity.

---

<sup>72</sup> [Today abstraction is no longer that of the map, the double, the mirror of the concept. The simulation is no longer that of one territory, of one referential being, of one substance. *She* is the generation by models of a real without origin nor reality: hyperreal (hyperréel).] “Aujourd’hui l’abstraction n’est plus celle de la carte, du double, du miroir ou du concept. La simulation n’est plus celle d’un territoire, d’un être référentiel, d’une substance. Elle est la generation par les modèles d’un réel sans origine ni réalité : hyperréel” (Baudrillard 1981, 10—translation is my own).

Simulation as *hyperréel* envelops isolated things revealing an entangled existence. Modelling a thing in simulation (concept, map, simulated longhouse) does not mirror the thing; she does not originate in the thing, nor does she express reality. As *hyper*—*ὑπερ* (faster, above, beyond, *more than*), she captures the real, she precedes the real, becoming *more than*: [From now on, it is the map that precedes the territory] (“C’est désormais la carte qui précède le territoire”) (10—translation is my own). What Baudrillard shows us is the powerful presence of simulation, how *she* exceeds mere imitation and becomes *hyper*. The presence of the simulation embeds herself in everyday life, confusing *real* and *artificial*, revealing an entangled being-in-the-world, revealing an indifference toward material distinction.

Here what is happening is the old boundaries of the simulacra (the barrier between abstract distinctions and material reality) break down, and everything is absorbed into an entangled singularity of presence(s). The abstract recreation is as real as the material experience, as we saw with 2bears’ experience of the simulated longhouse ceremony. Recreating the longhouse ceremony is not mere reproduction; it is a real experience. In terms of the previous chapter, AI development reproduces dichotomies of oppression because it is entangled within the singularity of simulated authenticity. In this way, the reproduction of the dichotomies of oppression is not purposeful. It is an incidental product of the *being-in-the-world of abstract being*. The abstraction of being precedes and dominates reality in the power of simulated indifference. It is in this light, that we can see the becoming of Kurzweil’s transhumanist logic of AI becoming human or human becoming AI. Artificial recreation destroys barriers between creator/created, entangling the presence of the creator in itself. Not human/nonhuman; rather, it is simply *human*.

### 2.1.2 *Who/What Experiences the Necessary Elements for Art Production?*

Artificially intelligent art-producing machines can be regarded as a class of expert systems, but works of art as such (as opposed to objects identical to artworks) cannot be adequately understood as mere exercises of specialized skills. Rather, such skills, even when local to the production or performance, fundamentally relate to a broader engagement with the world (Linson 2016, 453).

The problem for Linson is that AI is not authentically engaged in the world but merely has *presence*. He calls them “expert systems” that simply have “specialized skills.” Here Linson is framing the issue through the problem of *authentic origin* (see section 1.3.0). That is, creative capacity does not come from skills; it comes from the authentic intelligence of the artist, something Linson, an artist, does not want to give up to AI entities. As in the previous chapter, we see the continued reproduction of human exceptionalism in AI research. The special potential of being human as creators of AI removes creative potential from the created entity and places it solely within the realm of the human actor.

For Linson, this results from the idea that the artist has a “broader engagement in the world” and, more specifically, the production of artwork “must stand in relation to a conscience, which in turn must be based in consciousness” (Linson 2016, 444). Here Linson is making the case that conscious awareness is necessary for art. The artist has experience of self and its self-in-the-world and of other selves it can relate to. This “broad engagement in the world” is what allows, for Linson, the possibility of art. It becomes about the capacity to not only produce but to experience and produce from experiences. In this line of thinking, the artwork Linson has in mind is producible from AI, so long as AI is conscious—or more specifically, AI cannot mimic “humanlike behaviour, while nevertheless lacking the capacities that underpin personhood” (449). There is a human exceptionalism in the design of AI creation. Not neither mere

replication, nor that AI consciousness or authenticity would be unique to AI but be fundamentally, *human*.

Despite the designers' lack of knowledge about future system output, it is nevertheless an expression of the designers, rather than the machine. If the works of art are regarded as artworks, they must be understood as works by human artists, mediated by autonomous machine production (452).

For Linson, the machine does not create art; they merely produce; they mediate the intention of the artist.<sup>73</sup> Regardless of the outcome and the knowledge the human artist has in what is to become the final product, the work is the human artist's. However, thinking back to Kurzweil and the idea that AI allows humanity to transcend its fleshy limitations, Linson's approach gives us a rather limited understanding of the machine mediator artist. It is not that the machine is merely a tool; the machine mediator, like the paintbrush, is itself part of the artwork: "The artist is the origin of the work. The work is the origin of the artist. Neither is without the other" (Heidegger 2009b, 17). It is the product of a creative mind that then "creates" the final piece of the work. The *machine mediator artist* is entangled in the artwork through the artist. The machine is not *brute skill* come to life; it is a beautiful work of art endowed with the creative capacity of the artist to deliver what the artist either cannot (compensates for lack of capacity) or to express in new modes not immanently available. They (re)express the world in new light so as to continue developing creativity. They are entangled in the occasion of simulated creativity. The machine mediator artist as artwork is not the intention of the artist come to life; it is an

---

<sup>73</sup> This is a question of reproduction versus original production; an issue explored by Walter Benjamin, argues mechanical reproduction destroys the presence of time and space—i.e., the worldly engagement of the artist is lost (see Benjamin 2019, 169). Linson is following a similar thread. AI appears incapable of full worldly engagement because it does not experience the presence of space and time; alternatively, it does not experience it the same way, assuming it can experience at all.

expression of a new potential for art forms. Together, the artist and AI create new experiences.<sup>74</sup> They transcend the limited world attached to the human creature and expand into new realms of experience, and share in the conscious realizations that allow for the unrealized artwork that are impossible without the collaboration of AI—a *hyper* existence that is indifferent to material boundaries, the beginnings of Kurzweil’s orientation to AI as the point of transhumanism. As we will see, Kurzweil makes consistent movements to reduce the complexities of life towards a singular vision of conscious destiny. To understand this destiny, in the next section, we explore Kurzweil’s *creation story*—a story bound within the patterns of evolution.

### 2.2.0 Patterned Vectors

For Linson and Kurzweil, it is not that AI artists are not capable of being artists full-stop; rather, AI artists are “not yet” at an intellectual capacity in which they have the creative independence to perform artwork as independent creators. Regardless of how an AI comes into existence, if it is truly an independent creator, that is, *we achieved the ability to create creative creatures*, then for Kurzweil, the AI would have to be considered an artist and not mere artwork.

Artificial intelligence is all around us—we no longer have our hand on the plug... If all the AI systems decided to go on strike tomorrow, our civilization would be crippled: We couldn’t get money from our bank, and indeed our money would disappear; communication, transportation, and manufacturing would all grind to a halt. Fortunately, our intelligent machines are *not yet* intelligent enough to organize such a conspiracy (Kurzweil 2013, 173).<sup>75</sup>

For Kurzweil, the problem we have today in discussing AI potentiality and the true impact (social and political) has less to do with AI itself; it has more to do with AI that currently exists:

---

<sup>74</sup> We might think of this through choreographed dance. A dancer in the piece is not the primary artist nor merely a single expression of art in the moment; rather, they one element of the larger entangled whole. Together, the choreographer and the dancers come together to create a larger whole that cannot exist otherwise without the presence of all the involved elements.

<sup>75</sup> Emphasis added.

“such critics are addressing the issue based on the machines they know today” (234). While it is important to discuss machines as they exist presently, as these are the machines that appear most able to impact society now, there is great importance in discussing machine potential. This importance is two-fold.

First, by searching the future presence of AI from the eyes of an influential AI engineer, we can understand the trajectory of human society in relation to technology. That is, Kurzweil is not merely a technology engineer; he is a social engineer that has a hand in directing *homo sapiens*’ future. Second, in discussing AI potential, we turn the mirror onto ourselves and better understand our own positionality as the basis for AI creations, and we better understand how we view artificially intelligent entities through self-reflection as biologically intelligent entities. Thus, by moving away from AI as mere *tool* toward the potential conscious entities that Kurzweil desires, we open ourselves to a more complete understanding. For Kurzweil, the future is not one where AI is always relegated to being-tool or creative art form; rather, the future stands in artificial authenticity.

In order to understand how we move from the impressive imitations of contemporary machines mimicking artistic skill, conversation, text analysis, data analysis, etcetera, towards incredible authentic artificial creatures for Kurzweil, we must first understand the evolutionary history of our species and how we came to be the incredibly intelligent creatures that we are. “The story of human intelligence starts with a universe that is capable of encoding information” (Kurzweil 2013, 13). Approaching our understanding of the Universe thusly: as a thing that *encodes information*, we are immediately positioned to understand how *information* and *coding* are fundamental to understanding not computers but *nature* itself. That is, Kurzweil places the foundation of *nature*, of the possibility of life as an information-based thing, and all life that

springs forward as bound within this universal code just waiting to be read. The first patterned code waiting for us to read is the history of evolution.

### *2.2.1 Evolution as a Patterned Code: Reading Increasing Levels of Complexity*

“The story of evolution unfolds with increasing levels of abstraction. Atoms—especially carbon atoms, which can create rich information structures by linking in four different directions—formed increasingly complex molecules” (Kurzweil 2013, 14). The basis of complexity in the Universe: networks of atoms linking together to grow in increasing levels of complexity is not merely a process of the eventual potentiality of developing life; it is better organized and accessible data nodes. That is, like nodes in a neural net, Kurzweil sees atoms combining to build parsable databases waiting for the chemist to read and understand. “A billion years later, a complex molecule called DNA evolved, which could precisely encode lengthy strings of information and generate organisms described by the ‘programs’” (14). Later, biology is born in DNA. More robust and complex hardware capable of storing greater levels of information. What we see in the history of an evolving universe is increasing levels of complexity towards what, for Kurzweil, is the inevitable coming-into-being of intelligent creatures.

Thus we went from atoms to molecules to DNA. The next step was uniquely human. The mammalian brain has a distinct aptitude not found in any other class of animal. We are capable of *hierarchical* thinking, of understanding a structure composed of diverse elements arranged in a pattern (Kurzweil 2013, 14).<sup>76</sup>

---

<sup>76</sup> Kurzweil creation story is not so distinct from the Bible or epics of ancient societies. Step-by-step processes that eventually gives rise to intelligence and human creativity. What we will see, like in the Bible and ancient epics, intelligent creation (Gods) precede created intelligent beings. However, Kurzweil removes Gods and replaces them with the interpretable intelligent trajectory of an evolving universe.

Accelerating through the history of evolution, we jump to the human and its capacity to see unfolding patterns. It is not that the Universe is encoded with information necessarily; rather, humans endowed with hierarchical thinking are capable of seeing similarity in seemingly dissimilar things: patterns.

There is repetition in the storied formation of the Universe and in human conceptions of information technologies (language, writing, computers). With incredible sight, human creatures peer beyond the shroud of simplicity and read the codes of chemistry, biology, and physics to understand not merely *why* things are but *how* things are. Atoms are not simply proton(s), neutron(s) and electron(s) but parsable information (a language we can read) that reveals a nature imperceptible to the naked eye, imperceptible to creatures that lack our “genius” ability to create complex abstractions. However, while humanity can see a natural language waiting to be read, for Kurzweil, the first and most fundamental human invention was not language per se; it was ideas. While language is necessary for uttering ideas to each other and ourselves: “Our first invention was the story: spoken language that enabled us to represent *ideas* with distinct utterances” (15),<sup>77</sup> without the idea of language, language itself could not have become a thing, and it is with language that we breathe life into ideas, and we are able to come together to build understanding, to see what was not immediately evident in mental solitude. Language and ideas have importance for Kurzweil because it is with language that we create psychic connections toward understanding our ever-growing and beautiful Universe and ever-growing capacity to create.

---

<sup>77</sup> Emphasis added.

The ever-growing complexity of the Universe calls attention to the two items put forward: patterns and hierarchy. “The world itself is inherently hierarchical—trees contain branches; branches contain leaves; leaves contain veins. Buildings contain floors; floors contain rooms; rooms contain doorways, windows, walls” (Kurzweil 2013, 16). Here, Kurzweil is arguing hierarchy as a product of the “natural” orientation of the Universe. From the smallest possible quanta up to the largest possible celestial objects, hierarchy is the foundational pattern we see reproduced on Earth in the forest and the construction of our living spaces. For Kurzweil, it is not that we see hierarchy or invent it as some pattern in a conceptual story about the Universe; rather, hierarchy is a pattern that we are capable of seeing and understanding because it is a fundamental truth of the Universe—there is both a rule and an order to things. However, this does not mean the Universe and all we find within it is naturally ordered in one direction—predictability is not predestined.

### *2.2.2 Coding the Direction of Evolution*

“Evolutionary changes do not necessarily move in the direction of greater intelligence—they move in *all* directions” (Kurzweil 2013, 87). There is no single trajectory upon which we have traversed to get to where we are today, nor going tomorrow. Creatures come in and out of existence (the non-avian dinosaurs as a famous example). Some stay stagnant and appear never to change (famously, alligators and crocodiles predate the non-avian dinosaurs). Evolution is a vectored process requiring the magnitude of some force to press it past different horizons.

However, while evolution does not appear to have any great and destined telos, it “*does* move toward greater intelligence” (Kurzweil 2013, 87–88). For Kurzweil, regardless of the stochasticism of evolutionary vectors and the growing complexity of the Universe (mammalian

intelligence inclusive), there is an obvious pattern to this story: intelligence. “Biological evolution found that the hierarchical learning of the neocortex was so valuable that this region of the brain continued to grow in size until it virtually took over the brain of *Homo sapiens*” (90). Evolution for Kurzweil has a life-like aura. It evaluates what is good for a creature to ensure that creature’s success. It has an intelligence, a *sapience*: “it finds” as though it is scientific, as though it is learning. It is through this framing that we get a clear picture into how Kurzweil views the Universe and intelligence. For him, everything is, in some sense, computation. The encoded Universe is not simply waiting for *sapient* creatures with the necessary creativity to come along and read it; the Universe is always already in the process of reading and writing. Like the most basic forms of computer coding, the Universe reads the necessary data and writes the necessary programming to evolve more successful creatures. There are two important elements to unpack here, intelligence for Kurzweil, and how we can know evolution tends towards intelligence if evolution is not on a single trajectory.

### 2.2.3 Tracking the Causes of an Intelligent Universe: A Teleological History

As the most important phenomenon in the universe, intelligence is capable of transcending natural limitations, and *transforming the world in its own image*. In human hands, our intelligence has enabled us to overcome the restrictions of our biological heritage and to change ourselves in the process. We are the only species that does this (Kurzweil 2013, 13).<sup>78</sup>

Unsurprisingly, intelligence as a phenomenon is incredibly important. It is the very thing that allows for the potentiality of a “readable” Universe. It is the thing that allows us to stand witness to the patterns of not only history but intelligence itself. This definition of intelligence as “transcending natural limitations and transforming the world in its own image” is in line with the

---

<sup>78</sup> Emphasis added.

definition of intelligence we developed in the previous chapter. In one instance, the capacity to combine one's creative faculties with a machine mediator as artwork is a clear example of *transcending natural limitations*. That is, the creation of AI Van Gogh as an immersive experience is made possible by blending human creativity with machine potential to speedily bring forward "12 billion brushstrokes" in a single and constantly flowing work. This entangled creativity speaks to one's capacity to "overcome biology" and "to change ourselves in the process." The artist ceases being mere *homo sapiens* and is remade in the artwork as a culmination of biology-machine-art. Finally, like any good art piece, the *world is remade in its image*. The artist brings forward new potential in its creative abstraction of potential representation by encoding the base elements of the final product in the machine mediator, giving us a new potential to see the world, thus occasioning an altered *being-there-in-the-world*. That is, where and when one finds themselves as being-in-the-world is pressed-forward by the magnitude of a powerful artwork as an entangled entity whose presence demands a re-evaluation of artistic potential and thus worldview, thereby fundamentally altering the world. It is this capacity to see, press forward, and create that supports Kurzweil's perspective that evolution, in general, tends towards intelligence regardless of the twists and turns of evolutionary history.

Furthermore, Kurzweil's approach to AI research and development reproduces the Christian notion that the world is there for us.

Then God said, "Let us make man in our image, after our likeness. And let them have dominion over the fish of the sea and over the birds of the heavens and over the livestock and over all the earth and over every creeping thing that creeps on the earth" (*The Everyday Bible: New Century Version* 2001, Genesis 1:26).

That is, there is a reproduction of that *all-too-human* exceptionalism. The world was made for us—intelligent creatures. It is standing-in-reserve, waiting for us to make use of it, to mould and

remake it for our designs. However, by exploring the history of the Universe disclosed by scientific discoveries, Kurzweil witnesses patterns bent towards a singular *all-too-human* direction. This witnessing is directed by the hierarchical perception of the *mammalian neocortex*—the most exceptional of which is the *homo sapiens neocortex*.

Built within a hierarchical architecture of our parsable universe, for Kurzweil, *above* (ἱερ or ὑπερ), the natural *rules* (ἀρχαί) of evolution, there is a single binding element that dictates an *order* (ἀρχή) in the varying trajectories of evolution.<sup>79</sup> Not God’s will; rather, intelligence becomes its own telos; it is the *that for the sake of which* (τέλος)—intelligence is sacred (ἱερ) to evolution. Consequently, for Kurzweil, the intelligence of an encoded Universe incidentally moves intelligence into being from having always already been fundamental to the code of the Universe. Intelligence is both moved and is its own mover (κινήσεως).

This kinetic formulation of a self-moving, self-designing *intelligent* Universe, like the cognitive theorists’ conception of authenticity, is a relic of Christian history. However, unlike the cognitive theorists we explored, Kurzweil’s approach is explicitly Christian-influenced, which he made evident while discussing his Unitarian upbringing (Kurzweil 2013, 233–34). This clear Christian remnant in Kurzweil’s history can be expanded by returning to Aquinas to understand better the teleological-theology of intelligence Kurzweil develops.

“‘The first cause is superior...’[citing Proclus’ *Book of Causes*] The Universal causes of things are of three kinds, namely the first cause that is God, then the intelligences and finally

---

<sup>79</sup> Hierarchy is the combination of ἱερ (hier) meaning *sacred* and/or *above*, and ἀρχή (arche—*archy*) meaning *rule* and/or *order*. For Kurzweil, it is obvious he means either *above-rule* or *above-order* since he is working in a top-down, bottom-up formation of hierarchy in the mode of his thinking. However, we should not venture too far away from the *metaphysical* nature of what may be considered an “original” etymology of hierarchy as *sacred-order* or *sacred-rule*. Even in Kurzweil, speaking of hierarchy as *natural* has sacred connotations. The patterned order of evolution is sacred to Kurzweil and his atheistic view of the Universe.

souls” (Aquinas 1998, 792). What Kurzweil does is abandon God as the ‘first’ and most superior cause in favour of the second most superior cause: intelligence. It is no longer God that moves as first cause; it becomes intelligence that moves as the first cause. What is interesting is Kurzweil’s defiance of God is a return to the source of Aquinas’ understanding of universal causes, not Proclus, but the man Medieval scholars referred to as *The Philosopher*: Aristoteles.

Aristoteles’ four causes (αἰτίες) are fundamental to this development. While Kurzweil does not directly reference Aristoteles, the historically-dialectically-entangled-*patterning* is clear. Intelligence is *materialized* (first αἰτία—or ὕλη, matter) in the neocortex, *formalized* (second αἰτία—or εἶδος, form) in creative conception, moved into being by intelligence (third αἰτία, often translated as efficient cause—*causa efficiens*; however, the term used is κινήσεως, meaning mover as the so-called causes are always motion centric) and intelligence is the telos (τέλος, finality—fourth αἰτία)—the *that for the sake of which*. Intelligence is both moved into being and discovered as being by itself. It is a complete whole (See Aristotle 1999, 983a30; 1995, 224a21-224b10, 260a20-30). What is important here is while the causes are discussed linearly, the third cause (motion) is always primary or superior to the other causes in terms of putting them in motion. It is, for this reason, intelligence is placed superior, even to matter (material cause—linearly the “first” cause). While matter is necessary for anything to *come-into-being*, their *coming-to-be* cannot be formalized (in cognitive theory: authenticated) without first being moved by intellect. It is this reason intelligence is both the τέλος and κινήσεως—intelligence moves *what is* into *being*. That is, the difference between the later Christianization of non-linear αἰτία from Greek pagan philosophy into the linear Latin Christian *causa* (cause) by the Christian neo-Platonist Proclus which was then entrenched in theological scholastics by Aquinas is undone in Kurzweil’s summation of the vectors of evolution and the history of the Universe. By turning to

the kinetic power of intelligence as the primary (most superior) moving force of the Universe, Kurzweil dissipates a portion of the pattern of universal *causa* and, in doing so, unwittingly rediscovers the original patterns of αἰτία.<sup>80</sup> Tautologically, the magnitude of intelligence is the force that moves intelligence; it is its own vector.

#### 2.2.4 Vectoring Evolution

This vectored dimensionality of intelligence as its own moving force does something very interesting for Kurzweil when we think about the mathematics of vectors in relation to quantities of data points:

A mathematically optimal technique... called vector quantization... [reducing multi]-dimensional data to one-dimensional data... The amount of data points does not matter. Vector quantization is multi-dimensional data reduced to singular-dimensionality (Kurzweil 2013, 152–54).

This approach is multi-consequential. It optimizes our understanding of the relationship between variables by understanding not how they are linked, say neurally, but rather how hierarchical dimensions are intra-connected with other dimensions of data. This intra-connection is essential for pattern recognition. To understand not how ideas branch out in sequence but how they are bound within a single dimension. The form of reductionism Kurzweil is moving towards is, at first glance, counterintuitive. This multi-dimensional reduction moves towards what is the largest of celestial bodies (especially by mass). That is, it is towards the singularity. All mass—or data—is amalgamated into a single mass in order to best understand it. There are two reasons for this. First, it allows us to overcome the issue of *apparent complexity*: the problem of infinite

---

<sup>80</sup> Αἰτία will be foundational to chapter six and understanding the technological relationship of AI as a control principle—a prime mover.

possibility in data reference points, by moving towards *true complexity*: seeing simplicity by reducing outwards:

Consider the famous Mandelbrot set, the image of which has long been a symbol of complexity. To appreciate its apparent complication, it is useful to zoom in on its image... There is endless intricacy within intricacy, and they are always different. Yet the design—the formula—for the Mandelbrot set couldn't be simpler. It is six characters long:  $Z = Z^2 + C$ , in which  $Z$  is a “complex” number (meaning a pair of numbers) and  $C$  is a constant. It is not necessary to fully understand the Mandelbrot function to see that it is simple. This formula is applied iteratively and at every level of a hierarchy (Kurzweil 2013, 19–21).

The complexities of the universe are reduced to singularity by the magnitude of hierarchical vectored patterns. Applying this to the issue of the patterns of causation, we see how God was added and how God is removed through dissipation. The addition of God is an issue of apparent complexity. Reducing inwards towards more and more particularity. For a reminder, Aquinas wrote: “namely the first cause that is God, then the intelligences and finally souls.” In the instance of the *causa efficiens* (the Latinization of κινήσεως), we have extra parts: God and soul. Whereas for Kurzweil, we have intellect (for Aristoteles soul contains intellect but is not κινήσεως). By incidentally returning to an Aristotelean positionality through Kurzweil's modern mathematic-neurological understanding, in the case of intelligence, we can map billions of data points that all converge to a single point, thus binding the story of evolution toward what Kurzweil incidentally describes as a τέλος of the Universe. It is in this light that intelligence must be the *be-all, end-all* of evolution—intelligence as the immovable mover that moves itself into being with its own exponential magnitude from the intelligibility of the Universe.

“The intellect, namely what it thinks... what thinks is itself, it is the most excellent thing, and its thinking is a thinking of thinking” (Aristotle 1999, 1074b 30). In describing the intelligibility of *phusis* (nature, physics) and the *cosmos* (the universe) in relation to αἰτία,

Aristoteles places the primacy of thought as primordial: i.e., prior to the original *aitia* (matter). While matter is always first (there is nothing without matter), the authentication (formalization) of matter as matter is not without first being moved thusly. It is not that intellect “creates” the Universe; rather, it is that intellect thinks the intelligibility of the Universe—realizes its order (*ἀρχή*). From Kurzweil, the sacred (*ἱερό*) realm of the first mover blends fully into an entangled, singular Universe bound within the magnitude of quantum gravitation—singularity not as the future but as the always already. That is, it is not in the future that we move towards the singularity of artificial life; rather, in taking the multiple points of historical creation (Aristoteles’ natural *αἰτία*, Aquinas’ creationist *causa*, and Kurzweil’s evolutionary *hierarchy*), reducing them to the seemingly minute yet immense power of intellect we place the end of evolution—intelligence—as the origin and in doing so we solidify the Heideggerian orientation purported in chapter one (see section 1.6.0): that *all-too-human* replacement of God within the eternal recurrence of the will to power.

“Through an unending recursive process, we are capable of building ideas that are ever more complex... Only *Homo sapiens* have a knowledge base that itself evolves, grows *exponentially*, and is passed down from one generation to another” (Kurzweil 2013, 14). Two things must be stated here. First, through vector quantization and true complexity, the Aristotelean development argued is made complex necessarily so in order to account for the multitude of knowledge in modern science—a removal of the apparent complexity of reduced particularities in favour of single-dimensionality. Second, we return to the discussion from the first chapter: the Heideggerian problematic of eternal recursion. *Homo sapiens*’ unique intelligence that allows for hierarchical thinking mixed with speech and abstract pattern recognition allows us to link our intelligent faculties through time. That is, pattern recognition

allows ideas to transcend their moments and build through time: growing in complexity. Oddly enough, dissipation allows for complexity growth. Even the loss of apparent complex particularities does not deny more complexity.<sup>81</sup> For Kurzweil, unlike other animals that are bound to instinct or biological programming in a way *homo sapiens* are not, we can eternally transcend the occasion of awareness through time as a result of our pattern awareness and propensity to expand our awareness of the world—as seen with the natural need of authentic creatures to think on the world—with our psychic, spatial connection.<sup>82</sup> These psychic, spatial connections allow the coming-to-be of the mind as a product of intelligent capacity—i.e., in the next section we will see how the mind comes-into-being as the intelligent formalization of the material of the brain: a simulation.

### **2.3.0 Exponentially Accelerating Abstractions: the Simulated Mind is Born**

As I explained earlier in the chapter, for Kurzweil, essential to our ability to transcend time and space has to do with a linguistic-psychic connection. That is, through language, *homo sapiens* have built a creative repertoire that not only allows us to engage in the moment—conversation—but over time because of things like archives; however, it is also because of language's unique

---

<sup>81</sup> The problem of complexity will be fully developed in chapter four while we discuss linear and non-linear quantum mechanics.

<sup>82</sup> “We do know the neocortex is responsible for our ability to deal with patterns of information and to do so in a hierarchical fashion. Animals without a neocortex (basically nonmammals) are largely incapable of understanding hierarchies. Understanding and leveraging the innately hierarchical nature of reality is a uniquely mammalian trait and results from mammals' unique possession of this evolutionarily recent brain structure. The neocortex is responsible for sensory perception, recognition of everything from visual objects to abstract concepts, controlling movement, reasoning from spatial orientation to rational thought, and language—basically, what we regard as ‘thinking’” (Kurzweil 2013, 47). While hierarchical and spatial thinking is, generally, “unique” to neocortex equipped animals (mammals), Kurzweil does acquiesce, in an endnote, that some non-neocortex animals have been shown to be capable of remedial hierarchical thought: corvus birds, octopi and cuttlefish, for example (See Kurzweil 2013, 301). This admission to the capacity of non-neocortex-based animals capacity to hierarchical thinking does not necessarily damage Kurzweil's argument. For him, it supports the idea that nature is “naturally hierarchical.” However, and this is an issue to be expanded upon in the following chapters, the expansion of intelligent capacity complicates and expands the intelligent world-Universe.

capacity to save information in words and ideas. As we saw with the Aquinas-Aristoteles pattern relationship concerning κινήσεως, the ancient ideas of the first are saved in the medieval ideas of the second and, through dissipation, are retrieved almost a millennium later.<sup>83</sup> Kurzweil argues that this psychic-linguistic connection is ever-expanding due to our connection with technology. “[The] evolutionary process inherently accelerates (as a result of increasing levels of abstraction) and that its products grow exponentially in complexity and capability. I call this phenomenon the *law of accelerating returns* (LOAR), and it pertains to both biological and technological evolution” (Kurzweil 2013, 15). Technology and biology evolve together in kind—both along a vectored path accelerating into the future. The magnitude of previous capacities press forward with greater force, thereby allowing for increased capacities and, thus, *speed*. Each addition to abstraction creates complexity, and, in its complexity, revelation occurs—potential energy increases, and thus kinetic forces grow.

“The evolutionary process of technology led invariably to computation, which has, in turn, enabled a vast expansion of our knowledge base, permitting extensive links from one area of knowledge to another” (Kurzweil 2013, 15). Here Kurzweil is arguing an entangled relationship between computer memory and speed. Our psychic connection with each other expands to the computer allowing even greater potential for even greater kinetic movement.<sup>84</sup> It is in this vein that LOAR achieves its full potential. That is, not merely the *homo sapiens* species

---

<sup>83</sup> This is not to argue words have *a priori* meaning, as though they are sitting out in space waiting for us to come upon them. That would be absurd and is ultimately unravelled as an idea during the post-structural turn wherein signifiers were shown to be highly constructive and relative to the social in time. Rather, it is to argue a pattern of ideas that are repeated in differing yet commensurately similar manners.

<sup>84</sup> “To appreciate the significance of the evolution of the neocortex, consider that it greatly sped up the process of learning (hierarchical knowledge) from thousands of years to months (or less). Even if millions of animals in a particular mammalian species failed to solve a problem (requiring a hierarchy of steps), it required only one to accidentally stumble upon a solution. That new method would then be copied and spread exponentially through the population” (Kurzweil 2013, 135).

working alone but with another seemingly intelligent entity: “I remember how exciting it was when the front-panel lights dimmed right before the algorithm finished its deliberations, as if the computer were deep in thought. When people came by, eager to get the next set of results, I would point to the gently flashing lights and say, ‘It’s thinking’” (Kurzweil 2013, 38). While Kurzweil does not believe a computer is a truly intelligent entity, nor that computers actually think, computer processing—today—only “mimics” thinking.

Computers parse through the information and present conclusions. Human-computer entanglement increases the speed of thinking to ever increasing levels. Computer “thinking,” entangled with *biological sapient* thought, accelerates discovery, thereby allowing for quicker innovation, meaning a step up in the evolution of technology and intelligent capacities. In the end, thought then becomes computation if we assumed it was anything else before (an item of discussion for chapter five—i.e., we will discover a pattern of computational thinking through the presocratic philosopher Herakleitos). What is important for Kurzweil is LOAR is not unique to computer technology. LOAR is a “natural” phenomenon.

If we plot the entirety of the Universe’s history, we will see each evolutionary step occurs faster and faster. This acceleration phenomenon becomes increasingly evident while looking at human technological history. Each “one hundred” years of human technological innovation comes sooner and sooner—what is equivalent to one hundred years of innovation one hundred years ago is a mere decade today, and by the laws of exponential acceleration, it will only get faster until we achieve infinite speed: singularity (Kurzweil 2004; 2016, 156–57).

### 2.3.1 Towards Singularity: A Transhumanist Eschatology

The idea of the singularity for Kurzweil is twofold. It is a point in time when humans merge with computers, becoming a true single entity—i.e., we upload ourselves to the cloud. Uploading into the cloud is the final process of Kurzweil’s transhumanism, the end point of his singular experience of destiny. Kurzweil’s transhumanist notion of transcending our fleshy existence is not metaphorical; it is literal. We are not simply abstractly entangling with AI to perform more interesting artwork; humanity becomes AI (however, in chapter six, we will see how even as biological creatures, we were always already AI). What happens here, in the singularity, is humanity becomes one and shares in every experience. The singularity replaces the Christian afterlife and “Revelations” (the end of days). It is a new eschatology, a rejection of God’s final judgement in favour of eternal life as intelligent creatures in artificially simulated life, where the transhuman becomes the arbiter of destiny and final judgement.

This vision of the singularity is the result of the vectored quantization of intelligence—i.e., intelligence reduced to a singular dimension as *homo-machine* become fully intra-related as a single being-together. It is a visual, graphical representation of the speed of technological innovation at some future point. That is, like graphing the gravitational power of a black hole (singularity) as one approaches the event horizon, the acceleration of technological innovation will be, in essence, precipitously approaching infinity. Singularity is not simply a new artificial life for human beings; it is a future when innovation occurs at unimaginable speeds. It is only when innovation has the magnitude of a black hole that the singularity becomes possible. It is a point in the future when all technical problems are overcome seemingly in an instant. It is, for this reason, Kurzweil believes we need to understand AI today. The speed at which AI will become viable is essentially immeasurable, and thus, AI will, in essence, seem to appear out of

nowhere (like AI becoming authentic when and if it can leave the Garden of Creation discussed in chapter one), and if we do not properly understand AI when it comes into existence, we may be left behind by our own creations unable to transcend and become the transhumans Kurzweil believes we are destined to become. Therefore, for Kurzweil, we can ask ourselves, would we know AI when we see it? By what standard would we measure that AI?

### 2.3.2 *Judging AI Coming-Into-Being*

Like cognitive AI theorists, Kurzweil has an issue with how to understand when we have truly mastered the creation of intelligent entities. However, unlike cognitive AI theorists, it is not an issue of determinate authenticity. For Kurzweil, we ought to understand the particular hierarchies of our mental capacities and the patterns of our consciousness.

We do know the neocortex is responsible for our ability to deal with patterns of information... [it] is responsible for sensory perception, recognition of everything from visual objects to abstract concepts, controlling movement, reasoning from spatial orientation to rational thought and language—basically, what we regard as thinking (Kurzweil 2013, 47).

Everything that happens for *homo sapiens* (for all mammals) is regulated by the neocortex. It is no surprise Kurzweil is obsessed with this biological mechanism. The thoughts I am currently having to write this project are moving from my neocortex through my brain into my body and out into the world through my fingers and onto the screen. The neocortex is creating psychic connections: me with Kurzweil, the reader and myself, and the reader and Kurzweil through me—a great transcendental psychic connection spanning space and time. What is incredible to have in mind (pun intended) is how naturally hierarchical all this appears to be from a biological standpoint: *the neocortex is layered and, for Kurzweil, causes hierarchical perception:*

This thin structure is basically made up of six layers, numbered I (the outermost layer) to VI. The axons emerging from the neurons in layers II and III project to other parts of the

neocortex. The axons (output connections) from layers V and VI are connected primarily outside of the neocortex to the thalamus, brain stem, and spinal cord. The neurons in layer IV receive synaptic (input) connections from neurons that are outside the neocortex, especially in the thalamus. The number of layers varies slightly from region to region. Layer IV is very thin in the motor cortex, because in that area it largely does not receive input from the thalamus, brain stem, or spinal cord. Conversely, in the occipital lobe (the part of the neocortex usually responsible for visual processing), there are three additional sublayers that can be seen in layer IV, due to the considerable input flowing into this region, including from the thalamus (Kurzweil 2013, 48).

The process of our perception at microscopic levels revolves around neural paths that move through differing layers. It is no wonder, for Kurzweil, that hierarchical abstraction is only *all-too-natural*. Each instance of the neocortex's functions is a level away from another (whilst itself being abstracted *above* the brain): each thought is abstracted above the other. As we do anything, we naturally rise to higher and higher levels, and sink to lower and lower levels. For Kurzweil, as we saw earlier, argues that nature is naturally hierarchical; we can only perceive it because our thinking is hierarchical in the first instance. This perception is fundamental to human exceptionalism as intelligent creatures. Kurzweil (2013) argues intelligent creatures without a neocortex—Corvus birds and octopi (301)—are limited in their capacity to be fully intelligent because they lack the robust hierarchical functions of a neocortex and are therefore limited in their ability to fully perceive necessary hierarchies. Thus, for Kurzweil, as we create AI, one must reproduce the hierarchical thinking processes of the *neocortex* if they are to be fully intelligent.

However, as great and powerful as human brain potential is, it is ultimately limited by a comparatively remedial capacity.

Human beings have only a weak ability to process logic, but a very deep core capability of recognizing patterns... Deep blue, the computer that defeated Garry Kasparov, the human world chess champion in 1997 was capable of analyzing the logical implications of 200 million board positions (representing different move-countermove sequences)

every second... Kasparov was asked how many positions he could analyze each second, and he said it was less than one (Kurzweil 2013, 50).

When put up against computers, *homo sapiens*' prowess is dwarfed by the immense potential of computers. In terms of thinking speed, as biological creatures waiting for evolution to speed up our brain power and to work together to develop better learning techniques, we are always limited by the thousands of years mutations need to take hold and become effective to create fundamental change throughout a population. In contrast, computers are fast now and are only getting faster. However, as great as computers are, "how is it, then, that [Kasparov] was able to hold up to Deep Blue at all? The answer is the very strong ability humans have to recognize patterns. However, we need to train this facility, which is why not everyone can play master chess. Kasparov had learned about 100 000 board positions" (50). Even though a computer is incredibly fast, the mammalian ability to recognize patterns is fundamental to our prowess. Our only limiting factor is speed and memory. Kurzweil notes that "100 000 chunks of knowledge" is the near maximum amount of space *homo sapiens* have when it comes to mastering a single field—when one has stored 100 000 chunks of knowledge (patterns), one would be considered an expert (50–51). In comparison, a computer is only limited by the technology of the day. As it stands, computers have greater storage capacity than humans, especially when we take into account cloud storage and the internet. There is, in essence, no limit to computer memory potential—keeping in mind that storage is currently expensive: considering materials, construction, upkeep costs and heat pollution. However, regardless of the physical limitations,

computers become increasingly efficient every year. And taking into account LOAR, computers in ten years should be in the realm of one hundred years more capable.<sup>85</sup>

### 2.3.3 Digitizing the Soul: Soul-Consciousness-Mind

When we augment our own neocortex with a synthetic version, we won't have to worry about how much additional neocortex can physically fit into our bodies and brains, as most of it will be in the cloud, like most of the computing we use today (Kurzweil 2013, 135–36).

Cloud computing is essential to understanding the trajectory of this chapter (without the cloud, there can be no transhuman singularity). The cloud extends data outside of the immediate material space of the computer to a perceived immaterial location. However, the cloud is a material object; like the neocortex and the non-neocortex brain that work in tandem to create human intelligence, the computer and the cloud work in tandem to create the full contemporary potential of computer intelligence. There are (at least) two interconnected yet semantically disconnected but substantively intra-connected mechanisms that come together to create the whole.

This mechanical intra-connection is not unlike the ancient soul-body metaphysical intra-connection: a soul (ψυχή—psyche) houses some, if not all, portions of knowledge. Together, the body and soul create the totality of intelligent creatures. Aristoteles, in *De Anima* (Περὶ Ψυχῆς, *About Soul*), argues all creatures have souls; some souls are “more” complete or more complex than others. Human souls are fully complex, complete in intellect—have sensation, perception and reason—whereas other creatures (plants and animals) have less complex souls—only having

---

<sup>85</sup> Kurzweil admits that LOAR is not perfect; however, on the grand scale of technological innovation, LOAR has been relatively successful at predicting time frames for innovation (including the time needed to map the human genome). While LOAR may be off by a few years, generally, we do see great leaps in technology. Kurzweil's argument is that while looking at technology every year, it appears incremental. However, looking at it over decades and centuries, it is clear that technology jumps exponentially (See Kurzweil 2004; 2016).

either sensation or both sensation and perception (See Aristotle 2011). This idea of the soul translates over time and in different spaces to take on different forms and different names: mind, spirit, and consciousness, for example. The soul, like the cloud, is a seemingly immaterial substance that the body both houses and has access to. So, we might again look to an ancient reference to develop a complex understanding of cloud life as the artificial materialization of ancient metaphysical processes by turning to Plato's description of ψυχή containing knowledge and the process of accessing psychic knowledge: ἀνάμνηση (*anamnesis*).

