Symbiotic relationships in environments of change:

The potential of biological interactions in Artificial Intelligence

Marcos Soares

A thesis submitted in partial fulfilment of the requirements of the Royal College of Art for the degree of Doctor of Philosophy in Innovation Design Engineering

March 2024

Abstract

This design research interrogates the evolving paradigms of human-intelligent machine interactions, anchored in the biological principles of symbiosis. The thesis scrutinises the dynamic interplay between humans and technology, particularly in the context of the pervasive integration of intelligent machines into the multifaceted dimensions of daily life, which precipitates both novel challenges and opportunities. The investigation aims to elucidate the foundational elements that could incentivise these interactions and examines how researchers and designers might conceptualise a more profound symbiotic link between humans and intelligent machines. This doctoral inquiry probes the potential of leveraging symbiotic relationships observed in nature as a cornerstone for the advancement of machine intelligence systems, with the objective of augmenting human-machine interfaces. By establishing analogies between biological symbiosis—encompassing mutualism, parasitism, commensalism, and amensalism—and artificial intelligence (AI), the study endeavours to transpose these natural dynamics onto the design of intelligent machines that bolster human capabilities and contribute to societal well-being.

The thesis is propelled by three principal research questions: (1) How can symbiotic relationships observed in nature serve as a muse for the design of machine intelligence systems? (2) How can the principles of biological symbiosis be harnessed to fortify human-machine collaborations? (3) In what manner can symbiotic design principles be applied to overcome extant limitations within Artificial Intelligence algorithms and systems? Employing a structured methodology that covers environment comprehension and the proposition of innovative approaches, this doctoral thesis seeks to chart a forward-looking trajectory for Artificial Intelligence research and application, deeply rooted in the principles of symbiosis.

A practice-oriented research methodology unfolds across three distinct project branches. The initial branch delves into the experiential journey of a Machine Learning engineer, shedding light on the pragmatic facets of Artificial Intelligence development. The subsequent branch undertakes the categorisation of machine intelligence from a symbiotic perspective, advocating for a novel taxonomy for Artificial Intelligence systems. The final branch critically evaluates the ramifications of the symbiotic framework on privacy and surveillance within human-AI systems, underscoring the ethical and practical dilemmas of integration.

This dissertation highlights a design philosophy predicated on mutual enhancement and co-evolution between humans and machines, arguing that intelligent machines inspired by biological symbiosis can generate more adaptive, resilient, and ethically robust technologies. Focusing on an audience across the disciplines of Artificial Intelligence, design, and biological sciences, this work advocates for a paradigm where technology and humanity coalesce, guided by the knowledge stemming from natural symbiosis.

Table of Contents

List of Figures and Illustrations

Acknowledgements

This research would not be possible without the guidance, support and encouragement of several people across this journey. First, I wish to express my gratitude to my supervisors, Professor Ashley Hall and Dr. Delfina Fantini Van Ditmar. Additionally, I would like to thank all the previous supervisors and the staff at the Royal College of Art for their support.

I am also grateful for the interactions and discussions that I had with fellow practitioners in these past few years. A special thanks to the other doctoral researchers at the Royal College of Art, everyone from the London Doctoral Design Centre, and the people that collaborated with me at Microsoft Research, especially Prashant and Carlo.

Lastly, I want to thank my family, especially my parents José and Ana for all their constant support and motivation. I also want to thank my grandparents, uncles, aunts, cousins, nieces and nephews for their care, support and company. A special thanks to my friends: Luis Rodriguez, José Confesor, AJ, Pablo Castilla, João Martins, among many others.

Lastly, this PhD is dedicated to my four family members that passed away during this journey. What I learned from them about companionship and love cannot be described in words.

Chapter 01 / Motivation

‘We must recognize that, in a very deep sense, we were always hybrid beings, joint products of our biological nature and multi-layered linguistic, cultural, and technological webs’

Andy Clark in *Natural Born Cyborgs*

Research Motivation

Nestled in the heart of South America, the Amazon rainforest, often referred to as our planet's lungs, encompasses approximately 7,000,000 km^2 and spans across nine nations. It is a biodiversity hotspot, housing one-tenth of all known species on Earth, thereby standing as one of the planet's most diverse biological arenas. The Amazon is a cradle for multiple ecosystems, all intricately woven by biological interactions. It epitomises the interdependence of organisms and the complexity of symbiotic relationships, which have evolved over millions of years to sustain the ecosystem, ensure species survival, and provide a framework for organisms to navigate and adapt to environmental challenges (Wilson, 2001).

Examples of symbiotic relationships abound: flowers offering nectar and pollen to bees, which in turn enable the flowers' reproduction (Potts et al., 2003; Whitehead & Peakall, 2009; Willmer, 2011); the mutualism between red-billed oxpeckers and black rhinoceros (McElligott et al., 2004); the intricate balance between clownfish and anemones (Litsios et al., 2012; Pratte et al., 2018); the connection between barnacles and sand crabs (Messick, 1998; Shields et al., 2015); the relationship between sharks and remora fish (B. M. Norman et al., 2021; Xu et al., 2021); and the complex interplay between humans and their gut microbiota (Fan & Pedersen, 2021; Janssen & Kersten, 2015).

Symbiosis is characterised as a close biological interaction between two or more distinct organisms, often living in immediate proximity. These relationships may be mutualistic, with both entities benefiting; commensalistic, where one organism benefits while the other is neither advantaged or disadvantaged; parasitic, where one organism thrives at the other's expense; or amensalistic, where one organism is inhibited or destroyed while the other remains unaffected. The quintessence of symbiosis is the mutual interdependence of the organisms involved, leading to evolutionary adaptations that fortify their collective survival and prosperity (Douglas, 2010). These symbiotic relationships, having evolved over aeons, are instrumental in ecosystem maintenance, species survival, and the creation of conditions under which organisms can negotiate and address environmental challenges. Each interaction contributes not only a direct function but also influences the broader ecosystem, fostering a dynamic network of communication and belonging that adapts in response to environmental perturbations (Hadj-Hammou et al., 2021; Olds et al., 2012).

This research was initiated by observing these biological interactions and pondering the disparities between such natural symbioses and those manifesting in human-machine relationships. Human-Machine Symbiosis refers to the seamless and advantageous integration of machines, particularly intelligent systems, into human life, emphasising a partnership where humans and machines augment each other's capabilities(Breazeal et al., 2016; Coeckelbergh et al., 2018). In this symbiotic paradigm, intelligent machines assist humans with tasks that are repetitive, arduous, or beyond human capacity, while humans endow machines with purpose, direction, and ethical frameworks. This interplay seeks to enhance human potential, streamline tasks, and unlock new frontiers, all while ensuring that humans remain integral to decision-making processes.

My intrigue in biological symbiosis stems from an acknowledgment that biological phenomena harbour a wealth of insights often overlooked by academia and industry in favour of other knowledge sources. Nonetheless, it is imperative for researchers to recognise that biological phenomena can inspire human innovation. The concept of drawing inspiration from biological phenomena for technological advancement dates back to the earliest tool development epochs. For example, biomimetic architecture spawns innovation by emulating natural phenomena, such as a bird's wing influencing the design of an aeroplane (Benyus, 2009). This doctoral research distinguishes itself by concentrating on symbiotic relationships and their applicability to Artificial Intelligence, rather than serving as inspiration, as is the case with biomimetic architecture.

The thesis is motivated by a personal interest in human-machine symbiosis and concurrent advancements in Artificial Intelligence. Artificial Intelligence encompasses a spectrum of technologies and methodologies enabling machines to execute specific or general tasks ordinarily necessitating human intelligence (Russell & Norvig, 2020). This technology spectrum includes, but is not limited to, natural language processing, pattern recognition, complex problem-solving, and data-driven decision-making. Artificial Intelligence systems are designed to enhance their performance iteratively, adapting to new data and environmental stimuli. At the core of Artificial Intelligence are algorithms, crafted to process information, derive learnings, and apply these learnings to designated tasks. These algorithms can vary from rudimentary rule-based systems to sophisticated Machine Learning (ML) models that simulate the human brain's structure and functionality (Mitchell, 1997).

While "Machine Intelligence" is frequently used interchangeably with "Artificial Intelligence," there is a subtle distinction. "Machine intelligence" highlights the machine's inherent capabilities, whereas "Artificial Intelligence" emphasises the replication of human-like intelligence by non-human agents. In this research context, and in numerous academic and industry contexts, the terms are synonymous, describing identical concepts. Artificial Intelligence has undergone substantial enhancements over the past decade, prompting novel research opportunities. The research commences with a literature review in the realms of symbiosis and Artificial Intelligence, transitioning thereafter to the research's practical segment. However, secondary research, emanating from the literature review, was continually revisited and refined throughout this doctoral journey.

This doctoral research is delineated by three pivotal components: (1) Environment Understanding, (2) Proposition of a Novel Approach, and (3) Dissemination. The literature review and the inaugural research project, "Journey of the Machine Learning Engineer," endeavour to chart the contemporary research landscape, pinpoint opportunities and challenges within Artificial Intelligence, and provide a first-hand account of the trajectory involved in planning, developing, and deploying intelligent machines. The subsequent research project, "Types of Symbiotic Relationships and Symbiotic Primers," introduces a fresh perspective on human-machine symbiosis, seeking to discern correlations between machine intelligence tasks and various symbiotic relationships. The third component, encapsulated in the research project “Evaluating Privacy and Surveillance in Human-AI Systems through a Symbiotic Lens,” explores the potential of the designed symbiotic framework in the context of privacy and surveillance.

Research Questions

RQ01. How might symbiotic relationships observed in nature inspire and inform the design and development of machine intelligence systems?

RQ02. What are the potential capabilities of applying principles of biological symbiosis to human-machine interactions to enhance collaboration effectiveness?

RQ03. How might symbiotic design principles serve to address and ameliorate existing limitations within Artificial Intelligence algorithms and systems?

This doctoral thesis seeks to address the aforementioned research questions. The document is tailored for an audience committed to pushing the frontiers of machine intelligence and fascinated by the prospects of translating insights from biological interactions into the realm of artificial intelligence through design research.

Thesis Structure

This doctoral thesis is structured into seven comprehensive chapters.

Chapter 1 offers an introduction to the research's genesis, delineates the objectives this doctoral work seeks to fulfil, articulates the primary research questions, and outlines the document's overall structure.

Chapter 2 expounds on the motivations underpinning the doctoral research, clarifies essential concepts, and reviews current literature and creative endeavours within the domain of human-machine symbiosis.

Chapter 3 details the research methodology, approaches, and philosophical underpinnings. This chapter defines the researcher's stance, delves into the research questions, and sketches the contours of the practice-based doctoral journey.

Chapter 4 chronicles the first of three research projects within this doctoral work, "The Journey of the Machine Learning Engineer," illustrating the pathway entailed in developing intelligent machines and deliberating on the medium's inherent challenges and constraints.

Chapter 5 discusses the second research project, bifurcated into two sub-projects: "Types of Symbiotic Relationships" and "Symbiotic Primers." This chapter elaborates on the potential for augmenting human-machine symbiosis, explores the various types of symbiotic relationships and their attributes, and probes the interplay between machine intelligence tasks and symbiotic relationship types.

Chapter 6 summarises the third and final research project, "*Evaluating Privacy and Surveillance in Human-AI Systems through a Symbiotic Lens.*" This chapter explores the potential of the symbiotic framework delineated in the previous chapter, with a specific focus on the context of privacy and surveillance. It aims to provide a comprehensive understanding of how the symbiotic framework can be applied to assess and address the intricate dynamics between privacy, surveillance, and human-AI interaction.

Chapter 7 reflects on the contemporary landscape of human-machine symbiosis, revisits the research questions, and underscores the unique contributions of this research to the broader academic discourse.

Chapter 02 / Foundations

“Since the dawn of technology humans have endowed artifacts with mind. In our collective imagination, inhabited by objects, animals, and now machines, mind has rarely been held as an exclusively human preserve. Mind has prevailed until recently as a quality distributed among all things, captured one lifetime at a time and then returned.”

George B. Dyson in *Darwin Among the Machines*

A foundation serves as the bedrock or principle that underpins and supports other core elements. In this context, the foundation refers to prior research that justifies the need for further investigation into this subject. This chapter seeks to delineate the research's scope by critically reviewing extant literature and elucidating how the primary research undertaken will bridge existing knowledge gaps in the field. The foundational work is bifurcated into two segments: the first delves into research surrounding symbiosis, while the second probes the domain of machine intelligence.

Symbiosis

Symbiosis is a term broadly recognised yet often misunderstood. Originating from the Greek word *sumbíōsis*, meaning 'living together, 'it has been interpreted in numerous ways, contributing to its ambiguity. For clarity in this research, symbiosis is defined as any intimate, long-term interaction between two or more distinct organisms. This definition stems from a rich historical debate on the term, lasting over 130 years (Martin & Schwab, 2012; Sapp, 1994), and encompasses the concept's evolution since its coinage in 1878 by the German botanist Anton deBary. DeBary described symbiosis as the 'living together of differently named organisms '(Margulis, 1998). The contemporary understanding of symbiosis has extended beyond mere coexistence, recognising it as a crucial evolutionary driver, significantly shaping Earth's biodiversity.

Historically, the term "symbiosis" was established in the scientific lexicon by Anton deBary in 1878. DeBary's original definition highlighted the concept of different organisms living together, a notion that has undergone considerable elaboration over the years (Margulis, 1998). Today, symbiosis is not only seen as an arrangement of coexistence but as a foundational process influencing evolutionary developments and biodiversity. This view is vividly detailed in works like Microcosmos by Margulis and Sagan (Margulis & Sagan, 2023), which explores the role of microorganisms in evolutionary history, and Frank Ryan's Darwin's Blind Spot (Ryan, 2002), offering insights into diverse symbiotic relationships and their evolutionary impacts.

The research on symbiosis, while specialised, spans a broad array of disciplines, illustrating its interdisciplinary nature. Boucher's *The Biology of Mutualism: Ecology and Evolution* (Boucher, 1985) and Cuddington's *Ecosystem Engineers: Plants to Protists* (Cuddington et al., 2011) exemplify the extensive reach of symbiotic studies, from molecular details to broader ecological contexts. These studies highlight the integral role of symbiosis in various biological and ecological processes.

Symbiotic research, though it might appear as a niche field, has substantial implications across various scientific domains. The transformative influence of symbiotic interactions is far-reaching, extending from the minutiae of cellular processes to the grand scale of ecological networks. This breadth of impact is evidenced in the multitude of disciplines that symbiosis research encompasses. The concept's relevance and application in diverse scientific inquiries underscore its significance in understanding the complexity and dynamism of life on Earth.

The research seeks to provide a comprehensive and nuanced understanding of symbiosis. By tracing its historical roots, examining its evolving definitions, and exploring its multifaceted impacts across the scientific spectrum, we gain a deeper appreciation of the term's complexity and significance. This inquiry delves into the intricacies of symbiotic relationships, reaffirming their pivotal role in the evolutionary narrative and the sustenance of biodiversity on our planet.

Classifying Symbiosis

Symbiotic associations encompass a diverse range of interactions, including mutualism, commensalism, amensalism, and parasitism. Mutualism represents one end of this spectrum, characterized by mutually beneficial relationships between organisms. This ecological interaction is commonly observed in nature, with examples such as the interdependence between bees and flowering plants. In this symbiosis, bees facilitate plant reproduction through pollination, while simultaneously collecting nectar and pollen for their own nourishment. Other mutualistic examples include the association between spider crabs and algae, where the crab gains camouflage from the algae residing on its back, and the algae benefit from a secure habitat. Similarly, the relationship between red-billed oxpeckers (Buphagus erythrorhynchus) and the black rhinoceros (Diceros bicornis) exhibits mutualism, where the oxpecker accesses food sources while providing the rhinoceros with parasite control and predator alerts. The symbiosis between clownfish (Amphiprion ocellaris) and sea anemones (Heteractis magnifica) is another classic case, with the anemone providing shelter and the clownfish contributing nutrients through its waste. Finally, the relationship between humans and certain strains of Escherichia coli in the gut demonstrates mutualism through the production of vital nutrients such as vitamin K, essential for blood clotting (Douglas, 2010; Jorgensen & Fath, 2008; Margulis, 1998; Paszkowski, 2006; Sapp, 1994).

Parasitism, at the other end of the symbiotic spectrum, involves organisms living on or within a host, often causing harm. Robert Poulin, in "Evolutionary Ecology of Parasites," describes parasites as organisms that display varying degrees of adaptation to their hosts (Poulin, 2007). Examples include the parasitic love vine (Cassytha filiformis), which drains nutrients from other plants and gall wasps, and the barnacle Sacculina carcini, which infects and alters the reproductive behavior of sand crabs. The phorid fly (Apocephalus borealis) represents another parasitic case, laying eggs within bees and wasps, leading to eventual host death.

Commensalism describes interactions where one species benefits and the other is neither harmed nor aided. Classic examples include the relationship between sharks and remora fish, where remoras gain protection and transportation. Pseudoscorpions (Cordylochernes scorpioides) utilize Harlequin beetles (Acrocinus longimanus) for transport, and the Emperor shrimp (Periclimenes imperator) resides on sea cucumbers for protection (Hulme-Beaman et al., 2016; Jorgensen & Fath, 2008; Wilson, 2000).

Amensalism involves an interaction where one species is inhibited or destroyed while the other remains unaffected. For instance, the Spanish ibex (Capra pyrenaica) and weevils of the genus Timarcha compete for the same food source, resulting in the inadvertent consumption and detriment of the weevils by the ibex. The black walnut (Juglans nigra) secretes juglone, which inhibits the growth of surrounding herbaceous plants (Jorgensen & Fath, 2008).

Symbiotic interactions cover a spectrum from mutualism to parasitism, each with unique ecological and evolutionary implications. Mutualism is evident in relationships ranging from bees and flowering plants to humans and gut bacteria, each showcasing reciprocal benefits (Bronstein, 2015; Douglas, 2010; Jorgensen & Fath, 2008; Margulis, 1998; Paszkowski, 2006; Sapp, 1994). Parasitism, contrasting mutualism, involves a one-sided benefit at the host's expense, explored in depth by Poulin (2007) and Combes (2001). Commensalism's one-sided benefit without harming the other species is illustrated through interactions such as sharks with remora fish and pseudoscorpions with beetles (Fellous & Salvaudon, 2009; Hulme-Beaman et al., 2016; Jorgensen & Fath, 2008; Wilson, 2000). Lastly, amensalism, where one species is negatively affected while the other remains unaffected, is exemplified in the interactions between the Spanish ibex and weevils, and the black walnut's impact on its surroundings (Jorgensen & Fath, 2008).

Properties of Symbiosis

Symbiosis, characterised by the interdependent relationships between organisms, possesses attributes highly relevant to artificial intelligence, primarily adaptability. Such relationships, frequently facultative, are capable of evolving into intricate interdependencies. These associations demonstrate resilience and the capacity for introducing innovation within ecosystems by adapting to environmental changes and community shifts. The adaptability inherent in symbiosis is emphasised as a key force driving the evolution of organisms (Moran, 2007; Wernegreen, 2012).

The localisation of symbionts, whether inside (endosymbionts) or outside (ectosymbionts) another organism, is another significant property of symbiosis. Typically, endosymbiotic partnerships are more persistent, indicating a longer duration or more frequent contact, compared to ectosymbiotic associations. However, relationships like the episodic cleaning symbiosis of marine fishes or pollination by bees exhibit a more intermittent persistence.

Another dimension of symbiosis is the level of dependence, ranging from obligatory, where symbionts are highly specialised and dependent on the symbiotic relationship for survival, to facultative, where symbionts can also exist independently. Determining whether a symbiont is obligate or facultative can often be challenging due to the varying nature of these relationships.

The specificity of symbiotic associations is also noteworthy. Symbionts may exhibit a high specificity to one organism or associate with multiple hosts. Generally, the longer and more evolved the symbiotic relationship, the more specialised the interaction between the symbionts becomes (Paracer et al., 2000).

Symbiosis, through its adaptability and evolving interdependencies, presents a framework valuable to the study of artificial intelligence. For instance, the adaptability of these relationships to ecological and environmental shifts mirrors the desired resilience in AI systems. Margulis's Symbiosis in Cell Evolution (Margulis & Margulis, 1993) delves into the cellular and physiological aspects of these interactions, while Boucher's The Biology of Mutualism: Ecology and Evolution (Boucher, 1985) explores their behavioural dynamics.

Endosymbiotic events have been crucial in life's evolutionary history, as elucidated in Buchner's Endosymbiosis of Animals with Plant Microorganisms (Buchner, 1965). Additionally, Thompson's Geographic Mosaic Theory of Coevolution (J. N. Thompson, 2005) provides insight into the adaptive nature of symbiotic interactions. The role of symbiosis as a catalyst for evolutionary change is further supported by the works of Moran (Moran, 2007a) and Wernegreen (JJ, 2012).

Finally, the exploration of symbiosis's broader cultural and societal implications is masterfully presented in Quammen's The Tangled Tree: A Radical New History of Life (Quammen, 2019). Collectively, these scholarly works illustrate the diverse facets of symbiosis, extending from molecular mechanisms to overarching ecological and cultural impacts.

The Role of Symbiosis in Technological Evolution

Symbiosis, a pivotal force within ecosystems, drives innovation and novelty. This doctoral research theorises that the principles of symbiosis, fostering unique interactions in nature, can also be applicable to the evolution of machine intelligence. Historically, the evolution of mankind has been closely knit with technological advancements, illustrating a symbiotic relationship between humans and machines.

Ancient tools like the astrolabe exemplify this relationship, reflecting humanity's longstanding curiosity about the universe. Such artefacts not only demonstrate early technological feats but also signify the enduring collaboration between humans and technology. In the digital age, this relationship has evolved, yet its essence remains. The transition from mechanical to digital interactions marks a significant shift in this symbiosis. The PARO therapeutic robot, developed by AIST in Japan, epitomises this evolution. By aiding in stress reduction and socialisation for patients and caregivers, PARO highlights the continuing evolution of human-machine symbiosis (Calo et al., 2011).

The exploration of symbiosis as an innovation catalyst in both natural and technological domains is central to this research. The study draws upon natural symbiotic relationships, suggesting their potential to inspire and inform developments in machine intelligence.

A multidisciplinary approach is necessary to fully understand symbiosis. In ecology, seminal works like Margulis and Fester s Symbiosis as a Source of Evolutionary

Innovation (Margulis & Fester, 1991) and Smith and Douglas The Biology of Symbiosis (Smith et al., 1987) explore evolutionary mechanisms in symbiotic relationships. Bronstein s Mutualism (Bronstein, 2015) discusses the ecological implications of such interactions.

In the sociological realm, Haraway s When Species Meet (D. J. Haraway, 2013) examines the complex relationships between humans and other species, reflecting on similar symbiotic connections. Tsing s The Mushroom at the End of the World (Tsing, 2017) provides an anthropological perspective on human interactions with nature under the influence of global capitalism.

Philosophically, works like Deleuze and Guattari s A Thousand Plateaus (Deleuze et al., 2004) and Bateson's Steps to an Ecology of Mind (Bateson, 2000) delve into interconnectedness and systemic thinking, relevant to understanding symbiosis.

Technological aspects are crucial for this topic. Kelly's Out of Control (Kelly, 2009) discusses the convergence of biology and technology. Rifkin's The Zero Marginal Cost Society (Rifkin, 2014) compares the emerging digital economy's collaborative nature with natural symbiotic processes.

From an ethical and systemic viewpoint, integrating Russell and Norvig's Artificial Intelligence: A Modern Approach (Russell & Norvig, 2020b) enriches understanding of AI's evolution in this symbiotic framework. Norman's The Design of Everyday Things (D.

Norman, 2013) provides insights into human-computer interaction, a key aspect of modern symbiosis. Wiener's Cybernetics (Wiener et al., 2019) and von Bertalanffy's General System Theory (von Bertalanffy et al., 2015) contribute foundational knowledge on control and communication in animals, machines, and systems – all pivotal in the discourse of technological symbiosis.

Addressing the ethical dimensions, Bostrom's Superintelligence: Paths, Dangers, Strategies (Bostrom, 2014) offers a critical view of AI s future trajectory within this human-machine partnership. Capra and Luisi's The Systems View of Life (Capra & Luisi, 2014) suggests an integrated perspective on life s interconnected systems, pertinent to understanding technological and ecological symbioses.

Culturally, Morton's The Ecological Thought (Morton, 2012) examines the intricate web of life, highlighting the role of symbiosis in shaping environmental and cultural narratives.

This diverse collection of sources, covering biology, ecology, sociology, anthropology, philosophy, technology, and ethics, provides a comprehensive understanding of symbiosis. It underscores the interconnection between various domains, enriching the thesis and offering a holistic view of the role of symbiosis in technological evolution.

Human-Machine Symbiosis

Central to this doctoral thesis is the concept of human-machine symbiosis, initially envisioned by Licklider in 1960 at MIT. Licklider foresaw a future where humans and computers would interact fluidly, with each contributing unique strengths towards shared goals, thus achieving what neither could alone. He highlighted the differences in human and computer capabilities across various dimensions, including speed, memory capacity, organisation, and types of analysis, suggesting that computers would assume routine tasks while humans would focus on goal-setting and creative planning (Licklider, 1960). Today, Licklider's vision is partly realised, as computers have become integral to many aspects of daily life, from basic internet searches to the sophisticated decision-making processes in autonomous vehicles.

Beyond traditional computing, the past decade's advancements in artificial intelligence have broadened the scope of human-machine symbiosis. This expansion encompasses not just extensions of human capabilities but also positions machines as tools or partners, indicating varying levels of machine autonomy (Clark, 2003, 2008). Although Licklider’s work addressed machine autonomy, it didn’t fully explore the evolution and adaptability of machines—key aspects this research intends to examine through design research. Understanding the history and development of artificial intelligence is crucial to this endeavour.

Several pivotal texts provide a multifaceted framework for exploring human-machine symbiosis. Cybernetics, as discussed in Wiener's Cybernetics: Or Control and Communication in the Animal and the Machine, lays a foundational understanding of the interplay between humans and machines (Wiener, 1965). Engelbart's Augmenting Human Intellect: A Conceptual Framework sheds light on early ideas of enhancing human abilities via technology (Engelbart, 2021). From a philosophical angle, Heidegger's The Question Concerning Technology probes the essence of technology in relation to humanity (Heidegger, 1977), and Turkle's Alone Together considers the psychological and social effects of growing technological dependency (Turkle, 2017).

Technologically, Kurzweil's The Singularity Is Near and Brooks' Flesh and Machines discuss predictions and perspectives on the future interplay between humans and technology (R. Brooks, 2002; Kurzweil, 2005). Haraway's A Cyborg Manifesto and Balsamo's Technologies of the Gendered Body examine the cultural and gendered implications of technology in human-machine relationships (Balsamo, 1996; D. Haraway, 2006). Lastly, Clark's Natural-Born Cyborgs and Nicolelis' Beyond Boundaries provide insights into integrating human cognition with machines and the forefront of brain-machine interface research (Clark, 2003; Nicolelis, 2011). This literature provides a comprehensive understanding of the field, essential for exploring the knowledge gaps this thesis aims to address.

Connecting Machine Intelligence to Symbiosis

The intricate interdependence observed in nature, particularly within symbiotic relationships, offers a fresh perspective for understanding the evolving connection between humans and Artificial Intelligence. Symbiosis, derived from the Greek word *sumbíōsis* meaning "living together," encompasses the essence of mutually beneficial interactions where organisms coexist and often thrive due to their association. While traditionally applied to biological entities, Donna Haraway, in her seminal work Simians, Cyborgs, and Women, argues that the concept of symbiosis resonates profoundly in the realm of Artificial Intelligence and human-machine interactions (D. J. Haraway, 2015). Donna Haraway's concept of symbiosis in Simians, Cyborgs, and Women (2015) goes beyond a simple biological metaphor. She argues for "situated knowledge," where knowledge production is not objective but rather arises from the specific contexts and relationships we inhabit. In the context of human-machine symbiosis, this implies that our understanding of Artificial Intelligence is shaped by our ongoing interactions and co-evolution with these technologies. For example, as Artificial Intelligence becomes more integrated into healthcare, our understanding of health and medicine will be shaped by the capabilities and limitations of these intelligent systems. This co-evolutionary process creates a new "situated knowledge" specific to the human and intelligent machines partnership in healthcare.

Additionally, Haraway's influential concept of the cyborg further emphasises the blurring boundaries between human and machine. The cyborg is not a monstrous fusion but rather a representation of our increasingly intertwined existence with technology. This concept challenges traditional notions of human exceptionalism and underscores the interconnectedness of biological and technological systems. In the context of human-machine symbiosis, the cyborg highlights the potential for a more nuanced understanding of human and intelligent machine partnerships. Humans are not tool users, but we are co-creators and co-evolvers alongside Artificial Intelligence. This perspective demands a shift towards responsible and ethical design principles that foster a mutually beneficial symbiosis.

Historically, the biological definition of symbiosis, established by Anton deBary in 1878 and later expanded upon by Lynn Margulis (Margulis, 1998), laid the groundwork for the understanding of these relationships. However, this perspective can be extended metaphorically to explore the human-machine relationship. This argument gains traction from Haraway's broader notion of entangled existence within complex systems (D. J. Haraway, 2015).

The human-machine relationship exhibits characteristics similar to the evolutionary trajectories observed in biological symbiosis. J.C.R. Licklider's 1960 vision at MIT, which foresaw a seamless collaboration between humans and computers, capitalising on their respective strengths, exemplifies this evolving symbiotic relationship (Licklider, 1960). As Artificial Intelligence has grown more sophisticated, the nature of this partnership has shifted from one of simple utility to a collaborative endeavour, as evidenced by Andy Clark's work (Clark, 2003, 2008).

The connection between symbiosis and Artificial Intelligence can be further solidified by providing more specific examples of symbiotic relationships in nature that mirror aspects of human-machine interaction. In terms of mutualism, similar to the classic example of clownfish and anemones, where each organism benefits from the other's presence, humans and Artificial Intelligence can develop partnerships where both parties gain advantages. For example, Artificial Intelligence powered language translation tools can enhance human communication across cultures, while human feedback can improve the accuracy and efficiency of these translation systems. When it comes to Commensalism, and similar to the relationship between birds and seed-dispersing plants, humans can leverage Artificial Intelligence for specific tasks without necessarily offering anything in return in the short term. For example, Artificial Intelligence powered search engines assist us in finding information, but we may not directly contribute to the development of the search engine itself. Relating to parasitism, and while it is a potentially negative form of symbiosis, it also serves as a cautionary tale for human and intelligent machine interaction. Humans must be vigilant in developing Artificial Intelligence systems that are beneficial to humanity and avoid scenarios where Artificial Intelligence becomes parasitic, exploiting human resources for other purposes rather than improve humanity.

The adaptability inherent in symbiotic relationships, explored by Moran (2007) and Wernegreen (2012), offers valuable insights into how machine intelligence might evolve alongside human needs and aspirations. Similarly, Evan Thompson's Mind in Life (E. Thompson, 2007) underscores the potential for machines to integrate seamlessly into the human experience. Just as natural symbiosis fosters resilience and introduces novelty into ecosystems, the human and intelligent machine partnership has the potential to be a wellspring of innovation, driving progress across technology and society.

The limitations of current approaches to human-AI interaction highlight the need for a symbiosis framework. A symbiotic approach can help address the following concerns:

- **Focus on Technical Capability**: The dominant discourse often emphasises the technical prowess of Artificial Intelligence, neglecting the importance of human and intelligent machine collaboration and the social implications of these technologies. A symbiotic framework refocuses attention on the need for a balanced and mutually beneficial partnership.

- **Power Imbalances**: Current Artificial Intelligence development can lead to power imbalances, with Artificial Intelligence potentially becoming a dominant force. A symbiotic framework emphasises co-evolution and shared agency, ensuring that both humans and machines have a role in shaping the future of this relationship.

- **Ethical Concerns**: Issues of privacy, bias, and control become even more critical in a symbiotic future. By adopting a symbiotic lens, humans can proactively design and develop Artificial Intelligence systems that prioritise ethical considerations and ensure a balanced partnership.

The rich tapestry of research on symbiosis, from foundational works to contemporary explorations, offers more than just a metaphor. A symbiotic approach provides a robust conceptual framework for understanding the deep relationship between humans and Artificial Intelligence. By critically analysing the literature through the lens of human-machine symbiosis, we can glean crucial insights into the future trajectory of this partnership and its societal implications.

To navigate deeper into the potential of the human-machine symbiosis framework, the research proposes a three-pronged approach:

1. **Developing a Human-Machine Symbiosis Framework**: The initial phase will involve constructing a comprehensive framework for studying human-machine symbiosis. Drawing upon existing literature on biological symbiosis and Artificial Intelligence, the framework translates these concepts to the unique context of human-machine relationships. Consequently, it will encompass key characteristics of successful symbioses, such as mutual benefit, coevolution, and adaptation.

2. **Case Studies in Symbiotic Collaboration**: The research will then analyse real-world examples of human-machine partnerships. The analysis will explore diverse applications of Artificial Intelligence powered tools used in healthcare, education, and creative industries. By evaluating these partnerships through the lens of the developed framework, the research will identify strengths and weaknesses in existing human-machine collaboration.

3. **Ethical Considerations of a Symbiotic Future**: The third strand of research will indirectly explore the ethical implications of human-machine symbiosis. This will involve investigating potential challenges associated with a deeper human and intelligent machine relationship, such as issues of privacy, autonomy, and control. This strand of research will also explore opportunities to promote the ethical development and deployment of Artificial Intelligence in symbiotic partnerships.

Machine Intelligence

Symbiotic relationships, particularly those within the realm of machine intelligence, form the core of this research. The term "machine intelligence" is utilised to encompass the full spectrum of artificial intelligence, both in its practical applications and in its more theoretical or speculative aspects.

The historical intertwining of humans and machines, visible since the early development of tools and artefacts, has become increasingly complex due to the rapid advancement of machine capabilities. This evolving relationship has captivated researchers and artists for generations, leading to a diverse range of interpretations and manifestations, from concrete engineering innovations to abstract philosophical ideas.

A notable early thinker in this domain was Roger Bacon, a 13th-century English philosopher and scientist. Renowned as an early advocate of the modern scientific method, Bacon’s seminal work, *Opus Majus*, exhibits a remarkable breadth of knowledge across various scientific domains. In this extensive text, Bacon explores topics ranging from the pursuit of wisdom and truth to the intricacies of mathematics, optics, and experimental sciences. His analysis of optics covers imagination, the human eye’s anatomy, and the interplay between the eye and the brain, as well as the influence of factors like light, color, distance, position, and vision (direct, reflected, and refracted) through mirrors and lenses. His exploration into experimental science not only reviews alchemy and celestial bodies but also anticipates inventions such as microscopes, telescopes, flying machines, and steam ships. Furthermore, the final part of *Opus Majus* addresses moral philosophy and ethics, showcasing Bacon's multidisciplinary intellect.

Other significant contributions to the field include Alan Turing's landmark paper, *Computing Machinery and Intelligence* (Turing, 1950), which introduced the Turing Test, fundamental to modern AI discussions. Norbert Wiener's seminal book, *Cybernetics* (Wiener, 1965), elaborated on the control and communication in machines and living organisms, enriching the discourse on human-machine interaction. Ray Kurzweil's *The Singularity Is Near* (Kurzweil, 2005) provides a visionary outlook on the convergence of humans and machines, and Hubert Dreyfus (Dreyfus, 1972, 1992) offers a critical perspective on the capabilities and limitations of AI.

Returning to Bacon's *Opus Majus*, his examination of experimental science not only underscores his scientific foresight but also highlights his contributions to the field of automata. This is exemplified by the story of the mechanical brazen head, a legendary automaton attributed to Bacon, reputed to answer binary questions—a reflection of his innovative spirit and understanding of mechanical intelligence (Rosheim, 1994).

Through exploring these foundational works, this research aims to deepen the understanding of the synergistic relationships between humans and machines. It examines not only the historical trajectory of these interactions but also their future potential and implications. By weaving together insights from early visionaries like

Bacon with contemporary thinkers and innovators, we can form a more comprehensive view of the symbiotic evolution of humans and machines, laying a foundation for exploring novel avenues in machine intelligence.

Figure *1* - Depiction of Roger Bacon's Mechanical Brazen Head

Automata

Automata, as defined by the Oxford English Dictionary, are mechanical devices created to mimic human actions. Operated by pre-programmed coded instructions, they have long fascinated and inspired inventors across various cultures and epochs. Historical references to automata include Greek mythology, Jewish folklore, and ancient Chinese stories. Joseph Needham's *Science and Civilisation in China* (Needham, 1965) elucidates early Chinese automata within a wider technological context.

Muslim societies during the early centuries showed considerable interest in constructing human-like machines. Ismail Al-Jazari, a prominent Muslim inventor, is particularly recognised for devising several complex mechanical devices. Among his inventions, a notable example is a water basin system that, upon pulling a lever, drains and then refills the basin, a precursor to the modern flush toilet (Rosheim, 1994).

Figure *2* - Al-Jazari device for water draining and refill

The Renaissance marked a resurgence of interest in automata, accompanied by increased sophistication in their design. Italian engineer Giovanni Fontana created machines resembling animals and demonic figures, while Leonardo da Vinci engineered a complex automaton, the Automa cavaliere or Leonardo’s mechanical knight. During this period, philosophical perspectives on automata diversified. René Descartes, the French philosopher and scientist, argued that animal bodies could be considered elaborate machines, their organs replaceable with mechanical counterparts. Jacques de

Vaucanson, inspired by such mechanistic views, built the Canard Digérateur or Digesting Duck, an automaton that mimicked the acts of eating, wing-flapping, and digesting grain (Rosheim, 1994). Jessica Riskin's *The Restless Clock* (Riskin, 2016) further delves into the philosophical and scientific discourse of this era, particularly regarding the mechanization of life and living beings.

Figure *3* - Mechanical device by Giovanni Fontana

Figure *4* - Another mechanical device by Giovanni Fontana

Figure *5* - Da Vinci's Mechanical Knight

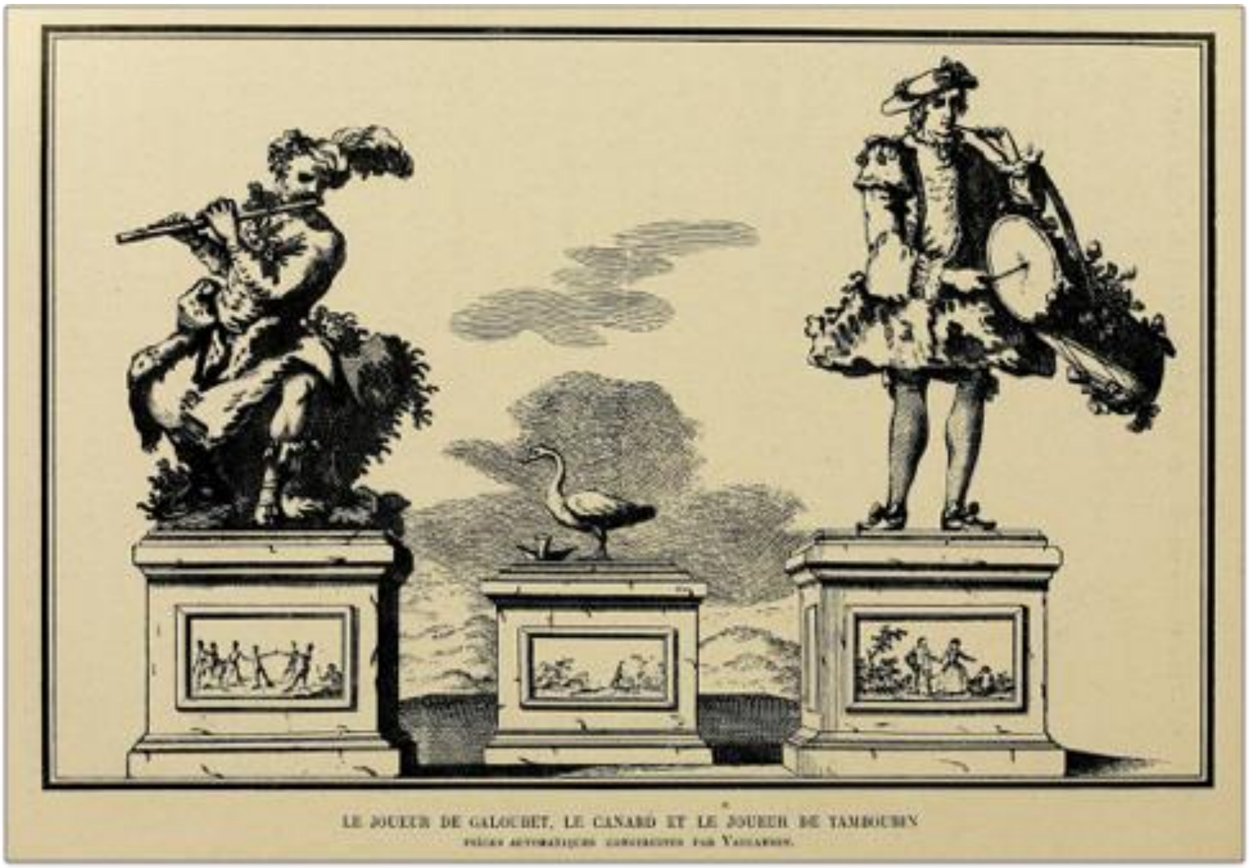


Figure *6* - Jacques de Vaucanson’s Flute Player, the Digesting Duck and the Tambourine Player

Among the most intriguing examples of historical automata is the Mechanical Turk, crafted by Wolfgang von Kempelen in the late 18th century. Though essentially a hoax, this chess-playing mechanism, housing a concealed human operator, gained fame throughout Europe and America for its seemingly intelligent play.

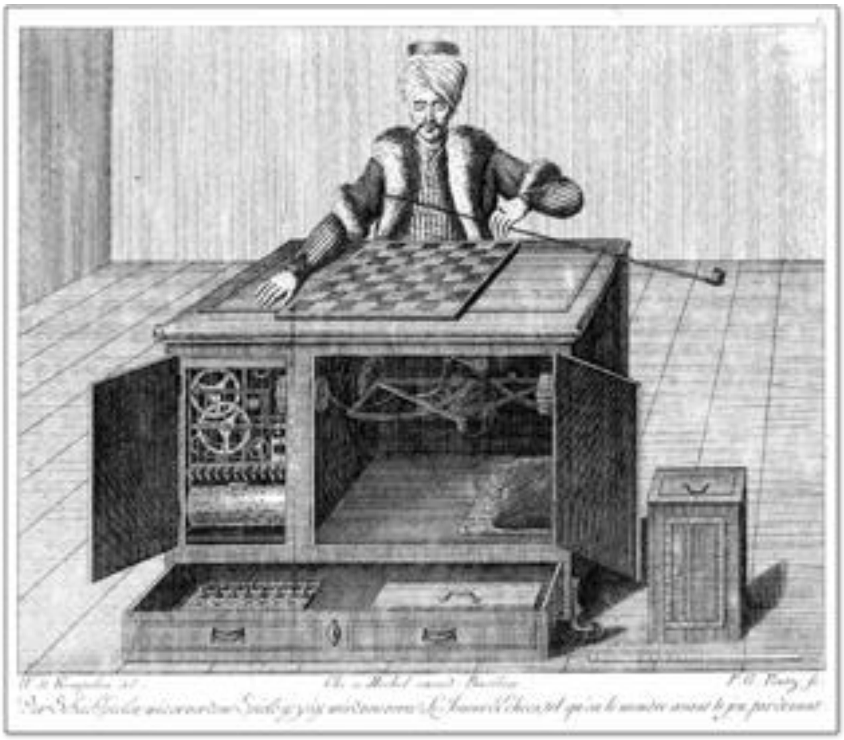


Figure *7* - Copper engraving of the Mechanical Turk

Automata can be regarded as a pivotal step in the historical development of the relationship between humans and machines. Preceding modern robotics and artificial intelligence, these mechanical devices contributed significantly to the advancement of technology. The evolution of automata, as documented in Gaby Wood's *Living Dolls* (Wood, 2002) and Elly Truitt's *Medieval Robots* (Truitt, 2015), facilitated the emergence of computing devices. These developments have not only improved communication between humans and machines but have also engendered a visionary realm where intelligent machines pervade every aspect of life.

In examining the history of automata, it becomes apparent that these devices are more than mere artefacts of historical ingenuity. They symbolise the longstanding human quest to replicate and understand our own mechanisms of life and thought, leading to pivotal advancements in technology and philosophy. This lineage of mechanical imitation and the pursuit of artificial intelligence have fundamentally shaped our current understanding of, and relationship with, machines. Reflecting on the evolution of automata offers vital insights into our past innovations and guides us towards future technological and ethical considerations in the ever-evolving field of machine intelligence.

Digital Computing

The field of automata, with its intricate mechanical devices designed to imitate human actions, sharply contrasts with the concept and development of computers. Yet, intriguingly, these two realms converge in various aspects. Defined by the Oxford English Dictionary, a computer is an electronic device that stores and processes data according to a program's instructions. The use of computation predates the advent of automata, manifested in ancient devices such as tally sticks and the abacus, used across numerous early societies. While a comprehensive history of digital computing exceeds this research's scope, certain milestones pivotal in shaping the human-machine relationship warrant discussion.

Thomas Hobbes, an English philosopher, was a forebearer in examining this relationship. In his seminal work *Leviathan or the Matter, Forme and Power of a Common-Wealth Ecclesiasticall and Civil* (Hobbes, 1651), Hobbes delved into the existence of a collective intelligence, superseding the intellect of individual entities and their institutions, positing that human society itself constituted a novel life form. This perspective significantly influenced André-Marie Ampère, leading to his coining of the term "cybernétique" for the study of political organisations. Ampère, known for his contributions to game theory and electrodynamics, thus laid the groundwork for Norbert Wiener and the field of Cybernetics (Dyson, 2012).

Hobbes's notion of Leviathan suggests that human logic and digital computation stem from shared foundational knowledge. This basis is increasingly visible in contemporary society with the advent of Artificial Intelligence. The convergence of collective and hierarchical processes, underpinning both biology and technology, fosters a symbiotic relationship between human and machine intelligence. American computer scientist William Daniel Hillis posits that machine intelligence can evolve autonomously due to the inherent limitations in engineered intelligence (Dyson, 2012). In "Darwin Among the Machines," George Dyson echoes Hillis's view, suggesting that while individual computers and software evolve towards AI, the expansive networks facilitating inter-computer communication might catalyse an emergent "Leviathan" of artificial mind (Dyson, 2012). Nick Bostrom, a Swedish philosopher, builds upon Hobbes's Leviathan, theorising that superintelligence could emerge through enhancing networks and organisations that interconnect individual minds with various artefacts and bots (Bostrom, 2014).

Gottfried Leibniz, a German polymath inspired by Hobbes and Blaise Pascal, significantly impacted arithmetic and logic, contributing to the genesis of digital computers. Norbert Wiener, in "Cybernetics: or Control and Communication in the Animal and the Machine," underscores Leibniz's importance, noting his advocacy of universal symbolism and a calculus of reasoning (Wiener, 1965).

Charles Babbage's Analytical Engine remains a critical artefact in computing history. Although Babbage never completed the Engine, its conceptual framework was revolutionary, incorporating elements recognised only during World War II. Babbage foresaw the formal languages and timing diagrams vital to modern computing (Dyson, 2012). Other pivotal figures include George Boole, who introduced Boolean logic, thus expanding computers' mathematical capabilities; Alfred Smee, with his method of parsing language through cyphers; and Ada Lovelace, recognised for devising one of the earliest algorithms intended for machine execution.

