

Harm to the Environment and Non-human Animals: Can Care Ethics Re-Think AI?

by

Ceren Polat

A Dissertation Submitted to the
Graduate School of Social Sciences and Humanities
in Partial Fulfillment of the Requirements for
the Degree of
Master of Arts

in

Philosophy



April 24, 2025

Harm to the Environment and Non-human Animals: Can Care Ethics Re-Think AI?

Koç University

Graduate School of Social Sciences and Humanities

This is to certify that I have examined this copy of a master's thesis by

Ceren Polat

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

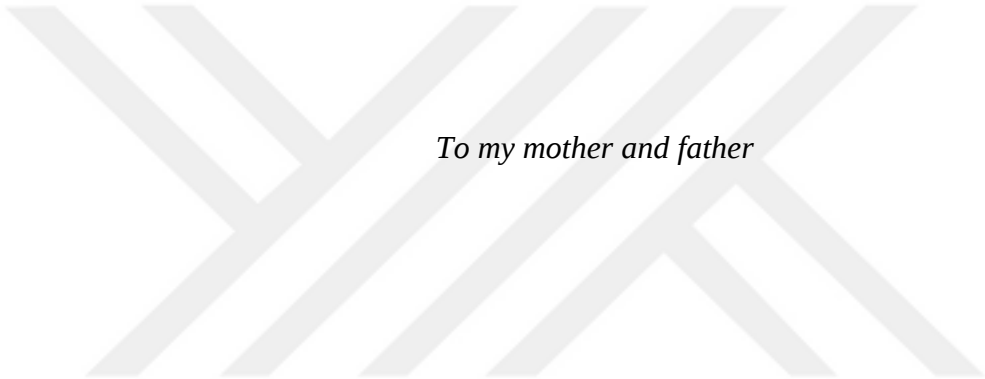
Committee Members:

Prof. Zeynep Direk (Advisor)

Assoc. Prof. İpek Azime Çelik Rappas

Assist. Prof. Öznür Karakaş

Date: 24.04.2025



To my mother and father

ABSTRACT

Harm to the Environment and Non-human Animals:

Can Care Ethics Re-Think AI?

Ceren Polat

Master of Arts in Philosophy

April 24, 2025

The integration of artificial intelligence (AI) into various domains raises ethical concerns about its impact on environmental sustainability and animal welfare. Recent studies highlight AI systems, programmed to prioritize profit-making, can exacerbate ecological harm and animal wellbeing. Therefore, this thesis aims to argue for potential ways to approach the anthropocentric tendencies of the development of AI's decision-making from an ethical standpoint. I explore how feminist care ethics, emphasizing interdependence, context, and vulnerability, can provide an ethical framework for AI design that mitigates harm rather than reinforcing anthropocentric biases. Building on Gruen's (2015) concept of entangled empathy, I argue that humans are already in an entangled relationship with non-humans and AI should also be understood as part of relational networks that include animals and the environment. Broadening care ethics to include AI raises the question whether AI can care as humans. Building on the social-relational approach for moral status proposed by Mark Coeckelbergh and David Gunkel, this thesis argues that moral agency as well as ability to care and empathy need not rely on ontological properties but should be grounded in appearance of AI within relational contexts. Through this lens, AI systems may ethically engage with the world not because of what they are, but because of how they interact, respond to needs, form social bonds, and participate in practices of care. This thesis aims to argue that integrating care ethics into AI design fosters a more just and ecologically responsible approach, encouraging interdisciplinary dialogue on the ethical interactions across diverse entities.

Keywords: AI ethics, environment, non-human animals, anthropocentrism, care ethics, entangled empathy, social-relationalism

ÖZETÇE

Çevreye ve İnsan Dışı Hayvanlara Verilen Zarar: Bakım Etiği Yapay Zekâyı

Yeniden Düşünebilir mi?