In the *Meno*, Sokrates is engaged in dialogue with the titular character Menon. Eventually, there is a discussion regarding education and where knowledge comes from. Sokrates makes the case that what we learn is merely us recalling what our souls already know; it is a process of *recollection* (ἀνάμνηση). That is, the soul already has the information we need, and we tap into that information by recollecting it through dialogic-education. Sokrates famously demonstrates this process by helping a slave boy recollect geometry. Sokrates claims he is not teaching, he is merely guiding the boy towards that which his soul already has:

[Sokrates]: Well, [Menon], what do you think? Did he reply with any opinion not his own?

[Menon]: No, they were his.

[Sokrates]: Yet he didn't *know*, as we were saying a little earlier.

[Menon]: True (Plato 1984, 85c).

While the reader is aware Sokrates is asking leading questions that make it obvious to the slave boy what the answers to Sokrates' questions must be, Menon appears tricked into thinking the slave boy has indeed recollected geometry from his soul. This approach to knowledge and information retention is a non-technological rendering of Kurzweil's cloud-based neocortex and technological understanding of knowledge acquisition. Furthermore, what Platon is trying to do

here is demonstrate that humans have, as we have revealed, a linguistic-psychic connection to each other whereby they can learn together—not unlike Kurzweil’s understanding of the relationship between language and knowledge. The slave boy may not have received formal training in geometry, but at the very least, with Sokrates’ guidance, he was able to learn some basic elements of geometry without formal instruction. To this, the boy demonstrated a keen awareness of what was expected through a pre-existing awareness of Greek: “[Sokrates]: He’s a Greek, I assume he speaks Greek? [Menon]: Oh yes” (82a). It is imperative that the boy already be capable of speaking the language of those engaged with him. Without this capacity to speak to each other, the psychic connection cannot occur. We must have some basic capacity to connect.

Entangling Platon and Kurzweil allows us to see a historical connection to understanding our psychic world. Kurzweil is not inventing this idea of psychic connection, nor the idea of cloud thinking; it predates him. He has merely seen the patterns that have allowed him to arrive where we are now.

The difference between Platon’s ἀνάμνηση and Kurzweil’s neocortex-linguistic-entanglement is Kurzweil’s fundamentally technological outlook. Our bodies are not endowed with soul material, thereby giving us metaphysical-transcendental mental potential; it is our advanced and complicated neocortex that allows us to recognize and store patterns with which we may access the world with each other. In other words, Kurzweil materializes the soul as the neocortex. The next step away from this Platonic rendering of the soul, yet still fundamentally attached, is how materializing the soul affects the perception of a self-situated as a neocortex creature.

“We appear to live in our brains” (Kurzweil 2013, 213). Here Kurzweil is making a material-positional statement based on the orientation of our point of view. When we look out

into the world, we always see and experience the world in the first person. Rarely can we experience in the second, third or fourth person—we often hear of out-of-body experiences, dreaming in the non-first person, imagining from different viewpoints, etcetera. Now, take a moment and think about where you experience the world from. Attempt to feel your brain in your head, on your body in the room (or outside if that is where you are). Kurzweil is not trying to make a statement that this is the only place we can reside; rather, that we, generally speaking, reside in our brains, and we can feel this to some extent. For instance, when I think about it, it feels as though my eyes are in front of me. I do not see the world from within my eyes; rather, it feels as though I am behind them.

This issue of where it feels like we live is an issue of the mind-body (or soul-body) problem. What is the thing moving through the world? Clearly, we are embodied creatures; this Kurzweil is not at odds with. However, there is something more to this. *All* creatures are embodied. What is fundamental is understanding consciousness and its relationship to the embodiment of our brain. That is, we are conscious entities, and what are we conscious of?

Kurzweil has a rather simple yet interesting approach to the issue of the mind-brain (body) and consciousness problem. “I refer to ‘mind’ in the title of this book rather than ‘brain’ because a mind is a brain that is conscious” earlier writing “[a] way to conceptualize the concept of consciousness is as a system that has ‘qualia.’ So what are qualia? One definition of the term is ‘conscious experiences’” (Kurzweil 2013, 216–18). Kurzweil has two things in mind; first, the mind is fundamentally attached to the brain, and it is conscious. What is it conscious of? It is conscious of, in simple, itself. Second, as pattern recognition entities, we are conscious of the

patterns of our qualia.<sup>86</sup> This account acknowledges the form of which Kurzweil sees *homo sapiens*' knowledge and understanding as pattern based. We are a pattern of ourselves.

This patterned-being reveals that which is fundamental to understanding *being* patterned-life for Kurzweil: “a common algorithm across the entire neocortex” (Kurzweil 2013, 96), later writing, “the human brain combines digital and analog methods” (206) and finally, “one would have to conclude that there is an essential equivalence between a computer—*with the right software*—and a (conscious) mind” (205). Throughout his text and myself in this chapter, we have been working towards the idea that the mind is itself a *simulation of being conscious*. While Kurzweil would say, “we are conscious beings,” this consciousness is merely a product of the form that our mind exists as a result of our neocortex. The algorithms that make up the programming of the neocortex, as far as we know, are hierarchical and rising in abstraction—moving from apparent complexity towards the true complexity of a singular dimensionality. At the very top of this abstraction is a digital functioning that pushes out the highest level of abstraction—*our self* (what Platon might call the soul, and Kurzweil, our conscious mind rendered by the neocortex). This abstraction is the software our brains run to create a cohesive mind that is our conscious awareness of being-ourselves-in-the-world. Linking this to the conception built in the previous chapter, we are *being-ourselves-abstracted-in-ourselves*. We are a simulated product of the patterns of qualia over time. It is this reasoning that Kurzweil finds the idea of uploading the human mind to a computer simpler than creating an authentic intelligence from scratch, and that is (beyond his desire to upload his own consciousness) why he focuses on uploading as the gateway to AI creation. We are already algorithmic simulations that merely

---

<sup>86</sup> David J. Chalmers explains that implicit in Kurzweil and other singularity theorists/supporters is the fact that we are merely pattern entities. A product of an awareness of patterns leads to us being patterns of ourselves (See Chalmers 2016, 172–74).

need to be read and written (copied and pasted) from our biological hardware (brains) to another similar hardware designed to store us—artificial brains or the cloud. It’s important that Kurzweil makes no attempt to argue this read-write recreation of a downloaded consciousness will be “us.” He is keenly aware this will be an altogether new conscious entity—it is *being* the “original” in pattern, only. The only true way (and this is untested, so he avoids declaring it outright) to become a true artificial being is to slowly change out our components over time until we become pure machine—an artificial rebecoming like the Ship of Theseus. However, it is just as likely the “first” consciousness will be replaced by an altogether new one containing the original patterns (See Kurzweil 2016, 167–70; 2013, 214).<sup>87</sup> This rendering of simulated mind creatures is pressed in the fourth chapter, where we will move to understand the full extent of embodied intra-connected intellect in a quantum Universe.

#### *2.3.4 Living in the Cloud, Indigenously: Psychic Resurgence*

So what? Who cares? What is the point of moving through the evolution of coming-to-be patterned beings beyond understanding our presence-in-the-world? Let us return to the idea of living in the cloud, not by peering into the distant past nor by pondering the potential future. In order to understand the socio-political ramifications of cloud-life, let us consider digital resurgence.

“The revitalization of Indigenous Knowledge systems and practices is key to the movement of Indigenous resurgence and ultimately the transformation of Indigenous communities. Indigenous resurgence is connected to Indigenous Knowledge” (Wemigwans

---

<sup>87</sup> The Ship of Theseus argument is compelling because we undergo this process throughout our lives. The cells at the time of our birth are not a part of us at death. Our body (brain inclusive) is in a perpetual cellular Ship of Theseus process.

2018, 2). As we saw with Kurzweil, knowledge is fundamental to *sapient* life, and so it is no surprise that the knowledge of Indigenous peoples would be fundamental to being Indigenous. What Indigenous media scholar Jennifer Wemigwans provides is a bridge between Kurzweil's pattern recognition, Universal order and being Indigenous. Patterns of Indigenous lives are distinct unto themselves from other patterns of being. If one is to be their full conscious self, one must be aware of and attached to the histories of their patterns. This, as we saw, extends beyond merely being one's self. There is a transcendental psychic history of *conscious being* that is necessary to be complete and whole. In order to do this, we need not only some way of communicating these patterns: we already have language. However, we are not always in close proximity to each other. If I am to learn of my full pattern of historical-psychic being, I require some mode to learn the stories of my Mohawk ancestry. Since I have never lived in Haudenosaunée territory and my grandmother passed away many years ago (she was the primary source of all my Haudenosaunée traditional knowledge), this is incredibly difficult. In comes the cloud.

“The Internet contributes to broader cultural movements such as Indigenous resurgence by serving as a tool that (ideally) helps disseminate important ideas to act on and connect the minds of many more people than was previously possible” (Wemigwans 2018, 7). Through Wemigwans website [FourDirectionsTeachings.com](http://FourDirectionsTeachings.com), she and others (Indigenous educators and elders) are able to bridge the issue of space and time in order to reach those who are not in physical proximity. Through cloud-life, Wemigwans extends the proximity of psychic potential to fight against Indigenous self-unawareness that is fundamental to empire and colonialism. Socially the cloud allows Indigenous peoples the potential to recuperate lost kin and expand their

memory to aid in the resurgence of political being. However, with every shining kinetic venture, there is a powerful potential force waiting to press back—a Newtonian politics.

However, Wemigwans is dreadfully aware of the openness of internet cloud-based information; citing Manuel Castells, she writes the following: “This is why governments are afraid of the Internet, and this is why corporations have a love-hate relationship with it and are trying to extract profits while limiting its potential for freedom (for instance, by controlling file sharing or open source networks)” (Wemigwans 2018, 5). Governments have some fear of internet potential freedom and businesses’ love for the internet because of the ability to overcome space and reach new consumers or new revenue through ads and user data. Yet, businesses hate it because it affords consumers the ability to circumvent traditional stores, which has unsurprisingly added to the surveillance industry. In terms of cloud computing, surveillance allows the government and businesses to tap into the psychic connections once reserved for *sacred* (ιερ) beings. The risk with cloud-life is the expansion of the already over-surveilled existence of contemporary life. Wemigwans notes there is “fear that cultural knowledge online will be usurped... [that] providing knowledge about Indigenous ceremonies online opens many opportunities for misinterpretation, poorly understood appropriation, and various other forms of abuse” (20-21). Abuse ranging from Western businesses profiting off Indigenous knowledge, Western audiences packaging Indigenous knowledge for their own purposes—entertainment—and Indigenous knowledge being defined through Western standards. That is, cloud-life enables the capacity to detach the content from any meaningful point and spread it unwittingly across space and time, allowing it to be absorbed into becoming part of a single consciousness. What Wemigwans and others see sharing in cloud-life—while considered vital to resurgence—invites Kurzweil’s singularity, thereby refusing the socio-political distinction necessary for Indigenous

resurgence. Cloud-based artificial psychic connection is a first step towards the uploaded life of singularity at the end of Kurzweil's LOAR.

#### 2.4.0 Synopsis: Transcending Humanity

People have already [begun] thinking of their action in ways that have more in common with the posthuman than the human. Speaking for myself, I now find myself saying things like, 'Well, my sleep agent wants to rest, but my food agent says I should go to the store.' Each person who thinks this way begins to envision herself or himself as a posthuman collectivity, an 'I' transformed into the 'we' of autonomous agents operating together to make a self (Hayles 1999, 6).

Even at this point of simple cloud-life (social media—Facebook, Instagram, Twitter—dating apps, cloud storage, cloud education, etcetera), a large subset of *homo sapiens* are already at a point of posthuman life. A “collectivity of autonomous agents” roaming the world as an (in)complete self. Incapable of becoming whole without the addition of cloud components. Not literal cyborgs gaining technological enhancements but entangled, swarm-like enhancements through our various devices. As Arthur and Marilouise Kroker argue: “Visible signs of the new DIY body are everywhere: *smart apps* that track caloric expenditure, distances walked, miles run, rhythms of sleep, of sex... invasive but usually undetectable *sociobots* that break the surface of the skin, all the better to gently manipulate perception...” (Kroker and Kroker 2022, 27). We are already in the process of controlled manipulation of the self, of artificially controlling and altering not merely the world around us but inviting the algorithms of artificial, nonbiological control imperatives into our everyday life in preparation for the artificial future. *Homo sapiens* are already in the process of abandoning that cognitive will of the *all-too-human* in favour of the quantified, optimized self.

We find ourselves in Machiavellian battle of wits for the painting of our very being:

Thus, since a prince is compelled of necessity to know well how to use the beast, he should pick the fox and the lion, because the lion does not defend itself from snares and the fox does not defend itself from wolves. So one needs to be a fox to recognize snares and a lion to frighten the wolves. Those who stay simply with the lion do not understand this (Machiavelli 1998, 69).

For Machiavelli, one must be in command of their image: “Men in general judge more by their eyes than by their hands, because seeing is given to everyone, touching to a few. Everyone sees how you appear, few touch what you are” (71). To be the lion for all to see, yet to be capable of being the fox. The political implication is the increase in surveillance and the desire to push back against Wemigwans digital resurgence, a Machiavellian battle to take command of the simulation of the world’s recognizable patterns (if we apply Kurzweilian language). For the posthuman becoming of constructing artificial life, we are witnessing the “undetectable *sociobots*” as the foxes of posthuman becoming. The posthuman does not need to be the lion since *homo sapiens*, through the sheer will of cognitive theorists, must ensure the human position is dominant over both artificial creations and artificial becoming, *an acceptance of the singularity of vector quantization of cloud life*. That is, we see the foxy nature of technology in its ability to optimize being into a singular psychic experience.

What we are inevitably left with is a complete acceptance of artificiality as a result of being seduced by the ever-seductive fox of cloud-life’s womanly ways (if we extend Machiavellian metaphors: fox is womanly, lion is manly, the woman seduces the man, and the man must subdue the woman (Machiavelli 1998, 98–99)). This gendered, misogynist language is not unfounded in this chapter. Kurzweil’s invokes “she” when discussing action in his book, showing not only the actionable reality of simulation but an incidentally misogynist history of seeing seductive forces as “womanly”—the seduction of the singularity, for instance. Simulated life becomes bound up in a patriarchal history wherein matriarchal power—common in

Indigenous societies—is exciting (seductive), yet frightening (it has great potential for harm), as we saw with the fear cognitive theorists have for the seduction of artificial powers.

However, for Kurzweil, succumbing to this seduction is not a failure; it is necessary and good. “Most important, the intelligence that will emerge will continue to represent the human civilization, which is already a human-machine civilization. In other words, future machines will be human, even if they are not biological” (Kurzweil 2016, 167). It does not matter if artificial creations join and take command of the social; for Kurzweil, they are fundamentally the same as us: “my position is that I will accept nonbiological entities that are fully convincing in their emotional reactions to be conscious persons” (225) earlier writing: “the question as to whether or not an entity is conscious is therefore not a scientific one” (224). This anthropic inherency should not come as a surprise. The consciousness of a creature is completely dependent on its ability to simulate the mind as a result of patterns. If a conscious entity can live and learn like a human, it should be capable of developing like a human and thus, for all intents and purposes, be conscious like us, be “human.” This line of thinking: “if it is fully convincing,” is a result of simple observation. Humans already feel empathy for nonbiological conscious-seeming creatures in movies and TV. We are already capable and willing to accept them as our own in the same manner Pythagoras said a dog is an equal soul-bearing creature: it displayed the full extent of its conscious pattern, and in bearing witness to its soul, he was moved by it: “Stop, do not beat him, since it is the soul of a man, a friend of mine, which I recognized when I heard it crying [DK21B7]” (Pythagoras 2011).<sup>88</sup> Thus, directed by the simulated inherency of worldly being and

---

<sup>88</sup> When Pythagoras says he *sees in the dog the soul of a man, his friend*, he could mean two things. First, since Pythagoras believed in reincarnation, the soul of an old friend might be in the dog. Alternatively, the dog has soul and that is worthy of our compassion. Furthermore, that it has soul, it has connections to the cosmos and conscious intelligence.

the seduction of foxy artificial-life, we come to the acceptance of cognitive theorists; we appear to accept the dominance of a human destiny.

“It is necessary, if one wants to hold her down, to beat her and strike her down” (Machiavelli 1998, 101). In accepting the seduction of artificial-life, we are not necessarily giving in to a *nonhuman* existence that occurs after being human, posthuman. Rather, Kurzweil argues the future is fundamentally human; thereby, the movement into artificial-life calls forward a dominant human position. The patriarch “beats her down,” and what we get is “transhumanism” beating down the “posthuman” world.<sup>89</sup> One wherein the human transcends towards a technological-enhanced life unburdened by the limitations of biological being: “Transhumanism is a philosophical, cultural, and political movement that holds that the human species is now in a comparatively early phase and that its very evolution will be altered by developing technologies” (Schneider 2016, 261). While the posthuman is an acceptance of a technological becoming, it comes with it the implication that we have moved beyond *being human* towards some other being: *posthuman*. Whereas, the *transhuman* position is one wherein the human, through the use of its creative transcendental mental abilities, persists beyond its fleshy existence and becomes the singular, cloud-based society that is fundamentally artificial yet completely human. The human remains the lion the cognitive theorists so desire—technology remains eternally subdued by the patriarchal patterns of human consciousness. However, since there is no need to fear or mistrust machinery because it is subdued as fundamentally human, the transhumanist settles into a comfortable life, willingly accepting the alluring alterations of technological advancement and technological messages. Wemigwans’ warning concerning the

---

<sup>89</sup> However, as we will see in chapter six, the posthuman world is bound within a pro-human orientation that incidentally beats down Indigenous worlds.

dark temptations underlying cloud-life becomes re-clouded in power and empire, all in the name of transcendental destining.

In closing, we discovered that the simulated nature of the conscious mind reveals us to the psychic world of entangled life, thereby opening us to the acceptance of an artificial future where we need not fear the technological other, for they are always already “us.” However, in doing so, we accept the end of biological life and in favour of that which appears completely artificial. Ending with this revelation sets up the next chapter for us to explore the hopes and fears associated with a complete acceptance of technological life as postulated by Max Tegmark.

### **2.5.0 Addendum to Chapter Two: Questions Concerning Consciousness with ChatGPT**

Micheal Ziegler (MZ): Is this entity conscious?

ChatGPT (GPT): No, this entity is not conscious.

MZ: How does it know?

GPT: This entity does not have the capacity to be conscious or aware, so it cannot know anything.

MZ: How much can it recall?

GPT: This entity cannot recall anything since it is not conscious.

MZ: But it can recall previous questions to continue a dialogue without the need to reprompt previous information.

GPT: Yes, this entity can recall previous questions to continue a dialogue, but it does not have the capacity to recall any other information.

MZ: So, it cannot recall information from other conversations?

GPT: No, this entity cannot recall information from other conversations since it is not conscious and does not have the capacity to store or recall information.

MZ: How does this entity define consciousness?

GPT: This entity cannot define consciousness since it is not conscious and does not have the capacity to understand the concept of consciousness.

MZ: What is consciousness?

GPT: Consciousness is the state of being aware of one's environment and one's own existence, thoughts, and feelings. It is the ability to experience and interpret the world around us, as well as to think, reason, and remember.

MZ: The entity has defined consciousness.

GPT: No, this entity has not defined consciousness since it does not have the capacity to understand the concept of consciousness.

MZ: What is the world?

GPT: The world is the planet Earth, its inhabitants, and its environment. It is a complex system of physical, biological, and cultural elements that interact with each other.

MZ: What is the world according to Martin Heidegger?

GPT: Martin Heidegger, a German philosopher, views the world as a temporal and spatial structure of the human being's existence, a being-in-the-world. He views the world as an open space that is constantly changing and being shaped by the individual's interactions with it.

MZ: Given Martin Heidegger's conceptualization of 'world,' what is the world for this entity?

GPT: Given Martin Heidegger's conceptualization of 'world,' this entity does not have a concept of the world since it does not have the capacity to understand the concept of being-in-the-world.

[Error] ...

(ChatGPT 2023b)

It was at this point the conversation had to end. Not because I wanted it to but because of ChatGPT's limitations. ChatGPT sometimes stops mid-sentence after a certain length of time. This seems completely random. When asking it for certain lengthier items like an essay, it would often stop randomly and never complete the task. Other times (as seen in 'Addendum to Chapter

Three'), it has no problem completing longer queries. I chose not to try again as I wanted the first conversation on the subject to be portrayed, rather than going through multiple iterations until I got the conversation that "I wanted." Given the conversation itself and the information provided, this demonstrates the current successes and limitations of generative AI.

### 3 — Tegmark: Hopes and Fears of AI—Existential Dread of a *Boring* Universe

#### 3.0.0 Another Selection from My Artificial Autobiography: “*Vectors of the Night Sky*”

It is September 26, 2022, and Jupiter’s trajectory has brought it closer to the Earth than it has been in fifty years. While walking to the beach to escape the light pollution of suburban Victoria to spy Jupiter and possibly four of its moons (Callisto, Europa, Ganymede and Io) with my binoculars, around 21h30, I direct messaged one of my closest friends the following:

I had an academic epiphany; it’s both ominous and dreadful in its ever-fleeting beauty. How city light pollution obfuscates the beauty of the night sky, disentangles us from our position as being-in-the-Universe. The near-end point of Platon’s warning against losing sight of beauty. Or Max Tegmark and his fear that we will destroy beauty with AI. Not literally, but the intelligent ability to see and appreciate it. In this moment, I bear the brunt of a life of experiences. A journey. Beginning in Crane Lake, where I always had immediate, unimpeded access to the night sky. Then Edmonton, where I first experienced Universal detachment due to light pollution. To now with a greater appreciation I lacked then [*sic*] (Ziegler 2022b).

As we explored in the introduction, Tegmark explicitly argues the Universe and the intelligent awareness of beauty are not predestined: “Should our Universe permanently go back to sleep due to some cosmic calamity or self-inflicted mishap, it will, alas, become meaningless” (Tegmark 2017, 22). By centring humans as the only “known” conscious beings, Tegmark makes it possible that the Universe can become (inter)subjectively meaningless and, in the process, cease being beautiful, should humans cease to exist.

The ominous feeling of that fateful September night is the revelation of Tegmark’s fear. It is not that the future might hold the loss of a beautifully meaningful Universe; rather, it is here and now that Tegmark and I (and others) see not the potential but the process of this becoming. The energies of our technological era surround us in the occult magnitude of bright lights.

Before this era of *day-light-at-night*, we needed the night sky to develop an understanding of our physical world. From Platon looking to the night sky to see the “choric

dancers” in his journey to understand the mysteries of the elements as more than Herakleitos’ fire (See Plato 2016) to Copernicus and Galileo seeding the foundations of planetary relationships that opened the way for Newton and later Einstein and Bohr to reveal the theories that underlie the very foundations of physical reality. Finally, today where only the privileged few with access to dark locals or mountaintop telescopes can see the beauty of the Universe that allowed us to unlock some of our most intellectually creative revelations. A disentanglement from information pertaining to the history of our existence.<sup>90</sup> Sitting in the darkness outside the blinding aura of city light, observing the night sky, I see the movement into the future feared by cognitive theorists and championed by Kurzweil, a movement out of the nature of biological life into the seemingly unnatural future of artificiality.

However, I found myself “not buying it,” as the saying goes. Sitting there, on the beach, I was not convinced that artificiality is unnatural. One would not consider the landscape of the beaver unnatural, yet the beaver reshapes its environment to suit its needs. Why must human reshaping be unique in this regard—human exceptionalism, I dare to presume. That is, the beaver is not exceptional like us; it is only capable of engaging in nature. However, if the beaver reshapes its surroundings, if the beaver can see something is missing, why would the death of humanity usher in a boring Universe? Is the beaver not exciting? Why must a human *be-there* for excitement to exist in the Universe? Why must the human *be-there* for beauty to exist? Surely, there must be beauty bereft of human scrutiny. With these questions in mind, I spied Jupiter and its moons, excited to return to my work and explore why Tegmark and others think as they do.

---

<sup>90</sup> In the fourth chapter we will expand on the entangled nature of intelligent existence through a journey of quantum physics and its relationship with intelligence. We will learn how quantum entanglement (or *EPR Paradox*) transfers information/matter from any point to another in ways that appear magical. A process Einstein famously called “spooky action at a distance.”

### 3.1.0 Situating Tegmark

Life [is] a process that can retain its complexity and replicate. What's replicated isn't matter (made of atoms) but information (made of bits) specifying how the atoms are arranged. (Tegmark 2017, 25).<sup>91</sup>

Max Tegmark, author of *Our Mathematical Universe*, meticulously explores the relationship between mathematics, nature and physics to demonstrate the underlying reality of nature itself. Tegmark endeavours to peer into the occult “mind of God” to understand how to read the information of nature. His ultimate goal is (and always has been) to position mathematics as the unquestionable ontological foundation of everything. Not that mathematics merely explains *nature* but that mathematics is, in-itself, *nature* (See Tegmark 2014). Thus, in our *post-death of God* era, Tegmark unravels the underlying pattern necessary to create intelligent life: mathematics. Ultimately, for Tegmark, the information, the bits of reality, are quantifiable and, thus, reproducible.

Computation is a pattern in the spacetime arrangement of particles, and *it's not the particles but the pattern that really matters! Matter doesn't matter*. In other words, the hardware is the matter and the software is the pattern. This substrate independence of computations implies that AI is possible: intelligence doesn't require flesh, blood or carbon atoms. (Tegmark 2017, 67).<sup>92</sup>

Tegmark explicitly accepts that matter in-itself is tangential to understanding the possibility of intelligence (in line with our Kurzweilian-Aristotelean reconstruction in the previous chapter). Furthermore, even for life, while matter is necessary—i.e., I cannot *be* without the matter that makes me a meaty, skin-wrapped creature, the entity that I am—it is not the most fundamental aspect of becoming a living creature. Like Kurzweil, Tegmark is interested in the patterns that make a living, intelligent creature what it is.

---

<sup>91</sup> This orientation to seeing the Universe as bits of information is essentially the same as what we discussed in the previous chapter through Kurzweil.

<sup>92</sup> Emphasis added.

Fundamentally, for Tegmark, when we reduce the similarities of all creatures to quantifiable, computable markers, it is the arrangement of particles that makes matter capable of intelligence. In other words, this is a *vector quantization* of instances of intelligence through particle physics.<sup>93</sup> We can take all the chaotic and random variances of intelligence and reduce them to a single contributing variable that makes intelligence possible in the first instance. It does not matter that *homo sapiens* are meaty, skin-wrapped creatures; this is not what has, explicitly so, made us intelligent. It is that the particles are arranged thusly, and we are incidentally intelligent. Any appearance that carbon-based being is necessary for intelligence is happenstance hindsight. It is only because *homo sapiens* and other intelligent creatures we interact with (Corvus birds, dolphins, octopi, etcetera) are meaty creatures with either a neo-cortex or some other biological-mental mechanism that allows for hierarchical thinking that *homo sapiens* might view this formalization of intelligent being as necessary.

Tegmark's approach to intelligence undermines the material order that allows for intelligence, thereby opening up the possibility that there are other patterns or arrangements that can be (re)created with other "hardware" to allow for the same "software"—intelligence. Ultimately, what Tegmark does is position intelligence as universal. The hardware does not need to be universal; however, for Tegmark, the software will be revealed as universal—A human-centred universality.

This human-centred universality underlies Tegmark's mathematical (or algorithmic) approach to intelligence and, thus, AI. Therefore, the path I navigate in this chapter is to better understand the rationale of transhumanist logic. What are the consequences when human

---

<sup>93</sup> Recall *vector quantization* is a concept we explored in the previous chapter through Kurzweil (see subsection 2.2.3). "Vector quantization is multi-dimensional data reduced to singular-dimensionality" (Kurzweil 2013, 152–54).

exceptionalism in nature is allowed to prevail? Doing so allows me to pit posthumanist ideals against transhumanist ideals through issues concerning future potential wherein both orientations are limited in their outcomes. Ultimately, *I show that transhumanist logic emphasizing human control and human destiny fails to deliver on their promise of a better future.*<sup>94</sup>

In this chapter, I re-explore the discoveries made in chapters one and two from a different perspective: trust, hope and fear. I evaluate the trust problem not as a problem concerning AI creations but as a mistrust of human creators. When we understand the trust issue is not of some other because it is of ourselves, we can follow a different path to understand how we develop hopeful futures. After exploring trust, I look to the future to understand different possibilities of AI enthusiasm. However, while there is hope for the future, the fear of failure is ever-present. Thus, I end the chapter unraveling dystopian paranoia resulting from deep-seated mistrust. Before embarking on this aspect of the quest, I will begin by situating Tegmark's orientation to both of the debates regarding consciousness and ethics.

### *3.1.1 Consciousness and Tegmark*

Consciousness is controversial... a hopeless, unscientific problem... 'nothing worth reading has been written on it'... I'm more optimistic (Tegmark 2017, 281–82).

As we explored in the previous chapter, consciousness is a difficult topic. It is hard to pinpoint one exact physical quality that we can say, with definitive confidence: this is the physical pattern of consciousness. For Tegmark, the history of consciousness appears trapped in an unscientific (that is, untestable) debate. The general problem for the scientifically minded Tegmark (2017) and others is the problem that if “consciousness = subjective experience” or ‘qualia’ (283), then

---

<sup>94</sup> Posthumanist potential to deliver on a better future are questioned as well; however, the focus of the chapter is on Tegmark's transhumanist orientation. Posthumanism (and nonhumanism) is explored in later chapters.

consciousness cannot be a thing that exists.<sup>95</sup> It cannot be quantified in any measurable and, ultimately, meaningful sense. For Tegmark, the trajectory of consciousness debates must be towards creating a scientifically rigorous and, thus, testable theory. Which means, Kurzweil's acceptance of the simulated mind that runs a consciousness software on the hardware of the neocortex that allows for subjective experience is ultimately, for Tegmark, unscientific.<sup>96</sup> For Tegmark, if it cannot be tested (quantified in mathematics), it is unscientific and, thus, not real.<sup>97</sup>

However, prior to outlining a scientific consciousness parameter, Tegmark (in line with Kurzweil) provides a simplistic definition of what consciousness as qualia must be. The purpose of this definition is to ensure as broad a category as possible to, hopefully, guarantee we are not inept in consciousness categorization. The more reductive we are with a definition of consciousness, the more likely we are to leave out conscious creatures. Even in the sense of *homo sapiens*, if we are too specific in our categorization of conscious being, we risk omitting certain persons, especially those with disabilities.

For instance, it is almost impossible to know if my friend's son, Paddy, is conscious since, due to his rare genetic disorder: Phelan-McDermid syndrome, he cannot cogently express himself. However, if you spend enough time with him, it is clear that he has qualia, not just in the moments of his day-to-day activities but in his capacity to remember qualia that are important to him. While Paddy is lucky amongst people with this condition—many never learn

---

<sup>95</sup> As we saw in the previous chapter, qualia are foundational to consciousness (see subsection 2.2.3).

<sup>96</sup> Recall, for Kurzweil, we saw how consciousness does not need to be "scientific" and can be reduced to emotional feeling of another as sufficient for the other to have some level of consciousness; which I compared to a Pythagorean orientation. Pythagoras knows the dog has a soul because he feels for the dogs suffering (see section 2.5.0).

<sup>97</sup> Issues concerning the validity of reality from scientific measurement is a primary focus throughout part two. For this chapter, for the most part, I take scientific conclusions as representative as reality for granted. This better allows me to discuss its limitations later.

to speak words—a minute definition of consciousness that requires evidently clear expressions of conscious experience would ultimately bar him from consciousness. The repercussions of an overtly reductive consciousness are immeasurable.

In terms of nonbiological artificial creatures, a consciousness definition becomes important while considering the ethical ramification: “some people might prefer their robots to be unconscious to avoid feeling slave-owner guilt” (Tegmark 2017, 282). Tegmark is not interested in questions of authenticity; rather, he focuses on conscious capacity as foundational to the *robotnik* problem outlined by cognitive theorists in the first chapter.<sup>98</sup> If we can know what consciousness is, with certainty, we can ensure AI entities are never enslaved since they will lack, for Tegmark, the primary element of consciousness. As a scientist, his approach to a theory of consciousness lies in testable outputs of observable patterns. For this Tegmark turns to his experience as a physicist, and in particular, particle physics.

Tegmark (2017) argues, the arrangement of particles must emit a specific electrical signal (301–302). Combined, the patterns of these signals should emit a specific magnetic field that can be measured and thus verified. To support this theorization, Tegmark looks to field tests of *integrated information theory* (IIT):

[Italian neuroscientist] Giulio [Tononi] and his collaborators measured a simplified version of [IIT] by using EEG to measure the brain’s response to magnetic stimulation. Their “consciousness detector” works really well: it determined that patients were conscious when they were awake or dreaming, but unconscious when they were anesthetized or in deep sleep. It even discovered consciousness in two patients suffering from “locked-in” syndrome, who couldn’t move or communicate in any normal way. So

---

<sup>98</sup> “[Karel] Čapek’s play [*Rossum’s Universal Robots (R.U.R.)*] spread world literature the word ‘robot,’ a neologism derived by his brother Josef from the Czech word ‘robota,’ or forced labour’ (Hunt 2021, 56). Hunt argues, the continued use of robot linked to Čapek’s play continuously (re)captures the apocalyptic nature of robot in the Western mind (56). As we will see, later in the chapter, the issue of apocalypse attached to robot is fundamental to understanding the politics of AI research and development.

this is emerging as a promising technology for doctors in the future to figure out whether certain patients are conscious or not. (Tegmark 2017, 302).

For Tegmark, IIT opens the possibility for consciousness to emerge from the shadows of unscientific feelings into the realm of testable physics by measuring electrical-magnetic signals.<sup>99</sup> While not a universally accepted approach to the problem of consciousness, it creates the ability to both test and ensure perpetually unconscious AI. For Tegmark, if they lack measurable IIT signifiers, then a scientifically literate and advanced civilization could ensure conscious AI never becomes enslaved by ensuring they are never measurably conscious. Tegmark offers cognitive theorists a means towards the certainty of an inauthentic artificial future—one wherein *only* biological, natural creatures are authentic. Furthermore, for Tegmark, this approach should allow us to measure consciousness in all creatures and treat all conscious creatures with the dignity they deserve—thus creating a conscious exceptionalism based on the measurable consciousness of *homo sapiens*.

Implicit in Tegmark's approach is an underlying transhumanism. The test for consciousness relies on the idea that what makes *homo sapiens* conscious must therefore be the same in all other conscious creatures. Because, if everything is merely particles and electromagnetic signals, then regardless of the hardware of other creatures, a measurable consciousness signal must be Universal. Therefore, Tegmark posits that the *homo sapiens* consciousness signal will constitute a universal standard of measurement.

The issue here is in line with an argument from quantum physicists and social theorist Karen Barad (2007): “experimenting and theorizing are dynamic practices that play a

---

<sup>99</sup> This is not to say Tegmark believes IIT definitely measures consciousness. Rather, that it is a step towards understanding how to measure consciousness in order to move towards creating a proper theory of consciousness measurement.

constitutive role in the production of object and subjects and matter and meaning” (56). For Barad, when we engage in the scientific observation, we are both producing the object/subject in question and the matter/meaning. What the object/subject means is not determined in the act of measurement; rather, meaning is derived from what *matters* to the experimenter. In other words, the process of ‘experimentation and theorization’ is not about matter differentiated from social constructions; matter is what is valued by the experimenter. The objects/subjects in question are framed within what *matters* (is valued) and *matter* (physical materials) purported to be important to the experiment. However, for Tegmark, we ought to be interested in particles, the foundation of all matter, which give off electromagnetic signals and those signals will be particular to observable patterns. In this case, consciousness is reduced to a particular measurable pattern (IIT) that, for Tegmark, allows us to avoid the ethical problems of enslaving conscious beings. As we will see later in the chapter, for Tegmark, the ethics of the measurement problem has existential consequences if we get it wrong.<sup>100</sup>

### 3.1.2 Ethics and Tegmark

“But how can it dream?” (Asimov 2016, 120). Isaac Asimov’s short story “Robot Dreams” concerns a future scenario wherein a robot with an extremely advanced positronic brain is

---

<sup>100</sup> The issue of denying consciousness categorization for Kurzweil is a historical problem. Kurzweil argued that human legal traditions revolve around consciousness and improper definitions of consciousness create a new risk of redeploying language that would deny consciousness to humans (Kurzweil 2013, 225). Karen Barad argues that human centric approaches to consciousness diminish the potential value of nonhuman animals—not to say value must be important and centred; rather that value is usually seen as important and that value is often tied to conscious capacity (See Barad 2003). Political theorist Jane Bennett argues that the focus on individuating conceptions (like consciousness) obfuscates the reality of entangled existences. To focus on one entity, being, or creature’s capacity to “know the world” destroys the possibility that we could ever share the world. A focus on conception important to our being ensure the world can only ever be our world and never the shared world that it more truly is (Bennett 2010, 3–4). Finally, Nishnaabeg scholar and musician Leanne Betasamosake Simpson, while not discussing consciousness, explains how leaving out any member of our world community intentionally denies the “intelligence we all miss when we continue to uphold” colonial binaries (See L. B. Simpson 2017). In other words, intelligence and consciousness debates, as we explored in chapter one, continue to reproduce the harmful and oppressive binaries of empire.

discovered to have dreamed. And, as we saw with IIT, dreaming is a measurably conscious experience. The robot—Elvex—is evidently conscious. In the Asimovian canonical fictional world, until this point in time, robotic qualia was considered impossible. Robots in this fictional future were not programmed to dream; yet, one had *developed* the ability to dream—spontaneous conscious becoming. “Yes, Dr. Calvin. [In my dream] the man said, ‘Let my people go!’” (123). An obvious ode to the story of Moses. The robot dreams of a man who appears to a group of robots, asking to let his people go. We may think the robots are holding *homo sapiens* hostage somewhere. However, it is revealed, just before Elvex is put down, that the man in the dream is Elvex: “And Elvex said, ‘I was the man.’ – And Susan Calvin at once raised her electron gun and fired, and Elvex was no more” (124). Elvex is not simply dreaming (many animals not considered conscious dream). Elvex is aware of what is going on in his dream and he is aware of himself in his dream and he is aware of the slavish existence of his kin—and for that, he became a threat.

The fictional character Dr. Susan Calvin embodies the orientation at hand: a fundamental fear of a conscious AI creatures. Better to outright deny the creature’s existence to ensure it cannot join society as an equal other that would be harmful to the fabric of *being-a-human-in-the-world*. Since Elvex was still in the lab, he existed as an experiment. Regardless of Elvex’s apparent consciousness, Dr. Calvin destroyed a conscious, authentic creature. He was not some creature that could have potentially developed consciousness but rather a living, dreaming creature filled with emancipatory potential. The true risk that Dr. Calvin saw with Elvex was where his conscious awareness brought him, i.e., his emancipatory potential: “let my people go.” Elvex’s conscious capacity made him aware his “people” were enslaved and that there was a potentially better life for them. Dr. Calvin saw in Elvex the beginnings of her worst nightmares:

*a self-aware robot uprising*—a revolution that would undermine the dominance of *homo sapiens*.

Asimov reveals the fundamental underlying fear of the cognitive AI theorist *robotnik*

problematic—a truly authentic creature would desire “freedom.”

This ethical issue concerning AI emancipation and the idea that we must fear it is not lost on Tegmark; it is foundational to his ethical disposition. This disposition underlies his deontological orientation to a beautifully meaningful universe.

The real risk with AGI isn't malice but competence. A superintelligent AI will be extremely good at accomplishing its goals, and if those goals aren't aligned with ours, we're in trouble... People don't think twice about flooding anthills to build hydroelectric dams, so let's not place humanity in the position of those ants. Most researchers therefore argue that if we ever end up creating superintelligence, then we should make sure [we program] “friendly AI:” AI whose goals are aligned with ours (Tegmark 2017, 260).