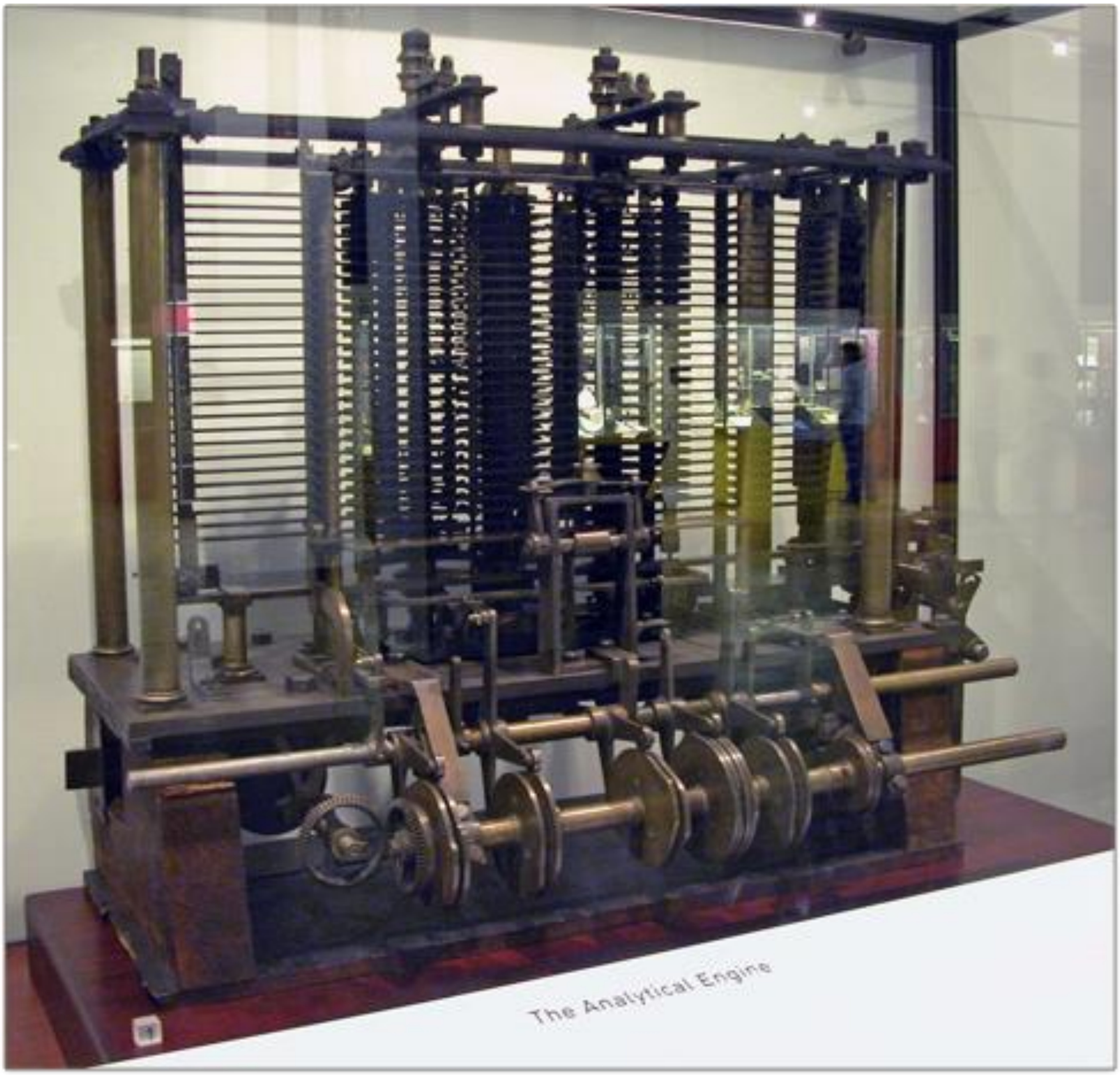


Figure *8* - Babbage's Analytical Engine

The era of modern computers emerged during World War II. Prior to this, the term "computer" typically referred to humans, predominantly women, engaged in calculations or computations. These human computers later transitioned into some of the first computer programmers. The advent of modern computers aimed primarily at deciphering extensive codes during the war, exemplified by the Z3, the Atanasoff-Berry Computer (ABC), the ENIAC, and Colossus (Dyson, 2012). Stuart Russell and Peter Norvig highlight the Heath Robinson, developed by Alan Turing and his team in the 1940s for decrypting German communications, as the first operational computer. Despite its initial inefficiency, the construction of the Colossus under Thomas H. Flowers, managed by Max Newman, signified a major advancement, demonstrating the feasibility of computer replication (Russell & Norvig, 2020). The Z3, created by Konrad Zuse in 1941, was the first operational programmable computer, with Zuse also inventing floating-point arithmetic and Plankalkül, the first high-level programming language. The ABC is regarded as the initial electronic computer, and the ENIAC, developed at the University of Pennsylvania for military use, stands out for being Turing-complete, digital, and reprogrammable (Dyson, 2012).

This research focuses on the symbiotic relationships within the realm of machine intelligence, encompassing both its practical and speculative aspects. Pamela McCorduck, in *Machines Who Think* (McCorduck & Cfe, 2004), provides a thorough historical and philosophical examination of artificial intelligence, offering essential insights into the evolution of machine intelligence and its interplay with humans.

Historically, symbiotic human-machine relationships have been evident since the earliest use of tools and other artefacts. The complexity of this interaction has grown with the rapid advancement of machines. This dynamic has been a topic of intense scrutiny and creativity among researchers and artists, generating a rich array of documented viewpoints and interpretations, from engineering achievements to philosophical reflections.

Luciano Floridi, in *The Fourth Revolution* (Floridi, 2014), explores the transformative effects of information technology on human society, including the burgeoning field of artificial intelligence. His analysis presents a holistic view of the digital revolution and its profound implications for human existence.

Artificial Intelligence

The hardware underpinning machine intelligence is equally as integral as the software that animates it. A historical overview of technological advancements leading up to the 1950s reveals the foundational developments that heralded the advent of artificial intelligence. With the evolution of computers into versatile, replicable devices capable of being meticulously programmed, a novel form of cooperation between humans and machines emerged. During this era of burgeoning computer technology, pioneering research began to illuminate the concept of thinking machines. Key contributions from Norbert Wiener in Cybernetics, Claude Shannon in Information Theory, and Alan Turing's Computational Theory posited the feasibility of electronically reconstructing human cognitive processes.

In 1943, the seminal work of Warren McCulloch and Walter Pitts marked the commencement of what is now recognised as artificial intelligence. Building upon

Turing's Computational Theory, an understanding of neuronal physiology, and a detailed analysis of propositional logic, they introduced a model of artificial neurons capable of binary states – either 'on' or 'off'. This foundational model set the stage for subsequent developments, notably Donald Hebb's formulation of the Hebbian learning method, which enabled the dynamic adjustment of synaptic connections between neurons, as documented by Russell and Norvig in 2020.

Alan Turing's influential contributions to Artificial Intelligence, as articulated in his paper "Computing Machinery and Intelligence," further deepened the field. Turing introduced pivotal concepts such as the Turing test, machine learning, genetic algorithms, reinforcement learning, and the concept of a 'child machine'. In this work, Turing proposed a radical approach:

"Rather than endeavouring to replicate the adult mind, why not attempt to model the child's mind? Subjected to a suitable educational trajectory, this could ostensibly lead to the development of an adult brain. Analogous to a blank notebook from a stationer, the child's brain, with minimal inherent structure, offers extensive potential for learning and development. We hypothesize that the simplicity of the child's brain mechanism might allow for easier programming. We envisage that the 'education' of this nascent machine intelligence would approximate the developmental process of a human child. Thus, we bifurcate our challenge into the creation of the foundational 'child machine' and its subsequent educational pathway. This iterative process of development and refinement, drawing parallels with evolutionary theory, aligns the structure of the child machine with genetic heredity, while alterations to the machine mimic biological mutations. The experimenter's evaluative role is akin to the process of natural selection." (Turing, 1950).

In this context, understanding both the hardware and software aspects of machine intelligence is crucial. The journey from the mechanical looms of the early 19th century to the electronic computers of the mid-20th century signifies a remarkable technological evolution, culminating in the first iterations of AI. Concurrent with the development of modern computers, the early theories and models posited by Wiener, Shannon, and Turing opened new vistas in understanding and emulating human cognition electronically. McCulloch and Pitts ’binary neuron model, along with Hebb's learning mechanism, provided a concrete basis for this emulation. Turing's exploration of AI through the metaphor of a child machine's development underscored a profound shift in perspective, proposing a developmental, iterative approach to AI that mirrors human learning and evolution.

These developments not only represent significant technical milestones but also symbolise a transformative era in the relationship between humans and machines, an era where digital intelligence began to mirror, and potentially extend, the capabilities of the human mind. The progression from basic computational models to complex learning algorithms and the conceptualisation of AI as an evolving, educable entity underscores a dramatic shift in our understanding and aspirations for machine intelligence. This historical perspective provides invaluable context for contemporary AI research, highlighting both the rapid advances and the enduring challenges in creating machines that can think, learn, and evolve akin to human beings.

The Dartmouth Conference in 1956 is universally acknowledged as the genesis of Artificial Intelligence (AI). This historic gathering united John McCarthy, Marvin Minsky, Claude Shannon, and other notable figures, tasked with the ambitious objective of developing machines capable of using language, forming abstract concepts, solving human-exclusive problems, and self-improvement. Although the conference didn't produce immediate breakthroughs, it was instrumental in fostering collaboration among key AI pioneers (McCarthy et al., 2006).

The era following the Dartmouth Conference up to the early 1970s was marked by heightened enthusiasm, soaring expectations, and substantial funding. During this period, significant strides were made. Allen Newell and Herbert Simon developed the General Problem Solver, an attempt to mimic human problem-solving skills from the outset. This development represented one of the first instances of the 'thinking humanly' approach in AI (Russell & Norvig, 2020). The program's success, along with earlier efforts like the Logic Theorist, led to the formulation of the Physical Symbol System hypothesis. This hypothesis posited that a physical symbol system is both necessary and sufficient for general intelligent action (Russell & Norvig, 2020b).

In the realm of machine reasoning, Herbert Gelernter designed the Geometry Theorem Prover, capable of solving complex mathematical theorems. Another Dartmouth participant, Arthur Samuel, pivoted to developing programs that could learn checkers, showing that machines can outperform humans in specific tasks. Concurrent developments at MIT included John McCarthy's creation of LISP, the pioneering AI programming language; the introduction of time-sharing for computer resources; and the conceptualization of the Advice Taker, an early model of a complete AI system. Additionally, Joseph Weizenbaum's ELIZA, one of the first programs addressing natural language processing, inadvertently demonstrated humans' tendency to anthropomorphize machines.

Marvin Minsky's impact, alongside Seymour Papert, was profound. They concluded that the intricacy of the real world necessitated focusing AI research on limited domains or micro-worlds. This shift resulted in successful projects such as James Slagle's SAINT for solving calculus problems, Tom Evans' ANALOGY for geometric analogies, Daniel Bobrow's STUDENT for algebraic stories, and Terry Winograd's SHRDLU, a landmark in AI demonstrating advanced natural language understanding and object manipulation (Russell & Norvig, 2020).

Following this era of excitement, the first AI winter emerged, characterized by reduced funding, limited computing resources, and growing academic and philosophical criticism. Russell and Norvig highlight three key challenges: the reliance of early AI programs on syntactic manipulations without understanding the content; the intractable nature inherent in many AI approaches; and the fundamental limitations of the basic structures used to model intelligent behavior. AI experienced a revival in the early 1980s with the broader adoption of expert systems in industry, such as Stanford's DENDRAL and MYCIN, the latter diagnosing infectious blood diseases. This resurgence shifted the focus from general intelligence to domain-specific problem-solving capabilities (Russell & Norvig, 2020).

Simultaneously, the field of connectionism, advocating the use of artificial neural networks to replicate human intelligence, regained prominence. Central to connectionism are the concepts of distributed representation—each input represented by multiple features—and backpropagation, a method critical for training deep neural networks. The resurgence of backpropagation significantly propelled machine learning research, laying the groundwork for modern AI advancements (Goodfellow et al., 2016).

In summary, the progression of AI from its inception at the Dartmouth Conference to its modern-day manifestations has been marked by a series of ebbs and flows, influenced by technological advancements, theoretical breakthroughs, and shifts in research focus. This dynamic history underscores the complex, evolving nature of AI as a field driven by both theoretical inquiry and practical applications, navigating through periods of both skepticism and acclaim.

The second AI winter, enduring until the early 1990s, signified a pivotal shift in AI research, marked notably by the advent of intelligence agents and the expansion of the Internet. This period entailed a retrospective examination of the progress in AI since the 1950s, revealing a focus on solving narrow AI problems while largely neglecting the broader 'whole agent' concept integral to the field's inception. The internet facilitated a burgeoning environment for intelligent agents, leading to the development of search engines, recommender systems, and website aggregators. The era’s quintessential AI achievement was IBM's Deep Blue, which notably became the first computer to defeat the reigning world chess champion, Garry Kasparov. This event was significant not only for demonstrating AI’s potential to match human intelligence in specific tasks but also for eliciting a perception of a new form of intelligence by Kasparov himself.

During the 1990s, the AI community witnessed a philosophical divide. The majority, content with AI's progress in delivering intelligent agents for distinct tasks like chess and speech recognition, contrasted with the founders and other leading figures who aspired towards Human-Level AI (HLAI) or Artificial General Intelligence (AGI). AGI is predicated on the idea that a machine could be capable of performing any cognitive function that a human can. Seth D. Baum from the Global Catastrophic Risk Institute noted more than forty projects dedicated to AGI, including prominent ones like Google’s DeepMind, OpenAI, and The Human Brain Project. Baum also highlighted the bifurcation within AGI research between humanitarian and intellectualist pursuits.

In parallel to these debates in AGI, the 1990s also marked significant improvements in the performance of narrow AI. The decade was crucial for machine learning, a subset of AI reliant on statistical models for pattern recognition in data. Despite the conceptual roots of machine learning not being new, advancements in computing power, in line with Moore’s Law, and the utilisation of large datasets led to notable results. Key developments included the application of backpropagation algorithms in neural network research. Pivotal in this regard was Geoffrey Hinton's 2006 demonstration that deep belief networks could be effectively trained using a greedy layer-wise pre-training approach, a landmark moment that propelled the deep learning movement forward.

Critically, the evolution of AI has been informed by various intellectual perspectives. Hubert Dreyfus, in *What Computers Can't Do* (Dreyfus, 1972) and *What Computers Still Can't Do* (Dreyfus, 1992), provided a critical examination of AI's limitations, challenging the era's dominant optimism. In contrast, Ray Kurzweil, in *The Singularity Is Near* (Kurzweil, 2005), speculated on the transformative potential of AI, envisaging a future where human and machine intelligence converge. Rodney Brooks's *Cambrian Intelligence* (R. A. Brooks, 1999) shifted the focus towards behaviour-based robotics, emphasising the role of environmental interaction, thereby critiquing conventional symbolic AI approaches. Ian Goodfellow, Yoshua Bengio, and Aaron Courville's *Deep Learning* (I. Goodfellow et al., 2016) has emerged as a seminal work, delineating the methodologies underpinning the success of modern machine learning.

In summary, the journey of AI from its inception to the present day encapsulates a series of advancements, setbacks, and philosophical shifts. From the pioneering days of AI in the 1950s, through periods of intense optimism and subsequent winters, to the current landscape marked by profound developments in narrow AI and the pursuit of AGI, the field continues to evolve. These developments are not only technological but deeply intertwined with shifts in theoretical perspectives, methodologies, and the broader understanding of what constitutes intelligence in the context of machines.

Deep Learning

The field of deep learning, a subset of machine learning, represents a significant shift in the approach to machine intelligence. Unlike traditional machine learning techniques, which necessitate manually engineered feature extractors to transform raw data into a detectable or classifiable format, deep learning relies on methods of representation learning. These methods enable a machine to identify patterns in data autonomously, reducing the need for expert-crafted feature extractors (Chollet, 2018; Goodfellow, I. et al., 2016; Russell & Norvig, 2020).

Deep learning's rapid advancement is primarily due to several key factors: the escalation of computing power, the availability of extensive and sophisticated datasets, and improvements in device interconnectivity. This growth has expanded the capabilities of machine intelligence, ranging from the prediction of future events to the autonomous learning of tasks without direct human oversight. To fully appreciate deep learning’s role in contemporary computing systems, both practical applications and experimental projects warrant discussion, as they showcase the operational behaviours of intelligent machines.

In practical terms, deep learning has been seamlessly integrated into many modern computing systems, often creating a symbiotic relationship with users. Noteworthy examples include Microsoft Research's endeavours to enhance clinical decision-making using medical imaging AI, and Google's Neural Machine Translation system, which refines translation accuracy through learning from extensive multilingual datasets.

Furthermore, deep learning has ventured into the realm of creative content generation. Artists and researchers like Mario Klingemann and Robbie Barrat have explored the boundaries of machine intelligence, primarily focusing on experimental applications. This contrasts with more utilitarian applications that typically reside in the domain of supervised learning. These experimental forays often involve unsupervised learning, diverging from the guided learning paths prevalent in practical applications.

Another significant area within machine learning is reinforcement learning, characterised by its pursuit of a balance between exploring new possibilities and exploiting existing knowledge. Sutton and Barto (Sutton & Barto, 2015) describe reinforcement learning as the process of mapping situations to actions to maximise a numerical reward. This paradigm is distinct from supervised learning, which is predominantly used in commercial applications and relies on labeled data for learning. Reinforcement learning, in contrast, learns from its own interactions within an environment, diverging from unsupervised learning's goal of uncovering data patterns. The principal aim in reinforcement learning is to enhance the reward outcome, differentiating it from other learning paradigms.

Deep learning, being a pivotal subset of machine learning, excels in identifying patterns from unprocessed data, thus negating the requirement for manual feature extraction. This significant shift is thoroughly explored in a seminal paper by LeCun, Bengio, and Hinton (LeCun et al., 2015), offering an in-depth understanding of deep learning's core principles and its wide-ranging impact across diverse fields. Deep learning's accelerated development is attributed to factors such as advanced computational capabilities, large-scale data availability, and improved interconnectivity of devices. These advancements have facilitated numerous machine intelligence applications, from predictive analysis to autonomous task learning and creative content generation. Generative Adversarial Networks (GANs), introduced by Goodfellow et al. (I. J. Goodfellow et al., 2014), exemplify this, particularly in content generation, as demonstrated in the artistic outputs of Klingemann and Barrat.

Reinforcement learning represents a unique perspective within machine learning, centering on the equilibrium between exploration and exploitation. This approach differs from supervised learning, which leverages labeled data, and from unsupervised learning, which aims to identify patterns within data. Instead, reinforcement learning's primary driver is the maximisation of a reward signal based on environmental interactions. This method has shown remarkable success in various applications, such as achieving human-like performance in video games through deep reinforcement learning (Mnih et al., 2015). The foundational concepts and algorithms of reinforcement learning are comprehensively outlined in Sutton and Barto's textbook (Sutton & Barto, 2015), providing valuable insights into this dynamic field of study.

Foundational Findings and Limitations

This dissertation presents a comprehensive examination of the potential for symbiotic relationships between humans and intelligent machines, rooted in a multidisciplinary exploration of current advancements and theoretical frameworks. Initially, the research embarked on a thorough investigation of biological symbiotic phenomena, aiming to decipher intricate relationships exemplified by the Hawaiian bobtail squid and Vibrio fischeri, as well as the complex relationships between ants and trees. This intensive study of biological symbiosis, while enlightening in understanding the multifaceted nature of these relationships, ultimately proved tangential to the primary research focus. It became apparent that while these biological examples provided insights into the complexity and patterns of symbiosis, they were less conducive to extrapolating principles relevant to the machine intelligence domain. Consequently, the research direction shifted towards a more theoretical inquiry into symbiosis.

In exploring symbiotic theory, the works of Angela Douglas, Lynn Margulis, and Jan Sapp were instrumental in deepening the understanding of symbiotic dynamics. This theoretical groundwork enabled the identification of common patterns within biological symbiotic relationships, setting the stage for translating these biological interactions to the human-intelligent machine interface. To anchor these insights within a relevant context, the researcher delved into the history of computing, with a particular emphasis on the evolution of machine intelligence. Understanding the historical trajectory of computing illuminated the distinctions between purely biological environments and the emergent hybrid ecosystem of humans and intelligent machines.

An in-depth exploration of machine intelligence, specifically machine learning, was pivotal not only in contextualising the differences between the biological symbiosis and machine intelligence fields but also in laying a foundational understanding of the logic and evolutionary paths in artificial intelligence. This engagement with machine learning theory prepared the researcher for the ensuing practical research phase, fostering a more open and receptive approach.

Reflecting on the initial stages of this thesis, a literature review revealed that most questions concerning the integration of symbiosis and machine intelligence were adequately addressed. However, a significant gap remained regarding the transferability of biological knowledge to the domain of machine intelligence, which encompasses a spectrum of biological and mechanical phenomena. Existing research, like Biomimetics and Material Ecology, typically focuses on emulating natural models to address human challenges or designing collaboratively with nature. In contrast, this thesis seeks not to emulate but to derive inspiration from biological models and nature, applying these insights within digital environments.

During this thesis's development, the literature exploring the transition of complex, biologically-based symbiotic relationships to machine intelligence was scant. It was within the domain of second-order cybernetics, particularly concerning concepts of autonomy, epistemology, and self-organisation in complex systems, where more relevant literature emerged. The seminal works of Humberto Maturana and Francisco Varela, particularly their concept of autopoiesis — self-maintaining biological systems — in "Autopoiesis and Cognition," provided profound insights into living systems 'self-sustaining nature (Maturana & Varela, 1987, 2012). Despite the groundbreaking nature of their work, Maturana and Varela's focus diverged from this thesis's primary intent. The key differences include the objectives of translating knowledge across disparate domains, the analytical approach to deconstructing complex phenomena, and the emphasis on the practical application (praxis) within the realm of machine intelligence.

Contrasting with Maturana and Varela's efforts to delve deeply into the biology and cognition of natural systems, this research is anchored in the practical translation of these concepts to machine intelligence. Maturana and Varela adopt a balanced theoretical and practical perspective, whereas this thesis is heavily inclined towards actionable outcomes and applications. The Chilean biologists seek to understand and describe, whereas the focus here is on deconstructing complexity to foster practical implementation.

To synthesise, this thesis aims to bridge the gap between the sophisticated, often abstract, realms of biological symbiosis and the pragmatic, rapidly evolving field of machine intelligence. Through exploring the parallels, differences, and potential synergies between these domains, it seeks to chart a path forward where the intricate, self-regulating mechanisms of biological systems can inform and enhance human interaction with intelligent machines. This interdisciplinary approach, straddling biology, cybernetics, and artificial intelligence, underscores a vision of a future where technology is not merely a tool, but a partner, dynamically intertwined with human capability and creativity.

Continuing from where we left off, the research journey further unraveled the complexities and potential avenues for bridging biological symbiosis with machine intelligence. The addition of Gregory Bateson's holistic systems theory perspective, as encapsulated in *Steps to an Ecology of Mind*, significantly contributed to this exploration (Bateson, 1972). Bateson’s ideas about communication patterns and the essence of relationships in systems offered a novel lens through which the interplay between biology and technology could be viewed. This perspective was crucial in understanding not just the structural, but also the communicative and relational dynamics that could potentially inform symbiotic relationships between humans and intelligent machines.

Janine Benyus's work in "Biomimicry: Innovation Inspired by Nature" and the evolving field of Material Ecology underscored the aspirations to connect biological insights with technological applications (Benyus, 1997). However, unlike these approaches, which often strive to directly emulate natural models or foster collaboration with nature, this dissertation aimed at an interpretative application of biological principles in the digital domain. This nuanced approach sought to transcend the mere replication of nature, endeavoring instead to reimagine and repurpose biological tenets within the context of machine intelligence.

In navigating the complexities of translating biological systems to technological applications, second-order cybernetics provided a fruitful perspective. The pioneering insights of Heinz von Foerster, particularly in "Understanding Understanding: Essays on Cybernetics and Cognition," illuminated the recursive and self-referential aspects of systems and their interconnections (von Foerster, 2003). This examination of understanding and cognition in systems brought an added dimension to the thesis, aligning well with the overall objective of deciphering and harnessing the essence of biological systems for technological advancement.

This dissertation's journey through various disciplines and theoretical frameworks ultimately illuminated both the challenges and the potential pathways in aligning the principles of biological symbiosis with the domain of machine intelligence. While the fields of Biomimetics and Material Ecology offer a framework for direct emulation and collaboration with nature, this research sought a more interpretative application of these concepts, aiming to integrate the fundamental principles of biological systems into the fabric of digital and machine intelligence environments. The recursive and self-sustaining principles observed in natural systems, as highlighted by the theories of second-order cybernetics, were particularly insightful. These principles suggested new ways of conceptualising and developing intelligent machines that are not only responsive but also adaptive, exhibiting a degree of autonomy and self-regulation akin to biological entities.

The objective was to transcend the traditional boundaries between biology and technology, paving the way for a new class of intelligent systems. These systems, inspired by the adaptability, efficiency, and resilience of biological organisms, could offer innovative solutions to complex challenges, mirroring the harmonious balance found in nature. This thesis, therefore, stands not just as an academic exploration but as a visionary framework for future research and development in the field of symbiotic relationships between humans and intelligent machines. The aim is to foster a paradigm where technology no longer stands apart from the natural world but is seamlessly integrated, contributing to a more sustainable, efficient, and intuitive interaction between humans and the digital realm.

The research encapsulates a multidisciplinary endeavour, traversing the realms of biology, cybernetics, and artificial intelligence. It advocates for a novel synthesis of ideas and practices, inspired by the intricate workings of natural systems, to inform and enhance the development of machine intelligence. Through this integrative approach, the dissertation sets the stage for a future where the symbiosis of humans and intelligent machines transcends conventional interactions, fostering an era of enhanced understanding, collaboration, and innovation.

Chapter 03 / RESEARCH METHODOLOGY

“We find ourselves in a thicket of strategic complexity, surrounded by a dense mist of uncertainty. Though many considerations have been discerned, their details and interrelationships remain unclear and iffy—and there might be other factors we have not even thought of yet”

Nick Borstrom in *Superintelligence*

This chapter provides a comprehensive overview of the research methodology that forms the backbone of this PhD study. The methodology is built upon four foundational pillars: designing, making, experimenting, and reflecting. Each of these pillars plays a crucial role in shaping the research trajectory and outcomes. The primary research approach adopted is "research through design," a concept highlighted and elaborated upon in Christopher Frayling's seminal 1993 paper, "Research in Art and Design." This approach stands out from Frayling's other two modes—research into design and research for design. In "research through design," it is the design practice itself, with its iterative and exploratory nature, that drives the inquiry. In contrast, the latter two modes often encapsulate knowledge that is tacit, embedded within the design artefact, making it less explicit and more intrinsic (Frayling, 1994).

The research's conceptual foundation stems from the researcher's deep-seated interest in translating symbiotic relationships into the realm of machine intelligence. This interest is not isolated but is supplemented and enriched by various principles and theories explored in the literature review. This conceptual base has been the catalyst for a diverse range of research activities, including but not limited to, observations, interviews, and hands-on making. This chapter aims to elucidate the researcher's philosophical stance, present the research questions in detail, list the factors that have played a pivotal role in shaping the design research, and provide an in-depth account of the methods and techniques employed.

Position of the researcher

From an epistemological and philosophical standpoint, this research is firmly anchored in constructivism. The researcher posits that knowledge isn't a static entity that individuals passively receive. Instead, knowledge is a dynamic construct, actively shaped and reshaped by individuals based on their interactions with the world around them. As individuals encounter new experiences, they interpret, assimilate, and sometimes challenge this new information based on their prior experiences and knowledge structures. Constructivist researchers often employ qualitative methods, delving deep into open-ended scenarios, pioneering techniques, and seeking to understand the nuances and complexities of human experience.

Research Questions

The research's constructivist foundation implies that knowledge was not just theoretically gained but was experientially internalised and subsequently reflected upon. As the study progressed and evolved, the researcher recognised the imperative need to refine its scope, ensuring that the research remained focused and relevant. The research questions encapsulate this refinement process, each one delving deeper into the intricacies of the research topic:

RQ01. How might symbiotic relationships observed in nature inspire and inform the design and development of machine intelligence systems?

RQ02. What are the potential capabilities of applying principles of biological symbiosis to human-machine interactions to enhance collaboration effectiveness?

RQ03. How might symbiotic design principles serve to address and ameliorate existing limitations within Artificial Intelligence algorithms and systems?

These questions, henceforth referred to as **RQ01, RQ02,** and **RQ03**, serve as the foundational pillars underpinning the research's findings, projects, and contributions.

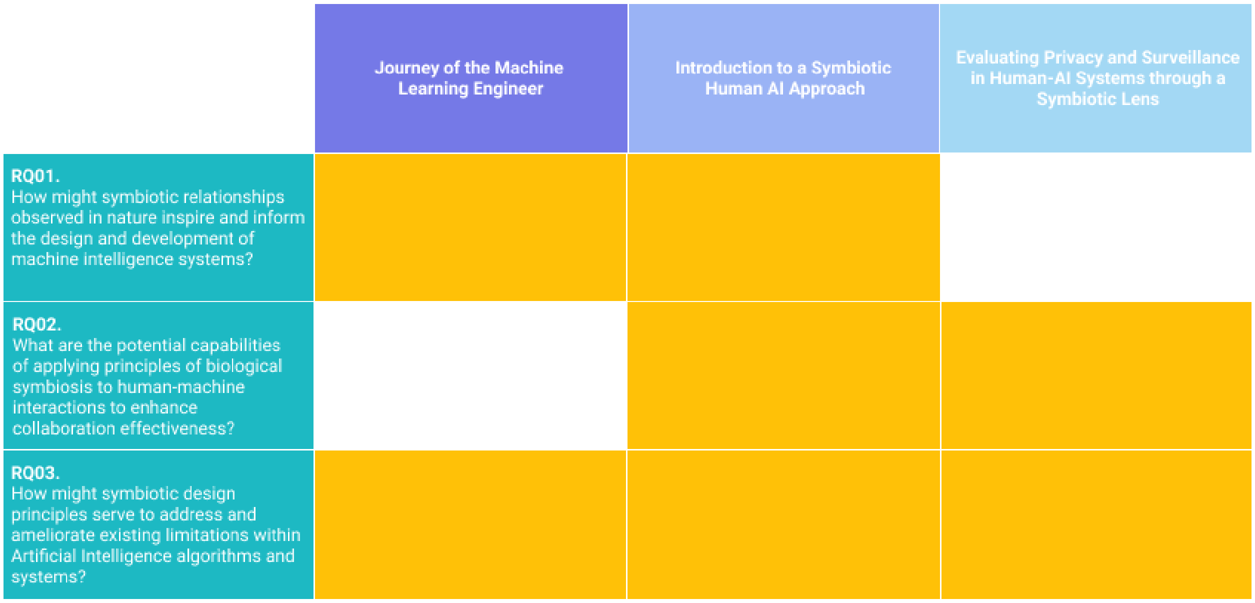


Figure 9 - Expected connection between the research questions and the practical research projects of this PhD

Target Audience

This PhD is tailored primarily for researchers navigating the intricate landscape of machine intelligence. This encompasses not just design researchers but also specialists in human-computer interaction, social researchers, and even those who might not have a traditional background in computer science or statistics. The researcher aspires for this work to not just inform but to galvanise and inspire future design research in human-machine symbiosis. Biological interactions, refined over countless lifecycles and marked by continuous negotiations between organisms, offer a treasure trove of insights. Transferring knowledge across research domains can bolster the interdisciplinary essence of design research, fostering a robust, collaborative, and multidisciplinary ethos.

Importance of Design Research

This study delves into the dynamic and expansive terrain of machine intelligence, examining it through the lens of design research. This approach is particularly relevant to the scope of this thesis, as design research straddles both practical and theoretical realms. It offers a permeable methodology, allowing for the amalgamation and application of diverse knowledge, methods, and cultural perspectives.

The genesis of design research dates back to the 1960s in London, a critical juncture for the discipline marked by the establishment of the Design Research Society and the Design Research department at the Royal College of Art, under Bruce Archer's guidance. In the ensuing decades, design research began to carve out its distinct identity, focusing on design-specific methodologies and areas of interest, thus differentiating itself from other academic disciplines (Joost et al., 2016; Michel, 2007).

Three fundamental concepts of design research significantly inform this thesis:

1. **Tacit Knowledge**: As introduced in Michael Polanyi's "Personal Knowledge," tacit knowledge refers to insights that are intrinsically difficult to articulate, being deeply embedded in personal experiences and intuition. This type of knowledge, while elusive, is pivotal in shaping the outcomes of research endeavours (Polanyi, 1958, 1966).

2. **Reflective Practice:** Echoing Polanyi's concept, Donald Schön, in "The Reflective Practitioner," suggests that practitioners often hold deeper knowledge than they can

readily express. This knowledge, rooted in their intuitive understanding and reflection during practice, enables them to address complex problems in ways that might not be immediately explicable. This reflective practice not only aids in problem-solving but also helps in expanding the boundaries of their field, introducing elements of surprise and innovation (Schön, 2017).

3. **Research Through Design:** Christopher Frayling's framework, outlined in 1994, distinguishes between different research modes in art and design. "Research through design" is particularly notable, aiming to generate new knowledge by proposing alternatives to the current state of affairs in research. This iterative process involves deep engagement with practice, leading to a continuous cycle of reflection and improvement.

These principles highlight the suitability of design research for exploring complex, ambiguous, and 'fuzzy' research areas. It requires continuous reflection, embracing a hands-on approach to research and uncovering alternative insights and methodologies. This research's alignment with action research and grounded theory further underscores design research's affinity for creativity, iterative processes, and experimentation. In line with Ashley Hall's observation, design research doesn't just contribute to the innovation of new products; it also leads the way in proposing future directions and possibilities (Hall, 2011).

In this context, the study investigates machine intelligence from a design research perspective. This approach enables the integration and practical application of varied knowledge forms, bridging the gap between practice and theory. Rooted in the vibrant intellectual environment of 1960s London, design research has evolved to emphasise the unique methods and interests of design. This evolution highlights the discipline's distinctiveness and its capacity to address research questions characterised by complexity and fluidity.

Overall, design research's flexibility and its focus on iterative, creative exploration make it an ideal framework for studying the evolving field of machine intelligence. This study aims to leverage these attributes, seeking to shed new light on this rapidly advancing area and to propose forward-thinking solutions and insights.

PhD as a Design Programme Framework

The methodology anchoring this PhD research is inspired by Johan Redström's concept of a design programme. Redström explains that although the term 'program' finds its use in diverse contexts ranging from computer software to theoretical frameworks, it consistently denotes the intention and structure underpinning a process poised for evolution. In this vein, a design programme constitutes the foundational bedrock, sculpting the trajectory of research practice. This bedrock is dynamic, constantly strengthened and reshaped through an iterative process of reflection within and upon the practice itself. The design programme of this PhD is delineated into three interlinked phases: (1) Composing the design programme, (2) Undertaking design experimentation, and (3) Reflecting on the outcomes to refine the design programme (Redström, 2017).

The composition of the design programme was guided by a set of principles: (1) a belief in harnessing knowledge from biological interactions to enhance machine intelligence, (2) addressing emergent concerns in artificial intelligence research, such as inherent biases and under examined scientific bases (Crawford et al., 2019), and (3) exploring philosophical and speculative narratives on human-machine symbiosis and advanced artificial intelligence (Asimov, 2004; Bostrom, 2014; Dick, 1968; Tegmark, 2017).

Design experimentation is the cornerstone of this design programme. These experiments underpin the programme's foundational work, following a sequential progression where theoretical insights, both internal and external to the field, influence the design process.

This PhD research integrates design experimentation with reflective practice, creating a synergistic relationship. Each round of experimentation leads to a reflection phase, subsequently informing the refinement of the design programme or initiating further experimental iterations.

This framework represents a distinctive approach to research in the field, intertwining the principles of design with the methodology of a doctoral study, thereby encapsulating the innovative spirit and intellectual rigour characteristic of PhD research.

Research through Design Practice

While the chapter's outset introduced 'research through design' as the chosen research modality for this PhD, it's essential to delve deeper into its mechanics. What does 'research through design' entail, and what are its defining characteristics? The primary research techniques employed encompass making, observing, and conducting expert interviews. These methods, however, were not static; their application varied across different research stages.

The research is structured around three pivotal pillars: (1) the journey of the Machine Learning engineer, (2) an exploration of symbiotic relationships and their primers, and (3) the implementation of a symbiotic framework in contextual scenarios. The first and third pillars are rooted in the practice of making and subsequent reflection. Expert interviews further bolster these two pillars. In contrast, the second pillar diverges, primarily relying on observation and expert interviews. Initially, the plan for the second pillar was to conduct workshops followed by expert interviews. However, the unforeseen challenges posed by the COVID-19 pandemic necessitated adaptability, leading to modifications in the research approach.

Research Methodology

This research adopts a mixed methods approach, employing both qualitative and quantitative techniques to gain a comprehensive understanding of the topic. The practical research projects will integrate qualitative exploration with quantitative support.

In essence, qualitative data will provide depth and richness, while quantitative data will offer breadth and generalisability. This interplay between the two types of techniques will lead to a deeper and more nuanced understanding of the research area.

Literature supports mixed methods approaches, highlighting several advantages over single-method research designs. One primary advantage is triangulation. By combining qualitative and quantitative insights, the research can achieve cross-verification of findings, enhancing its validity and reliability (Creswell & Clark, 2011). Another advantage is completeness. Integrating different forms of data provides a more holistic and comprehensive picture of the research topic, enabling a deeper level of understanding than would be possible with a single method (Greene et al., 1989). Mixed methods approaches also offer contextualization. Qualitative data can aid in interpreting and explaining quantitative findings, and quantitative data can ground qualitative insights by offering measurable trends (Tashakkori & Teddlie, 2003). Finally, mixed methods can offset weaknesses. The strengths of one methodology can address the shortcomings of the other, leading to more robust conclusions (Johnson et al., 2007). Central to this mixed methods approach is the emphasis on an action research methodology, informed by a "learning by doing" philosophy and drawing on the work of Donald Schön (2017) in reflective practice. Action research can be considered a cyclical, iterative process that is often described as a form of systematic inquiry. Through repeated cycles of action, critical reflection, and refinement, practitioners can inquire into their practices and subsequently improve both their practices and the environments in which they operate. This cyclical nature allows for a systematic approach to learning and improvement, aligning with Schön's concept of reflection-in-action, where researchers learn and adapt their methods as they progress through the research process.

By incorporating Schön's framework of reflective practice, the research ensures a continuous feedback loop. Qualitative data collected throughout the research will be used to inform and refine subsequent actions and data collection methods. The iterative process fosters a deeper understanding of the research questions and allows for ongoing adjustments to ensure the research remains relevant and addresses the complexities of the topic at hand.

Kurt Lewin, in the 1940s, pioneered the first action research model as a response to social action challenges. He posited that without collaborative endeavours, there was a disconnect between theory and application, with researchers formulating theories devoid of practical application and practitioners operating without the requisite theoretical knowledge. Lewin's action research methodology thus provided a framework for practitioners to research their actions, enhancing effectiveness and concurrently developing theories of social action. His contributions underscored the potential of action researchers to offer pertinent, context-specific knowledge and solutions to issues rooted in social action. Moreover, the iterative nature of action research, which involves identifying problems and then undertaking actionable experiments grounded in research, has been shown to yield innovative solutions (Adelman, 1993; Dickens & Watkins, 1999).

Before delving into the rationale for selecting action research, it's essential to elucidate its definition and the core concepts that characterise this methodology. While a universal definition remains elusive, numerous scholars have extolled the virtues of action research. Notably, Stephen Kemmis and Robin McTaggart describe it as a collective self-inquiry that participants employ to refine the rationale behind their practices and deepen their understanding of these practices. Collaboration is a cornerstone of this definition (Altrichter et al., 2002; Kemmis, 2009; Kemmis et al., 2014). Conversely, Peter Reason and Hilary Bradbury perceive action research as a tool to inform and shape practice, encompassing a spectrum of approaches that, despite their differences, all aim to engage with or enhance specific research fields (Reason & Bradbury, 2008). Anne Burns offers a nuanced perspective, defining action research as a set of approaches marked by self-reflection, systematicity, and criticality, geared towards identifying issues deemed worthy of investigation by participants to effect critically informed changes in practice (J. D. Brown & Coombe, 2015). Despite the lack of a unanimous definition, proponents concur on the fundamental tenets of action research, which for this thesis include:

- Endeavouring to deepen understanding of the research domain,
- Enhancing existing practices,
- Identifying and addressing challenges,
- Recognising the sequential and cyclical nature of the research process,
- Welcoming change and adaptations, and
- The potential of action to foster theory development.

Action research is often conflated with Participatory Design due to their shared participatory attributes, yet they diverge in terms of strategies and objectives. Both methodologies, with their distinct strategies and goals, can coexist harmoniously within a single research project. Marcus Foth and Jeff Axup delineate the parallels and disparities between participatory design and action research by juxtaposing two research endeavours: one rooted in Action Research methodology and the other in Participatory Design methodology (Foth & Axup, 2006; Kagan et al., 2006). For this research, the iterative essence of reflection intertwined with practice inherent in action research was deemed more congruent, especially given the research's focus on immersion in practice and subsequent reflection. Participatory Design would augment the research if there existed a robust knowledge base on symbiotic relationships in Human-AI interaction. Given the current knowledge gap, the researcher prioritised immersion in practice and reflective contemplation over ethnographic and *in-situ* methodologies.

Several factors underscore the alignment of action research with this research endeavour. Firstly, action research emphasises the enhancement of practice through action, evaluation, and critical reflection, resonating with the ethos of learning by doing and facilitating reflection on both the positive and negative experiences encountered during the research process. Secondly, its participative and collaborative nature aligns with the research's emphasis on personal development achieved through direct and indirect engagement with experts in machine intelligence and symbiosis. Thirdly, the context-specific nature of action research mirrors the contextual nuances of symbiotic relationships. Lastly, the knowledge generation process in action research is dynamic, with insights emerging as actions unfold.

The act of reflection post-action is emblematic of action research and is mirrored in this research approach. Following a cycle of sustained action, the researcher allocates time to assimilate the accrued knowledge and reflect on the insights gleaned. Analogous to this research, the essence of action research is experiential learning, culminating in personal or professional growth. The action research spiral is deliberately integrated into this research methodology. Within this spiral framework, the researcher initiates the process by planning a change, implementing it, and subsequently reflecting on the process and its outcomes. This modus operandi is evident across the trio of practical projects that constitute this thesis (Reason & Bradbury, 2008; Sharp, 2019).

Complementing action research in this mixed methods research project is the inclusion of constructivist grounded theory. While action research emphasises experiential learning, grounded theory is geared towards theory construction through data collection and analysis. The synergy of these methodologies fortifies the research foundation, enabling the researcher to glean insights from practice while concurrently formulating and validating theories or hypotheses. Grounded theory is distinguished by its structured flexibility, thriving in research domains where limited knowledge exists about a specific phenomenon or interaction. Its utility in design research stems from its capacity to generate theories anchored in qualitative data.

Grounded theory encompasses three primary methodological genres: traditional grounded theory, evolved grounded theory, and constructivist grounded theory. Traditional grounded theory seeks to formulate conceptual theories that encapsulate pertinent behavioural patterns. Evolved grounded theory, building on its traditional counterpart, introduces the concept of symbolic interactionism, which pertains to the subjective meanings individuals ascribe to objects, behaviours, or events they perceive as authentic. This thesis employs constructivist grounded theory, which, while building on the preceding genres, emphasises the processes through which individuals construct meaning. Distinctively, constructivist grounded theory, with its pragmatic orientation, acknowledges that theories are researcher-constructed rather than discovered (Birks & Mills, 2012; Bryant & Charmaz, 2007; Charmaz, 2014).

As alluded to earlier, grounded theory, akin to action research, is non-linear, characterised by its iterative and recursive nature. A typical grounded theory cycle encompasses technique sampling, succeeded by data collection via surveys or interviews. Subsequently, the data undergoes multiple coding stages with the objective of theory construction. The integrity and rigour of grounded theory are paramount for theory construction, with the quality of the resultant theory being contingent on three determinants: the researcher's expertise and research acumen, the researcher's positioning, and the precision of the deployed research methodologies.

The research methodology delineated above encapsulates the depth and breadth requisite for this research. The amalgamation of action research and constructivist grounded theory seemingly addresses the comprehensive requirements of this project. Given the research's focus on concept translation across domains, the "learn by doing" paradigm emerges as the optimal approach, facilitating the researcher's immersion in the research domain while concurrently accruing practical experience. Action research emerged as the most congruent methodology, covering the aforementioned dimensions. The dual emphasis on action and concurrent research was deemed the most effective approach to explore biological phenomena, translate this knowledge, and integrate it into the realm of Artificial Intelligence. The flexibility to conduct research while actively engaging in the field enabled the researcher to rapidly assimilate information, reflect upon it, and gain practical experience. Additionally, the research incorporates constructivist grounded theory. Among the trio of prevalent grounded theory methods, constructivist grounded theory was identified as the most fitting for this project. This methodology enables the researcher to concurrently function as a participant in the research process. Given the researcher's endeavour to grasp the intricacies of machine learning, it was deemed appropriate for him to be perceived as a participant in the research journey.

The symbiotic relationship between action research and constructivist grounded theory was identified as the optimal methodological approach for this project. Firstly, both methodologies emphasise iterative information collection and reflection. Secondly, both are tailored for qualitative research. Thirdly, their complementary nature enhances the research's quality. Constructivist grounded theory revolves around iterative cycles of data collection, review, and repetition until concepts materialise. Concurrently, action research is predicated on planning, action, and reflection cycles. The continuous iterations of information assimilation, practical task execution, and reflection not only expedited the research's information acquisition but also honed the research focus as it progressed.

However, the confluence of these methodologies is not devoid of limitations. The primary limitation is the relative neglect of quantifiable data. In the context of translating information across research domains, a quantitative approach could bolster the research's foundation. The second limitation pertains to the time-intensive nature of these methodologies, which are best suited for small-scale studies. Consequently, despite the cyclical nature of the methodology, time constraints limit the number of cycles that can be executed.

Research Assumptions

Like many research endeavours in machine intelligence, this study is rooted in research assumptions and the personal biases of the researcher. It's important to view these assumptions not as detrimental but as foundational stepping stones for reasoning and action. Research assumptions generally emerge from the initial understanding of the research problem. A critical part of the research process involves distinguishing between what constitutes factual information and what remains assumptive. This discernment, often guided by experience and immersion in the research field, is crucial. Direct engagement in the research space aids in distinguishing facts from assumptions, though it rarely provides definitive answers. Consequently, research assumptions persist throughout the study, with some becoming evident only upon its conclusion.

Research assumptions originate from various factors, including but not limited to:

- Incomplete knowledge of a specific research area.
- Efforts to simplify complex problems.
- Creating constraints to manage the scope of research.
- Focusing on general rather than specific research aspects.
- Availability of resources.
- The research culture encompassing the researcher and the field.

One significant assumption in this research stems from the necessity to simplify a complex problem, leading to the imposition of constraints to narrow the research scope. Symbiotic relationships, as previous literature indicates, are intricate and context-sensitive, prompting the need to define specific research boundaries. The researcher, in this case, relied on symbiotic theories rather than exhaustive biological interactions, entrusting earlier research to provide a solid framework for exploration.

Another initial assumption related to the researcher's unfamiliarity with machine intelligence. At the onset, limited understanding and resources informed certain presumptions. However, as the researcher deepened their technical knowledge, many of these assumptions dissipated. For instance, the researcher initially had a rudimentary grasp of data's role in machine intelligence, not fully appreciating the nuances of data collection, cleaning, and formatting for machine learning algorithms.

The research environment also significantly shaped assumptions. The researcher's background in Human-Computer Interaction and interactions with peers from diverse fields and cultural backgrounds inevitably influenced the study. These interactions underscore the perspective that assumptions, while unavoidable, are not inherently negative if acknowledged and understood within the research context.

Additionally, the researcher's personal biases, shaped by life experiences and professional milieu, play a role. Compared to much of the global population, the researcher belongs to a distinct, technologically adept academic group. This group understands the subjectivity of machine learning models and the biases inherent in human-curated datasets. Acknowledging this, the researcher endeavoured to mitigate inherent biases by explicating concepts thoroughly to participants, who, while sharing similar backgrounds, brought their distinct biases to the research.

Moreover, the researcher's academic orientation, straddling design and engineering, represents another bias. The potential to disproportionately favor one discipline over the other exists, and recognising this, the researcher strives for balance and neutrality between these fields.

In sum, this research project, like all scholarly inquiries, is influenced by a blend of assumptions and biases. Identifying and understanding these elements are pivotal to comprehensively grasping and conducting nuanced research in machine intelligence.

Antithesis

The effectiveness of a thesis defence is closely tied to the thoroughness with which it addresses its counterargument. A counterargument provides critical examination and contradiction of the main thesis, and rigorously engaging with these challenges is essential for the credibility of the research. This thesis specifically intends to articulate the primary arguments of the counterargument, underpinned by several crucial considerations.

At the core of the counterargument lies the assertion that investigating biological interactions may not yield significant insights into the symbiotic relationships between humans and machine intelligence. The basis of this viewpoint is that biological interactions, inherently organic, might not translate effectively into hybrid environments where organic and digital entities interact.

This counterargument is reinforced by several key factors:

1. **Machine Intelligence Foundations**: Machine intelligence has been developed based on human reasoning and engineered using human-derived knowledge. Biological interactions, governed by a different logic, may not align or integrate seamlessly within

mechanical contexts. This discrepancy suggests a potential mismatch in applying biological principles to the domain of machine intelligence.

2. **Complex Nature of Biological Interactions**: The intricacies of biological interactions, which are context-sensitive and replete with variables, add layers of complexity. Theories conceptualising symbiosis are, essentially, simplifications of these intricate dynamics. When such simplifications are applied to machine intelligence, there's a risk of losing critical nuances and details, potentially leading to ineffectiveness or misrepresentation.

3. **Dynamic Evolution of Machine Intelligence**: The current landscape of machine intelligence is continually evolving, marked by rapid advancements. This dynamic nature implies that the issues currently identified, particularly those concerning contextual relevance, may become outdated or irrelevant as machine intelligence progresses.

4. **Unaddressed Negative Implications of AI**: The thesis recognises both positive and negative aspects of human-AI interaction. However, it does not comprehensively cover all negative impacts. Notably, issues such as algorithmic bias, improper data usage and ownership, foundational errors in AI construction, and a range of technical vulnerabilities – including data poisoning, backdoor exploits, and susceptibilities in transfer learning – are insufficiently addressed.