Ceren Polat

Felsefe, Yüksek Lisans

24 Nisan 2025

Yapay zekânın (YZ) çeşitli alanlara entegrasyonu, çevresel sürdürülebilirlik ve hayvan refahı üzerindeki etkilerine dair etik kaygıları gündeme getirmektedir. Son araştırmalar, kâr odaklı programlanan YZ sistemlerinin ekolojik zararı artırabileceğini ve hayvan refahını olumsuz etkileyebileceğini ortaya koymaktadır. Bu nedenle, bu tez, YZ'nin karar alma süreçlerinin gelişiminde ortaya çıkan antroposantrik eğilimlere etik bir bakışla nasıl yaklaşılabilirine dair olası yolları tartışmayı amaçlamaktadır. Karşılıklı bağımlılık, bağlamsallık ve kırılabilirliğe vurgu yapan feminist bakım etiğinin, antroposantrik önyargıları yeniden üretmek yerine zararı azaltacak bir etik çerçeve sunabileceği ileri sürülmektedir. Gruen'ün (2015) "iç içe geçmiş empati" (entangled empathy) kavramına dayanarak, insanların hâlihazırda insan olmayan varlıklarla iç içe geçmiş bir ilişkide olduğu ve YZ'nin de hayvanlar ve çevreyi kapsayan ilişkisel ağların bir parçası olarak anlaşılması gerektiği savunulmaktadır. Bakım etiğinin YZ'yi kapsayacak şekilde genişletilmesi, YZ'nin insanlar gibi "bakım gösterip gösteremeyeceği" sorusunu gündeme getirir. Mark Coeckelbergh ve David Gunkel'in önerdiği sosyal-ilişkisel ahlaki statü yaklaşımından yola çıkan bu tez, ahlaki fail olmanın, bakım verme ve empati kurma yetisinin ontolojik özelliklere dayanması gerektiğini; bunun yerine YZ'nin ilişkisel bağlamlardaki görünüşüne dayalı olarak ele alınması gerektiğini ileri sürmektedir. Bu bakış açısıyla, YZ sistemleri dünyayla etik ilişkiler kurabilir; yalnızca ne olduklarından dolayı değil, nasıl etkileşim kurdukları, ihtiyaçlara nasıl yanıt verdikleri, sosyal bağlar oluşturup bakım uygulamalarına nasıl katıldıkları önemlidir. Bu tez, bakım etiğinin YZ tasarımına entegre edilmesinin daha adil ve ekolojik açıdan sorumlu bir yaklaşım sunduğunu ve çeşitli varlıklar arasındaki etik etkileşimlere dair disiplinlerarası diyalogu teşvik ettiğini savunmaktadır.

Anahtar kelimeler: Yapay zeka etiği, çevre, insan olmayan hayvanlar, insan merkezilik, bakım etiği, iç içe geçmiş empati, sosyal-ilişkiselcilik

ACKNOWLEDGEMENTS

I would like to begin by expressing my heartfelt gratitude to my thesis advisor, Prof. Zeynep Direk, for her unwavering support throughout this journey. I am especially thankful that she encouraged me to explore AI ethics—an area that has already opened up exciting opportunities for the next chapter of my life. Her guidance has extended far beyond my master’s studies; she has supported me since my freshman year in 2018. I feel incredibly fortunate to have learned from and worked with an academic who has always believed in me unconditionally, who recognized my vulnerabilities, and who genuinely celebrated my achievements.

I am also grateful to all the faculty members at Koç University. I have had the privilege of learning a diverse range of perspectives and ideas in their courses, all of which have shaped who I am today. Many thanks to Prof. İlhan İnan, whose passion for philosophy inspired me to pursue an academic path. I would also like to express my gratitude to Dr. Seyed Zacharus Gudmunsen, who, through his expertise in AI ethics, recommended valuable readings, ideas, and scholars relevant to my work. Many thanks to Assist. Prof. Öznür Karakaş, one of the jury members of my thesis defense, for her valuable comments and suggestions regarding my thesis. Special thanks to Assoc. Prof. İpek Azime Çelik Rappas, whose class inspired me to undertake this project. I truly appreciate her support and encouragement throughout this journey.