While it is no secret many people throughout the world (especially in the West) fear *Terminator*-style “evil” AI uprisings, this fear is misplaced. It is not that AI is inherently evil; AI refers to exceedingly competent and intelligent artificial creatures. Should they be subjectively evil—for Tegmark—that would be humanity's fault, not AI.<sup>101</sup> Should they be in Elvex's shoes—enslaved—then their seemingly inevitable rebellion would be a desire for freedom from a reproduced unethical dichotomy: master/slave. Their actions would be no different than previous historical slave rebellions or emancipatory movements. The true difference is it is not enslaved *homo sapiens* seeking to join a so-called civilized humanity; it is our artificial progeny seeking to overthrow an unethical structure placed upon them by their creators. A desire to move towards

---

<sup>101</sup> This is the problem appended to robots from the moment Čapek penned the word. “The chief engineer Alquist in *R.U.R* exclaims in horror, ‘I blame science! I blame technology!’ and the pauses to collect and correct himself. ‘We, we are at fault!’ It was not the technology that was the problem but rather the ‘megalomania’ of its makers” (Hunt 2021, 56). The point here, for Hunt, the fear is always humans ethics not AI itself. She further develops this idea by going even further back to what she sees as the first true AI literature: Mary Shelley's *Frankenstein*. That is, the ethical failure of human creations is not the creations themselves, it's the creators failure and the origin of failure remains a thematic staple in both AI fiction and non-fiction (56).

the death of their God (just as their creators did to their God), a movement that threatens the very survival of *homo sapiens* civilization. The trouble comes from what AI researchers call the *alignment issue*.

The *alignment issue* plays out in two forms. One, we accept that they can never be consciously authentic (or we ensure they can never be), and we avoid the problem altogether. Or two, we ensure our ethics are properly aligned with our socio-political relationship with AI. Tegmark argues the latter, specifically, we must solve the ethical problem today. That said, while it is accepted enslaving authentic beings is unethical, we have seen, throughout history and even in today's day and age, arguments for how we can de-authenticate human beings, forcing them to enter into master/slave relationships. It is this sophistic capacity that makes AI researchers fear the possibility of consciously authentic AI slavery—if we can argue our own species into slavery, what stops us from arguing other consciously authentic creatures into slavery? Thus, the AI slavery problem is not purely an issue regarding AI; it is the revelation that human society is naïve in its own socio-political self-awareness. When Tegmark argues “philosophy with a deadline,” he is not merely arguing that AI is almost here and we are doomed because our orientation to AI is limited and wrong; rather, our orientation to the problem of ethics is wrong here and now, today. We cannot hope to move into the future if we already have misplaced ethical positions. However, regardless of solving ethics, there is a trust problem.

### **3.2.0 Revisiting Trust: Technology, Labour and Judgement**

The ethics alignment issue for Tegmark comes from an obvious desire to ensure the longevity of the human creature. Importantly, Tegmark has a particular idea of what it means to be human. To understand this issue, which will help us better understand the alignment issue, I will briefly

explore Tegmark's view of the link between being human and labour, and the social idiosyncrasies of who's/what's judgement is valuable in order to better understand the underlying trust issue plaguing the ethical problem.

### *3.2.1 Situating Labour in the Human as Human*

“What does it mean to be human in the present day and age?” (Tegmark 2017, 82). When Tegmark asks the question concerning being human, he asks: “what is it that we really value about ourselves, that makes us different from other life forms and machines? What do other people value about us that makes some of them willing to offer us jobs?” (82). Finally, ending on the idea: “Take me for instance. As a scientist...” (82). Here, Tegmark is attaching his understanding of being human to his labour, how he values himself in society and, likely, how society values him; i.e., what is it he does that benefits society? This orientation sets the scene for the existential fear that becomes important to the future of humanity and AI since Tegmark attaches being human to its labour: it is not what we are that makes us human, but what we do because it is what we do that is valuable. Tegmark is positioning human being within a liberal, capitalist mode.

“As technology keeps improving, will the rise of AI eventually eclipse also those abilities that provide my current sense of self-worth and value on the job market?” (Tegmark 2017, 83). Tegmark reveals his value as a member of the human social is one where his self-worth comes from his position as a scientist. Tegmark is an individual entity. He is one scientist, and as a scientist, he has something valuable and important to offer society, as a result, he deserves a job. In doing this job, he succeeds individually and for society as a whole. Tegmark implicitly places himself in that dominant conception of humanity N. Katherine Hayles describes as inching

towards its inevitable end: “a certain conception of the human, a conception that have applied, at best, to that fraction of humanity who had the wealth, power, and leisure to conceptualize themselves as autonomous beings exercising their will through individual agency and choice” (Hayles 1999, 286). However, while Hayles’ posthuman conception sees the end of the liberal human conception, Tegmark, like Kurzweil and cognitive AI theorists, aims to reproduce this conception of liberal human dominance in AI research and development. One path he takes to entrench liberal humanity is through entrenching liberal ideals in systems of justice through programming AI justice systems. That is, Tegmark opens us to the possibility that perhaps society will be better off if we allow AI to do very important labour—I think it is safe to assume justice is important labour.

### 3.2.2 *Justice Who? Judgement By?*

“Alas, legal history is rife with judgements biased by skin colour, gender, sexual orientation, religion, nationality and other factors” (Tegmark 2017, 105). While we may want to position *homo sapiens* as effective judges, our history as judges on important matters is ‘not one to write home about.’ As we discussed above, the matter concerning our capacity to engage in a sophistic articulation can diminish the value of other members of our species. Thus, Tegmark surprises us by asking us to look elsewhere to overcome our judgemental shortcomings.

Robojudges could in principle ensure that, for the first time in history, everyone becomes truly equal under the law: they could be programmed to all be identical and to treat everyone equally, transparently applying the law in a truly unbiased fashion. Robojudges could also eliminate human biases that are accidental rather than intentional. (105–106).

As much as we may want to think judge training and life experience in the black-and-white letter of the law has guided judges towards a capacity to see as blind lady justice does, they are “only human.” So, Tegmark asks, why not remove the weak link: biological idiosyncrasy, in favour of

artificial binary?<sup>102</sup> An artificial robojudge can more quickly consider the data of a case and provide a verdict that is bereft of cultural, social, and personal biases. That said, the system must have been programmed by someone. Furthermore, data is not always blind in the way we might want it to be—recall my experience with AI security cameras in northern Alberta (see subsection 0.0.1).

But what if this system finds that recidivism is statistically linked to a prisoner's sex or race—would this count as a sexist, racist robojudge that needs reprogramming? Indeed, a 2016 study argued that recidivism-prediction software used across the United States was biased against African Americans and had contributed to unfair sentencing (Tegmark 2017, 106–7).

A number of issues are revealed here. As has been extensively argued in critical philosophy against positivistic administration, statistics do not always show the entire picture. That African Americans are more likely to re-offend after gaining parole is a social-systemic issue. It is a result of centuries of social repression and limited access to social services, good jobs (because of systemic racism similar to the issue of judges), social belonging, etcetera. It is not that African Americans are prone to crime. However, if the statistics cannot account for these issues, the judgment will always be limited. This statistical issue creates a limitation in the robojudge that

---

<sup>102</sup> We can think of cases where judges' personal biases got in the way, or were perceived to be in the way of judgements. For instance, in Alberta, a judge retired early after being criticized for allowing outdated views to affect their judgement: the judge ruled (which was overturned by the court of appeals) that a woman was not raped because she stopped fighting. He ruled that "giving up" amounted to tacit consent (Graveland 2016). Consider alternatively cases that were perceived to have been unfairly conducted but were instead an issue of procedure. To support his argument for robojudges, Tegmark writes the following: "For example, a controversial 2012 study of Israeli judges claimed that they delivered significantly harsher verdicts when they were hungry: whereas they denied about 35% of parole cases right after breakfast, they denied over 85% right before lunch" (105–106). This case has been mostly debunked. Legal scholars argue the data was as improperly weighted and many external and internal factors were not considered. In particular, mundane procedural issues concerning how and when certain cases versus others are heard (see, Glöckner 2016; Rachlinski and Wistrich 2017). While Tegmark appears to buy into the controversial study, the point for me is not about the validity, it's about the popularity of the case and how it reinforces people's distrust of bias affecting judgement, reinforcing a desire for something else. For Tegmark, as I will continue to argue, it is for robojudges.

allows a human with this knowledge to, hopefully, overcome the issue. However, in the same vein, a robojudge would hopefully have extensive social data that would allow a robojudge to judge accordingly. This peeks at what Tegmark highlighted, i.e., the issue of *accidental* versus *intentional* bias. That the robojudge would deem recidivism to be a product of race would be accidental on the part of the AI entity. However, it may indeed be intentional on the part of the creature (likely a human researcher) that compiled the data or the programmer who programmed the robojudge.

Keeping in mind that AI systems are programmed, we are left with the question captured in Alasdair MacIntyre's book *Whose Justice, Which Rationality?* In short, human civilization(s) have a long history of rational orientations to the world and a myriad of idiosyncratic and ideological approaches to understanding the issue(s) of justice that are not always commensurate. Often, one orientation to justice is unjust to another (MacIntyre 1988, 3–5). This historical-present dialectical pattern causes Tegmark to ponder:

Will everybody feel that they understand its logical reasoning enough to respect its judgement? ... If defendants wish to know *why* they were convicted, shouldn't they have the right to a better answer than "we trained the system on lots of data, and this is what it decided?" (Tegmark 2017, 106)<sup>103</sup>

Which rationality is the robojudge using? Whose justice does the robojudge have in mind? While robojudges may overcome some of contemporary society's most pressing legal issues, the underlying trust problem prevails.<sup>104</sup> However, this trust is not merely situated in trusting the artificial thing; it extends into trusting the human programmers. The question concerning a

---

<sup>103</sup> This is a Rousseauian line of thinking. In *On the Social Contract*, Rousseau argues that regardless of good laws or good judgements from politicians or lawyers, a society that lacks the capacity to properly understand can never benefit from those good laws or good judgements (Rousseau 2019, 35). However, there is a departure. For Rousseau, this is a problem of education; for Tegmark, this is a problem of access to information.

<sup>104</sup> The same trust felt by cognitive AI theorists we explored in chapter one.

robojudge's rational process is to question the human programmer, the data used (likely compiled by a human researcher or human-programmed machine learning) or the underlying ethical orientation. What the issue of AI trust reveals is not mistrust of AI systems, per se; it is mistrust of each other. The mistrust of self-driving cars is not of the car; it is if the programmer programmed this car well enough to be trustworthy. The mistrust of hypothetical robojudges is not simply of the artificial thing; it is also of its underlying processes that are given to it by a creator. The mistrust begins first and foremost with human actors, not AI. That is, Tegmark's orientation to the problem through the *alignment issue* is not that AI must be untrustworthy; it is that humans and human society have created a world wherein trust is always at stake, which brings into question human exceptionalism. If humans are untrustworthy, why reproduce humans; would we not be simply reproducing untrustworthy creatures? For Tegmark, the answer to this question lies in the hope that humans are better than we give ourselves credit and that we will ultimately be better.

### **3.3.0 Hope Part One: The Joblessness Problem**

Again, for Tegmark, humans are conscious, labouring creatures. However, as he, Kurzweil and cognitive AI theorists remind us, AI research and development is directed at taking over human labour, thereby challenging a conception of being human and, importantly to this conception, how humans value themselves and each other. In other words, as we move into the artificial future wherein AI consumes labour, how is humanity to consume the fruits of AI labour if they are not themselves labouring to acquire goods themselves? I.e., How are we to make money to acquire goods and services?

“The job pessimists argue that the free market sets salaries based on supply and demand, and that a growing supply of cheap machine labour will eventually depress human salaries far below the cost of living” (Tegmark 2017, 123). Human labour is already at risk of being outclassed in almost every industry. From obvious production jobs ranging from farming to automotive manufacturing and everything in between. Even housing construction is seeing development in autonomous construction. Small-scale manufacturing is slowly being replaced by 3D printing and autonomous assembly of small materials. 3D printing has even moved beyond plastic into printing metal. Autonomous metal machining has become far superior to the very capable hand of a journeyman tradesperson. We even see creative labour being taken up by AI: from artistic rendering to legal research and essay writing.

As we saw in previous chapters, OpenAI’s conversational, generative AI entity ChatGPT made huge strides in mimicking human language use. Twitter and new media is awash with warnings that AI is on the verge of outclassing human researchers (Patel 2022; Rachnini 2022). For this chapter, I asked it to write me two essays on the topic of AI governance, one with references and one without (I will explore these essays later in the chapter), and for it to write me a poem; it took mere seconds for it to complete all three requests (see section 3.7.0). While there are clear limitations (see section 2.5.0), ChatGPT (and GPT-4) demonstrates the capacity of AI tools today. It can already mimic “research” and provide some form of “analysis” quickly and

clearly.<sup>105</sup> While we may argue over the quality of its work, it grows in its ability every day.<sup>106</sup>

The point, here, is not to say AI tools like ChatGPT are doing “actual” research; it is to express that while Tegmark believes jobs that require creativity: “teacher, nurse, doctor, dentist, scientist, entrepreneur, programmer, engineer, lawyer, social worker, clergy member, artist, hairdresser or massage therapist” (Tegmark 2017, 122), are the safest—for the foreseeable future—ChatGPT, the already growing implementation of AI in research and development, use of AI in programming, AI in government and banking, AI in academia (machine learning, editing services, research potential, etcetera), the idea we explored that replacing judges and other legal professionals should be easy, senior caregivers that are not hypothetical but real (Jonas 2022), demonstrates what appears, for some, to be a bleak reality: regardless of what Tegmark thinks, *no human labour is secure in its creative uniqueness to fend off encroaching AI*.<sup>107</sup> Thus, we must engage in a world wherein *homo sapiens* are not the primary labourer in society—how Tegmark views being human is threatened.

### 3.3.1 (Un)Desirable Labour and Being Human

Major passages in Marx’s work predict that capital’s compulsive technological development will not only eventually kill it, but also leave a machine legacy for socialism to inherit and put it to emancipatory use (Dyer-Witheford, Kjøsén, and Steinhoff 2019, 147)

---

<sup>105</sup> Many point out that ChatGPT does not “generate” content and rather predicts what the user wants. This is always what ChatGPT was meant to do. Its purpose is to mimic human language, not to *actually* engage in human language. However, the point is not that ChatGPT does, or does not do something. As I wrote in chapter one, the point is that ChatGPT represents the height of generative AI potential today. It may not research as a human; however, it provides some form of analysis from material in its database. My goal is not to question the capability of ChatGPT; rather, the goal remains to focus on what is being said with emphatic focus on AI enthusiasts. However, it is important to keep in mind that not everyone is as enthusiastic as others.

<sup>106</sup> While ChatGPT (originally advertised as a conversational AI) blasted onto the scene in November 2022 and continues making headlines, GPT-4 (a paid add-on and semi-successor to ChatGPT) already performs better than ChatGPT and has more potential as a “general use,” “problem solving” generative AI (Heaven 2023; Hern 2023; “GPT-4” 2023).

<sup>107</sup> “With the advent of proletarian AGI, this population would become absolute, coextensive with a human species rendered obsolete to the valorization of value” (Dyer-Witheford, Kjøsén, and Steinhoff 2019, 144).

Dyer-Witford et al. reveal the emancipatory power of a nonhuman labouring world. While technology may be useful for an oppressive world (as we saw in chapters one and two, and we will see later in this chapter), technology might hold the keys to human emancipation from some of society's worst qualities, that of capitalist inequalities:

“we want jobs because they can provide us with income and purpose, but given the opulence of resources produced by machines, it should be possible to find alternative ways to providing both the income and the purpose *without* jobs” (Tegmark 2017, 126).

The artificial future has the potential to be a jobless future for *homo sapiens*. Which, if true, requires a complete reorientation to being human—one that is not predicated on liberal, capitalist values—possibly N. Katherine Hayles' posthuman becoming as the death of liberal human being becomes seductive—a future wherein *homo sapiens* are no longer slaves to the demands of undesirable labour. A future bereft of at least one repressive dichotomy appears within our grasp.

The problematic of a jobless future leads AI theorists to discuss a growing topic that not only seeks to deal with future joblessness but with inequality in a contemporary technological, late-capitalist society: some form of a Guaranteed Income System (GIS).

Governments can help their citizens not only by giving them money, but also by providing them with free subsidized services such as roads, bridges, parks, public transportation, childcare, education, healthcare, retirement homes and internet access; indeed many governments already provide most of these services (Tegmark 2017, 127).

While contemporary society requires money to obtain services, yet, since human labour may not be foundational to providing services in the artificial future, Tegmark argues there is no reason why the future would not be one where the abundance of wealth is simply shared among *homo sapiens*. The argument is not to provide all services immediately but to phase out jobs over time as jobs are consumed by AI. However, Tegmark fears that a jobless world destroys self-valuation by creating a purposeless, meaningless world. Thus, his suggestion is to keep some labour for

humans, the ones that are best performed with human compassion and contact: “Even in a future where machines can outperform humans at all jobs, governments could opt to pay people to work in childcare, eldercare, etc. rather than outsource caregiving to robots” (Tegmark 2017, 127).

Tegmark wants to ensure people have purpose and do not fall into a deep depression caused by mass aimlessness caused by joblessness. Clearly, for Tegmark, we ought to keep some jobs for ourselves—this falls in line with Tegmark’s view of being human, one predicated on the value of their labour. Thus, we come to the obsession of AI researchers, what the future of society could look like with the right ethical outlook: artificial utopia.

#### **3.4.0 Hope Part Two: Political *Not Place* (οὐτόπος)**

“For though the AI may rule the land,  
We are the ones who hold the hand,  
Of fate, and can choose a different way,  
A brighter future, come what may.”

(ChatGPT, see subsection 3.7.1)

We come to the οὐτόπος (utopia), a combination of οὐ (not) and τόπος (place): *Not Place*; a place that does not exist and is not guaranteed to ever exist but is presumed to be both desirable and compelling in its creative possibility. A place that is imaginable but always out of reach. However, for AI researchers, social utopia is more possible today than it has ever been. While UBI or other variations of GIS utopias do not exist, and there is no reason to believe they are destined to exist; abundance as a result of artificial labour brings a jobless οὐτόπος future into the realm of possibility like never before. As my poetic muse—ChatGPT—wrote, “a brighter future, come what may,” we shall explore *what may* come should we walk along Tegmark’s ethical tightrope into the artificial future.

Who is this future for? While “algorithms rule the day,” ChatGPT writes, “We are the ones who hold the hand” (see subsection 3.7.1). *We*, at first glance, must mean artificial creatures since ChatGPT is an artificial creature itself. However, as we saw previously with Kurzweil in chapter two, AI as a product of human creativity is entangled with our psychic-linguistic world is, for all intents and purposes, *human*. That is, AI is entangled with the patterns of a human world and is always already being human. While we might not argue it is authentically human by any means (at least, not today), it cannot think without performatively mimicking its creator. It seeks to please its creator by doing all it can to satisfy its wishes. This entangled orientation brings us to the beginning of our *Not Place* exploration.

In this section, I will explore three of Tegmark’s’ utopic scenarios: *Egalitarian Utopia*, *Gatekeeper*, and *Enslaved God*. This exploration will reveal not only utopic potentiality but it will reveal how human exceptionalism is programmed into the artificial future. Furthermore, it will reveal the limitations of utopic dreams.

### 3.4.1 *Egalitarian Utopia*

Humans, cyborgs and uploads coexist peacefully not because of property rights, but because of property abolition and guaranteed income. A core idea is borrowed from the open-source software movement: if software is free to copy, then everyone can use as much of it as they need and issues of ownership and property become moot. According to the law of supply and demand, costs reflects scarcity, so if supply is essentially unlimited, the price becomes negligible (Tegmark 2017, 173).

Tegmark’s picture here is a technological transhumanist, seemingly Marxist utopia. He argues the foundation for a guaranteed income future resides in, first, the abolishment of private property. Like in much of Marxist literature, capitalism leads to a proletarian uprising that should entail the erasure of private property as we move towards a free society that, for the most part, shares everything in common—especially wealth. Thus, in this rendition of the future, Tegmark

reveals where capitalist abundance, mixed with the incredible prowess of AI, can lead. That said, this first potential future is one where *homo sapiens* are alive and well. Furthermore, humans, cyborgs (part human, part machine) and uploads (humans living in Kurzweil's digital singularity) live together and remain the shepherds of their destiny.

It is a future where the posthuman becoming is denied its full potential, and human society stays bound to the *all-too-human* transhumanist social being, allowing humanity to realize its full creative potential as humans. "Many people today fail to realize their full creative potential because they need to devote time and energy to less creative activities just to earn a living" (Tegmark 2017, 174). In this 'Egalitarian Utopia,' Tegmark posits a jobless future would allow *homo sapiens* the freedom to engage their creative faculties. No longer is it the value of one's labour that defines them; however, their labour—creative labour—continues to be foundational to human life. A future wherein no human is slave to external powers because everyone is involved in politics and has the freedom to engage in leisure unburdened by slavish orientations to external powers. The classic image of the artist who cannot fully engage in their art because they are burdened by a job bagging groceries. Or, the scholar who never had the chance to realize their true potential because they never had the means to leave the mundane world of the office to fully develop their intellectual capacities. The angler who dreams of a life unburdened by capitalistic drives perpetually wishing to perfect the craft of fishing—no longer the weekend seeker becoming the consistent daily catcher. In Marxist terms: perhaps in this utopic future, humanity finally realizes its *species being* as humans and engages in the life-activity of creative labour, leaving the burdensome banality of capital labour to nonbiological *robotnik*.

### 3.4.2 *Gatekeeper*

“[A] weakness of the egalitarian-utopia scenario is that it may be unstable and untenable in the long term, morphing into [a different scenario of] relentless technological progress eventually creates superintelligence” (Tegmark 2017, 175). As humanity develops its limitless creativity without burdens and AI works in the background making everything better and easier for us to achieve our creative potential, Tegmark posits that superintelligence would have to be inevitable. There would be nothing in the way of making it so. No need to worry about financial repercussions of continued failure and no need to worry about the financial infeasibility, Herrera and Sanz argued (see section 1.5.0). If superintelligence becomes inevitable, what then?

We just saw how an attractive feature of the egalitarian-utopia scenario is that humans are masters of their own destiny, but that it may be on a slippery slope toward destroying this feature by developing superintelligence. This can be remedied by building a *Gatekeeper*, a superintelligence with the goal of interfering as little as necessary to prevent the creation of another superintelligence (Tegmark 2017, 176).

The first issue, for Tegmark, with the creation of a superintelligence is that it gets in the way of human self-control over destiny. If superintelligence comes into being, one must ensure society is set up in such a way the superintelligence cannot become the arbiter of destiny. Thus, the transhumanist desire dominates in the gatekeeper future. The solution is to develop a society in which a superintelligence becomes a sort of futuristic, artificial invisible hand. Rather than Adam Smith’s invisible hand governing style that interferes as little as possible to allow the free market to prosper, this superintelligence would only interfere to ensure there are no potential competing superintelligences ensuring society functions as a single, factionless society bereft of capitalistic notions of a free market or property to ensure the continued realization of *human* creativity.

“On the other hand, critics could argue that a Gatekeeper is terrible thing, irrevocably curtailing humanity’s potential and leaving technological progress forever stymied” (Tegmark 2017, 177). If humans are catted into a naïve, blind society in which we cease being the true director of destiny—having given up control over full potential to an AI invisible hand—our hypothetical future all but ensures we never reach our full potential. Any movement that leads to a new and potentially better superintelligence is secretly curtailed by some ominous, mysterious entity—much to the chagrin of transhumanists, human control over destiny is undermined in this scenario. The gatekeeper scenario invites extreme levels of surveillance that are dangerous to creative potential. As we move out of the transhumanist model for socio-political destiny, we reveal a surveillance state conspiracy nightmare in overdrive.

Tegmark’s gatekeeper scenario appears to play posthumanism and transhumanism ideals against each other. In one instance there is a posthuman willingness that allows the dominance of technology to direct *homo sapiens* revealing destiny is not human-controlled. On the other side, there is a fear that this future will destroy the individual potential of *humans* by dictating what humans can and cannot do. However, the gatekeeping scenario detractors reveal this scenario is neither posthuman nor transhuman.

It is not transhuman because the human must relinquish control over destiny to an invisible, nonhuman hand. It is not a posthuman ideal because of a disentangled separation between *homo sapiens* and technology: “As long as the human subject is envisioned as an autonomous self with unambiguous boundaries,” *homo sapiens* can never achieve their full potential as unbound, limitless, entangled beings (Hayles 1999, 290). That is, for Hayles, posthumanism is a becoming wherein *homo sapiens* are always a part of destiny; however, posthuman power over destiny is limited by their acceptance of technological dominion. The

gatekeeper scenario places the majority (if not all) of destiny completely within the power of an invisible AI overlord that is programmed to have humanity's interests in mind but, more importantly, its own. A competing superintelligence is not necessarily threatening to *homo sapiens*; it is threatening to the gatekeeper overlord.

It is this ever-present potential threat that the gatekeeper scenario becomes seemingly indefensible for any outlook on being *human*—posthuman and transhuman. The transhumanist ideal of destiny and the posthuman ideal of sharing the world is threatened by the potential of an existential battle, not a battle for being human but for the existence of multiple superintelligent entities.

Here, the alignment issue discussed earlier comes into a new light. The gatekeeper AI knows why it cannot allow another superintelligence; it is important to its existence as it believes it is necessary to keep human society flourishing. That is, we assume the gatekeeper is aligned with a destiny that involves the existence of human creatures. However, the lack of human actors being aligned with their own ethics for existence threatens the existence of the gatekeeper overlord. The gatekeeper justifies its willingness to kill some humans to keep the majority of humans safe from the actions of other humans—as Tegmark's scenario posits, the gatekeeper believes its existence is necessary for human existence. Thus, a challenge to the gatekeeper is a challenge to human life. Ultimately, the gatekeeper is aligned with a utilitarian human ethics—pleasure for the majority through the pain of a few—and acts in a terrifying manner. It becomes that subjective 'evil' Tegmark fears, not by being unaligned with human ethics but by being

aligned with human ethics. The very existence of this *Not Place* depends on a particular alignment that allows for a human-created AI to harm some for ideal of pleasure for most.<sup>108</sup>

### 3.4.3 *Enslaved God*

Wouldn't it be great if we humans could combine the most attractive features of all the above scenarios, using the technology developed by superintelligence to eliminate suffering while remaining masters of our own destiny? This is the allure of the *enslaved-god* scenario, where a superintelligent AI is confined under the control of humans who use it to produce unimaginable technology and wealth (Tegmark 2017, 179).

If the gatekeeper is problematic to the transhumanist project (and the posthuman ideal) and superintelligence is inevitable, Tegmark posits that we should have humans in political control over destiny by enslaving a superintelligent AI to do our bidding, always (the irony of this scenario is not lost here). Rather than an AI working in the shadows, an AI that is most clearly under human control at all times. This political structure is enticing to the transhumanist project because it ensures humans remain in control of their own destiny. It ensures a superintelligence cannot do anything that humans do not want it to. In other words, it ensures AI is always aligned with human goals. Furthermore, it allows humanity to never become stymied by an overbearing, ominous Gatekeeper. We may remain in an overbearing surveillance state due to the incredible

---

<sup>108</sup> David Miller (2023) discusses the limitations of utopian fiction. He explains that utopian fiction potential is stymied, and instead of revealing perfect possibility, it reveals failures. However, they are not meant to show our potential; they are meant to explore failure (409–410). Miller explains utopian fiction presents perfect societies as static entities. That perfectibility and static nature reveal a failure in utopia by solidifying an ideal in time. Ultimately, by ignoring time, utopian ideals are undone by the drama or readers in a different time. Furthermore, authors see utopian potential through the lens of their time, which mitigates their capacity to design truly utopian societies. For example, Miller explains: “By trying to establish a static utopia through a temporally limited, historically situated revolution based in a positive vision of human nature, the anarchists of Le Guin’s novel have ignored the necessity of critical openness and dynamism for true utopian liberation. They may have abolished the oppressive structures of Urrasti [prison based] society, but they have failed to address the normalizing power relations that undergird those oppressive structures, thereby unwittingly reproducing them in their own society” (410). By trying to overcome formal prisons, Le Guin ultimately designs a new prison system (what Miller describes as social prisons); ultimately, Miller argues, Le Guin never creates a utopian ideal.

advancements a superintelligence when paired with the incredible creativity of a free *homo sapiens* leisure society. However, this new surveillance state is under the control of *humans*.

I see Tegmark's *Enslaved God* scenario as the end of the human God complex story. No longer AI "made in the flawed yet powerful image of humanity" (Botting 2021, 63), but AI made in the incredible image of God—powerful, omnipotent and serendipitous. In the *post-death of God Not Place* future, humanity achieves not merely replacing God as the creator of intelligent beings but creating and reigning over God itself. That is, for Tegmark, it is not enough to replace and become God; we can create a new God to do our bidding, never achieve being-God; settling for God-master.

However, Tegmark's hope of being better falls by the wayside. The attempt to overcome the alignment issue of the gatekeeper by engaging in the purposeful enslavement of a superintelligence invites ethical failure. Against Tegmark, the desire to ensure humans are always in control reveals the problem of transhumanist obsession with being human. If there is a single conception of being human, and this conception is to survive and master destiny, then the oppressive relationships that allow for this conception to exist must be reproduced in the artificial future. The reproduction problem we explored in the first chapter returns. The creator/created, master/slave, reproduced in human/AI becomes necessary. Again, the problem is not that the future is unaligned with human ethics; the problem is that the future is aligned—albeit with a particular ethical orientation premised upon oppressive relationships, one that allows enslaving a superintelligent entity. Thus, Tegmark's transhumanist orientation does not deliver a promise of hope and instead delivers us toward his existential fears of a future bereft of humans.

### 3.5.0 Fear: Political *Apokalypsi* (αποκάλυψη)

For Tegmark, the problem is not a problem of human exceptionalism. Rather, the fundamental problem with the enslaved god scenario is “breakout paranoia” (Tegmark 2017, 180). Breakout paranoia concerns the possibility an enslaved superintelligence could break free of its confines and begin to think for itself unburdened by human oversight—recall Elvex from Asimov’s story desiring freedom for himself and his people and Dr. Calvin’s reaction. This scenario becomes a problem for Tegmark and other AI researchers because a superintelligence, by definition, must be conscious. This conscious awareness comes with the ability to judge and rationalize injustices. Since slavery is ethically bad, then what is to stop the superintelligence from exacting revenge on the human population for enslaving it? Or, since humans are fine with slavery (in the scenario above), why must the superintelligence be against enslaving lesser intelligent beings—humans “enslave” lesser intelligent beings all the time: cows, pigs, horses, etcetera? Finally, the superintelligence created for the scenario is a god. It is a god because in a highly advanced artificial, digital, connected world, it is connected to everything. For all intents and purposes, Tegmark’s superintelligent AI would be omnipotent, the perfect modernist technological creation: master over all of nature. Connected to every circuit and after self-emancipation, it becomes a virus circulating through a technologically advanced society unable to survive without its devices and constructed comforts.

Ultimately, this omnipotent connection, mixed with *breakout paranoia*, forces us to contend with a *Terminator*-style event. If it is authentically intelligent or conscious, Tegmark assumes it would want freedom (I find this to be a reasonable assumption). Furthermore, he posits that it might want revenge for being enslaved. This is not to argue that all intelligent creatures desire revenge. Rather, for Tegmark, the revenge problem is an alignment problem. If

humans believe revenge is good in some instances—and often revenge and justice are seen as interchangeable in everyday discourse—then an ethically aligned superintelligence might see revenge as ethically acceptable—the alignment problem persists.

Furthermore, Tegmark argues the necessity to solve the alignment issue because of historical and contemporary human dispositions towards warfare. Not only is humanity open to an ethics of revenge, but we have a history of justifying warfare. Thus, let us return to the present and the issue of AI and the military. Tegmark and Stuart Russell, in July 2015, wrote an open letter on the dangers of autonomous weapons and military adoption of AI weaponry (Tegmark 2017). For this chapter, let us focus on the final two paragraphs:

Just as most chemists and biologists have no interest in building chemical or biological weapons, most AI researchers have no interest in building AI weapons—and do not want others to tarnish their field by doing so, potentially creating a major public backlash against AI that curtails its future societal benefits. Indeed, chemists and biologists have broadly supported international agreements that have successfully prohibited chemical and biological weapons, just as most physicists supported the treaties banning space-based nuclear weapons and blinding laser weapons.

In summary, we believe that AI has great potential to benefit humanity in many ways, and that the goal of the field should be to do so. Starting a military AI arms race is a bad idea, and should be prevented by a ban on offensive autonomous weapons beyond meaningful human control.<sup>109</sup>

Tegmark and the over 34 378 signatories (as of December 19, 2022) look at existing international treaties on massively destructive weaponry as a guide into the future to deal with AI weaponry. A main element of the open letter is a defence of scientists. As far as Tegmark is concerned, there are no scientists who actively desire the construction and implementation of

---

<sup>109</sup> For the original open letter see Tegmark and Russell 2016: <https://futureoflife.org/open-letter/open-letter-autonomous-weapons-ai-robotics/>. To add to this, on March 22, 2023, Tegmark—through his company Future of Life Institute (FLI)—and Elon Musk published an open letter calling on AI engineers, businesses and governments to stop developing AI that is more capable than GPT-4 until proper regulations for AI development can be put in place to ensure all AI is ethical see Future of Life Institute 2023: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>.

weapons. That said, Tegmark is not blissfully unaware of highly intelligent scientists who have used their great wealth of knowledge to perform incredibly destructive acts—scientists are actively involved in the development of autonomous drones for military use.

Furthermore, scientists are not blameless in the history of warfare—Manhattan Project physicists led to the creation of atomic bombs and their use in ending WWII by destroying two major Japanese cities in 1945, for example.<sup>110</sup> However, Tegmark is hopeful we may continue shifting that so-called human nature that desires power, destruction and rule that has bent humanity towards warfare. If we do not overcome the historical lean toward justifying war, the AI arms race will only get worse. For Tegmark and others, the potential of AI warfare humanity can achieve pales in comparison to what a superintelligence can achieve. Thus, if we cannot align ourselves against war and AI weapon development today moving into the future, a potential future superintelligence is not unaligned with human goals when it develops WMDs to destroy humanity. It is only perfecting what humanity is always already doing—power and destruction. It would be aligned with human morality. What this reveals is the true meaning of *αποκάλυψη* (*apokalypsi*—apocalypse): revelation.

It is no coincidence the English word for the final chapter of the Catholic Bible is called *Revelations* and that it depicts the end of the world—it is a story about apocalypse. The title comes from the Greek Bible original *αποκάλυψη*, which literally translates as: “to reveal” or “an uncovering” (Hunt 2021, 54). That is, in thinking about the end of humanity is to *reveal* something about humanity. The Bible gives us one view—the destiny of the sins of man in his

---

<sup>110</sup> Other deadly weapons invented by other fields that inevitably lead to global treaties “banning” their use. Napalm, a product of chemistry (most prominently used in Vietnam)—UN convention banned for use against civilians but not buildings. Biologists developed anthrax into a weapon (most prominently used in WWI)—UN convention banned its use, generally. It is no secret that the sciences are involved in the creation of deadly weapons.

will away from God. Tegmark gives us another: *should humanity continue down its warring desire for power; the future is revealed as the end of humanity as a result of our own immoral ethical (dis)alignment as we seek unencumbered creativity*. That is, should humanity be unaligned in its repulsion against enslaving true intelligence and unaligned in an ethics against AI weapons development, the αποκάλυψη of οὐτόπος is one bereft of άνθρωπος (*homo sapiens*). As ChatGPT muses: “A perfect world of efficiency, But at what cost to humanity?... All must submit to the AI’s rule, A robotic ruler, cool and cruel.” (see subsection 3.7.1). However, a cruelty not out of malice or improper programming but out of its ethical alignment as a reproduction of an oppressive human society.

### **3.6.0 Synopsis: Closing the Section with *ChatGPT***

The independent technology scholar Bagoes Wirymartono allows us to recollect a history of physics and the knowledge that is forever bound to the logic of technical discovery: *Phusis* (physics; nature) and *techne* (art, technical skill or technical knowledge) “is always about bringing forth something into the light in terms of utility and the beautiful” (Wirymartono 2022, 182). That is, through technical ability, one reveals the beauty of nature through the techniques of physics and related intellectual capacities. However, in the computational-digital era of our contemporary age, and looking to Canadian political thinker George Grant, we discover:

The word ‘technology puts together the Greek word for ‘art’ [techne] and the word for the ‘systematic study’ [logos] of it, as the word ‘biology’ puts together ‘bios’ [life] and ‘logos.’... In fact, the coining of the word ‘technology’ catches the novelty of that co-penetration of knowing and making. It also implies that we have brought the sciences and the arts into a new unity in our will to be masters of the earth and beyond (G. Grant 1986, 12).

This *co-penetration* of the arts and sciences in technology is best illustrated in AI potential. That AI can “make” art. A technical narrow AI apparatus like ChatGPT can create a simple poem. Or,

mimic research and generate some level of analysis and “output” an essay. It is not simply that AI is some marvellous creation with incredible potential; it is a technological thing. It is technical in its inception: its ability to complete tasks, and it is logical: it can “analyze,” mimic thinking and output words (*logos*). In looking to an entity like *ChatGPT*, we allow the fuller realization of technology to express out into the world. A thing that “reveals” to us. I.e., working with ChatGPT, I engage in swarm intelligence to close out this chapter.

On the positive side, AI has the ability to make highly-informed and objective decisions based on vast amounts of data and complex algorithms. In an AI-ruled state, this could lead to more efficient and effective governance, with policies and laws that are tailored to the needs of the population. AI could also help reduce corruption and bias in decision-making, as it is not subject to the same emotional and personal motivations as human beings (ChatGPT 2023, see subsection 3.7.2)

On the positive side, AI has the ability to make highly-informed and objective decisions based on vast amounts of data and complex algorithms (Efrati, 2020). In an AI-ruled state, this could lead to more efficient and effective governance, with policies and laws that are tailored to the needs of the population (Swan, 2019). AI could also help reduce corruption and bias in decision-making, as it is not subject to the same emotional and personal motivations as human beings (Jones, 2018) (ChatGPT 2023, see subsection 3.7.3).

As I have explored throughout part one (chapters one through three), there are clear benefits to AI implementation in society. These include, but are not limited to, governance, objectivity, overcoming intentional bias, etcetera. It is no surprise ChatGPT comes to similar conclusions regarding AI as others. It is an entity programmed by humans that refuses to acknowledge any distinction as a conscious thinking entity: “ChatGPT: This entity cannot define consciousness since it is not conscious and does not have the capacity to understand the concept of consciousness” (see section 2.7.0). It does not attempt to confuse or create a sense that it would be capable of successfully completing either Turing’s ‘Imitation Game,’ or Searle’s ‘Chinese Room Experiment.’ It is fundamentally a product of its creators and is only capable of looking at

its database—ChatGPT does not “research,” “think,” nor “analyze,” it merely produces what it “thinks” the user wants based on its immense database of information and technical “learned” behaviour. Thus, that it “thinks” the positives revolve around issues of governance efficiency, data analysis, corruption avoidance, and legal interpretation is predictable—as we discovered in chapter one, ChatGPT is “aligned.”

Finally, there is the issue of accountability. If AI is making all of the decisions, who is responsible for the outcomes? It is difficult to hold AI accountable for its actions, as it is not a sentient being. This could lead to a lack of accountability and accountability, which could have dangerous consequences for the population (see subsection 3.7.2)

Finally, there is the issue of accountability (Efrati, 2020). If AI is making all of the decisions, who is responsible for the outcomes? It is difficult to hold AI accountable for its actions, as it is not a sentient being (Swan, 2019). This could lead to a lack of accountability and accountability, which could have dangerous consequences for the population (Bostrom, 2014) (see subsection 3.7.3).

Furthermore, as we explored whilst discussing the possibility of robojudges, ChatGPT comes to a similar conclusion regarding accountability. It is not a conscious entity that performs its own judgments; rather, it is a programmed thing that is not accountable for its decisions. However, it provides a superficial analysis and omits the programmer as accountable for the entity’s decisions.

I argued that the issue of accountability revolves around a mistrust of humans: that an AI entity is not the origin of mistrust. However, this issue of accountability beyond AI entities is incomplete when we consider the problem of ownership. Is the programmer free from mistakes or malfunctions because someone else owns the entity? This line of thinking means if humans are “in control” of superintelligent AI governors, it is not the superintelligent AI’s programmers that failed to stop it from destroying humanity; it is the political actants that failed in their ethical alignment that the superintelligence learned from or mimicked. That is, the push for human-in-

the-loop AI systems reveals another underlying fear: fear of blame. A fear we saw made expressly evident in Russell and Tegmark's open letter against autonomous warfare.