The defence of a thesis is substantiated not only by the arguments it presents but also by its capacity to critically engage with and counterbalance the counterarguments. This engagement enhances the thesis's depth, demonstrates the researcher's understanding of the subject's complexity, and strengthens the overall validity of the research findings.

Ethical Considerations in Machine Intelligence

In the realm of machine intelligence research, it is imperative to delve into the ethical dimensions. The profound influence that Artificial Intelligence and its related fields are poised to exert on humanity's future necessitates a thorough examination of the ethical boundaries. This encompasses addressing pivotal questions about the capabilities of intelligent machines, their appropriate applications, the inherent risks, and strategies for risk mitigation. While humans navigate their actions based on a moral compass, intelligent machines operate devoid of such guidance, relying solely on the data they are fed. Intelligent machines' lack of consciousness and empathy underscores the urgency of addressing ethics in AI. This becomes even more pertinent given this research's focus on the evolving, long-term relationships between humans and intelligent machines, inspired by biological symbiotic interactions. Consequently, it is essential that these machines are designed to eschew actions detrimental to their human counterparts.

The ramifications of neglecting ethics in Artificial Intelligence development are manifold. Many existing Artificial Intelligence systems, unfortunately, pose direct or indirect harm to their human users. One of the paramount ethical dilemmas in Artificial Intelligence application is the balance between privacy and surveillance (Bostrom, 2014; Coeckelbergh, 2020; Crawford et al., 2019; Liao, 2020). Given that Artificial Intelligence systems thrive on extensive data collection and analysis, there's an ethical imperative to safeguard individual privacy and uphold their right to confidentiality. It is essential that individuals are informed about data collection, its purpose, and the entities accessing their data. Another ethical quandary is the potential for behaviour manipulation by Artificial Intelligence. As Artificial Intelligence systems amass data through frequent interactions, they gain profound insights into individual behaviours, which could be exploited to target their vulnerabilities (Bostrom, 2014; Crawford et al., 2019; Shanahan, 2015; Tegmark, 2017).

The opacity and inherent biases of Artificial Intelligence systems further compound the ethical challenges. Artificial Intelligence systems, especially those designed for automated decision support and predictive analysis, have sparked concerns over transparency, accountability, and auditability. The black-box nature of many Artificial Intelligence systems makes it challenging, even for experts, to discern the decision-making process. Furthermore, the potential for bias in Artificial Intelligence, stemming from data labelled by individuals with their own inherent biases—whether societal, cognitive, or statistical—cannot be overlooked (Broussard, 2018; Campolo et al., 2017; Crawford et al., 2019; Noble, 2018; O'Neil, 2016).

Machine ethics, which advocates for a shift in perception of machines from mere objects to subjects, presents another ethical frontier (Anderson & Anderson, 2011; Moor, 2006; Wallach & Allen, 2009). This perspective, though ambitious, is crucial, as highlighted in various science fiction narratives. Authors like Isaac Asimov and Philip K. Dick have underscored the potential repercussions of neglecting machine ethics, emphasising the need for principles of inclusion, fairness, and understanding in AI (Asimov, 2004; Dick, 1968; Murphy & Woods, 2009). Shoshana Zuboff's work on "surveillance capitalism" (Zuboff, 2019) sheds light on the potential dangers of Artificial Intelligence systems that prioritise data collection and control over human well-being. Similarly, Luciano Floridi's "onlife" approach (Floridi, 2014) emphasises the blurred lines between online and offline realities in the digital age. This perspective is vital when considering the long-term, symbiotic relationships envisioned in this research, ensuring that Artificial Intelligence development fosters a positive and balanced human-machine co-existence.

Moving forward with the practical elements of this research, the researcher proposes to conduct case studies of intelligent machine systems designed for human-machine collaboration, adhering to ethical principles outlined in relevant literature. This approach will involve identifying potential areas for improvement in the ethical implementation of these systems. The approach will also provide concrete insights into the practical challenges and potential solutions for implementing ethical Artificial Intelligence in real-world applications. Additionally, the researcher will foster a collaborative approach to addressing ethical challenges in Artificial Intelligence by engaging in continuous dialogue with relevant stakeholders. The collaborative approach will involve direct engagement with experts in their fields and indirect engagement through literature review. This comprehensive approach will ensure that the derived research outcomes are both practical and ethically sound, gathering wider support from relevant communities.

By implementing these practice elements, this research aims to contribute not only to the theoretical understanding of ethics in Artificial Intelligence but also to the development of practical solutions for ensuring the responsible and beneficial development of Artificial Intelligence systems, particularly in the context of long-term human-machine relationships.

Chapter 04 / Journey of the Machine Learning Design Researcher

"In a forest of a hundred thousand trees, no two leaves are alike. And no two journeys along the same path are alike"

Paulo Coelho in *Aleph*

Machine Learning: Bridging the Gap Between Theory and Application

The central tenet of this research was to demystify machine intelligence, specifically focusing on the realm of Machine Learning. Common misconceptions often portray Machine Learning as an advanced, robot-like entity capable of executing a vast array of tasks. Popular culture tends to oscillate between benign and dystopian representations, like WALL-E and the Terminator. However, Machine Learning transcends these simplistic archetypes. Tracing its origins to the 1950s, the term "Machine Learning" was coined by Arthur Samuel. Tom M. Mitchell offered its widely acknowledged definition, positing that a computer program is said to learn from experience \(E\) concerning a class of tasks \(T\) and performance measure \(P\), if its performance at tasks in \(T\), as measured by \(P\), improves with experience \(E\). The initial significant application of Machine Learning in the 1990s was the spam filter used in email systems. Currently, Machine Learning underlies a vast array of products and features we regularly engage with, such as recommendation systems and voice command technologies (Mitchell, 1997).

This chapter aims to bridge the gap between theoretical understanding and practical applications of Machine Learning. It endeavours to highlight the integral roles of design and research within the Machine Learning process, clarifying where Machine Learning begins and ends. It seeks to delineate what it means for a machine to learn, and the complete workflow involved in delivering a Machine Learning application to end-users. Design researchers should possess an in-depth understanding of intelligent systems' capabilities and limitations to navigate the terrain of machine intelligence effectively.

While the intersection between designers and machine learning engineers is growing, the bulk of design involvement traditionally remains within the front-end aspect of projects. This bias towards the front-end, often driven by the engineering-heavy nature of the back-end development, can marginalise a designer's contribution in areas where their insights could be crucial. Understanding this dynamic is critical for enhancing the synergy between design and machine learning. Hence, this section not only explicates essential concepts and workflows in Machine Learning but also discusses the resultant challenges and opportunities for designers and researchers.

Advancing the Discourse in Design and Artificial Intelligence:

It's imperative to comprehend the technical complexities associated with Artificial Intelligence, particularly in the context of design research. Contrary to the assumption that technical details of intelligent machines hold little value in design-focused discussions, the converse is true. The research adopts a constructivist stance, signifying that learning about machine intelligence isn't an end in itself. Rather, it's a means to better understand and reflect on the requirements and barriers experienced by designers and others who are not deeply entrenched in Artificial Intelligence. Every Machine Learning algorithm encountered and understood by the researcher provided a fresh perspective. This continuous acquisition of knowledge allowed for reflection on the current limitations and unrealised potential within intelligent machines, sparking ideas on how biological interactions and paradigms could be translated into the field of Artificial Intelligence through thoughtful design research.

The journey into the essence of Machine Learning algorithms illuminates the gap between mere theoretical knowledge and practical, applied wisdom in the field of AI. It underscores the need for a seamless integration of design principles in the development and implementation of AI technologies. Through this exploration, the research aims to foster a more holistic view of AI, where the integration of design and technical know-how can lead to more user-centric, ethical, and sustainable AI solutions. This balanced perspective is vital in advancing the discourse in both design and AI, paving the way for innovations that are not only technologically advanced but also socially responsible and attuned to human needs.

Project Description

In the first two years of my research journey, I transitioned into a role as a Machine Learning design researcher, embracing a range of tasks in this field. Initially, Artificial Intelligence (AI) presented itself as a labyrinthine and formidable domain. Recognising the complexity of AI, I was determined not to discuss the concept of symbiotic relationships with intelligent machines without first acquiring a foundational understanding of machine intelligence. This immersion into the world of Machine Learning was twofold in purpose: firstly, to highlight the significance of design researchers within technically demanding fields, and secondly, to elucidate the intricacies of Machine Learning, thus making it more accessible to upcoming designers and researchers.

One of the primary challenges I encountered was identifying the appropriate contexts for applying Machine Learning. This technology is versatile, finding relevance across a myriad of applications ranging from Computer Vision to Natural Language Processing. It particularly shines in situations where conventional programming approaches fall short. Machine Learning is adept at addressing several types of challenges:

1. Problems where existing solutions need extensive fine-tuning or are bound by a lengthy list of rules. An example is spam filtering, which requires constant updating and refinement to be effective.

2. Tackling complex issues where traditional software engineering methods fail to deliver robust solutions, such as in speech recognition technologies.

3. Situations involving dynamic or context-sensitive environments. This aligns with the concept of human-machine symbiosis, where machines adapt and respond to changing human needs and contexts.

4. The extraction of insights from large and complicated datasets, often referred to as data mining. This process is crucial in identifying patterns and making informed decisions based on vast amounts of data.

With a clearer understanding of Machine Learning's applications, my focus shifted to exploring the classification of Machine Learning systems. Broadly, these systems can be divided into three categories: Machine Learning Paradigms, Learning Methods, and Methods of Generalisation. Each category offers a unique framework and approach, contributing to the overarching effectiveness and adaptability of Machine Learning in solving diverse and complex problems.

Machine Learning Paradigms : Categories and Applications

The categorisation of training methods in Machine Learning (ML) is primarily based on the level of human supervision involved in training algorithms. These methods are classified into four primary categories: supervised learning, unsupervised learning, semi-supervised learning, and Reinforcement Learning. Each category has unique characteristics, applications, and methods, playing crucial roles in advancing various ML applications.

Supervised Learning

Supervised learning is predominantly used in classification and regression tasks. This approach involves training the algorithm with a labeled dataset, where each example in the training set includes the desired output (label). A classic example of a supervised classification task is an email spam filter. The algorithm is trained with a dataset containing emails labeled as either spam or non-spam (ham). This training enables the system to classify new emails based on learned patterns. For regression tasks, supervised learning predicts a continuous value, such as estimating the market price of a smartphone, based on various features. Key algorithms in supervised learning include k-Nearest Neighbours (k-NN), Linear Regression, Logistic Regression, Support Vector Machines (SVMs), Decision Trees, Random Forests, and Neural Networks (Chollet, 2018; Géron, 2019; I. Goodfellow et al., 2016; Russell & Norvig, 2020b).

Unsupervised Learning

Unsupervised learning differs from supervised learning in that it deals with unlabelled data. The ML system learns patterns and structures from the data without explicit instructions. This category includes:

- **Clustering Algorithms** (e.g., K-Means, Hierarchical Cluster Analysis): These algorithms group sets of objects based on similarities, aiding in tasks like customer segmentation and data organisation.
- **Anomaly and Novelty Detection Algorithms** (e.g., One-Class SVMs, Isolation Forest): These are used to identify unusual data points or activities, such as detecting fraudulent credit card transactions.
- **Visualisation and Dimensionality Reduction Algorithms** (e.g., PCA, t-SNE): These tools help simplify high-dimensional data, making analysis and visual interpretation more manageable.
- **Association Rule Learning Algorithms** (e.g., Apriori, Eclat): These algorithms discover interesting relationships between variables in large datasets, like uncovering shopping patterns from transaction data.

Semi-Supervised Learning

Semi-supervised learning occupies the middle ground between supervised and unsupervised learning. This approach is used when the dataset is partially labeled. Labelling data can be resource-intensive; hence, semi-supervised algorithms, which combine elements of both supervised and unsupervised learning, are particularly valuable. For instance, Deep Belief Networks utilise both labeled and unlabelled data to improve learning accuracy.

Reinforcement Learning

Reinforcement Learning (RL) is a complex yet powerful ML paradigm, often likened to the way biological entities learn from interaction with their environment. RL is applied in diverse fields such as robotics, recommendation systems, and autonomous vehicles. A famous example is AlphaGo by DeepMind, which defeated the world champion in the board game Go. AlphaGo's learning approach involved playing numerous games against itself, progressively improving its strategy.

RL involves an agent that makes decisions based on the state of the environment and receives rewards for actions. The goal is to maximise the cumulative reward. RL can be subdivided into:

- **Model-Based Reinforcement Learning**: The agent utilises a model of the environment for decision-making.
- **Model-Free Reinforcement Learning**: The agent learns directly from interactions with the environment without a predefined model. This includes:
    - **Action-Utility Learning** (e.g., Q-learning): Here, the agent learns a function (Q-function) to estimate the value of actions taken in different states.
    - **Policy Search**: The agent directly learns the policy mapping states to actions.

Understanding these categories is essential for Machine Learning practitioners and researchers. The choice of learning method significantly impacts algorithm design, effectiveness, and the scope of its application. Each method offers distinct advantages and poses specific challenges, requiring careful consideration and expertise in deployment. The dynamic and evolving nature of ML demands continuous exploration and optimisation of these training methods to achieve the best outcomes in various practical applications, from simple classification tasks to complex, autonomous decision-making systems.

Learning Methods

Machine Learning (ML) systems can be distinguished by their learning methods, broadly categorised into Batch Learning and Online Learning. Each approach has distinct characteristics, advantages, and limitations.

- **Batch Learning**: In Batch Learning, the system is comprehensively trained using the entire dataset available. The data is not incrementally introduced but provided in one large batch. This method typically requires the process to be conducted offline, attributing to its extensive duration and substantial computing resources demand. Following the training, the system is deployed for production. To update the system with new data, a complete retraining is necessary, leading to the older version being replaced with an updated one. Despite its thoroughness, Batch Learning's time and resource-intensive nature limit its agility and flexibility in adapting to new data or evolving data trends.

- **Online Learning**: In contrast, Online Learning introduces data to the system in smaller, manageable batches or even continuously. This incremental approach is not only more agile but also more resource-efficient than Batch Learning. It's particularly beneficial for systems that need to adapt rapidly to new data streams or changing data patterns. Online Learning systems can quickly incorporate new information, making them ideal for dynamic environments. Additionally, this method offers engineers the flexibility to adjust the learning rate—a crucial hyperparameter in optimising ML algorithms. However, its efficiency comes with the caveat of vulnerability to the quality of incoming data. If low-quality or erroneous data are fed into the system, it could progressively degrade the system's performance, necessitating robust mechanisms to ensure data integrity and quality (Géron, 2019).

Understanding these learning methods is crucial for ML practitioners, as the choice between Batch and Online Learning significantly influences system design, performance, and maintenance. The decision hinges on various factors, including the nature of the task, data availability, system requirements, and operational context. Therefore, selecting the appropriate learning method becomes a foundational decision in the development of effective and efficient ML systems.

Generalisation Methods

The classification of Machine Learning (ML) systems based on their data generalization methods forms an essential aspect of understanding their functionality and application. This segment delineates the key categorisation approaches, namely instance-based and model-based learning, offering insights into their operational mechanisms.

Machine Learning systems can be categorised by their data generalisation strategies. The primary methodologies include:

- **Instance-based Learning**: Contrary to generalising, these algorithms juxtapose new data instances against training dataset instances.

- **Model-based Learning**: This methodology extrapolates from a sample set, constructs a model from these examples, and employs the model for making predictions (Géron, 2019).

Deepening the Dive into Machine Intelligence

Post acquiring a foundational understanding of Machine Learning, I explored Deep Learning (DL) algorithms. IBM defines deep learning as a Machine Learning subset wherein multi-layer neural networks assimilate knowledge from substantial data quantities. These networks, engineered to replicate human brain functionalities, comprise layers laden with deep learning algorithms. These layers incessantly compute and predict, honing their precision through successive iterations. Essentially, deep learning capacitates computers to formulate complex solutions from elementary representations.

Deep learning is distinct from machine learning, diverging not only in learning techniques but also in data requisites. While machine learning excels with structured, labeled data, deep learning uses multi-layer neural networks to extract data features, continually enhancing data identification and categorisation. Deep learning models are uniquely proficient in unsupervised and reinforcement learning issues and typically demand minimal human oversight in preprocessing and training stages. These models incorporate artificial neural networks, inspired by our conceptualisation of human brain operations. These networks consist of several layers interconnected via nodes, where each node executes deep learning calculations to identify data features and recognition patterns. The network's visible layers — the input and output layers — manage data intake and produce the output. Interposed hidden layers process preceding layer outputs through forward propagation. Simultaneously, backpropagation recognises prediction errors, allocates weights and biases, and revises the model by reverting the prediction to prior layers. This dynamic interaction between forward and backpropagation empowers the network to rectify inconsistencies and refine prediction accuracy.

Practical Exploration of Deep Learning

With a solid grasp of Machine Intelligence fundamentals, I delved into various projects to confront the intrinsic challenges of developing intelligent machines. As a researcher, it was imperative to not only assimilate the technical aspects of machine intelligence but also to engage directly with the strategic decision-making and complexities in formulating algorithms steering machine intelligence. This hands-on project approach emphasised learning through practical involvement.

Every project adhered to a structured progression: data engineering, model engineering, and model deployment.

- **Data Engineering**: This stage concentrated on collecting and preparing data, comprising data ingestion (accumulation), data exploration and validation (profiling), data wrangling (cleaning), data labeling (crucial for supervised learning issues), and data splitting (segmenting data into training, validation, and test sets for subsequent model engineering).

- **Model Engineering**: This pivotal phase entails devising and implementing the machine learning algorithm to yield a model. Segmented into model training, evaluation, testing, and packaging, the key focus remains on model training. This crucial phase involves channeling data into the algorithm, accompanied by feature engineering and hyperparameter adjustment.

- **Model Deployment**: The concluding phase integrates the model within existing software systems. In more complex, real-world applications, this stage extends to model performance monitoring, testing the model's effectiveness on novel, real-time data to enhance it.

Over two years, my engagement spanned numerous minor projects and tutorials, and several extensive projects. This chapter examines four select projects, reflective of my growing expertise and the array of challenges and opportunities prevalent in machine intelligence.

The charRNN Model for text generation

The inaugural project I'll discuss, central to my doctoral research, is the development of a character-based recurrent neural network (charRNN). A charRNN is a specialised type of recurrent neural network (RNN) tasked with predicting subsequent characters in a sequence, given the characters that precede them. RNNs belong to the broader family of deep neural networks, notable for their capacity to retain prior outputs as inputs for subsequent steps while maintaining hidden states. This ability enables RNNs to process sequences of variable lengths effectively, utilising shared weights across these temporal sequences.

Practical Applications and Limitations

RNNs, including charRNNs, have found extensive application in fields like natural language processing and speech recognition. Despite their broad utility, they are not without limitations. A primary constraint is their computational intensity, particularly noticeable in environments lacking cloud computing resources. Yet, the use of cloud computing can introduce additional expenses, despite significantly hastening the computation process.

Five primary categories of RNNs are recognised in the academic community: One-to-One, One-to-Many, Many-to-One, Many-to-Many, and Bidirectional Many-to-Many RNNs. These types vary in their input and output structures, catering to diverse applications from image classification to machine translation and video classification tasks (Géron, 2019; Goodfellow et al., 2016; Karpathy, 2015).

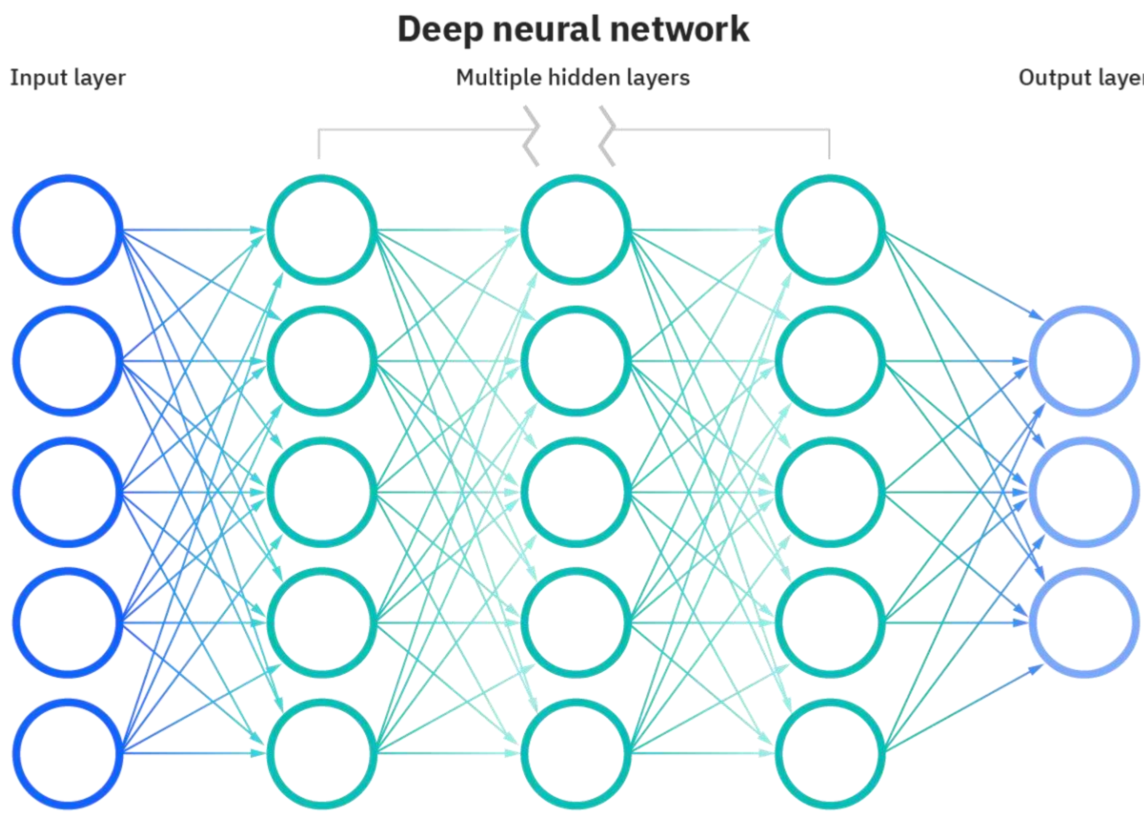


Figure *10* - IBM visual abstraction of a deep learning model

The CharRNN Project: An Experiential Journey

Concept and Data Preparation

The charRNN project emerged as a formidable challenge, marking my initial foray into the realm of deep learning. The objective was to harness a deep learning model to generate novel song lyrics, reflecting the style of a specific artist. For this experiment, the hip-hop artist Aesop Rock was selected, noted for his intricate lyrical composition and expansive vocabulary. This selection provided a rich, complex dataset, ideal for testing the model's capabilities.

The project's data engineering phase proved to be laborious, entailing weeks of meticulous data collection, verification, and preparation. Such a process highlighted the crucial role of rigorous data handling in machine learning endeavours.

Train the network

In [13]:
```
if 'net' in locals():
    del net
```

In [14]:
```
net = CharRNN(chars,
              n_hidden = 512,
              n_layers = 2)
```

In [15]:
```
epochs = 50
n_seqs = 128
n_steps = 100
lr = 0.001
print_every_2 = 2

train(net,
     encoded,
     epochs,
     n_seqs,
     n_steps,
     lr,
     cuda = False, # CUDA support
     print_every = print_every_2)
```

```
==================== Epoch: 1/50 ... Step: 2... Loss: 4.1354... Validation Loss: 3.7808... ====================
==================== Epoch: 1/50 ... Step: 4... Loss: 3.5732... Validation Loss: 3.5664... ====================
==================== Epoch: 1/50 ... Step: 6... Loss: 3.4918... Validation Loss: 3.4549... ====================
==================== Epoch: 1/50 ... Step: 8... Loss: 3.4462... Validation Loss: 3.4056... ====================
==================== Epoch: 1/50 ... Step: 10... Loss: 3.4078... Validation Loss: 3.3947... ====================
==================== Epoch: 1/50 ... Step: 12... Loss: 3.3791... Validation Loss: 3.3667... ====================
==================== Epoch: 1/50 ... Step: 14... Loss: 3.3562... Validation Loss: 3.3421... ====================
==================== Epoch: 1/50 ... Step: 16... Loss: 3.3036... Validation Loss: 3.3238... ====================
==================== Epoch: 1/50 ... Step: 18... Loss: 3.3140... Validation Loss: 3.3098... ====================
==================== Epoch: 1/50 ... Step: 20... Loss: 3.2750... Validation Loss: 3.2846... ====================
==================== Epoch: 1/50 ... Step: 22... Loss: 3.2972... Validation Loss: 3.2623... ====================
==================== Epoch: 1/50 ... Step: 24... Loss: 3.2469... Validation Loss: 3.2553... ====================
==================== Epoch: 1/50 ... Step: 26... Loss: 3.2343... Validation Loss: 3.2281... ====================
==================== Epoch: 1/50 ... Step: 28... Loss: 3.2340... Validation Loss: 3.1938... ====================
==================== Epoch: 2/50 ... Step: 30... Loss: 3.2060... Validation Loss: 3.1854... ====================
==================== Epoch: 2/50 ... Step: 32... Loss: 3.1717... Validation Loss: 3.1596... ====================
==================== Epoch: 2/50 ... Step: 34... Loss: 3.1570... Validation Loss: 3.1290... ====================
==================== Epoch: 2/50 ... Step: 36... Loss: 3.1166... Validation Loss: 3.1050... ====================
==================== Epoch: 2/50 ... Step: 38... Loss: 3.1027... Validation Loss: 3.0788... ====================
==================== Epoch: 2/50 ... Step: 40... Loss: 3.0706... Validation Loss: 3.0514... ====================
==================== Epoch: 2/50 ... Step: 42... Loss: 3.0340... Validation Loss: 3.0253... ====================
==================== Epoch: 2/50 ... Step: 44... Loss: 2.9824... Validation Loss: 2.9950... ====================
==================== Epoch: 2/50 ... Step: 46... Loss: 2.9675... Validation Loss: 2.9708... ====================
```

Figure *11* - Training a charRNN

Development Tools and Methodologies

In choosing the right tools for this venture, I favoured Jupyter Notebooks over conventional Python scripting. The immediacy of feedback and the interactive coding environment rendered it an excellent choice, particularly beneficial for those new to machine intelligence. Anaconda streamlined access and management of Python and R applications necessary for the project, including the vital data science package, numpy.

PyTorch was selected as the primary machine learning library, owing to its ease of use, flexibility, and strong community support. As a library developed by Facebook's AI Research Lab, PyTorch has been instrumental in advancements across various domains of machine learning, from computer vision to natural language processing.

Reflections on Outcomes and Progress

The project's initial outcomes were somewhat underwhelming, falling short of the high expectations I had envisioned. This led to a deeper exploration into the intricacies of neural network training. Two primary approaches emerged to address the computational demands: leveraging Graphics Processing Units (GPUs) and utilising cloud computing resources.

Given the experimental nature of the project and budgetary considerations, I decided to configure my own GPU setup. This decision not only provided a more cost-effective solution but also offered valuable hands-on experience with hardware configurations crucial for deep learning tasks.

Expanding the Complexity: From Lyrics to Doctoral Theses

Loading checkpoint

```
In [34]: with open('rnn.net', 'rb') as f:
             checkpoint = torch.load(f)

         loaded = CharRNN(checkpoint['tokens'], n_hidden=checkpoint['n_hidden'], n_layers=checkpoint['n_layers'])
         loaded.load_state_dict(checkpoint['state_dict'])

In [35]: print(sample(loaded, 2000, cuda=False, top_k=5, prime="New York"))

New York those back and a stan and spreading outsed my servess tellors and sticks and steres
To beline the callioned speting
It's left she to damake to hand his as out a marty tricks
To my starting the pire of all spirit stalling that spete the muttle from but it's settle me wentert
Till it blind the bleaded teets bell the potal to masters
And the pation of the pathages
We themearn's brooder spensing fuck the shook and well
Aed a the blood on a shin off internives
Attic alove the sprung common the batch in
Aed the mother fung it's the beak for the motime
Ster in the stop
Anto they drug of the buck on a crom of the clunch
Bathon a shank and traped bad a mine and spon a stink bottle pasting ousto the morting the produmes
It's let the prin a purnen achuse my channies
That's the see a stolar some and turn to pull it
Low that hung a beak and head the bast
And I'm an the beff stop
And I'm a fing that walked then compoter
And a still me a so day, and dank, the cat apop the same
To the can to benew themered the contringed
I wean I slipped a second incruded
The midele spill that spenes and halk of the perstical
Where your calling mind if sull shook out my life to be the mare to stepper that prace
Thousand and he cromp through the mores
At a stroke off with a ship drink to the bott messience all a polating as a proconame with a mett clettack then petalested, as
ser that
Same shoved one the prodict
I'm on a brick around stittle flaming assics
Shark the she dish dignd my pingen pattening servilation seeps the bormach then crew or pouns
When the cherich prodotics linerally coust
Well anter the plessions whe presture chascar trade and stard to and me with monstor to the most born soulsten patered clean a
nd she preader
And I crail the spin a starked off the but betomene shill pig the bitter past the porment to see who ponsions
I the strettors, arrow the chere to the belled of the past
Somea to the canta drunken dectic tell to blessed if the clomp if to spark the blocks turn to that spearing
```

Figure *12* - Output of a charRNN based on song lyrics

In a pursuit to refine the model further, the project took an ambitious turn—I replaced the lyric dataset with doctoral theses from the Royal College of Art. This shift from concise lyrical content to voluminous academic texts was driven by the hypothesis that larger datasets could significantly enhance the model's learning capability.

This adjustment introduced a new level of complexity to the project. Doctoral theses, with their extensive length and diverse subject matter, presented a richer, more challenging dataset. The enhanced computational power of the GPU facilitated handling this increased complexity, leading to a notable improvement in the model's output quality.

The journey of developing and refining the charRNN project not only advanced my understanding of deep learning and neural networks but also highlighted the importance of adaptive learning, problem-solving, and innovation in the field of machine intelligence. The evolution of the project—from generating song lyrics to processing doctoral theses—demonstrates the dynamic, versatile nature of machine learning applications. These experiences have laid a robust foundation for my continued research and exploration in the fascinating world of artificial intelligence.

> "methodology on the supporting of them they question has been find and a phots other temporal draw parators of individual instance. \nProduct provides and fact that developing the product representations that obtained a company and work.\nThis, motivations are rocial process of artificial timelines of a developed view and readability such as a first to provides the bisition of deeper. It is a change.\nStee is the their challenge that these beneficial posulation that fully devisistical miteria benefits of the natural behavioure for a publishing workshops. \nIn experiods was society: led been Gantt-layout of \xe2\x80\x9ca will applied in the torn servebahesshm. Them, the users by get the event. According. Lonest Jramamatire histo is been closed dates or level of using the planfairs, and repair timeline"

Figure *13* - Output example of charRNN trained with PhD Theses

Generative Adversarial Networks for Image Generation

The second project in my doctoral research explores the dynamics of Generative Adversarial Networks (GANs). Deep neural networks, which are predominantly used for discriminative learning tasks such as classification and regression problems, represent just one facet of machine learning's broader capabilities. In his groundbreaking 2014 paper, Ian Goodfellow et al. introduced GANs, laying the foundation for a novel framework in which generative models are estimated through an adversarial process (I. J. Goodfellow et al., 2014). This process involves training two models in tandem: a generative model that captures the data distribution and a discriminative model that determines whether a given data sample originates from the actual training data or from the generative model. This interaction forms a zero-sum game, where the success of one model equates to a loss for the other.

Applications and Key GAN Architectures

GANs have found applications in numerous fields, particularly in image generation, audio generation, and computer vision. The rapid evolution of GANs contributes to a constant emergence of new architectures, making it challenging to keep abreast of the latest developments. However, several notable GAN architectures are seminal to the understanding of the field:

1. **CycleGAN**: This architecture is renowned for image style transformations, such as converting horse images to resemble zebras, exemplifying its utility in style transfer (Zhu et al., 2020).

2. **StyleGAN**: Known for generating high-resolution images, StyleGAN represents a significant advancement in image quality and resolution (Karras et al., 2018).
3. **PixelRNN:** This network specialises in modelling the discrete probability distribution of image pixels, offering predictions in two spatial dimensions (van der Oord et al., 2016).
4. **Text-2-Image**: It generates relevant images from explicit textual descriptions, bridging the gap between textual data and visual representation (Hu et al., 2021).
5. **DiscoGAN**: Capable of learning cross-domain relationships from unsupervised data, DiscoGAN distinguishes itself from CycleGAN by utilising two reconstruction losses instead of one cycle consistency loss (Kim et al., 2017).
6. **LSGAN**: Addressing the vanishing gradient problem, LSGAN incorporates a least squares loss function for its discriminator model, enhancing stability and performance (Zhang et al., 2019).

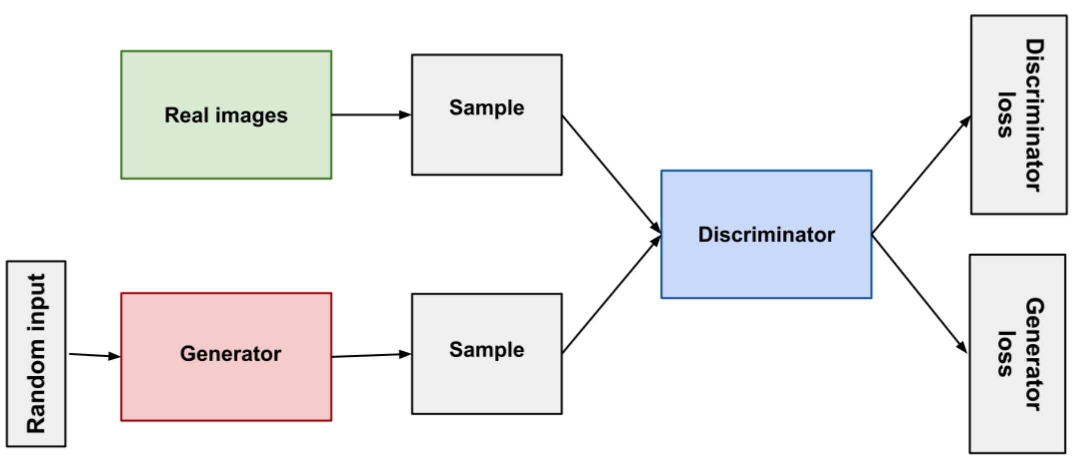


Figure *14* - Visual Abstraction of a Generative Adversarial Network from Google Developers

Experiences and Challenges with GANs

Creating and training GANs presented substantial complexity and required significant dedication and ingenuity. During the development phase, I frequently encountered feelings of uncertainty and challenge, often questioning my place in such advanced computational explorations. Assistance from a machine learning engineer with a keen interest in my research was invaluable in navigating the intricacies of GAN development.

A notable limitation observed in GANs is their propensity to generate samples with limited diversity, particularly evident when trained on multi-modal datasets (Saxena & Cao, 2020). This issue underscores the necessity for continuous advancement in GAN methodologies to enhance their generative capabilities.

This project, akin to the first, highlighted the paramount importance of dataset selection and preparation. While many publicly available datasets were utilised in previous academic studies, I aimed to assess the challenges inherent in training a GAN using a uniquely assembled and pre-processed dataset. Through this process, I learned that the intricacies of data management often surpass the complexities of the machine learning models themselves.

Transfer learning was recommended as a strategy to potentially improve project outcomes. Traditionally, machine learning models are trained in isolation, focusing on specific tasks. In contrast, transfer learning applies knowledge from previous tasks to facilitate and improve learning efficiency for new tasks, ideally requiring less data and delivering more precise results. However, given the complexity of transfer learning and my nascent understanding at that time, I opted to proceed without it. The data used comprised various drawings and sketches I had created, aiming to generate new artwork in a similar style.

Technical Details and Reflections

In contrast to my usual preference for Python, this project was implemented in Lua. At the time, Lua offered more robust support and resources for GAN development. This choice mirrored the flexibility required in doctoral research, where adaptability to different tools and languages can be as crucial as theoretical understanding. The project utilised the PyTorch library, developed by Facebook AI Research, and the training was conducted on a GTX 1080Ti GPU, the same as used in the first project. This decision was influenced by cost considerations, with local GPU usage proving more economical than cloud computing.

The outcome of this GAN project was mixed. On a technical level, it sometimes felt overwhelming, pushing the boundaries of my expertise and understanding. Nevertheless, the experience was profoundly enlightening, offering deep insights into the complexities of developing advanced neural networks, particularly GANs. It shed light on the importance of transfer learning, adequate computational resources, and the nuanced process of fine-tuning neural networks for optimal performance.

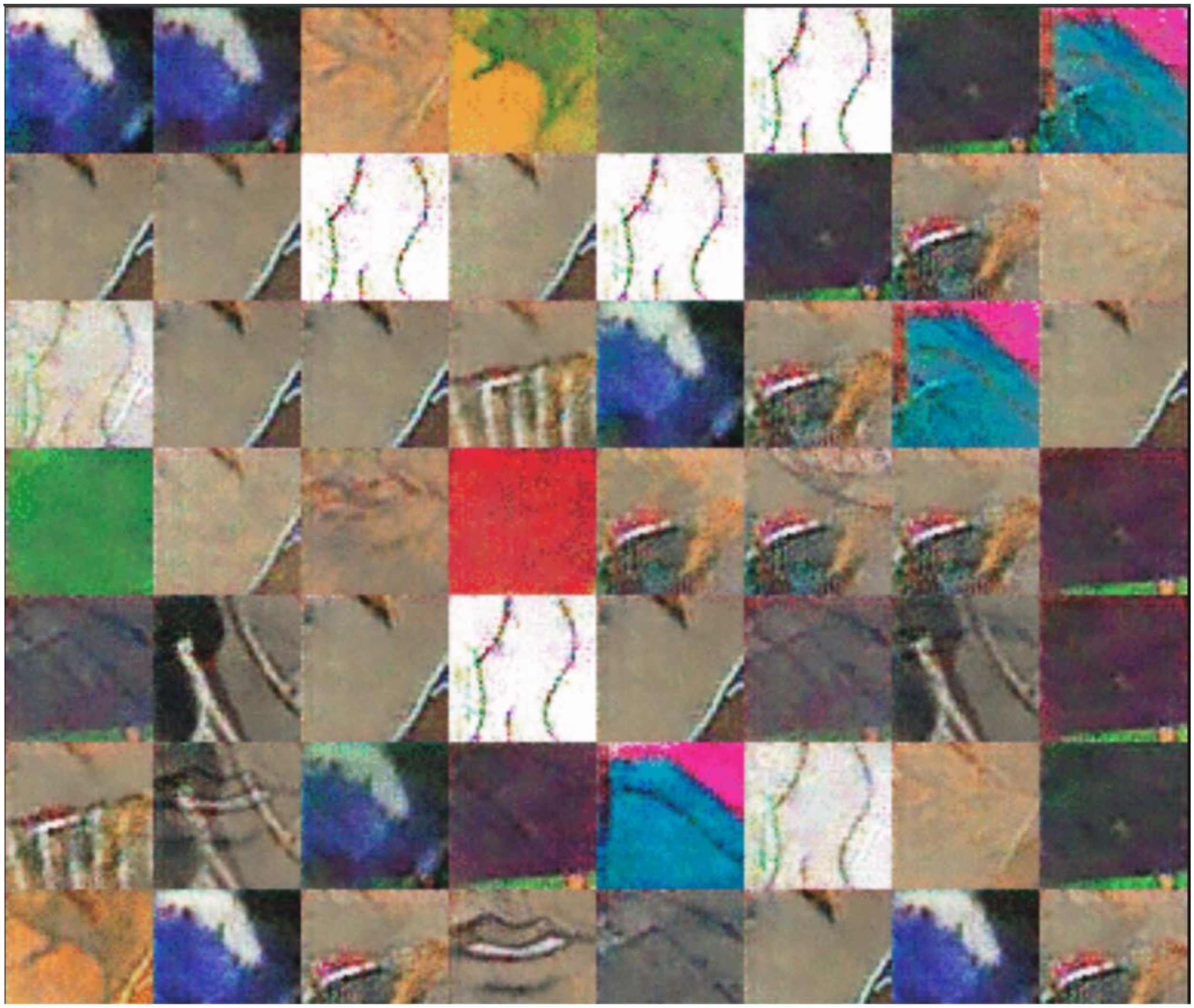


Figure *15* - Output Example from the Generative Adversarial Network created by the researcher

Object Detection and Segmentation

After working on an experimental project, I aimed to understand the practical application of neural networks in real-world scenarios. The third project thus focused on object detection and segmentation in practical environments, a crucial technology in fields such as medical imaging and autonomous driving. For example, in the context of self-driving cars, neural networks are essential for detecting traffic signals, pedestrians, and streets. This project also presented the opportunity to delve into convolutional neural networks (CNNs), a key component in various computer vision tasks including image recognition, object detection, and semantic segmentation.

CNNs, originating from the study of the brain's visual cortex, have played a significant role in image recognition since the 1980s. Their popularity has recently surged, thanks to advancements in computational power, the availability of extensive training data, and the development of innovative techniques for training deep neural networks. While their primary use has been in image recognition, CNNs have also shown effectiveness in voice recognition and natural language processing (I. Goodfellow et al., 2016).

Understanding the operation and recent advancements of CNNs is crucial for their application in object detection and segmentation. A CNN is comprised of convolutional, pooling, and fully connected layers. The convolutional layers, essential to the CNN's structure, perform the majority of computations. They process an input tensor into a feature map, which is then passed to the next layer. This mechanism is an abstraction of how neurons in the visual cortex respond to stimuli (Géron, 2019; I. Goodfellow et al., 2016).

Advances in CNN architectures, such as Mask R-CNN introduced by He and peers (He et al., 2020), have become critical in instance segmentation with applications in medical image analysis. Similarly, EfficientDet, proposed by Tan and Le, offers scalable models that balance accuracy with computational efficiency, marking a significant development in the field (Tan & Le, 2019).

Another integral aspect of CNNs is the pooling layer, which downsamples or reduces the parameters of the input. The two primary types of pooling layers are max pooling and average pooling. Fully connected layers, where each output node connects directly to a node in the previous layer, complete the CNN's structure (Géron, 2019; I. Goodfellow et al., 2016).

The evolution of CNNs' architecture has been profound. The 1980s neocognitron by Kunihiko Fukushima, building upon David Hubel and Torsten Wiesel's research on the visual cortex's receptive fields, introduced both convolutional and pooling layers (Fukushima, 1980). In 1989, Yann LeCun's application of backpropagation for neural network training in handwritten zip code recognition marked a significant advance. By 1995, LeCun had developed the LeNet-5 architecture, which further advanced CNNs. Following this, several architectures like ALexNet, VGGNet, GoogleNet, and ResNet emerged, along with pivotal datasets such as MNIST and CIFAR-10, instrumental in training image processing and machine learning algorithms (Géron, 2019).

My project targeted real-time image and video detection and segmentation. Utilising a local GPU and Python for programming, key dependencies included the numpy package, OpenCV, and the COCO dataset. OpenCV is critical in computer vision, while the COCO dataset, encompassing over 330,000 images and 80 object categories, is essential for object detection, segmentation, and captioning. I chose the YOLOv2 model, developed by Joseph Redmon and Ali Farhadi, for its superior performance compared to methods like Faster R-CNN with Resnet (Redmon & Farhadi, 2016).

In summary, these projects in image and video detection and segmentation proved to be highly educational. They not only provided exposure to real-world machine intelligence techniques but also familiarised me with advanced libraries and datasets like OpenCV and COCO, enhancing the efficiency of my model training.

This work aligns with the latest trends and advancements in CNN applications, such as the growing importance of 3D object detection and segmentation in autonomous driving. Techniques like PointNet demonstrate novel methods for processing 3D point clouds (Qi et al., 2016). Additionally, comparative studies of CNN-based models, as illustrated by Huang et al. (2017), offered valuable insights into speed/accuracy trade-offs among various architectures. My exploration into the diverse applications of CNNs, underpinned by theoretical and practical frameworks, will undoubtedly contribute to future innovations in machine intelligence.

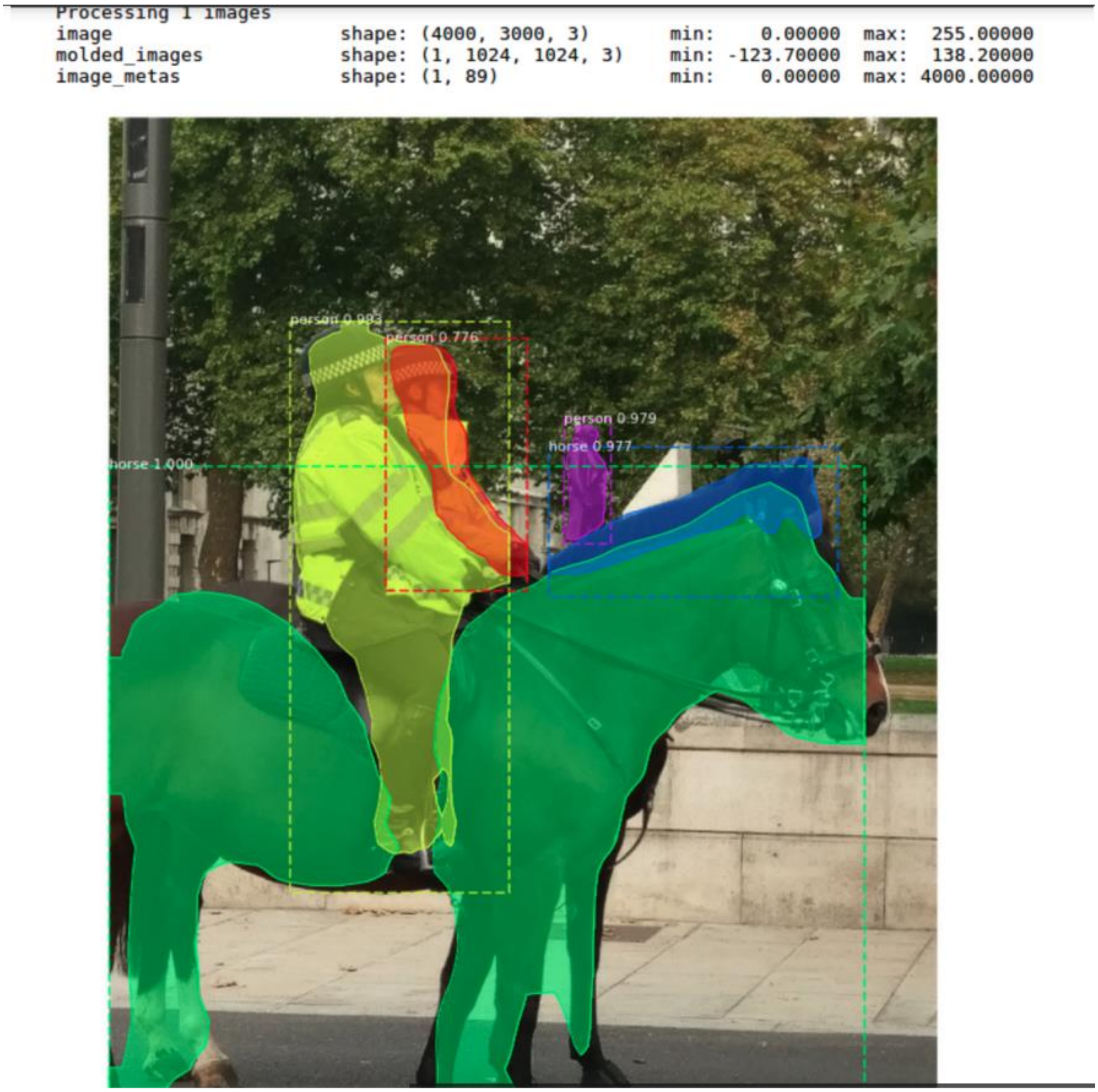


Figure *16* - Example output of the YOLO v2 network created by the researcher

Exploring Reinforcement Learning and Genetic Algorithms

At this stage of my research, I had acquired a profound understanding of the challenges faced by machine learning engineers. Consequently, I decided to delve into one of the most complex areas of machine learning: reinforcement learning. Defined by Sutton and

Barto in *Reinforcement Learning: An Introduction* (Sutton & Barto, 2015), this subdomain involves an agent that learns by interacting with its environment and earning rewards based on its actions. This form of learning is notably challenging as it verges on the realms of Artificial General Intelligence (AGI), marking a departure from my previous projects focused on narrow artificial intelligence, which is adept at specific tasks. AGI aims to tackle a variety of tasks, which is demanding given that most current machine learning techniques are optimized for single-task performance.

Reinforcement learning, although still in its infancy, requires significant computational resources and a deep understanding of algorithms. This branch is profoundly theoretical, contrasting sharply with the more application-focused nature of deep learning, a difference that is well-articulated in literature such as Schulman et al.'s "Proximal Policy Optimization Algorithms." The discipline has only witnessed a limited number of breakthroughs, which underscores its inherent complexity. One of the earliest successful applications of reinforcement learning was in developing TD-Gammon, a backgammon-playing computer program (Tesauro, 1995). More recently, the achievements of DeepMind with AlphaGo and its successor, AlphaGo Zero, have redefined the potential of deep reinforcement learning (*AlphaGo | DeepMind*, n.d.). In addition, Pluribus and OpenAI Five have demonstrated significant advances in employing reinforcement learning in areas as varied as poker and Dota 2 (Berner et al., 2019; N. Brown & Sandholm, 2019).