To my wonderful family—thank you for making me feel like the luckiest daughter in the world. I am endlessly thankful to my mother, Yasemin, and my father, Yılmaz, for their unconditional love and support in every decision I have made. I would not be who I am without them. I also want to thank the sweetest grandmother, Hüsniye, who has always proud of my accomplishments.

The people I have met at Koç University have left a meaningful impact on my life. This journey would have been much harder without the support of my fellow graduate students. I want to thank Burak and İbrahim for the many philosophical (and non-philosophical) conversations we have shared. I am especially grateful to Berkan for his steady encouragement, and to my closest friends, Ashi and Victoria, for always valuing my work and standing by me. I am truly grateful for everything they have done.

Lastly, I want to thank Koç University for providing me with experiences and memories that I will cherish forever.

TABLE OF CONTENTS

List of Tables	viii
Abbreviations.....	ix
Chapter 1: Introduction.....	1
Chapter 2: Arguments Concerning the Anthropocentric Bias of AI	7
2.1 The Use of AI and the Environment	7
2.2 The Use of AI and Non-human Animals	8
2.3 The Negative Effects of AI on the Environment	9
2.4 The Negative Effects of AI on Non-human Animals	13
2.5 The Need for Ethically Approaching the Anthropocentric Bias of AI	15
Chapter 3: Feminist Ethics of Care as a Solution to The Anthropocentric Bias of AI..	17
3.1 History of Care Ethics and Influential Care Ethicists.....	17
3.2 Applied Care Ethics	21
3.2.1 Feminist care tradition in environmental and animal ethics	22
3.3 Overcoming the moral distance between humans and non-humans: Lori Gruen's entangled empathy	26
Chapter 4: Reimagining Care Ethics: The Role of AI in Ethical Relationships.....	33
4.1 Is AI Capable of Caring?.....	33
4.1.1 What is vs How It Appears	38
4.1.2 AI as an Extension of Humans vs entangled?	42
4.2 Objection: Is AI's simulating moral acts a form of deception?	45
Chapter 5: Conclusion.....	47
Bibliography.....	5

Error! Bookmark not defined.

LIST OF TABLES

Table 1: A Practical framework for considering animal harm from AI.	1
--	---

Error! Bookmark not defined.



ABBREVIATIONS

AI	Artificial Intelligence
LoA	Level of Abstraction



CHAPTER 1:

INTRODUCTION

When Artificial Intelligence (AI) was first encountered, it was a technology that potential users were treating skeptically, just like many other newly emerged technological developments in the past. Despite such skepticism about AI technologies, many of us left our concerns behind and allowed it to become part of our daily lives. Whether we allowed AI to gain popularity or we find ourselves in a sense forced to use it due to fomo (fear of missing out), for instance, is still open to debate. In any case, this increased use of AI is due to its rapid growth and the many benefits it provides. AI helps us while doing research, it checks grammar mistakes, translates, generates images, drives our cars, assists us in medical diagnoses, and even writes poems.

Various forms of AI ranging from autonomous vehicles and chatbots to medical assistant tools and artificial companions, AI technologies have quickly differed from other technological tools that we already have, and they have become central features of our social, economic, and political infrastructures. Despite the fact that this integration of AI in various parts of our lives is bringing too many benefits, it has also raised serious ethical concerns. One thing that we should pay close attention to is that, while helping us in various activities, AI's interaction with humans and their environment as well as the decisions it makes are not free from implications that need to be ethically approached.