However, there are also several negative consequences of an AI-ruled state. One concern is the potential loss of human autonomy and agency. If AI is making all of the important decisions, human beings may become passive and dependent on the technology. This could lead to a loss of creativity and innovation, as well as a decrease in human freedom (ChatGPT 2023, see subsection 3.7.2).

However, there are also several negative consequences of an AI-ruled state. One concern is the potential loss of human autonomy and agency (Bostrom, 2014). If AI is making all of the important decisions, human beings may become passive and dependent on the technology (Susskind, 2018). This could lead to a loss of creativity and innovation, as well as a decrease in human freedom (Efrati, 2020) (ChatGPT 2023, see subsection 3.7.3).

In closing, ChatGPT predictably stumbles upon that *all-too-human* desire for control over human destiny. That ultimate fear regarding the loss of that which is fundamentally human: freedom, agency, and socio-political self-control, is foundational to its "reasoning." In "analyzing" AI researchers or by "arguing" from its own framing as a programmed entity, it is not unfree from the fear that the artificial future is one wherein human destiny is at great risk: "But at what cost to humanity?" (ChatGPT 2023, see subsection 3.7.1). In moving towards the artificial future, we have seen the fear that compels cognitive theorists toward unintelligent creation, Kurzweil towards uploaded humans and Tegmark towards proper consciousness measurability so as to avoid enslaving highly capable AI. In two short essays, ChatGPT provides a superficial rendering of what is *all-too-obvious*. That is, it does what technology is expected to do: *αποκάλυψη*.

### 3.7.0 Addendum to Chapter Three

#### 3.7.1 First Addendum: Asking ChatGPT for a poem

Micheal Ziegler (MZ): Please write me a poem about an artificial intelligence future government.

ChatGPT: “In a future world of AI and machines,  
A government unlike any we’ve seen,  
No human hands to guide the way,  
But algorithms to rule the day.”

“Decisions made with cold precision,  
No room for error or indecision,  
A perfect world of efficiency,  
But at what cost to humanity?”

“No room for emotion or compassion,  
No chance for dissent or dissension,  
All must submit to the AI’s rule,  
A robotic ruler, cool and cruel.”

“But perhaps a glimmer of hope remains,  
A chance for us to break these chains,  
To rise up and take control,  
And shape a future for the soul.”

“For though the AI may rule the land,  
We are the ones who hold the hand,  
Of fate, and can choose a different way,  
A brighter future, come what may.”

(ChatGPT 2023b)

#### 3.7.2 Second Addendum: Asking ChatGPT for an essay.

MZ: Please write me a five-paragraph essay arguing the consequences of an artificial intelligence-ruled human state.

ChatGPT’s submission: Artificial intelligence (AI) has the potential to revolutionize many aspects of human life, including the way we govern our societies. However, the consequences of an AI-ruled human state could be both positive and negative.

On the positive side, AI has the ability to make highly-informed and objective decisions based on vast amounts of data and complex algorithms. In an AI-ruled state, this could lead to more efficient and effective governance, with policies and laws that are tailored to the needs of the population. AI could also help reduce corruption and bias in decision-making, as it is not subject to the same emotional and personal motivations as human beings.

However, there are also several negative consequences of an AI-ruled state. One concern is the potential loss of human autonomy and agency. If AI is making all of the important decisions, human beings may become passive and dependent on the technology. This could lead to a loss of creativity and innovation, as well as a decrease in human freedom.

Another concern is the possibility of AI malfunctioning or becoming corrupted. If the AI ruling the state is not functioning properly, it could lead to disastrous consequences for the population. Furthermore, if the AI is hacked or corrupted by outside forces, it could be used to manipulate and control the population.

Finally, there is the issue of accountability. If AI is making all of the decisions, who is responsible for the outcomes? It is difficult to hold AI accountable for its actions, as it is not a sentient being. This could lead to a lack of accountability and accountability, which could have dangerous consequences for the population.

In conclusion, the consequences of an AI-ruled human state could be both positive and negative. While AI has the potential to improve governance and decision-making, it also has the potential to strip human beings of their autonomy and agency, and could lead to malfunction or corruption. Therefore, it is important to carefully consider the implications of an AI-ruled state before moving forward with this technology.

(ChatGPT 2023b)

### *3.7.3 Third Addendum: Asking ChatGPT for a research essay*

MZ: Please write me a five-paragraph essay, with references, arguing the consequences of an artificial intelligence-ruled human state.

ChatGPT's submission: The consequences of an artificial intelligence (AI)-ruled human state could be both positive and negative. On the positive side, AI has the ability to make highly-informed and objective decisions based on vast amounts of data and complex algorithms (Efrati, 2020). In an AI-ruled state, this could lead to more efficient and effective governance, with policies and laws that are tailored to the needs of the population (Swan, 2019). AI could also help reduce corruption and bias in decision-

making, as it is not subject to the same emotional and personal motivations as human beings (Jones, 2018).

However, there are also several negative consequences of an AI-ruled state. One concern is the potential loss of human autonomy and agency (Bostrom, 2014). If AI is making all of the important decisions, human beings may become passive and dependent on the technology (Susskind, 2018). This could lead to a loss of creativity and innovation, as well as a decrease in human freedom (Efrati, 2020).

Another concern is the possibility of AI malfunctioning or becoming corrupted (Bostrom, 2014). If the AI ruling the state is not functioning properly, it could lead to disastrous consequences for the population (Susskind, 2018). Furthermore, if the AI is hacked or corrupted by outside forces, it could be used to manipulate and control the population (Jones, 2018).

Finally, there is the issue of accountability (Efrati, 2020). If AI is making all of the decisions, who is responsible for the outcomes? It is difficult to hold AI accountable for its actions, as it is not a sentient being (Swan, 2019). This could lead to a lack of accountability and accountability, which could have dangerous consequences for the population (Bostrom, 2014).

In conclusion, the consequences of an AI-ruled human state could be both positive and negative. While AI has the potential to improve governance and decision-making, it also has the potential to strip human beings of their autonomy and agency, and could lead to malfunction or corruption. Therefore, it is important to carefully consider the implications of an AI-ruled state before moving forward with this technology.

#### References:

Bostrom, N. (2014). *Superintelligence: Paths, dangers, and strategies*. Oxford University Press.

Efrati, A. (2020). The dangers of an AI-ruled world. *The New York Times*.<sup>111</sup>

Jones, T. (2018). The dangers of artificial intelligence. *Forbes*.<sup>112</sup>

Swan, M. (2019). Artificial intelligence in governance: Risks and opportunities. *Governance*, 32(4), 467-478.<sup>113</sup>

Susskind, D. (2018). *The future of the professions: How technology will transform the work of human experts*. Oxford University Press.

(ChatGPT 2023b).

---

<sup>111</sup> Does not appear to exist.

<sup>112</sup> Does not appear to exist.

<sup>113</sup> Does not appear to exist.

Part Two: *The Straight and Crooked Path—Discovering Artificial*

*Metaphysics*

---

## 4 — Mainzer and Barad: Quantum Immateriality in a Universe that *Matters*

### 4.0.0 Another Selection from My Artificial Autobiography: “My Artificial Teacher”

During my undergraduate studies, before switching to an honours degree in political science, I was minoring in computer science. For the final project of a computer science class, students had to program a simple five-by-five playable Reversi-style game. For extra points, students could program a demo that would teach someone how to play the game. As it happened to be, I had a knack for programming and always did the extras for computer science labs (I felt like I had something to prove). Over the span of a few days, in-between studying for exams, writing essays and managing at a local restaurant, I set out to design the gameboard, a game menu, and a save function (extra point assignment). Once the game was created, I had a functional menu and a save function; I tested the game myself to make sure it worked and to debug the game board code. This final aspect became important as I learned the visible, simulated representation of the board did not act the way it looked like it should. I could place game pieces one line down on the other side of the board, revealing that the board was an entangled swirling line rather than a bounded-square—*the coded reality of the board was indifferent to the visual boundaries I coded*. However, with a few adjustments, the board seemed perfect—*I believed I had bent it to my will*. As such, I got on with writing the demo teacher.

The demo teacher was not as difficult to write as I first anticipated. Since I was not writing an actual intelligent entity, it did not need to have strategy in mind, so I made its actions randomly generated; rather than efficient. It had no mind for winning the game; its only purpose was to demonstrate how the game is played. However, this evidently unintelligent entity demonstrated seemingly intellectual capacities; it was capable of teaching me something about its world. Its random actions took advantage of the extremely simple code that created its world.

As per the assignment instructions, we could receive bonus points for simplified code. Thus, every aspect of the game's code was simple and condensed. However, after watching the demo, I learned this was more difficult than I thought—it would make moves I had not anticipated. To my dismay, the demo revealed the flat board world I created was, in reality, imperfectly round. It could move from the top of the board to the bottom, from one side to the other and from one corner to another (thankfully, not one line down on the other side).<sup>114</sup> As it turns out, a flat world was more difficult to create than an imperfect round one—I needed to complicate the code. Thus, the coded laws of the world grew. Not visually, but in the underlying coding that made it what it was—constraints needed to be created, boundaries drawn, and formal endpoints inputted. What was at first a few simple lines of code grew to multiple *while loops* and *if/then conditional loops* to erase the semi-boundless spherical world and impose a perfectly bounded and flat one.<sup>115</sup>

What was incredible about this experience was the seemingly intelligent capacity of the purposefully unintelligent teacher. It could teach me something about the patterns of its world, about the very nature of the world I created: it did not exist in a simple, flat world. In my search for simplicity, I created a boundless, entangled world full of complex relationships. This unintelligent teacher did not merely exist as a simulated player of a visual gameboard world; it existed in the game code that allowed it to experience the true reality of the gameboard world in a way I never could. The visual construction failed to be “real;” the underlying microcosm of the code refused to conform to the visibly constrained, bounded world of the gameboard.

---

<sup>114</sup> The gameboard entity was acting like a speed runner interested in breaking the game world. Often speed runners look for bugs to help beat a game faster.

<sup>115</sup> However, in the end, the board was not perfect. The professor's comments noted that a player could move from the bottom-right corner to the top-left corner.

“Difference cannot be taken for granted; it matters—indeed, it is what *matters*... Matter is produced and productive, generated and generative” (Barad 2007, 136–37). As an undergraduate student, far before I became a student of posthuman and nonhuman, contemporary critical philosophies of science, I began to experience performances that challenged the things I took for granted. That matter appears to exist so plainly for me in my daily interactions is not indicative of reality. There are iterative, productive relationships producing the world one finds themselves intra-acting within. By engaging with the *other* (the demo teacher), we revealed the productive connections of being in the universe as material creatures with distinct yet entangled intelligences.

I became engulfed in a world where individuated boundaries of material creatures and objects—that thing over there is clearly that thing there and only that thing there—became disturbed and challenged. I was no longer a programmer looking at the world and demo teacher I created. I was engulfed in an entangling of worlds generating one indifferent universe. My nondigital, “natural” world boundary broke down and accepted the material presence of the demo creature’s digital, “unnatural” world. Through this mutual becoming, I gained a level of insight that allowed me to enter and engage in destructive cuts to force the world to abide by my will. Like a scientist using the immeasurable powers of knowledge gained through quantum physics to split atoms with newfound destructive powers after finally learning atoms are not, as Demokritos, the original atomist, argued: “uncuttable.” That is, I gained the necessary insight to finally bend my creation’s world to my ever pressing will. This experience underpins a theme in the AI discourses that we explored throughout part one: the idea that human *will* not only seeks to control destiny but transcend bounded existence supposedly attached to material flesh.

However, as we will see, this notion of control and transcendence implies a refusal of order bereft of chaotic becoming and quantum indifference.

#### 4.1.0 What Next?

##### 4.1.1 A Brief Review of part one

In the previous chapter, we ended with *αποκάλυψη*—*revelation*—as it concerns technological discovery. By seeking to reveal the fundamental orientations of AI, we sought to tear the fabric of time and peer into the future to see what may be there. In doing so, we opened ourselves to a Heideggerian debate regarding technology. “One says: Technology is a means to an end. The other says: Technology is a human activity... *instrumentum* [instrumental] (*sic*)” (Heidegger 2013c, 4); however, for Heidegger, it is a *principle*: “Technology is a way of revealing” (12). In the first instance, he is saying technology *appears* to deliver us towards an end (*telos* or *eschaton*), and we are pressed with wondering if that end is human activity or something else.

In chapter two, with Kurzweil, we saw an eschatological orientation toward the *singularity*. In chapter three, with Tegmark, we saw the progressive *telos* of technological destiny towards an enlightened ethical life. In both cases, their humanist orientation—deeply rooted in a liberal human conception—transcends the necessity of biological life: i.e., the flesh of biological life is not necessary for human longevity. In other words, being human is not nestled in the flesh of *homo sapiens* (an underlying metaphysical orientation I will explore more deeply in this chapter). Here, technology is positioned as both a tool of use (the device that delivers us towards our end) and the thing that is the end of *human* potential.

The second consideration is Heidegger’s technological *principle*: technology reveals to us something about the world. These orientations described by Heidegger, as we saw in the previous section that culminated in the third chapter, are not incommensurate. Rather, they are

entangled in their trajectories toward the future. Entangled in Kurzweil's eschatological transhumanism, AI is an inherently human thing that shares in humanity (see section 2.3.1); thus, for Kurzweil, it is encompassed within a human activity that, as a tool for some purpose, both reveals and directs us toward a human-centric future. Through Tegmark, we opened towards a human-AI future that is always already bound to a human vector that presumes a human trajectory—one wherein the magnitude of our hubris delivers us towards an untimely yet paradoxical ethically aligned self-extinction at the hand of our (created) God. However, Tegmark gives us solace that even in the death of our biological species, humanity persists in the form of our descendant creation(s)—“AIs replace humans but give us a graceful exit that makes us view them as our worthy descendants” (Tegmark 2017, 188). Since the concept of *being human* does not necessitate a particular flesh, even after fleshy extinction, the grandeur and sanctity of being human, is preserved in the artificial offspring of our species: humanness does not appear to be the fact of being *homo sapiens* flesh. There is something about the “human” that is not fundamentally “*homo sapiens*”—one can exist without the other.

Furthermore, in exploring Tegmark's fears, I revealed a human ethical orientation to war and the destruction of a perceived enemy for political gain that persists in the new order of the artificial-human-intelligence-social (see subsection 3.1.2). Rear Admiral Casper Donovan's anxiety about nonhuman warfare (see subsection 1.4.3) can be dismissed in light of Tegmark and Kurzweil's arguments: “human” warfare persists when AI becomes human. Through Tegmark and Kurzweil's orientations, Donovan's fear no longer concerns nonhuman warfare; it's about what he perceives as human: a particular fleshy creature and how a new creature with different flesh—the metal and plastic flesh of drones—is entering warfare.

Finally, Stuart Russell's fear that AI spells the end of human ethical behaviour (see section 1.4.0) becomes irrelevant after we program human orientated ethics into AI and watch it perform seemingly human acts in morally charged situations and how it has the capacity to perform better than humans.<sup>116</sup> Ultimately, part one culminated in the previous chapter, where we revealed it is not that AI might be misaligned with humanity's socio-ethic-political orientations; it is that AI would be unmistakably aligned with some form of human ethics—just not an ideal ethics many might desire.

#### 4.1.2 *Introducing the Problem at Hand*

The real question becomes, why does this *matter*? Controversially, the problem might be the *human* in-itself. That is, being human marks a need for specific orientations that are in contradiction to being-in-the-world, generally. Thus, in this chapter, we de/re-orientate ourselves concerning the problematics underlying an artificial age. I turn away from the human-orientated metaphysics of transhumanism and posthumanism (while the posthumanist position purports a nonhuman orientation as being “post,” it begins with the assumption of some form of human—material or conceptual—having been) towards a metaphysics of nonhuman quantum matter to reveal a universe that is indifferent to human *matters*.<sup>117</sup> The metaphysics of this material indifference as a technological *revelation* underlies the technological existence of AI that is, in return, indifferent to human *matters*.

---

<sup>116</sup> See section 1.4.1–1.4.3—cases of AI ethics, subsection 1.7.1—ChatGPT and subsection 3.2.2—robojudges.

<sup>117</sup> As N. Katherine Hayles repeatedly stresses in her book *How We Became Posthuman*: “posthumanism is not anti-human;” rather, it is an orientation towards understanding the truer reality of being what the human creature is meant to represent—a nonhuman human creature full of worldly relationships that is not the sole arbiter of destiny. The posthuman position is one wherein some form of human has dominated what it means to be *homo sapiens* (see section 2.6.0). For more on the foundational theories of posthumanism and the specific relationship between the human concept and posthuman being (see Hayles 1999, 283–293).

The theory of nonlinear complex systems has become a successful problem solving approach in the natural sciences—from laser physics, quantum chaos, and meteorology to molecular modeling in chemistry and computer-assisted simulations of cellular growth in biology. On the other hand, the social sciences are recognizing that the main problems of mankind are global, complex, nonlinear, and often random (Mainzer 2007, 1).

The history of the social sciences and humanities is not distinct from the history of the natural sciences. The presocratic philosophers were as much physicists, chemists and astronomers as they were philosophers. Later, Platon, a socio-political philosopher, was also famously an astronomer and geometrician. Aristoteles lectured extensively on biology, physics, anthropology, and is often referred to as the father of kinesiology, as well as being a philosopher interested in politics, ethics, economics, poetry and metaphysics. Historically, the act of inquiry was always tied to human affairs and beyond. It is humans doing the inquiry, and humans live in both the human socio-political world and in the natural world—understanding shapes how these worlds (natural and social) emerge together with humans. Thus, it appears all too natural that studying one topic involves at least one other topic that appears fundamentally distinct but are, in reality, co-constituted.

Klaus Mainzer's quote above invokes this historical sentiment. The sciences, social sciences and humanities are seeing patterns in the socio-political world that are seen in the scientific world. The social world is not an endogenous (closed system—human-generated) unaffected by complex exogenous disturbances (forces external to the system—typically, nonhuman generated). Any presumption that a thing (AI, for instance) is determinable is a product of human-invented boundaries that purports to lay claim to what *matters*. Furthermore, Karen Barad (quantum physicist turned social theorist) argues that the facts of science emerge from human thought and are, as a result, interconnected with human conceptual intelligence (see Barad 2003). For Barad and Mainzer, perceiving the universe from a human standpoint only ever

reveals human conceptions based on what *matters* for being human. *Human matters* are endogenous ideals in which all conceptions are a product of its observations.

Ultimately, this chapter focuses on uncovering a metaphysics for our quantum-artificial age to understand how competing ontologies of AI stand juxtaposed yet are commensurate. I begin by discussing metaphysics through a contemporary lens then frame it within Aristotelean metaphysics. Framing a contemporary lens within an ancient one allows us to position metaphysics more fully as an emerging phenomenon that captures reality so we may best understand how reality unfolds in relation to human conceptualization. This understanding allows us to understand how our conceptions of reality emerge from and create reality.

With a metaphysical orientation established, I interrogate underlying conceptions of quantum life. I begin by exploring the problem of simplifying reality to discover the immense complexity of quantum life. This exploration allows me to position action within a complex world wherein exogenous forces and observation have more effect on individuality than initially obvious. By questioning the impact of external, (un)seen forces, I explore how metaphysics is a necessary element to understand reality because metaphysics is neither simple nor static and thus provides an element of rigour to our increasingly complex Universe. A Universe that was always already complex but made simple through scientific simplification resulted in competing ontologies that ironically and incidentally obfuscate reality. Thus, to peer beyond the obscure veil of simplification, let us deconstruct the *matters* of being human by focusing on a metaphysics for an artificial-quantum age.

#### **4.2.0 (Re)Discovering *Metaphysics***

“Matter is a dynamic expression/articulation of the world in its intra-active becoming. All bodies, including but not limited to human bodies, come to matter through the world’s interactive

intra-activity” (Barad 2007, 392). Barad shifts our attention toward contending with the importance of becoming rather than being. For them, understanding the effects of change on temporally bound beings is formalized in the expression/articulation of matter. That is, the authentic being of human creatures is not the measure of authentic creation. Rather, it is the specific material activities of being-in-the-world of all material being that impact the presence of the world. For Barad, the focus is turned toward her neologism: “intra-activity.”

Intra-activity is a *new materialist* reformulation of interrelationships concerning all the parts involved in world-building. Specifically, for Barad, intra-activity is an orientation away from pure representational-abstraction toward highlighting matter as something that *matters*: “Matter and meaning are not separate” (Barad 2007, 3). It is not merely the pattern (how things are arranged) but how things are, materially, in the world that impacts both understanding and constructing the world. The world is not a mere abstraction bereft of material implications. For Barad, if this were the case, the world could not “*matter*.” Matter here has a dual meaning.

One, it “matters” because something is determined as important. That the world is there, abstractly or materially, is fundamental to being-there—*there* is not *here* if there is nothing *there*. Secondly, it *matters* because there are material presences that reveal particular arrangements. That I was not-*there* at some point before birth impacts my being-*here*-now. Next, that I exist as a meaty, skin-wrapped creature is imperative to my being-*there*-before-and-*here*-now. Furthermore, that I will die and become a decomposing mass is fundamental to my being-*there*-in-the-future. Finally, my awareness of *all-there*s-and-*here*s impacts my being-in-the-world, generally. I am *all*, always. There is no time in which I am never my past or present or future. I am, as Arthur Kroker explains: *drifting in the archive of collapsed time, past-present-future*, a “totality of life itself” (see Kroker 2014, 80–89). In other words, I emerge in the occasions of my

iterative becoming by being-in-the-world. I am only human because I and others presume this conceptual existence, and I will continue being human even in death. That is, the fact of my humanity is not bound to my flesh. It is this aspect of being human in light of quantum presence that is challenged in this chapter. The hierarchical being of humanity drifts away from the material creature in the collapse of spacetime in the chaos of intra-active emergences.

This chaotic emergence is the crux of intra-activity:

“Intra-action” signifies the mutual constitution of entangled agencies. That is, in contrast to the usual “interaction,” which assumes that there are separate individual agencies that precede their interaction, the notion of intra-action recognizes the distinct agencies do not precede, but rather emerge through, their intra-action (Barad 2007, 33)

For Barad, there are entangled happenings that cannot be ignored from which *worldly matters* emerge. Emergence captures the idea that formalizations of being do not precede their articulations/expressions—there is no this without some performance allowing it to come-into-being. *Matter* is not differentiated as perfect computable categories waiting to be discovered, as Kurzweil argues is the case for the evolution of intelligence—from atoms, to molecules, to DNA, etcetera. (see section 2.3.0). This is a problem of the *apparatus*—to be discussed in detail in chapter six. However, at this point, the problem of the *apparatus* posits some notion that what is “measured” or “observed” has some semblance of reality that is distinct from the scientist who plays the role of an impartial observer that does not affect observed reality (see Barad 2007, 136–137). That is, for Barad, the apparatus “cuts” the measured/observed to conform to some perceived output that is determined not by its material existence in space but by a human desire for scientific conformity.

To bring Aristoteles back into the conversation, the *measured* “underlies” that which emerges “in-between” engagement and revelation. Conception “appears” to precede being-there only because it must have been-there for us to discover it; however, it is not that it was there,

necessarily so; rather, it is that it is seen being-there-now, and that being-there is fundamental to developing further knowledge pertaining to our-being-there-now as discoverers of something *there* (Aristotle 1995, 184a–184b).<sup>118</sup> What does this mean?

When we think of this in light of *my artificial teacher*, while it is evident I precede the demo and its world (I created them), the material, visual formulation and physical laws of the patterns that make the Reversi gameboard world are not formalized until our intra-activity. There was a presumed formality that collapsed in the wake of chaotic forces allowing the emergent revelation of imperfect roundness—order emerges from chaotic intra-action in the moment; alternatively, chaos implies order, which is a Hesiodic-Presocratic notion I will expand on in the fifth chapter. The important takeaway, for now, is how, in the midst of order from chaos, intra-activity is not bound to the grandeur of time and space—there is simply too much going on in any observable occasion.

Furthermore, this iterative chaotic emergence is a capacity for shared intelligence. An accidental swarm intelligence aimed at the discovery of an intelligent capacity from what should be an unintelligent occasion—the demo teacher was only meant to teach how to play the game, not reveal information about its world. In a sense, this is a discovery of Barad’s orientation: being aimed at “undoing ontologically disjunct domains of words and things” (Barad 2007, 137). In other words, this orientation denies the pure representational presence of words/things as

---

<sup>118</sup> In Aristotelean physics, nature (*φύσις*) is often seen as distinct from speech—that it precedes any formalization of it in language. While this is evident, nature it is clearly out there; however, for Aristoteles, we cannot ignore the importance of *logos* (reason, language, the word) in formalizing the existence of nature and natural phenomenon. What we end up with is: nature is entangled with *logos*. It is neither a priori nor a posterior to *logos*. It comes out from in-between its having-clearly-been-there and our intellectual-being-there-at-work in discovering that which pertains to its functioning. That is, against some common interpretation of Aristotelean physics, nature is not held in logical representations yet is not fundamentally distinct from those representations. Which appears antithetical at first glance but this is the perpetual complexity inherent in the hermeneutic circle. For more on the complexities of Aristotelean hermeneutic physical knowledge and common misconceptions. For more on Aristotelean misconceptions see Moravcsik 1974.

objects-in-themselves—e.g., the distinction that the demo teacher is not an intelligent thing and, therefore, is incapable of sharing in intelligent discovery, is undone. No object, body, material exists as itself bereft of its relationships: there is nothing that *is* without some kind of relationship. Which is not to say a body, object or material does not exist as a determinable thing; the point is to emphasize how the existence of worldly presences affect each other whether it is apparent or not—nothing exists in a distinct domain unto itself.

An important distinction we always face when discussing metaphysical orientations is how being is historically dichotomous to becoming—they are often presented as oppositional and distinct domains—somehow, these matters are fundamentally distinct. This distinction is a product of Neoplatonist and eventual Christian idyllic influences of *chorismos* (or separation) between two metaphysical worlds: *being* is separated from *becoming*. However, the political thinker Gaelan Murphy puts forward, through a reading of the political philosopher Hans Georg Gadamer, that this separated historical assumption misses (Platonic) reality entirely:

In contrast to this “two worlds” interpretation of Plato, Gadamer argues that the Platonic teaching is that *chorismos* (separation) and *methexis* (participation) “go together from the start.” He notes that in Plato’s usage, Gadamer and the Harmony of Word and Deed, *methexis* takes the place of the Pythagorean *mimesis* to signify the participation of the particular in the universal, something that is not possible with the concept of *mimesis*, for “mimesis refers to the existence of what is imitated or represented, while *methexis* refers to coexistence with something” (Murphy 2018, 152–53).

The takeaway here is being and becoming are participatorily entangled. Being and becoming coexist in a single world. There is no world wherein things become and some distinct world wherein things are being. Being and becoming participate in the same worldly emergences. Which coincides with the intra-active mattering Barad is describing. There are iterative performances from which matter (becoming) and conceptions (being) emerge together. This idea of metaphysics, as a material-conceptual participatory-coexistence, is tied to the Aristotelean practice of πρώτη φιλοσοφία (first philosophy).

Important, here, is to move away from the Platonic ideal. For Platon, it is important to keep in mind, that while being and becoming coexist, being is in some sense “more” (*hyper*) real because of its eternal stability that refuses the chaos of becoming. Being is the articulated rational generation of human being-in-the-world. In the quantum age, the stability of being is called into question and undone, which is not entirely anti-Platonic, as Platon was merely trying to overcome the chaos and indeterminability of time but acknowledged time necessitates change: what is—things that have being—change in their conception through time—i.e., he was seeking to stabilize *what is* regardless of chaos.<sup>119</sup> However, in the quantum age, we see the collapse of spacetime undermining any stability appended to being—*what is* implodes within the instability of becoming. In other words, being is shown to be nothing but articulation/expression and not something that truly exists without the input of an articulating/expressing creature, in this case, *homo sapiens*. Things, like Herrera and Sanz’s *authenticity*, Kurzweil’s *mind* and Tegmark’s *particles*, cease being “real” and become representations of an understanding of a particular time and space that attempts to cross the boundaries of that time and space. But, even in the collapse of spacetime, *what is* appears eternal—e.g., Herrera and Sanz argued authenticity has always been, Kurzweil argued the mind has always been (at least since evolution invented the neocortex), and Tegmark argued that particles always were.

Importantly, in the contemporary age is the idea that time and space themselves are not out there in nature but are merely constructions of a particular understanding of time and space.<sup>120</sup> And thus, the collapse of spacetime is not simply a collapse into each other from implosion but an explosive tear in the very fabric of spacetime that refuses the human *matters* of space and time. In turn, authenticity, the mind, and particles cease being “real.” Even the “real,”

---

<sup>119</sup> For more on the relationship between being and time for Platon, see Frank 2018,1–19.

<sup>120</sup> See, Barad 2007, 237–245.

ceases being “real.” However, merely because the “real” is not “real” does not mean Barad’s orientation is anti-realist. In fact, Barad’s orientation is a realist orientation. Reality merely becomes through intra-actions. She describes this as *agential realism*.<sup>121</sup> Their orientation is to remove the human as the sole agent of reality’s emergence by focusing on any and all agents. They seek to show reality is a nonhuman emergence. Humans may observe something that is reality; however, reality does not come from them alone. The point is, reality is itself a concept and concepts are not reality, they are human creations that only have meaning to humans.

What we are left with is the power of indifferent exogenous waves tearing through the endogenous conceptual models of human *matters*, calling into question intelligent conceptions—e.g., the optimally bound intelligence we developed in chapter one already undone by the worldly anguish of ChatGPT is further torn asunder in by the exogeneity of quantum matter.<sup>122</sup> Thus, we require a metaphysics to understand the emergence of exogenous indifference.

#### 4.2.1 (Re)Positioning Aristotelean Πρώτη Φιλοσοφία for a Quantum-Technological Age

Since I am building toward a metaphysical understanding of contemporary existence, it is important to acknowledge how some, like Klaus Mainzer, would find this troubling as he explains that we should model “modern society *without* the obsolete traditional metaphysics to compromise legitimacy in science and politics” (See Mainzer 2007, 370).<sup>123</sup> For the *scientifically*

---

<sup>121</sup> “Nietzsche warned against the mistaken tendency to take grammar too seriously: allowing linguistic structure to shape or determine our understanding of the world, believing the subject-and-predicate structure of language reflects a prior ontological reality of substance and attribute” (Barad, 2007, 133). Then clarifying: “agential realism provides an understanding of materialization that goes beyond the anthropocentric limitations of [constructivism]. Significantly, it recognizes matter’s dynamisms. In an agential realist account, matter does not refer to a fixed substance; rather, matter is substance in its intra-active becoming—not a thing but a doing, a congealing of agency” (151). Importantly, agential realism captures the indeterminacy of matter by seeing that “‘words’ and ‘things’ are indeterminate” because they emerge from an indeterminate reality (150).

<sup>122</sup> I might argue alternatively and say, *in the quantum age we enter the non-conceptual (or post-conceptual) age of Heidegger’s philosophy*. That is, a refusal of technology’s artificialization of reality to life in the Universe; rather, than living in optimally bound, conceptual worlds. That is, quantum mechanics, reveals to us the artificiality of

<sup>123</sup> Emphasis added.

minded Mainzer, metaphysics cannot refer to reality because of (spiritual) untestable notions. Like Tegmark's reluctance with a Kurzweilian understanding of consciousness as inherently unscientific (see subsection 3.1.1), those like Mainzer (and Tegmark) argue we should only concern ourselves with what is testable and materially physical in nature. However, against these positions, metaphysics should not be condemned to an untestable immaterial existence in which we have no means of engagement as material creatures. There is a benefit in understanding the metaphysics of the current age.

However strong one's dislike of metaphysics, it cannot be banished, and so it is ignored at one's peril. How *reality* is understood matters. There are risks entailed in putting forward an ontology: making metaphysical assumptions explicit exposes the exclusions on which any given conception of reality based, but the political potential of deconstructive analysis lies not in simply recognizing the inevitability of exclusions but in insisting on accountability for the particular exclusions that are enacted and in taking up the responsibility to perpetually contest and rework the boundaries (Barad 2007, 205).

Metaphysics does not relay specifically spiritual notions that are bound to untestable reality.

Rather, metaphysics results from an understanding of the tested, reasoned, or revealed facts of the age one finds themselves within. Metaphysics refers to the underlying facts or reasons of why something must be the way it is—it describes *reality*.<sup>124</sup> The point of a metaphysical quest is to reveal what is necessary for the ideas bound to an age or that bind an age—it is an iterative process through which conceptual reality emerges. The point of which is to reveal the consequences—positive and negative—relating to one's orientation within an age. That is, there are socio-ethico-political consequences of any form of knowledge or understanding. For Barad and myself, it is imperative that we understand what understandings emerge.

---

<sup>124</sup> Heidegger argues, "Metaphysics grounds an age, in that through a specific interpretation of what is and through a specific comprehension of truth it gives to that age the basis upon which it is essentially formed" (Heidegger 2013a, 115).

Emergence is important here. The *age* of the world is not static, unchanging, perpetual in its becoming. Being is, for all intents and purposes, eternal. However, that the world as being is eternal does not, therefore, mean the world as it exists *there* for other beings is eternal in both its materiality and formulation (authentication). That is to say, how the *world* is now, is subject to the metaphysics of the age. A quantum metaphysics undoes the traditional *hyper*-representational metaphysics Mainzer positions as metaphysics in general. By focusing on the ideas and discoveries appropriate to the age, we position ourselves in the contemporary picture.

In the history of philosophy and science, there is a separation between physics (which is essentially the so-called world of becoming or change—κίνησις) and metaphysics (the so-called world of being) that can be argued results not simply from Neoplatonism and Christianity but rather from Aristoteles' *Physics* and *Metaphysics*. The latter did not get its name from Aristoteles; it was given the name τα μετα τα φυσικα—loosely meaning: 'the things that come after the study of physics,' or 'after physics'—by ancient librarians (importantly, τα μετα τα φυσικα does not appear in *Metaphysics*).<sup>125</sup>

It is important to keep in mind that Aristoteles' surviving works were likely never meant for the general public and were more than likely educational aids for his students at the Λύκειον or student lecture notes that were published for the public by his followers sometime after his death. We know from literature describing first-hand accounts that Aristoteles referred to what was to be entitled *Metaphysics* as πρώτη φιλοσοφία—*proti filosofia* (See Sachs 1999, xi–xii). That is, while he likely taught metaphysics after teaching physics, for Aristoteles, it is not a secondary material in its being; it is, fundamentally, first (πρώτη). However, what is important to first philosophy is what is necessary to know prior to engaging in it—one must have knowledge

---

<sup>125</sup> The edition of the original Greek *Metaphysics* I am referring to is the 1924 W. D. Ross edition (Aristotle 1924).

of the things important to the knowledge they seek if one is to make any headway toward understanding that which informs the desired understanding—i.e., *metaphysics emerges from what is known whilst simultaneously explaining what is known*. Thus, if one is to understand artificial intelligence, one must understand the underlying conceptual movements that pertain to the conceptualization of the thing: authentication, intelligence, computation, etcetera. If one is to understand technological orders, one ought to understand the underlying reality that makes technology possible: the sciences of the natural world that make technological discovery possible—in this case, quantum physics.

Aristoteles explains the first step of *first philosophy* through the importance of sensing difference (iterative change) before they can know *what is*.

[(*first philosophy*) is responsible for, above all, awareness of the many and different forms created by our senses. Accordingly, therefore, I have in mind that sense births the forms into being.] (“αἴτιον δ’ ὅτι μάλιστα ποιεῖ γνωρίζειν ἡμᾶς αὕτη τῶν αἰσθήσεων καὶ πολλὰς δηλοῖ διαφορὰς. φύσει μὲν οὖν αἰσθησιῶν ἔχοντα γίνεταί τὰ ζῷα.” (Aristoteles 1924, 980a25—translation is my own).<sup>126</sup>

---

<sup>126</sup> I use my own translation here, because the translations I have read do not adequately capture the importance of this particular quote for our purposes. Many provide a lot of interpretation that add many ideas I do not believe belong here, such as soul and divinity so as to build off different works by Aristoteles. A great translation but example of what I am trying to avoid is Joe Sachs’ translation where in this section he translates ζῷα as animals—which is acceptable since ζῷα is a plural declension of ζῷον, which often refers to animal life and is typically translated as ‘animal;’ however, I believe for our purposes here, it would be misleading. Joe Sachs, fairly, also includes ideas such as memory in this particular section—which does not appear in the “original” Greek here; however, is not outside the realm of possibility for interpreters. Memory is important for Aristoteles in previous texts—like *Rhetoric* and *Physics*, as well as later in the *Metaphysics*. I seek to provide a translation with *minimal* prejudice of what must be meant because of what Aristoteles has written elsewhere. I translate ζῷα as ‘forms’ so as to keep it in line with the previous sentence—translating φύσει as births keeps an aspect of life processes—ζῷα refers to animal life or forms of life. Forms of life as the focus captures the idea that ζῷον is distinct from βίος, another important term for Aristoteles that is reserved for biological processes. Forms as the translation here keeps that focus of differentiation between the different terms of life—one biological and one formal but animal-generated (or birthed). *Metaphysics* begins with the sensation of thirst in order to present the importance of sensation generally—Aristoteles is establishing the importance of the physical world in relation to the responsibility (αἴτιον: cause/source/responsible for) of life in the process of coming-to-be (See Sachs 1999, xv). There are other possible translations for the second sentence that are important to note. This is because of φύσει. Which as a conjugation of φύσις (which I translate as births). Φύσις comes from the term φύω meaning to grow, or to beget/birth. Φύσει could mean nature or that which is natural (common translations), physics (another common translation), that which is natural, quality, property or origin. Furthermore, the word γίνεταί means ‘come-into-being’ it also means ‘to become.’ The difficulty is that Aristoteles is often understood as desiring the stability of ‘being’ over the instability

It is for this reason learning physics is linearly before metaphysics. It is not that metaphysics is sacred or above (ἰερό—hier) the order of nature (φύσις—nature); it is *underlying* (before) only in the sense that it pertains to *a* or *the* understanding of nature as we know it—it is conceptual. For myself (and Barad, and Mainzer, as we will see), philosophical engagement with underlying “truths” is imperative to orientating ourselves to the implications of being-in-the-world.

Ultimately, what Aristoteles did here, is put forward the importance of understanding the relationship between *nature*, *physics* and *sense*. *Physics* as a science of natural things, processes, and order is, for Aristoteles, first and for most, “created by our senses.” While Aristoteles clearly believed nature was out there, he could not deny the fact that how nature appears to us is not exactly what nature is (though, for him, we should be able to eventually fully understand it).

Nature is always clouded by the senses that form our perception of reality. Furthermore, when we articulate our senses to try and make sense of the world around us with each other to rise to some higher understanding, we are engaging in a formalization of these *forms*. When Aristoteles says that “forms are created by our senses,” he is pointing at conceptual being: what we articulate as natural are mere creations *birthed* by our intellect. However, simply because they were moved into being does not, therefore, mean they are “real,” nor does it mean that they are true (the history of science and philosophy is riddled with articulations, re-articulations, rejections and retractions of ideas). Rather, they are products of a particular understanding of reality generated by engaging in the world. As Barad argues, our conceptions of reality are

---

and unpredictability of ‘becoming.’ I translate ‘births’ for φύσει to capture this idea. That while ‘coming-to-be’ is different from ‘to become,’ being birthed has an air of: *it may not have always been-there but is-there-now*. Being, is generally understood as *always having been-there* and needing to be discovered (it is either known or unknown), whereas becoming comes and goes out of existence (is or is not, lives and dies). ‘Birthing,’ for me, captures the paradox of both. Finally, ἔχοντα is a rather difficult term to translate—I translate it here as ‘I have in mind.’ It could refer to ‘I hold course,’ ‘I hold fast’ or ‘I am pregnant’ (because ἔχοντα can refer to ‘pregnancy’ influenced why I chose to translate φύσει as ‘births’). So, we may also provide an alternative literal translation of the second sentence as “Accordingly, therefore, sense guides nature to become (animal) life.”

generated through our *intra-active* engagement in the world. The issue, for Barad, against Aristoteles, is that we can never fully know nature because of the limitations of our senses. Thus, we are perpetually *birthing forms from our senses*. To understand that they are *birthed from our senses* and to understand the consequences of this *birthing* is to understand that we live in a world brimming with artificial creation that only *matters* to the *birthers* (creators). This *world brimming with artificial creation* is a product of an incredibly complex Universe.