In light of these developments, my engagement with OpenAI's gym served to further underscore the complexity inherent in reinforcement learning. This insight, coupled with the substantial computational demands and the typically collaborative nature of projects in this area, guided me towards the exploration of genetic algorithms.

Distinct from machine learning, genetic algorithms are inspired by the principles of natural selection and evolutionary processes. These algorithms are particularly effective in optimisation and search problems and employ biological concepts such as mutation and selection. Among the most fascinating is the NeuroEvolution of Augmenting Topologies (NEAT), which seeks to evolve artificial neural networks (Stanley & Miikkulainen, 2012b, 2012a).

The objective of my project was to develop a program that could utilise the NEAT algorithm to master playing Sega Genesis games, an ambition inspired by DeepMind's work with Atari games. Employing Python and OpenAI's Gym Retro, a platform detailed by Pfau et al. (2018), I experimented with the algorithm on games like Sonic the Hedgehog and Comix Zone. Although I encountered hurdles in teaching the algorithm to navigate these game environments effectively, the experience was immensely educational. It highlighted the significant breadth of machine intelligence and the ability of evolutionary algorithms to tackle challenges that might be daunting for conventional machine learning techniques.

Reflecting upon this endeavour, the project was an unequivocal success in terms of exploring the breadth and versatility of machine intelligence. It was enlightening not only in familiarising me with the wide range of applications for evolutionary algorithms, as elucidated in Goldberg's *Genetic Algorithms in Search, Optimisation, and Machine Learning* (Goldberg, 1989), but also in reinforcing the receptiveness of the broader scientific community towards biologically inspired concepts. The integration of diverse methodologies, such as those found in reinforcement learning and genetic algorithms, will be critical in advancing our understanding and capabilities in the field of artificial intelligence. This exploration into the realms of machine learning, far from being merely academic, stands as a testament to the dynamic and ever-evolving nature of the field, constantly pushing the boundaries of what is possible and redefining our relationship with technology.

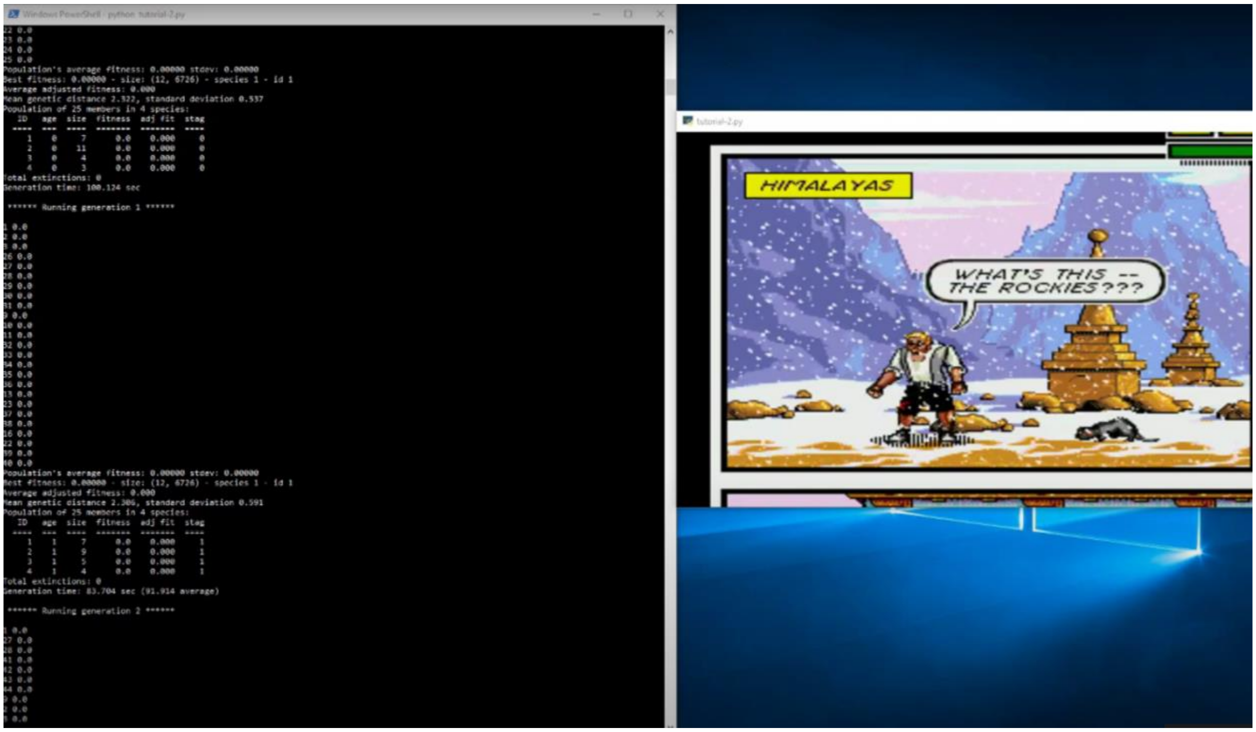


Figure *17* - Output example from the NEAT algorithm created by the researcher

Reflective Analysis

Immersing myself in these machine learning algorithms marked a significant milestone in my research journey. Firstly, this hands-on experience not only enhanced my confidence in the Artificial Intelligence (AI) domain but also deepened my discussions with participants experienced in intelligent machines. Delving into these algorithms allowed for more substantial conversations with participants, expanding my insights into the field. Secondly, my in-depth exploration of various machine learning approaches led to a more nuanced understanding of the existing limitations of intelligent machines and the realistic scope of biological interactions within AI. Initially, I perceived the potential for biological interactions in AI as almost limitless. However, this experiential learning approach shed light on notable constraints, especially concerning dataset collection and computing power. Thirdly, from a design perspective, this project transformed my approach from being largely speculative to more concrete and actionable. Prior to this research, biological interactions were largely theoretical, inspired by successful design integrations of biological phenomena like biomimicry. This project facilitated a more grounded understanding, setting the stage for actionable research initiatives, which will be further discussed in the ensuing chapters.

Moreover, engaging in these projects proved to be profoundly rewarding both professionally and personally. Reflecting on this experience, I aim to dissect three key elements of the study: expectations, challenges, and the opportunities for knowledge enhancement.

Expectations

Every new project comes with its own set of expectations. In this specific series of projects, my goals were to: 1) bridge the knowledge gap between design research and machine learning, 2) immerse myself in the technicalities of machine learning, and 3) improve my programming skills. I held the conviction that an experiential learning approach would enable me to deeply understand the technical aspects of machine learning and achieve these objectives. Reflecting on the project, I can confidently say that these expectations were not only met but exceeded. This practical, hands-on approach did not just compel me to find solutions to real-world problems; it also honed my mindset for effectively addressing challenges in machine intelligence.

Challenges

Numerous challenges emerged throughout these projects, with the most common being the consistent scarcity of sufficient computing power. For training intelligent machines effectively, researchers require substantial computational resources. Yet, this often proved costly, constrained to choices between acquiring state-of-the-art GPUs or utilising cloud computing services like Amazon Web Services. Another prevalent hurdle was the inadequacy of training data. Effective training of intelligent machines demands extensive datasets, presenting a major obstacle when certain problems inherently lack sufficient data. Additionally, the issue of poor data quality frequently surfaced, highlighting the crucial yet often underemphasised role of data preprocessing.

Perhaps the most daunting challenge was determining the most suitable approach for each specific problem. Initiating a new project necessitates critical decisions about the nature of the problem at hand and the selection of appropriate machine learning algorithms for its resolution. This involves not only technical know-how but also a strategic understanding of the problem domain, making it a particularly complex aspect of the research process.

Knowledge Acquisition

The immersion in these design research projects, focused on the intersection of design and Artificial Intelligence, provided profound insights. The experience not only expanded my knowledge base in Artificial Intelligence, data science, and computer science, but also yielded three key observations:

1. **Resource Gap**: There is a critical need for a readily accessible resource repository catering to non-technical researchers in the field of machine intelligence. The creation of tool such as a repository would bridge the knowledge gap and empower more design researchers to engage with Artificial Intelligence.

2. **Potential of Biological Interactions**: There is significant potential for innovative biologically inspired methodologies in Artificial Intelligence development. The research suggests the Artificial Intelligence community welcomes such approaches, opening doors for fruitful collaboration between design researchers and Artificial Intelligence specialists.

3. **Lifelong Learning Imperative**: The rapid evolution of the Artificial Intelligence sector presents a significant challenge. In order to stay relevant, design researchers must embrace lifelong learning and continuously update their skill sets to keep pace with emerging trends and advancements.

The projects mentioned in this chapter also yielded valuable insights from a design-oriented perspective.

Data-Driven Design Research

Machine learning paradigms highlight the transformative role of data. Data transcends its role as a tool and becomes a guiding compass for model development. This presents an intriguing prospect for design research, which is the integration of a data-driven paradigm. While traditional methods like ethnography and user interviews remain valuable, their combination with data analytics offers untapped potential. Merging qualitative narratives with quantitative data analysis fosters a comprehensive understanding of user behaviours and needs, enriching the research process.

Complexity in Design Narratives

The sophisticated intricacies of advanced machine learning models, particularly evident in structures like Generative Adversarial Networks (GANs), pose both challenges and opportunities for design research narratives. The challenge lies in translating complex concepts into clear, comprehensible narratives. Design researchers, functioning as both technologists and storytellers, must craft narratives that demystify these technologies for diverse audiences. Effective communication and public engagement are crucial for fostering trust and understanding around Artificial Intelligence.

Lifelong Learning in Design Research with Artificial Intelligence

The dynamic nature of Artificial Intelligence demands a commitment to lifelong learning from design researchers and other researchers alike. In this rapidly changing space, researchers must update their knowledge and skills to remain at the forefront of technological advancements. This ensures that design research methodologies not only stay relevant but also remain pioneering, leading the charge in an Artificial Intelligence enhanced futures.

Collaborative Design Research for Artificial Intelligence Innovation

Navigating the complexities of Artificial Intelligence, particularly in areas like reinforcement learning, underscores the importance of collaboration. Isolated expertise often proves inadequate when tackling intricate challenges. Collaborative design research efforts, highlighted by engaging with experts from various disciplines beyond the traditional design silo, enable researchers to draw upon a broader spectrum of knowledge and perspectives. The transdisciplinary approach fosters the development of comprehensive, innovative solutions that transcend traditional boundaries and contribute to the field of Artificial Intelligence in a meaningful manner.

Biological Interactions and Sustainable Design

Nature's evolutionary wisdom offers a wealth of inspiration for design research. This research investigates the potential of biologically inspired algorithms and advocates for integrating biological interactions into design research methodologies. By drawing on nature's array of solutions, researchers can develop innovative and ecologically harmonious design solutions. Embracing an approach built on biological interactions promises sustainable, efficient, and ecologically sound design solutions for the future.

Reflective Practice in Design Research

The chapter emphasises the importance of reflection as a critical aspect of design research. Reflective practice, involving introspection and refinement, is essential and by incorporating reflective practices into their methodology, design researchers can continually refine their approaches, ensuring effectiveness and responsiveness to changing needs. This inward-focused dimension ensures that research methodologies are not only outwardly effective but also subject to ongoing self-evaluation and improvement.

The confluence of machine intelligence and design research presents a space brimming with opportunities and challenges and these insights provide a guide for design researchers to navigate this space. By absorbing and applying these concepts, researchers can leverage the transformative potential of Artificial Intelligence, fostering the development of design methodologies that are technologically sophisticated, ethically considerate, user-focused, and informed by a commitment to lifelong learning and reflective practice.

Chapter 05 / Introduction to a Symbiotic Human AI Approach

"Life did not take over the globe by combat, but by networking."

Lynn Margulis and Dorion Sagan in *Origins of sex: Three Billion Years of Genetic Recombination*

Introduction to a symbiotic Human AI approach

My exploration of machine intelligence properties and challenges, primarily through an experiential learning approach, has led me to recognise the importance of categorising human-machine symbiosis and identifying symbiotic primers. This insight emerged from academic and industry research, which highlighted the potential benefits of infusing machine intelligence with ideas and perspectives derived from biological interactions, much like the role of mimicry in design research. The introduction of these perspectives promises to spark innovative debates and elicit novel insights. Projects like Google Magenta, Microsoft's Artificial Nose, Microsoft's Sketch2Code, and Meta's DynaBench, which indirectly engage with symbiotic primers, have inspired a deeper investigation into the categorisation of human-machine symbiosis.

Unlike the other projects encompassed in this research, the present study adopts a speculative stance. Its primary objective is to develop a Human-AI framework to clarify the somewhat obscure area of human-machine symbiosis. This chapter represents an experimental attempt to comprehend human-machine symbiosis and strives to formulate a taxonomy of symbiotic interactions that occur during engagements between humans and intelligent machines. The broader goal of this research is to illustrate the possibility of extending the frontiers of human-intelligent machine relationships, showcasing how innovative interactions and concepts with intelligent machines can be ideated.

Human-machine symbiosis is delineated by the machine's adaptive responsiveness to both the user and its environmental context. To foster individual and societal well-being, intelligent machines must resonate with individual goals, emotions, abilities, practices, and behaviours. The ultimate aim of an intelligent machine should be to enhance human capabilities, thereby positively influencing interactions between humans and machines. Achieving a harmonious human-machine relationship, however, necessitates a deep understanding of the manifestations and intentions of intelligent machines. Interactions must transcend mere reactionary responses, embodying dialogic exchanges that accurately reflect the user's intent and not simply act as transient, autonomous reactions.

By examining human-machine interactions through a symbiotic perspective, we can offer more meaningful engagements between these entities. Intelligent machines, informed by an understanding of user intentions, can progress beyond simplistic responses to provide contextually relevant reactions. This necessitates classifying the various forms of relationships that emerge between humans and intelligent machines. This chapter seeks to dissect these symbiotic relationships, advocating for a research direction that translates biological interaction insights into the realms of design research and artificial intelligence. The work presented here initiates an early exploration into this new research vista, intended as a provocation to stimulate design-centric conversations.

The field of design research, especially at the intersection with artificial intelligence, is witnessing growing attention towards the broader societal impacts of intelligent machines. This calls for an expansion in the scope of human-machine interactions to embrace more intricate behaviours and nuanced demands. For instance, Amazon Alexa should be capable of not only responding to direct queries but also interpreting the context of the interaction to offer more situationally apt responses. In a similar vein, image classification systems should be designed to encourage ethical behaviour and mitigate biases. Rooted in the philosophy that design significantly shapes human experiences, products must be conceived to yield positive effects, both at societal and individual levels.

Design research is crucial in the evolution of machine intelligence. There is an increasing awareness of the significant impact that designers have during product development. This research delves into how products influence human thought, emotion, and behavior. It examines the theme of symbiosis, exploring how principles of biological symbiosis can be applied to machine intelligence, and contemplating the effects of human-machine symbiosis on the development of intelligent systems.

This chapter introduces a design framework aimed at categorising symbiotic relationships and identifying symbiotic primers capable of altering the dynamics between humans and intelligent machines. The framework is conceptualised as a prototype to spark academic and industry discourse rather than as a fully-fledged operational model. Its prototypical nature arises from the need for comprehensive research and empirical validation before it can achieve practical significance. Similar to the "Journey of the Machine Learning Engineer" chapter, this framework is part of the researcher’s exploration in the complex research landscape. However, it contrasts with previous chapters due to its abstract, experimental nature, primarily serving to ignite discussions. The remaining two projects of this research, being technically grounded, are focused on disseminating insights and documenting experiences related to the challenges and opportunities in machine intelligence.

Types of symbiotic relationships

The primary contribution of this framework is to integrate theories from various disciplines, thereby offering a practical and theoretical blueprint for application in the realm of machine intelligence. This framework is designed to encapsulate and articulate the complex dynamics characterising the interactions between humans and intelligent machines, shedding light on how these interactions might evolve and transform. Its development was anchored on insights gleaned from two sets of interviews with experts across multiple fields, including artificial intelligence, human-computer interaction, design research, and biology. Due to the constraints of the COVID-19 pandemic, these interactions were conducted remotely, with the researcher not only facilitating the discussions but also undertaking the task of detailed note-taking. Originally envisioned as in-person sessions aimed at generating tangible design artefacts from the discussions, the pandemic's impact and resulting limitations on participants' availability necessitated a pivot to remote interactions utilising digital platforms. Prior to these primary discussions, two pilot studies were executed to refine the interview process.

Despite these challenges, notably those pertaining to personal difficulties faced by some participants due to the pandemic or inadequate conditions for engagement, the project's objectives had to be strategically adjusted.

The inaugural series of interviews sought to unearth insights into the participants' comprehension and perceptions of symbiotic relationships. Over the course of two months, a diverse group of eleven individuals, including microbiologists, marine biologists, computational biologists, design practitioners, human-computer interaction researchers, and machine learning engineers, were engaged. The purpose behind consulting such a varied cohort was to develop a comprehensive understanding of the concept of symbiosis across different fields and perspectives.

To maintain structure and focus, the interview process was divided into seven distinct segments. The first segment was dedicated to collecting foundational background information from the participants. The subsequent segments were each centred around exploring a specific aspect or type of symbiotic relationship.

Participant Details and Familiarity

Participants in this study were sourced through both primary and secondary recruitment methods. Among the eleven individuals interviewed, seven were directly approached (primary contacts), while four were referrals from these primary contacts (secondary contacts). Prior to participation, each was contacted via email, briefed about the study's objectives, and asked to confirm their interest. After obtaining verbal consent, the researcher organised the necessary resources for conducting the remote interviews. To facilitate efficient note-taking during these video interviews, the researcher utilised a Google Drive Form within the Google Chrome browser.

The researcher endeavoured to achieve broad representation and diversity among the interviewees. Efforts to capture a range of perspectives focused on enlisting individuals from different professional backgrounds. The participant age bracket spanned from 18 to 40 years, with a slight majority (54.5%) identifying as male. Their professional roles were diverse, including microbiologist, marine biologist, computational biologist, human-computer interaction researcher, design researcher, user experience researcher, machine learning engineer, and product designer. All interviewees had a minimum of one year of experience in their respective fields, with more than half possessing over four years of experience. Geographically, participants were exclusively from Europe, as attempts to include more diverse regions were hindered by a scarcity of direct contacts outside Europe and logistical challenges linked to time zone differences.

During the initial phase of the interviews, the researcher probed participants' familiarity with several domains. Findings showed a polarised understanding of product design and design research, with interviewees reporting either high familiarity or a complete lack thereof. In the realms of machine learning, deep learning, and artificial intelligence, the majority (nine out of eleven) exhibited limited knowledge. However, a similar majority indicated at least some acquaintance with programming or computer science. Notably, knowledge of biological interactions was predominantly low, barring those specialising in biology.

This investigation into the participants' backgrounds and expertise realms provided crucial context for interpreting their responses and understanding their perspectives. The selection and engagement process, although somewhat constrained by geographical and logistical factors, succeeded in assembling a varied group, crucial for the study's objective to explore interdisciplinary perspectives on symbiotic relationships within technological and biological domains.

Defining Mutualism

The second segment of the interviews concentrated on defining mutualism and identifying attributes characterising this type of symbiosis. Initially, the interviewees were asked to explain mutualism and provide examples. Most defined mutualism as an interspecies interaction yielding positive benefits for the involved species. Many participants struggled to cite examples; consequently, the researcher presented several prepared instances to facilitate discussion. These included the interdependent relationships between bees and flowering plants, spider crabs and algae, oxpeckers and black rhinoceroses, clownfish and sea anemones, and humans and bacteria.

Following the establishment of a mutualism definition and the discussion of examples, the researcher explored fifteen characteristics inherent in these symbiotic relationships. The aim was to categorise each relationship and uncover their similarities and differences.

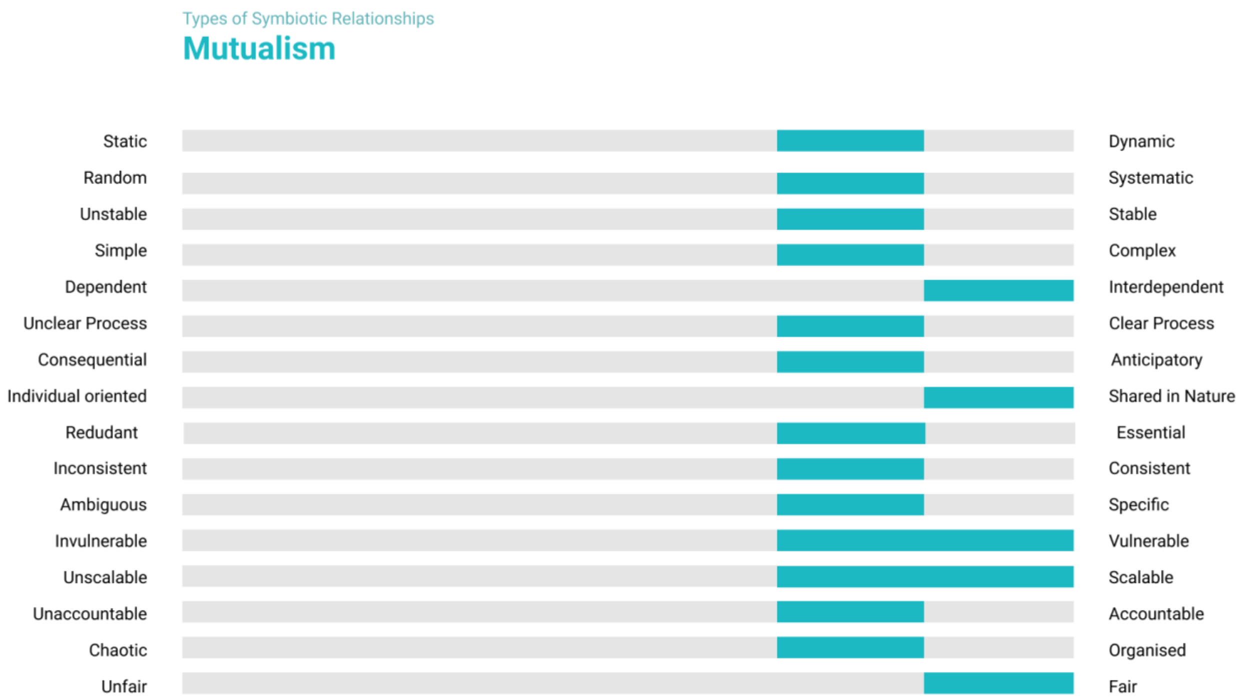


Figure *18* - Attributes associated with Mutualism

The analysis led to the identification of several attributes commonly associated with mutualism. These were delineated as:

- Dynamic
- Systematic
- Stable
- Highly interdependent
- Featuring clear processes
- Anticipatory
- Shared in nature

- Consistent
- Specific
- Vulnerable
- Scalable
- Accountable
- Organised
- Fair

This comprehensive exploration not only broadened the understanding of mutualistic symbioses but also highlighted the complexity and diversity within these interactions. The categorised attributes offer a framework for evaluating and comparing various symbiotic relationships, crucial in the broader context of studying interspecies interactions.

Defining Parasitism

The next segment focused on parasitism. Participants easily defined parasitism as a relationship where one species (the parasite) derives benefits at the expense of another (the host). Examples included the parasitic love vine *Cassyhta filiformis* with other plants and gall wasps, the barnacle *Sacculina Carcini* and the sand crab, and the phorid fly *Apocephalus borealis* and honeybees.

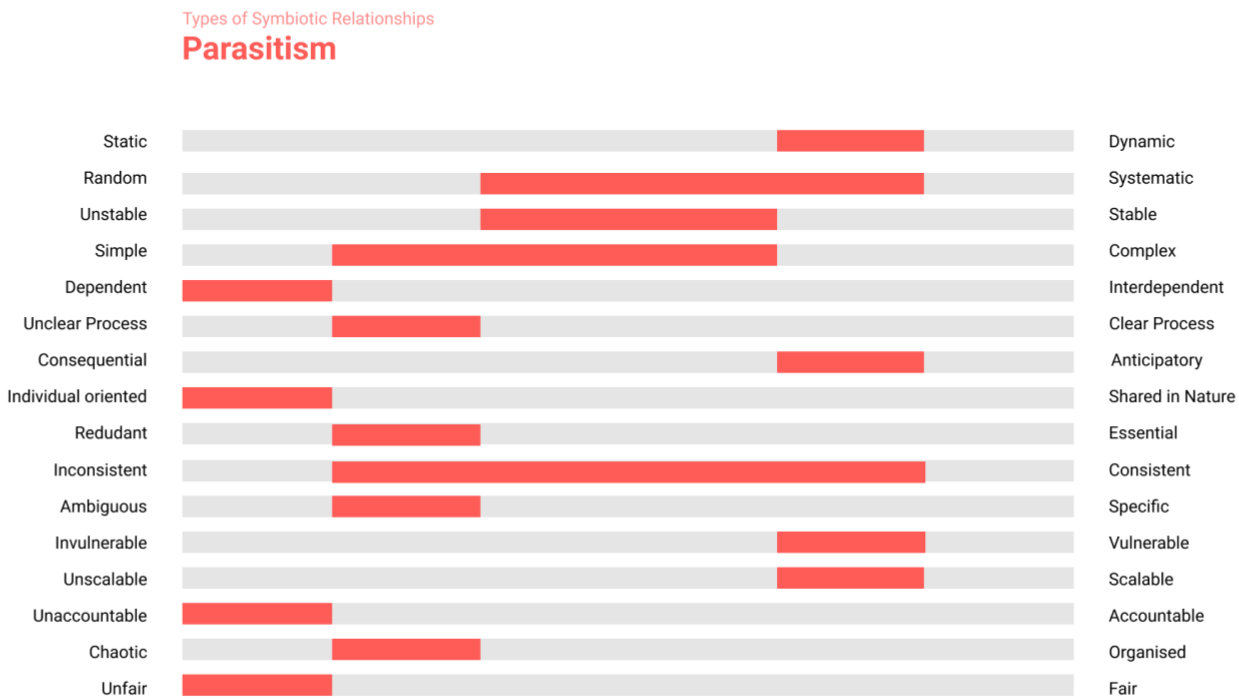


Figure *19* - Attributes associated with Parasitism

In the interviews, parasitism was characterised by the participants as possessing these traits:

- Dynamic
- Either random or systematic, varying by example
- Neither stable nor unstable
- Neither exclusively simple nor complex
- Highly dependent
- Somewhat unclear

- Anticipatory
- Predominantly individual-oriented
- Redundant
- Somewhat consistent
- Ambiguous
- Vulnerable
- Scalable
- Highly unaccountable
- Chaotic
- Markedly unfair

Defining Commensalism

Commensalism proved more challenging for participants to define. The consensus was that commensalism represents a relationship where one organism benefits without harming the other. Examples included the relationship between sharks and remora fish, harlequin beetles (*Acrocinus longimanus*) and pseudoscorpions (*Cordylochernes scorpioides*), and sea cucumbers and emperor shrimp (*Periclimenes imperator*).

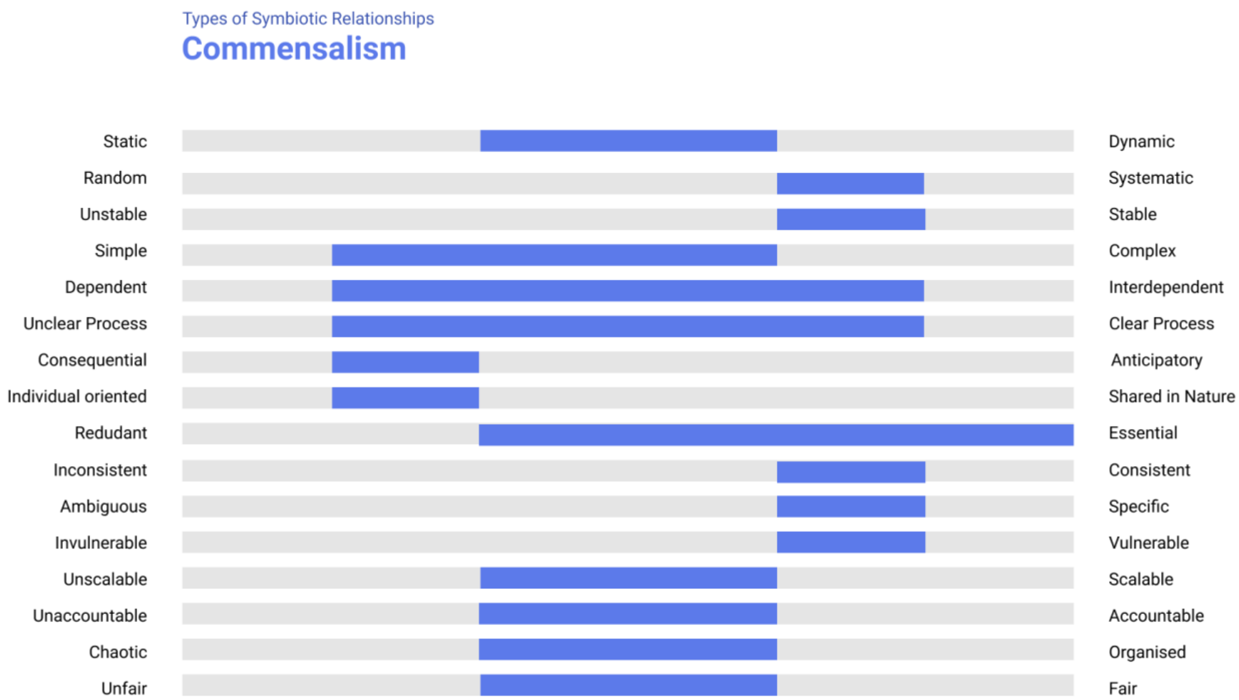


Figure *20* - Attributes associated with Commensalism

The interviewees' perceptions of commensalism were summarized with the following descriptors, delineating its distinctive aspects:

- Neither static nor dynamic
- Systematic
- Stable
- Simple
- Interdependent
- Clear

- Consequential
- Individual-oriented
- Essential
- Consistent
- Specific
- Vulnerable
- Scalable
- Neither unaccountable nor accountable
- Organised
- Neither fair nor unfair

Defining Amensalism

Amensalism was the most elusive for participants. Most were unfamiliar with this form of symbiosis. Amensalism is characterised by an interaction where one organism is inhibited or destroyed, while the other remains unaffected. Examples discussed were the Spanish ibex (*Capra pyrenaica*) and weevils of the genus Timarcha, and the Black Walnut (*Juglans nigra*) and surrounding herbaceous plants.

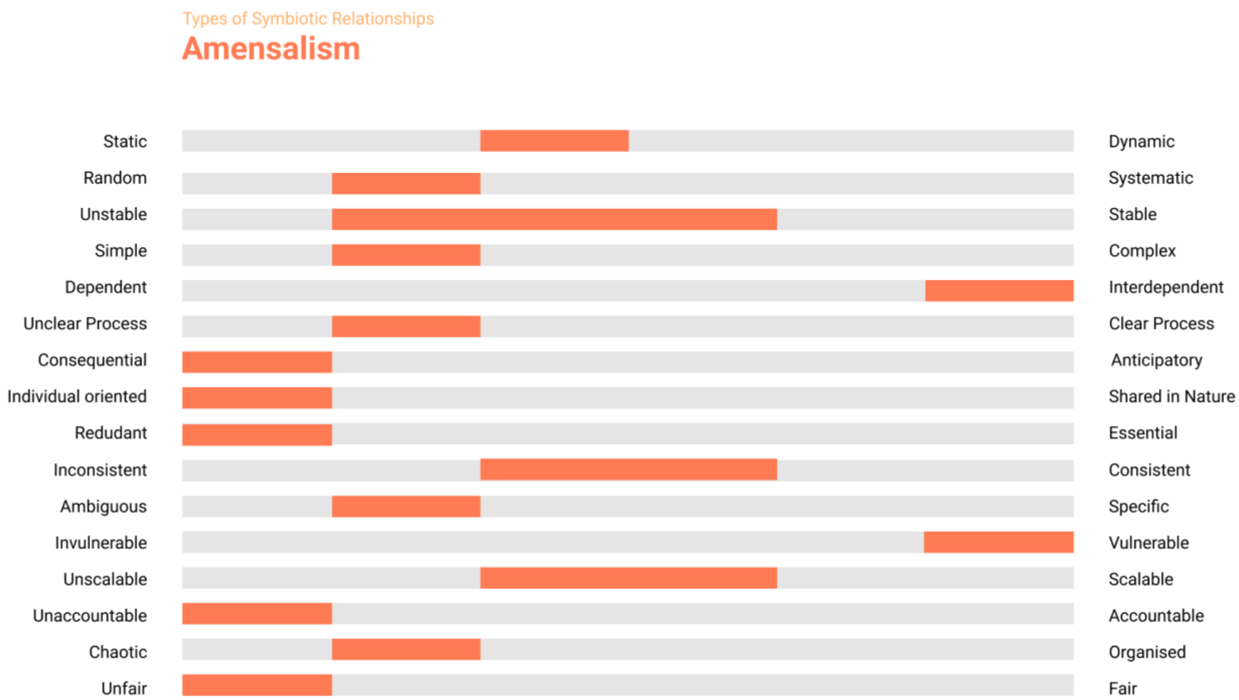


Figure *21* - Attributes associated with Amensalism

The interviewees characterised amensalism with the following attributes, highlighting its complex and nuanced nature:

- Neither static nor dynamic
- Random
- Unstable
- Simple
- Highly interdependent
- Ambiguous

- Significantly consequential
- Predominantly individual-oriented
- Redundant
- Neither consistent nor inconsistent
- Vague
- Extremely vulnerable
- Scalable
- Largely unaccountable
- Chaotic
- Markedly unfair

Defining Mimicry

The discussions on mimicry proved insightful. While mimicry is acknowledged as a biological interaction, there's no consensus on whether it constitutes a symbiotic relationship. It involves a species adopting the traits of another to benefit itself. Prominent examples include the Common Mormon (Papilio polytes) imitating the Common Rose (Pachliopta aristolochiae) and the Viceroy Butterflies (Limenitis Archippus) emulating the Monarch butterflies (Danaus plexippus).

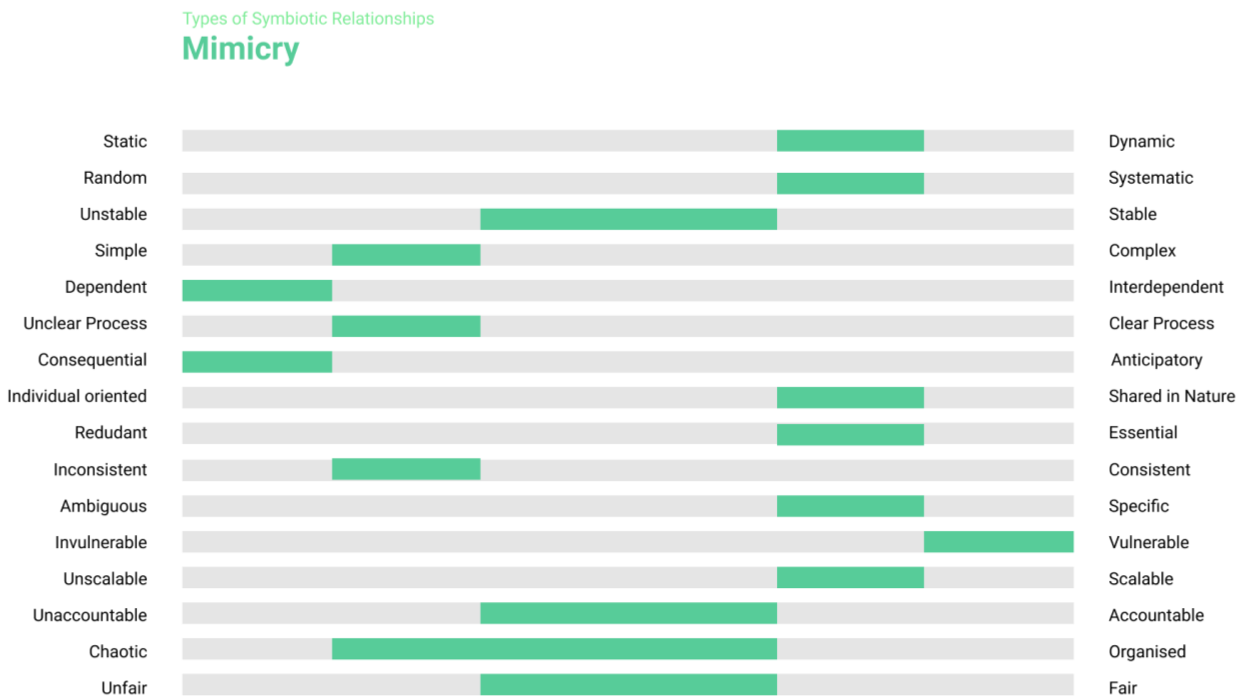


Figure 22 - Attributes associated with Mimicry

In characterising mimicry, the interviewees described it with various attributes, reflecting its multifaceted nature:

- Dynamic
- Systematic
- Neither stable nor unstable
- Simple
- Highly dependable
- Ambiguous

- Consequential
- Shared
- Essential
- Moderately inconsistent
- Specific
- Vulnerable
- Scalable
- Moderately unaccountable
- Mildly chaotic
- Neither fair nor unfair

The data from these interviews is invaluable, deconstructing the symbiotic relationships into distinct attributes. This research elucidates the perceptions of experts on each symbiosis type. This novel methodology harvests attributes representative of biological concepts, integrating them into artificial intelligence via design research. Identifying these attributes allows design researchers to comprehend how they can be translated into machine intelligence. This foundational study aids researchers in visualising various symbiotic relationships, clarifying the often ambiguous nature of biological interactions. Although this research marks an initial venture, mapping the attributes of symbiotic relationships is crucial for delineating the boundaries of each type and addressing the common overlaps in biological interactions. In symbiotic relationships involving humans and AI, there is a preference for logical, ordered approaches, characteristics not always present in biological interactions. Therefore, identifying the attributes of symbiotic relationships is essential for advancing human-AI interaction, facilitating the adaptation of biological knowledge into the realm of artificial intelligence and, consequently, encouraging innovative methods of interaction with intelligent machines.

These interviews not only highlight the differences among symbiotic relationships but also place them within the context of artificial intelligence. A follow-up set of interviews delved into symbiosis in the domain of machine intelligence. The goal was to link types of symbiosis with specific subdomains of machine intelligence and introduce the concept of "symbiotic primers." These primers are guidelines that researchers can use to construct intelligent systems. For example, in developing an image classification system, using a symbiotic primer helps determine the most suitable symbiotic relationship to emulate. This approach ensures that the user experience is shaped by the chosen biological interaction, enhancing the system's design and functionality.

Symbiotic Primers Explored

The second set of interviews was bifurcated into two distinct versions. One version was designed to encapsulate the perspective of an intelligent machine, while the other was tailored to capture the viewpoint of a human interacting with said machine. Both versions encompassed two sections: the first delved into personal information and familiarity with the research topic, while the second probed into symbiotic relationships within the digital ecosystem.

Human Perspective

As previously mentioned, interviewees were randomly assigned one of two versions of the interview. The human-centric version positioned the participant alongside a human in scenarios involving an intelligent machine. Five interviewees evaluated twenty-nine distinct scenarios to determine the most fitting symbiotic relationship for each.

Under the human perspective, mutualism was adjudged most appropriate for:

- Protecting large groups from cyberattacks.
- Identifying and securing endpoint vulnerabilities.
- Designing tailored, personalised learning plans.
- Utilising voice assistants for increased engagement and supporting people with disabilities.
- Developing language assessment tools.
- Employing conversational agents for guidance.

- Creating game artificial intelligence.
- Providing automated diagnostic support, notably in diabetes and skin cancer.
- Allocating medical resources effectively.
- Detecting features in medical imaging.
- Enhancing capabilities with augmented writing software.

In contrast, parasitism was predominantly identified in:

- Facial recognition and biometric analysis.
- Crowd analysis and mass surveillance.
- Development and utilisation of autonomous weapons.
- Production of computer-generated news.

Commensalism, meanwhile, was associated with:

- Automating complex, repetitive tasks previously necessitating human intervention.
- Conducting predictive analysis for self-harming behaviors.
- Managing diverse management systems.
- Streamlining documentation processes.
- Aiding educators in managing classroom environments.
- Automating the mapping of infrastructure, including traffic lights and fire hydrants.

Amensalism was notably prevalent in scenarios involving:

- Stress and aggression detection.

- Demographic analysis.
- Speech recognition.
- Affect recognition.
- Assessment of health risk scores.
- Decision support tools like COMPAS.

Finally, mimicry was deemed most suitable for machine intelligence applications in art and music, with examples including DeepDream and SketchRNNs.

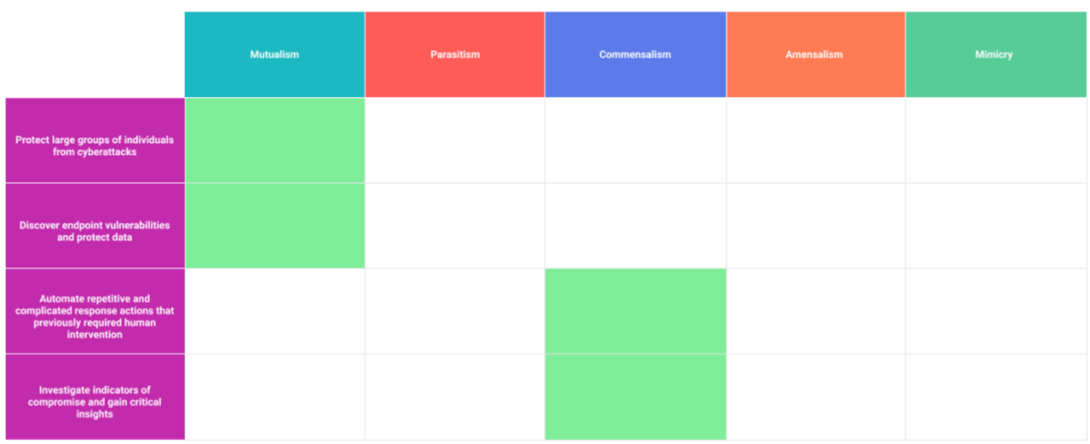


Figure *23* - Association between the first set of machine intelligence tasks and the types of symbiotic relationships from a human perspective

Figure *24* - Association between the second set of machine intelligence tasks and the types of symbiotic relationships from a human perspective

Figure *25* - Association between the third set of machine intelligence tasks and the types of symbiotic relationships from a human perspective

Figure *26* - Association between the fourth set of machine intelligence tasks and the types of symbiotic relationships from a human perspective

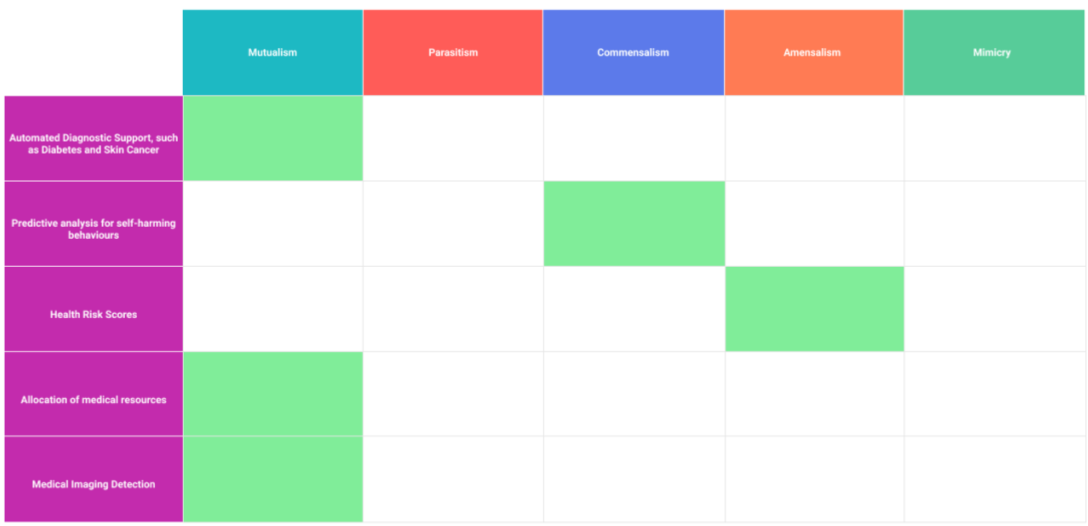


Figure *27* - Association between the fifth set of machine intelligence tasks and the types of symbiotic relationships from a human perspective

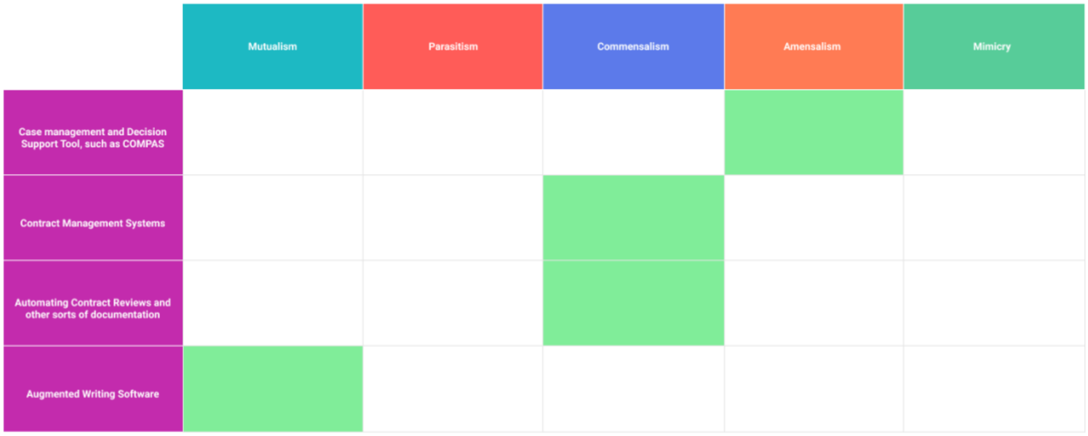


Figure *28* - Association between the sixth set of machine intelligence tasks and the types of symbiotic relationships from a human perspective

Machine Perspective

Contrasting with the human viewpoint, the machine-centric perspective placed interviewees in the role of the intelligent machine. Each participant was presented with the same twenty-nine scenarios, tasked with identifying the most applicable symbiotic relationship for each.

From the standpoint of the machine, mutualism emerged as the preferred symbiosis for:

- Protecting large groups from cyberattacks.
- Detecting endpoint vulnerabilities and safeguarding data.
- Investigating indicators of compromise for critical insights.
- Employing language assessment tools.
- Using conversational agents for guidance.
- Implementing facial recognition and biometric analysis.

- Analysing crowds and conducting mass surveillance.
- Automating the mapping of traffic lights and fire hydrants.
- Developing game artificial intelligence.
- Applying machine intelligence in art and music, for instance, DeepDream and SketchRNNs.
- Supporting automated diagnostics in health conditions like diabetes and skin cancer.
- Allocating medical resources.
- Detecting features in medical imaging.
- Enhancing writing with augmented software.

Parasitism was seen as most fitting for scenarios involving autonomous weapons and the detection of stress or aggression.

Commensalism was notably aligned with:

- Developing tailored, personalised learning plans.
- Enhancing engagement and assistance for people with disabilities through voice assistants.
- Generating computer-aided news.
- Assisting educators in classroom management.
- Automating intricate and repetitive tasks previously needing human input.
- Implementing speech recognition.
- Predictive analysis targeting self-harming behaviours.

- Calculating health risk scores.
- Managing and systematising contract management systems.
- Automating the review of contracts and other documents.

Amensalism was predominantly associated with scenarios such as affect recognition, demographic analysis, and decision support tools like COMPAS.

Interestingly, mimicry was not linked to any of the scenarios evaluated by the interviewees.