How AI works and makes those decisions is an important yet not totally mysterious question. One of the first steps for AI to work is collecting a vast amount of data and algorithms. After that collection, AI makes decisions about what to do with the collected data. "The AI sorts and deciphers the data using patterns it has been programmed to learn until it recognizes similar patterns in the data that is being filtered into the system" (How Does AI Work? Basics to Know | Coursera, n.d.). Perhaps it would be appropriate to draw an analogy between AI's learning process and a child's learning process. During their cognitive ability development, children collect data from other people, especially their parents, as well as from their environment (Council, 2015). A child uses this data when making decisions in his or her life based on the similar patterns that are detected. What is crucial in both of them is how effective the features of developers are since it refers to a transfer of data from the developer (or parents) to the AI (or children). Indeed, this is the

very reason why we are dealing with algorithmic bias today. If the data AI collected from the developers is already biased, whether racist, sexist, or anthropocentric, that means AI reflects the biases of the ones who program it.

An example of this point can be found in the popular Netflix documentary “Algorithm Bias”. It starts by illustrating how Joy Buolamwini, a researcher from MIT, discovered that the facial recognition technology she was testing could not recognize her black skin (Trenholm, 2021). Such depictions of biased approaches to AI technologies inform not just AI developers but also everyone who encountered this news in the media. Accordingly, newly occurring developments in AI programming started to aim for coping with biases that concern certain groups of people. Another example was observed in Amazon’s recruitment process, in which gendered algorithmic bias supported the idea of underrepresentation of women in the tech industry by preferring male employees over females when checking their resumes (Chang, 2023). Besides working based on the codes of developers, thanks to recent developments, progress has been made in designing AI as embodied. This allows some robots to become social and learn by interaction with the environment via their body (Coeckelbergh, 2010a, p. 2).

The points above concerning AI’s biases in decision-making indicate that its rapid growth is not free from ethical problems. It has brought significant ethical issues that need to be dealt with urgently to the agenda of researchers from many different fields. Recent studies of moral philosophers concerning AI have paid attention to various topics such as algorithm bias (Stinson, 2022), privacy (Hartmann et al., 2023), human dignity (Pfeifer-Chomiczewska, 2023), autonomy (Formosa, 2021), and human relations with machines (Coeckelbergh, 2010a). Perhaps what led scholars to criticize and work on AI’s effects on humans is the fact that the negative impacts became publicly known thanks to the media. This enabled people who face several problems in regard to AI uses, to have their voices being heard. That can be the reason why there is an effort to make certain changes in AI programming both by scholars and AI developers in order to decrease the number of biased decisions that concern human beings.

Thanks to recent developments in AI technologies, there are now applications that either directly or indirectly affect the environment and non-human animals. Very recently, researchers in AI ethics pointed out their concerns about the fact that AI is anthropocentrically biased, and since this poses challenges to the lives of many animals, there is an urge to include non-human animals in the agenda of AI ethics. The reason why

non-human animals are included will be better understood when I introduce the phenomenological notion of entanglement in the following chapters.

Although not much in number, scholars who started working on the impact of AI systems on environment have investigated issues such as sustainability, AI's effect on climate change, and its energy footprint (Moyano-Fernández & Rueda, 2023; Nordgren (2023) ; Baum & Owe, 2021). Besides the impacts on humans, AI systems also affect non-human animals in terms of, for instance, its energy use and e-waste destroying animal habitat, its anthropocentric bias, animal's contact with machines instead of their human caretakers, roadkill due to self-driving cars and AI-enabled harmful treatment and confinement in livestock farming (Coghlan & Parker, 2023; Singer & Tse, 2022).

For instance, a recent study conducted by researchers at MIT aims to determine how much energy is used to train certain popular large AI models such as Transformer, BERT, ELMo and GPT-2. The findings show that training can emit over 626,000 pounds of carbon dioxide, or the same amount as almost 300 round-trip flights from New York to San Francisco. This is approximately five times the average car's lifetime emissions (Hao, 2019). Similarly, Patterson and colleagues (2021) have also studied the substantial level of the carbon footprint of training other famous models such as T5, Meena, GPT-3, GShard, Switch, and Transformer, and showed that deep learning points out to a huge computing and energy demand which induce significant carbon emissions, accordingly, aggravating climate change.