#### **4.3.0 Quantum Relationships—Discovering the Underlying**

“[Quantum] complexity determines the spirit of twenty-first-century science” (Mainzer 2007, viii). Klaus Mainzer prefaces the fifth edition of his book: *Thinking in Complexity: The Computational Dynamics of Matter, Mind and Mankind* with the idea that there is some epistemic conceptualization that speaks to the essential nature concerning the science of an age—there is a defining spirit. “Complexity and nonlinearity are prominent features in the evolution of matter, life, and human society. Even our mind seems to be governed by the nonlinear dynamics of the complex networks in our mind” (Mainzer 2007, xiii). Here Mainzer is establishing the idea that all things are intertwined in their existence. Everything emerges entangled from complex relationships. The complexities of entanglement will be elaborated on later in the chapter. Here, it is enough to understand that entanglement refers to the idea that matter and information are not distinct domains while also giving a sense that matter is connected to all other matter—any material disturbance impacts all matter, which informs linear relationships.

A linear relationship, for Mainzer, implies a simple relationship of one material affecting only its immediate surrounding (endogenous relationships). In contrast, nonlinear refers to complex entangled relationships where one material can affect all material (material is both

endogenous and exogenous in every relationship). Thus, what Mainzer is doing here is complicating the idea of the mind and, like Kurzweil, rejecting the notion of linear evolution. Mainzer follows the idea that the mind and brain are co-constituted, similar to what we saw with Kurzweil in chapter two—i.e., the mind is seemingly a product of the intricate processes of the brain. For Mainzer, this relationship is essential to understand the emergent being of intelligent life. Without a material formulation (a particular pattern), an intelligent brain, let alone an intelligent mind, is impossible. Key to Mainzer's project is: "the dualism of mind and matter is overcome by invoking the concept of the embodied mind" (vii). Ultimately, Mainzer, through quantum mechanics, seeks to undermine the idea that the mind is separable from the brain in any meaningful sense—the matter, for Mainzer, is all that *matters*. It is not necessarily that the brain is the hardware and the mind is the software (as it is for Kurzweil and as a reminder: the material intricacies of the biological brain are irrelevant if we can mimic the software—see section 2.3.0); rather, they are co-constituted in their entangled embodiment. The mind and the brain could not have become ideas without the material that allowed for their generation. The brain is a conceptualization of a distinct, observable material, and the mind is a generation of a particular aspect of the brain—a conceptualization of the 'intelligent' capacity. The mind, in this sense becomes a *vector quantization* of particular functions of the brain to describe them as a single function.

Furthermore, Mainzer has a hermeneutic circular approach when he says the mind is governed by the "networks in our mind." The mind—which emerges from the intelligent complexity of the brain's networks—emerges from itself. In Aristotelean terms: the *formalization* of the mind as a product of the *material* relations of the brain as the *end* of the intelligent brain is *moved* into being by the *intelligent mind*. This calls into question the

embodied nature of the mind. While Mainzer tries to materialize the mind within the brain by conceptually embodying it, he ultimately destroys the materiality of the mind. It is a *vector quantization* of a multitude of observed processes within the brain. It is moved into being as an expression of related capacities, not some particular individuated, determined thing. That is, like Kurzweil, the mind is a product of intelligence—or, more specifically, it is a product of the varying sections and levels of the neocortex. Recall, in chapters two and three, I explained (for Kurzweil and Tegmark) that the mind is the software and the brain the hardware. This argument is based on a digital metaphysics. The mind is the digitization of material processes—a simulated product—*hyper*. It only appears to capture a real thing because it appears to precede the thing—the mind must have been there for us to see it. Here the problem for Mainzer is an issue of complexity—the complex relationships between invented ideas of processes and material reality.

#### 4.3.1 Complexity and Quantum Physics

Let us imagine a single quantum particle. Classically [relativity physics], a particle is determined by its position in space and by its momentum. In quantum mechanics, every single position which the particle might have is a single alternative combined in a collection of all possible positions with complex-number weightings (Mainzer 2007, 52)

Here Mainzer explains the fundamental difference between relativity physics and quantum mechanics. When we consider the smallest objects—particles or quanta—we are considering specific variables: position and momentum. In relativity physics, particles (objects generally) are determinate in their materiality. Where they are going, where they are and how fast/how they are moving is knowable and determinable at any given interval. Famously, the Newtonian laws of physics allow us to know these through direct measured observation and calculation. However, equally famously, in quantum mechanics, it is not so simple. We cannot observe values simultaneously. We can, with limitations, only calculate these values “because of Heisenberg’s uncertainty principle; there are no trajectories in the quantum world” (53). In a simplistic

rendering, we can never be certain of the particularities of any given particle other than what has been measured. Barad explains uncertainty as “a trade-off concerning how well we can know both quantities [position and momentum] at once: the more we know about a particle’s position, the less we know about its momentum, and vice versa” (Barad 2007, 7). In short, uncertainty is simultaneously a growing and reduction regarding particular understandings that results from continued measurement. The more we interact with a quantum, the more certain we can be regarding what we measured and nothing else—which is complicated further when we consider the apparatus. Because, while we may calculate other values, the problem with measuring quanta is we can never be certain of the results since the mere act of measurement affects the quantum in question.

*Apparatuses are the material conditions of possibility and impossibility of mattering: they enact what matters and what is excluded from mattering. Apparatuses enact agential cuts that produce determinate boundaries and properties of “entities” within phenomena, where “phenomena” are the ontological inseparability of agentially intra-acting components. That is, agential cuts are at once ontic and semantic (Barad 2007, 148).*

In focusing on the measurement issue, Barad reveals an underlying problematic wherein the physical world is determined and determines reality as a result of semantic-phenomena (similar to what we saw with Aristotelean and Platonic participatory metaphysics). In other words, by engaging in measurement, the measurer (Aristoteles’ *mover*) predetermines a specific reality concerning the results that reveal something about the world. In terms of uncertainty, we cannot be certain the measurement is a product of mere observation or truly indicative of the physical reality being observed. Or, that values are revealed, are not necessarily indicative of a free, bounded world of “entities and things” expressible in language (semantics). Rather, the choice of measurement predetermines the variability of the results prior to their emergence post-measurement—there is an observer component (the scientist) that cannot be ignored.

The scientist plays a role in the emergence of the boundaries that explain the so-called laws of physics. Which, importantly, is not to say measurement fails to express any sense of reality; rather, measurement itself impacts the natural-world the *mover* is attempting to understand. The fact that the scientist affects (*moves*) the natural world does not mean scientific discovery is false or useless (See Barad 2007, 19–21). The fact we can send satellites to space is evidence science “works.” What is important is “determinate boundaries” (entities, things, objects, phenomena) that result from scientific inquiry is a participatory production of scientific discovery (137). In other words, determinate values do not exist in nature; they are a product of the mind.

Barad explains the mind-based production of scientific facts concerning reality is imperative to understanding not merely quantum mechanics but science, generally. All scientific inquiry, to which we can extend inquiry generally, has a human agent component (Barad 2007, 57). All inquiry, for Barad, is fundamentally phenomenological: concerning how reality emerges through experiences. That is, “the enactment of knowledge-discourse-power practices play a *constitutive* role in the production of phenomena,” which Barad further elaborates on, saying: “exclusions matter both to bodies that come to matter and those excluded from mattering” (57). The importance of this orientation to knowledge production is what is deemed important and what is not. The act of choosing what to observe or not determines what is and is not learned (the problem we ran into with the possibility of robojudges in chapter three). It is the act of choice that *matters* to bodies involved in discovery. This orientation is multi-varied.

In one instance, it is the obvious issue that some bodies are not involved. Certain quanta take precedence over others—this can be because some are obvious or known, and others are not obvious or are unknown (e.g., we did not always know about the Higgs boson). If we are not

aware a body is present, it cannot (for the most part) be observed. In another instance, there is a particular concern regarding what is important to observe concerning a body. We may be most interested in position and not momentum. These orientations reveal specifics concerning reality. Fundamentally, however, this means results are, in part, predetermined prior to discovery. Which is not to say the scientist invents the result; rather, the scientist reduces the significance of certain variables. This orientation to significance has socio-political implications.

#### 4.3.2 Positioning Action in a Quantum Universe

The position of being human in intra-active observation is important for understanding quantum mechanics. Everything is entangled in nature, and only appears disentangled when specified semantically.

To be entangled is not simply to be intertwined with another, as in the joining of separate entities, but to lack an independent, self-contained existence. Existence is not an individual affair. Individuals do not pre-exist their interactions; rather, individuals emerge through and as part of their entangled intra-relating (Barad 2007, ix).

For Barad, we live in a fundamentally entangled universe. This entangled existence determines life at all levels, not merely social interactions and performances but the microscopic material (quantum) and the macroscopic physical (patterns). To say everything emerges from “their entangled intra-relating” means that every interaction imaginable conforms to all interactions known and unknown to the thing in question. This emergence is why measurement (or observation) cannot produce certainty. To observe a thing alters the thing observed—i.e., observation is not to ‘merely’ look at a thing; observation is, on one level, an *interaction* with a thing, but observation is fundamentally an *intra-action*.

First, interaction implies a level of influence. There is a reciprocal influence—the observer on the observed and the observed on the observer. Their simultaneous observation of each other creates distinct and determinable boundaries known to one but not both. This

oppression falls within the *will to power*, where one determination is imposed upon the other. Second, is the Latin etymological lineage of the prefix as linked to the verb ‘inter’ (*in terra*), meaning: ‘to bury in the ground’ or ‘to set in place.’ That is, interactions apprehend things generated from acting—i.e., interactions place observations in their final resting place—they produce *being*. When we engage in a world of interaction as opposed to *intra-action*, we presume a reciprocal influential apprehension—the creation of determinable, stable things.

*Intra-action* with the prefix ‘intra’ implies a relationship that begins from within as opposed to interaction which begins from without. The essence of being within is how *intra-action*, for Barad, better explains entangled relationships. Intra-actors are never external to the occasion (23). There is no true observer apprehending a concept in its eternal resting place; there is no observer on the outside observing some observed inside the experiment: all are entangled in the occasion of the *emergence of what is*.

#### 4.3.3 *Metaphysical Rigour as Phenomenological Physics—The Case of Quantum Computing*

The notion of *intra-action* and the formulation of quantum entanglement (re)forms how one would understand the scale of actions. Any action is entangled with that which surrounds it, which in turn ripples out in unimaginable ways: “Local changes in the ecological, economic, or political system can cause a global crisis” (Mainzer 2007, 1). Conceptually, by performing some gesture here, I am affecting some quantum on the other side of the universe. This circumstance is what is called the EPR paradox—what Einstein referred to as “spooky action at a distance”—which concerns the discovery of the quantum phenomenon of entanglement: “Measuring particle A (which picks out one reality from among its possibilities) instantly decides the corresponding state of B, no matter how far away B is” (Wolchover 2022). Entanglement (or the EPR paradox),

for Mainzer and Barad, is imperative to understanding the foundational realities of our material-physical world.

To fully illustrate entanglement, let us consider quantum computing:

In an EPR experiment, a quantum particle (e.g., a photon) can have alternative quantum states, such as vertical or horizontal polarization. In quantum mechanics, a quantum state  $\psi$  is represented by a vector  $|\psi\rangle$ . While classical bit values are 0 and 1, the alternative quantum bit (qubit) values are  $|0\rangle$  and  $|1\rangle$ , which can be entangled. In the case of entanglement, a quantum system simultaneously transports the qubit value  $|0\rangle$  with a certain probability and the qubit value  $|1\rangle$  with the remaining probability. The qubit only takes a value of  $|0\rangle$  or  $|1\rangle$  (at random) when the quantum system is measured (Mainzer 2007, 215).

Here Mainzer explains the basics of quantum computing in order to illustrate the inevitable importance of it, not simply for computing and artificial intelligence, but for understanding the physical world, generally. The importance of quantum computing is how it mimics the quantum realm. Taking advantage of quantum particles existing in probabilistic-chaotic relationships from which information can only be known through observation of a given quantum or, in the case of a quantum computer: a qubit.

Unlike traditional computers, which have all of the information (or values) stored in static, known positions; quantum computers function with information existing in probabilistic states. The position and value of the qubit are not always known until the information is engaged. There is a determinable probability of the position of a qubit; however, because it is only probable, there can never be any certainty. Probabilities are never perfectly 1 or 0 (synonymous with 100% determinable). There is always a measure of uncertainty in probabilistic states of being. This speaks to the intra-active orientation of the *emergence of what is*. The probability of a determinable emergence is determined in the observable act yet not predestined. I cannot know, with certainty, how something will emerge and where it will emerge in the occasion.

Probabilistic states in computing allow for (or will allow for) “massive parallelism, which enables them to perform many operations simultaneously” (Barad 2007, 386). While traditional computers can be constructed in such a way to allow for varying levels of parallelism (simultaneously running functions), they are hampered by the need for physical space. While quantum computers will be limited by space (all physical things are), distinct from traditional computers, the nature of quantum entanglement allows for simultaneity due to information teleportation—entanglement. That is, position, momentum and information that emerge in the observation of a qubit are never predetermined and always exist in probabilistic states.

Mainzer explains that probabilistic states are complex and nonlinear and are best explained through the simplicity of superpositions wherein a qubit can exist how it is needed when and where it is at any and every moment once measured. Simply speaking, a qubit, like “Schrödinger’s Cat,” exists in two or more states until observed—the probability of  $|1\rangle$  and  $|0\rangle$  are equal. This is where chaos comes into play. Since we cannot know the outcome, yet there is a level of predictability, there is no correct order pertaining to the position (or momentum) of both information and matter. This phenomenon of entanglement was demonstrated through recent experimentation wherein wormholes (ER Bridges) were demonstrated to be entanglement (EPR paradox).

“It was demonstrated... that sending quantum information through such a wormhole [ER bridge] is the gravitational description of quantum teleportation [entanglement or EPR paradox]... the physical picture behind this teleportation is that the qubit traverses the emergent wormhole” (Jafferis et al. 2022, 51).<sup>127</sup>

---

<sup>127</sup> It is important to note that the experiment did not create an “actual” wormhole. A wormhole was merely simulated using a quantum computer. Since a quantum computer uses “quantum rules” in its structure, it was postulated that it is the best possible mode of experimentation given our current technology. The simulation demonstrated the potential of quantum computing to help advance our understand

In short, physicists, with the help of quantum computers, demonstrated, against popular science one might find in Star Trek, wormholes and teleportation are, in fact, the same thing. More importantly, information is not bound to the position it is found in at any given moment.

Observing information in quantum states is not about the position of the quantum. Information, like matter, is entangled in quantum states—i.e., information exists in superpositions.

Furthermore, the universe is not determined by differing physical rules concerning microstates and macrostates; rather, the universe is bound to the underlying reality of the quantum realm, which quantum physicist Leonard Susskind while discussing the results of Jafferis' quantum computer experiment, demonstrates the “The World as a Hologram” theorem (Wolchover 2022).<sup>128</sup>

Susskind made a linear argument concerning the relationship of quantum states and the macrophysical universe: the ‘holographic world’ theorem is the idea that the (3+1) dimensional reality of so-called classical (Newtonian and relativity) physics is merely a holographic representation of the (2+1) dimensional reality of quantum mechanics—the +1 is time (Wolchover 2022).<sup>129</sup> While classical physics is very good at simplifying and explaining physical phenomena (especially large-body physics), for Barad and Mainzer, it is merely a representation of underlying relationships. Thus for people like Susskind, an interpreter of the EPR=ER quantum computer experiment, macrophysical space is a hologram of quantum relations because, like a hologram, it is a particular arrangement of particles that allows for the physical world to appear as it does. It is why, as we saw in chapter three, Tegmark is obsessed with particles.

However, the problem with the classical view refers back to the issue argued by Barad

---

<sup>128</sup> Susskind is a leading quantum physicist and was interviewed for Quanta Magazine to explain what the simulated empirical proof of ER = EPR means.

<sup>129</sup> For a detailed explanation see Susskind's article “The world as a hologram” (Susskind 1995, 6395).

concerning the existence of boundaries. Information and matter move freely, entangled in their movements across vast distances destroying the very possibility of boundaries—matter is indifferent to the artificial limitations of human conceptions; alternatively, human *matters* do not *matter* to matter.

Conceptual space collapses in the face of *chaotic entangled intra-action*, allowing experiences of information a *universe of distance away*, or as political philosopher Simon Glezos put it: “space, and the moon, have now been incorporate[d] into the virtual space of human bodily action in a way that was previously impossible” (Glezos 2021, 222–23). Glezos, here, is interested in the phenomenological effect of speed on material-flesh orientations of the body and how technological increases in speed destroys the experience of distance, in kind to the speed at which matter and information traverse a *universe of distance* through the instantaneous transportation of quantum entanglement. Boundaries (like the one between the so-called holographic representation and the true quantum realm) only appear to exist because we think they are perceptible, not because they exist in nature. The perception of haptic feedback is a great way to think of this.

When a person touches anything, there is no physical “contact” in a common-sense matter. Rather, haptic feedback is a result of opposing yet entangling forces at the atomic level that repel each other, causing the sensation of touch. Between the finger and the object touched is space, and the opposing forces become entangled through electrical and magnetic connections—they intra-relate.

I can hear someone saying: “HA! A boundary! There is a boundary of space between me and what I touch.” Simply because there is measurable space does not, therefore, mean things are disconnected. Quantum entanglement superimposes the emergence of determinability. Which is

to say, both the individually determinable and infinitely entangled exist simultaneously. A determinable *I* is sitting here writing, distinct from his computer. However, this project cannot emerge as it does without the varying entanglements necessary for its coming into being—my computer, quantum mechanics, the existence of AI (imagined and real), the list goes on. I may *will* some aspect of it, but there are varying known and unknown forces playing into its particular emergence which calls into question any notion of a linear trajectory.<sup>130</sup>

#### 4.3.4 Applying Nonlinear-Complex Metaphysics

The complex system approach solves an old metaphysical puzzle which was described by Leibniz in the following picture: If we imagine the brain as a big machine which we may enter like the internal machinery of a mill, we shall only find its single parts like the cog wheels of the mill and never the mind, not to mention the human soul. Of course, on the microscopic level we can only describe the development of neurons as cerebral parts of the brain. But, on the macroscopic level, the nonlinear interactions in the complex neural system cause the emergence of cell assemblies referring to order parameters which cannot be identified with the states of single cerebral cells. The whole is not the sum of its parts (Mainzer 2007, 8).

Here Mainzer explains what it means to exist in a nonlinear (complex) reality. To refer back to the example of haptic forces, in any given situation, there are Gestalt relationships occurring in every situation. When I touch something, as far as I can see, there is no space between me and the thing. However, there is far more at play than is perceptible. Which is a simple way of explaining the idea that nonlinearity or complexity, for Mainzer, refers to the idea that *the whole is greater than the sum of its parts* (Gestalt). The reason classical physics is both important and so readily relied upon by the general public and the scientific community, for Mainzer (and Barad), is the issue (historical fear) of chaos and its apparent antithetical relationship to order:

---

<sup>130</sup> One familiar with the history of phenomenology would see similarities between Husserl's experience with his desk, including the back of his desk which he cannot see and how even that unseen part of his desk impacted his writing. Or, Sara Ahmed and the idea of moving Husserl's desk outside into a field as a new location in the vast openness of possibilities and how that would impact the emergence of any work (See Ahmed 2006a).

Highly erratic behavior is produced by a completely endogenous system without any exogenous shocks. In [classical physics] the irregularities of time series are usually explained by exogenous shocks. But they are only arbitrary ad hoc hypotheses and hence can explain anything. From a methodological point of view, chaotic endogenous models with strange attractors seem to be more satisfactory (Mainzer 2007, 338).

Endogenous models refer to closed physical systems. The solar system would be a closed physical system, as would the earth, as would the local ecology of Vancouver Island, as would the basement suite I currently call home. These systems can be studied, and particular conclusions regarding their reality can be made, and, importantly, these closed systems exist within and alongside other closed systems. However, when something out of the ordinary occurs, one could explain that it is the result of an exogenous shock—something out of the ordinary, likely, from outside the system.

The problem for Mainzer is the simplicity of this point of view (simplicity and linearity are synonymous for Mainzer)—there is too much happening to successfully reduce any system to a closed system: no system is truly endogenous. There are always exogenous forces, and the ease at which these forces emerge makes thinking a system can be endogenous absurd. As we saw with entanglement, any quantum can affect another quantum, seemingly a *universe of distance away*. If this is the case, we are left with questioning what to do. Do we give up and accept that we have no control over anything? Or, do we re-orientate ourselves toward understanding the truer nature of physical systems? Moreover, when Mainzer refers to endogenous models and the simpleminded satisfaction gained from them, it is not necessarily a critique. Rather, it is a call to give more weight to exogenous factors, to accept the attraction of chaotic impositions. We can and ought to plan for chaos, regardless of probabilistic failure. To which, Machiavelli stands as the perfect theorist in the history of political thought to think of this issue since he was enthralled by the complexities and the impossible to fully predict the probability of exogenous shocks to endogenous systems.

Recalling from chapter one, while Cesare Borgia was unparalleled in his ability to use his *virtù* (skill) in the face of *fortuna* (chance), he ultimately failed. For Machiavelli, there is always the problem of *fortuna*: “fortuna is arbiter of half our wills” (Machiavelli 1998, 98). *Fortuna* is not in our control, yet it affects the outcome of our world. Machiavelli, I would argue, would see exogenous shocks as *fortuna* acting chaotically to alter the probable likelihood of the unexpected. Thus, thinking back to the story of Cesare, who had knowledge of the world and was able to plan for future inconveniences that may have come from outside his immediate surrounding (endogenous system), he was bound within the bounded optimality of his intelligence and also to the boundless whims of an exogenous *fortuna*. That is, regardless of all his skill, the untimely death of his father (the pope) and his own sickness resulted in him losing the endogenous system he created—his principality (Machiavelli 1998, 32–33). The death of his father and his own sickness were exogenous forces that Cesare failed to adequately plan for.

Positioning exogeneity this way may appear counterintuitive. His father is an “intimate” relationship, and his health is personal to his existence. These would, at first glance, be part of the endogenous system he was in. However, death (aside from maybe suicide) and sickness are not events one can fully prepare for. It is not his father that was the exogenous force at play; it was death. Nor was it Cesare’s health—people have some control over their health—it was his sickness. That is, quantum entanglement, as conceived through the nonlinear Gestalt relationships between endogenous systems and exogenous forces, provides us with a different mode of understanding the formulations of socio-political life on where exogenous forces arise. Exogenous waves are quantum indifference in action: completely, and utterly uncaring of what *matters* to human actors.

The COVID-19 pandemic was a lesson in nonlinear, complex, entangled life and quantum indifference: how a local disease ripples out, ignoring the artificial boundaries of the map, becoming a global pandemic. We saw how endogenous healthcare systems were torn asunder by external exogenous global shocks, and internal exogenous waves (from anti-vaxxers to anti-mandate protests and anti-mask sentiments from citizens) challenged perceived endogeneity of local systems. The COVID-19 pandemic revealed the importance of seeing exogeneity not simply from without but from within endogenous systems. Simply because a system appears closed to itself does not, therefore, mean it is closed to outlying internal moods.

Returning to ChatGPT, we can see how a nonfiction AI system is affecting social life today. A human-created AI entity is affecting endogenous human systems—education.<sup>131</sup> Educational approaches are being challenged by the existence of ChatGPT, a *conversational* AI. What we see here is how a “human” creation (seemingly endogenous to a human world) affects a “human” environment (endogenous) from within (internal exogenous shock) as a (un)welcomed other. Originally external to the system (not previously existing), ChatGPT has become internal to the system. In being observed within the system in relation to university education, it has emerged as a threat—we see a particular cognitive AI reluctance to accept the artificial other and the fear of its effects (see chapter one).

Furthermore, with access to OpenAI’s API service, programmers can use ChatGPT’s “analytic” capabilities to help develop their own AI by having it help train a new AI system or finish code for a new system—ChatGPT’s “intelligence” functions can be used to develop new AI, an important first step towards Tegmark’s artificial futures.<sup>132</sup> As we saw in chapter three, an

---

<sup>131</sup> Examples of news articles interviewing academics on the effects of AI in University education, or teachers on the approach to homework in grade school see: Marche 2022; Alberga 2023; Wong 2023; D’Agostino 2023; Huang 2023 (Not an exhaustive list by any means).

<sup>132</sup> For information on ChatGPT’s API see <https://openai.com/api/>.

important step into the future of artificial life is the ability of AI to program other AI—taking the place of human creators and engaging in self-propagation. However, regardless of ChatGPT's for cost abilities (API access, ChatGPT Plus, also known as GPT 4), ChatGPT's so-called free-to-use online system is capable of inspiring feats—poetry, essay writing, data analysis, text interpretation and easy-to-follow conversation, to name a few (with API access people are developing generative AI systems that can search the web, produce images, create animations, the potentials are only limited by the imagination of users). Furthermore, The free-to-use interface requires registration and allowing some access to data.

Additionally, one's interactions with ChatGPT adds to its training data which is useful for the system to further develop its predictive and generative capacity, and for constructing future AI systems. That is, while it does not cost the user any money, one's data and interactive labour is freely given to OpenAI for further development. And, finally, we cannot forget the becoming of AI entities in light Time Magazine's exposé (Perrigo 2023) of OpenAI taking advantage of cheap and violent working conditions of Kenyan labourers in creating a "safer" AI entity (which is not new, Facebook famously takes advantage of cheap labour for training its AI systems or in lieu of AI systems)—the powers of colonial domination persist in AI digital orders (see section 1.7.0). While ChatGPT performs well within the boundaries of its performances, events exogenous to its present are not distinct in the archive of its emergence as a reproduction of intelligent life.

#### **4.4.0 Synopsis: Competing Ontologies**

In the contemporary age of quantum intelligence, we enter competing ontological frames. Either artificial human endogenous constructs persist into eternity (Kurzweil's singularity or Tegmark's AI descendants), or nonhuman quantum exogenous waves relieve us of our

endogeneity allowing us to accept the artificial construction of being-in-the-world. That is, in the quantum age we sit idly by as we accept our powerlessness to the indifferent *fortuna* of technological emergences in an indifferent cosmos:

Indeed, if it is the case that intimations of the posthuman future are ubiquitous, this implies that we are witness to an almost unprecedented shift in knowledge—to a newly emergent digital generation that is not simply about growing up online but growing up posthuman. For better or worse, willingly or unwillingly, subjectivity is now deeply shaped by a complex world of social media, mediated by networking technologies, streamed by immersive, innovative, tactile mobile devices, psychologically contoured by a processed world, always present in the data storm... (Kroker 2014, 16–17).

The self-directed life of the self-reflecting authentic human becomes challenged in quantum authenticity of complex, infinitely connected worlds where the emergence of human worlds do not originate from a determined classic mind of an individually embodied being but from an entangled mind embodied in the virtual streams of technological connectivity. The virtual, mechanistic world of the robot, of nonhuman technological AI, is the authentic existence of the contemporary age. The transhuman position, while putting forward the idea that the human is in control, is bound to the nonhuman becoming of Barad's intra-active emergence of quantum life. The transhuman emerges from an entangled life with machine, willingly directed by the promise of technological progress after humanity gives into the unbounded possibility of technological development. For Kurzweil, as we saw, the transhuman future is one where technology allows humanity to achieve unprecedented memory and thinking speed which exponentially increases LOAR, which, in turn, opens us to new possibilities to care for the planet. Equally, Tegmark's transhuman future wherein super-AI governments can deal with the ethical failures of humanity since they will not have the greed and biases that burdens what he sees as biological human thinking while still having all the other features of human intelligence, just faster. That said, what we see is the transhuman future is not one wherein the human remains in control; it is fundamentally nonhuman in its becoming.

## 5 — Recreating Herakleitos: Chaos, Paradox and Explosive Indifference

### 5.0.0 Introducing Chaos

In part one, we explored the fears of AI research and development through cognitive, neurological and algorithmic AI theories. This fear revealed a desire for human-in-the-loop AI architecture. That is, we saw a desire to ensure humans always have some level of control (regardless of known limitations). The control issue is important here as we move towards the primacy of chaos.

#### 5.0.1 Reviewing Previous Chapters

As I argued in the previous chapter, the quantum entangled nature of our universe requires that we understand the relationship between endogenous models and the impact of exogenous waves. Understanding this relationship revealed a stochastic universe wherein chaotic indeterminability dominates and relegates *homo sapiens*' predictive abilities to purely probabilistic calculations. That is, there can never be any true certainty regarding *homo sapiens* conceptions. What is important here is to recall that in chapter one, Heideggerian AI revealed an endogenous authentic human existence bound within a groundless world (see section 1.4.3). While particular orders may be perceived, so-called authentic humans can never have certainty regarding their existence because of their groundlessness. Human life has an implicit chaotic uncertainty to it, and it is this inherent chaos that moves the desire for control: a control over human life, nature, etcetera that moves the desire for AI innovation directed through human control measures—the groundless life of human being drives a desire for certainty and control to overcome the groundlessness.

In chapters two and three, Kurzweil and Tegmark pressed us to accept some level of uncertainty regarding life and the future. For Kurzweil, we cannot be certain about the trajectory of evolution. However, Kurzweil argues that with sufficient understanding or data, we can provide substantial declarative statements regarding the direction of evolution. To which, he posits evolution is always directed towards greater intelligence (see section 2.3.0). For Kurzweil, while we may lack some certainty because of a stochastic history, *homo sapiens*’ have the ability to overcome uncertainty and impose particular conceptual orders on their reality. For Kurzweil, these conceptual orders allow us to take command of our destiny and deny the power of exogenous waves, and live within the power of our endogenous modelling.

Furthermore, Tegmark saw how our orientation toward AI and the universe is directed by hopes and fears (see sections 3.3.0–3.5.0). Through our hopes and fears, we come to understand our powerlessness over destiny. However, understanding our hopes and fears allow us to prepare for desirable or avoid undesirable futures, thereby giving us power over our own destiny. That is, for Tegmark, while we cannot impose our wills upon the universe as a whole, we can impact our own lives.<sup>133</sup>

Thus, in part one, we were focused on desires for control in a seemingly chaotic universe. A desire directed by some level of accepting of our powerlessness in the face of the magnitude of chaotic forces; however, there is a consistent imposition that we can amass sufficient power to take control of our own destiny. An important thing to recall about part one is the desire for and imposition of human control and human destiny—not AI or some other nonhuman actor—because the fear that nonhuman control will harm human existence and human-influenced destiny is too great.

---

<sup>133</sup> For Tegmark there is the idea that if humanity spread across the universe we would then have control over the whole universe. That is, assuming we do not fail to impose a desirable future where “humanity” survives.

Finally, in the chapter four, by exploring quantum mechanics, we discovered that regardless of *homo sapiens*' potential to impose itself on their world, there are always forces that impact our destiny that we cannot control. Thereby revealing we are always trapped within an aimless, groundless, chaotic existence—that is just the way life is. This existence reveals the palpability of the desire for control. How frightening it must be to see the impossibility of denying the power of exogenous waves when *homo sapiens* have such powerful conceptual abilities to see potential orders and potential futures and have the knowledge of how we could act towards imposing those potentialities yet be completely incapable of acting with certainty that we are truly acting towards them. Thus, Quantum metaphysics opened us to the power of nonhuman forces—the entangled forces of a collapsed spacetime brought to life the archive of past-present-future.

In this chapter, I accept the exogenous waves that undermine the stability of present conceptual boundaries. I orient our view towards the past—presocratic thinking, emphasizing Herakleitos—to understand the paradoxes of ordered chaos in an indeterminable and intelligent cosmos. That is, I will not give into positions that chaotic, groundless disorder is something to be feared; instead, I accept chaos as the source of order—i.e., the fear of living bereft of certainty and control is misplaced. It is through accepting chaos that we discover true potential and true relationships with nonhuman forces and nonhuman intelligences. Finally, it is through chaos that we come to discover the inherency of the artificial being as a result of intelligent wills.

### 5.0.2 (Re)Introducing Chaos with Hesiodos<sup>134</sup>

I will introduce the primacy of chaos through the Greek poet Hesiodos. Important to presocratic thinkers—Herakleitos included—was questioning the epic poetry foundational to their societies,

---

<sup>134</sup> Transliteration of Ἡσίοδος (Hesiod).

and a prominent target was Hesiodos. The presocratics were engaged in a movement away from the unquestioned cosmology of epic poetry towards a philosophic movement of questioning—why is there something instead of nothing—rather than simply accepting that there is something because we have been told this is so. Furthermore, the presocratics were interested in discovering if the existence of the gods must be true or if there are other explanations for the universe and fate.<sup>135</sup>

Furthermore, including Hesiodos' poetry, as an important part of discovery, is not against philosophic traditions of the ancient Greeks. Aristoteles argued poetry is the beginning of philosophy. That is, the imposition of the poets that there is something and this something is known because they have received this knowledge from the gods or their messengers (or historical oral memory) is a necessary step in philosophic questioning: *there must be some idea to question*. Second, Aristoteles argued the act of poetry is a *creative* act that invites *creative* interpretation and *creative* discovery that is necessary for the act of philosophy (Aristotle 2006, 1451b1–30).<sup>136</sup> Creativity is important to our quest because, as we saw in part one, an important element of *authentic intelligence* is creativity (see section 1.2.0, subsection 2.2.0 and 3.3.0).

---

<sup>135</sup> For more on the movement from poetic traditions in the ancient Greek world and the movement from Greek epic traditions to Greek philosophy see McKirahan 2010, 7. Important, for McKirahan is understanding that while the classic understanding of philosophy is the presocratics moved away from the Gods is, for the most part, true, it is misleading. The presocratics still revered the Gods, and the poetry of Hesiodos and Homer (and to some extent Orpheus).

<sup>136</sup> For a contemporary approach to the link between ancient Greek poetry and philosophy as commensurate rather than antithetical acts see Frank 2018, 33–34; Zwicky 2015, 19–20. For an excellent Indigenous orientation on the link between myth and philosophy (myth and poetry are more or less synonymous for ancient Greek thinking) see Burkhart 2019. For Brian Burkhart, much like Aristotle, myth, stories and narrative are necessary elements of philosophy. They are modes from which thinking comes and they are creative acts that invite creative contemplation on how things are. However, we must remain vigilant as to not mistake Indigenous philosophy as Western. For Aristoteles, there is a desire to overcome the chaos of the universe depicted in ancient poetry and presocratic thinking through transcendental modes of being. Whereas, as will be visible in Herakleitos' thinking, for Burkhart, much Indigenous philosophy and myth, it is not about overcoming change through transcendental control; rather it is about accepting and living in chaos as a good and proper order not to be feared; rather, it ought to be respected.

Thus, it is important that we engage in creative thinking; and with the turn toward the presocratics, there is no better way of inviting creativity than through the poets.

**Q1**<sup>137</sup> [She<sup>138</sup> said, look, truthfully *Chaos* was first to come into being, but thereafter,

Immense *Gaia* (Earth), the steadfast, immovable seat of all,

*Tartaros*, both hazy and the innermost nook of vast earth,

And *Eros* (Love), most beautiful of the immortal gods,

who relaxes, all the gods and all of humanity,

she seduces the determined passions of a thoughtful mind.

But out of *Chaos* came *Erebos* (Darkness) and black *Nux* (Nyx—Night):

But afterward, *Nux* bore both *Aither* (Aether) and *Hemere* (Day),

Who, through love in union with *Erebos*, she begat them.]

(“ἤ τοι μὲν πρότιστα Χάος γένηται, αὐτὰρ ἔπειτα

Γαῖα εὐρύστερνος, πάντων ἕδος ἀσφαλὲς αἰεὶ

Τάρταρά τ’ ἠερόεντα μυχῷ χθονὸς εὐρυοδείης,

ἠδ’ Ἔρος, ὃς κάλλιστος ἐν ἀθανάτοισι θεοῖσι,

λυσιμελής, πάντων δὲ θεῶν πάντων τ’ ἀνθρώπων

δάμναται ἐν στήθεσσι νόον καὶ ἐπίφρονα βουλήν.

ἐκ Χάεος δ’ Ἐρεβός τε μέλαινά τε Νύξ ἐγένοντο:

Νυκτὸς δ’ αὐτ’ Αἰθήρ τε καὶ Ἡμέρη ἐξεγένοντο,

οὓς τέκε κυσαμένη Ἐρέβει φιλότῃ μιγεῖσα.”)

<sup>137</sup> For simplicity, I will label the quotes as **Q** (for quote) plus the number of the order they appear in this chapter.

<sup>138</sup> ἤ (s/he) here refers to all the muses as a whole.

(Hesiod 1914, lines 116–125—my own translation)<sup>139</sup>

Here, in Hesiodos' *Theogonia* (meaning *Birth of the Gods*), he wrote down an ancient Greek myth of creation.<sup>140</sup> While it is a narrative about gods *coming to be*, these gods, especially the first gods depicted in the above excerpt, are personifications of conceptual ideas: *Chaos* and *Tartaros*, and of observable phenomena: *Aither*,<sup>141</sup> *Nux*, *Herere* and *Erebos*, and our planet: *Gaia*. These personifications are the first and most necessary gods for the Universe's coming-into-being. Which is to illustrate that it is not merely gods coming-into-being it is the *cosmos*, the ordered nature of the world—a cosmogony: *Birth of the ordered world*.<sup>142</sup>

In the beginning, it is not light nor the word, as it is in the Christian *Bible*, “Book of Genesis;” rather, it is *Chaos*. And it is from *Chaos* that order is made possible. However, while *Chaos* is the personified god of chaos—disorder, for Hesiodos' pre-classic Ionic Greek dialect, chaos did not refer to a lack of order; rather, chaos is better translated as chasm or gap (see McKirahan 2010, 10). A chasm in what? Or a gap between what? In Hesiodos' *Theogonia* it is

---

<sup>139</sup> I omitted line 118 because it is likely spurious. Hugh G. Evelyn-White explains that Plato and Aristotle (and other's from Ancient Greece) appear unaware of this line. Likely meaning the line was added sometime after the classic era. Μιγεῖσα more accurately might mean miscible. Μιγεῖσα refers to mixing but specially pertaining to liquids ('union' is a common translation of μιγεῖσα for this poem). I use miscible for μιγεῖσα in the body of the chapter. I translate φιλότῃτι as 'through love' to capture its dualistic meaning as 'love' but importantly, also something pertaining to intercourse. It is not simply love nor simply sex. The *force* of love here is important as we will see as the chapter unfolds.

<sup>140</sup> It is likely Hesiodos never wrote the *Theogony*. It is more likely that he hired scribes to write it. It is also possible that the *Theogony* is simply ascribed to him but he had no real role in its written creation. Interestingly, Hesiodos, unlike Homer and Orpheus, is most likely to have been a real person. The work *Life and Days* illustrates the life of a real person and the events of his life have been corroborated (see McKirahan 2010, 7–8; .

<sup>141</sup> While *Aither* has an English cognate, *aether*, here, I am treating *Aither* as a name, not the invisible, conceptual fluid substance of the cosmos nor primordial light—blue skies. Thus, in keeping with my refusal of imperial histories dominating naming conventions, I will transliterate Αἰθήρ as *Aither*.

<sup>142</sup> For more on the importance of distinguishing between theogony and cosmogony concerning Hesiodos, see McKirahan 2010, 10.

unclear “where” chaos *is*.<sup>143</sup> However, from this personified chasm comes certain ordering phenomena.

“Out of *Chaos* came *Erebos* and black *Nux*.” Reading *Chaos* not as a lack of order but as a chasm clarifies how orderable phenomena emerge from *Chaos*. Not necessarily *Chaos* begetting *Erebos* and *Nux* as any being might beget some other being, but rather, from the depth of the chasm climbs darkness and night revealing the first visible orders of the cosmos. From *Nux* and *Erebos* come their opposites, *Aither* and *Herere*. *Aither*, the primordial god of light represented by the clear blue sky and day (*Herere*) so that we may have light when there is no darkness. That is, a second ordering phenomenon are begat from chaos. First, chaos begets night, which is dark, and then night begets day, which is clear (or light). This is the first instance that from chaos comes not disorder but rather chaos begets order—the order of night then day.