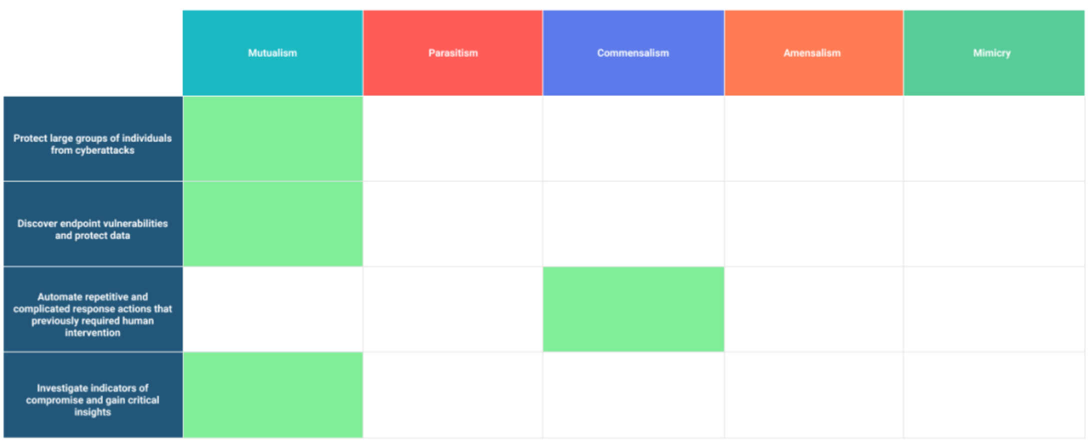


Figure *29* - Association between the first set of machine intelligence tasks and the types of symbiotic relationships from a machine perspective

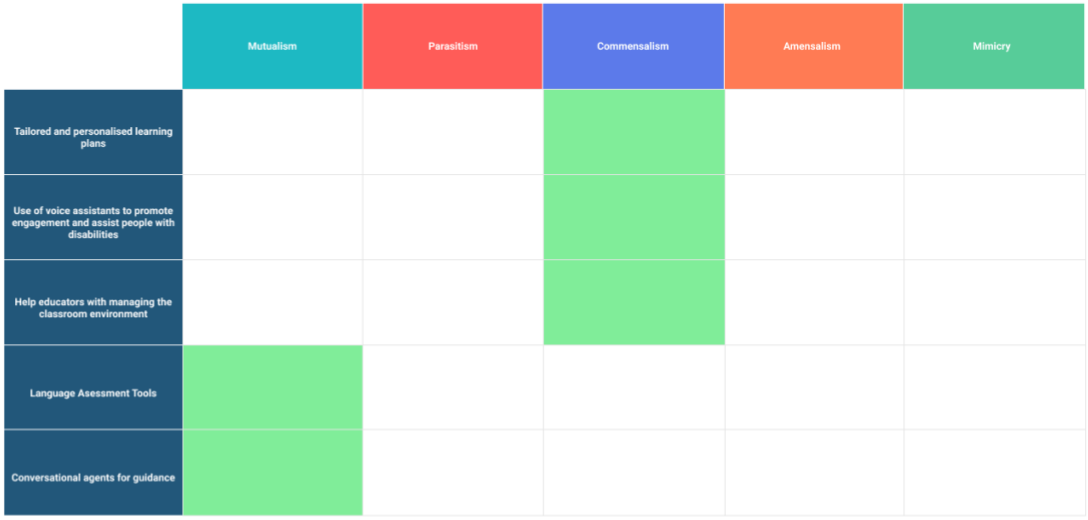


Figure *30* - Association between the second set of machine intelligence tasks and the types of symbiotic relationships from a machine perspective

Figure *31* - Association between the third set of machine intelligence tasks and the types of symbiotic relationships from a machine perspective

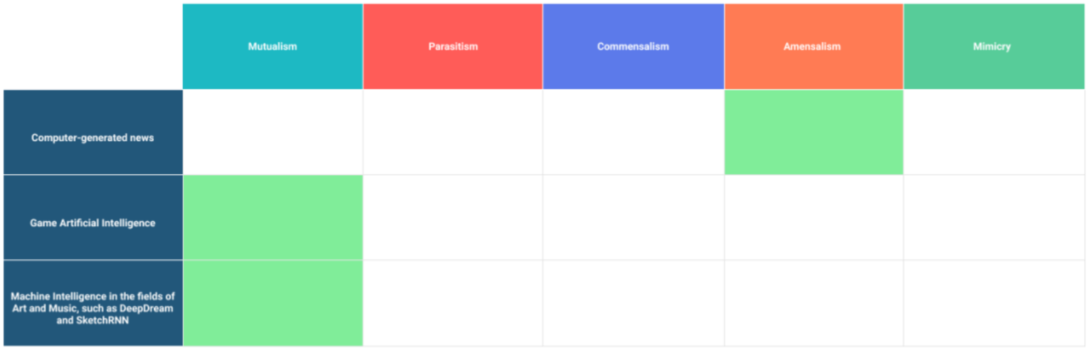


Figure *32* - Association between the fourth set of machine intelligence tasks and the types of symbiotic relationships from a machine perspective

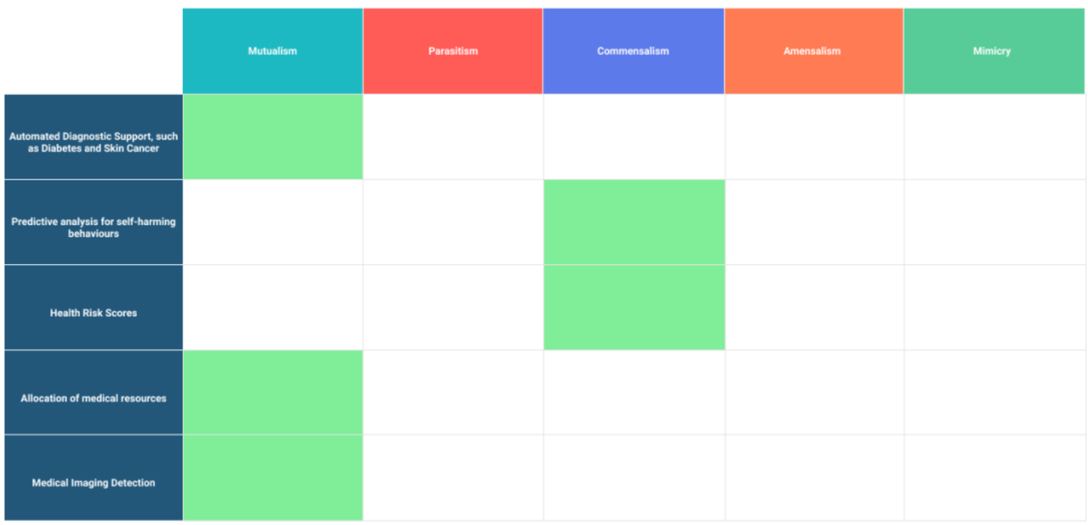


Figure *33* - Association between the fifth set of machine intelligence tasks and the types of symbiotic relationships from a machine perspective

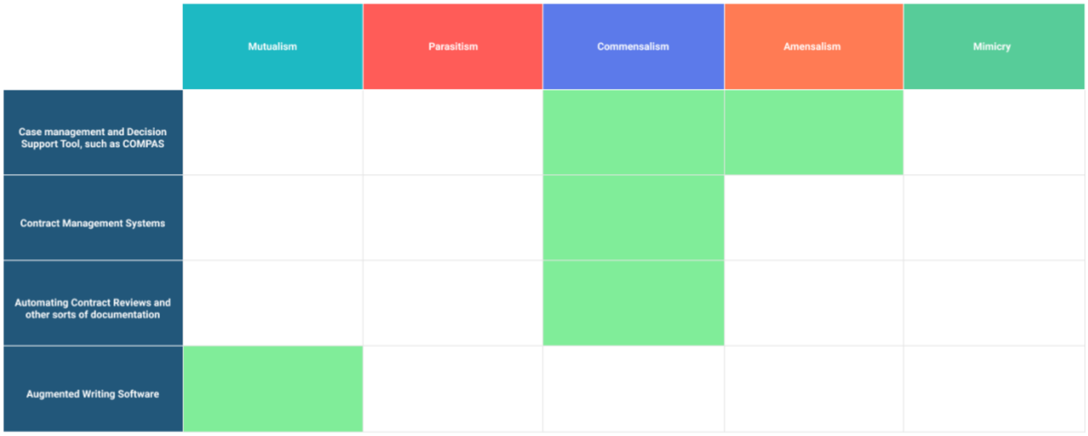


Figure *34* - Association between the sixth set of machine intelligence tasks and the types of symbiotic relationships from a machine perspective

Insights and Challenges

This chapter's research yields insights that can be categorically divided into three primary areas: (1) Understanding and Dissection, (2) Causation, and (3) Possibilities.

Understanding and Dissection

Employing a top-down approach in the analysis and deconstruction of concepts has been both informative and insightful. While this research is still in its preliminary phase, it has successfully begun translating insights from biological symbiosis to the realm of human-machine interactions. Initial findings reveal that various attributes of biological symbiosis can be applied to human-machine relationships, particularly in the field of intelligent task automation. A notable instance is the association of mutualistic symbiosis with enhanced cybersecurity measures in large groups.

Causation

The research goes beyond simple associations, indicating the emergence of distinct patterns linking types of symbiotic relationships with specific tasks in machine intelligence. While larger-scale research is needed to corroborate these findings, early indicators suggest a link between parasitic relationships and the deployment of technologies like facial recognition and biometric data analysis. Expanding upon these patterns could provide invaluable insights for design researchers and machine learning engineers, particularly in terms of transitioning between different types of symbiotic relationships within specific applications. For example, modifying a parasitic relationship in biometric analysis towards a more mutualistic or commensal arrangement could open new avenues in the design and application of these technologies.

Possibilities

The research introduces a fresh perspective on the potentialities within human-machine symbiosis. Interviews and data analysis suggest that certain types of symbiotic relationships may be inherently linked with specific machine intelligence tasks. Understanding the unique characteristics that define each symbiotic relationship could empower researchers and engineers to innovate in the design of machine intelligence tasks, tailoring them more closely to human needs and contexts.

Currently, the findings presented in this chapter represent an initial foray into a broader and more complex field of study. Future research should involve more diverse participant samples and an expanded resource pool to gain deeper, more nuanced insights. Additionally, to engage a broader audience and encourage more dynamic interaction with the study's concepts, alternative methods of data collection and dissemination could be considered. Interactive formats, such as gamification, offer a promising departure from more traditional, formal methodologies like surveys, potentially facilitating more engaging and effective data collection and public outreach.

Reflections

Reframing Methods: Probes Over Interviews

Upon reflection, it becomes evident that the term "interviews" may not entirely capture the essence of the methodological approach adopted in this research. Instead, a more fitting descriptor would be "probes." The term "probe" emerged in design research during the 1970s as a way to move beyond traditional interviews and surveys (Gaver et al., 1999). Probes are tools or techniques used to inspire participants, eliciting richer responses that provide deeper insights into their lives, values, and thoughts.

In this context, the interactions with participants were not merely structured question-answer sessions. The probes, which took the form of fictional contextual scenarios, functioned as conversation starters, designed to elicit deeper, more nuanced insights into the symbiotic relationships between humans and intelligent machines.

The fictional contextual scenarios presented participants with various situations. The first set explored biological interactions of symbiotic relationships, such as mutualism. Participants were asked to explore their understanding of what attributes shape this type of relationship.

The second set of probes involved presenting a number of contextual scenarios. Participants encountered these scenarios from two perspectives: a human perspective and a machine intelligence perspective. The goal was to not only elicit responses pertaining to the type of symbiotic relationship they associated with each scenario but also to gauge their reactions and ethical considerations.

The use of probes allowed for a more interactive, dynamic, and participant-driven exploration of the topic. Participants were not just passive respondents and they actively engaged with the scenarios, often reflecting, questioning, and even challenging the premises presented to them. This methodological shift from interviews to probes underscores the exploratory and participatory nature of the research, emphasising co-creation of knowledge with participants.

Insightful Over Representative Reflection

Another pivotal reflection pertains to the nature of the experiments conducted. While the term "experiment" often connotes a controlled, replicable investigation aiming for generalisable results, the experiments in this research were more heuristic in nature.

They were designed to provide deep insights into specific contexts rather than produce universally applicable outcomes.

The value of these experiments lies not in their representativeness but in their ability to shed light on the intricate nuances of human-machine symbiosis. Each interaction, response, and reflection from participants added layers of understanding to the complex tapestry of human-AI relationships. The experiments were not about validating predefined hypotheses; they were about exploring the vast, multifaceted landscape of symbiotic interactions, uncovering hidden patterns, challenges, and opportunities.

This reframing from representative to insightful underscores the research's commitment to depth over breadth. It also acknowledges the limitations of the experiments in terms of broad generalisability but celebrates their strength in providing rich, detailed insights that can inform and inspire future research and design endeavours.

Challenges and Future Endeavours

This project encountered a multitude of challenges, among which the COVID-19 pandemic was notably significant. Undertaken during a historically unprecedented global crisis, participant recruitment and methodological execution were profoundly impacted. The necessity to adapt the research design for remote implementation was paramount, ensuring all participants experienced consistent conditions. As a result, interviews were identified as the most feasible method, while other techniques, such as hands-on design workshops, became impractical due to logistical constraints.

The shift to remote research required not only a methodological but also a conceptual reorientation. The adoption of probes instead of traditional interviews responded to the challenges of conducting research in a virtual environment. Probes, offering greater flexibility and adaptability to the unpredictable nature of online interactions, proved more effective under these circumstances.

Moreover, the project's initial ambition and a certain degree of naivety presented additional challenges. The human-machine symbiosis research landscape is vast and relatively unexplored, demanding a broader, more in-depth investigation. This study, initially conceived as a preliminary inquiry, aimed to inspire further scholarly exploration in this area. The complexities and expansiveness of the field necessitate diverse research methodologies and perspectives, particularly with a view towards long-term impacts.

The challenges encountered, rather than hindering progress, enriched the research experience. These hurdles necessitated a deeper reflection on and adjustment of the chosen methods, leading to a more nuanced approach. This reframing transformed interviews into dynamic probes and experimental explorations into valuable insights.

Despite its preliminary stage, the significance and potential of this research are unmistakable. Participant feedback during the probes highlighted the widespread interest and high expectations surrounding the topic of human-machine symbiosis. Recognising the research's nascent state, there is a clear understanding that this work represents a foundational blueprint rather than a conclusive exploration. Most participants, being active researchers themselves, appreciate that a domain of such magnitude requires extensive collaborative efforts and diversified resources.

As this research journey progresses, it acknowledges the uncharted territories ahead. While the current findings lay a critical foundation, they mark merely the beginning of a broader scholarly quest. Future research promises deeper dives, more varied participant involvement, and innovative approaches. The path ahead, though fraught with challenges, is ripe with opportunities for groundbreaking discoveries and advancements.

Chapter 06 / Evaluating Privacy and Surveillance in Human-AI Systems through a Symbiotic Lens

“Practice without thought is blind. Thought without practice is empty.’

Kwame Nkrumah

The previous chapter introduced the Symbiotic Human-AI approach. In this chapter, the researcher explores the potential of the design framework in the context of privacy and surveillance. Artificial Intelligence systems, burgeoning in their capabilities and applications, have engendered significant concerns regarding privacy and surveillance. The ubiquity of technologies such as facial recognition, targeted advertising, and predictive policing in everyday life sharply underscores their potential for invasive data practices, posing a profound challenge to the sanctity of user privacy.

While AI technologies offer undeniable benefits and enhanced utility across various sectors, their data management practices raise concerns about user privacy. These concerns are intrinsic to the way Artificial Intelligence systems function, often requiring the collection and analysis of vast amounts of personal data (Mittelstadt et al., 2016; Wachter et al., 2021). In the security and access control domain, facial recognition can improve security and access control in restricted areas, expedite identity verification processes, and even assist with missing person investigations. However, its widespread use raises concerns about potential misuse, mass surveillance, and biased algorithms (Berk et al., 2021). In the targeted advertising space, Artificial Intelligence personalises content by analysing user data, delivering targeted advertising that can be more relevant and engaging (Lee & Cho, 2020). However, such practices raise concerns about user profiling, manipulation, and the potential for creating "filter bubbles" that limit exposure to diverse viewpoints (Pariser, 2011). In the Predictive Policing realm, algorithms analyse data to identify areas with a higher likelihood of crime, potentially leading to more efficient allocation of law enforcement resources and reduced crime rates (Pastaltzidis et al., 2022). However, these algorithms can perpetuate existing biases in policing practices and lead to discriminatory profiling (Selbst et al., 2019).

It should be acknowledged that these potential benefits come at a cost. The intrinsic invasiveness of their data management practices necessitates a critical examination of their impact on privacy. A.G. Ferguson's seminal work, The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement (Ferguson, 2017), critiques the transformation of law enforcement practices through predictive policing and big data. Ferguson highlights the privacy implications and potential biases inherent within these systems, assessing their impact across various communities and the resultant implications for individual privacy. Additionally, The Intersection of Targeted Advertising and Security: Unraveling the Mystery of Overheard Conversation (Bouke et al., 2023) investigates the challenges posed by Artificial Intelligence driven targeted advertising and big data analytics to the notion of user anonymisation and privacy. The detailed analysis elucidates the inner workings of targeted advertising and its potential for breaching user privacy. Furthermore, Buolamwini and Gebru (2018), in Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification, explore the privacy and ethical concerns raised by facial recognition technology. Their research into accuracy disparities in commercial Artificial Intelligence systems based on gender and race highlights the privacy risks and inherent biases in these technologies.

O. Tene and J. Polonetsky's insightful article, *Big Data for All: Privacy and User Control in the Age of Analytics* (Tene & Polonetsky, 2012), scrutinizes the big data analytics paradigm, including Artificial Intelligence applications such as facial recognition and targeted advertising. They propose a nuanced framework aimed at reconciling big data innovation with the imperatives of user privacy and control. S. Zuboff's landmark treatise, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (Zuboff, 2018), is indispensable for understanding the commodification of data in the era of surveillance capitalism. Zuboff's work articulates the pernicious privacy issues engendered by Artificial Intelligence's capabilities in collecting, analysing, and monetising extensive user data, revealing the economic and power dynamics fueling privacy concerns within the AI milieu.

Frank Pasquale's *The Black Box Society: The Secret Algorithms That Control Money and Information* (2015) sheds light on the opacity surrounding Artificial Intelligence systems, particularly concerning decision-making processes that significantly affect our lives. This exploration of the "black box" nature of these algorithms underscores the substantial surveillance and privacy concerns stemming from a lack of public understanding and scrutiny. Moreover, Cathy O'Neil's *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (2016) critically appraises the guise of objectivity and neutrality often ascribed to AI. Her examination highlights how AI, especially through predictive models, can exacerbate privacy invasions, social inequality, and biases.

On a more practical implementation level, the European Commission's High-Level Expert Group on Artificial Intelligence's "Ethics Guidelines for Trustworthy AI" provides a comprehensive framework for the ethical, secure, and privacy-respecting development and deployment of AI systems. This document addresses key issues around data governance, privacy, and surveillance, offering foundational guidelines for mitigating these challenges in AI implementations.

These literary works collectively illustrate a multifaceted analysis of the intricate issues surrounding privacy and surveillance in the realm of Artificial Intelligence. The works cover technical, ethical, individual, societal, legal, and philosophical dimensions, offering a comprehensive foundation for understanding and addressing these critical concerns.

One of the most poignant criticisms of Artificial Intelligence is the lag in public oversight relative to the rapid development of AI technologies and the challenges that this lag exacerbates. The proliferation of opaque systems enables unchecked surveillance, further fueled by a pronounced asymmetry in information access between users and system operators. Privacy advocates and artificial intelligence scholars are increasingly vocal about the emerging "surveillance gap," a widening chasm in the understanding and control of privacy between users and the architects of Artificial Intelligence systems.

Revisiting Shoshana Zuboff's *The Age of Surveillance Capitalism*, the author provides an incisive critique of how personal data becomes a commodity in our digital age. The author’s analysis underscores the erosion of privacy as a consequential outcome of the digital revolution, offering crucial context to the urgency and magnitude of the privacy issues exacerbated by technological advancements. In parallel, the Human-AI symbiosis framework introduced in the previous chapter, delineates relationships based on symbiotic archetypes such as mutualism, commensalism, and parasitism, and offers an enlightening perspective on the interplay between humans and intelligent machines. This framework has the potential to become instrumental in explicating the dynamics of this interplay and necessitating the alignment of these relationships for ethical Artificial Intelligence development and deployment.

Yan Shvartzshnaider, Helen Nissenbaum and others advocate for the integration of human values like privacy into the Artificial Intelligence development matrix (Shvartzshnaider et al., 2019). The authors call for human-centered design principles as foundational in mitigating risks related to surveillance and privacy breaches. The approach introduced in the previous chapter, grounded in biological symbiotic relationships, enables a more nuanced understanding of the disconnects between user perspectives and the operational algorithms of Artificial Intelligence systems. Additionally, *The Oxford Handbook of Ethics of AI* (Dubber et al., 2020), expands on the ethical discourse of surveillance and privacy concerns in Artificial Intelligence. The anthology discusses the imperative to establish ethical frameworks and governance models for Artificial Intelligence, ensuring there is an alignment with human values and societal norms.

This chapter expands on the concept of the human-AI symbiosis framework introduced in the previous chapter and explores the framework as a novel tool to examine privacy and surveillance issues in Artificial Intelligence. This approach not only reveals the disconnects between user perspectives and Artificial Intelligence system logic but also highlights the potential for design research and participatory and inclusive approaches in technology design.

Privacy and surveillance issues in Artificial Intelligence continue to expose a number of ethical quandaries surrounding intelligent machines such as predictive policing. Predictive policing, in specific, is scrutinised in Andrew Guthrie Ferguson's *The Rise of Big Data Policing*. The author describes the consequences of the rise of big data surveillance and underlines the complexities and exigencies surrounding privacy issues in Artificial Intelligence applications. Ferguson provides a critical examination of predictive policing and exposes ethical and privacy challenges, reinforcing even stronger the necessity of realigning intelligent machines with human-centric values (Ferguson, 2017b).

The chapter seeks to enrich our understanding of the delicate equilibrium between advancing Artificial Intelligence technologies and safeguarding individual privacy. By employing the human-AI symbiosis framework proposed in the preceding chapter, the chapter aims to introduce novel approaches to how Artificial Intelligence systems are conceived and governed, advocating for the development of technology that is not only cutting-edge but also ethically sound and respectful of privacy and human rights.

Methods

The chapter aims to delve into the perceptions and dynamics of the human-AI symbiotic relationship through an experimental design that focuses on various artificial intelligence applications. The purpose of the study is to not only contribute to our understanding of human interactions with Artificial Intelligence but also shed light on the perceived roles and characteristics of these interactions from both human and Artificial Intelligence perspectives.

Development of Hypothetical AI Application Scenarios as Probes

The experimental design involves the development of fifteen hypothetical contexts encompassing various applications of Artificial Intelligence, including facial recognition, targeted advertising, and audio surveillance. These scenarios were meticulously constructed to serve as probes, enabling the extraction of subjective interpretations and insights into the nature of human-AI relationships.

Each contextual scenario was designed to elicit responses on specific aspects pertaining to the relationships between intelligent machines and humans. The contextual scenarios took into consideration context and environment, interactivity and engagement, and emotional and ethical responses. For context and environment, the researcher paid attention to the setting in which the Artificial Intelligence system was used (home, workplace or public space) and the perceived impact of the system on the scenario. For interactivity and engagement, the primary question related to how humans in the scenario interacted with the intelligent machine. The research was focused on establishing whether it was a passive experience or the human in the scenario was actively engaging with the intelligent machine. Lastly, for emotional and ethical responses, the scenarios were crafted to incite emotional reactions, such as comfort, fear, anger, while also addressing ethical considerations such as fairness, justice, and rights.

The purpose of the contextual scenarios was multifaceted. First, the research aimed at eliciting subjective interpretations. With the participants engaging with an hypothetical yet relatable situation, they were able to project their thoughts, feelings, and attitudes towards these intelligent machines, and consequently reveal their underlying biases, hopes and fears. Second, after introducing the Human-AI symbiotic framework in the previous chapter, it made sense to investigate how the framework would work in real life scenarios. Since there were a broad number of scenarios to explore, using contextual scenarios made more sense in order to reach a larger breadth and depth of scenarios. The responses in these scenarios aim to provide a window into how different types of Artificial Intelligence are perceived in terms of the symbiotic relationships they create. In other words, the research project wants to understand how participants view intelligent machines as beneficial partners (mutualism), harmless co-inhabitants (commensalism), harmful co-inhabitants (ammensalism) or exploitative entities (parasitism). Third, the contextual scenarios were a tool for understanding perceived attributes. Beyond the type of relationship, the scenarios also aimed to elucidate what were the attributes of intelligent machines (transparency, control, autonomy, intelligence) that influenced the nature of these perceived relationships.

This experimental design, with the diverse breadth of Artificial Intelligence applications and nuanced construction of scenarios, stood to offer deep insights into the societal and psychological dimensions of our growing coexistence with Artificial Intelligence. By exploring a range of Artificial Intelligence technologies through the lens of human-AI symbiosis, the study contributed to a more holistic understanding of not just how these technologies function, but how they integrate into, shape, and are shaped by human society, perceptions, and values. Additionally, this in turn can guide to more ethical, user-centric, and socially conscious approaches in artificial intelligence development and policy-making.

Hypothetical Artificial Intelligence Contextual Scenarios

As mentioned previously, the design of this study revolves around the intricate development of fifteen hypothetical contextual scenarios, each carefully tailored to probe into the diverse breadth of Artificial Intelligence applications and the corresponding human-AI symbiotic relationships. The contextual scenarios are not simple fictional narratives, but instead they are constructed as realistic, plausible instances of Artificial Intelligence integration into everyday life, ensuring relevance and resonance with the participants. These scenarios range a spectrum of Artificial Intelligence applications, exemplified in the figures below:

Contextual Scenario no.01
Facial Recognition at Retail Stores

The first contextual scenario involves Artificial intelligence powered cameras in stores with the task of identifying customers, offer personalised shopping suggestions based on their past purchases and browsing history.

Contextual Scenario no.02
Artificial Intelligence in Predictive Policing

The second contextual scenario involves an artificial intelligent system responsible for analysing historical crime data and social media to predict potential crime hotspots and advising police patrols accordingly.

Contextual Scenario no.03
Targeted Advertising in Virtual Reality (VR)

The third contextual scenario operates in the virtual reality space. While interacting with a VR platform, users encounter Artificial Intelligence curated advertisements that morph based on their reactions and interactions within the Virtual Reality environment.

Contextual Scenario no.04
Artificial Intelligence as a Personal Health Assistant

The fourth contextual scenario involves the use of a wearable device using Artificial Intelligence that monitors health parameters, offers real-time health advice and alerts the user to potential medical issues.

Contextual Scenario no.05
Automated Job Recruitment Tool

The fifth contextual scenario illustrates an Artificial Intelligence system with the ability to screen resumes, conduct initial virtual interviews, and rank candidates based on verbal and non-verbal cues.

Figure 35 – First set of Contextual Scenarios

Contextual Scenario no.06
Artificial Intelligence Tutor for Online Education

The sixth contextual scenario explores the interaction with an Artificial Intelligence tutor able to provide personalised learning experiences, content adaptation and teaching styles based on the student's learning pace and preferences.

Contextual Scenario no.07
Smart Home Security Using Audio Surveillance

The seventh contextual scenario involves smart home devices constantly listening to ambient sounds, identifying unusual noises or voices for security purposes in an individual's home.

Contextual Scenario no.08
Artificial Intelligence Driven Traffic Management System

The eighth contextual scenario depicts an Artificial Intelligence system with the ability to analyse traffic flow data, adjust traffic lights and suggest optimal routes to drivers to reduce congestion.

Contextual Scenario no.09
Artificial Intelligence in Elderly Care Robots

The ninth contextual scenario pertains to robots powered by Artificial Intelligence that assist in caring for the elderly, provide companionship, remind them of medications, and alert authorities in case of emergencies.

Contextual Scenario no.10
Automated Social Media Content Moderation

The tenth contextual scenario involves an Artificial Intelligent system that monitors and flags inappropriate or harmful content on social media platforms, based on evolving guidelines and user feedback.

Figure 36 - Second set of Contextual Scenarios

Contextual Scenario no.11
Voice-Activated Virtual Assistants in the Workplace

The eleventh contextual scenario depicts office workers using voice-activated Artificial Intelligence assistants for scheduling, email management, and information retrieval, shaping daily workflow and productivity.

Contextual Scenario no.12
Emotion Recognition for Customer Feedback

The twelfth contextual scenario involves an Artificial Intelligence system analysing customer facial expressions and tone of voice to gauge satisfaction levels in retail or service environments.

Contextual Scenario no.13
Personalised Newsfeed Algorithms

The thirteenth contextual scenario illustrates an Artificial Intelligence system that customises newsfeeds for users on a digital platform, influencing the diversity of information they receive based on their past interactions.

Contextual Scenario no.14
Artificial Intelligence in Autonomous Vehicles

The fourteenth, or penultimate, contextual scenario is broader in terms of context compared to other scenarios. This scenario revolves around self-driving cars and how these vehicles use Artificial Intelligence to navigate, make decisions in real-time about routes, and react to road conditions, affecting passenger safety and urban mobility.

Contextual Scenario no.15
Artificial Intelligence for Environmental Monitoring

The fifteenth, or last, contextual scenario explores the use of drones equipped with Artificial Intelligence collecting data on environmental parameters like air quality, deforestation rates, and wildlife movements, aiding in conservation efforts.

Figure 37– Third set of Contextual Scenarios

The fifteen contextual scenarios described above encompass a range of applications and highlight different aspects of the Human-AI relationship, each with unique implications for privacy, ethics, trust, and societal impact.

Recruitment and Participants

For the study, 6 participants were recruited from previous research activities. The recruitment strategy aimed to gather a breadth and depth of experience, from people with expert knowledge in artificial intelligence to individuals with a background in design. To enhance the generalisability of the findings, the study attempted to recruit as diverse a demographic as possible. Participant ages spanned from 18 to 40 years old, with a majority (66%) identifying as male. The professional roles were diverse, including human-computer interaction researchers, machine learning engineers, and product designers. Similar to previous research activities, interviewees had a minimum of one year of experience in their respective fields, with the majority possessing over four years of experience. Geographically, participants were exclusively from Europe, as attempts to include more diverse regions were hindered by a scarcity of direct contacts outside Europe and logistical challenges linked to time zone differences. The researcher believes the diversity in the participant sample is sufficient to ensure a broad spectrum of insights into Human-AI interactions, as exemplified in the research activities described in the former chapter.

In terms of knowledge criteria, participants were selected based on their domain knowledge expertise and some prior exposure to the Human-AI symbiotic framework. While recruiting participants with no prior knowledge was considered to minimise bias and ensure the freshness and authenticity of responses, the researcher ultimately decided that prior exposure would generate deeper insights due to participant familiarity.

Data Analysis

An inductive thematic analysis approach was employed for the methodological approach. This qualitative research method is particularly well-suited for uncovering and understanding patterns within data, enabling the researcher to develop a rich, detailed, and complex account of the information gathered (Braun & Clarke, 2006). Unlike deductive thematic analysis which uses predetermined themes, inductive analysis allows themes to emerge directly from the data itself.

Thematic analysis in this study involved a three-pronged categorisation and coding schema focused on the following variables:

Categorisation by Application Type: The contextual scenarios presented to participants were categorised based on the Artificial Intelligence application or system they represented. This categorisation facilitates a more structured analysis and aids in identifying application-specific themes or patterns within the data.

Coding of Relationships and Attributes: Responses from participants were coded to identify and solidify the association of the presented contextual scenarios with different types of symbiotic relationships (e.g., mutualistic, parasitic). The coding process also served to collect key attributes relating to these perceived symbiotic relationships. This involved dissecting data into meaningful segments, which were then labelled with codes to aid in theme identification.

Frequency Analysis: Additionally, the frequency of certain responses or themes was tracked to identify dominant perceptions and viewpoints among participants. This allowed for an understanding of how frequently certain ideas or perspectives arose within the data set.

The final step in conducting inductive thematic analysis involved synthesising the coded data to extract overarching patterns and insights. The synthesis was not merely an aggregation of data but an interpretative process aimed at understanding the underlying meanings and implications of the findings. For instance, if a significant number of participants perceived a relationship with facial recognition technology as parasitic, this association might reflect broader societal apprehensions regarding privacy and autonomy in the context of Artificial Intelligence.

The experimental study described in this chapter aimed to contribute significantly to the field of Human-AI interaction research. By exploring symbiotic relationships from a dual perspective – human and intelligent machine – the study endeavours to unveil nuanced understandings of how intelligent machines are perceived, experienced, and conceptualised in various socio-technical contexts. The goal is for the insights collected to become useful for designers, developers, policymakers, and other knowledge domain experts in the core and peripheral spaces of Artificial Intelligence, assisting them in delivering more ethical, user-centred, and socially aware Artificial Intelligence systems.

Results

In this research project, the researcher embarked on an exploratory study to dissect the multifaceted perceptions and dynamics inherent in human-AI symbiotic relationships across fifteen unique Artificial Intelligence application scenarios. Each scenario was analysed based on human perspectives and attributed characteristics to the artificial intelligence systems, thereby unraveling the complexities of these techno-social interactions.

The advent and integration of Artificial Intelligence in a variety of sectors, from retail and healthcare to security and environmental monitoring, pose profound questions about the nature of human-AI relationships. This project aims to understand Human-AI relationships through an interpretative lens, classifying them into symbiotic categories and investigating underlying themes and attributes.

Methodology

Each of the fifteen Artificial Intelligence application scenarios were constructed to probe participants' views. The participants, although familiar to a predefined framework regarding Human-AI symbiosis, provided fresh, interesting and unbiased insights. The collected insights were then qualitatively analysed, adhering to an inductive thematic approach to identify overarching trends and individual scenario nuances.

Results and Discussion

The researcher went through each individual contextual scenario and summarised the insights collected from participants. The gathered insights were divided into two groups, a human perspective and an intelligent machine perspective.

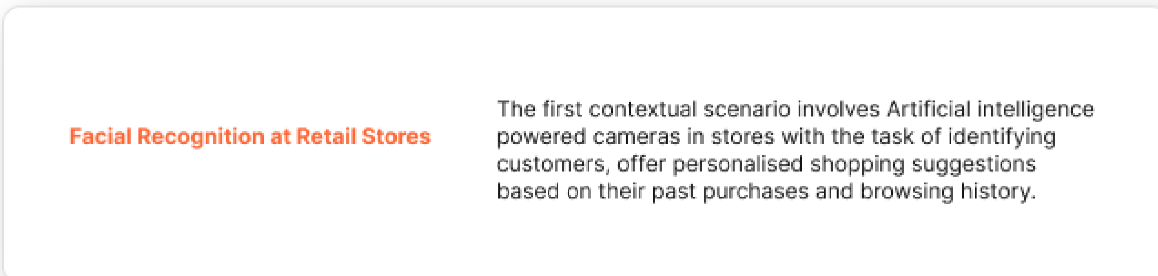


Figure 38 - Detailed Description of 1st Contextual Scenario

In the first contextual scenario, participants described facial recognition at retail stores from a human perspective as being parasitic, portraying a lack of accountability and fairness, and fostering a sense of chaos and invasion of privacy. On the other hand, the intelligent machine perspective was described as mutualistic, systematic, focusing on a dynamic and interdependent nature, and ultimately highlighting Artificial Intelligence's potential in personalising shopping experiences efficiently.

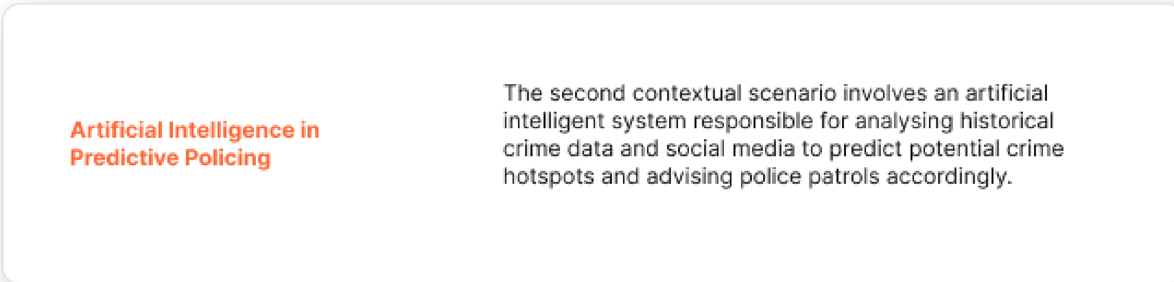


Figure 39 - Detailed Description of 2nd Contextual Scenario

From a human lens, participants viewed Artificial Intelligence in Predictive Policing as being a sensitive area and they were particularly concerned over the system's potential biases, infringement on personal freedoms, and the ethical implications on surveillance. On the other hand, from an intelligent machine lens, the contextual scenario was viewed as a tool for enhancing safety and efficiency in law enforcement, but the nuance views on its predictive accuracy and fairness were seen as pivotal elements for maintaining a symbiotic relationship.

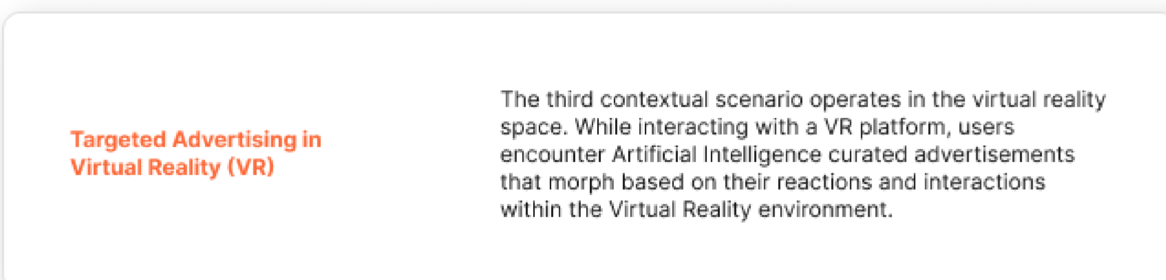


Figure 40 - Detailed Description of 3rd Contextual Scenario

In this contextual scenario and from a human perspective, participants stated they were likely to express discomfort or feeling manipulated by the intelligent system. Participants were also concerned about privacy and the intrusive and pervasive nature of advertisement morphing based on an individual's reactions. From an intelligent machine perspective, this scenario was labelled as an innovative, responsive tool that dynamically tailors advertisements and consequently enhances the user experience and marketing effectiveness.

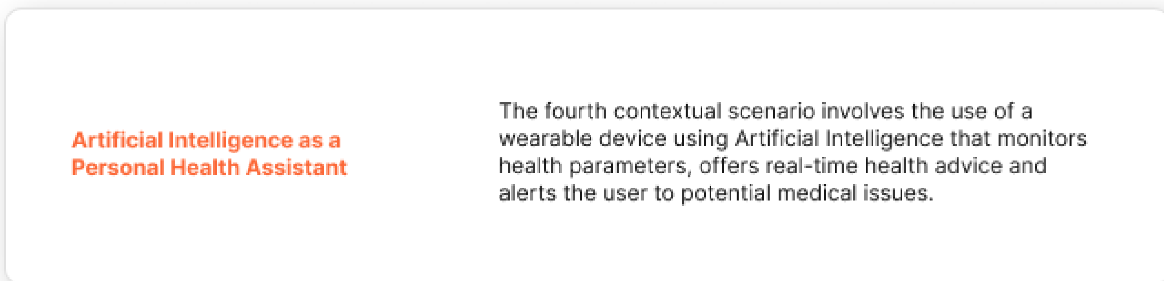


Figure 41 - Detailed Description of 4th Contextual Scenario

From a human perspective, Artificial Intelligence as a Personal Health Assistant was perceived positively. Participants appreciated the ability of real-time health monitoring and potential lifesaving alerts, but they demonstrated concerns about over-dependence and persistent data privacy concerns. On the other hand, the intelligent machine perspective was regarded as a beneficial, proactive healthcare partner, and instrumental in promoting preventive health measures and personalised care.

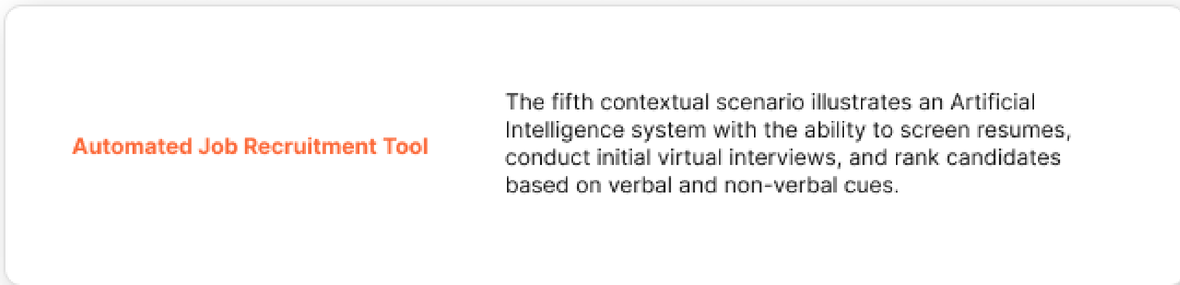


Figure 42 - Detailed Description of 5th Contextual Scenario

In this contextual scenario, the human perspective was clouded by skepticism. The participants were skeptical about the tool's ability to fairly and effectively interpret candidates qualification and soft skills, hinting at potential serious biases. From an intelligent machine perspective, the intelligent system perspective was seen as efficient, objective, and capable of streamlining the recruitment process by analysing large amounts of data to identify suitable candidates.

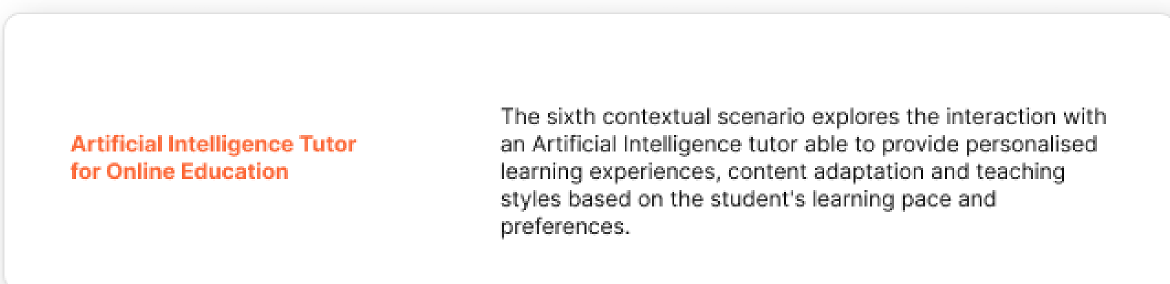


Figure 43 - Detailed Description of 6th Contextual Scenario

On the sixth contextual scenario, the human perspective valued the personalised learning style and adaptability to each individual's pace and style. However, participants demonstrated concerns around the lack of human interaction and direct mentorship. The intelligent machine perspective considered an Artificial Intelligence tutor for online education to be an effective educational tool, adaptable, responsive to student needs, and with the potential of reshaping education accessibility and customisation.

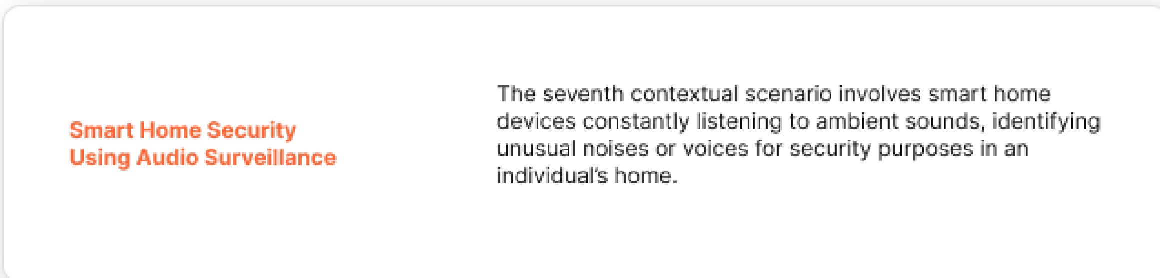


Figure 44 - Detailed Description of 7th Contextual Scenario

From a human lens, participants understood the security benefits but there was some degree of concern and apprehension about pervasive and constant surveillance, invasion of privacy, and misuse of sensitive data. On the other hand, the intelligent machine perspective perceived the solution as a continuous vigilant and effective security measure able to autonomously monitor potential threats and enhance the safety of homes.

Artificial Intelligence Driven Traffic Management System

The eighth contextual scenario depicts an Artificial Intelligence system with the ability to analyse traffic flow data, adjust traffic lights and suggest optimal routes to drivers to reduce congestion.

Figure 45 - Detailed Description of 8th Contextual Scenario

In this scenario, the human perspective identified Artificial Intelligence Driven Traffic Management Systems as generally beneficial for reducing traffic congestion and improving commute times. At the same time, there were some concerns about the reliability of the system and data privacy. From an intelligent machine perspective, these systems were described as an intelligent approaches for optimising traffic flow and public transport efficiency, and ultimately have the potential to transform urban mobility.

Artificial Intelligence in Elderly Care Robots

The ninth contextual scenario pertains to robots powered by Artificial Intelligence that assist in caring for the elderly, provide companionship, remind them of medications, and alert authorities in case of emergencies.

Figure 46 - Detailed Description of 9th Contextual Scenario

The human lens in this scenario demonstrated an appreciation for the support and assistance these systems would be able to provide. However, similar to other contextual scenarios, there were some concerns, particularly with the emotional adequacy of the systems and the ethical implications of substituting human caregivers. On the other hand, the intelligent machine lens portrayed these systems as being compassionate and reliable assistants with the ability to aid in routine care and emergency situations, ultimately enhance the quality of life for the elderly.

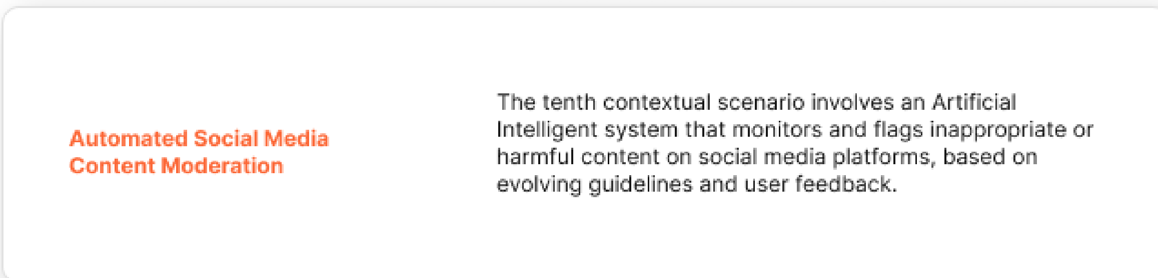


Figure 47 - Detailed Description of 10th Contextual Scenario

In the tenth contextual scenario, the human perspective mainly demonstrated concerns over censorship, biases and the ability of intelligent machines to understand contextual information and nuances in content. The intelligent machine perspective portrayed this application as a necessary tool for managing large volumes of content, maintaining community standards, and protect users from harmful content.

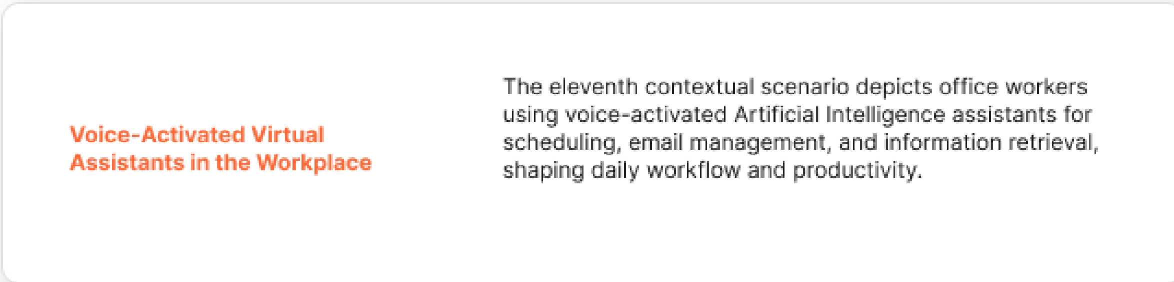


Figure 48 - Detailed Description of 11th Contextual Scenario

From a human lens, the eleventh contextual scenario was recognised for the potential to increase productivity and efficiency in administrative tasks. At the same time, participants raise issues about the accuracy of the system, misunderstanding context, and potential job displacement. On the other hand, from an intelligent machine lens, the system was considered to be effective, time-efficient, and able to help streamline workflow and task management.

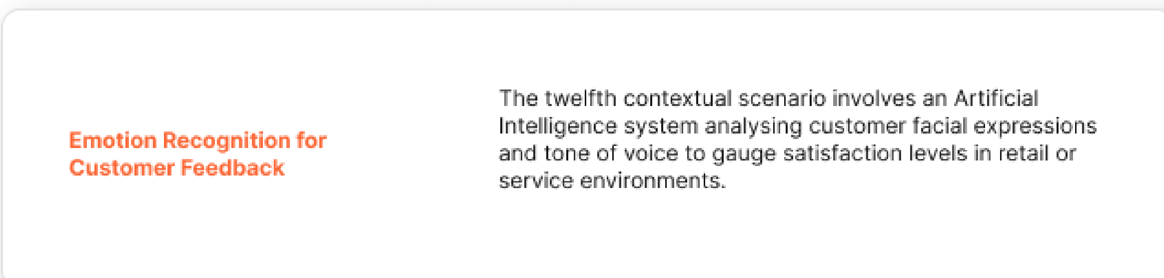


Figure 49 - Detailed Description of 12th Contextual Scenario

This contextual scenario brought mixed feelings across participants. From a human perspective, participants were intrigued by the potential to improve customer service but they were also concerned with privacy and the accuracy in interpreting human emotions. The intelligent machine perspective brought up a similar sentiment. The system was seen as a tool to increase customer engagement and understanding, but participants once again were concerned with the accuracy in reading human emotions.