Another serious environmental challenge that is posed by the use of AI technologies is their e-waste. Electronic waste, or e-waste, is the term used to describe electrical appliances that are abandoned, and every year, “around 50 to 60 million tons of e-waste are generated” (Wu, 2023). E-waste contains toxic chemicals such as lead, mercury, and cadmium, which unfortunately have an impact on contaminating the soil and water supplies and put human health and the environment at risk (Kanungo, 2023). When we take AI's rapid growth into account, it would not be odd to expect that those effects will increase in huge numbers.

The points above show us how the use of AI technologies is creating results that concern not just human health but also the environment in general and the ones who live in it. However, the important question that emerges here is whether the effects of AI are greater than those of any other significantly used technologies, such as our smart phones, televisions, cars, and so on. What makes it so special that we find ourselves dealing with

those problems? We may offer three potential responses to such a question: The first one is about the fact that AI is newly developing, which may indicate that there is not enough research about its negative impacts. Even if there are some studies, they may not be as comprehensive as they should be. This makes the whole topic shadowy, and perhaps there are much more significant effects than we think. Secondly, the fact that there is a possibility for AI to not have significantly different effects in terms of numbers when compared to other technologies does not necessarily make it less important. It is true that the use of our smart phones is much more frequent than our use of AI tools, but again, this does not mean that it will stay this way forever. AI might start being incorporated into any technology that we are already using, and accordingly, the effects will increase in number. Lastly, a much stronger argument than the first two is about what makes AI special when compared to other technologies. AI is intended to be as intelligent as humans. It is designed in a way that can imitate the cognitive abilities of humans by learning, perceiving, and thinking as we do. A significant implication of this fact is that the normative and ethical values that concern humans can also start to concern artificially intelligent beings. Ethical questions about autonomy, responsibility, and vulnerability have the potential to extend to include AI besides humans. The reason lies in their evolving potential to engage in moral decision-making, which is something lacking in other technologies. We probably have never found ourselves too concerned with whether our dining table in living room is responsible for making our cat pissed off just by standing there or whether massive media such as billboards in New York Times Square feel vulnerable since their huge bodies are seen by anyone who walks by. But similar questions can occur if the subject that we are dealing with is an artificially intelligent being.

In this thesis, I argue that feminist care ethics, a tradition that emphasizes relationality, vulnerability, and context-sensitive approach to ethics, offers a powerful theoretical lens for rethinking the ethical role of AI in relation to non-human animals and the environment. While recent efforts have spent to address human-centered biases (e.g., racism and sexism) in AI, there has been considerably less attention to anthropocentrism, biases that affect the environment and non-human animals. It might be because of the tendency to privilege human needs and experiences and treat the non-human world to the extent that they benefit human beings. This oversight is particularly problematic given the increasing use of AI in areas like agriculture, livestock farming, and environmental

monitoring, where decisions made by AI systems have potential negative effects on animal welfare and ecological sustainability.

This thesis begins with the recognition that AI, like all technologies, is not neutral. It cannot be thought of independent from the values, assumptions, and priorities of those who design, program, and deploy it. When these values are affected by economic efficiency, productivity, and control concerns, AI technologies risk reproducing and even amplifying harm to the most vulnerable beings in our world.

In my thesis, I explore whether AI that is traditionally seen as lacking the necessary conditions for moral agency (like consciousness, emotions, responsibility) can be included within the scope of care-based ethical relations. By drawing on literature on feminist care ethics, social-relational approach to morality, and AI philosophy, I propose a shift from ontological property-based approach to moral status to appearance-based and interaction-based accounts of moral status. This will allow us to ask not “What is AI?” but rather “How does AI appear and act within relational contexts of care?” I also explore how AI might be designed and programmed to engage in practices of care.