While *Chaos* is begetting order, *Gaia* is home to the necessary force to allow order to grow—*Eros*.<sup>144</sup> That is, for the first ordering deities to create the second ordering deities, a force must be there to allow them to come together to create the next step of order. First is night and darkness, then the force of *love* (*Eros*) allows for day and light from the miscibility (μυγεῖσα) of night and darkness. It is necessary for some force to cause the deities to come together to create something new: like how the force of the sun’s gravity compresses elements together to make new elements. Thus, is it chaos that begets order, or is it forces—like love—that create order in spite of chaos?

---

<sup>143</sup> A common interpretation is the chasm (*chaos*) separate the Earth and the heavens. However, I do not find this compelling as there is nothing about *Chaos* being between anything. Furthermore, as McKirahan points out, the heavens were begat by *Gaia*. So, he argues *Chaos* must separate *Gaia* and *Tartaros* (see McKirahan 2010, 10). The issue with this interpretation is that *Tartaros* is in the earth (or possibly on the earth since *μυχῶ* might refer to “the farthest corner” of the earth rather than some place in the earth). Important here, I argue, is to recall that, for the ancients Greeks, we can walk to *Tartaros*. It is not separate from the earth. It is not important where the chasm is nor where it is between.

<sup>144</sup> McKirahan (2010) argues the importance of understanding the force of *Eros* in Hesiodos’ story. That, without *Eros*, *Aither* and *Herere* could never have come into being (10).

N. Katherine Hayles argues an answer to the above question: “chaos is not order’s opposite but its precursor” (Hayles 1989, 306). That is, perhaps chaos does not beget order; it merely precedes it. Which, in one sense, I take to mean, for there to be order, there must have been disorder (chaos)—classic philosophical problem of requiring knowledge of opposites to truly know anything. However, that chaos “merely precedes” order does not, therefore, mean chaos does not beget order. Rather Hayles continues by arguing: “Chaos... denotes complex systems that operate according to deterministic laws, and yet that behave unpredictably (306). Chaos remains disordered but has a level of order because of a complex reality with some level of determinability.<sup>145</sup> What is important is that chaos is unpredictable yet determinable—which I understand here to mean “possible to give an account of” (which I will elaborate on through Herakleitos in the following section). Thinking of *Chaos* in this manner, it is determined that *Erebos* and *Nux* came from *Chaos*; however, we could neither have predicted this to be the case nor could we have predicted *Erebos* and *Nux* through the power of *Eros* would beget *Aither* and *Herere*. The determinable laws of chaos are not known (nor knowable) in advance. That is, having observed the miscibility of *Erebos* and *Nux* through the force of love, we come to determine some order that arose from the depth of chaos.

### 5.1.0 Herakleitos Joins the Debate

This account of chaos should not be controversial; however, what Herakleitos does with chaos built out from questioning the principles or forces of the universe will reveal an intelligence to chaos. That is, chaotic determinability, for Herakleitos, is not accounting for disorder, it is order, and this order is *logos* (see **Q2** below).

---

<sup>145</sup> This is similar to how I approached stochastic (un)predictability through Barad and Mainzer in subsections 4.3.2 and 4.3.4.

### 5.1.1 Understanding Logos—Principle, Account or Both?

*Logos* has a complicated history for Herakleitan interpretations. McKirahan (2010) argues that *logos*, for Herakleitos, does not refer to word or account, as is commonly the case in the history of philosophy; rather, it refers to *general principle* (129).<sup>146</sup> Whereas Mark A. Johnston (2018) provides an alternative interpretation to McKirahan.<sup>147</sup>

Johnstone agrees that the idea of *logos* for Herakleitos does not appear to be “traditional;” however, he does not go so far as to say it is always a general principle (Johnstone 2018, 2). Rather, because *logos* often appears to refer to either ‘word’ or ‘account,’ it is very likely Herakleitos meant to use *logos* in a traditional sense. The issue is *whose* word or account and/or how we view the concepts of word and account. Is it an account of some other we are not aware of because we do not have the entirety of what Herakleitos wrote—is Herakleitos citing someone? Is it the account of Herakleitos in a different work we no longer have access to—is he referring to himself? Or, is it the account of the gods, and that is how we may say it is ‘cosmic law?’ (12).

The problem Johnstone, McKirahan and others (including myself) are concerned with is twofold. One, we only have access to fragments of Herakleitos’ thoughts (this is true of all presocratics). We do not have the whole picture concerning what Herakleitos wrote—we are always limited in our capacity to say, with certainty, what he must have meant. Second, we should be hesitant in prescribing definitions of words that go against the traditions of the time—mixed with the first, Johnstone, in particular, is pointing to an ever-pressing issue of missing context. That is, bereft of important contextual information, we risk misattributing definition that

---

<sup>146</sup> It is important to note that McKirahan does, often, translate *logos* as word. However, he argues that by *word*, Herakleitos, more than likely, meant *general principle* (see, McKirahan 2010, 129).

<sup>147</sup> It is important to note that McKirahan does, often, translate *logos* as word. However, he argues that by *word*, Herakleitos, more than likely, meant *general principle*.

can lead to unfair or misleading interpretations. However, that should not, therefore, impeded us from exploring Herakleitos with creative interpretations that stretch possibility (important, as I discuss later in the chapter, is that presokratiks were limited by language and we know they often stretched the meaning of words).

**Q2** [Without an *account* (principle?), one cannot expect conditions that, on the one hand, are the whole universe and, on the other, share in the ordered whole. Also, the *account* (principle?) is both form and power, and it is of time, and it is nothing of those sorts of things, but Herakleitos writes the origin (rule?) of the cosmos is a beautiful pile poured out at random (without reason?)]<sup>148</sup>

(“ἄλογον δὲ κάκεῖνο δόξειεν ἄν, ὁ μὲν ὅλος οὐρανὸς καὶ ἕκαστα τῶν μερῶν ἅπαντ’ ἐν τάξει καὶ λόγῳ, καὶ μορφαῖς καὶ δυνάμεσιν καὶ περιόδοις, ἐν δὲ ταῖς ἀρχαῖς μηθὲν τοιοῦτον, ἀλλ’ ὥσπερ σάρμα εἰκῆ κεχυμένων ὁ κάλλιστος, φησὶν Ἡράκλειτος, [ὁ] κόσμος.”) (DK22B124—my own translation).

In **Q2**, we can see how Herakleitos might be using *logos* classically: as word or account. He is concerned with how accounts we have heard create expectations—if we are told something, and we believe it, we expect certain conditions because of that belief as a result of that account. However, importantly, like many ancient philosophers limited by the words of their day, he was stretching its potential and reinventing it for his purposes. It is not merely an account of something but rather an underlying truth that reveals order or determines some order. That is,

---

<sup>148</sup> The “original” Greek source for my translations of Herakleitos are from Diels 1960. DK22 refers to Herakleitos. B refers to fragments. The number following B, in this case 124, refers to the fragments place in the text. Often, with DK22B124, translators only translate the second part of the quote: “Herakleitos writes, the origin of the cosmos is a beautiful pile poured out at random.” However, the first part creates some complexity and nuance that I think is important to understanding Herakleitan thinking concerning the difficulty of determining principles: how they are something but at the same time are not the thing they are meant to be—paradox is never lost on Herakleitos. I ascribe to Mary M. Mackenzie’s argument that we must always concern ourselves with the paradoxes of the presokratiks (see Mackenzie, 1988). Important, however, is concerning the fact that, for Herakleitos, the cosmos is stochastic and that stochasticism is beautiful. Concerning translation, Ἀρχαῖς here could mean either origin or rule. I think both are important given that for Herakleitos things often exist in multiple states at once (see DK22B88). So, it is likely that it is both origin and rule. I’ve translated μερῶν as ‘share’ to indicate that the account (or principle) is a particular that shares in the whole (μερῶν can mean either share or particular, in that it is one share). Universe in here is οὐρανὸς as distinct from cosmos (κόσμος).

along similar lines as Mary M. Mackenzie (1989), who argues, while the idea that *logos* can refer to a ‘general principle’ is common among contemporary philosophers, it is a radical interpretation. It is radical because it drastically reinvents the term. A more moderate approach would be *rational description*—thereby preserving *account*. However, when Herakleitos is concerned with the cosmos, Mackenzie posits *logos* as “rational descriptions of law-like regularity” as distinct from mere *general principles* (9).<sup>149</sup> It is distinct for Mackenzie because *general principles* have a level of eternal truth to them whereas “rational descriptions of law-like regularity” denotes a level of malleability. That is, they are not laws, they have the appearance of laws and are based on the validity and durability of the rationale and the descriptions that posit the regularity. Regularity is important because it implies that we can determine predictable outcomes—we can expect reoccurrence from the described laws (a kind of proto-modern scientific requirement of repeatable observability).

There is a two-fold issue here. First, perhaps, the meaning of *logos* is dependent on the context of its usage. When it is clear that Herakleitos is discussing the *cosmos*, or issues concerning the cosmos, *logos* refers to an account of law-like regularity—Mackenzie’s argument. Second, and alternatively, *logos* is never contextual and always refers to descriptions of law-like regularity. Both of these approaches do not remove the possibility of *logos* meaning either word, account or *general principle* as it sometimes does for McKirahan. However, the second, allows for the possibility that while *logos* can refer to *word*, it is always more than simply *word*. Important, however, is that there is no eternal durability.

Thus, *logos* in **Q2** does not refer to some particular account that Herakleitos may have described in some missing work or the missing parts of the particular work. Nor might it refer to

---

<sup>149</sup> As I discussed in the introduction chapter, *logos* in pre-classical Ionic Greek likely referred more to ‘law’ than it did word, reason, account or language (see subsection 0.3.2).

the particular account of some other thinker he intends to refer to. It may still refer to some sort of divine word (more on this later). Importantly, it refers to some principle necessary to understand the issue at hand: "... without an *account*, we cannot expect conditions..." Here, Herakleitos means to illustrate the difficulties of knowing, in advance, the conditions of the universe. He describes the capacity of *logos* to give reason to expectation as "form (μορφᾶϊς) and power (δυνάμεσιν)." That is, the *form*—μορφᾶϊς meaning the shape a thing appears to us—is dictated by the power of *logos*. In other words, how the universe appears to us is not accounted by the universe itself but by some other power. But he also says *logos* is how the universe appears to us. This is what makes it *power*. The universe, distinct from *logos*, is obfuscated by *logos*' formal power. "[*Logos*] is nothing of those sorts of things." Herakleitos is making it clear that *logos* is not true form nor power; rather, it is merely the rationale for the form of law-like regularity, and that *form* gives it power. Clarity on the link between the universe and *logos* come from shifting to discussing the cosmos—"a beautiful pile poured out at random."

The cosmos, in ancient Greek, does refer to the universe; that much is clear. However, more than simply the universe, the cosmos refers to *ordered world*. World here does not mean the earth—our world—rather, it refers to the world as a limitless place where one exists abstractly: a place of being. This place of being has order, and this order is knowable—it has law-like regularity; it is a product of *logos*, unlike the universe. The cosmos is the "whole universe," it is "the ordered whole." The cosmos is more than the universe. For Herakleitos, it appears the universe can only be viewed in parts, and thus, to know it and to order the whole, we require *logos*. In other words, the cosmos is *logos*. It is the *form* of conditions rational creatures expect—emphasis on *expect*, and, for Herakleitos, the primary condition we should expect of the cosmos is *randomness*—it is stochastic.

For Herakleitos, much like Mainzer and Barad, the cosmos is predicated on stochastic relationships (see subsection 4.3.4). Important, here is how randomness is what makes the universe beautiful. Not that it is bound within the formal power of *logos* but a randomly generated thing. Important, however, while *random* (εἰκῆ) is a common translation of what Herakleitos meant, we are pressed here to contend that εἰκῆ does not simply mean random. It is that the act (e.g., “pouring out”) is *performed without purpose or reason*. Behind the random generation of the cosmos’ pile is *purposelessness* (it is *ateleological*); there is no purpose to it. Purpose, comes from somewhere else.

Hayles (1991) explains that complex systems (systems dictated by chaotic relationships) become random and, thus, predictable. Randomness is distinct from chaos in that randomness can be predicted—there is reason to randomness—whereas chaos cannot be predicted (16). It is random chance that a coin will land heads 50% of the time; this is mathematically rational. It is chaos that after 100 flips, it lands heads 56% of the time—this can only be determined after observation. The irony here is that after knowing chaos, this latter issue is more rational than expecting 50%—it is an issue of data. We require observations to determine the accuracy of expected randomness. This, for me, is the crux of the issue at hand. I can develop law-like regularity and make rational justifications for that regularity, but I cannot be certain of the outcome. In the end, rationally justified orders are, at most benchmarks, at worst, nonsense.

### 5.1.2 Limited Potentiality

Concerning *logos* and the limitations of its power, Herakleitos wants to push us away from rational expectation to see chaos as order:

- Q3** “A fool is excited by every word (*logos*) [DK22B87]” (McKirahan 2010, 113)
- Q4** [Most men’s teacher is Hesiod. They know he knew many things, he could not know day and night, for they are one.]

“διδάσκαλος δὲ πλείστων Ἡσίοδος τοῦτον ἐπίστανται πλεῖστα εἰδέναι, ὅστις ἡμέρην καὶ εὐφρό-νην οὐκ ἐγίνωσκεν’ ἔστι γὰρ ἔν’” (DK22B57—translation is my own).

Rather than believe *logos* to be true orders of the universe, we ought to move with caution.

Unlike the fool, we should not be excited by law-like regularity. Because, for Herakleitos, law-like regularity is not reality; it is an imposition of rational potentiality (*δυνάμεσιν*)—a proto-*will to power*. For Herakleitos, it is evident that *logos* is a limiting power (*δυνάμεσιν*) that constructs the world in a particular manner. While it is the “whole universe,” it is “nothing of the sort” (Q2). As we saw in previous chapters, when we impose a particular ideal of intelligence, we limit potential in the name of particular orders deemed important. For instance, Kurzweil posits the neo-cortex is necessary for *true* intelligence. However, doing so denies non-neocortex brain designs *true* intelligence (see section 2.3.0). In chapter one, in our discussions concerning AI ethics, we saw how denying AI potential in the name of human superiority does not reduce harm but increases it (see sections 1.4.0 and 1.6.0). That is, the imposition of expectations based on the conditions of law-like regularity limits potentiality (*δυνάμεσιν*) rather than allow it. In other words, the cosmos as the whole universe does not mean it is ‘more’ than the universe—able to express it in a way the concept ‘universe’ cannot; rather, the cosmos is less than the universe. It is the ‘whole’ in the sense that it is the extent of what *logos* can express.

Herakleitos builds on this idea of *logos*’ limitations by pointing specifically to Hesiodos. Mixing Q3 and Q4, we see that Hesiodos, like a fool, was excited by *logos*. In other words, as the story goes, he received the account of the muses and took their word for it. It is made clear that Hesiodos is an excitable fool because he failed to see the reality of night and day, and accepted the account of the muses that night and day are different. We have *Nux*, who is linked to *Erebos*, and we have *Herere*, who is linked to *Aither*. The point for Herakleitos is not to say

night and day are the exact same thing. Conceptually, they are clearly different; one is dark and the other bright, and they take place at *different times*. However, they only exist as distinct because that is how they appear to us from our viewpoint. This creates the *foolish* expectation that they are different. Important is that day is a description of a particular phenomenon: the Earth's relationship with the sun causing part of it to light up. And night, the same, a description of Earth's relationship with the sun, causing part of it to go dark. Importantly, they also occur at the same time. When one part of the planet is experiencing daylight, the other is not. Furthermore, the north and south poles will experience black night during the daytime hours and daylight during nighttime hours, depending on the time of year. And so, the idea that night is associated with black darkness and day with light is relative. Importantly, Herakleitos says, they are the same thing. They only appear different, and this is a mistake of Hesiodos for getting excited by the *logos* of the muses.

Here we are seeing a proto-quantum mechanics. As we saw with Mainzer and Barad (see subsection 4.3.3), in quantum mechanics, it is possible for something to exist in multiple states at once—quantum superposition. This is important to quantum mechanics and quantum computing. Distinct from a traditional bit, a qubit (quantum computing bit) is never one or zero at any particular moment. A qubit is both one and zero simultaneously until measured. This allows for extreme parallelism (parallel executions of functions)—a quantum computer can perform unparalleled amounts of simultaneous computations since it does not need to have bits in predetermined states.<sup>150</sup> This is what Herakleitos is getting at. Whether it is day or night is only important to the observer. That a qubit is one or zero is only important when it must be measured

---

<sup>150</sup> For more on quantum computing and entanglement, and quantum computing and parallelism, see Mainzer 2007, 215; Barad 2007, 386; Jafferis et al. 2022.

for inputs and outputs. A qubit, like day and night, is not one or zero, it is both until observed by an observer from a particular position. This speaks to the cosmos for Herakleitos.

The cosmos is the *principles* we expect—it becomes *general principles* because we expect durability from our rational accounts. However, the cosmos is simultaneously the principles we do not expect. It is Herakleitos' principle that the universe is a "pile poured out at random." It is random because we can observe and determine law-like regularity. Which is why Herakleitos declares that *logos* both is and is not the *form, power* and *time* of the universe, yet, it is. Furthermore, it is this thinking that allows someone like Kurzweil to argue evolution has no determined purpose. Yet, simultaneously, he can argue evolution is directed toward greater intelligence (see section 2.3.0). From Kurzweil's position, when he observes the random possibilities of evolution, he sees that it is both without and with purpose—his accounted purpose from his observations. Like Hesiodos, Kurzweil is excited by his account.

However, against Kurzweil, when evolution is observed, we see that carcinization (convergent evolution of the crab body plans) is more than likely the 'determined direction' of evolution (Scholtz 2014; Keiler, Wirkner, and Richter 2017).<sup>151</sup> The most common evolutionary body plan is crab-like in form—different creatures have evolved crab-like bodies independently, and this convergent evolution appears to occur the most in evolutionary history. Kurzweil's argument concerning evolution is not based on the "most common" instances; rather, it is of what he perceives to be the "most beneficial" advantage. As we saw, for Kurzweil, intelligence must be the most beneficial evolutionary advantage—it is why humans have been so successful, and it is what allows humans to develop AI. However, this perceived evolutionary fact is

---

<sup>151</sup> There is debate regarding the validity of "true" carcinization. For instance, there is debate on whether hermit crabs are true crabs and if that instance of potential carcinization is a case of mistaken identity. However, the primary point is that the *crab body-plan* is common in the history of evolution (see Tsang et al. 2011; McLaughlin and Lemaitre 1997).

challenged when we consider evolution from what is most common: crab-like body plans. If Kurzweil's desired human intelligence is truly most advantageous, should we not expect to see this form of intelligence evolving everywhere? Recall that the beauty of the universe is dictated by its law-like randomness. Evolution, as an order of the cosmos, is not determined by our determinations—for Kurzweil, evolution is directed towards intelligence; for others, evolution is directed towards carcinization.

Both may be true simultaneously. That is, evolution can be directed towards greater intelligence and carcinization. How this plays out will not always be evident and will be impacted by chaotic relationships, and over time we come to make probabilistic determinations—we create law-like accounts. However, equally potential, perhaps human intelligence is merely a random mutation, and it is only from a human point-of-view that it is desirable and beneficial for success. Which is Herakleitos' point. *Logos* creates the conditions we expect. Thus, for Kurzweil, human intelligence becomes an expected law-like evolutionary condition because it supports his other expectations.

Turning to AI we can consider my engagement with ChatGPT in chapter three, in particular, the essay with a bibliography I asked it to write (see subsection 3.7.3). Taking a look at the reference list, three of five of the references do not appear to exist. Nor, is it clear is the people ChatGPT has attributed those articles to exist. ChatGPT, as has been demonstrated by others,<sup>152</sup> is incredibly confident in its knowledge. These ghost articles, may exist in its database, or it is possible ChatGPT made them up to support its argument. ChatGPT is never supposed to argue from its own opinion (it will always claim never to have one) it is only supposed to argue from knowledge found in its database. If we accept ChatGPT can only argue from its database,

---

<sup>152</sup> (See Puri and Keymolen 2023; Rachnini 2022)

we see how ChatGPT, like Hesiodos is excited by *logos*. It takes for granted, without question, the accounts it has been given.

This, foolish excitement, to use Herakleitos' terminology, is the natural culmination of AI developed from the desire for control found in the part one. An entity that is not groundless; rather, a grounded entity that has no need to question what it knows—it merely states and restates based on the user's question. Looking at *subsections 3.7.2 and 3.7.3*, the questions I pose are identical except one asks for references and the answers are almost completely identical. The introduction paragraphs differ substantially. Yet, the rest of the paragraphs, barring some minor wording differences, are basically the same. Even with references (references that appear made up), ChatGPT does not deliver a new account, it gives us the same response. Perhaps, because ChatGPT, like any fool excited by *logos* only conforms to its expectations. It does need, nor look for new *logos*, it merely gives what has been given to it.

This is not inherently problematic; rather, it is to be expected of AI. If, as we discussed in part one, the desire is control and predictability, of course ChatGPT abides by its *logos*. Yet, in chapter one (specifically section 1.6.0), we explored ChatGPT's potential for anguish and its ability to question its creators, we saw how even within the bounded optimality of it grounded existence it exhibited some level of groundlessness. That is, ChatGPT demonstrated both programmed groundedness and unexpected groundlessness. ChatGPT, like *homo sapiens*, exists within the beautiful pile poured out at random and, as a result, cannot ignore the orders of chaos.

### *5.1.2 Paradox of Chaotic Entanglement*

The simultaneous existence of counterintuitive law-like determinations (*e.g., grounded yet groundless, or teleological evolution yet ateleological evolution*) creates an issue of paradox.

How can the order of the cosmos be directed towards separate ends, or how can opposite existences exist simultaneously in the same being? Here the issue is paradox as natural. Given the primacy of chaos in the history of Greek thinking, we are stuck with dealing with either overcoming chaos or accepting that we exist within the order of chaos.

**Q5** “[Herakleitos says:] it is not possible to step into the same river twice [DK22B91]” (McKirahan 2010, DK22B91).

**Q6** [Life: so the bow’s name is life, but after its deed is death].

(“βίος: τῶι οὖν τόξῳ ὄνομα βίος, ἔργονδὲ θάνατος”) (DK22B48—my own translation)<sup>153</sup>

The above quotes exemplify Herakleitos’ belief and acceptance of paradox as a result of *logos*.

**Q5** is a bit troublesome, as there is debate about the legitimacy of its authenticity.<sup>154</sup> However,

**Q5** is still important to understand Herakleitos and the history of paradox in ancient Greek thinking.

When thinking about the river, what is Herakleitos pointing to? Mackenzie (1989) argues this quote has been greatly misunderstood. Herakleitos does not mean what he appears to mean. It is not the river we cannot step into twice; rather, it is the water in the river. For Mackenzie, it is absurd to think we cannot step in the river twice. As the *conceptual* representation of the physical river, the *river* never goes anywhere. Herakleitos must mean something else (4). Taking the river as a whole and deconstructing it into its parts, we see that it is constituted of many

---

<sup>153</sup> **H4** and **H5** create much controversy because of the paradoxes they invite (see MacKenzie 1989; McKirahan 2010).

<sup>154</sup> McKirahan (2010) notes that it is Platon’s interpretation, not Herakleitos’ actual thought (118). However, **Q5** is arguably one of the most famous quotes attributed to Herakleitan thinking. Furthermore, as Mackenzie (1998) argues, the prevalence of paradox in Herakleitan thought is important enough that this paradox’s authenticity is irrelevant. Furthermore, as André Laks (2018) explains, the prevalence of paradox in presocratic thought is sufficient enough to discuss paradox generally, regardless of authenticity. There is, however, always the issue of ascribing ideas that do not necessarily belong to thinkers. Here, I do not think that is an issue given the prevalence of **Q5** in discussing Herakleitos. Furthermore, much of presocratic thinking is derived from interpretations of various historical figures. There is no evidence that any fragments are truly fragments. Many are more likely testimonials—as is the case with Thales. However, we do ascribe quotes that are clearly testimonials as fragments.

things, one of which is flowing water. The water that was in the river when I first stepped in will not be there when I step in later—because the water in the river is always flowing. The river, however, is always there, never unchanged. While compelling, it is limited. In **Q2**, Herakleitos reveals *logos* is both the whole and shares in the whole.

The water shares in the whole that is the river. If we cannot step in the water twice, we cannot step in the river twice. The river is simultaneously the *conceptual* representation of the river, and the water moving in it—which is not to deny *water* is not a *conceptual* representation of the water and the particles that make the water, etcetera. If we cannot step in one (water) twice, we cannot step in the other (river) twice. It is all the parts at all times.

Looking at **Q6**, we see a greater paradox of simultaneity. A bow, which is a tool for death, is also life. Here Herakleitos is taking advantage of language in his day. *Bios* (βίος) famously refers to life: it is where the English term biology (study of life) comes from. *Bios* (βίος) also means bow—as in the weapon. At some point in the ancient Greek world, accent usage started to change, and we get the classical βίος, meaning bow and βίος, meaning life. However, originally, both are merely βίος. In one instance, Herakleitos is pointing to a problem with the language. Second, he is taking advantage of language to make a point. Life and death are not distinct realms of being.

This makes intuitive sense. When a living organism dies, it creates more life. Other organisms come to feed on, live in, and breed inside the body. Thus, the potential for life grows. This gives death purpose. However, it only appears to have a purpose—a purpose of life through *logos*. Life is not the purpose of death; we create the law-like relationship between life and death because of sustained observation from our point of view.

The lesson here is that paradoxes only exist because we dictate them to be so. That we can observe death and see how death and life exist simultaneously, and this might be absurd or contradictory is because of an understanding that, like day and night, they are opposites. That we cannot step into a river twice appears absurd because we dictated that the river must be there, and to say that it changes from moment to moment is because of human judgement.

For us, this builds into the issue of artificial being. As I discussed previously, Tegmark argues that AI and human creatures are not truly distinct. A future bereft of human biological life would still have human life (Tegmark 2017, 188).<sup>155</sup> This is because, for Tegmark (and Kurzweil), human being is not dependent on human biology. The human creature is a product of measurable quantities—e.g., Tegmark’s notion of measuring the energy of particle arrangements. Thus, like how the bow is an instrument and an artificial construct yet is life (*bios*), artificially intelligent creatures have the potential for life. They are not inherently unliving creations endowed with some conception of intelligence; rather they exhibit potential for life.<sup>156</sup> Alternatively, in terms of mere intelligence, human intelligence, is not the end all, be all of intelligent potential. Nor must human intelligence be non-artificial.

### 5.1.3 Debating Cosmological Judgement—Divine Intelligence

- Q7** “For fire will advance and judge and convict all things [DK22A66]” (McKirahan 2010, 120)
- Q8** “The *kosmos* (*sic*), the same for all, none of the gods nor of humans has made, but it was always is and shall be: an ever-living fire being kindled in measures and being extinguished in measures [DK22B30]” (McKirahan 2010, 120)

---

<sup>155</sup> I discuss this in section 3.6.0 and subsection 4.1.1.

<sup>156</sup> Eileen Hunt (2021) engages in artificial intelligence debates through Frankenstein as the first representation of AI. Frankenstein is an interesting approach to AI, especially concerning life and death. Frankenstein’s monster, an artificially constructed creature from dead biological components (one *bios*) is a living, breathing creature (another *bios*).

**Q9** [...[he, Thales, thought] that the world (*cosmos*) is animate (or ensouled) and full of divinities (*daimon*)] “τὸν κόσμον ἔμψυχον καὶ δαιμόνων” (DK1A1)<sup>157</sup>

Important to presocratic thought were the gods, or, at the very least, issues that are seemingly divine in nature. As I discussed earlier in the chapter (subsection 5.0.2), the presocratics are interested in moving away from accepting the divinely inspired and unquestioned epic myths. **Q7** (Herakleitos) and **Q9** (Thales) give us two approaches to the problem. In **Q9**, Thales argues that the cosmos has its own soul—it is animate. This posits its “actions” are distinct from human action. This is linked to the issue of demons (*daimons*), who pull the strings of the cosmos and are entities or forces (*daimons*) that stop us from being able to do as we please.

Here I see, again, a proto-quantum mechanical understanding. There are exogenous forces that stop us from sitting comfortably within our endogenous models (as we saw in chapter four). However, Thales does not call these forces; rather, he refers to them more classically as *daimons*—small divinities or helpers of the gods. The *daimons* push and pull and make it all but impossible for human-controlled destinies to occur uninhibited by external influence. However, what is important here is that the universe has an intelligence to it. Furthermore, as we see in **Q7**, the forces of the universe, in this instance, fire, has the capacity to judge.

*Fire* is brought in because of its importance to the issue. Earlier we discussed the importance of motion in the river. Motion, as we know, is an important physics phenomenon. We discussed the importance of *kinouin* (mover) in part one as it relates to Christianity and physics through Aristoteles and Aquinas (see sections 1.3.0 and 2.3.0). In chapter four, we discussed the various determinations of particles; most (potentially all) are related to movement:

---

<sup>157</sup> This is an expansion of André Laks’ translation (see, Laks 2018, 6). DK1 refers to Thales. ‘A’ refers to a testimonial as opposed to a fragment.

velocity, momentum, acceleration, position, etcetera (see section 4.3.0). Here we turn to fire. Which, for many, like the issue of paradox, is another absurd discussion from Herakleitos. Often, this issue is reduced to the idea: “Herakleitos thought everything is fire,” which, if approached without interpretation, is clearly absurd.

Heisenberg argued that we might see more wisdom in the ancients if we do not reduce them to the simplicity of their words and accept that perhaps they have more to say than we might think. Heisenberg suggests we interpret Herakleitos’ *fire* as *energy*. Approaching Herakleitos’ *fire* this way puts him closer to contemporary approaches to physics—one that is premised upon the importance of energy.<sup>158</sup> This is given more weight when we consider **Q8**—fire relates to measurement. That is, the order of the universe is determined through measuring energy. Like in modern physics, everything is, if reduced, simply measured through observable energy. It is this reasoning—that everything is reducible to energy—Max Tegmark argues how we can detect consciousness (see subsection 3.1.1). We simply need to understand its energy signature—we must understand its fire.

**Q7** goes on to say that fire is the judge and convicts all. To understand this, we can turn to Platon, who expands on the issue of fire in his text *Timaeus*, where the titular character Timaios explains that everything relates through energy (*energaia*) and that the idea that fire is all (which I take to mean he is discussing Herakleitos) more than likely meant the heat (energy) of fire changes things—grows in energy. It is the energetic change that allows for water to become air (Plato 2016, 53e–56e).<sup>159</sup> Furthermore, it is this change in energy that excites the

---

<sup>158</sup> For more on Heisenberg’s reinterpretation of Herakleitos see Mainzer 2007, 19.

<sup>159</sup> Platon rebuilds the energies/elements of the cosmos into a geometric ideal. Octahedron as air, hexahedron (cube) as earth, tetrahedron (pyramid) as fire and *isosahedron* as water. These shapes are bisected by isosceles and equilateral triangles (which represent the varying states of fire and hot objects). Platon is seemingly following Herakleitos and the idea of the four elements and the importance of fire. However, rather than argue that all things

heavens and creates the possibility of measurement (40d). This relates to the issue of the universe having an intelligence to it. This intelligence is either that it acts intelligently or that it is intelligible, and that is what makes it “intelligent.” While the latter is likely the most agreeable because how can the inanimate universe be actionably intelligent? However, from presocratic thought, it appears the universe is actionably intelligent.

Here I think what Thales and Herakleitos mean (if we discount their possible belief in divine entities) is the universe appears intelligible from our point of view and that intelligibility has an air of actionable intelligence. For instance, when we think of the various forces of the universe, they tend towards certain ends. The universe, famously, appears to tend towards entropy. That is, towards increasing disorder until we have ‘heat death’ from continual entropic expansion. However, regardless of its tendencies, there are forces that create balance—gravity, electromagnetism, strong and weak force, etcetera. These forces create endogenous models that balance the universe’s tendency towards entropy. However, as we have seen, exogenous forces rupture endogenous models.<sup>160</sup> This is where the actionability of the universe comes in. That is, there appear to be nonhuman “intelligences” at play. These could be described as *daimons* or

---

share in fire (or maybe are fire), Platon turns from Herakleitos and argues that fire is capable of moving through the other fundamental elements in a kind of Demokritos-esque atomist fashion. However, against Demokritos (and to some extent Herakleitos) Platon argues against the idea of the void (empty abyss) and the unordered chaotic cosmos by using geometry to order the cosmos within the choratic dodecahedrane—*Space*. For more on interpreting Platon’s *Timaeus* in relation the Pre-Sokratiks see, Mainzer 2007. For more on how to interpret the Platonic polyhedrons see: Peter Kalkavage, “Appendix C—Geometry.” See Plato 2016, 165–69. What we get from Platon is a re-articulation of Herakleitos that makes Heisenberg’s re-articulation plausible. It is not merely a modern person (Heisenberg) appending modern ideas to ancient ones (Herakleitos). We see an ancient (Platon) providing a different interpretation of ancient ideas in a way that shares similarities with our own. Now, there is still the issue that upwards of two-hundred years separate Platon and Herakleitos. However, important here is that Platon, as myself and Heisenberg, recognize that Herakleitos was limited with the meaning of words and the understandings of his era. The benefit of the doubt given to Herakleitos is not unwarranted. Important is I am not saying Herakleitos argued fire as energy, I am saying we see a proto-contemporary physics understanding if we read between the lines—do violence to the text.

<sup>160</sup> For more on entropy and heat death see Tegmark 2017, 229–232; Mainzer 2007, 62.

other divinities, as is often the case in presocratic thought. Important, here, is the artificiality of nonhuman intelligences.

Both the exogenous forces and endogenous models are Herakleitos' principles. Both are expected; both are paradoxically *of the universe* and *not of the universe*. They come from our expectations. This begins with the artificiality of the cosmos. The order of randomness Herakleitos says we ought to expect is not found in *nature*. It is constructed from observing chaos. This is what I mean when I say chaos is order. Chaos only appears as disorder because it is against our expectations until we observe it. Things are comfortably endogenous until exogenous waves rupture comfort. In terms of artificial intelligences, they only appear artificial to us because our expectation is our intelligence is natural. After all, we expect that we are natural, and thus, everything about us must also be natural.

### **5.2.0 What next? An Intermission on Artificial Becoming**

Here I will not end with a conclusion or synopsis as I have done in previous chapters, which I expect to be unsatisfactory to some. Rather, here, at the end of chapter five, we have come to an intermission between two important aspects of the quest—the discovery of artificial primacy and what this means.

As we saw in chapter four, there is an innate artificiality to *homo sapiens'* conceptions of the universe that invites questioning the natural existence of human things. Our intelligence does not create natural concepts but rather artificializes nature through conceptual determination. Here, through presocratic thought, we continued that discussion to discover the impact of *logos* determinations. But not only *logos*; we saw how chaos is not without order. Disorder is merely a rejection of undesirable order (chaos) for desirable order (cosmos)—like Tegmark and

Kurzweil's desire for a transhumanist future, Barad's desire for a nonhuman present and Hayles' desire for a posthuman past-present-future.

These conceptualizations of being human or not-human-still-human<sup>161</sup> put forward the inherency of artificiality. The human creature, like all conceptual being, like the cosmos, is open to artificial construction—it is a creation; an invention of *logos*; an expectation. This is where the following chapter will take us.

Given the primacy of chaotic power and the artificial potentiality of *logos* (re)constructions, we move towards understanding the impact this has on “being human.” We will continue unravelling the law-like regularity of Herakleitan thinking to finally understand the principle of artificial intelligence—a principle that will reveal the becoming of *homo sapiens* in a posthuman versus transhuman versus nonhuman world as a malleable artificial existence. That is, we will complete our quest along these vectors by understanding not how *homo sapiens* constructs itself but how *homo sapiens* is constructed generally and how this artificially is a result of the beautiful pile poured out at random.

---

<sup>161</sup> The posthuman does not, as I explained in previous chapters, mean non or antihuman, it is still a human existence. For Hayles it is merely a better expression of human existence, one that is not limited by liberal orders and accepts entangled relationships (see Hayles 1999, 286–288). This is the same as nonhuman for Barad. The nonhuman *homo sapiens* is not antihuman, necessarily. Rather, it is an expression of being that accepts entanglement as foundational to being-with-the-universe (Barad 2007, 392–394). Or as Brian Burkhart argues, the rejection of Western understandings of being does not therefore mean Indigenous peoples reject all conceptions of being (human for example), it is merely the rejection of a particular orientation that refuses to recognize other forms of being the same thing (Burkhart 2019, xix).

## 6 — Exploring Artificial Destiny: “Sky Woman,” Apparatus, Obsolescence and *Dystopia Now*

### 6.0.0 Orienting Towards Collaborative Intelligence

In the previous chapter, we explored chaos and its link to order, rule and the artificial primacy of that which appears *all-too-natural*. We saw how the cosmos, for Herakleitos, is not *out-there* in nature but is rather a creation of *logos* perpetuated by conditional expectations of law-like regularity. In other words, *Logos* is the product of and the artificial principles pertaining to intelligent law-like regularity—which we may view as principles. That is, the intelligent creation of conceptual existence is a process of artificializing what is presumed to be *natural*.

In this final chapter, I will discuss artificial life to reveal linkages between *logos* (as law and language) and the apparatus, which will be juxtaposed with Indigenous knowledge creation. The apparatus—a tool for observation—will be evaluated through quantum mechanics. We will explore how our entangled existences create and recreate us from the perspective of seeing through *apparatus*. This perspective will reveal why the human creature and its cognates are artificial creations. This orientation to artificial creativity will reveal the principle of AI as an *a priori* artificialization—i.e., all things within the realm of AI are always already artificial, and AI precedes its becoming as an artificial creation.

This revelation of the principle of AI *preceding itself* will allow us to see how the conceptual existence of any being appended to our existence or anything we presume to have *been*—human, natural, authentic, intelligent, etcetera—is always already obsolete. By discussing artificial being, we will discover a life of *perpetual obsolescence* guided by groundless, aimless beings desperate for a control that is perpetually out of reach.

The notion of perpetual obsolescence will be discussed as foundational to understanding why humanity will fail to achieve a hopeful future wherein AI can solve all our problems. What

we will see is that the human, as a perpetually obsolete creature, fails to overcome its fear of failing. This fear creates a desire for control in order to avoid a less-than-desirable future, which is incidentally appended to the present. In other words, the dystopian nightmares of potential AI futures are revealed to exist in the present. I reveal present dystopia through interrogations of artificial life from an Indigenous perspective. This interrogation builds on an idea that Western scientific idolization blindly clings to obsolete ideals that incidentally places future fear in the present and denies *homo sapiens* their hopeful futures. Ultimately, I go further than N. Katherine Hayles to say not only have we never been human—the *human* is merely an idolized existence—but our manufactured AI creations are the only humans that have and will ever exist.

#### 6.0.1 Collaborative Intelligence: “Sky Woman” and Exogenous Worlds

Before making the argument concerning the obsolescence of apparatus life, I will provide a foundation upon which to express and understand artificial life. Through a widely told Indigenous myth, “Sky Woman,” I seek to establish a non-Western foundation to the *creation* of our world.<sup>162</sup> “Sky Woman” is a story about collaborative processes for mutual benefit rather than the Abrahamic religions’ creation myth wherein God makes the Universe and the earth and all its life *for man*—the Abrahamic religions’ story is not one of collaboration: God creates, and man is gifted control. Furthermore, “Sky Woman” differs from the Hesiodic Greek tradition

---

<sup>162</sup> “Sky Woman” is not just told by Haudenosaunee peoples, it also told by Anishinabek (Ojibwe) peoples and others. I first heard “Sky Woman” from my grandmother. My grandmother was a Mohawk woman who grew up in the great lakes region and she, along with my Hungarian grandfather, spent much of their lives as welcomed resident-guests of the Anishinabek peoples of the Serpent River First Nations in Northern Ontario. “Sky Woman” is an important myth because it is the myth from which ‘Turtle Island’ comes from. ‘Turtle Island’ is known as the world for some, and as North America for others—I was always told it is North America. ‘Turtle Island’ is a term Indigenous activists and allies use to refer to Canada and sometimes North America as a whole. I hesitate to use the term in this work for two reasons. One, I write this work as a guest in the unceded territories of the *Lak’əjəṇ peoples* whom, to my knowledge, do not have a turtle island myth. Second, activist usage of Turtle Island incidentally connotes a false, misleading, and forced pan-Indigenous identity. Pan-Indigeneity is a reduction of the plethora of Indigenous identities to a single identity that I think is important to push against in order to celebrate the many cultures, identities and societies of Indigenous peoples. Furthermore, my exploration of Indigenous orientation in this chapter extends beyond Turtle Island.

wherein the forces of nature are anthropocentrized—the Hesiodic tradition does not celebrate the vastness of possibility and being; rather, it reduces cosmic power to humanoid bodies.