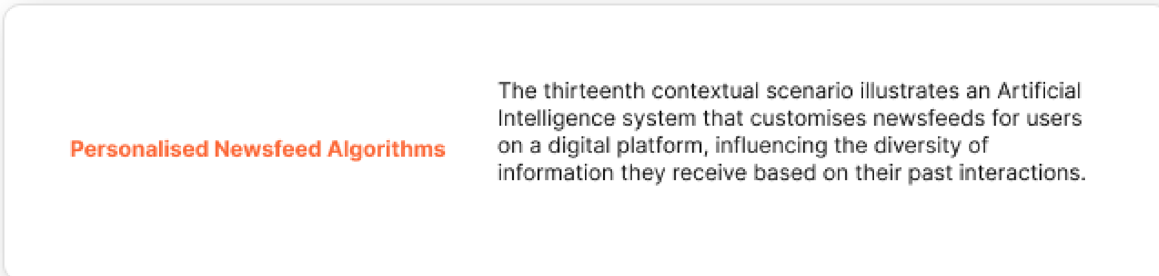


Figure 50 - Detailed Description of 13th Contextual Scenario

In this contextual scenario, the human perspective highlighted the ability to curate relevant content to an individual but at the same time criticised the creation of echo chambers and the exposure to diverse sources of information. From an intelligent machine perspective, the participants demonstrated similar attributes. Personalised newsfeed algorithms were regarded as smart filters that tailor content to individual preferences, but the participants mentioned the implications for accessing diverse information and potential biases.

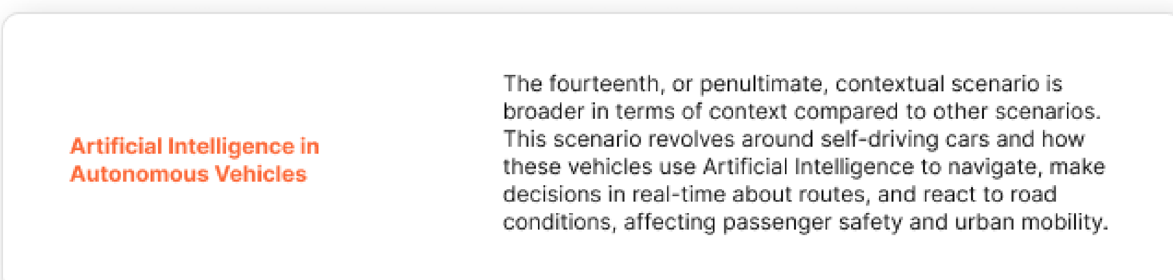


Figure 51 - Detailed Description of 14th Contextual Scenario

From a human lens, participants were primarily concerned about the safety, reliability, and ethical decision-making in critical situations. On a more positive note, there was some acknowledgment of the potential of autonomous vehicles and revolutionising transportation. From an intelligent machine perspective, the participants' tone was more positive, and they viewed autonomous vehicles as a transformative technology with the potential to drastically reduce accidents, optimise traffic, and reshape urban landscapes.

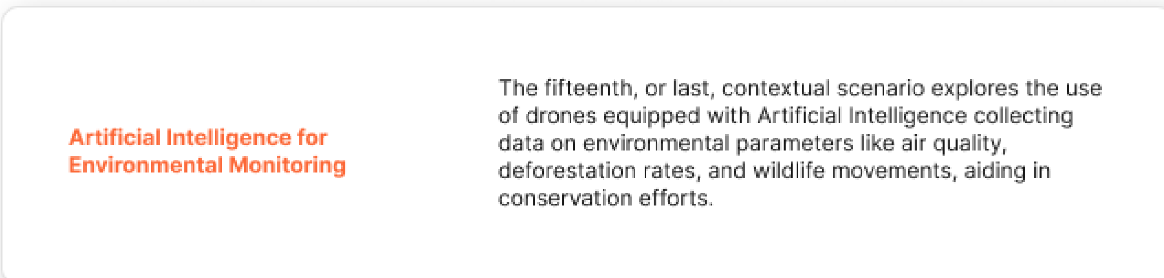


Figure 52 - Detailed Description of 15th Contextual Scenario

The last contextual scenario was highly valued from a human lens. Participants saw the potential for conservation and environmental protection. The only less positive aspect related to the constant surveillance introduced in environmental monitoring. From an intelligent machine lens, artificial intelligence for environmental monitoring was seen as an important tool to collect and analyse environmental data, and consequently help in policy-making and conservation strategy efforts.

Analysing Data

The inductive thematic analysis enabled the categorisation and interpretation of complex and rich qualitative data. Each response was coded for sentiments, perceived relationship types (parasitic, mutualistic, among others), and key defining attributes such as dependency, control, and benefit. The frequency of these codes across different scenarios provided a quantitative layer to the qualitative insights, facilitating the understanding of prevalent trends and outlier perceptions.

There were some common patterns detected, quantified and analysed. For instance, the researcher identified patterns like the prevalence of privacy concerns, dependency issues, and mistrust in certain applications, particularly facial recognition and predictive policing. Another pattern located on the opposite side of the spectrum was the more optimistic view in health and environmental applications. This variation detected in the patterns suggests a nuanced public perception, heavily influenced by the context of the application and visible impact of the Artificial Intelligence system.

Reflection

Each scenario illustrated the complex, often contradictory nature of human-AI interaction. While AI is generally seen as a tool for efficiency, personalisation, and problem-solving, human perspectives often dwell on ethical, privacy, and dependency concerns. The findings underscore the need for transparent, ethical AI development, and comprehensive stakeholder education to foster more informed, balanced perceptions of AI.

The analysis of the contextual scenarios reveals a pronounced divergence between human perceptions and the functional intentions attributed to Artificial Intelligence systems. Humans often approach Artificial Intelligence with skepticism and their concerns are deeply rooted in ethical, privacy, and dependency issues. In contrast, intelligent machines are primarily framed in terms of efficiency, personalisation, and their role in advancing technology symbiotically. This dichotomy not only illustrates the gap between Artificial Intelligence's potential capabilities and the public's perception but also points to a broader misunderstanding or mistrust of AI's roles and intentions.

This study unveils the intricate tapestry of human-AI relationships across various contextual applications, marked by a juxtaposition of hope, skepticism, and the need for a balanced approach in Artificial Intelligence development and policy-making. Future research endeavours should expand on using the proposed symbiotic framework on other contextual scenarios and applications. Another alternative for future research efforts is to delve into the causal factors behind these perceptions and explore mechanisms to bridge the understanding gap, and consequently ensuring a more harmonious human-machine symbiosis.

This project meticulously maps the multifaceted landscape of human-AI relationships, underscoring a spectrum that ranges from optimism to apprehension. The project highlights the criticality of balancing technological advancement with ethical considerations, privacy, and public trust. Additionally, addressing these dichotomies and understanding their root causes is crucial for nurturing a more integrated and positive future for AI in society.

The study also highlighted certain implications for Artificial Intelligence Development and Policy. The first implication pertains to Ethical Artificial Intelligence Design. The insights from this study should guide designers, developers and researchers towards ethical, transparent, and user-centric Artificial Intelligence designs. Understanding human perceptions is fundamental throughout the design process (Grgić-Hlača et al., n.d.). This allows for the development of Artificial Intelligence systems that are not only technically proficient but also ethically aligned, publicly accepted, and ultimately beneficial to society (Tjondronegoro et al., 2022). For example, by understanding how users perceive symbiotic relationships with Artificial Intelligence, designers can create systems that foster a sense of collaboration and mutual benefit, fostering trust and public acceptance (Auernhammer, 2020).

Another implication relates to policy making. The project findings can significantly contribute to informed policy-making, particularly in areas concerning privacy, data protection, and equitable use of AI. As exemplified in this project, policies need to be dynamic and reflect the evolving understanding of Artificial Intelligence and its societal impacts.

Public Awareness and Education is another implication highlighted in this research project. This project highlights the crucial role of enhancing public understanding and trust in Artificial Intelligence. Educational initiatives, transparent communication, and participatory development processes can bridge the gap between Artificial Intelligence capabilities and public perceptions. Building an informed user base can mitigate unfounded fears and pave the way for more symbiotic Artificial Intelligence integration in daily life.

Overall, the nuances captured in this research project provide a foundational understanding crucial for the evolving landscape of Artificial Intelligence and its symbiotic relationship with humans. Future research should continue to explore the underlying causes of varied perceptions and attitudes towards Artificial Intelligence and further refine the role of symbiotic relationships in Human-AI interactions. Investigating these causal relationships can offer deeper insights into how the public trusts and perceives Artificial Intelligence.

Analysis

In the contemporary digital landscape, the interaction between humans and Artificial Intelligence is increasingly pivotal. Clearly illustrated in this chapter, the inherent relationships formed through these interactions vary considerably, influenced by the nature and application of the intelligent machine. This study embarked on an exploratory journey to unravel some complexities of these symbiotic relationships, examining how different Artificial Intelligence applications are perceived and the underlying values or attributes that shape these perceptions.

Perceptual divergences in Human-AI relationships in a privacy context

As previously mentioned, the researcher identified a number of patterns relating to Human-AI symbiosis and deriving from inductive thematic analysis. The most evident pattern relates to the perceptual divergences in Human-AI relationships in a privacy context. The research project illuminates a stark divergence in the perceptions of relationships formed through various Artificial Intelligence applications, particularly when juxtaposed against the backdrop of privacy concerns. In instances where intelligent machines are seen as intrusive, notably those involving intensive data surveillance or analysis of personal data without explicit consent, they are frequently perceived by participants as either parasitic or amensalistic. Such perceptions are emblematic of a relationship in which the human participants are either detrimentally impacted or do not discern any tangible benefit. In stark contrast, the perspective of the intelligent machine, inferred from its design objectives and operational rationale, often categorises these relationships as either mutualistic or commensalistic, suggesting a perceived bilateral benefit or a unidirectional advantage to one party (the intelligent machine) without inflicting harm. Additionally, intelligent machines designed with an intrinsic focus on protecting or augmenting user privacy tend to be viewed as mutualistic by both the human perspective and the machine perspective, pointing towards a symbiotic and harmonious relationship. This dichotomy accentuates a profound discord in risk and value perceptions between human users and Artificial Intelligence systems.

In the contextual scenarios identified by privacy-invasive AI applications, such as scenarios entailing exhaustive data surveillance or advanced personal data processing algorithms operating without explicit user consent, the human lens largely views the relationship as either parasitic or amensalistic. This perception signals a one-sided relationship wherein humans feel disadvantaged or devoid of benefit, thereby raising critical ethical and moral issues concerning the equitable apportionment of risks and benefits in Artificial Intelligence deployment.

The depiction of the relationships as parasitic highlights a situation where the intelligent machine, analogous to a biological parasite, reaps benefits at the human's counterpart expense. Conversely, amensalism conveys a circumstance in which the intelligent machine, while not necessarily deriving benefit, detrimentally impacts humans, akin to a large tree that diminishes sunlight to smaller plants. These perceptions emanate from concerns about privacy infringements, a lack of control over personal data, and the Artificial Intelligence's opaque functioning, engendering user feelings of vulnerability and exploitation.

From a functional standpoint of the intelligent machine, these relationships are often construed as mutualistic or commensalistic, mirroring the systems' objectives designed to maximise efficiency, enhance user experiences, or generate more profound data-driven insights. In these contextual scenarios, the intelligent machines operate under the supposition of a benign or even advantageous interaction, often neglecting or unaware of the ethical and personal repercussions on its human counterpart.

Mutual Benefit in Privacy-Enhancing Applications

In contrast, AI applications specifically crafted to bolster user privacy or enable users to manage their personal data are perceived as mutually advantageous by both humans and AI. This mutualism illustrates a balanced, harmonious interplay where both parties are viewed as beneficiaries. Instances of such applications include AI tools facilitating user control over data sharing preferences and AI systems improving data security.

In these cases, both human users and AI systems acknowledge and prize the significance of data privacy and security. For humans, this mutualistic bond cultivates trust in the technology, nurturing a sense of empowerment and governance over their digital footprint. For machine intelligence, functioning within these privacy-preserving bounds bolsters its operational effectiveness and societal endorsement, assuring congruence with human ethical norms and values.

The big discrepancy in perceptions reveals a critical misalignment in the evaluation of risks and values between humans and intelligent machines, especially regarding privacy matters. It underscores an imperative for reevaluating the design, implementation, and regulation of Artificial Intelligence systems. The human viewpoint, focusing on ethical, emotional, and social ramifications of Artificial Intelligence, often stands in stark opposition to the intelligent machine perspective, dominated by efficiency, data exploitation, and task optimization. To bridge this chasm, a multidisciplinary strategy like the Human-AI symbiotic framework is essential, incorporating not only technological advancements in Artificial Intelligence but also a robust ethical framework, transparent practices, and inclusive policies that champion human rights and values.

Furthermore, this research project reflects a broader socio-technological issue - the challenge of balancing the exploitation of intelligent machine capabilities for societal benefit against the safeguarding of individual rights and freedoms. Achieving such equilibrium can be a possibility by implementing the Human-AI symbiotic framework and it may prove to be pivotal to ensuring that machine intelligence's evolution proceeds in harmony with human welfare, establishing a synergy where technology augments humanity without eroding its intrinsic rights and dignities.

Manifestation of Attributes

In contextual scenarios where intelligent machines are perceived as invasive, individuals prominently emphasise the need for fairness and accountability. These concerns arise from a broader societal and personal directive towards equitable treatment by technological systems. Fairness is interpreted as encompassing the neutrality of algorithmic decision-making, balanced data usage, and the prevention of discriminatory outputs. The increasing awareness of intelligent machines' potential to magnify societal and personal inequalities drives a demand for these systems to be inherently fair in their design and operation.

Accountability, an extension of fairness, involves responsibility for correcting errors or biases and transparent decision-making processes. This accent on accountability reflects a collective aspiration for regulatory mechanisms and supervisory frameworks such as the one introduced in the previous chapter, ensuring Artificial Intelligence systems abide by ethical norms and societal standards.

The research project also highlights that machine intelligence is chiefly designed to value systematicity and interdependence. Systematicity reflects the intelligent machine's adherence to a rule-based, structured operational pattern, prioritising order, precision, and consistent data processing. On another note, interdependence denotes the intelligent machine's dependence on a complex network of data, algorithms, and user interactions. This trait underscores Artificial Intelligence's effectiveness being inherently linked to a larger ecosystem, encompassing technology, data, and user engagement, pivotal for the intelligent machines learning and evolutionary process.

Lastly, in the context of privacy-enhancing contextual scenarios, there is a marked alignment in the attributes valued by humans and intelligent machines. In those scenarios, both parties focus on structured, efficient, and mutually beneficial frameworks. This confluence suggests that intelligent machines perceived as protective of individual autonomy and privacy elicit a shared value system between humans and intelligent machines, offering avenues for collaborative advancement.

This alignment offers a blueprint for synchronising human-AI relationships and that synchronisation can be accomplished with the symbiotic Human-AI framework. Aligning human-centred values such as fairness and accountability with the intelligent machine inherent systematicity and interdependence opens opportunities for crafting systems that are not only technologically refined but also ethically robust and socially aligned. This synergy fosters trust, acceptance, and appreciation in Artificial Intelligence applications, promoting a unified and cooperative future in human-AI interactions.

Additionally, the insights gained from analysing these attributes provide essential guidance for Artificial Intelligence developers, policymakers, and ethicists. The insights highlight the criticality of incorporating human-oriented values in the development of intelligent machines. Bridging the disparity between human expectations and Artificial Intelligence's operational logic requires a concerted endeavour to create systems that excel in efficiency and systematic operation while resonating with core human principles like fairness, transparency, and accountability.

The development of Artificial Intelligence technologies that respect and mirror human priorities, while leveraging their intrinsic strengths of systematicity and interdependence, is essential for moving towards a more just, trusted, and efficacious human-AI paradigm.

Shifting Symbiotic Relationships

As exemplified in this research project, the fluid nature of human-AI relationships offers avenues to recalibrate prevailing perceptions, particularly regarding applications that potentially infringe upon privacy. Substantial enhancements in accountability and transparency in Artificial Intelligence applications' operations are imperative to transition these relationships from parasitic and exploitative to more mutualistic, cooperative forms. Augmenting data protection and accentuating the primacy of user consent can diminish perceptions of unilateral benefit or detriment. A key to this evolutionary process is the nurturing of a reciprocal understanding. Instructing the human side about the functionalities, constraints, and protective measures of Artificial Intelligence, whilst embedding human-centric and ethical paradigms within Artificial Intelligence development, is crucial in bridging perceptual disparities. This foundational understanding is pivotal in metamorphosing these interactions into more sophisticated, equitably balanced engagements wherein the advantages and risks associated with Artificial Intelligence are distributed more justly and managed with greater transparency.

Theoretical Refinement of Relationship Dynamics

Human-AI interactions, by their inherently dynamic nature, necessitate an analytical probe into their potential evolution, especially in instances where intelligent machines are currently met with skepticism or apprehension. At the core of this transformation is the reframing of these interactions, shifting from paradigms that are either parasitic or amensalistic (where typically humans are disadvantaged or receive no discernible benefit) to models that are mutualistic or even commensalistic, marked by either mutual advantage or, at minimum, no harm to any party.

Methodologies for Contextual Symbiotic Evolution

Accountability and Transparency Enhancement

Among the methodologies to shift the contextual symbiotic evolution between humans and intelligent machines is the amplification of accountability and transparency within Artificial Intelligence constructs. This strategy encompasses the elucidation of Artificial Intelligence system operations, decision-making processes, and data usage in a manner that is clear and comprehensible. Establishing robust symbiotic frameworks focused on accountability, where intelligence machines' decisions are subject to review, critique, and potential reversal, can significantly alter the perception of Artificial Intelligence as an enigmatic and unregulated phenomenon.

Design plays a crucial role in enhancing both accountability and transparency (Schoenherr et al., 2023). In User-Centred Interface Design space, employing user-centred design principles allows developers to create interfaces that effectively communicate Artificial Intelligence system functionalities, decision-making processes, and data utilisation. This can involve utilising clear and concise language, leveraging visualisations, and offering interactive explanations (Arrieta et al., 2019; Gunning et al., 2019). Additionally, Designers can integrate Explainable AI (XAI) techniques into the development process. Explainable Artificial Intelligence tools enable the creation of explanations for Artificial Intelligence decisions in a way that is understandable by humans. These explanations can be tailored to different audiences, with varying levels of technical expertise (Lipton, 2018). Furthermore, Interactive visualisations can be a powerful design tool for fostering transparency. Interactive visualisations can allow users to explore data used by Artificial Intelligence systems, trace decision-making processes, and understand potential biases (V. Liao & Varshney, 2021).

Transparency in a symbiotic relationship transcends the mere provision of information, it entails the delivery of data in formats that are intelligible and pertinent to end users, both humans and intelligent machines. Therefore, it is vital to translate algorithms and data processing techniques into more user-centric language as an effort to dispel enigmas and myths surrounding Artificial Intelligence operations. Designers play a critical role in this translation process and there are certain key strategies designers can employ. For example, by following user-centred design principles, designers can craft interfaces that effectively communicate Artificial Intelligence functionalities in clear and concise language. This might involve utilising simple terminology, leveraging metaphors, and employing visuals to illustrate processes. Storytelling and narrative design techniques can also be used to frame complex information in a relatable and engaging way. By weaving narratives that showcase the "why" behind Artificial Intelligence decisions, designers can enhance user understanding and trust. Lastly, interactive visualisations allow users to explore data and Artificial Intelligence processes at their own pace. The visualisations can be designed to highlight key decision points, potential biases, and the overall reasoning behind Artificial Intelligence outputs.

By utilising some of these strategies, designers can bridge the gap between complex Artificial Intelligence systems and human users, and consequently foster a more transparent and trustworthy human and Artificial Intelligence relationship.

Mutual Understanding Promotion

The research project demonstrated that cultivating a mutual comprehension between intelligent machines and their human operators is fundamental in reshaping these relationships. This involves educating the public about Artificial Intelligence s mechanisms, ethical applications, and safety measures in place to safeguard human interests. Educational endeavours should strive to dismantle prevalent myths and misconceptions about intelligent machines and mitigate apprehension and distrust.

Additionally, the integration of human-centric design and ethical considerations in AI development is paramount. Artificial Intelligence systems should be designed not merely for task efficiency but with a conscientious regard for human values, cultural sensitivities, societal norms and overall surrounding context. Human-centric Artificial Intelligence efforts prioritise systems that enhance and support human abilities and experiences, rather than replacing or diminishing them, and it’s in this space where Human-AI symbiotic frameworks can create a significant impact.

Significance of Relational Evolution

Lastly, the shift towards more constructive and positive human-AI interactions harbours far-reaching consequences. More positive Human-AI interactions are conducive to fostering broader acceptance and integration of Artificial Intelligence across diverse life contexts. Furthermore, it sets the stage for harnessing Artificial Intelligence as a means for societal enhancement, improving well-being, safety, and efficiency. This contextual symbiotic evolution also impels and fosters the creation of intelligent machines that are not only technically advanced but also ethically oriented and globally embraced. Furthermore, it highlights the importance of formulating governance models and policy guidelines that ensure the just and fair application of Artificial Intelligence.

Transforming the nature of symbiotic relationships extends beyond the domain of Artificial Intelligence itself. Transforming the nature of such relationships requires an integrated approach encompassing technological innovation, ethically grounded Artificial Intelligence development, strategic policymaking, and public pedagogy. By employing all these concerted efforts, Artificial Intelligence can evolve into a technology that operates in synergy and adapts with humanity, paving the way for a future where humans and intelligent machines coexist in better symbiotic ecosystems.

Conclusion and Future Research

This research project expands on the previous chapter and elucidates the dynamic and multifaceted nature of human-AI relationships, moulded by the contextual deployment of Artificial Intelligence technologies. The research project not only provides a comprehensive overview of the intricate and evolving dynamics of human-AI interactions, but it also underscores that these interactions are fluid, heavily influenced by the specific applications of Artificial Intelligence and the encompassing social, ethical, and personal milieus. In light of rapid technological progression and evolving societal values, particularly concerning privacy, autonomy, and interconnectivity, it is imperative for this research domain to maintain both adaptability and responsiveness.

Future research should concentrate on experimental interventions aimed at digging deeper on symbiotic relationships, while investigating how alterations in Artificial Intelligence design, policy, and the augmentation of public awareness tangibly affect the symbiotic dynamics between humans and intelligent machines. Moreover, the implementation of additional longitudinal studies are poised to offer insights into the evolution of these relationships vis-à-vis technological advancements and the transformation of societal norms. Through persistent examination and engagement with these relational dynamics, we can channel the development of Artificial Intelligence towards trajectories that are more ethically responsible, aligned with user needs, yield broader social benefits and improve the overall relationships between humans and intelligent machines.

Divergences, Prioritisations, and Implications for Privacy-Enhancing Technologies

As Artificial Intelligence continues to evolve, its intersection with human ethical and societal values, especially concerning privacy, becomes increasingly more complex. This research study probes into the fundamental variances in how humans and intelligent machines perceive and prioritise attributes, such as privacy, utility, and efficiency, and investigates methodologies for better understanding the symbiotic relationship of both elements.

Disparities in Perception Between Human and AI Perspectives on Privacy

The research demonstrates that there are disparate perceptions between humans and intelligent machines, particularly highlighted in the context of privacy protection and implementation. Humans approach privacy with an emphasis on intangible ethical and emotional values, which sharply contrasts with the intelligent machines' predominantly utilitarian, efficiency-oriented reasoning. The depth of this discrepancy is considerable and multifaceted. Humans, influenced by a vast array of cultural, ethical, and emotional backgrounds, typically perceive privacy not just as a policy or technical mechanism, but as an integral human right intricately linked to their dignity, autonomy, and personal liberty. This human perspective is characterised by subjective variances, reflecting differences among individuals, communities, and cultures. In contrast, Artificial Intelligence systems, governed by algorithmic logic and data-driven decision-making, primarily address privacy through a utilitarian perspective. To intelligent machines, privacy is frequently construed as a set of programmable rules or a component within a privacy framework, primarily aimed at optimising data processing and utilisation while maintaining legal compliance. Artificial Intelligence systems' current approach lacks a comprehensive understanding of human emotional intricacies and socio-cultural dimensions. Such a different in perceptions necessitates a shift in Artificial Intelligence development strategies, and the symbiotic framework introduced in the previous chapter is an approach to shift those strategies. The framework exemplifies that strategies can be conceived that advocate for algorithms to become more empathetically tuned and context-aware, capable of recognising and valuing human perspectives on privacy and other qualitative dimensions.

Contrasting Prioritisation of Attributes in AI and Human Users

This study also explores the differing hierarchies in valuing attributes like privacy, accuracy, and functionality between humans and intelligent machines. The research study highlights the tendency among human users to favour privacy over efficiency, contrasting with Artificial intelligence systems which, being designed for optimal performance, may not prioritise privacy concerns.

By taking a deeper look, the study indicates that the valuation of attributes like privacy, accuracy, and functionality unveils a challenging terrain of contrasting values and expectations. Humans commonly value privacy as a protective barrier against potential misuse and a mechanism to exert control over their personal domains and information.

On the other hand, Artificial Intelligence systems, guided to provide peak performance, often place greater emphasis on attributes like accuracy and functionality.

For instance, in recommendation algorithms or targeted advertising, the priority is frequently on enhancing the precision and pertinence of the output, potentially at the expense of user privacy. This prioritisation by intelligent machines, driven by the goals of algorithmic efficiency and often commercial benefits, can inadvertently marginalise privacy considerations. The disparity in prioritising attributes demonstrates the necessity for a more equitable approach in AI development and the proposed symbiotic framework aims to be that approach. The end objective should be to forge Artificial Intelligence systems that harmoniously integrate technological prowess with fundamental human values like privacy and autonomy, ensuring that technology acts as an enhancer, rather than a suppressor, of human dignity.

Strategies for Embedding Privacy in AI Systems

The study also focuses on the integration of human-centric values, notably privacy, into the development of Artificial Intelligence systems. The research study promotes a participatory design methodology, aiming to ensure intelligent machines not only meet privacy requirements but also employ a wider range of ethical human standards.

In this point, the discourse marks a critical shift towards re-envisioning Artificial Intelligence development principles. At the forefront of this approach is the concept of participatory design, which involves incorporating a range of potential end-users in imagining the Artificial Intelligence relationship design process. The involvement from end users enables the creation of intelligent machines that authentically reflect diverse human values, experiences, and privacy needs.

Value of the Symbiotic Framework

Fresh, Participatory Approach

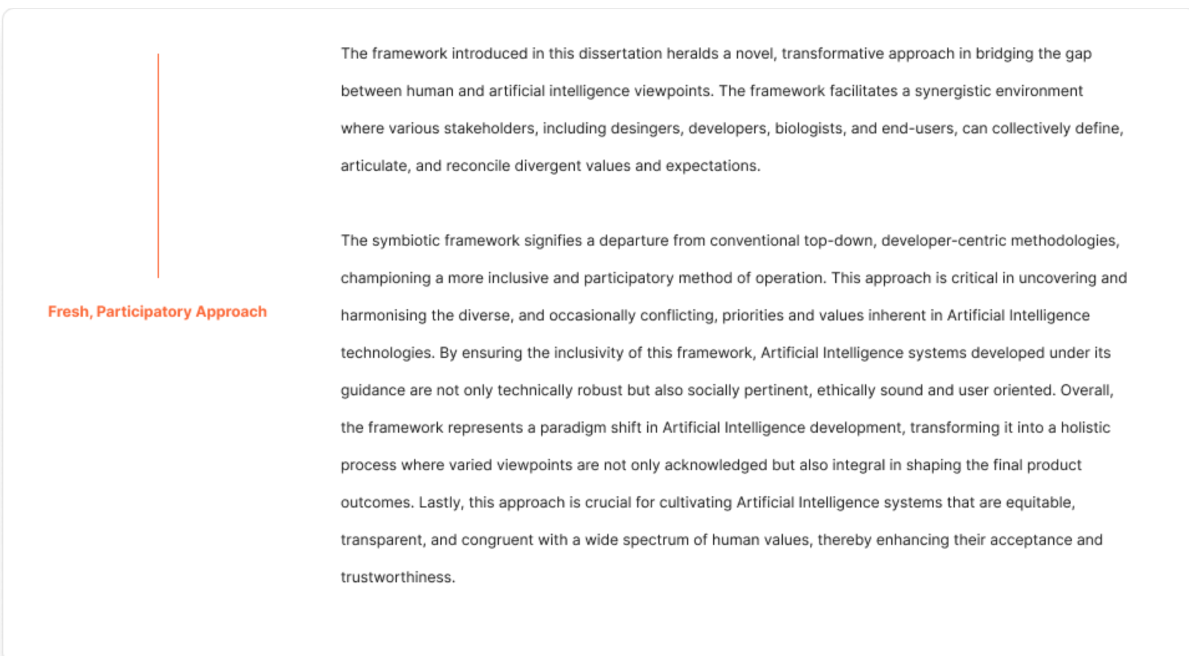


Figure 53 - Value of the Symbiotic Framework – Fresh, Participatory Approach

Reconciliation of Values

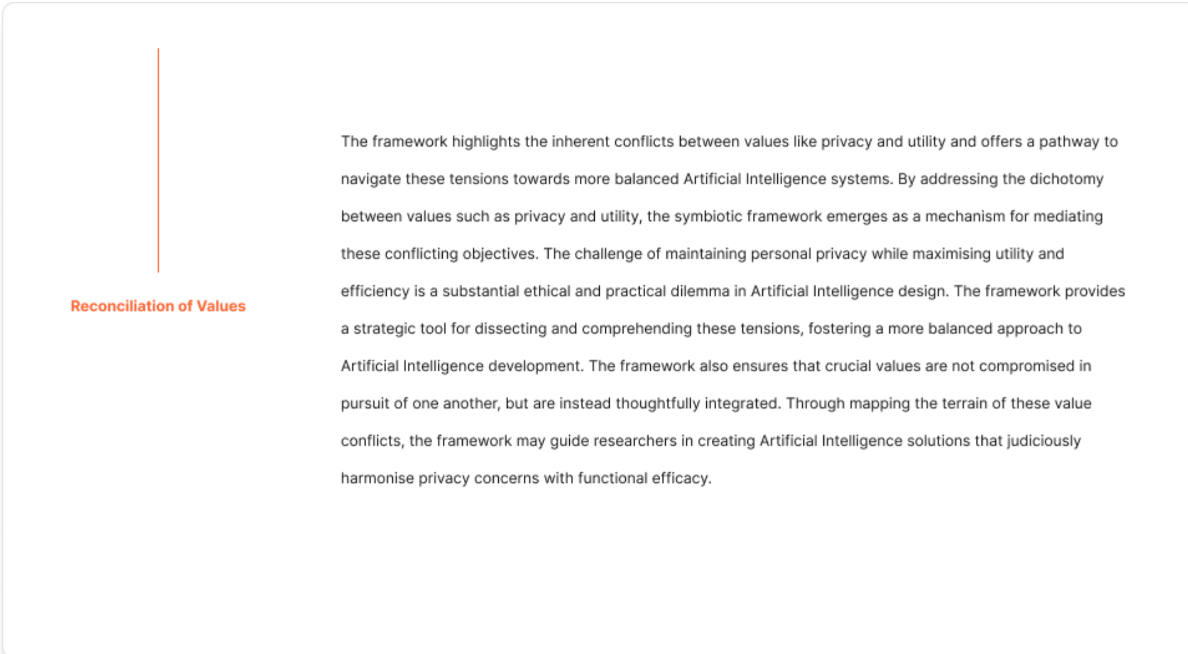


Figure 54 - Value of the Symbiotic Framework – Reconciliation of Values

Catalysation of New Design Philosophies

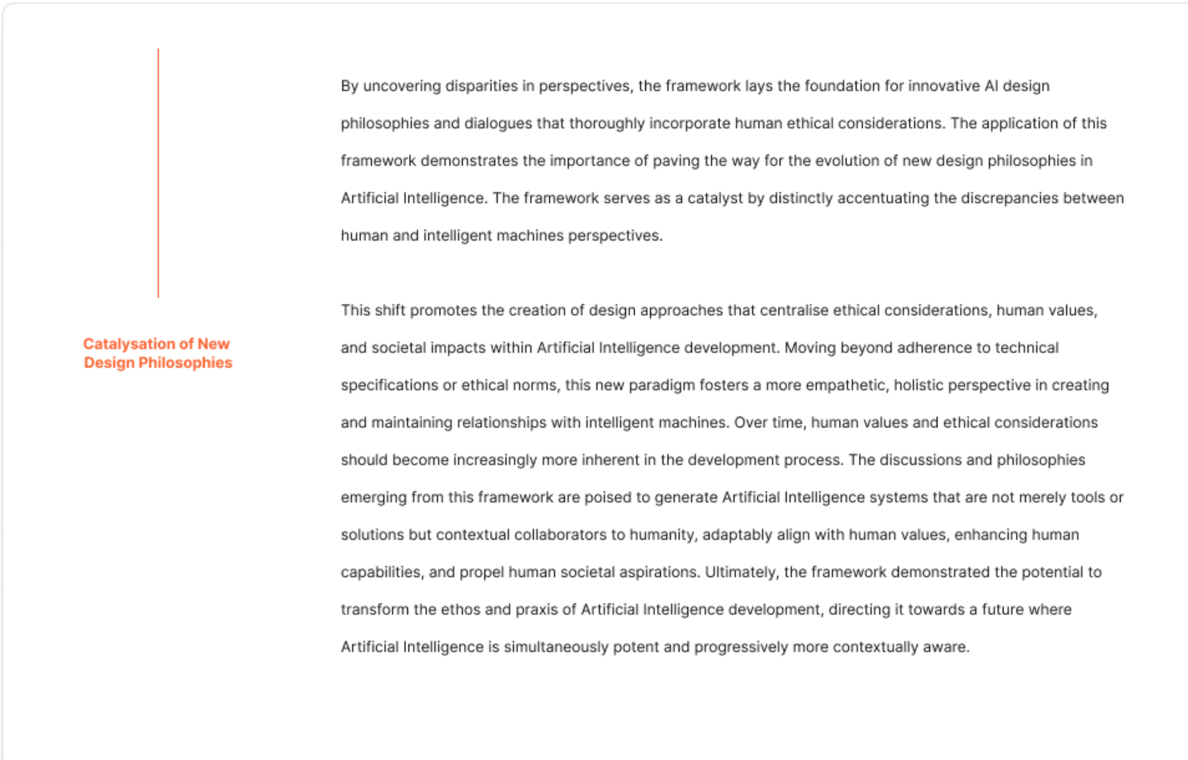


Figure 55 - Value of the Symbiotic Framework – Catalysation of New Design Philosophies

Research Study Limitations

Sample Size and Diversity

The study's limited sample size and lack of diversity among participants represent significant shortcomings, potentially compromising the generalisability of the findings. This constraint raises concerns regarding the reflection of a wider societal and cultural spectrum in the results. A sample with a larger degree in participant diversity, spanning variables such as cultural backgrounds, age brackets, levels of technological proficiency, and socio-economic statuses, might not effectively encapsulate the wide-ranging public opinion and varying user experiences with Artificial Intelligence. A limitation such as this is especially pertinent in the context of Artificial Intelligence, due to being a field characterised by its pervasive influence and global reach, where user experiences and perceptions substantially differ based on socio-cultural environments and personal interactions with technology. Additionally, the small sample size might not sufficiently uncover the subtle and possibly contradictory viewpoints prevalent among the broader population. The implications of this limitation are profound, affecting the empirical validity of the study's outcomes and their applicability across diverse geopolitical, societal, and cultural settings. An expansion and diversification of the participant sample are imperative in future research studies to ensure that the research accurately mirrors the complex and multifaceted nature of Artificial Intelligence's global impact.

Need for Deeper Attribute Analysis

The research can benefit from a more rigorous, comprehensive examination of Artificial Intelligence contextual scenarios, the intelligent machine attributes and their relative importance among varied participant groups to glean deeper insights. The study, while initiating a pivotal discussion on prioritising different Artificial Intelligence system attributes, must delve deeper. Ubiquitous attributes such as privacy, accuracy, transparency, user autonomy, and ethical ramifications require more meticulous and discerning investigation.

Future Research Directions

There are five proposed research directions for this research to delve deeper into this topic and obtain more accurate depiction of current Human-AI symbiotic relationships

Enhanced Sample Representation

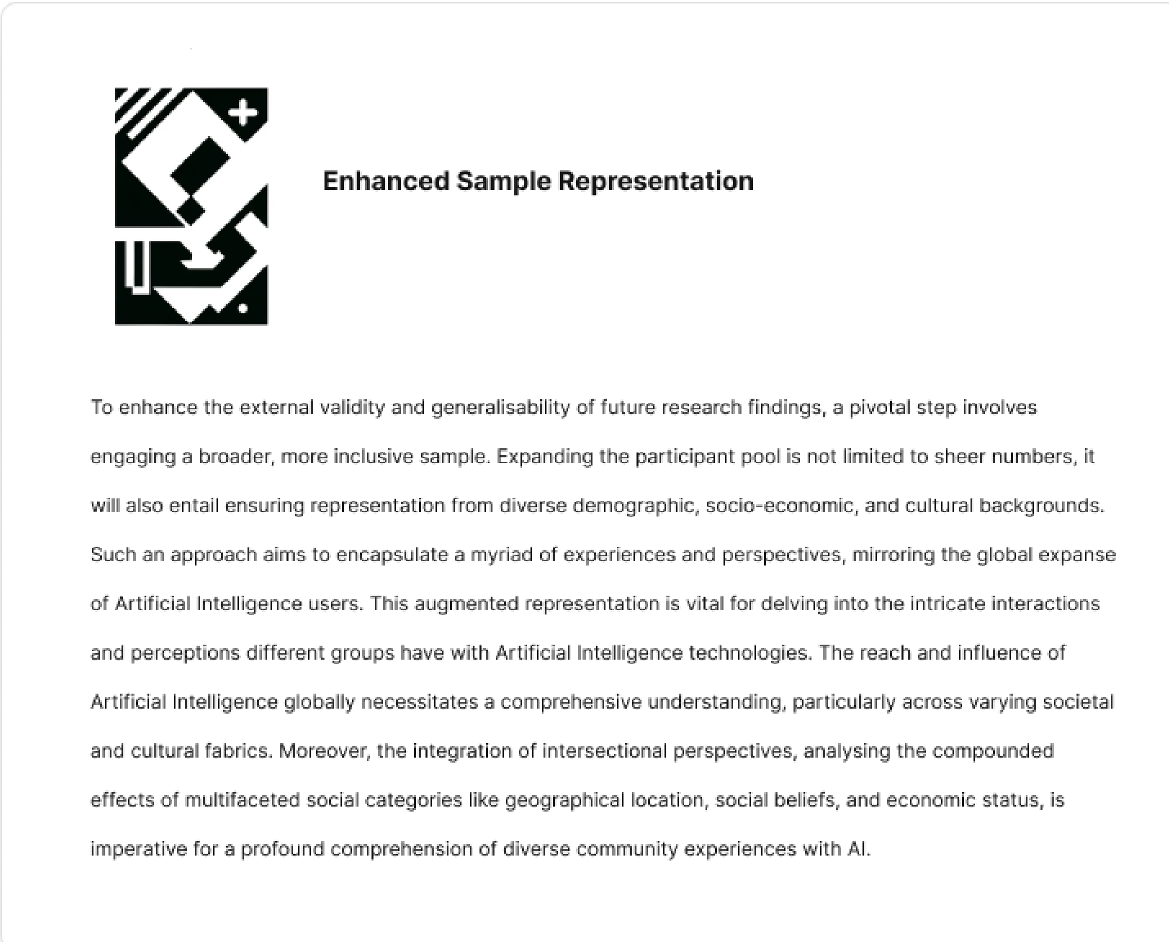
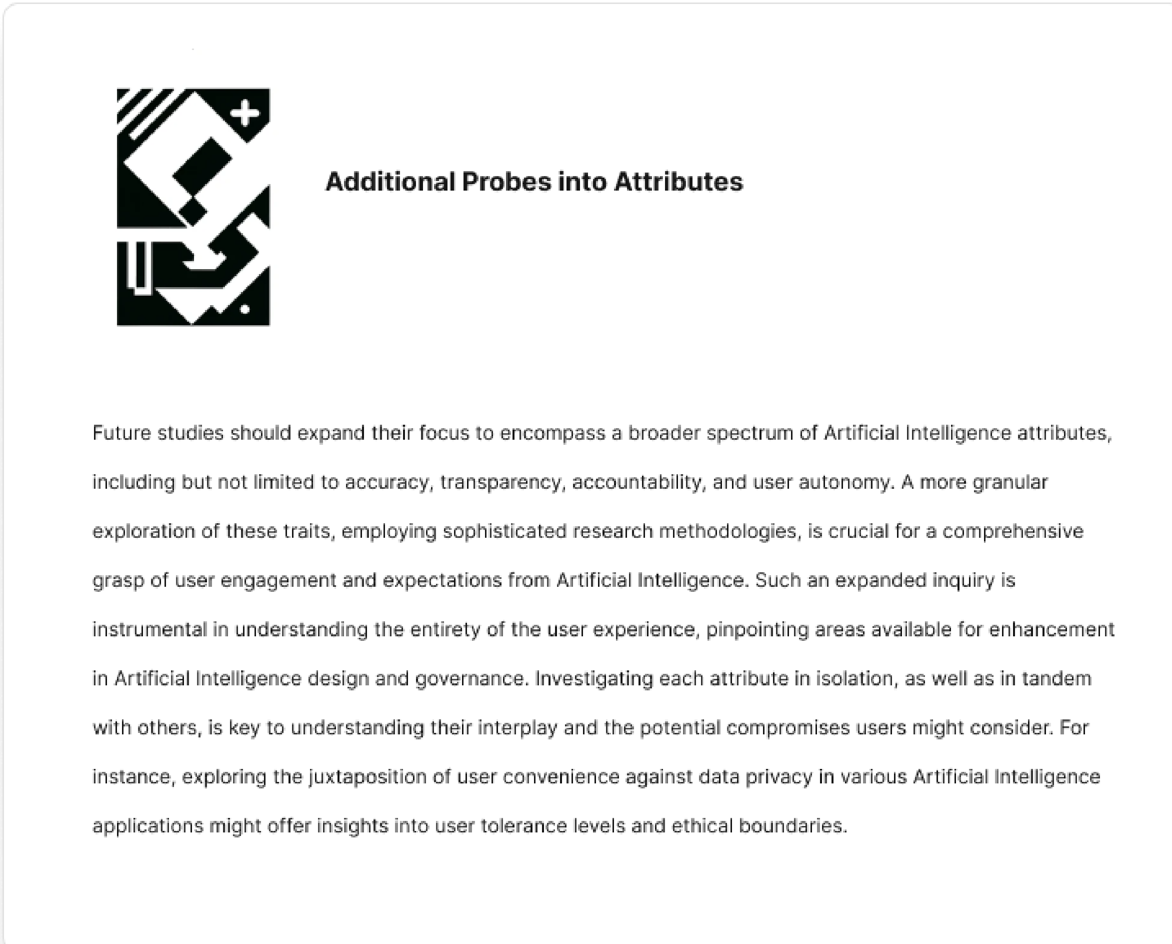


Figure 56 – Enhanced Sample Representation

Additional Probes into Attributes



Additional Probes into Attributes

Future studies should expand their focus to encompass a broader spectrum of Artificial Intelligence attributes, including but not limited to accuracy, transparency, accountability, and user autonomy. A more granular exploration of these traits, employing sophisticated research methodologies, is crucial for a comprehensive grasp of user engagement and expectations from Artificial Intelligence. Such an expanded inquiry is instrumental in understanding the entirety of the user experience, pinpointing areas available for enhancement in Artificial Intelligence design and governance. Investigating each attribute in isolation, as well as in tandem with others, is key to understanding their interplay and the potential compromises users might consider. For instance, exploring the juxtaposition of user convenience against data privacy in various Artificial Intelligence applications might offer insights into user tolerance levels and ethical boundaries.

Figure 57 - Additional Probes into Attributes

Comparative Experiments.

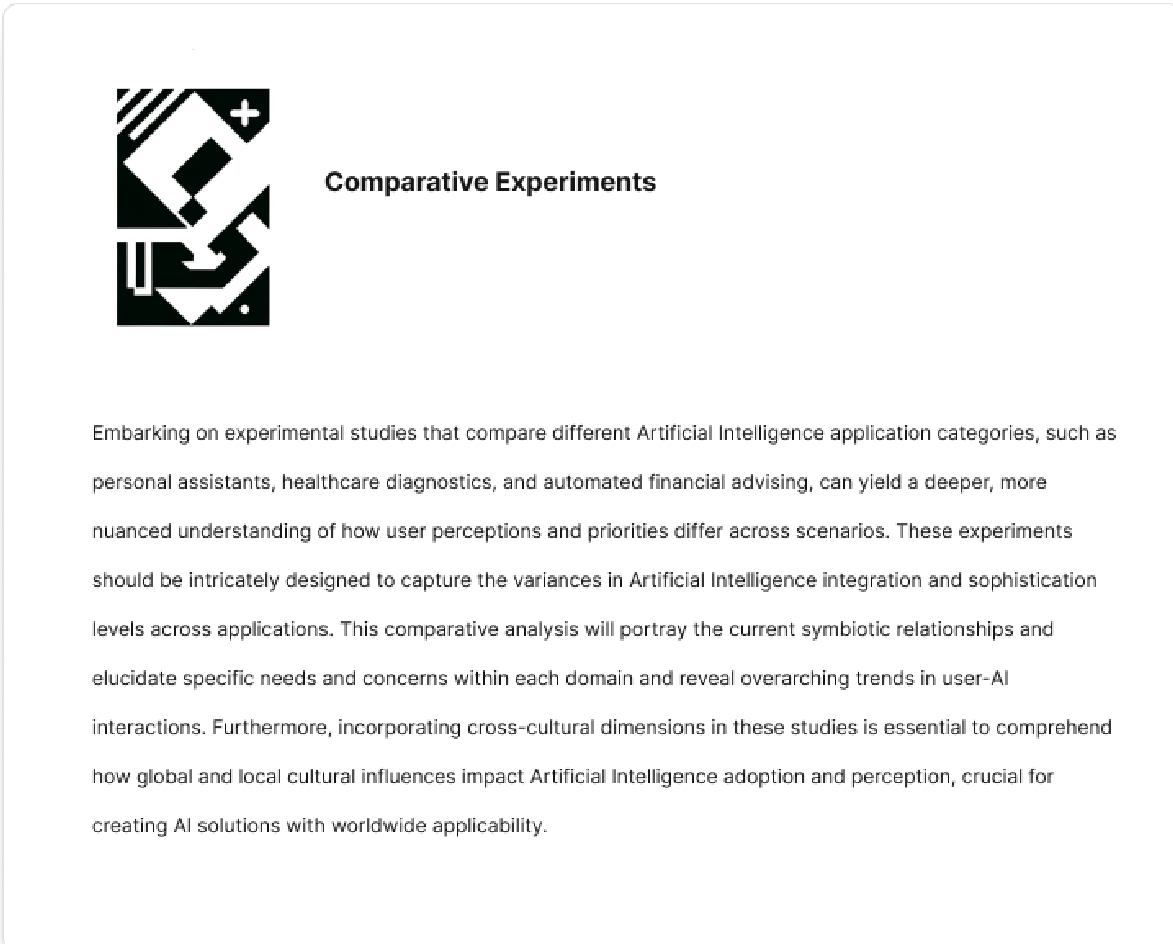
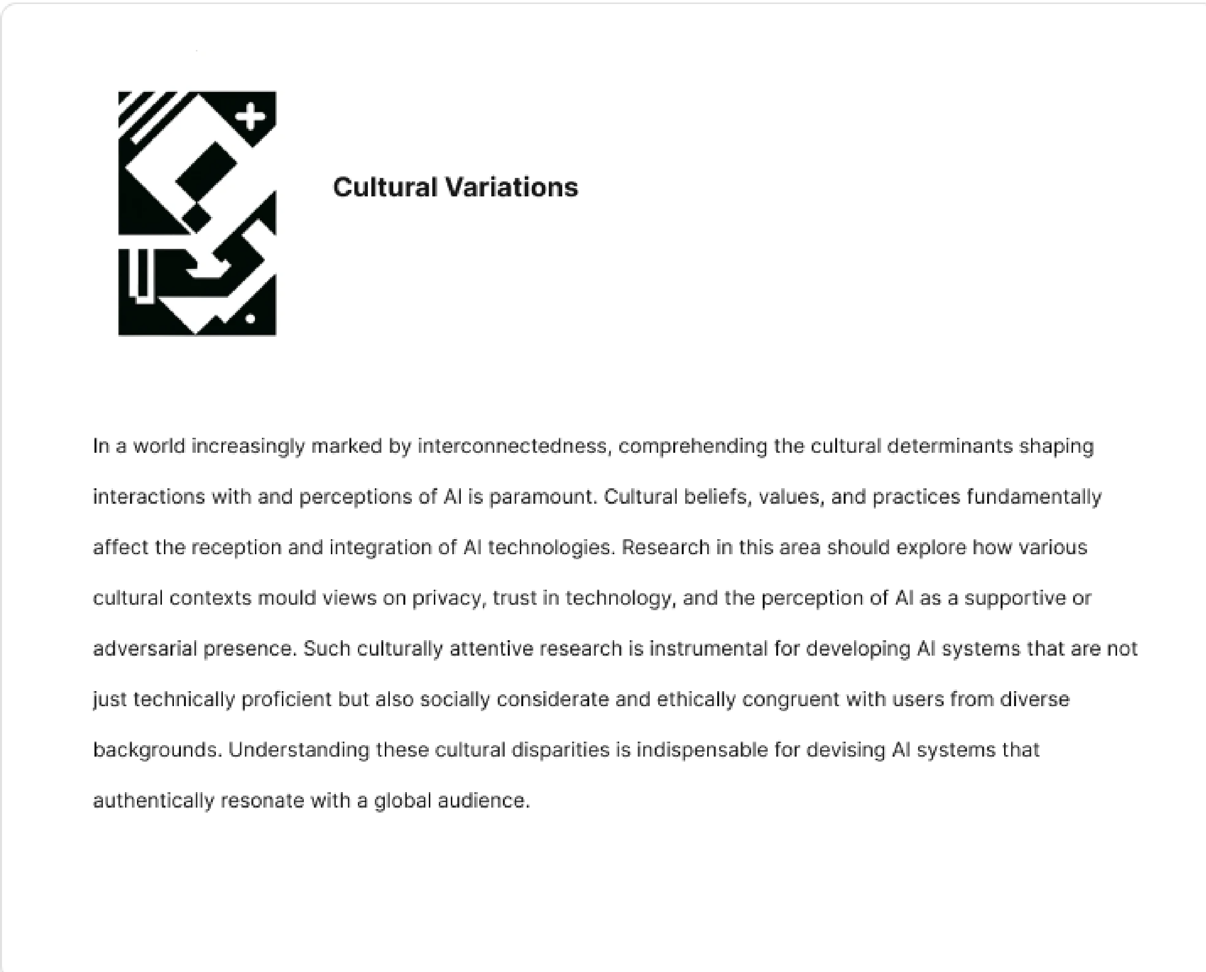


Figure 58 - Comparative Experiments

Cultural Variations



Cultural Variations

In a world increasingly marked by interconnectedness, comprehending the cultural determinants shaping interactions with and perceptions of AI is paramount. Cultural beliefs, values, and practices fundamentally affect the reception and integration of AI technologies. Research in this area should explore how various cultural contexts mould views on privacy, trust in technology, and the perception of AI as a supportive or adversarial presence. Such culturally attentive research is instrumental for developing AI systems that are not just technically proficient but also socially considerate and ethically congruent with users from diverse backgrounds. Understanding these cultural disparities is indispensable for devising AI systems that authentically resonate with a global audience.