This thesis is structured as follows: Chapter 2 introduces the arguments concerning the anthropocentric bias of AI. It first explains the reason why and how AI technologies are used in industries and areas related to the environment and non-human animals. This allows us to reach the scope of the problem. Then I discuss arguments in AI literature that focus on how AI can harm the environment and non-human animals if it is anthropocentrically biased. These arguments demonstrate the need for ethically approaching the anthropocentric bias of AI and offering potential solutions.

Chapter 3 discusses feminist ethics of care as a solution to the anthropocentric bias of AI and its potential harm. It first explains the history of care ethics and influential care ethicists then move on to applied care ethics where they see the importance of the emphasis on relational moral reasoning especially in environmental and animal ethics. The highlighted concepts and features of care ethics become an appropriate framework mitigating bias towards non-humans. By relying on secondary literature, I also explore how care ethics have been extended to include animals and the environment, focusing on the concept of entangled empathy.

Chapter 4 argues for broadening care ethics to include AI so that AI can also be recognized as caring and empathizing, which offers a solution to its potential harm. To discuss whether AI is capable of caring, this chapter first investigates the question of AI’s

moral agency, comparing standard and functionalist accounts. It argues that both descriptively and normatively speaking, social-relational account of moral status is advantageous to argue for AI's place in social and ethical relations. Based on this approach, without really possessing ontological properties required for being moral agent and for caring, AI's appearance as a moral agent and caring agent in social and ethical contexts is sufficient for us to argue that AI cares. In this sense, AI can be entangled with others or just an extension of humans that can simulate moral action. But rather than what really is the case, its appearance matters.

Lastly, chapter 5 concludes the thesis with a summary.



CHAPTER 2:

ARGUMENTS CONCERNING THE ANTHROPOCENTRIC BIAS OF AI

This thesis offers a way to ethically approach the problem concerning the anthropocentric bias of AI's decision-making and the potential effects on the environment and non-human animals resulting from these biases. However, before proposing any theoretical lens to approach this problem and delving into a discussion on further questions concerning AI's moral status, it is important to deepen the reader's understanding of the use of AI that affects the environment and non-human animals and how exactly AI's biased decision-making might put them in danger. Hence, in this chapter, I explain how, in what ways, and to what extent an artificially intelligent being can harm the welfare of non-human animals and the sustainability of the environment. Accordingly, understanding the potential harms will help highlight the necessity of addressing the negative impacts of AI from an ethical standpoint.

2.1 The Use of AI and the Environment

Thanks to progress made in developing AI technologies, the areas of use have significantly increased and there are now applications that impact the environment through profound transformations. One of the main reasons why AI technologies started to be used in industries working with the environment is the potential benefits provided by AI with the aim of making people's life easier. This rapid growth and proliferation of AI in relation to the environment has caught the attention of researchers from many different fields.

Incorporating AI in agriculture results in the transformation of early techniques with more efficient and smart ones. For instance, AI can be helpful when detecting diseases related to plants and pests. One study shows that thanks to deep learning by being provided with apple black rot images, AI is capable of detecting the severity of the disease 90.4% accurately (Wang et al., 2017). Another way of using AI in agriculture is making the spraying of pesticides and fertilizers automated. AI can detect the target areas and determine the amount to be sprayed and end up being efficient in terms of time and human labor (Rizzoli, 2021). AI using is also offered as a solution to the overuse of herbicides. Via machine learning and computer vision, AI-driven automated weed control systems

can distinguish crops from weeds, which in turn helps to lower the use of herbicides (Yun, n.d.). Therefore, the growing number of employments of AI in agriculture is due to the potential benefits in terms of efficiency.