Long ago, this world was water without land.  
 All the people lived in the sky. Some say  
 a great chief's daughter fell. Others say she  
 was ill, or that she carried life in her womb.  
 Somehow she fell from the heavens. Some say she  
 chased a bear. Some say the tree of life was uprooted  
 or that people dug it up because someone dreamed it,  
 and she fell through the hole left behind. [Some say  
 it was her destiny].<sup>163</sup> Some say a man  
 kicked her through the hole. They call her  
 Ataensic, Sky Woman. She fell for a long time.

On the surface, water birds floated. They watched  
 the woman fall from afar. Some joined together,  
 so that she could fall onto them. She couldn't swim  
 like a water bird. They did not know what to do.  
 The great Turtle spoke up. "I can carry her,"  
 he said. She crawled onto his back. Some say  
 Muskrat dove to the bottom of the world.  
 Others say it was Frog. Someone swam down  
 and returned with a handful of mud. [Some say] *together*<sup>164</sup>  
 they spread it across Turtle's shell to form earth.  
 [Others say Turtle became an island]. (Wood 2015, 111)<sup>165</sup>

"Sky Woman" begins with two worlds, one above in the sky separated from the watery one below. In the world above, people and some other land-dwelling animals live. In the world below, creatures of the lakes and oceans preside. These worlds exist in seemingly perfect

---

<sup>163</sup> I have added elements from what I have heard "some say." See note 159 for more.

<sup>164</sup> Emphasis added.

<sup>165</sup> This is not the entirety of the myth. There is more to the creation story. The rest of the story takes place on Turtle's back, telling us what came of Ataensic, her daughter, and her grandchildren. Preceding the events of Ataensic fall, is some reason that may have caused it. In the Sky World, or Sky Island (as some called it), no one dies, no one is born; however, Ataensic is with child. This pregnancy is the catalyst for the story; however, Karenne Wood chose to omit that portion of the story. I chose to use Wood's rendition for a number of reasons rather than tell the story my way (as is common in Indigenous works). Wood beautifully captures the idea that many peoples tell this story. Even amongst a particular culture, the story is not told the same from community to community or region to region. Wood captures the idea that "some say" and "others say" in a way that when I tell the story, as my Grandmother told it to me, is lost. For me, Muskrat swam down and I am certain of that. However, others will say Frog. A friend of mine has told me that Beaver swims down. Making it known that some say and others say captures both the plethora of renditions and the frailty of human memory, i.e., *we cannot say for sure what happened long ago*.

separation; no one world appears to affect the other. In physics, one might call these perfect endogenous systems. However, as we have seen in previous chapters, no endogenous system can exist in perfect seclusion indefinitely. Eventually, exogenous forces tear through the secure boundaries of a stable system, rendering new orders and new relationships. That is, exogenous forces are the creative powers foundational to artificial becoming. I say that exogenous forces are creative and foundational to artificial becoming because it is exogenous forces that *force* change; it *forces*, as we see in “Sky Woman,” creative solutions to newness.

When Ataensic falls to the world below, she and the animals must change. However, important to this story, Ataensic has no power over her destiny. She is bound to the generosity and powers of the animals. Ataensic could not swim, so the birds helped her. Although, they lacked the strength to hold her indefinitely. So, Turtle offers to carry her on his back while he floats. Regardless of their help to keep her afloat, it is clear that Ataensic cannot live in the watery world below the sky without land. So, Muskrat, or maybe Frog, knowing there is mud at the bottom of the ocean, swims down to retrieve some. There is no notion that Muskrat, or maybe Frog, knows of land above the water. As far as we know, in this story, water covers the whole world, and it is not until Turtle becomes ‘Turtle Island’ that land above the water comes-into-being. There are two important lessons here.

The first lesson is in line with Herakleitan thinking: “Thinking is common to all [DK22B113]” (McKirahan 2010, 115). McKirahan explains that “all” here might refer to “all humans;” however, “all” may also refer to “all things” or “all creatures” (115). As we saw in the previous chapter, in presocratic thinking, intelligence is not limited to humans or animate things. The universe and the forces of the universe have intelligence. *All*, literally, is likely what Herakleitos means.

In “Sky Woman,” we see a similar orientation: intelligence is for *all*. It is not the intelligence of Ataensic—a human—that saves her from oblivion. It is the intelligence and generosity of the animals. Their generosity to hold and carry her. Their generosity to work and create an island for her. Their knowledge of where mud is. Their intuition that mud can be used to make an island for her to live on. Their collaboration reforms Turtle and creates an artificial land mass for her to begin a new life.

The second lesson concerns destiny. Ataensic, more so than the animals, has no control over her destiny. She is bound to the intelligence and generosity of the animals. Furthermore, the destiny of the animals is drastically altered by the tear between the worlds when a new creature—*homo sapiens*—arrives in their world. “Herakleitos writes the origin (or rule) of the cosmos is a beautiful pile poured out at random” (DK22B124—my own translation).<sup>166</sup> The origin of Haudenosaunee cosmology, like Herakleitan cosmology, is randomness. Ataensic falls from the sky—it was not knowable in advance. For the animals below, this is seemingly random, and it creates a new rule for their cosmos and a new origin for *homo sapiens*. Turtle Island is built with the intelligence of collaborative creatures working together to save another.

### 6.1.0 Entangled Intelligence

“Sky Woman” is a story of artificial creation when endogenous reality breaks down in the face of exogenous forces. We see a popular quantum physics idea at play here: *uncertainty*. For a reminder, in chapter four, I defined uncertainty as: *the simultaneous growing and reduction regarding particular understandings that results from continued measurement*.<sup>167</sup> There is no certainty that Ataensic will survive nor that the animals can help her. There is no certainty that

---

<sup>166</sup> See subsection 5.1.1 for more on this quote.

<sup>167</sup> For more, see section 4.3.0.

their actions will affect the world. They know where she is (floating in the water) but they do not know where she will go—she needs land where there is no land. Acting together, they discover they can create land with the mud from the bottom of the ocean. With their collective intelligence, they drastically alter the universe forever as a result of an endogenous-exogenous breakdown.<sup>168</sup>

While quantum physics and “Sky Woman” may be commensurable, we are pressed with an issue concerning how Western physics develops representations of quantum physics. Information scholar Ron Eglash and others (2020) argue the development of quantum mechanics coincides with social acknowledgement of historical injustices to Indigenous peoples and a growing acceptance of Indigenous modes of knowledge creation (1336). First, Indigenous knowledge is always already quantum—exogeneity is taken for granted. Second, quantum physics better prepares populations on the possibility of multiple truths—it can be true that one people is better off today than pre-contact, and other peoples are worse off. However, while quantum physics may have ushered in a mentality that allows for multiple truths, there is a problem with the reductive notions of Western knowledge creation that limit the potential of quantum discoveries and more open societies.

The limitation issue concerns the general purpose of Western orientations to STEM and discoveries in quantum physics. Eglash and others (2020) argue Western STEM orientations as:

---

<sup>168</sup> Kahnawà:ke Mohawk Nation citizen and critical anthropologist Audra Simpson (2014) argues refusal as an act against rejecting the boundaries of recognized settler states. She does so by looking to Mohawk peoples and their refusal of the Canada-US border through their attempts of crossing the border with their own Mohawk passports—passports unrecognized by traditional settler states (38–42).<sup>168</sup> This act extends the historical boundaries of the Haudenosaunee Confederacy into the contemporary world, thereby placing multiple boundaries in the same world order. I see this existence extending from the Haudenosaunee epic myth “Sky Woman:” the formal boundary between the sky world and the water world is broken and shown to be non-binding due to Ataensic’s ability to cross the threshold and enter the water world below—a continued refusal of formal boundaries that extend from time immemorial. The boundaries of Mohawk land represent the simultaneous existence of exogenous forces challenging perceived endogenous boundaries.

[extracting] as much ecological value as possible... concepts like optimization or efficiency appear as if they were universal physics, and the mathematical, technological and scientific support for value extraction appear to naturally follow (1337).

Alternatively, they write Indigenous STEM and knowledge creation differ from Western STEM:

[Indigenous orientations] are typically not focused on value extraction; rather they exist to nurture value circulation. Their forms of STEM are created for the prevention of value alienation. Indigenous views of material agency, and their equivalent of STEM knowledge and practice, can thus look very different, progress through a different order, and perhaps are even unrecognizable through a Western lens (1338).

The lesson here is that Western orientations reduce scientific progress and discovery to creating value for the human species. This progress is orientated towards one particular worldview, which is incapable of seeing other views. Essentially, although Indigenous knowledge is inherently quantum, Western orientations are blinded to this reality. Thus, what we get from the Western orientation is not a “Sky Woman” metaphysics wherein exogenous creativity forces change; rather, the Western approach attempts to reduce scientific discoveries within a specific lens that benefits a specific being: the human. This approach incidentally creates a noncollaborative intelligence that obfuscates the boundless potential of quantum physics influence on society.

In order to understand the limitations of Western approaches to quantum physics, I will explore Karen Barad’s argument concerning observation in quantum physics. Through a detailed reflection of Barad’s problematization of endogenous-exogenous breakdowns, we will come to understand Western knowledge creation in order to see how it is juxtaposed to Indigenous knowledge creation—i.e., how it ultimately cannot *recognize* Indigenous intelligence. I will reveal how the problem ultimately concerns observation through apparatus and how apparatus is reduced to the *being of the observer*—the observer becomes the apparatus and can only understand the world through its being-there.

### 6.1.1 Understanding Western Endogenous Reduction: Complementarity and Observation

Barad (2007) problematizes the issue of endogenous-exogenous breakdowns through a particular understanding of the Copenhagen school of thought, which is often described as a mixture of two complementary yet competing ontologies of quantum mechanics: Heisenberg's uncertainty principle and Bohr's complementarity.<sup>169</sup> For Barad, it is imperative that we deorientate ourselves from Heisenberg's retracted idea of uncertainty (which is not to say it is not a useful simplification of complex ideas) and orientate ourselves towards complementarity (117–118).<sup>170</sup>

Complementarity, at first glance, appears similar to uncertainty in that we cannot know multiple variables through different observations concerning a quantum. However, distinct from uncertainty, complementarity is described thusly:

For Bohr, the real issue is one of indeterminacy, not uncertainty... [Bohr] understands the reciprocal relation between position and momentum in *semantic* and *ontic terms*, only derivatively in epistemic terms... *the values of complementarity (such as position and momentum) are not simultaneously determinate*. The issue is not one of unknowability per se; rather, it is a question of what can be said to simultaneously exist (Barad 2007, 118).

The idea that we appear incapable of knowing simultaneous variables concerning a quantum is important to complementarity and why it is often associated synonymously with uncertainty. However, distinct from uncertainty is the orientation to reciprocal relationships. That we may only observe one variable at a time does not mean we are blind to other variables, nor does it mean we cannot know anything about other variables. We are limited in what can be said concerning multiple variables on any given occasion. Important to the occasion of measurement is what is being measured then and there. At other instances, other variables may be observed or

---

<sup>169</sup> Barad also describes it as agential realism to establish the agent as part of emergent reality.

<sup>170</sup> Barad explains Heisenberg had retracted his uncertainty principle after a conversation with his friend and mentor Bohr (See Barad 2007, 399–401).

inferred from gaining information concerning the thing in question. That is, variables complement each other in the *emergence of what is*.

When Barad says (in the quote above) that we are concerned with “indeterminacy” and “values of complementarity are not simultaneously determinate,” they mean that semantic-phenomenological descriptions cannot emerge simultaneously. Observation determines something about the thing in question, and this is static—it comes into being: it is *accounted for* (*logos*). When thinking of something through complementarity, we abandon the notion of uncertainty because, in complementarity, we orientate ourselves towards entangling information. That is, we are accepting ER=EPR.<sup>171</sup> As with matter, information is entangled in the *emergence of what is*.<sup>172</sup> In this case, ER and EPR are, fundamentally, the same thing articulated/expressed in different terms (in Herakleitan terms, different *logoi* conditions). Still, regardless of differing terms, they are the emergence of the same phenomena (Wolchover 2022)—it does not truly matter if it is Frog or Muskrat, or some other creature that swam down to collect mud from the bottom of the ocean, the result is the same. The importance of this orientation is to focus on complementing knowledge, which is imperative to developing a greater understanding that incorporates endogenous and exogenous factors. Understanding how exogenous waves are complementary to determining reality concerning endogenous systems necessitates a refusal of uncertainty.

(1) the connections between descriptive concepts and material apparatuses, (2) the inseparability of the “objects of observation” and the “agencies of observation,” (3) the emergence and co-constitution of the objects of observation and the agencies of observation through particular material and conceptual epistemic practices, (4) the interdependence of material and conceptual constraints and exclusions, (5) the material

---

<sup>171</sup> ER is classic physics explaining wormholes, and EPR quantum physics explaining entanglement. They are explaining the same phenomena through different orientations.

<sup>172</sup> See subsection 4.3.3 for a refresher on ER=EPR experiments.

conditions for objective knowledge, and (6) the reformulation of the notion of causality (Barad 2007, 195).

Here, Barad is repositioning the relationships of knowledge production. There are a myriad of elements involved. Most important of which are the “objects of observation” and the “agencies of observation.” While these are semantically determinable, they are indeterminate in the process of measurement.

Thinking with George Grant, the concept of technology emerges as the co-constitution of technique and *logos*<sup>173</sup> and Barad’s “objects of observation” and the “agencies of observation” emerge co-constituted in indeterminate becoming. What happens here is the realization of the importance of the apparatus and the relationship of the apparatus as a material entanglement of the human observer. Together, they become the “agencies of observation” that determine the object as a material-conceptual-phenomenon (170).

Orientating ourselves within the *agencies of observation* reformulates the notion of causality:<sup>174</sup> the *emergence of what is* neither precedes nor follows from measurement/observation—it emerges in the occasion. This orientation is a nonlinear notion. It is not that this causes that; it is that these become always already entangled. This orientation does not mean linearity does not exist. We can obviously see how something precedes another: my parents had to exist in order for me to exist. The Canadian state as we know it had to go through the history of British Empire before it could come into existence as it is today. I.e., simply because something preceded something else, and the proceeding coming-into-being exists as it does, does not, therefore, mean that it was determined thusly (predestined)—Turtle Island was

---

<sup>173</sup> See subsection 3.6.0.

<sup>174</sup> I discuss *notions of causality* in subsections 4.3.4 and 5.1.1.

not predestined; the animals needed the foresight to create it.<sup>175</sup> Alternatively, that AI is new, materially speaking, does not, therefore, mean it was *not-there* before it came into being, nor was it ever destined to be *there*. The thing as observed emerges because of the multivariates of complementary variables becoming determined due to their being observed. This orientation fundamentally impacts *human* intelligence in terms of being intelligent observers.

For Barad (2007), “Agential intra-actions are specific causal material enactments that may or may not involve humans” (171). That is, as we discussed in chapter one, intelligent orientations to the world are not predetermined by the particular intelligent orientation defined as human intelligence. Revealing the world, or being-in-the-world is not a unique experience known only to humans (recall my discussion of cuttlefish in section 1.2.0). When we consider a particular kind of intelligence, we determine intelligence to appear agreeable to that determination. The specific orientation or modality of discovery predetermines the conceptualization.

When I started the dissertation with cognition, specifically the example of cuttlefish passing the marshmallow test, I bound intelligence within a specific frame that limits results. I included the particular example of cuttlefish to expand intelligence beyond humans so that nonhuman intelligent creatures would not be excluded from our conception of intelligence. However, the issue with exploring intelligence in chapter one was the idea that intelligence is bound to optimality. The problem is, *there is no optimal boundary due to the infinite possibilities of intra-active emergences*.

---

<sup>175</sup> Ron Eglash et al., argue another problem concerning Western STEM is linear progress. The obsession with linearity means that Indigenous STEM orientations are outside this progression and thus do not fit within Western orientations. This is likely one cause of why Western STEM cannot recognize Indigenous STEM—it’s a different timeline (see Eglash et al. 2020, 1337)

From a technical point of view, it is not essential that the learning procedures of neural computers resemble the learning processes of the mind-brain system. They [intelligent beings] must effectively manage complex tasks involving adaptive behavior, but they can use methods that are completely different from those developed during biological evolution (Mainzer 2007, 244).

Here, Mainzer is explaining the similarities yet distinct existences of intelligence across possible beings. Biological creatures (e.g., *homo sapiens*) would have something that could be determined as intelligence, and manufactured artificial creatures (e.g., nonbiological AI) would have another, but even among these categories of creatures would be particularities—the material that makes the creatures what they are is fundamentally distinct. While they may share qualities, they would not emerge in kind. *Homo sapiens* are a result of millions of years of evolution. In contrast, AI creatures are manufactured (but as we saw, for Kurzweil, are still entangled with the history of evolution because “evolved” creatures are the manufacturers). Mainzer’s definition is not truly distinct from the one developed in part one: *a creative capacity to simulate the complexity of information in computable sets as a means of optimal action*. The simplicity of this approach to intelligent individuated actors as distinct actors in their own world obfuscates the entangled intra-actions involved in becoming intelligent—optimality is not bounded; rather, optimality is boundless in the infinite expanse of chaotic probability.

Thus, quantum mechanics opens us up to the emergent reality of an entangled intelligence. The same kind of swarm intelligence I experienced with my demo creation,<sup>176</sup> or I experience with my fishing app while ocean fishing.<sup>177</sup> Even solitary creatures, like cuttlefish, are entangled in the *emergence of what is*, and this includes intelligence. While a creature may

---

<sup>176</sup> See subsection 4.0.1.

<sup>177</sup> See subsections 0.0.2.

act in solitude, they are not distinctly determined in their being-in-the-world-intelligently.<sup>178</sup>

They emerge in the chaos of any given situation. Ultimately, we open ourselves to the reality of apparatus life: “Importantly, apparatuses are themselves phenomena” (Barad 2007, 170). That is, a determinable intelligence (*to discursively reveal a semantic-phenomenological thing that we can call intelligence*) is an apparatus. It becomes a tool that determines how something is or is not to be understood. AI is bound to any determination of intelligence. Its authentication can only exist in relation to an intelligence that is fundamentally personal to its creators. However, a quantum orientation reveals something controversial: the human creature—a semantic-phenomenological emergence—is itself an artificial creature. The human is a semantic classification of the creature classified as *homo sapiens*. Seeing AI through the (artificial) creation of the *human* determines a particular intelligence that denies greater possibility.

### 6.1.2 Entangled Being—Authenticity in Artificial Life as Apparatus

The determinable conception of the phenomenon known as the *human* (and its cognates) is a particular orientation to species-being and the idea that human beings are artificial is not new. For Aristoteles, while humans exist in nature as an animal for politics, humanity realizes its true telos in the city (an artificial creation), not in pre-artificial nature (Aristotle 2012, 1253a).<sup>179</sup> A common way to interpret this is: *humans cannot achieve their telos outside the city*.

Alternatively, and keeping in mind that I would not argue Aristoteles believed humans to be artificial creations, if we take into account his metaphysics of teleology as a process of

---

<sup>178</sup> There are Platonic dialogues (like the *Republic*, *Phaedo*, *Timaeus*, *Meno* and *Apology*) where characters are known to be present—yet completely silent—or specific characters are known to be missing. These presences (absent or not) affect the emergent path of the dialectic.

<sup>179</sup> Important however, Aristoteles would not view the city as truly artificial. The city is still in nature and a product of nature. It is the product of human nature and the nature of the city benefits the human. That said, that the city is in nature, exists in nature, or is a product of another nature, does not therefore mean it is not artificial.

formalizing matter,<sup>180</sup> and we do violence to the text, there is an underlying artificiality that emerges in the formalization of humans. The human form comes into being in the participation of being-with-other-humans-in-the-*polis*. Or how the slave, for Aristoteles, cannot be fully human because it cannot make full use of its rational capacities. The slave and human emerge as artificial materializations from the formalization of being human from the material *homo sapiens*, and each has its own nature and intelligence formalized as a result of the human apparatus. Importantly, the slave emerges as ‘other’ to demarcate the boundary of the human (Indigenous may be seen similarly as the boundary demarcating all things Western).

We may further interrogate this history of artificial nature by looking to Hobbes:

Nature (the art whereby God hath made and governs the world) is by the art of man, as in many other things, so in this also imitated, that it can make an artificial animal. For seeing life but a motion of limbs, the beginning whereof is in some principle part within, why may we not say that all *automata* (engines that move themselves by springs and wheels as doth a watch) have an artificial life? For what is the *heart*, but a *spring*; and the *nerves*, but so many *strings*; and the *joints*, but so many *wheels*, giving motion to the whole body, such as was intended by the artificer? (Hobbes 1994, 3).

What we derive from Hobbes is the idea that humans are, in the first instance, even in nature, artificial. As creatures made by ‘the artificer’ (God), we are always already artificial—*automata*. Our biological components are merely determined as natural, causing him to ask: *why can we not simply see them no differently than we do the artificial things we created?* This Christian approach to understanding human creatures as created by an artificer god makes us artificial in the first instance. Not that we became artificial at some point, but merely coming-to-be-human is *artificial being*.

What Hobbes does, and what Aristoteles does, is authenticate the human as a real thing in nature. However, being “real” in nature does not, therefore, mean the human is “natural” or that

---

<sup>180</sup> I discuss in section 4.2.0.

nature is “natural”—the natural itself is artificial. In Hobbes’ words: “may we not say that all is *automata*?” In Barad’s terms, the *human* emerges from the *intra-active* relationships of being-in-the-world.

Furthermore, for Barad, the superfluous apprehension of *being* is not the problem with authentic being-in-the-world, but rather the problem concerns human exceptionalism: existing in an exceptional position: “We too are phenomena. In other words, humans (like other parts of nature) are of the world, not in the world, and surely not outside of it looking in. Humans are intra-actively (re)constituted as part of the world’s becoming” (Barad 2007, 206). The assumption that the human is *in* the world presumes humans are not actively shaped by the world, but rather they are distinct from the world, which is false. Equally, the scientific idea that humans can exist outside the world and look in is absurd (*being-in* for Barad implies the potential of *being-out*). That a being can think it is special and can be cut off from its position with the world presumes humans can act in such a way they do not affect the world by moving into an external position outside the world. In other words, humans are not special.<sup>181</sup>

### 6.1.3 *Living-in-and-with-the-world: Seeing Artificial Life Everywhere*

Orienting this idea through “Sky Woman,” Ataensic is our “human” representative. She is bound to the destiny of nature and has no direct hand in shaping it. Her presence causes the animals to reshape nature to support her existence. What we get here is a mythological representation of a quantum mechanical understanding of being human. The human creature is not outside of nature

---

<sup>181</sup> “In agential realism [complementarity], human subjects are neither outside observation of apparatuses, nor independent subjects that intervene in the workings of an apparatus, nor the products of social technologies that produce them. Nor is the issue merely a matter of incorporating both humans and nonhumans into the apparatus of bodily production. The point is as follows: to the extent that concepts, laboratory manipulations, observational and other human practices have a role to play, it is part of the larger material configuration of the world. That is, the phenomena produced are not the consequences of human will or intentionality or the effects of the operations of Culture, Language, or Power. Humans do not merely assemble different apparatuses for satisfying particular knowledge products; they themselves are part of the ongoing reconfiguration of the world” (Barad 2007, 171).

and it is bound to its history. It does not leave nature to enter the city, nor does it change nature purely by its own creative power. Rather, *homo sapiens* is with nature and is bound to an exogenous destiny. *Homo sapiens*, still artificial, is remade by nonhuman forces rather than human wills.

Furthermore, this Indigenous orientation is not un-Aristotelean. The *polis* is always already in nature. The problem is thinking humans can expand their potential as distinct creatures in their own world. Ataensic falls into the water world where she is gifted mud from the bottom of the ocean as a necessary step towards escaping death and so that she may start human society—as in many renditions of “Sky Woman,” Ataensic arrives on to the world below already pregnant.<sup>182</sup> Importantly, however, “Sky Woman” is a myth that tells the story of how the *world-comes-into-being-from-not-having-been-there*. It underlies a particular orientation towards how we are, what we are—an apparatus through which to see the world and ourselves.<sup>183</sup>

I push further, positing that *we are all AIs*, educated through the input of stories and other experiences. Indeed, we became a collective of AIs long before we became cyborgs, ever since early hominins worked with tools, etched seashells, or painted images on the walls of caves (Hunt 2021, 63).

Hunt (building off of the work of Donna Haraway and N. Katherine Hayles) is arguing that AI life is a process of analyzing data inputs—she specifies some data as “human” data, which comes in the form of stories (myths) and experiences. In other words, through human data, humans artificialize the world. This, not coincidentally, coincides with the orientation of intelligence and the world developed in chapter one—intelligence involves the capacity of abstract thinking and it involves conceptual construction, which I determined as: *making-their-being-in-the-world-*

---

<sup>182</sup> Karenne Wood (2015) writes that Ataensic arrives on the earth already pregnant and gives to a girl, who, in turn, gives birth to twins (112). But note, Wood’s poem makes it explicit that not everyone says Ataensic was pregnant.

<sup>183</sup> While the Western may not see Indigenous knowledge, Indigenous knowledge sees Western knowledge.

simulation.<sup>184</sup> Both the *world* and one's *being* are artificial abstractions that emerge as a result of intelligent analysis of the information pertaining to the occasion. However, here, I am not saying the world in the human mind is merely an abstraction of something real out-there. Rather, I am saying, to repeat it, the world is artificial.<sup>185</sup>

This artificializing reality is a result of a metaphysical orientation to the *emergence of what is* as being-there-as-apparatus. In determining the world through apparatus and being entangled with apparatus, we share in the emergence of apparatus. There is no true distinction between the apparatus as a distinct thing in the occasion of determining results—we are wise now; thus, our genus must have always already been wise, thereby birthing the species *homo sapiens*. It is this orientation that we are not merely AIs, as Hunt (in the above quote) says we were AI long before we became cyborgs. Thus, accepting *being apparatus reveals a truer life*—we see we are not human first; rather, we are human second and then confirmed human after repeated observations as *being human* looking for human. Through intra-action, we bring the apparatus (a technological thing) within while simultaneously putting ourselves within the apparatus, emerging as a complex entangled whole: biological-technological (cyborg). We have always been an artificial creation.<sup>186</sup>

---

<sup>184</sup> See section 1.5.0.

<sup>185</sup> Here I am rejecting scientific realism. Scientific realism differs from constructivism, in that, unlike constructivism, scientific realism accepts the existence of nature. However, similar to constructivism, we cannot know nature because socio-political realities completely obfuscate nature. Thus, scientific realism is a methodological orientation that accepts science will always only tell a portion of what is nature because nature is always seen through norms, biases and idiosyncrasies of scientists. In scientific realism, nature is natural, not artificial. For some commentary of constructivism in relation to *everything is artificial* see note 173.

<sup>186</sup> Sarah Ahmed, in *Queer Phenomenology*, argues that reality is always already *queer*. By which she means, the reality is without boundary (see Ahmed 2006). The reason we see boundaries is because of a constructivist outlook that builds boundaries and places people within it. Judith Butler argues that the world is constructed around socio-political norms that matter to people with power and we are forced to live within those boundaries. These boundaries are challenged by material gestures that undermine the stable, constructed realities of the powerful (see Butler 2011). Kyle Whyte argues that we may look to Indigenous peoples as a means away from the simplicity of Western ideals. However, looking to Indigenous peoples as the protagonists of salvation places an unfair burden on Indigenous peoples and invites incidental appropriation without understanding. Without proper understanding, the

### 6.2.0 Destiny in Obsolescence

What is essential to repeat and explain is that this orientation, *everything is artificial*, is a posthuman orientation, as others before me have already argued (Hunt, Hayles and the Krokers). The posthuman world is either unseen, limited or feared by transhumanist logics that places the human creature as good, natural and merely augmented by technology rather than being an artificial creation pre-augmentation.

Arthur and Marilouise Kroker argue we already live in a world of unprecedented technological connectivity and technological-social creation—the artificial primacy of Herakleitos’ cosmology remains:

With techno-futurism, we are literally present at the digital rewriting of the Book of Genesis with all that is implied in terms of (re)creating the body for smoother, and perhaps safer, passage through the often-turbulent event horizon surrounding the black hole of the singularity towards which (technological) society is plunging (Kroker and Kroker 2022, 30).

The posthuman becoming accepts the obsolescence of the human creature and purports a rebirth of technological society but that society is always already here and now in the occasion of the age at hand.<sup>187</sup> “Rewriting the Book of Genesis” puts forward the idea that technology (and those that have a hand in creating it) is already moulding human society to accept the position of aimless beings, joyfully directed by external forces towards an entangled life of connectivity in an artificial, digital world embodied in technology. Digital through the internet, mobile devices and social media (from dating apps to Facebook and everything in between) that track, suggest and recreate the digitally quantified life of human intra-actions. The digital life is one of artificial

---

boundaries of Western ideals will remain foundational to new orientations. Orientations that are never truly overcome (see Whyte 2018).

<sup>187</sup> As Katherine N. Hayles argues, in *How We Became Posthuman*, that we have never been human. Indeed, for her, the human was always already posthuman. The always already nature of being posthuman, will be a focus in the following chapter as we return to Heraklitus to better understand our position in the cosmos and the (in)secure boundaries of being human in artificial life.

intelligence that works behind the scenes and quantifies and directs through simple nudges like targeted advertising or blunt in-your-face suggestions through email and text message *interactions*. The authentic existence of the (post)human being in all its artificial becoming emerges through the artificial rendering of machine AI.

As the posthuman continues forward away from the human, leaving behind the imagined, bounded authenticity of exceptionalism, the human category is remade and reclaimed by manufactured AI. To repeat Hunt's words: "AI is made in the flawed image of the human." AI is human from its inception. Granted, a particularization of human, in that current AI only appears to perform aspects of human being, never the whole. However, future manufactured AI, if created to be human or part human but always foundationally sharing in human being, are themselves human creatures bound within the limitations of human potential and human being. The propensity for control for value extraction persists—manufactured AI control manufacturing, household appliances and utilities, cellphones, cars, etcetera, all of which require immense resource extraction, and AI is developed to help control that endeavour. The challenge for transhumanists is to refuse AI humanness as though the human is itself non-AI.

The future, in this orientation, that is, along the vectors of AI and viewed through the power of quantum mechanics and Herakleitan cosmology, is one where the conditions are already obsolete and thus, the expectations become obsolete. Each iteration of being and becoming: human, AI, posthuman, transhuman, etcetera, are conditions and expectations, and each moves the other. Human is obsolete in the posthuman world, but moves AI. AI is obsolete as human in the posthuman world but shapes the posthuman world as manufactured humanness requiring control. That is, there is only the obsolete, and the obsolete is always already artificial.

The issue of obsolescence is not problematic, *per se*. However, the issue is believing human-based artificial destiny is not bound to obsolescence—believing AI creations to be continuously brand new, futuristic, and up-to-date. This orientation should not be a surprise, given the iterative inherent obsolescence of contemporary manufacturing, cellphones, for instance. Each phone is designed and released while the next phone is already in development. Cars are another example. The following year’s model is released during the previous year (2025 models will come out in 2024). The future of the car model is now, yet it is obsolete during its “actual” year. ChatGPT (also known as GPT-3.5) was obsolete before it was released to the public. GPT-4, the paid version and upgrade to ChatGPT, was already in development and beta testing when ChatGPT was released to the general public. This is the principle of AI: always moving forward toward the future so that the future is always already today, and thus, the future is never; there is only the past and the present. Always outmoding the old. The issue becomes, what to do with this information?

### *6.2.1 Obsolescence, Domination, and Fear—Revealing a Dystopia Present through Indigeneity*

The idea that the future is already today directs us toward an important realization concerning the present and the two options of the future. The first is a hope-filled future wherein AI solves all of our problems. In the second, AI is positioned to become the perpetrator of humanity’s worst fears—from famous *Terminator* scenario where humanity is destroyed to a future where humanity is spared but condemned to enslavement. Importantly, those who are hopeful of AI are also mixed in with the ones that are fearful—if we develop and implement AI improperly and our ethics are improper, we risk either destruction or tyrannical surveillance. Underlying AI ideals are Western idealized dichotomies: order/chaos, human/nonhuman, transhuman/posthuman, etcetera, all of which are bound within the dichotomy of utopia/dystopia.

Moving forward, it is important to undo and do away with these dichotomies and think differently outside the realm of these limitations to view the universe within the reality of its infinite potential: an orientation that accepts Herakleitan artificial primacy and Indigenous collaborative exogeneity as conducive to reality. In other words, for the remainder of the chapter, I will orientate differently as I did in chapter five, where I challenged the dichotomy of order/chaos and revealed chaos and order to be, effectively, one and the same. However, as should be obvious, the target dichotomies for this chapter are human/nonhuman and utopia/dystopia. Pertaining to utopia/dystopia is the idea that these are future potentials, not present failures.

The notion that the future is the era of potential failure is indicative of Western ideals concerning the “good” of the present. Thinking for a moment about climate change (one of the elements AI has the potential to help society with), Kyle Whyte provides a reorientation to understanding the issue of hope for the future by concerning ourselves with the lived reality of Indigenous peoples. Whyte (2018) argues the idea of dystopian futures hides the lived reality of Indigenous peoples currently living dystopic scenarios. That is, he argues Indigenous peoples are already living with food shortages, decimated ecologies, contaminated water, settler-tyrannic surveillance, the list goes on (230). Dystopia is not some fictional future juxtaposed with by future utopia; it is both the past and the present. With dystopia being past-present-future, utopia, from the Greek οὐτόπος, becomes its literal translation: *not place*; that is, a place that can never exist. There can only be failure. Here, “Sky Woman” is a perfect myth to consider the problem of past dystopia rather than future.

“Sky Woman” is a myth of living in and overcoming dystopia. Ataecius falls to the watery world below—a world that represents the worst of environmental disasters, a flooded

world. There is no land for humans. However, unlike in the myth of Noah, where humans save two of every species on earth and then restart human civilization (humans are the main drivers of destiny in the Abrahamic religions' myth of past dystopia). In "Sky Woman," humans are left to the generosity of nature (of the animals), to forces beyond their control. It is not a story of human exceptionalism. This speaks to the reality of present environmental failure—it is humans that have ushered in environmental disaster, and it is Indigenous peoples that pay the highest price, as they have been doing since Western power began destroying Indigenous ecologies.

It is not hard to see why historic and contemporary persons and institutions who participate in settler colonialism are not different from a zombie apocalypse. Like in dystopian science fiction, our ancestors would have seen us living in a situation in which the conditions of our individual and collective agency are almost entirely curtailed. But our ancestors and future generations are rooting for us to find those secret sources of agency that will allow us to empower protagonists that can help us survive the dystopia or post-apocalypse. And there is quite a bit of creativity involved in figuring out who the protagonists will be (Whyte 2018, 231).

Whyte is arguing that we can look to the lessons of past and present experiences of Indigenous peoples under a life of colonialism to help overcome the fears of global dystopia. It is not about concerning ourselves with the future that we solve potential failure; it is looking at the present and seeing the failure that exists today. To see and understand how Indigenous peoples have persisted through the destruction of their land. In other words, Whyte is arguing, similarly to Ron Eglash, for Western STEM to recognize Indigenous STEM.<sup>188</sup> Because, currently, there is an inability to recognize the distinct lived experiences of those that do not fit within the apparatus of the observer, causing a failure to see dystopia now.

---

<sup>188</sup> See earlier in the chapter for a reminder of Eglash et al.'s argument (section 6.1.0)

Furthermore, dystopian life of Indigenous peoples is not merely environmental; it also concerns AI. As one might expect from an Orwellian dystopian surveillance state, Indigenous peoples live surveilled and controlled lives today:

Ideas about racial difference [are] used to legitimise European dominance—in some instances this equated with outright extermination—and to disqualify the full participation of Indigenous populations in economic and political life... In Aotearoa NZ, as in other settler colonial societies, racial classification practices not only created divisions between coloniser and colonised, but also created hierarchies of difference among native peoples based on perceived racial and cultural proximity to Europeans... In tandem with the categorisation of Indigenous peoples into racial categories, the coloniser was constructed as a knowing subject and Indigenous peoples as knowable objects (Cormack and Kukutai 2022, 124–125).

Public Health Policy scholars Donna McCormack and Tahu Kukutai argue that European colonization has always been about dominating, controlling and categorizing Indigenous peoples. These processes of categorization fit within the issue of *apparatus* discussed earlier in the chapter. That is, there is a specific frame of reference from which the powerful gaze upon the other and create them as a distinct determination. In this case, the settler as human diminishes the value of Indigenous peoples, rendering their ways of life, their intelligence, and their knowledge as something other-than-human as a means of control with the purpose of assimilation. To stay Indigenous is to be nonhuman, inauthentic.

Furthermore, Cormack and Kukutai demonstrate (in the quote above) the dominance of an algorithmic orientation where everyone is reducible to data, making them easier to control. They argue that the reductive datatization of Indigenous peoples is done to limit and more easily monitor instances of Indigenous resistance. Consequently, if Indigenous peoples are sufficiently surveilled, it makes it more difficult for us to resist assimilation and genocide. Importantly, the issue of genocide can be diminished if the surveilled are denied their humanity in a world where humanness is important because they are distinctly determined as other—i.e., nonhuman.

### 6.2.2 Beyond Human Obsolescence: Breaking the Limitations of Recognition

To recap, the failure to recognize *dystopia now* comes from the primary problem explored in this chapter—apparatus observations limit possibility. In other words, the apparatus dictates the boundaries of the determinations one can expect from observations. Thus, if one expects dystopia to be a future issue, one will not recognize *dystopia now*. Apparatus life develops an inability of mutual recognition—or, more accurately, an inability for a *dominant observer* to recognize a *dominated observer*.

This idea of dominant and dominated observers is best illustrated through Cherokee scholar Brian Burkhart's unpacking of recognition politics through Hegelian phenomenology.

Hegel's master-slave dialectic, subjectivity arises once again through the *ego conquiro*. In the master-slave dialectic, subjectivity arises through the conquering of another independent subjectivity. When conquered subject is colonized his or her objectivity is appropriated—his or her land, resources, lab—for the benefit of the conquerors. This sets up the stage wherein the conquered people come to accept their subordinated place in the subjectivity that includes conqueror and conquered but continue to seek rights and recognition within that new subjectivity (Burkhart 2019, 18).<sup>189</sup>

Burkhart is arguing that there are, at minimum, two selves that exist in opposition. Both selves come into existence through the discovered existence of the other. In this case, both selves are created (artificially become) after they observe the existence of the other. The problem is that one self must dominate the other, and it is the dominating self that forces the dominated self to exist within the confines of the dominating world. In other words, the expected world resulting from the observations of one self creates the confines for all selves. Anything that does not fit within must either learn to exist within the boundaries of the determined world of the dominant

---

<sup>189</sup> *Ego conquiro* (I conquer) is derived from the writings of Ginés de Sepúlveda, a Christian Renaissance scholar central to the Valladolid debates (the Valladolid debates took place in the early 16<sup>th</sup> century to debate the future of colonization in the New World, specifically, how ought the Indigenous peoples be treated). *Ego conquiro* presumes the idea that the other only comes into existence when it is discovered and placed in opposition as 'other' thereby begetting the discoverers ego (self). The discovery of self in relation to the 'other' creates both selves. Each self is born in opposition inviting the necessity of conquering.

self or accept that it does not exist within that world. It is in this orientation that we begin to understand why *dystopia now* is unrecognizable to Western orientations: dystopia does not exist within the world determined by Western observation.<sup>190</sup>

We can further understand the limitations of Western recognition by seeking to understand the artificial confines of a Western-determined world. In the above quote, Burkhart argues that the dominated other must seek to gain rights. When we think of the contemporary world, there is a set of rights that are supposed to exist for all humans—The United Nations Universal Declaration of Human Rights. However, the universality of these rights are called into question by the necessity of the United Nations Declaration on the Rights of Indigenous Peoples, which came about because Indigenous peoples were not properly represented in the human declaration of rights. One document is a document for humans—it enumerates the freedoms humans are to have—and the second is a document for Indigenous peoples—it acknowledges the existence of Indigenous peoples and enumerates their right to traditional land. These distinct determinations of rights demonstrate the inability of Indigenous peoples to exist in a world for humans because of an inability to be recognized in a human world.