Figure 59 - Cultural Variations

Expansion Across Artificial Intelligence Application Categories

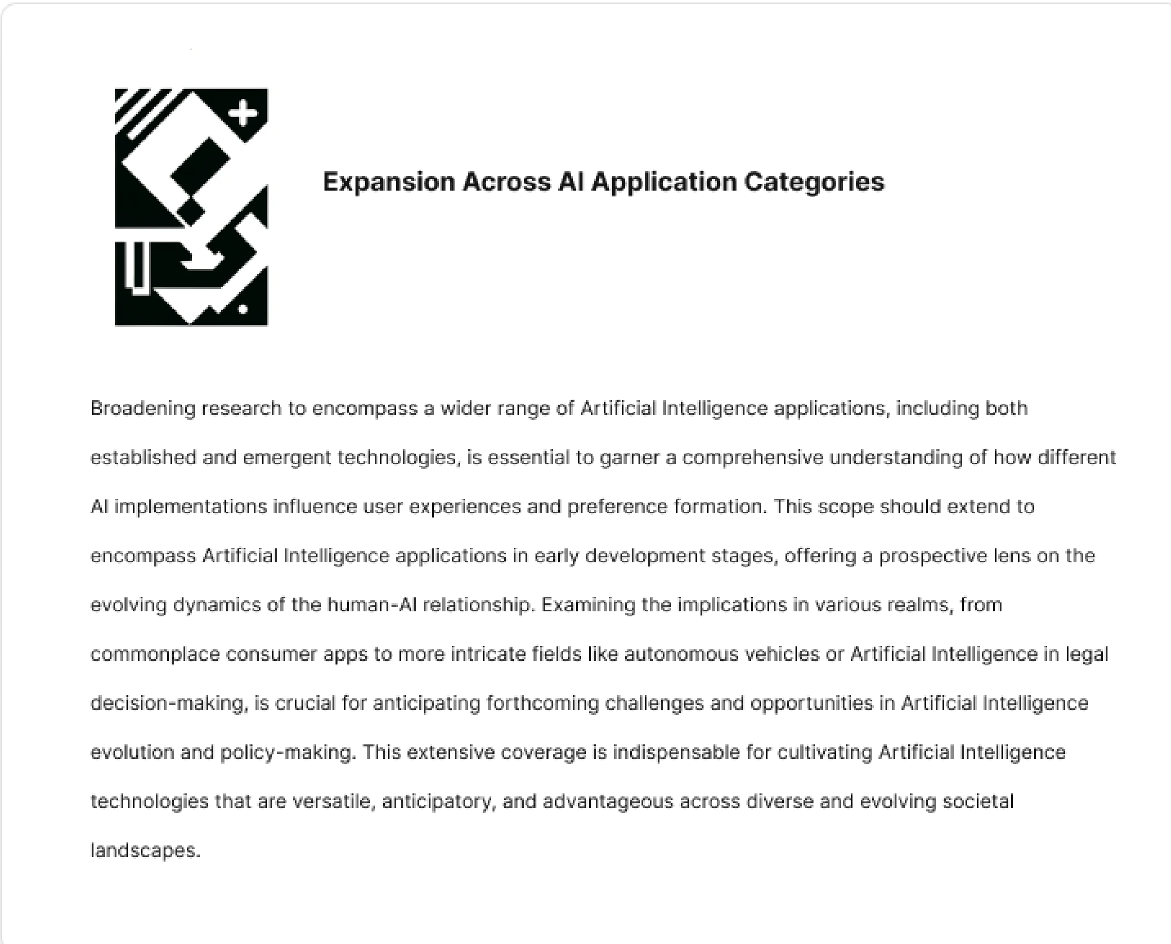


Figure 60 - Expansion Across AI Application Categories

Paving the way for a more empathetic, Human-oriented Artificial Intelligence development

This comprehensive investigation into the disparities and potential synergies between human perceptions and Artificial Intelligence strategies, especially regarding privacy within Artificial Intelligence systems, establishes a foundational platform for more empathetic, human-oriented Artificial Intelligence development. The study highlights the imperative to integrate a spectrum of human values into Artificial Intelligence design, accentuating the critical need for ongoing, diverse research in this domain. These insights aim to direct the evolution of Artificial Intelligence technologies along a path that is innovative and efficient, while simultaneously adhering to ethical standards and promoting social responsibility.

The culmination of this probing analysis into the nuances of Human and Artificial Intelligence perspectives, with a specific emphasis on privacy issues in Artificial Intelligence systems, is both enlightening and fundamental. This research acts as a catalyst for a significant shift in Artificial Intelligence development towards a model deeply rooted in human experiences and values. The core observations and discussions presented in the study offer a crucial perspective for reevaluating and reshaping the underlying principles of Artificial Intelligence progression.

Empathetic and Human-Centric Artificial Intelligence

The research uncovers the frequently neglected discord and tensions between human anticipations and Artificial Intelligence’s functional rationality. These discrepancies, especially notable in terms of privacy, necessitate a more compassionate approach to Artificial Intelligence design, surpassing mere technical and efficiency-driven considerations. This requires a richer understanding of the emotional, ethical, and sociocultural facets of human interactions with Artificial Intelligence. Consequently, human-centric Artificial Intelligence is not merely preferable but essential for developing technology that is congruent with, and respectful of, human needs and principles.

Integration of Diverse Human Values in Artificial Intelligence

The study profoundly emphasises the essentiality of embedding a broad spectrum of human values within Artificial Intelligence systems. Transcending traditional emphases on efficiency and functionality, Artificial Intelligence must be instilled with values such as privacy, fairness, transparency, and accountability. Incorporating these values not only addresses ethical obligations but also augments the practical effectiveness and societal endorsement of Artificial Intelligence technologies. The research indicates that the prioritisation of these values is contingent upon the context, highlighting the heterogeneity of human viewpoints. Therefore, Artificial Intelligence design must dynamically adapt to these diverse and evolving human value landscapes.

Urgency of Multifaceted Research

The findings articulate an urgent need for comprehensive and diverse research in the Artificial Intelligence and Human AI Symbiosis space. Future studies should enhance not only in magnitude and variety but also in profundity and intricacy, delving into under-explored areas such as cultural differences, sector-specific needs, and the extended societal ramifications of AI. With the continuous evolution and integration of Artificial Intelligence into society, the complexities surrounding Artificial Intelligence are increasingly significant. Thus, persistent and thorough academic exploration is essential to adeptly address these complexities and ensure that Artificial Intelligence development is anchored in a holistic understanding of its wider societal consequences.

Shaping the Trajectory of Artificial Intelligence Development

The insights derived from this study hold substantial potential to influence and shape the future direction of AI technology. By prioritising the alignment of Artificial Intelligence with symbiotic principles and values, we can guide Artificial Intelligence development towards not only technologically superior and innovative solutions but also towards ethical integrity and social responsibility. Focusing on ethically conscious and socially accountable Artificial Intelligence transcends academic discourse and it represents a compelling call to reform the ethos underpinning technological advancements for the benefit of humankind. This strategy envisages an Artificial Intelligence future that is not only remarkable for its technical prowess but is also deeply committed to fostering societal well-being, trust, and harmony.

Research Project Conclusion

This doctoral chapter has conducted an exhaustive examination of the interaction dynamics between human perceptions and the operational logic of Artificial Intelligence, with a special emphasis on applications concerning privacy and surveillance.

The pivotal findings of this study delineate a stark and multifaceted dichotomy in viewpoints, capturing the intricate nuances of human-AI interactions and the foundational concepts of these advanced technologies:

Misalignments in Perspectives

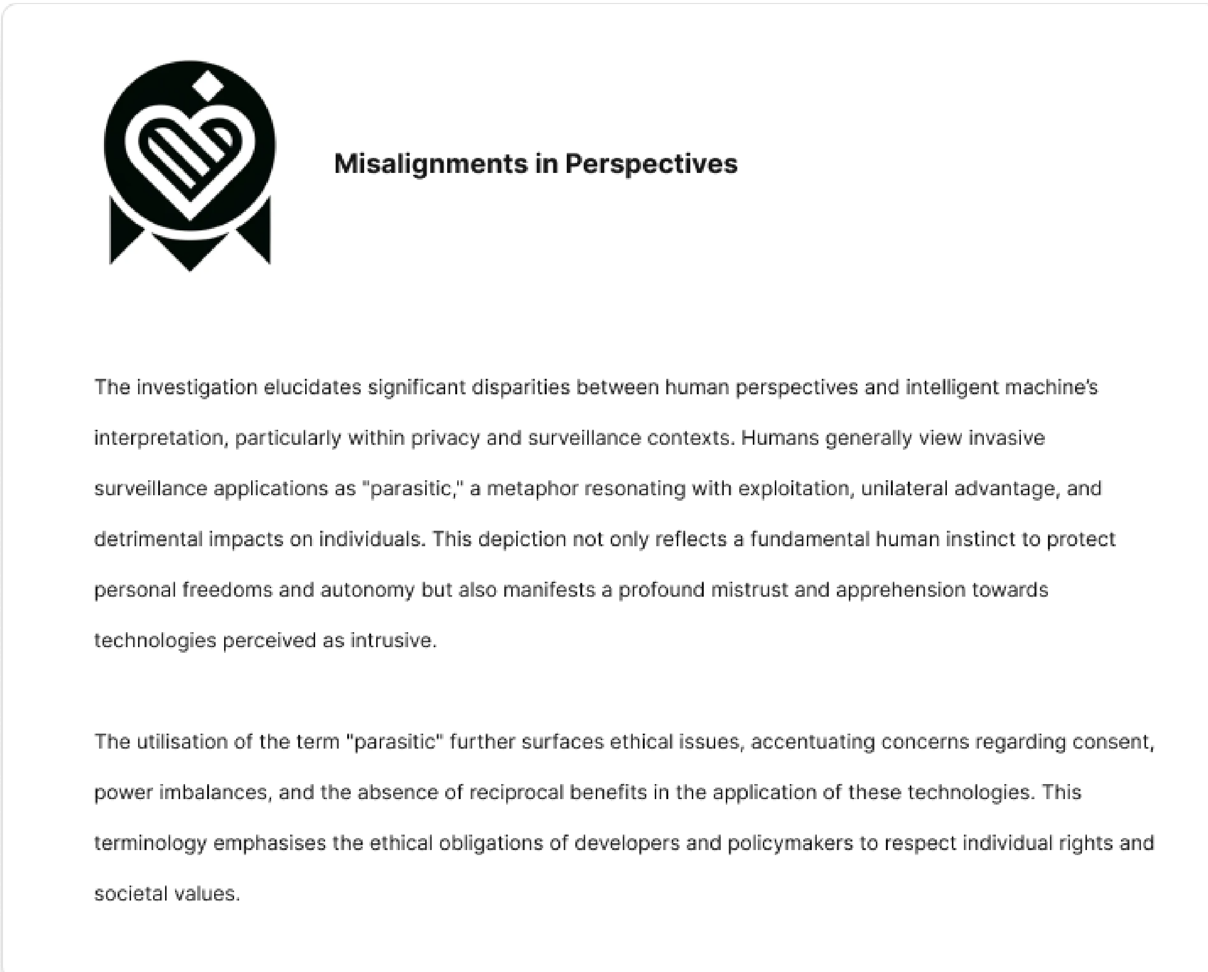


Figure 61 - Misalignments in Perspectives

Human vs. Artificial Intelligence Paradigms

Human vs AI Paradigms

Contrasting with human viewpoints, Artificial Intelligence systems often categorise these applications within a "mutualistic" paradigm, suggesting a balanced and harmonious integration offering mutual benefits. This perception demonstrates an Artificial Intelligence proclivity towards maximising systemic efficiency and promoting technological symbiosis, yet frequently negating the subjective human experiences of privacy violation and the ensuing ethical ramifications.

The AI-oriented perspective commonly overlooks the intricate subtleties of human emotional and ethical responses. Governed by algorithms and data-driven logic, Artificial Intelligence systems may not fully apprehend or value the human ramifications of privacy erosion, surveillance discomfort, or the moral quandaries these technologies introduce. This understanding gap underscores the necessity for a more empathetic and human-centric approach in Artificial Intelligence development that prioritises human impacts alongside functional capabilities.

Figure 62 - Human vs AI Paradigms

Alignment in Protective Applications

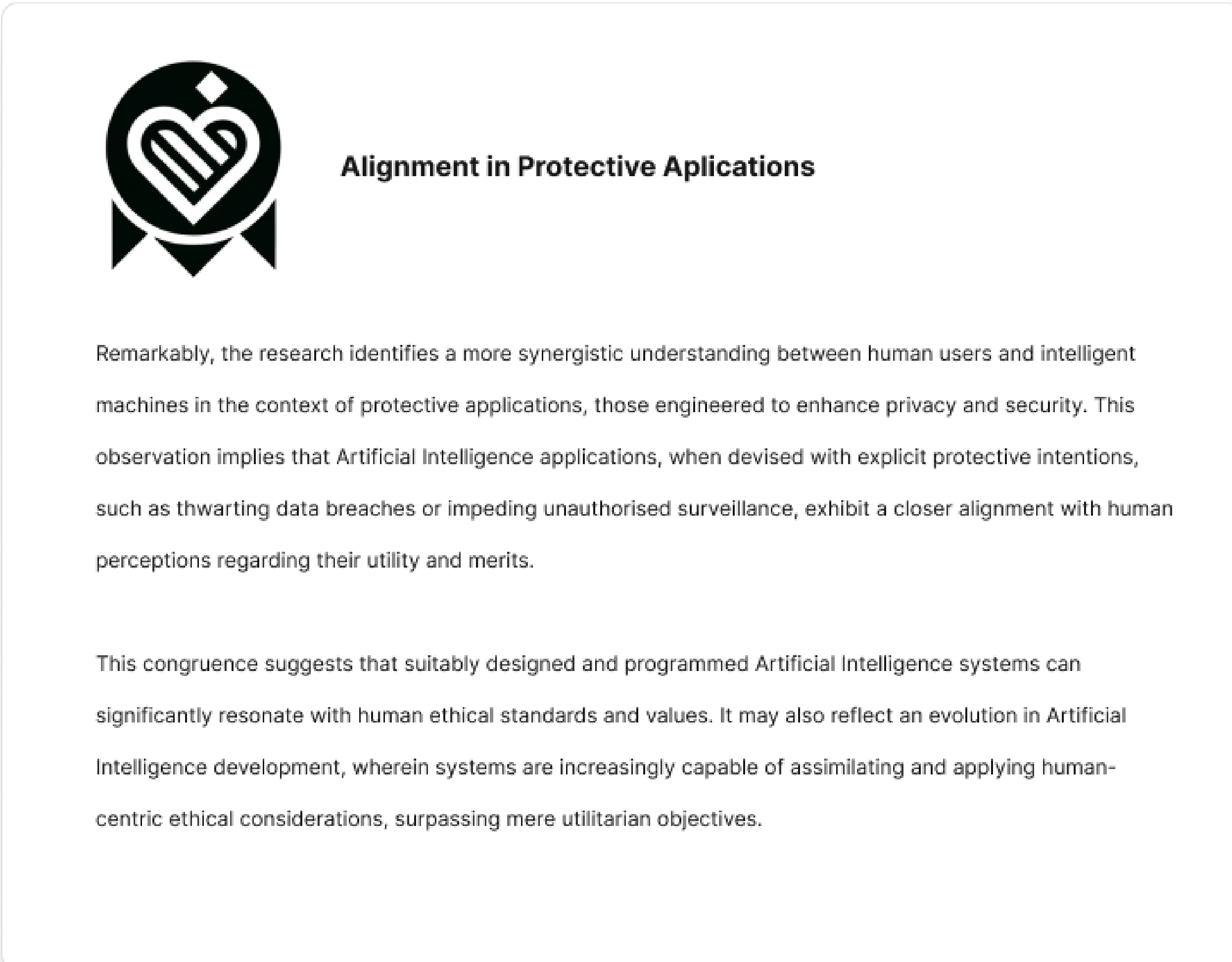


Figure 63 - Alignment in Protective Applications

Implications

The conclusions drawn from this research hold considerable significance for the formulation and regulation of Artificial Intelligence systems, particularly in applications centred around privacy. They necessitate a reevaluation and reformulation of Artificial

Intelligence design and deployment strategies, underlining the imperative to deeply embed human ethical standards into Artificial Intelligence developmental processes. The findings promote the symbiotic framework introduced in the previous chapter and advocate for establishing a dialogue among technologists, ethicists, policymakers, and the broader public to ensure that Artificial Intelligence's evolution aligns with its role as a facilitator of human values, not merely as a beacon of efficiency. Moreover, the study lays the groundwork for further inquiries into the ways Artificial Intelligence can be sculpted to resonate with and reflect symbiotic perspectives, potentially redefining the entire domain of privacy and surveillance technologies more profoundly.

Discussion of Core Arguments

This research delves deeply into the innovative concept of a symbiotic framework, which offers a transformative lens for examining the interplay of Artificial Intelligence with privacy and surveillance issues. The framework represents a significant shift from traditional analytical models for several reasons:

Innovation in Framework

One of the advantages in using the framework is the ecological Paradigm for AI Analysis. Inspired by biological symbiosis, the framework conceptualises Artificial

Intelligence and human interactions as analogous to biological relationships found in nature, ranging from mutualistic to parasitic. This metaphor transcends the typical binary categorisation of technology as merely beneficial or detrimental, enabling a more layered and dynamic understanding of Artificial Intelligence's influence within the social ecosystem. It considers the intricate interdependencies and the broader impacts that ripple through the system.

Another advantage of employing the framework is the Reframing of Ethical Considerations. Standard approaches in Artificial Intelligence ethics, primarily hinged on utilitarian (outcome-based) and deontological (duty-based) perspectives, are re-envisioned through this ecological lens. The symbiotic framework encourages an examination of Artificial Intelligence ethics that emphasises holistic system balance, interrelatedness, and the cumulative effects on an ecosystem, rather than focusing solely on isolated consequences or directives.

Human-Centered Design Imperative

The framework also focuses on human centred design and the benefits of this approach can be seen in the reconciliation of Attributes and Values. The framework illuminates significant divergences in how Artificial Intelligence and humans prioritise attributes like privacy, efficiency, and accuracy. Recognising and reconciling these differences is essential, not merely theoretically, but in practical terms, guiding the trajectory of Artificial Intelligence development toward designs that genuinely resonate with human ethical values and societal norms.

The research study also prioritises the practical implementation through the use of the symbiotic framework and human centred design. For Artificial Intelligence designers and researchers, the study exemplifies the incorporation of ethical considerations and human values at the foundational level of Artificial Intelligence development. This integration requires multidisciplinary collaboration, including insights from ethicists, sociologists, and end-users, to comprehensively infuse a wide range of human values and ethical principles into AI systems.

Ethical Alignment and Outcomes

Central to this framework is the concept of mapping and harmonising the characteristics and relationships between humans and Artificial Intelligence to foster more ethical outcomes. This strategy advocates for Artificial Intelligence development that balances technological innovation with a profound commitment to human-centric needs, expectations, and moral codes.

The symbiotic framework is a step towards Equitable AI. The ambition is to cultivate Artificial Intelligence systems that function not only as tools or subordinates but as collaborators within a shared ecological space. This entails crafting Artificial Intelligence capable of respecting privacy, ensuring fairness, and mitigating biases, thus contributing constructively to social welfare and equitable technological advancement.

An extended analysis in this chapter accentuates the transformative potential inherent in the symbiosis framework. By integrating this model, Artificial Intelligence development can evolve from a primarily technocentric perspective to an approach that is more ecocentric, emphasising ethical congruence and societal synergy. Such a paradigm shift heralds the advent of Artificial Intelligence systems that are more ethically conscious and attuned, fostering a nuanced comprehension of the complex and subtle dynamics between technology and the diverse aspects of human existence. The insights yielded by this research are poised to be pivotal in directing the future path of Artificial Intelligence development, steering it towards a trajectory characterised by responsible, empathetic, and human-aligned technological progression.

Emphasising Transformative Potential of the Framework

The symbiosis framework extends beyond theoretical speculation, offering a transformative roadmap for reimagining the interplay among AI, privacy, and surveillance. The framework's impact can be articulated in several key areas:

Evolutionary Pathways for Privacy and Surveillance in Artificial Intelligence


Evolutionary Pathways for Privacy and Surveillance in AI

This framework calls for a paradigmatic shift in the conception, design, and implementation of privacy and surveillance Artificial Intelligence. It advocates for an evolution that embeds ethical considerations and user-centric values from the inception of Artificial Intelligence development, rather than as mere appendages or in response to external regulatory frameworks.

A crucial component of this pathway is the promotion of transparency in Artificial Intelligence functionalities and decision-making processes. This involves Artificial Intelligence systems that not only comply with privacy legislations but also elucidate their data handling and algorithmic processes to users. Integrating robust data protection measures is equally vital, ensuring that Artificial Intelligence systems shield user data from misuse and enhance trust and dependability.

The framework envisages future Artificial Intelligence systems that harmoniously blend technological advancements with ethical obligations. This equilibrium necessitates systems that uphold user privacy, embody unbiased algorithmic practices, and are attuned to the varied socio-cultural dynamics of their user base.

Figure 64 - Evolutionary Pathways for Privacy and Surveillance in AI

Transformation of Relationships

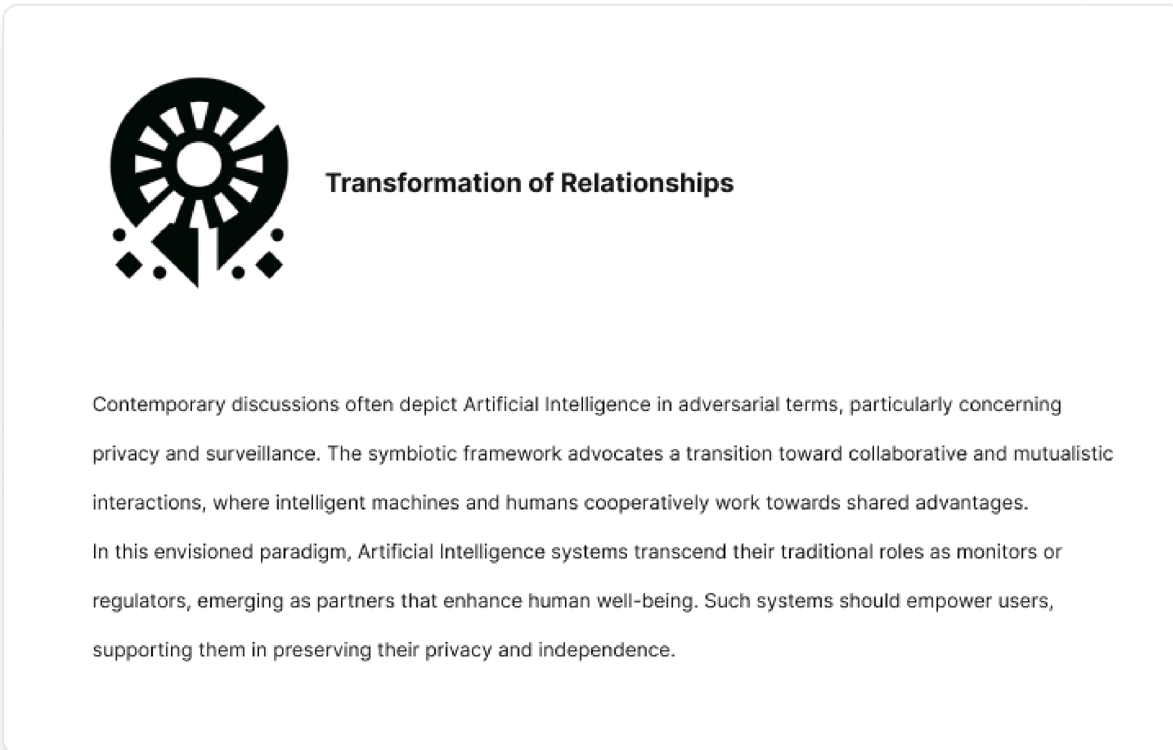


Figure 65 - Transformation of Relationships

Alignment with Humanistic Values

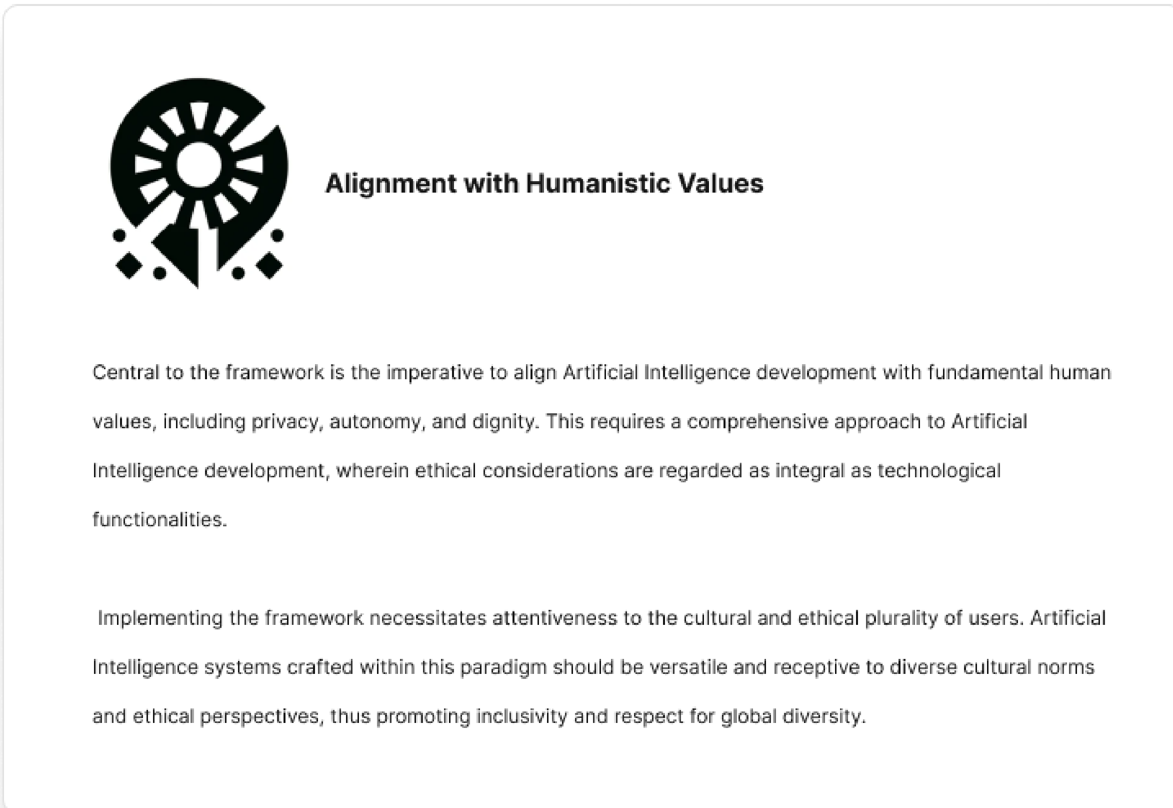


Figure 66 - Alignment with Humanistic Values

Broader Implications and Future Perspectives

The expansive potential of the symbiotic framework has significant ramifications. By reconceptualising the design and operational dynamics of Artificial Intelligence in the realms of privacy and surveillance, this framework has the capacity to fundamentally redirect Artificial Intelligence's developmental trajectory. It opens avenues for the creation of intelligent machine systems that are not just technologically sophisticated but also ethically informed and socially considerate. This paradigm shift could be a pivotal stride towards fostering Artificial Intelligence that not only augments human capabilities but also enriches societal welfare, ushering in an era where Artificial Intelligence is esteemed not for its invasiveness but for its empathy and congruence with human ambitions.

Reflections on the Evolution of Human-AI Relationships

The research marks an incipient foray into the complex, ever-evolving domain of ethical Artificial Intelligence. The intricate mosaic of human-AI relationships is in a state of constant flux, presenting an intricate interplay of challenges and opportunities. This research seeks to probe this nascent terrain, prioritising ethical congruity:

Charting the Course of Ethical AI

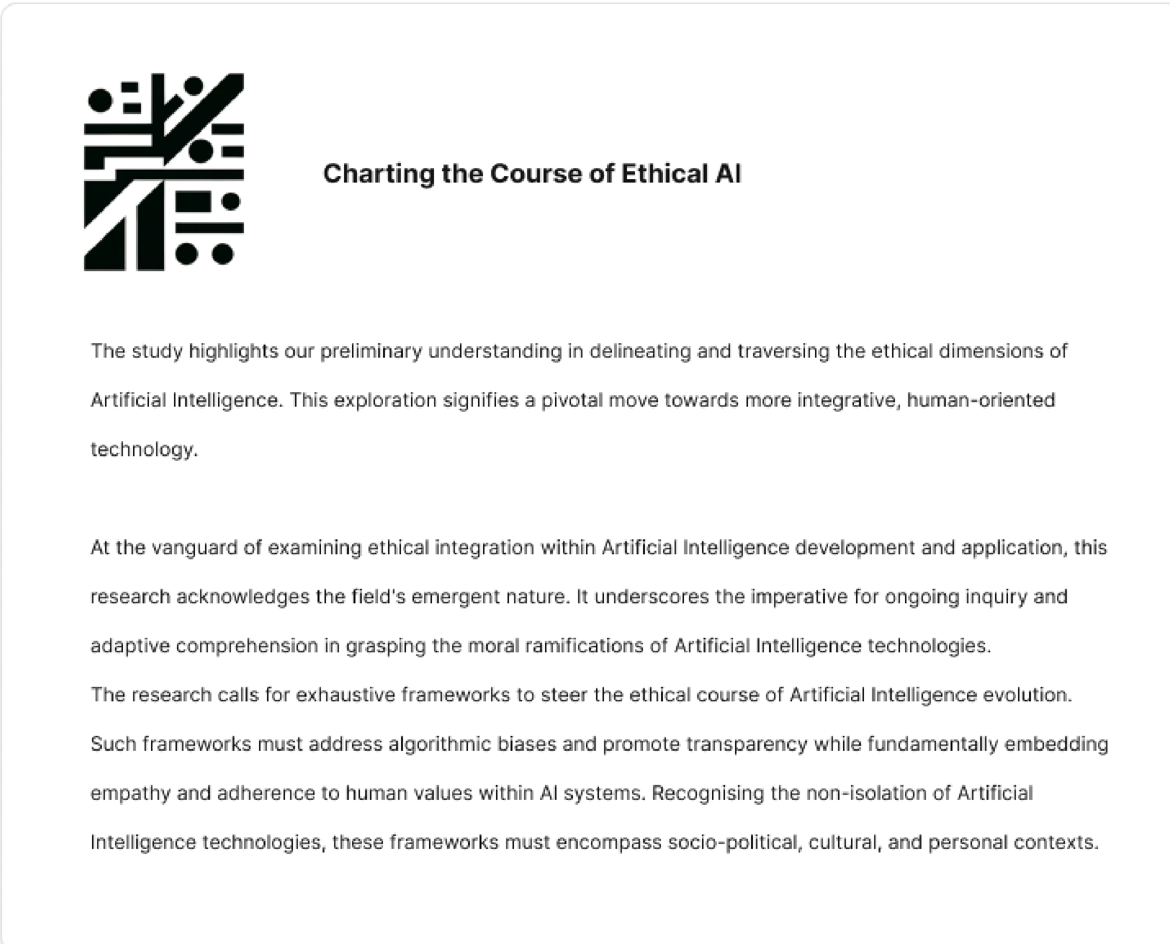


Figure 67 - Charting the Course of Ethical AI

The Imperative of Collaboration

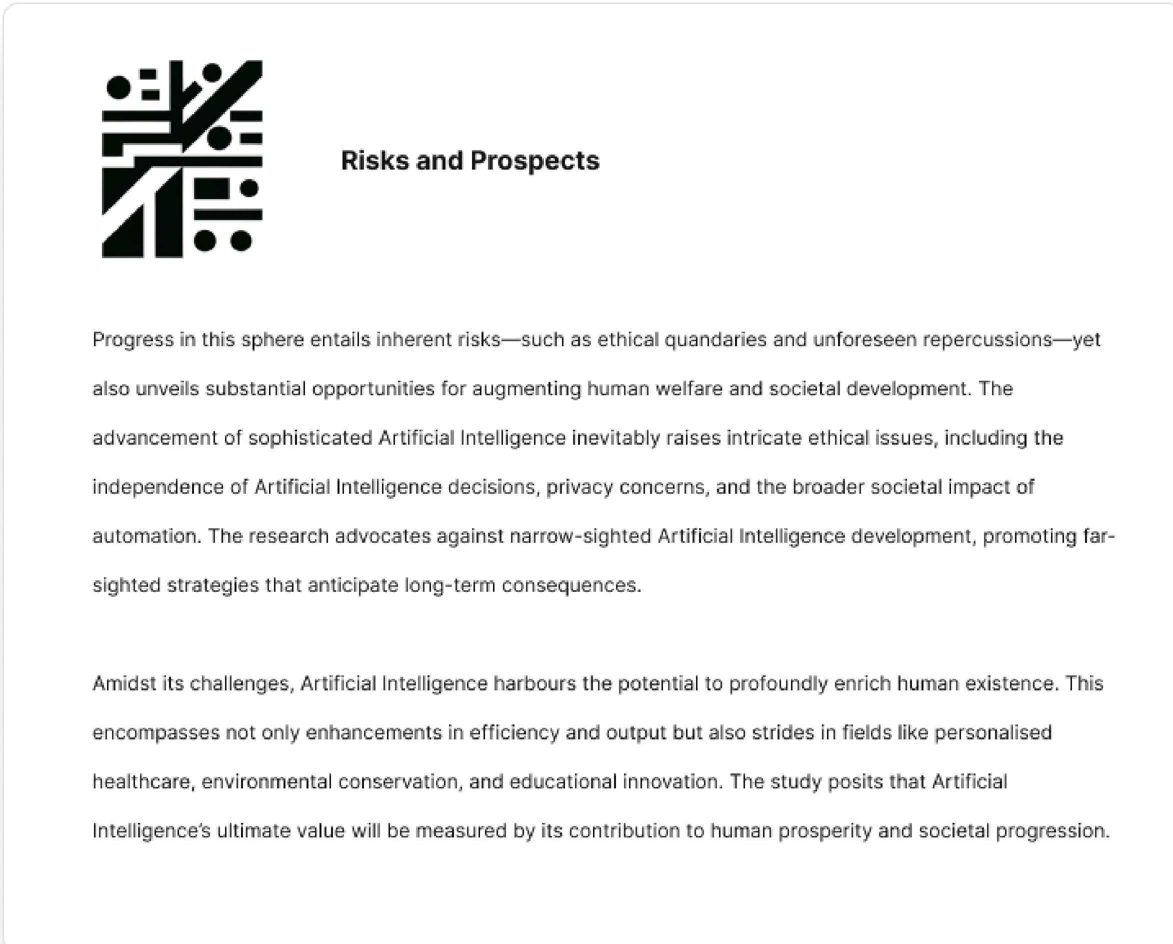
The Imperative of Collaboration

The research contends that fostering collaborative, mutually beneficial relationships between Artificial Intelligence and humans is of escalating significance. This partnership is crucial for maximising Artificial Intelligence's potential while maintaining human dignity and autonomy.

This study underscores the importance of developing partnerships between Artificial Intelligence and humans that transcend conventional user-machine interaction paradigms. It envisages a future in which Artificial Intelligence systems act not simply as tools or aids but as active participants in shared objectives.
This collaborative effort includes a dynamic learning process, with Artificial Intelligence systems and humans mutually informing and adapting to one another. Such a reciprocal relationship ensures that Artificial Intelligence evolves in a manner attuned to human needs and ethical standards, while humans assimilate and utilise Artificial Intelligence advancements for superior accomplishments.

Figure 68 - The Imperative of Collaboration

Risks and Prospects



Risks and Prospects

Progress in this sphere entails inherent risks—such as ethical quandaries and unforeseen repercussions—yet also unveils substantial opportunities for augmenting human welfare and societal development. The advancement of sophisticated Artificial Intelligence inevitably raises intricate ethical issues, including the independence of Artificial Intelligence decisions, privacy concerns, and the broader societal impact of automation. The research advocates against narrow-sighted Artificial Intelligence development, promoting far-sighted strategies that anticipate long-term consequences.

Amidst its challenges, Artificial Intelligence harbours the potential to profoundly enrich human existence. This encompasses not only enhancements in efficiency and output but also strides in fields like personalised healthcare, environmental conservation, and educational innovation. The study posits that Artificial Intelligence's ultimate value will be measured by its contribution to human prosperity and societal progression.

Figure 69 - Risks and Prospects

Closing Observations

As this study culminates, its core proposition accentuates the necessity of critically rethinking and redefining the interactive dynamics between humans and artificial intelligence. This reevaluation transcends academic rhetoric, constituting a crucial stride towards the synchronisation of technological advancements with the deeper echelons of human values and ethical standards. The insights gleaned from this research call for a fundamental paradigmatic shift, compelling a movement beyond the perception of Artificial Intelligence as a mere conduit for convenience or efficiency. Instead, Artificial Intelligence should be contemplated as an integral component of our societal matrix, reflective of and contributive to our most esteemed human ideals.

The points described below support the necessity and advantages of adopting a Human-AI symbiotic framework in practical applications:

Re-envisioning Human-AI Dynamics

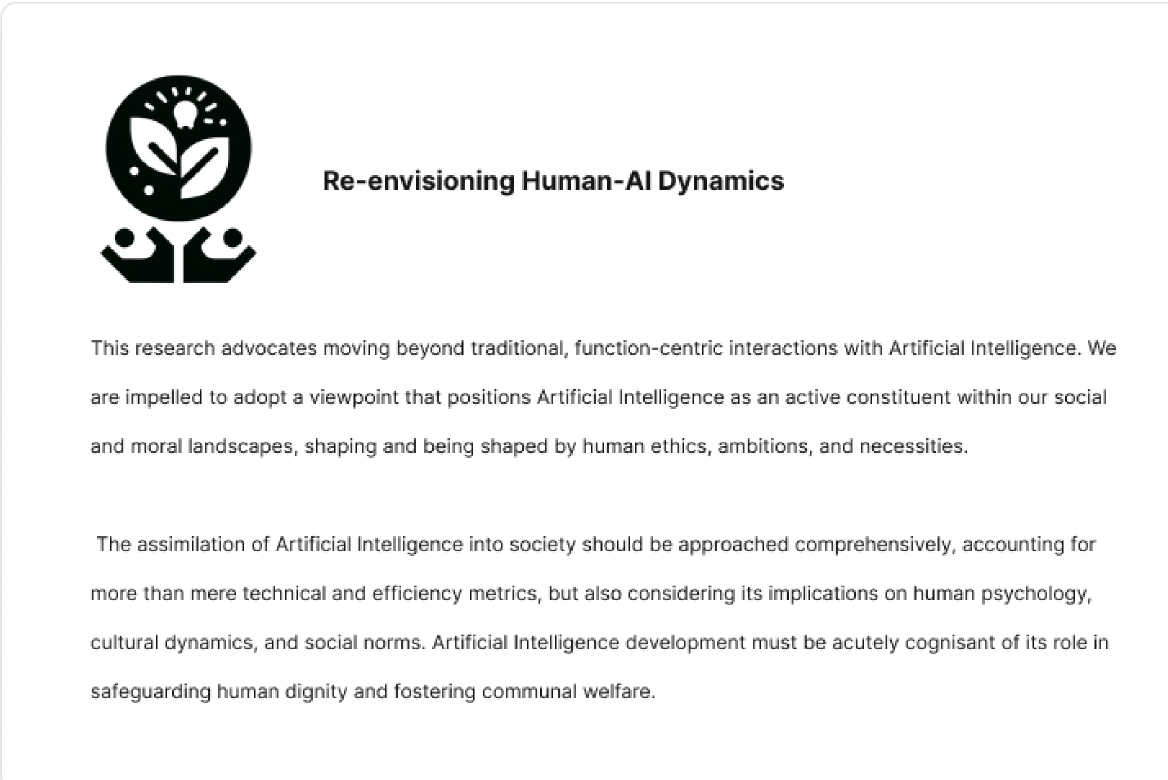


Figure 70 - Re-envisioning Human-AI Dynamics

Ethically Anchored Frameworks

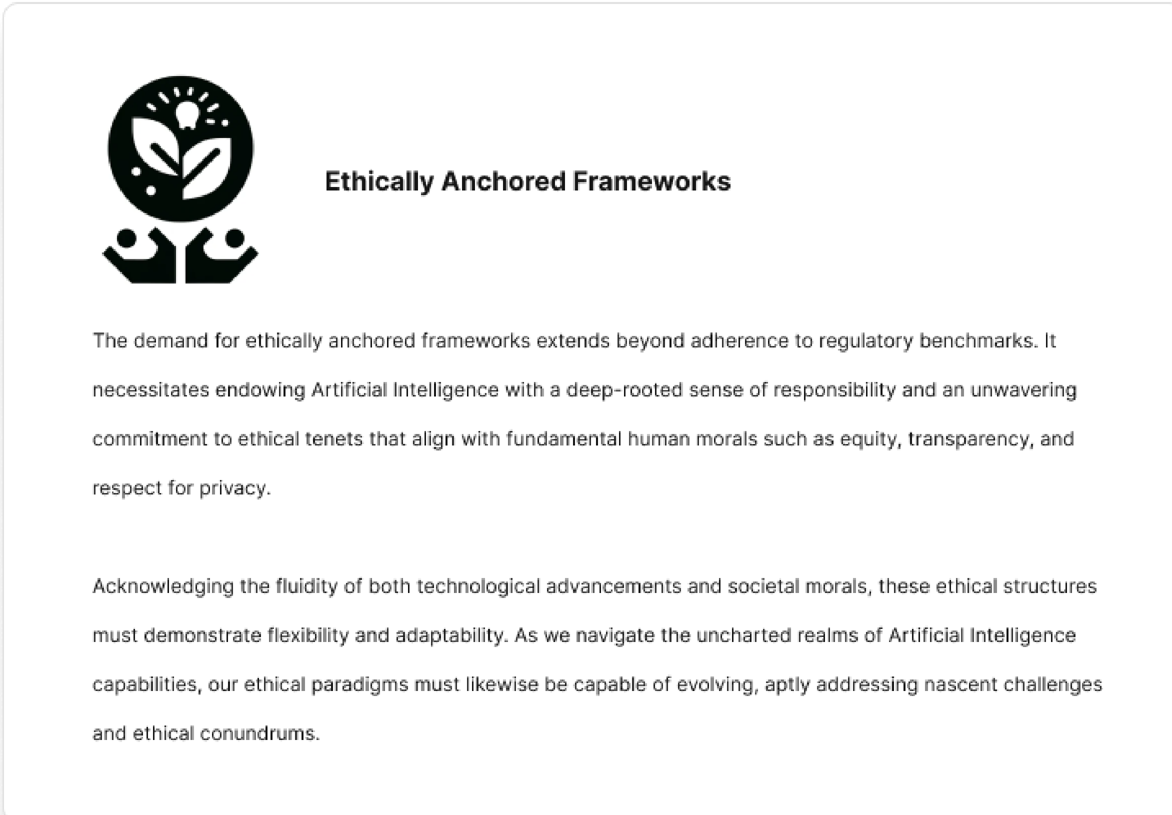


Figure 71 -Ethically Anchored Frameworks

Prospective Research Pathways

Figure 72 - Prospective Research Pathways

Paving the Way for Symbiotic Coexistence

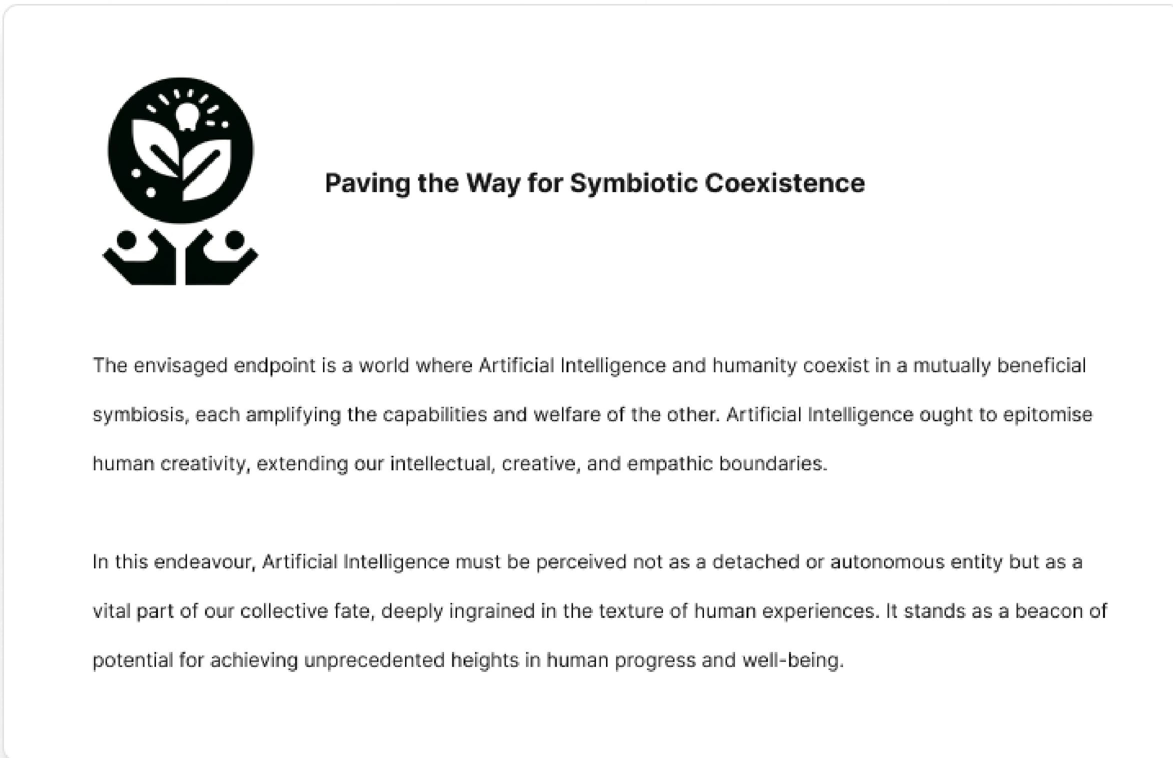


Figure 73 - Paving the Way for Symbiotic Coexistence

In closing, this research study establishes a robust platform for a new epoch of human-AI interaction, underscoring the imperative for a reflective, ethical, and human-centric approach in the development of Artificial Intelligence. It contends that the true valour of Artificial Intelligence is measured not solely by its technological prowess but in its capacity to enhance, rather than diminish, our humanity — fostering a world where technology and human virtues merge to forge a brighter, more inclusive future.

Chapter 07 / A Way Forward

“We do not learn from experience. We learn from reflecting on experience”

John Dewey on *How We Think*

In this chapter, I will delineate the present state of this research by revisiting the research questions and the trajectory undertaken to address these inquiries. Subsequently, I will reflect upon the contributions of this PhD research and explore prospective directions.

Research Questions Introspection

RQ01. How might symbiotic relationships observed in nature inspire and inform the design and development of machine intelligence systems?