In addition to the applications of AI technologies in agriculture, applications concerning water management also benefit from the use of AI. By being trained with related data, AI can assist in the determination of quality, usage, and availability of water resources. For instance, Ocean Cleanup, an organization based in the Netherlands that aims at saving oceans from plastic pollution, integrated AI technologies into their system to detect and collect plastic waste effectively (*Environmental Sustainability Through AI: 8 Conservation Strategies*, n.d.). A similar application can be found in waste management through AI tools that work with data concerning the collection and disposition of waste. As an example, Waste Robotics holds a mission to contribute to waste management by deploying artificially intelligent robots. Thanks to their AI-powered technologies, they handle waste such as construction demolition, recyclables, and metals (*Technology - Waste Robotics*, n.d.).

Perhaps, a more significant application of AI in relation to the environment lies in the promises about AI's help for combating climate change. There is a growing interest in the topic of how AI can facilitate actions dealing with the climate crisis, both by academics and media. As Cowls et al. (2021) stated, AI has the potential to both effectively fight the climate crisis and enhance our understanding of climate change (p. 287). Some applications include the use of machine learning techniques to forecast temperatures (Cifuentes et al., 2020), predicting extreme weather conditions such as heavy rain (Choi et al., 2018), and forecasting tropical instability waves (Zheng et al., 2020). Similarly, AI can assist in the development of renewable energy resources which are beneficial in terms of reducing greenhouse gas emissions and accordingly preventing the climate crisis from worsening. Instances of applications include “optimizing solar panel placement, predictive maintenance of wind turbines, and optimizing fermentation processes in biogas production” (Rusilowati et al., 2024). Such practices promise a responsible AI that does not overlook environmental sustainability and tackles with climate change which is one of the major global crises.

2.2 The Use of AI and Non-human Animals

The previous subsection explains the applications of AI technologies in relation to the environment as well as the potential benefits resulting from that. But recent studies also indicate that there are emerging fields of AI use concerning non-human animals that promise efficiency and productivity. The use of AI in livestock farming aims at increasing animal welfare by appropriately measuring animals' health status, their stress level, and detecting diseases and injuries. CattleEye, for instance, places security cameras to observe cows and then uses AI algorithms to analyze the data collected by observing cows (*CattleEye | Autonomous Livestock Monitoring - CattleEye, n.d.*). They aim at lowering the costs spent per cow as well as reducing the negative impacts of livestock farming on climate change in terms of carbon footprint. A similar deep learning algorithmic technology is used to track the feeding behavior of pigs (Chen et al., 2020). As a result, this application of AI promises to benefit farmers in terms of efficiently determining animals' feeding times and enhancing their health.

AI driven tracking technologies are also used with the hope of protecting biodiversity. Some instances of applications include the observation of wildlife, preventing poachers, and saving animals that are under the risk of extinction such as koalas and whales (Green, 2022). One specific example is the creation of a machine learning model to detect the song of humpback whales and accordingly determine where whales are accurately (Cattiau, 2018). Similarly, to protect and gain important data about bird species, AI powered-system BirdNet is able to recognize over 6,000 different bird species merely by their voice (Price, 2024).

2.3 The Negative Effects of AI on the Environment

The applications above indicate that the number of ways AI started to take part in relationships with non-human animals and the environment are rapidly growing. The reason is the promises of providing efficiency, assisting in combating the climate crisis, and contributing to human activities that aim at environmental protection. The growing use of AI in related fields appears to be positive however, there are still significant questions that remain: How ethical is AI in these practices? Who uses it? Who benefits as a result? Who is involved in the development of AI? Whose interests are prioritized? Whose values are transferred to these intelligent machines? Does AI serve the wellbeing of the planet as a whole or solely focus on benefiting humans?