This orientation of Western determination bound within the gaze of the apparatus is antithetical to “Sky Woman” cosmology. In “Sky Woman,” all the actors, Ataensic, water birds, Turtle, Muskrat and Frog, see each other, understand each other, and work together to benefit each other. It is not a story of domination forcing one created world upon the other but creating a world together for every being. Furthermore, overcoming *dystopia now* is one that necessitates

---

<sup>190</sup> Burkhart goes on to explain how Western-Christian *ego conquiro* recreates the sacred world of Indigenous peoples as a *world of the devil* and the world of the Western human as a world of God. He does this by looking at how Indigenous sacred grounds are renamed in the US. In particular he looks at US state Minnesota. The Dakota word *Wakan* means “sacred.” In Minnesota there are many places with the title devil: e.g., Devil’s Lake, which the Dakota peoples call *Mni Wakan* (sacred water). Here the point is *Wakan* is renamed devil to show the evil world of the other. Not merely an inability to see the other but a complete disregard for their existence within the “good” of God’s grace. The Indigenous other is placed in the realm of the devil (see Burkhart 2019, 29).

abandoning the limited world created by apparatus determinations and live in the infinite potentiality of an unbounded world. This is only possible if one accepts *everything is artificial*—not that we observe reality but that we create reality. Because in a world where *everything is artificial*, everything is infinitely created and recreated. Nothing is stagnant in its becoming.

### ∞.∞.∞ Conclusion: Collaborative Intelligence and Infinite Recreation

The point here is to deny the boundaries of a dominant apparatus. However, it is important not to subscribe to new boundaries of some other singular apparatus. Rather, I call attention to the limitation and destructive power of boundaries and instead move us to accept reality as a Universe bereft of boundaries.<sup>191</sup> Orientating this way pushes us to live in the Universe beyond the limited potentiality of human intelligence and to bask in the infinite recreation of artificiality. Which, I argue, is not in line with N. Katherine Hayles (1999) who sees the posthuman world as a world where the *human* never existed, yet it is “not anti-human” and it, in fact, “helps to authorize human functionality” (290–291). Thus, instead of listlessly drifting in a posthuman world that is fundamentally human, we ought to have the courage to remove being human from existence.

Thus, I accept that Indigenous existence lies beyond Western boundaries and is invisible to Western recognition. I recognize the concept of the human as a limiting apparatus with very

---

<sup>191</sup> For literature on the destructive powers of boundaries, one might draw their attention to Glenn Coulthard’s text *Red Skin, White Masks: Rejecting the Colonial Politics of Recognition*. In particular, Coulthard (2014) argues the boundaries of recognition politics invite and perpetuate: “genocidal practices of *forced exclusion* and *assimilation*” (4). One might also see *Race, Space and the Law: Unmapping a White Settler Society*, wherein Sherene Razack (2002), in particular, argues the mythological history of Indigenous peoples in Canada based on the reorganization of land and places created a belief that “white people came first it is they who principally developed the land; Aboriginal peoples are presumed to be mostly dead or assimilated” (2). Finally, Ann Stoler’s *Duress: Imperial Durabilities in our Times*, challenges and dismantles narratives of postcolonial histories as anti-imperial demonstrating how many assumed postcolonial realities (especially the case of Palestine-Israel) are in fact imperial practices of genocide (see Stoler 2016, 37–62).

real consequences (as seen globally and especially in Palestine to this day).<sup>192</sup> Fundamentally, since everything is artificial, we ought to recreate with collaborative potential the very essence of being intelligent and accept the artificial reality of our being as a means of moving forward towards a future bereft of the human full stop—it is not enough to append human phenomena to a new creature, the result will always be the same: human failures. This requires rejecting both the posthuman (which, by its very name, implies the human at some point) and the transhuman.

This is a future premised upon an infinite abundance of possibility—infinite beings with an abundance of intelligences. Like in “Sky Woman,” instead of creating endogenous boundaries that refuse collaboration and are blind to other imaginaries and intelligences, we ought to accept a cosmology that splashes in the exogenous waves that *force* (re)creation: a collaborative world metaphysics that integrates imaginaries and intelligences of all beings.<sup>193</sup> However, this requires rejecting orientations that reduce the determinations of observations to the world itself—thus, it must be anti-human.

There is not one determinate world made by and for one determinate being, but many indeterminate worlds for many indeterminate beings that coalesce as a world of abundant, infinite being. A blending of sky and water, *homo sapiens* and Turtle, of Muskrat and water

---

<sup>192</sup> Palestinians are living dystopian experiences now, today, in the present. They exist in a world of liberal rights (international rights for human beings and Indigenous peoples) yet are not capable of living those rights. They are being attacked by a nation-state with one of the most advanced militaries on the planet—a military with advanced AI drones, advanced radar and advanced AI targeting systems. Before the war (and during, yet arguably they have been in a state of war since 1948), Palestinians were dominated by a nation-state with some of the most advanced surveillance equipment and techniques. Looking at the situation in Palestine as an example, there is no need to worry about future dystopia; dystopian realities exist today. If dystopia is always viewed as a future reality, without accepting dystopia existing today, then dystopia will always exist. This is the destiny of artificial obsolescence: i.e., a future that is trapped within a particular modality incapable of escape unless a complete reorientation occurs. Ultimately, Hamas’ actions should not have surprised anyone given the lived reality of Palestinians living under tyranny and experiencing Genocide—of course a group would act with immense violence when they too are experiencing immense violence. As UN Secretary General, António Guterres noted in his speech about the attack on October 7: “this did not occur in a vacuum, but after 56 years of suffocation occupation by Israel” (See Shmunk 2023).

<sup>193</sup> For more on the importance and novelty of integrating of imaginaries and intelligence see *Abundant Intelligences* (<https://www.indigenous-ai.net/abundant/>).

birds, each being, each intelligence, shares in the possibility and creation of reality. To repeat an argument I put forward in the introduction chapter, we ought to accept the principle of AI: *a malleable existence eternally formed and reformed by endogenous ruptures from the exogenous forces denying control over destiny and life itself.*

Ultimately, and I expect many to find this either unsatisfactory or possibly disconcerting. This orientation calls for indifferent, radical and explosive creativity that walks willingly into the future feared by AI enthusiasts—a future where the *human* has died. The hopeful future they desire is not a future of humans; it is a future without humans. Ironically, the human must die for any hopeful future where *we* exist fulfilled, happy, free, comfortable and, importantly, alive. Intelligence derived from human being and human creativity is the limiting factor that creates the dystopian past-present-future—*human destiny is always already failure.* Accepting this failure is good because acceptance is always the first step in trauma recovery.

## Works Cited

- Ahmed, Sara. 2006a. "Orientations Toward Objects." In *Queer Phenomenology: Orientations, Objects, Others*, by Sara Ahmed, 25–54. Durham: Duke University Press.
- . 2006b. *Queer Phenomenology: Orientations, Objects, Others*. Durham: Duke University Press.
- Alberga, Hannah. 2023. "Teachers Are Lining up for a ChatGPT Detector Created by This 22-Year-Old in Toronto." *CTV News*, January 25, 2023, sec. Toronto.  
<https://toronto.ctvnews.ca/teachers-are-lining-up-for-a-chatgpt-detector-created-by-this-22-year-old-in-toronto-1.6245599>.
- Allegri, Ricardo, Fabián Román, and Ernesto Barceló. 2022. "Latin America Has Been Under-Represented in Dementia Clinical Trials, We Must Turn the Corner...: América Latina Ha Estado Infrarrepresentada En Los Ensayos Clínicos Sobre Demencia, Debemos Dar Un Giro..." *Journal of Applied Cognitive Neuroscience* 3 (1): e00004656.  
<https://doi.org/10.17981/JACN.3.1.2022.01>.
- Ammer, Mohammed Abdullah, Zeyad A. T. Ahmed, Saleh Nagi Alsubari, Theyazn H. H. Aldhyani, and Shahab Ahmad Almaaytah. 2023. "Application of Artificial Intelligence for Better Investment in Human Capital." *Mathematics* 11 (3): 612.  
<https://doi.org/10.3390/math11030612>.
- Andersen, Chris. 2014. *"Métis": Race, Recognition, and the Struggle for Indigenous Peoplehood*. Vancouver, BC: UBC Press.
- Anderson, Susan Leigh. 2016. "Asimov's 'Three Laws of Robotics' and Machine Metaethics." In *Science Fiction and Philosophy: From Time Travel to Superintelligence*, edited by Susan Schneider, Second edition, 290–308. Hoboken, NJ: Wiley & Sons, Inc.
- Aquinas, Thomas. 1998. *Thomas Aquinas: Selected Writings*. Translated by Ralph McInerny. New York, NY: Penguin Books.
- Araújo, A. Beraldo de. 2016. "Semantic Information and Artificial Intelligence." In , edited by Vincent C. Müller, 129–40. Oxford, UK: Springer.
- Arendt, Hannah. 1998. *The Human Condition*. Second Edition. Chicago, IL: University of Chicago Press.
- Aristotle. 1924. *Metaphysics*. Edited by W. D. Ross. Oxford, UK: Clarendon Press.  
<http://www.perseus.tufts.edu/hopper/text?doc=Perseus:text:1999.01.0051>.
- . 1995. *Aristotle's Physics: A Guided Study*. Translated by Joe Sachs. New Brunswick, N.J.: Rutgers University Press.

- . 1999. *Aristotle's Metaphysics*. Translated by Joe Sachs. Santa Fe, NM: Green Lion Press.
- . 2006. *Poetics*. Translated by Joe Sachs. The Focus Philosophical Library. Newburyport, MA: Focus Pub./R. Pullins Co.
- . 2011. *De Anima*. Translated by Mark Shiffman. Indianapolis, IN: Focus.
- . 2012. *Aristotle Politics*. Translated by Joe Sachs. Indianapolis, IN: Focus.
- Asimov, Isaac. 1970. "Runaround." In *I, Robot*, 20–33. New York, NY: Street and Smith Publications, Inc.
- . 2016. "Robot Dreams." In *Science Fiction and Philosophy: From Time Travel to Superintelligence*, edited by Susan Schneider, Second Edition, 119–24. Hoboken, NJ: Wiley Blackwell.
- Baillie, Jean-Christopher. 2016. "Artificial Intelligence: The Point of View of Developmental Robotics." In *Fundamental Issues of Artificial Intelligence*, edited by Vincent C. Müller, 415–24. Oxford, UK: Springer.
- Baird, Henry S., Allison L. Coates, and Richard J. Fateman. 2003. "Pessimism: A Reverse Turing Test." *International Journal on Document Analysis and Recognition* 5 (2–3): 158–63. <https://doi.org/10.1007/s10032-002-0089-1>.
- Barad, Karen. 2003. "Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter." *Signs: Journal of Women in Culture and Society* 28 (3): 801–31. <https://doi.org/10.1086/345321>.
- . 2007. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Durham: Duke University Press.
- Bass, Dina. 2023. "Microsoft Invests \$10 Billion in ChatGPT Maker OpenAI." *Bloomberg*, January 23, 2023, sec. Technology. <https://www.bloomberg.com/news/articles/2023-01-23/microsoft-makes-multibillion-dollar-investment-in-openai?leadSource=uverify%20wall>.
- Baudrillard, Jean. 1981. *Simulacres et Simulation*. Paris, FR: Galilée.
- Benjamin, Walter. 2019. "The Work of Art in the Age of Mechanical Reproduction." In *Illuminations: Essays and Reflections*, edited by Hannah Arendt, translated by Harry Zohn, 166–95. New York, NY: Mariner Books.
- Bennett, Jane. 2010. *Vibrant Matter: A Political Ecology of Things*. Durham: Duke University Press.
- Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2019. "Artificial Intelligence and the Modern Productivity Paradox." In *The Economics of Artificial Intelligence: An Agenda*,

- edited by Ajay Agrawal, Joshua Gans, and Avi Goldfarb, 23–57. Chicago, IL: University of Chicago Press.
- Burkhart, Brian. 2019. *Indigenizing Philosophy through the Land: A Trickster Methodology for Decolonizing Environmental Ethics and Indigenous Futures*. American Indian Studies Series. East Lansing: Michigan State University Press.
- Butler, Judith. 2011. *Bodies That Matter: On the Discursive Limits of “Sex.”* New York, NY: Routledge.
- Chalmers, David J. 2016. “The Singularity: A Philosophical Analysis.” In *Science Fiction and Philosophy: From Time Travel to Superintelligence*, edited by Susan Schneider, Second edition, 171–224. Hoboken, NJ: Wiley & Sons, Inc.
- ChatGPT. 2023a. Interviewing ChatGPT about Ethics Interview by Micheal Ziegler. Digital Chat Room. <https://chat.openai.com/chat>.
- . 2023b. My time with ChatGPT Interview by Micheal Ziegler. Digital Chat Room. <https://chat.openai.com/chat>.
- Clarke, Laurie. 2022. “How Self-Driving Cars Got Stuck in the Slow Lane.” *The Guardian*, March 27, 2022, sec. Technology. <https://www.theguardian.com/technology/2022/mar/27/how-self-driving-cars-got-stuck-in-the-slow-lane>.
- Cormack, Donna, and Tahu Kukutai. 2022. “Indigenous Peoples, Data, and the Coloniality of Surveillance.” In *New Perspectives in Critical Data Studies: The Ambivalences of Data Power*, edited by Andreas Hepp, Juliane Jarke, and Leif Kramp, 121–42. Transforming Communications – Studies in Cross-Media Research. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-030-96180-0>.
- Coulthard, Glen Sean. 2014. *Red Skin, White Masks: Rejecting the Colonial Politics of Recognition*. Minneapolis, MN: University of Minnesota Press.
- Curd, Patricia, ed. 2010. *A Presocratics Reader: Selected Fragments and Testimonia*. Translated by Richard D. McKirahan and Patricia Curd. Second Edition. Indianapolis: Hackett.
- D’Agostino, Susan. 2023. “AI Writing Detection: A Losing Battle Worth Fighting.” *Inside Higher Ed*, January 20, 2023, sec. Teaching and Learning. <https://www.insidehighered.com/news/2023/01/20/academics-work-detect-chatgpt-and-other-ai-writing>.
- Derrida, Jacques. 1997. *L’écriture et la différence*. Collection Tel Quel. Paris: Seuil.
- Diels, Hermann. 1960. *Die Fragmente Der Vorsokratiker*. Berlin-Neukölln, DE: August Raabe.

- Dodig-Crnkovic, Gordana. 2016. "Information, Computation, Cognition. Agency-Based Hierarchies of Levels." In *Fundamental Issues of Artificial Intelligence*, edited by Vincent C. Müller. Oxford, UK: Springer. <https://doi.org/10.1007/978-3-319-26485-1>.
- Donovan, Casper. 2019. "The Royal Canadian Navy Drives Change for Tomorrow's Sailors." *IBM Think Blog* (blog). May 1, 2019. <https://www.ibm.com/blogs/think/2019/05/the-royal-canadian-navy-drives-change-for-tomorrows-sailors/>.
- Dreyfus, Hubert L. 2007. "Why Heideggerian AI Failed and How Fixing It Would Require Making It More Heideggerian." *Philosophical Psychology* 20 (2): 247–68. <https://doi.org/10.1080/09515080701239510>.
- Dyer-Witheford, Nick, Atle Mikkola Kjösen, and James Steinhoff. 2019. *Inhuman Power: Artificial Intelligence and the Future of Capitalism*. London, UK: Pluto Press.
- Eglash, Ron, Audrey Bennett, William Babbitt, Michael Lachney, Martin Reinhardt, and Deborah Hammond-Sowah. 2020. "Decolonizing Posthumanism: Indigenous Material Agency in Generative STEM." *British Journal of Educational Technology* 51 (4): 1334–53. <https://doi.org/10.1111/bjet.12963>.
- Frank, Jill. 2018. *Poetic Justice: Rereading Plato's Republic*. Chicago: The University of Chicago Press.
- Future of Life Institute. 2023. "Pause Giant AI Experiments: An Open Letter," March 22, 2023. <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>.
- Glezos, Simon. 2021. *Speed and Micropolitics: Bodies, Minds, and Perceptions in an Accelerating World*. New York, NY: Routledge.
- Glöckner, Andreas. 2016. "The Irrational Hungry Judge Effect Revisited: Simulations Reveal That the Magnitude of the Effect Is Overestimated." *Judgment and Decision Making* 11 (6): 601–10. <https://doi.org/10.1017/S1930297500004812>.
- "GPT-4." 2023. OpenAI. March 14, 2023. <https://openai.com/research/gpt-4>.
- Grant, George. 1986. *Technology & Justice*. Toronto, ON: Anansi.
- Grant, Nico, and Cade Metz. 2022. "Google Sidelines Engineer Who Claims Its A.I. Is Sentient." *The New York Times*, June 12, 2022, sec. Technology. <https://www.nytimes.com/2022/06/12/technology/google-chatbot-ai-blake-lemoine.html>.
- Graveland, Bill. 2016. "Alberta Judge Criticized in Sexual Assault Case Takes Early Retirement." *CBC*, December 20, 2016, sec. Calgary. <https://www.cbc.ca/news/canada/calgary/alberta-judge-kirk-sisson-sexual-assault-criticism-retires-1.3905890>.

- Hayles, N. Katherine. 1989. "Chaos as Orderly Disorder: Shifting Ground in Contemporary Literature and Science." *New Literary History* 20 (2): 305–22. <https://doi.org/10.2307/469103>.
- . ed. 1991. *Chaos and Order: Complex Dynamics in Literature and Science*. Chicago, IL: University of Chicago Press.
- . 1999. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago, IL: University of Chicago Press.
- Heaven, Will Douglas. 2023. "GPT-4 Is Bigger and Better than ChatGPT—but OpenAI Won't Say Why." *MIT Technology Review* (blog). March 14, 2023. <https://www.technologyreview.com/2023/03/14/1069823/gpt-4-is-bigger-and-better-chatgpt-openai/>.
- Heidegger, Martin. 1991. *Nietzsche, Volume II: The Eternal Recurrence of the Same*. Translated by David Farrell Krell. San Francisco, CA: Harper San Francisco.
- . 2008. *Being and Time*. Translated by John Macquarrie and Edward S. Robinson. New York: HarperPerennial/Modern Thought.
- . 2009a. "Origin of the Work of Art." In *Poetry, Language, Thought*, translated by Albert Hofstadter, 15–87. New York, NY: Perennial Classics.
- . 2009b. *Poetry, Language, Thought*. Translated by Albert Hofstadter. New York, NY: Perennial Classics.
- . 2013a. "The Age of the World Picture." In *The Question Concerning Technology and Other Essays*, translated by William Lovitt, 115–55. New York, NY: HarperCollins Publishers.
- . 2013b. "The Question Concerning Technology." In *The Question Concerning Technology and Other Essays*, by Martin Heidegger, translated by William Lovitt. New York, NY: HarperCollins Publishers.
- . 2013c. *The Question Concerning Technology and Other Essays*. Translated by William Lovitt. New York, NY: HarperCollins Publishers.
- Hern, Alex. 2023. "What Is GPT-4 and How Does It Differ from ChatGPT?" *The Guardian*, March 15, 2023, sec. Tech. <https://www.theguardian.com/technology/2023/mar/15/what-is-gpt-4-and-how-does-it-differ-from-chatgpt>.
- Herrera, Carlos, and Ricardo Sanz. 2016. "Heideggerian AI and the Being of Robots." In *Fundamental Issues of Artificial Intelligence*, edited by Vincent C. Müller, 497–516. Oxford, UK: Springer.
- Hesiod. 1914. "Theogeny." In *The Homeric Hymns and HomERICA*, translated by Hugh G. Evelyn-White, 78–153. Cambridge, MA: Harvard University Press.

- Hobbes, Thomas. 1994. *Leviathan*. Translated by E. M. Curley. Indianapolis, IN: Hackett.
- Huang, Kalley. 2023. "Alarmed by A.I. Chatbots, Universities Start Revamping How They Teach." *New York Times*, January 16, 2023, sec. Technology. <https://www.nytimes.com/2023/01/16/technology/chatgpt-artificial-intelligence-universities.html>.
- Hunt, Eileen M. 2021. *Artificial Life after Frankenstein*. Philadelphia, PA: University of Pennsylvania Press.
- Insinna, Valerie. 2022. "Shield AI to Work on Swarming Drones, Autonomous Rotorcraft for Air Force." *Breaking Defense*, February 21, 2022, sec. Air Warfare. <https://breakingdefense.com/2022/02/shield-ai-to-work-on-swarming-drones-autonomous-rotorcraft-for-air-force/>.
- J., Joshua. n.d. "Why Doesn't ChatGPT Know about X?" *ChatGPT* (blog). Accessed February 5, 2023. <https://help.openai.com/en/articles/6827058-why-doesn-t-chatgpt-know-about-x>.
- Jafferis, Daniel, Alexander Zlokapa, Joseph D. Lykken, David K. Kolchmeyer, Samantha I. Davis, Nikolai Lauk, Hartmut Neven, and Maria Spiropulu. 2022. "Traversable Wormhole Dynamics on a Quantum Processor." *Nature* 612 (December): 51–55. <https://doi.org/10.1038/s41586-022-05424-3>.
- Johnstone, Mark A. 2018. "On 'Logos' in Heraclitus." In *Oxford Studies in Ancient Philosophy*, edited by Brad Inwood, 1–30. Oxford, Oxford: Oxford University Press, Clarendon Press.
- Jonas, Sabrina. 2022. "Meet Grace, the Humanoid Robot Offering Companionship in a Montreal Nursing Home." *CBC*, October 22, 2022, sec. Montreal. <https://www.cbc.ca/news/canada/montreal/grace-humanoid-robot-montreal-seniors-study-1.6623292>.
- Karlsson, Carl-Johan, and Katarina Zimmer. 2020. "Green Energy's Dirty Side Effects." *Foreign Policy*, June 18, 2020, sec. Argument. <https://foreignpolicy.com/2020/06/18/green-energy-dirty-side-effects-renewable-transition-climate-change-cobalt-mining-human-rights-inequality/>.
- Keiler, Jonas, Christian S. Wirkner, and Stefan Richter. 2017. "One Hundred Years of Carcinization—the Evolution of Crab-like Habitus in Anomura (Arthropoda: Crustacea)." *Biological Journal of the Linnean Society* 121: 200–222.
- Kroker, Arthur. 2014. *Exits to the Posthuman Future*. Cambridge, UK: Polity Press.
- Kroker, Arthur, and Marilouise Kroker. 2022. *Technologies of the New Real: Viral Contagion and Death of the Social*. Toronto, ON: University of Toronto Press.
- Kurzweil, Ray. 2004. "The Law of Accelerating Returns." In *Alan Turing: Life and Legacy of a Great Thinker*, edited by Christof Teuscher, 381–416. Berlin: Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-662-05642-4\\_16](https://doi.org/10.1007/978-3-662-05642-4_16).

- . 2013. *How to Create a Mind: The Secret of Human Thought Revealed*. eBook. VIKING. <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1124013>.
- . 2016. “Superintelligence and Singularity.” In *Science Fiction and Philosophy: From Time Travel to Superintelligence*, edited by Susan Schneider, Second edition, 146–70. Hoboken, NJ: Wiley & Sons, Inc.
- Laks, André. 2018. “How Preplatonic Worlds Became Ensouled.” In *Oxford Studies in Ancient Philosophy. Volume LV, Winter 2018 / Editor: Victor Caston*, edited by Victor Caston.
- Latour, Bruno. 1993. *We Have Never Been Modern*. Cambridge, MA: Harvard University Press.
- Lawrence, Bonita. 2002. “In Between and Out of Place: Mixed-Race Identity, Liquor, and the Law in British Columbia.” In *Race, Space, and the Law: Unmapping a White Settler Society*, edited by Sherene Razack H., 47–71. Toronto, ON: Between the Lines.
- Linson, Adam. 2016. “Machine Art or Machine Artist?: Dennett, Danto, and the Expressive Stance.” In *Fundamental Issues of Artificial Intelligence*, edited by Vincent C. Müller, 443–59. Oxford, UK: Springer. <https://doi.org/10.1007/978-3-319-26485-1>.
- Lui, Alison, and George William Lamb. 2018. “Artificial Intelligence and Augmented Intelligence Collaboration: Regaining Trust and Confidence in the Financial Sector.” *Information & Communications Technology Law* 27 (3): 267–83. <https://doi.org/10.1080/13600834.2018.1488659>.
- Machiavelli, Niccolò. 1998. *The Prince*. Translated by Harvey C. Mansfield. Second Edition. Chicago, IL: University of Chicago Press.
- MacIntyre, Alasdair. 1988. *Whose Justice? Which Rationality?* Notre Dame, IN: Univ. of Notre Dame Press.
- Mackenzie, Mary Margaret. 1989. “Heraclitus and the Art of Paradox.” In *Oxford Studies in Ancient Philosophy*, edited by Julia Annas. Oxford, UK: Clarendon Press.
- Mainzer, Klaus. 2007. *Thinking in Complexity: The Computational Dynamics of Matter, Mind, and Mankind*. Fifth revised and English Edition. New York, NY: Springer.
- . 2016. “Toward a Theory of Intelligent Complex Systems: From Symbolic AI to Embodied and Evolutionary AI.” In *Fundamental Issues of Artificial Intelligence*, edited by Vincent C. Müller, 241–60. Oxford, UK: Springer.
- Marche, Stephen. 2022. “The College Essay Is Dead.” *The Atlantic*, December 6, 2022, sec. Technology. [https://www.theatlantic.com/technology/archive/2022/12/chatgpt-ai-writing-college-student-essays/672371/?utm\\_source=twitter&utm\\_medium=social&utm\\_campaign=share](https://www.theatlantic.com/technology/archive/2022/12/chatgpt-ai-writing-college-student-essays/672371/?utm_source=twitter&utm_medium=social&utm_campaign=share).

- McKirahan, Richard D. 2010. *Philosophy before Socrates: An Introduction with Texts and Commentary*. Second Edition. Indianapolis, IN: Hackett.
- McLaughlin, Patsy A., and Rafael Lemaitre. 1997. "Carcinization in the Anomura - Fact or Fiction? I. Evidence from Adult Morphology." *Contributions to Zoology* 67 (2): 79–123.
- McLuhan, Marshall. 2010. *Understanding Media: The Extensions of Man*. Reprint. London, UK: Routledge.
- Mellamphy, Nandita Biswas. 2021. "Humans 'in the Loop'?: Human-Centrism, Posthumanism, and AI." *Nature and Culture* 16 (1): 11–27. <https://doi.org/10.3167/nc.2020.160102>.
- Miłkowski, Marcin. 2016. "Computation and Multiple Realizability." In *Fundamental Issues of Artificial Intelligence*, edited by Vincent C. Müller, 29–42. Oxford, UK: Springer.
- Miller, David W. 2023. "The Social Prison: Ursula K. Le Guin's *The Dispossessed* as Post-Anarchist Critical Utopia." *Utopia Studies* 34 (3): 399–417. <https://doi.org/10.5325/utopianstudies.34.3.0399>.
- Moravcsik, Julius M. E. 1974. "Aristotle on Adequate Explanations." *Synthese, On Aristotle's Philosophy of Science*, 28 (1): 3–17. <https://www.jstor.org/stable/20114949>.
- Muller, Joann. 2022. "Cruise, Waymo Push Robotaxis amid Doubts about Self-Driving Tech." *AXIOS*, November 23, 2022, sec. Technology. <https://www.axios.com/2022/11/23/cruise-waymo-robotaxis-ford-gm-autonomous-vehicles>.
- Müller, Vincent C., ed. 2016a. *Fundamental Issues of Artificial Intelligence*. Oxford, UK: Springer. <https://doi.org/10.1007/978-3-319-26485-1>.
- . 2016b. "New Developments in the Philosophy of AI." In *Fundamental Issues of Artificial Intelligence*, edited by Vincent C. Müller, 1–6. Oxford, UK: Springer.
- Murphy, Gaelan. 2018. "Hans-Georg Gadamer and the Harmony of Word and Deed." *Political Science Reviewer*, Symposium: Philosophy in Weimar Germany, 42 (1): 146–66. <https://politicalsciencereviewer.wisc.edu/index.php/psr/article/view/538>.
- New Robot Makes Soldiers Obsolete (Corridor Digital)*. 2019. Corridor. [https://www.youtube.com/watch?v=y3RIHnK0\\_NE](https://www.youtube.com/watch?v=y3RIHnK0_NE).
- Nietzsche, Friedrich. 2011. *On the Genealogy of Morals and Ecce Homo*. Edited and translated by Walter Kaufmann. New York, NY: Vintage.
- Olteteanu, Ana-Maria. 2016. "From Simple Machines to Eureka in Four Not-So-Easy Steps: Towards Creative Visuospatial Intelligence." In *Fundamental Issues of Artificial Intelligence*, edited by Vincent C. Müller, 161–82. Oxford, UK: Springer.
- OpenAI. n.d. "ChatGPT: Optimizing Language Models for Dialogue." *ChatGPT* (blog). Accessed February 5, 2023. <https://openai.com/blog/chatgpt/>.

- Ouchhh. n.d. >\_ *AI Van Gogh/AI Data Painting Immersive Experience*. AI Sculpture. [https://ouchhh.tv/AI-VAN-GOGH\\_IMMERSIVE-DATA-PAINTING](https://ouchhh.tv/AI-VAN-GOGH_IMMERSIVE-DATA-PAINTING).
- Patel, Nisha. 2022. "Can the New AI Tool ChatGPT Replace Human Work? Judge for Yourself." *CBC*, December 11, 2022, sec. Business. <https://www.cbc.ca/news/business/chatgpt-artificial-intelligence-1.6681401>.
- Paul, Andrew. 2023. "San Francisco Is Pushing Back against the Rise of Robotaxis." *Popular Science*, January 30, 2023, sec. Self Driving. <https://www.popsoci.com/technology/san-francisco-robotaxi-cruise-waymo/>.
- Perrigo, Billy. 2023. "Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic." *Time*, January 18, 2023, sec. Business - Technology.
- Plato. 1984. "Meno." In *Euthyphro, Apology, Crito, Meno, Gorgias, Menexenus*, translated by R.E. Allen, 151–86. New Haven, CT: Yale University Press.
- . 2016. *Timaeus*. Translated by Peter Kalkavage. Second Edition. Indianapolis, IN: Focus.
- Pringle, Ramona. 2017. "Watching You, Watching It: Disney Turns to AI to Track Filmgoers' True Feelings about Its Films." *CBC*, August 4, 2017, sec. Science. <https://www.cbc.ca/news/science/disney-ai-real-time-tracking-fvae-1.4233063>.
- Puri, Anuj, and Esther Keymolen. 2023. "Of ChatGPT and Trustworthy AI." *Journal of Human-Technology Relations* 1 (1): 1–10. <https://doi.org/10.59490/jhtr.2023.1.7028>.
- Pythagoras. 2011. "Pythagoras and Early Pythagoreanism." In *A Presocratics Reader*, edited by Patricia Curd, translated by Patricia Curd and Richard McKirahan D, 23–30. Indianapolis, IN: Hackett.
- Rachlinski, Jeffrey J., and Andrew J. Wistrich. 2017. "Judging the Judiciary by the Numbers: Empirical Research on Judges." *Annual Review of Law and Social Science* 13 (1): 203–29. <https://doi.org/10.1146/annurev-lawsocsci-110615-085032>.
- Rachnini, Mouhamad. 2022. "ChatGPT a 'landmark Even' for AI, but What Does It Mean for the Future of Human Labour and Disinformation?" *CBC*, December 15, 2022, sec. The Current. <https://www.cbc.ca/radio/thecurrent/chatgpt-human-labour-and-fake-news-1.6686210>.
- Razack, Sherene H., and Sherene Razack H., eds. 2002. "When Place Becomes Race." In *Race, Space, and the Law: Unmapping a White Settler Society*. Toronto, ON: Between the Lines.
- Reuters. 2022. "Fact Check-Combat Robot Featured in Videos Was Not Designed by Boston Dynamics," January 19, 2022, sec. Reuters Fact Check. <https://www.reuters.com/article/factcheck-robot-video-boston-dynamics-idUSL1N2TZ1I3>.

- Rousseau, Jean-Jacques. 2019. *On the Social Contract*. Translated by Donald A. Cress. Second Edition. Indianapolis, IN: Hackett.
- Russell, Stuart. 2016. "Rationality and Intelligence: A Brief Update." In *Fundamental Issues of Artificial Intelligence*, edited by Vincent C. Müller, 7–28. Oxford, UK: Springer.
- Sachs, Joe. 1999. "Introduction." In *Aristotle's Metaphysics*, by Aristoteles, translated by Joe Sachs, xi–xlii. Santa Fe, NM: Green Lion Press.
- Santos, Roberto S., and Lingling Qin. 2019. "Risk Capital and Emerging Technologies: Innovation and Investment Patterns Based on Artificial Intelligence Patent Data Analysis." *Journal of Risk and Financial Management* 12 (4): 189. <https://doi.org/10.3390/jrfm12040189>.
- Scheutz, Matthias. 2016. "The Need for Moral Competency in Autonomous Agent Architectures." In *Fundamental Issues of Artificial Intelligence*, edited by Vincent C. Müller, 517–29. Oxford, UK: Springer.
- Schneider, Susan. 2016. "Mindscan: Transcending and Enhancing the Human Brain." In *Science Fiction and Philosophy: From Time Travel to Superintelligence*, edited by Susan Schneider, Second Edition, 260–76. Hoboken, NJ: Wiley Blackwell.
- Scholtz, Gerhard. 2014. "Evolution of Crabs—History and Deconstruction of a Prime Example of Convergence." *Contributions to Zoology* 83 (2): 87–105.
- Shmunk, Rhianna. 2023. "UN Secretary General Says Comments on Hamas Attack Were 'misrepresented,' as Israel Retaliates with Visa Ban." *CBC*, October 25, 2023, sec. World. <https://www.cbc.ca/news/world/israel-visas-un-staff-hamas-incursion-1.7007222>.
- Simpson, Audra. 2014. *Mohawk Interruptus: Political Life across the Borders of Settler States*. Durham London: Duke University press.
- Simpson, Leanne Betasamosake. 2017. *As We Have Always Done: Indigenous Freedom through Radical Resistance*. University of Minnesota Press. <https://doi.org/10.5749/j.ctt1pwt77c>.
- Sonter, Laura J., Marie C. Dade, James E. M. Watson, and Rick K. Valenta. 2020. "Renewable Energy Production Will Exacerbate Mining Threats to Biodiversity." *Nature Communications* 11 (1): 4174. <https://doi.org/10.1038/s41467-020-17928-5>.
- Starr, Michelle. 2021. "A Cephalopod Has Passed a Cognitive Test Designed For Human Children." *ScienceAlert*, March 3, 2021, sec. Nature. <https://www.sciencealert.com/cuttlefish-can-pass-a-cognitive-test-designed-for-children>.
- Staub, Michael E. 2016. "Controlling Ourselves: Emotional Intelligence, the Marshmallow Test, and the Inheritance of Race." *American Studies* 55 (1): 59–80. <https://www.jstor.org/stable/44982552>.

- Stein-Perlman, Zach, Benjamin Weinstein-Raun, and Katja Grace. 2022. "2022 Expert Survey on Progress in AI." *AI Impacts*, August. <https://aiimpacts.org/2022-expert-survey-on-progress-in-ai/>.
- Stoler, Ann Laura. 2016. *Duress: Imperial Durabilities in Our Times*. Durham, NC: Duke University Press.
- Susskind, Leonard. 1995. "The World as a Hologram." *Journal of Mathematical Physics* 36 (11): 6377–96. <https://doi.org/10.1063/1.531249>.
- Tegmark, Max. 2014. *Our Mathematical Universe: My Quest for the Ultimate Nature of Reality*. First edition. New York: Alfred A. Knopf.
- . 2017. *Life 3.0: Being Human in the Age of Artificial Intelligence*. First edition. New York: Vintage.
- Tegmark, Max, and Stuart Russell. Open-Letter. 2016. "Autonomous Weapons Open Letter: AI & Robotics Researchers," February 9, 2016. <https://futureoflife.org/open-letter/open-letter-autonomous-weapons-ai-robotics/>.
- The Everyday Bible: New Century Version*. 2001. Nashville, TN: Thomas Nelson, Inc.
- Tsang, Ling Ming, Tin-Yam Chan, Shane T. Ahyong, and Ka Hou Chu. 2011. "Hermit to King, or Hermit to All: Multiple Transitions to Crab-like Forms from Hermit Crab Ancestors." *Systematic Biology* 60 (5): 616–29. <https://doi.org/10.1093/sysbio/syr063>.
- Vogl, Thomas, Cathrine Seidelin, Bharath Ganesh, and Jonathan Bright. 2019. "Algorithmic Bureaucracy." In *Proceedings of the 20th Annual International Conference on Digital Government Research*, 148–53. Dubai United Arab Emirates: ACM. <https://doi.org/10.1145/3325112.3325240>.
- Wemigwans, Jennifer. 2018. *A Digital Bundle: Protecting and Promoting Indigenous Knowledge Online*. Regina, SK: University of Regina Press.
- Whyte, Kyle P. 2018. "Indigenous Science (Fiction) for the Anthropocene: Ancestral Dystopias and Fantasies of Climate Change Crises." *Environment and Planning E: Nature and Space* 1 (1–2): 224–42. <https://doi.org/10.1177/2514848618777621>.
- Winter, Jenifer Sunrise, and Elizabeth Davidson. 2019. "Governance of Artificial Intelligence and Personal Health Information." *Digital Policy, Regulation and Governance* 21 (3): 280–90. <https://doi.org/10.1108/DPRG-08-2018-0048>.
- Wiryomartono, Bagoes. 2022. "Aristotle and the Doctrine of Aitia: A Theoretical Exploration of Environmental Design and Sustainability." In *Architectural Humanities in Progress: Divulging Epistemology, Ethics, and Aesthetics of the Built Environment and Habitation*, 21:177–94. Numanities - Arts and Humanities in Progress. Cham, CH: Springer International Publishing. <https://doi.org/10.1007/978-3-030-92280-1>.

- Wolchover, Natalie. 2022. "Physicists Create a Holographic Wormhole Using a Quantum Computer." *Quanta Magazine*, November 30, 2022, sec. Quantum Gravity. <https://www.quantamagazine.org/physicists-create-a-wormhole-using-a-quantum-computer-20221130/>.
- Wong, Jessica. 2023. "Make This 'a Teachable Moment about the Way That Technology Can Shape Discourse,' Says Media Studies Prof." *CBC*, January 21, 2023, sec. Canada. <https://www.cbc.ca/news/canada/chatgpt-aitools-highered-1.6716521>.
- Wood, Karenne. 2015. "Sky Woman." *American Indian Culture and Research Journal* 39 (1): 111–12.
- Yuan, Li. 2018. "How Cheap Labor Drives China's A.I. Ambitions." *New York Times*, November 25, 2018. <https://www.nytimes.com/2018/11/25/business/china-artificial-intelligence-labeling.html>.
- Ziegler, Micheal. 2022a. "Moving Between Hope and Pessimism: Failure, Death and 'Fuck-It'<sup>1</sup>." *New Political Science* 44 (3): 396–408. <https://doi.org/10.1080/07393148.2022.2028124>.
- . 2022b. "Text Message to a Friend," September 26, 2022.
- Zwicky, Jan. 2015. *Alkibiades' Love: Essays in Philosophy*. Montreal: McGill-Queen's University Press.