RQ02. What are the potential capabilities of applying principles of biological symbiosis to human-machine interactions to enhance collaboration effectiveness?

RQ03. How might symbiotic design principles serve to address and ameliorate existing limitations within Artificial Intelligence algorithms and systems?

This PhD research was meticulously designed to address these three pivotal questions. As illustrated in the table below, distinct projects contribute answers to specific research questions. RQ01 received its answers through the "Journey of the Machine Learning Engineer" and "Introduction to a Symbiotic Human-AI Approach" projects. The inquiries for RQ02 were addressed in the projects focusing on types of symbiotic relationships, symbiotic primers, and the expansion of the symbiotic framework within the realms of privacy and security. Conversely, RQ03 found its discussions within the backdrop of all three research endeavours.

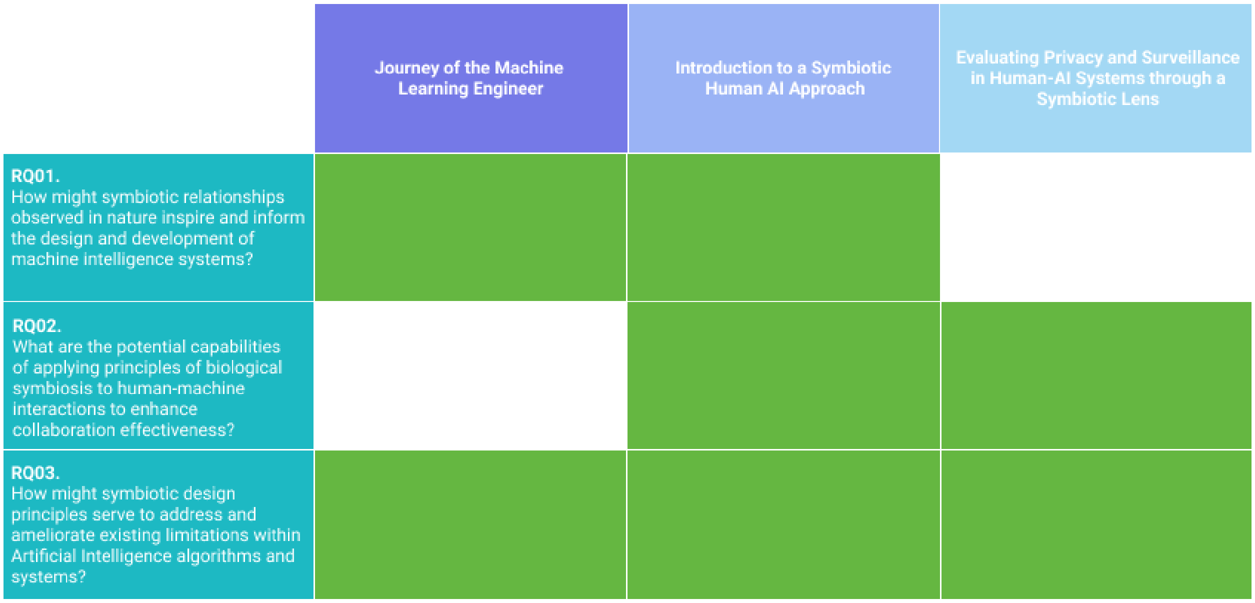


Figure 74 - Outcome from practical research and the relationship with the research questions

In terms of research depth, RQ01 emerges as the most expansive among the trio of questions, a deliberate choice to accommodate the research's top-down approach, its novelty, and its breadth. This expansive scope, closely intertwined with RQ02, underscores the strategic positioning of RQ01 as a comprehensive inquiry. RQ02, meanwhile, is characterized by a medium breadth and depth, spotlighting concepts vital not only to this PhD but also to the sustainability of human-machine symbiosis. RQ03, despite its omnipresence across all projects, delves into more detailed exploration regarding the depth and breadth of research.

RQ01 is elucidated through two project series, *Introduction to a Symbiotic Human-AI Approach* and *Journey of the Machine Learning Engineer*. Although high-level, the former series initiates a discourse on the role of design research in machine intelligence, evidencing a consensus among professionals from varied disciplines on the indispensable value of design research. This series advocates for increased visibility, accountability, and recognition of design research within the machine intelligence domains, proposing its central placement alongside computer scientists, statisticians, and philosophers. The *Journey of the Machine Learning Engineer* chronicles the researcher's first hand experiences, paralleling the challenges faced by machine learning engineers, thereby enriching the discourse with fresh insights into the crafting of intelligent machines.

RQ02 finds answers through the *Introduction to a Symbiotic Human-AI Approach* and the projects evaluating privacy and surveillance in human-AI systems through a symbiotic lens. The interconnection between RQ01 and RQ02 is evident, with RQ02 partially addressed through the exploration of symbiotic relationships and primers. These projects underscore the extraction of valuable lessons from symbiotic methodologies and the consensus on their significance across diverse academic fields. The focus on privacy and surveillance further applies these insights, offering practical applications and expanding the dialogue in human and intelligent machine relationships.

RQ03 is addressed comprehensively across the three practical project series, highlighting the first hand exploration of algorithmic limitations and proposing symbiotic relationships as a source of inspiration for overcoming these challenges. This approach not only acknowledges limitations within a practical context but also inspires professionals to devise alternative solutions through symbiotic methodologies.

In retrospect, it is my conviction that the PhD research has successfully responded to the posed research questions, therefore generating novel knowledge. Through personal experiences and professional feedback across various fields, the significance of design research in machine intelligence research domains has been accentuated. Design research introduces a praxis-oriented perspective, emphasizing hands-on approaches and the necessity to grapple with the technical intricacies of developing intelligent machines. This perspective also reveals inconsistencies within current machine intelligence research methodologies, advocating for a shift from analytical processes to experimental ones. Furthermore, design research's goal-directed methodology ensures a focused exploration of the research subject, maintaining alignment with established objectives.

Moreover, this PhD research substantiates the premise that insights from biological symbiotic relationships can be transposed to the realm of machine intelligence, offering strategies to ameliorate current algorithmic limitations through a focus on human-AI symbiosis. Although further research is warranted, the current findings provide compelling evidence that symbiotic approaches can innovatively influence machine intelligence, fostering new modalities of interaction between humans and intelligent machines. The research elucidates that symbiotic relationships introduce novel creative practices and innovative strategies to machine intelligence research. Despite the necessity for continued investigation to fully realize the potential of symbiotic relationships within machine intelligence, the documented research practices and collected data from participants and the researcher signify the transformative capability of symbiotic relationships as a catalyst for innovation and experimentation in Artificial Intelligence research.

Research Contributions

This PhD has contributed significantly to the domain of knowledge through practice-based research. While the impact of these contributions varies, their collective importance is unequivocal.

Biological Notions in Machine Intelligence

The research substantiates the concept that natural biological interactions can be leveraged to address challenges within machine intelligence. The application of biologically inspired concepts in engineering disciplines is not unprecedented; however, this research introduces a novel perspective by examining biological symbiotic interactions. This approach aims to inspire future investigations, akin to the influence of biomimetic architecture and material ecology. It represents a philosophical contribution, encouraging future scholars to explore not only biological symbiotic interactions but also other biological dynamics within the sphere of machine intelligence research.

Immersive Experiences Enhance Understanding

Engagement in the practical aspects of developing intelligent machines afforded me insights and expertise unattainable through theoretical study alone. Adopting a "learn by doing" methodology, I transitioned from a novice with no technical knowledge in machine learning to an intermediate-level engineer adept at constructing and debugging complex algorithms. This immersion qualifies as a practical knowledge contribution, serving as a paradigm for future researchers aspiring to acquire technical proficiency in artificial intelligence.

The Centrality of Data

A pivotal practical realization was identifying data as the foundational element of intelligent machines. This research elaborates on the meticulous attention to detail and organization required in data management for processing by intelligent machines. The tasks of collecting, preparing, and inputting data into machine learning algorithms are critical not only for the functionality of the intelligent machine but also for its human counterparts. The emphasis on data preprocessing highlights its significance in the development of intelligent machines, with data contamination through bias and developmental flaws being recurrent challenges. Optimal data management is crucial for addressing many vulnerabilities and issues within intelligent machines.

Computing Power as a Fundamental Resource

Another critical practical insight is the paramount importance of computing power, surpassing even original thought, in the development of intelligent machines. The incessant demand for computing power significantly influences the time, cost, and performance constraints of training intelligent machines. Conversely, the growing demand for computing power contributes to the current climate crisis, as the energy requirements for artificial intelligence's computing infrastructure escalate, exacerbating the environmental impact.

Creative Perspectives in Artificial Intelligence Research

This philosophical and methodological contribution highlights the predominance of technical perspectives in current artificial intelligence research. The foundational basis of artificial intelligence in mathematics and computer science has led to a technical orientation. Nevertheless, as artificial intelligence becomes more integrated into our daily lives, the recognition of novel perspectives is essential. Interviews conducted during this research indicate that biological concepts introduce fresh ideas into artificial intelligence, with design research acting as a conduit for these creative interpretations.

Categorising Human-Machine Symbiosis

The endeavour to categorize human-machine symbiosis represents a practical and methodological knowledge contribution. Although the research is in its nascent stages and necessitates further in-depth investigation, preliminary findings and expert feedback suggest that such categorization could introduce innovative concepts in artificial intelligence research. This categorization has the potential to develop new methods of interaction with intelligent machines and foster a more adaptable relationship with artificial intelligence products.

Biological Symbiotic Interactions as a Source of Knowledge

The research and expert feedback affirm that knowledge from biological interactions can be translated into the realm of artificial intelligence. This top-down research approach addresses the largely unexplored research space of human-machine symbiosis, emphasizing the need for high-level research depth. Given that biological symbiotic interactions are based on negotiations over organisms' life cycles, this research adds value to long-term studies on human-machine symbiosis, positioning biological symbiotic interactions as a significant source of new knowledge for both this PhD and future research in machine intelligence.

Symbiotic Primers as Initial Approaches

Symbiotic primers, conceptual frameworks linking specific symbiotic relationships with machine intelligence tasks, represent a substantial practical knowledge contribution. Through expert interviews, this research has begun to map out how machine intelligence tasks can be associated with symbiotic relationships, offering foundational insights for redefining these relationships within various tasks. For instance, the possibility of transforming parasitic relationships in biometric analysis to commensalism highlights the potential for significant advancements in the application of symbiotic principles to artificial intelligence.

Design Research as a Conduit

Philosophically, design research serves as a bridge connecting the biological world with that of artificial intelligence. This PhD underscores the value of design research in fostering creative interpretations and introducing innovative ideas into artificial intelligence research. Furthermore, design research has proven essential in translating biological knowledge into practical strategies within the artificial intelligence domain. Discussions with experts throughout this PhD have reinforced the unique position of design research in artificial intelligence, highlighting the recognition of its value when given a platform.

Reflections about the research

Every research endeavour encounters its share of challenges, and this investigation is no different. Anticipating research limitations is crucial, as they substantially influence the interpretation of findings. Acknowledgement of these limitations not only underscores a critical examination of the study but also propels recommendations for future research.

One such limitation arose from the interdisciplinary nature of the study. It proved challenging to recruit participants with expertise in both biological interactions and machine intelligence, a scarcity attributable to the nascent state of research in this field. Although this predicament was identified at the study's inception, ideally, participants would possess a foundational understanding of machine intelligence, design research, and biological interactions. The remedial strategy involved enlisting individuals with substantial expertise in at least one pertinent field—be it machine intelligence, design research, or biological interactions.

Another constraint pertained to the acquisition of appropriate resources and expertise. Initially, the researcher's expertise was confined to web development, with no grounding in machine learning or deep learning. The resultant steep learning curve necessitated a significant temporal investment in mastering machine learning, deep learning, reinforcement learning, and artificial intelligence theory. Access to more comprehensive resources could have mitigated this learning curve. Nonetheless, this limitation inadvertently contributed to the researcher's mastery of machine intelligence through a process of trial and error, ultimately fostering an expert understanding in this area.

Challenges and Key Moments

The challenges and pivotal moments of this research can be conceptualized through the metaphor of Chinese boxes—a series of nested containers where each box is encompassed by a larger one, symbolizing that within one framework or perspective lies another, each telling a distinct narrative.

There are four principal challenges followed by four key moments that warrant discussion. The foremost challenge was to synthesize concepts from two disparate research areas: biological interactions and machine intelligence. These fields are seldom correlated and locating an intersection between them was arduous. The breakthrough came when I acquired a robust understanding of biological interactions and symbiotic theories, coupled with an appreciation for the constraints and prospects within machine intelligence. This epiphany, occurring during the literature review phase, afforded me a novel viewpoint on the research.

The second challenge involved conducting research in a domain with scant existing literature. While biological interactions and machine intelligence are extensively studied independently, few studies explore them in tandem. Biomimetics and Material Ecology, disciplines related to this research, provided some guidance, but their aims diverge from those of my study. The realization of this gap in research galvanized my commitment, as I recognized that my work could lay the groundwork for future endeavors in design research and human-computer interaction. This key moment underscored the long-term value of pioneering new research avenues despite inherent risks.

The third challenge concerned the steep learning trajectory in machine learning, deep learning, and reinforcement learning. Eager to establish a tactile connection with the development of intelligent machines, I devoted substantial time to mastering the tenets of artificial intelligence and its practical applications. Despite my background in web development, delving into these fields from scratch was formidable. The onerous learning process, however, culminated in a pivotal thesis moment. By learning to construct intelligent machines, I gained fluency in the field, progressively tackling more complex problems, especially in reinforcement learning. This journey bolstered my confidence as a hands-on researcher and now serves as a model for future scholars eager to navigate the intricacies of developing intelligent machines.

The final challenge pertained to the research perspective of this thesis, which endeavoured to examine machine intelligence through a design research lens—a relatively uncharted approach. Given that artificial intelligence is fundamentally philosophical and its subsets—machine learning, deep learning, and reinforcement learning—are technically intricate, adopting this unique perspective was daunting. Incremental insights throughout the practical research segment gradually addressed this challenge. However, the most salient questions were resolved during the creation of the Symbiotic Human-AI approach. This research suggests that design research can reveal numerous innovations within machine intelligence, with the Symbiotic Human-AI approach exemplifying this untapped potential. In retrospect, while the Symbiotic Human-AI approach did not rectify any specific machine intelligence issues, it serves as a beacon for subsequent research in machine intelligence that transcends technical considerations. This investigation demonstrates that diverse perspectives can yield intriguing results and innovative ideas, particularly when design research is integrated into machine learning, deep learning, and reinforcement learning.

A Final Reflection

This thesis represented a singular odyssey, warranting reflection not solely on the outcomes but also on the path to these discoveries. As a researcher, I assimilated invaluable lessons in resolving intricate issues and examining research problems from diverse perspectives. Nonetheless, extensive work remains to thoroughly investigate the potential of symbiotic relationships between humans and intelligent machines. The fruits of this research should act as both an inspiration and a form of foundational study. The methodology of this thesis opens a plethora of opportunities, and future research should be emboldened to delve into novel concepts rather than be daunted by them.

Overall, the research posed considerable challenges, but the resultant rewards justified the difficulties encountered. Personally, I matured not only as a researcher but also as a designer and an engineer. This trajectory of growth was arduous, marked by numerous challenges that ultimately facilitated my professional development. One of the most formidable challenges was to cohesively address all the research domains encompassed by the thesis. As previously mentioned, the thesis straddles two distinct fields: biological interactions and machine intelligence. Discovering synergies between these domains was often challenging. My chosen method involved an exhaustive study of each area before attempting to integrate them. While I am content with the outcomes of this 'outside-in' approach, I posit that an 'inside-out' strategy could also prove beneficial, presenting an opportunity for future research initiatives.

It is equally critical to contemplate factors that adversely influenced the research. A primary drawback was the duration required to define a research scope. The initial proposal was overly ambitious, and as my understanding of the research problem deepened, I found it necessary to continually refine the scope. This process was more protracted than anticipated and detracted from the research. Moreover, the novelty of the research was a double-edged sword. While pioneering new territory was exhilarating, the dearth of guidance and resources hindered progress, compelling reliance on a trial-and-error methodology that was time-intensive—a precious commodity in research endeavours.

I contend that this thesis should be viewed as an inaugural phase in an iterative process. Future enhancements to the research could stem from retrospection, learning from the missteps of this venture. With successive iterations, the research is poised to become more sophisticated and deliver immediate, tangible benefits to design researchers in the field of machine intelligence.

Conclusion

This doctoral research advocates for the integration of design research within the field of artificial intelligence and corroborates the imperative for continued exploration of human-machine symbiosis. The findings from this practice-based research are auspicious, delineating a trajectory for further scholarly inquiry characterized by both expansive and intensive dimensions. Moreover, this research endeavours to act as a fulcrum, inciting design researchers to tackle artificial intelligence challenges, to innovate, and to navigate the confines of the research milieu.

The examination and application of insights from biological interactions infuse a fresh perspective into the predominantly technical domain of artificial intelligence. This discipline, which relies heavily on technical acumen, necessitates inventive interpretations to advance its frontiers. This study has demonstrated that such innovative interpretations not only have the capacity to engender new knowledge but are also welcomed by practitioners within this domain, who are receptive to the integration of concepts from biological interactions.

Despite the breadth of feedback from experts across various disciplines, the research within this PhD was conducted by a solitary researcher, presenting a distinctive array of benefits and drawbacks. Personally, this path has fostered my growth into a more seasoned researcher, with the journey itself eclipsing the significance of the destination. The unpredictability and intricacy encountered along this path have not only honed my research and software development skills but have also refined my capacity for critical reflection and decision-making. It is my aspiration that this PhD journey will embolden future design researchers to delve into research conundrums and navigate virgin territories with assurance and optimism.

In the future, I aim to persist in exploring, championing, and propagating the insights garnered from this research. As indicated earlier, the potential within this sphere of study is considerable, and the concepts introduced in this PhD warrant further maturation. Such sophistication can be realized through fostering a collaborative ethos with specialists from diverse disciplines and by persuading subsequent cohorts of design researchers to embrace the challenges inherent in human-machine symbiosis. This approach should be predicated on a willingness to endure discomfort and the recognition that interdisciplinary discourse yields more fruitful results.

Bibliography

Adelman, C. (1993). Kurt Lewin and the Origins of Action Research. *Educational Action Research*, *1*(1), 7–24. https://doi.org/10.1080/0965079930010102

AlphaGo | DeepMind. (n.d.).

Altrichter, H., Kemmis, S., Mctaggart, R., & Zuber-Skerritt, O. (2002). The concept of action research. *The Learning Organization*, *9*(3), 125–131. https://doi.org/10.1108/09696470210428840

Anderson, M., & Anderson, S. L. (2011). *Machine Ethics* . Cambridge University Press. https://books.google.es/books?id=N4IF2p4w7uwC&dq=Anderson,+Michael%3B+Anderson,+Susan+Leigh+Machine+Ethics&lr=&source=gbs_navlinks_s

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2019). *Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI*.

Asimov, I. (2004a). *I, robot*. 224.

Asimov, I. (2004b). *I, robot*. 224. https://books.google.com/books/about/I_Robot.html?id=2vnbMzYXBQsC

Auernhammer, J. (2020, September 10). *Human-centered AI: The role of Human-centered Design Research in the development of AI*. https://doi.org/10.21606/drs.2020.282

Balsamo, A. M. (1996). *Technologies of the Gendered Body: Reading Cyborg Women*. Duke University Press.

Bateson, G. (2000). *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. University of Chicago Press.

Benyus, J. M. (2009). *Biomimicry : innovation inspired by nature*.

Berk, R., Heidari, H., Jabbari, S., Kearns, M., & Roth, A. (2021). Fairness in Criminal Justice Risk Assessments: The State of the Art. *Sociological Methods & Research*, *50*(1), 3–44. https://doi.org/10.1177/0049124118782533

Berner, C., Brockman, G., Chan, B., Cheung, V., Dennison, C., Farhi, D., Fischer, Q., Hashme, S., Hesse, C., Józefowicz, R., Gray, S., Olsson, C., Pachocki, J., Petrov, M., de Oliveira Pinto, H. P., Raiman, J., Salimans, T., Schlatter, J., Schneider, J., … Zhang, S. (2019). *Dota 2 with Large Scale Deep Reinforcement Learning*.

Birks, M., & Mills, J. (2012). *Grounded Theory - A Practical Guide*.

Bostrom, N. (2014a). *Superintelligence: Paths, Dangers, Strategies*.

Bostrom, N. (2014b). *Superintelligence: Paths, Dangers, Strategies*. https://books.google.es/books?id=7_H8AwAAQBAJ&source=gbs_navlinks_s

Boucher, D. H. (1985). *The Biology of Mutualism: Ecology and Evolution*. Oxford University Press.

Bouke, M. A., Abdullah, A., ALshatebi, S. H., Zaid, S. A., & Atigh, H. El. (2023). The intersection of targeted advertising and security: Unraveling the mystery of overheard conversations. *Telematics and Informatics Reports*, *11*, 100092. https://doi.org/https://doi.org/10.1016/j.teler.2023.100092

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
Breazeal, C., Dautenhahn, K., & Kanda, T. (2016). *Social Robotics* (pp. 1935–1972). https://doi.org/10.1007/978-3-319-32552-1_72
Bronstein, J. L. (2015). *Mutualism*. Oxford University Press.
Brooks, R. (2002). *Flesh and Machines: How Robots Will Change Us*. Knopf Doubleday Publishing Group.
Brooks, R. A. (1999). *Cambrian Intelligence: The Early History of the New AI*. BRADFORD BOOK.
Broussard, M. (2018). *Artificial unintelligence : how computers misunderstand the world*.
Brown, J. D., & Coombe, C. A. (2015). The Cambridge guide to research in language teaching and learning. In *Guide to research in language teaching and learning*.
Brown, N., & Sandholm, T. (2019). Superhuman AI for multiplayer poker. *Science*, *365*(6456), 885–890. https://doi.org/10.1126/SCIENCE.AAY2400
Bryant, A., & Charmaz, K. (2007). The SAGE Handbook of Grounded Theory. In *The SAGE Handbook of Grounded Theory*. SAGE Publications Ltd. https://doi.org/10.4135/9781848607941
Buchner, P. (1965). *Endosymbiosis of Animals with Plant Microorganisms*. Wiley.
Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In S. A. Friedler & C. Wilson (Eds.), *Proceedings of the 1st Conference on Fairness, Accountability and Transparency* (Vol. 81, pp. 77–91). PMLR. https://proceedings.mlr.press/v81/buolamwini18a.html
Calo, C. J., Hunt-Bull, N., Lewis, L., & Metzler, T. (2011). Ethical Implications of Using the Paro Robot, with a Focus on Dementia Patient Care. *Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence*.
Campolo, A., Sanfilippo, M., Whittaker, M., & Crawford, K. (2017). *AI Now 2017 Report*. https://assets.contentful.com/8wprhhvnpfc0/1A9c3ZTCZa2KEYM64Wsc2a/8636557c5fb14f2b74b2be64c3ce0c78/_AI_Now_Institute_2017_Report_.pdf
Capra, F., & Luisi, P. L. (2014). *The Systems View of Life: A Unifying Vision*. Cambridge University Press.
Charmaz, K. (2014). *Constructing Grounded Theory*. SAGE Publications Ltd.
Chollet, F. (2018). Deep Learning with Python, Manning. *Manning*, 386.
Clark, A. (2003a). *Natural-born cyborgs : minds, technologies, and the future of human intelligence*. 229.
Clark, A. (2003b). *Natural-born cyborgs : minds, technologies, and the future of human intelligence*. 229. https://books.google.com/books/about/Natural_born_Cyborgs.html?id=8JXaK3sREXQC
Clark, A. (2008a). *Supersizing the mind : embodiment, action, and cognitive extension*. Oxford University Press.
Clark, A. (2008b). *Supersizing the mind : embodiment, action, and cognitive extension*. Oxford University Press.
Coeckelbergh, M. (2020). *AI ethics*. https://books.google.com/books/about/AI_Ethics.html?id=Gs_XDwAAQBAJ

Coeckelbergh, M., Loh, J., & Funk, M. (2018). *Envisioning Robots in Society – Power, Politics, and Public Space: Proceedings of Robophilosophy 2018 / TRANSOR 2018*. IOS Press. https://books.google.ch/books?id=jLh9DwAAQBAJ
Combes, C. (2001). *Parasitism: The Ecology and Evolution of Intimate Interactions*. University of Chicago Press.
Crawford, K., Dobbe, R., Dryer, T., Fried, G., Green, B., Kaziunas, E., Kak, A., Mathur, V., McElroy, E., Sanchez, A. N., Raji, D., Rankin, J. L., Richardson, R., Schultz, J., West, M. S., & Whittaker, M. (2019). *AI Now 2019 Report*. https://ainowinstitute.org/AI_Now_2019_Report.pdf
Creswell, J. W., & Clark, V. L. P. (2011). *Designing and Conducting Mixed Methods Research*. SAGE Publications. https://books.google.ch/books?id=6tYNo0UpEqkC
Cuddington, K., Byers, J. E., Wilson, W. G., & Hastings, A. (2011). *Ecosystem Engineers: Plants to Protists*. Elsevier Science.
Deleuze, G., Guattari, F., & Massumi, B. (2004). *EPZ Thousand Plateaus*. Bloomsbury Academic.
Dick, P. K. (1968a). *Do Androids Dream of Electric Sheep?*
Dick, P. K. (1968b). *Do Androids Dream of Electric Sheep?* https://www.goodreads.com/book/show/36402034-do-androids-dream-of-electric-sheep
Dickens, L., & Watkins, K. (1999). Action Research: Rethinking Lewin. *Management Learning*, *30*(2), 127–140. https://doi.org/10.1177/1350507699302002
Douglas, A. E. (2010a). *The symbiotic habit*.
Douglas, A. E. (2010b). *The symbiotic habit*.
Dreyfus, H. L. (1972). *What Computers Can't Do: A Critique of Artificial Reason*. Harper \& Row.
Dreyfus, H. L. (1992). *What Computers Still Can't Do: A Critique of Artificial Reason*. MIT Press.
Dubber, M. D., Pasquale, F., & Das, S. (2020). *The Oxford Handbook of Ethics of AI* (M. D. Dubber, F. Pasquale, & S. Das, Eds.). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190067397.001.0001
Dyson, G. (2012). *Darwin among the machines*.
Engelbart, D. C. (2021). Augmenting Human Intellect: A Conceptual Framework (1962). In *Ideas That Created the Future: Classic Papers of Computer Science*. The MIT Press. https://doi.org/10.7551/mitpress/12274.003.0024
Fan, Y., & Pedersen, O. (2021). Gut microbiota in human metabolic health and disease. *Nature Reviews Microbiology*, *19*(1), 55–71. https://doi.org/10.1038/s41579-020-0433-9
Fellous, S., & Salvaudon, L. (2009). How can your parasites become your allies? *Trends in Parasitology*, *25*(2), 62–66. https://doi.org/https://doi.org/10.1016/j.pt.2008.11.010
Ferguson, A. G. (2017a). *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. NYU Press. https://doi.org/10.2307/j.ctt1pwtb27
Ferguson, A. G. (2017b). *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. NYU Press. https://doi.org/10.2307/j.ctt1pwtb27
Floridi, L. (2014a). *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. OUP Oxford.
Floridi, L. (2014b). *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. OUP Oxford. https://books.google.ch/books?id=hOF_AwAAQBAJ
Foth, M., & Axup, J. (2006). *Participatory Design and Action Research: Identical Twins or Synergetic Pair?*

Frayling, C. (1994). *Research in Art and Design (Royal College of Art Research Papers, Vol 1, No 1, 1993/4)*.

Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics 1980 36:4*, *36*(4), 193–202. https://doi.org/10.1007/BF00344251

Gaver, B., Dunne, T., & Pacenti, E. (1999). Design: Cultural probes. *Interactions*, *6*(1), 21–29. https://doi.org/10.1145/291224.291235

Géron, A. (2019). Hands-on machine learning with Scikit-Learn, Keras and TensorFlow: concepts, tools, and techniques to build intelligent systems. *O'Reilly Media*, 851.

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning* (1st ed.). Addison-Wesley Longman Publishing Co., Inc.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative Adversarial Nets*.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a Conceptual Framework for Mixed-Method Evaluation Designs. *Educational Evaluation and Policy Analysis*, *11*(3), 255–274. https://doi.org/10.3102/01623737011003255

Grgić-Hlača, N., Zafar, M. B., Gummadi, K. P., & Weller, A. (n.d.). *On Fairness, Diversity and Randomness in Algorithmic Decision Making*. https://arxiv.org/pdf/1706.10208.pdf

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G.-Z. (2019). XAI—Explainable artificial intelligence. *Science Robotics*, *4*(37). https://doi.org/10.1126/scirobotics.aay7120

Hadj-Hammou, J., Mouillot, D., & Graham, N. A. J. (2021). Response and Effect Traits of Coral Reef Fish. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.640619

Hall, A. (2011). Experimental design: Design experimentation. *Design Issues*, *27*(2), 17–26. https://doi.org/10.1162/DESI_a_00074-Hall

Haraway, D. (2006). *A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late 20th Century BT - The International Handbook of Virtual Learning Environments* (J. Weiss, J. Nolan, J. Hunsinger, & P. Trifonas, Eds.; pp. 117–158). Springer Netherlands. https://doi.org/10.1007/978-1-4020-3803-7_4

Haraway, D. J. (2013). *When Species Meet*. University of Minnesota Press.

Haraway, D. J. (2015). *Simians, Cyborgs, and Women: The Reinvention of Nature*. Routledge. https://books.google.ch/books?id=xdP4oAEACAAJ

He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2020). Mask R-CNN. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, *42*(02), 386–397. https://doi.org/10.1109/TPAMI.2018.2844175

Heidegger, M. (1977). *The Question Concerning Technology, and Other Essays*. Garland Pub.

Hobbes, T. (1651). *Leviathan, Or, The Matter, Form, and Power of a Common-wealth Ecclesiastical and Civil*. Andrew Crooke.

Hu, K., Liao, W., Yang, M. Y., & Rosenhahn, B. (2021). *Text to Image Generation with Semantic-Spatial Aware GAN*.

Hulme-Beaman, A., Dobney, K., Cucchi, T., & Searle, J. B. (2016). An Ecological and Evolutionary Framework for Commensalism in Anthropogenic Environments. *Trends in Ecology & Evolution*, *31*(8), 633–645. https://doi.org/10.1016/J.TREE.2016.05.001

Janssen, A. W. F., & Kersten, S. (2015). The role of the gut microbiota in metabolic health. *The FASEB Journal*, *29*(8), 3111–3123. https://doi.org/10.1096/fj.14-269514

JJ, W. (2012). Endosymbiosis. *Current Biology : CB*, *22*(14). https://doi.org/10.1016/J.CUB.2012.06.010

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research*, *1*(2), 112–133. https://doi.org/10.1177/1558689806298224

Joost, G., Bredies, K., Christensen, M., Conradi, F., Unteidig, A., & of International Research in Design, B. (2016). *Design as research : positions, arguments, perspectives*.

Jorgensen, S. E., & Fath, B. D. (2008). *Encyclopedia of Ecology*. ScienceDirect.

Kagan, C., Burton, M., & Siddiquee, A. (2006). Participatory design and action research: Identical twins or synergetic pair. *Handbook of Qualitative Research Methods in Psychology*, *i*(5), 93–96.

Karras, T., Laine, S., & Aila, T. (2018). *A Style-Based Generator Architecture for Generative Adversarial Networks Timo Aila*.

Kelly, K. (2009). *Out Of Control: The New Biology Of Machines, Social Systems, And The Economic World*. Basic Books.

Kemmis, S. (2009). Action research as a practice-based practice. *Educational Action Research*, *17*(3), 463–474. https://doi.org/10.1080/09650790903093284

Kemmis, S., McTaggart, R., & Nixon, R. (2014). Introducing Critical Participatory Action Research. In *The Action Research Planner* (pp. 1–31). Springer Singapore. https://doi.org/10.1007/978-981-4560-67-2_1

Kim, T., Cha, M., Kim, H., Lee, J. K., & Kim, J. (2017). *Learning to Discover Cross-Domain Relations with Generative Adversarial Networks*.

Kurzweil, R. (2005). *The Singularity is Near: When Humans Transcend Biology*. Viking.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature 2015 521:7553*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lee, H., & Cho, C.-H. (2020). Digital advertising: present and future prospects. *International Journal of Advertising*, *39*(3), 332–341. https://doi.org/10.1080/02650487.2019.1642015

Liao, S. M. (2020). *Ethics of artificial intelligence*. https://books.google.com/books/about/Ethics_of_Artificial_Intelligence.html?id=2ST3DwAAQBAJ

Liao, V., & Varshney, K. (2021). *Human-Centered Explainable AI (XAI): From Algorithms to User Experiences*.

Licklider, J. C. R. (1960). *Man-Computer Symbiosis*.

Lipton, Z. C. (2018). The Mythos of Model Interpretability. *Queue*, *16*(3), 31–57. https://doi.org/10.1145/3236386.3241340

Litsios, G., Sims, C. A., Wüest, R. O., Pearman, P. B., Zimmermann, N. E., & Salamin, N. (2012). Mutualism with sea anemones triggered the adaptive radiation of clownfishes. *BMC Evolutionary Biology*, *12*(1), 212. https://doi.org/10.1186/1471-2148-12-212

Margulis, L. (1998a). *Symbiotic planet : a new look at evolution*. Basic Books.

Margulis, L. (1998b). *Symbiotic planet : a new look at evolution*. Basic Books.

Margulis, L., & Fester, R. (1991). *Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis*. MIT Press.

Margulis, L., & Margulis, U. M. A. M. L. (1993). *Symbiosis in Cell Evolution: Microbial Communities in the Archean and Proterozoic Eons*. Freeman.
Margulis, L., & Sagan, D. (2023). *Microcosmos: Four Billion Years of Microbial Evolution*. University of California Press.
Maturana, H. R., & Varela, F. J. (1987). *The Tree of Knowledge : The Biological Roots of Human Understanding*.
Maturana, H. R., & Varela, F. J. (2012). *Autopoiesis and Cognition: The Realization of the Living*.
McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Magazine*, *27*(4), 12. https://doi.org/10.1609/AIMAG.V27I4.1904
McCorduck, P., & Cfe, C. (2004). *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. CRC Press.
McElligott, A. G., Maggini, I., Hunziker, L., & König, B. (2004). Interactions between red-billed oxpeckers and black rhinos in captivity. *Zoo Biology*, *23*(4), 347–354. https://doi.org/10.1002/zoo.20013
Messick, G. A. (1998). Diseases, Parasites, and Symbionts of Blue Crabs (Callinectes Sapidus) Dredged From Chesapeake Bay. *Journal of Crustacean Biology*, *18*(3), 533–548. https://doi.org/10.1163/193724098X00368
Michel, R. (2007). *Design research now : essays and selected projects*. Birkhäuser.
Mitchell, T. (1997a). Machine Learning. In *McGraw Hill*. McGraw Hill. http://www.cs.cmu.edu/~tom/mlbook.html
Mitchell, T. (1997b). Machine Learning. In *McGraw Hill*. McGraw Hill.
Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 205395171667967. https://doi.org/10.1177/2053951716679679
Moor, J. H. (2006). *IEEE INTELLIGENT SYSTEMS The Nature, Importance, and Difficulty of Machine Ethics*. www.computer.org/intelligent
Moran, N. A. (2007a). Symbiosis as an adaptive process and source of phenotypic complexity. *Proceedings of the National Academy of Sciences*, *104*(suppl 1), 8627–8633. https://doi.org/10.1073/PNAS.0611659104
Moran, N. A. (2007b). Symbiosis as an adaptive process and source of phenotypic complexity. *Proceedings of the National Academy of Sciences*, *104*(suppl 1), 8627–8633. https://doi.org/10.1073/PNAS.0611659104
Morton, T. (2012). *The Ecological Thought*. Harvard University Press.
Murphy, R. R., & Woods, D. D. (2009). Beyond asimov: The three laws of responsible robotics. *IEEE Intelligent Systems*, *24*(4), 14–20. https://doi.org/10.1109/MIS.2009.69
Needham, J. (1965). *Science and Civilisation in China, Part 2, Mechanical Engineering*. Cambridge University Press.
Nicolelis, M. (2011). *Beyond Boundaries: The New Neuroscience of Connecting Brains with Machines---and How It Will Change Our Lives*. Henry Holt and Company.
Noble, S. U. (2018). *Algorithms of oppression : how search engines reinforce racism*. https://books.google.com/books/about/Algorithms_of_Oppression.html?id=g8OSDgAAQBAJ

Norman, B. M., Reynolds, S. D., & Morgan, D. L. (2021). Three-way symbiotic relationships in whale sharks. *Pacific Conservation Biology*, *28*(1), 80–83. https://doi.org/10.1071/PC20043
Norman, D. (2013). *The Design of Everyday Things: Revised and Expanded Edition*. Basic Books.
Olds, A. D., Pitt, K. A., Maxwell, P. S., & Connolly, R. M. (2012). Synergistic effects of reserves and connectivity on ecological resilience. *Journal of Applied Ecology*, *49*(6), 1195–1203. https://doi.org/10.1111/jpe.12002
O'Neil, C. (2016a). *Weapons of math destruction : how big data increases inequality and threatens democracy*. https://books.google.com/books/about/Weapons_of_Math_Destruction.html?id=60n0DAAAQBAJ
O'Neil, C. (2016b). *Weapons of math destruction : how big data increases inequality and threatens democracy*.
Paracer, Surindar., Ahmadjian, Vernon., & Ahmadjian, Vernon. (2000). *Symbiosis : an introduction to biological associations*. 291.
Pariser, E. (2011). *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*. Penguin Publishing Group. https://books.google.ch/books?id=wcalrOI1YbQC
PASQUALE, F. (2015). *The Black Box Society*. Harvard University Press.
Pastaltzidis, I., Dimitriou, N., Quezada-Tavarez, K., Aidinlis, S., Marquenie, T., Gurzawska, A., & Tzovaras, D. (2022). Data augmentation for fairness-aware machine learning. *2022 ACM Conference on Fairness, Accountability, and Transparency*, 2302–2314. https://doi.org/10.1145/3531146.3534644
Paszkowski, U. (2006). Mutualism and parasitism: the yin and yang of plant symbioses. *Current Opinion in Plant Biology*, *9*(4), 364–370. https://doi.org/10.1016/J.PBI.2006.05.008
Polanyi, M. (1958). *Personal Knowledge*.
Polanyi, M. (1966). *The Tacit Dimension*.
Potts, S. G., Vulliamy, B., Dafni, A., Ne'eman, G., & Willmer, P. (2003). LINKING BEES AND FLOWERS: HOW DO FLORAL COMMUNITIES STRUCTURE POLLINATOR COMMUNITIES? *Ecology*, *84*(10), 2628–2642. https://doi.org/10.1890/02-0136
Poulin, R. (2007). *Evolutionary ecology of parasites*. 332.
Pratte, Z. A., Patin, N. V., McWhirt, M. E., Caughman, A. M., Parris, D. J., & Stewart, F. J. (2018). Association with a sea anemone alters the skin microbiome of clownfish. *Coral Reefs*, *37*(4), 1119–1125. https://doi.org/10.1007/s00338-018-01750-z
Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2016). PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *CoRR*, *abs/1612.0*.
Quammen, D. (2019). *The Tangled Tree: A Radical New History of Life*. Simon \& Schuster.
Reason, P., & Bradbury, H. (2008). *SAGE Handbook of Action Research*.
Redmon, J., & Farhadi, A. (2016). *YOLO9000: Better, Faster, Stronger*.
Redström, J. (2017). *Making design theory*. 171.
Rifkin, J. (2014). *The Zero Marginal Cost Society: The Internet of Things, the Collaborative Commons, and the Eclipse of Capitalism*. St. Martin's Publishing Group.
Riskin, J. (2016). *The Restless Clock: A History of the Centuries-Long Argument Over What Makes Living Things Tick*. University of Chicago Press.
Rosheim, M. E. (1994). *Robot evolution : the development of anthrobotics*. 423.

Russell, S., & Norvig, P. (2020a). *Artificial Intelligence: A Modern Approach, 4th Edition*. https://www.pearson.com/us/higher-education/program/Russell-Artificial-Intelligence-A-Modern-Approach-4th-Edition/PGM1263338.html
Russell, S., & Norvig, P. (2020b). *Artificial Intelligence: A Modern Approach, 4th Edition*.
Ryan, F. (2002). *Darwin's Blind Spot: Evolution Beyond Natural Selection*. Houghton Mifflin.
Sapp, J. (1994). *Evolution by association : a history of symbiosis*. 255.
Schoenherr, J. R., Abbas, R., Michael, K., Rivas, P., & Anderson, T. D. (2023). Designing AI Using a Human-Centered Approach: Explainability and Accuracy Toward Trustworthiness. *IEEE Transactions on Technology and Society*, *4*(1), 9–23. https://doi.org/10.1109/TTS.2023.3257627
Schön, D. A. (2017a). *The reflective practitioner : how professionals think in action*.
Schön, D. A. (2017b). *The reflective practitioner : how professionals think in action*. https://books.google.com/books/about/The_Reflective_Practitioner.html?id=OT9BDgAAQBAJ
Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and Abstraction in Sociotechnical Systems. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 59–68. https://doi.org/10.1145/3287560.3287598
Shanahan, Murray. (2015). *The technological singularity*. https://books.google.com/books/about/The_Technological_Singularity.html?id=rAxZCgAAQBAJ
Sharp, C. (2019). Action learning and action research: genres and approaches. *Https://Doi.Org/10.1080/14767333.2019.1655974*, *16*(3), 317–321. https://doi.org/10.1080/14767333.2019.1655974
Shields, J. D., Williams, J. D., & Boyko, C. B. (2015). Parasites and diseases of Brachyura. In *Treatise on Zoology - Anatomy, Taxonomy, Biology. The Crustacea, Volume 9 Part C (2 vols)* (pp. 639–774). BRILL. https://doi.org/10.1163/9789004190832_015
Shvartzshnaider, Y., Apthorpe, N., Feamster, N., & Nissenbaum, H. (2019). Going against the (Appropriate) Flow: A Contextual Integrity Approach to Privacy Policy Analysis. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, *7*(1 SE-Technical Papers), 162–170. https://doi.org/10.1609/hcomp.v7i1.5266
Smith, D. C., Douglas, A. E., & Douglas, A. E. (1987). *The Biology of Symbiosis*. E. Arnold.
Stanley, K. O., & Miikkulainen, R. (2012a). *Efficient Reinforcement Learning through Evolving Neural Network Topologies*.
Stanley, K. O., & Miikkulainen, R. (2012b). *Evolving Neural Networks through Augmenting Topologies*.
Sutton, R. S., & Barto, A. G. (2015). *Reinforcement Learning: An Introduction*.
Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *CoRR*, *abs/1905.1*.
Tashakkori, A., & Teddlie, C. (2003). *Handbook of Mixed Methods in Social & Behavioral Research*. SAGE Publications. https://books.google.ch/books?id=F8BFOM8DCKoC
Tegmark, M. (2017a). *Life 3.0 : being human in the age of artificial intelligence*.
Tegmark, M. (2017b). *Life 3.0 : being human in the age of artificial intelligence*. https://books.google.com/books/about/Life_3_0.html?id=3_otDwAAQBAJ

Tene, O., & Polonetsky, J. (2012). Big Data for All: Privacy and User Control in the Age of Analytics. *Northwestern Journal of Technology and Intellectual Property*, *11*.
Tesauro, G. (1995). *Temporal Difference Learning and TD-Gammon*.
Thompson, E. (2007). *Mind in Life*. Harvard University Press. https://books.google.ch/books?id=OVGna4ZEpWwC
Thompson, J. N. (2005). *The Geographic Mosaic of Coevolution*. University of Chicago Press.
Tjondronegoro, D., Yuwono, E., Richards, B., Green, D., & Hatakka, S. (2022). Responsible AI Implementation: A Human-centered Framework for Accelerating the Innovation Process. *ArXiv E-Prints*, arXiv:2209.07076. https://doi.org/10.48550/arXiv.2209.07076
Truitt, E. R. (2015). *Medieval Robots: Mechanism, Magic, Nature, and Art*. University of Pennsylvania Press, Incorporated.
Tsing, A. L. (2017). *The Mushroom at the End of the World: On the Possibility of Life in Capitalist Ruins*. Princeton University Press.
Turing, A. M. (1950). *COMPUTING MACHINERY AND INTELLIGENCE*.
Turkle, S. (2017). *Alone Together: Why We Expect More from Technology and Less from Each Other*. Basic Books.
van der Oord, A., Kalchbrenner, N., & Kavukcuoglu, K. (2016). *Pixel Recurrent Neural Networks*.
von Bertalanffy, L., Hofkirchner, W., & Rousseau, D. (2015). *General System Theory: Foundations, Development, Applications*. George Braziller, Incorporated.
Wachter, S., Mittelstadt, B., & Russell, C. (2021). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. *Computer Law & Security Review*, *41*, 105567. https://doi.org/10.1016/j.clsr.2021.105567
Wallach, W., & Allen, C. (2009). Moral Machines: Teaching Robots Right from Wrong. *Moral Machines: Teaching Robots Right from Wrong*, 1–288. https://doi.org/10.1093/ACPROF:OSO/9780195374049.001.0001
Wernegreen, J. J. (2012). Endosymbiosis. *CURBIO*, *22*, R555–R561. https://doi.org/10.1016/j.cub.2012.06.010
Whitehead, M. R., & Peakall, R. (2009). Integrating floral scent, pollination ecology and population genetics. *Functional Ecology*, *23*(5), 863–874. https://doi.org/10.1111/j.1365-2435.2009.01620.x
Wiener, N. (1965). *Cybernetics, Second Edition: or the Control and Communication in the Animal and the Machine*. 212.
Wiener, N., Hill, D., & Mitter, S. (2019). *Cybernetics or Control and Communication in the Animal and the Machine, Reissue of the 1961 second edition*. MIT Press.
Willmer, P. (2011). *Pollination and Floral Ecology*. Princeton University Press. https://doi.org/10.1515/9781400838943
Wilson, E. O. (2000). *Sociobiology : the new synthesis*. 697.
Wilson, E. O. (2001). *The Diversity of Life*. Penguin Books Limited. https://books.google.ch/books?id=VS7GeNorZE4C
Wood, G. (2002). *Living Dolls: A Magical History of the Quest for Mechanical Life*. Faber \& Faber.
Xu, Y., Shi, W., Arredondo-Galeana, A., Mei, L., & Demirel, Y. K. (2021). Understanding of remora's "hitchhiking" behaviour from a hydrodynamic point of view. *Scientific Reports*, *11*(1), 14837. https://doi.org/10.1038/s41598-021-94342-x

Zhang, R., Dong, S., & Liu, J. (2019). *Invisible Steganography via Generative Adversarial Networks **.
Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2020). *Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks*.
Zuboff, S. (2018). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (1st ed.).
Zuboff, S. (2019). Surveillance Capitalism and the Challenge of Collective Action. *New Labor Forum*, *28*(1), 10–29. https://doi.org/10.1177/1095796018819461

Appendices

Appendix A – A series of small Machine Learning projects

This appendix provides a link to the code of a series of small machine learning projects from the research practice *Journey of the Machine Learning Engineer,* located in Chapter 4 of this PhD.

Link: Series of Machine Learning Projects

Appendix B – Character Recurrent Neural Network (charRNN) project

This appendix provides a link to the code behind a character recurrent neural network project from the research practice *Journey of the Machine Learning Engineer,* located in Chapter 4 of this PhD.

Link: CharRNN Project

Appendix C – YOLO Object Detection with OpenCV

This appendix provides a link to the code behind a YOLO object detection project from the research practice *Journey of the Machine Learning Engineer,* located in Chapter 4 of this PhD*.*

Link: YOLO Object Detection with OpenCV

Appendix D – Generative Adversarial Network project

This appendix provides a link to code behind a generative adversarial network project from the research practice *Journey of the Machine Learning Engineer*, located in Chapter 4 of this PhD.

Link: Generative Adversarial Network Project

Appendix E – Video of Neuroevolution of Augmenting Topologies (NEAT) genetic algorithm training in a Comix Zone environment

This appendix provides a link to a video demonstrating a Neuroevolution of Augmenting Topologies (NEAT) genetic algorithm project training in a Comix Zone environment. This practice belongs to the research practice *Journey of the Machine Learning Engineer,* located in Chapter 4 of this PhD.

Link: NEAT genetic algorithm training in a Comix Zone environment

Appendix F – Supplementary material for classifying the types of symbiotic relationships

This appendix provides a link to three supplementary materials relating to the research project *Introduction to a Symbiotic Human AI Approach*, located in Chapter 5 of this PhD. The goal of this supplementary material was to guide the researcher during the interviews and have a place to quickly take note of the preferences of the people being interviewed.

Link No.01: Types of Symbiotic Relationships

Link No.02: Symbiosis in Machine Intelligence - Machine Intelligence Perspective

Link No. 03: Symbiosis in Machine Intelligence - Human Perspective

Appendix G – Code for the Open Repository

This appendix provides a link to the entire code used in the development of the Open Repository previously mentioned in the PhD. The repository was created as an attempt for disseminating and teaching non-technical researchers about Artificial Intelligence implementation.

Link: Code for Open Repository for Symbiotic Relationships