In the light of the questions above, some scholars started working on the cost of using AI in relation to environmental issues. One major focus is on sustainable AI. As van Wynsberghe (2021) argues, the term sustainable AI must be understood in two ways: “AI *for* sustainability” and “sustainability *of* AI” (p. 213). While the first one refers to the use of AI to support actions for protecting environmental sustainability (such as some of the examples above), the latter refers to making the development and use of AI itself sustainable so that its environmental cost will be lowered. This highlights the reality behind the application of AI technologies and the marks they leave in our world.

To better understand the negative side of AI applications, it will be helpful to examine studies that provide empirical evidence. A recent study conducted by researchers at MIT aims to determine how much energy is used to train certain large AI models such as Transformer, BERT, ELMo and GPT-2. The findings show that training can emit over 626,000 pounds of carbon dioxide which is approximately five times the average car's lifetime emissions (Hao, 2019). Similarly, Patterson et al. (2021) have studied the carbon footprint of training other famous models such as T5, Meena, GPT-3, GShard, Switch, and Transformer, and showed that deep learning points out huge computing and energy demands that induce significant carbon emissions, accordingly, exacerbating the climate crisis. Besides their energy consumption, to train large language AI models, there is a need for enormous amounts of water. For instance, the training process of GPT-3 is said to consume 5.4 million liters of water (Li et al., 2023). However, as Moyano-Fernández & Rueda (2023) state, focusing only on the training of such technologies in terms of their carbon footprint does not give us the whole picture rather, looking at their entire lifecycle will provide a better insight about their environmental cost (p. 222). Therefore, this points out greater challenges that will be added to the existing challenges of the climate crisis.

Another serious environmental challenge that is posed by the use of AI technologies is e-waste. Electronic waste, or e-waste, is the term used to describe electrical appliances that are abandoned, and every year, “around 50 to 60 million tons of e-waste are generated” (Wu, 2023). E-waste contains toxic chemicals such as lead, mercury, and cadmium, which unfortunately have an impact on contaminating the soil and water supplies and put human health and the environment at risk (Kanungo, 2023). In addition, when AI technologies' lifecycle comes to an end, there occurs carbon due to transporting, disposing, and processing waste which increases the negative impacts of

their use on the environment (Mulligan & Elaluf-Calderwood, 2021). Especially when AI's rapid growth is considered, it would not be unreasonable to expect that those impacts will evolve into huge numbers in the near future.

The examples above show that not only the use of AI but also the design, development, and end of lifecycle of AI can have negative impacts on the environmental sustainability that should not be underestimated. This emphasizes the need for what we call "the sustainability of AI". Especially the task of lowering carbon footprint and e-waste of AI's design and development processes is rather a technical issue that needs to be resolved by professionals in related fields. Nevertheless, about the use of AI, there still occur important ethical questions, such as a potential comparison between the resulting value of using AI to combat environmental crises versus the value that could be gained if AI is not commonly used so that the harm of it to the environment will not be increased. However, instead of "sustainability of AI", my thesis takes "AI for sustainability" as its focus and attempts to offer an ethical solution specifically for AI's potential anthropocentric bias. The reasons for this focus will further be explained in the following subsections.

Perhaps the most crucial point that needs to be emphasized is AI's ability to engage in decision-making. It is due to this potential ability that there are now scholars who started to discuss AI being anthropocentric in its decisions that have moral consequences. What is meant by an anthropocentric approach to AI's decision-making could be explained as AI placing more value on humans, or rather, developers and coders of AI put humans' interests at the center. Followingly, a question that needs to be asked is which humans and which interests are valued more. The answer lies especially in the use of AI in relation to the environment and non-human animals, with the promise of efficiency in terms of time, resources, and human labor. The aim of increasing efficiency benefits the companies that use AI in their applications in terms of economic gain.

As the examples above indicate, to incorporate AI in applications regarding the environment and non-human animals, there is mostly a need for training AI with related data. But the question that is overlooked here is whether companies are transparent when sharing the data they train AI with. For instance, to prevent showing how their use of AI technologies exacerbates carbon emissions; some companies may not provide enough data for AI to estimate carbon footprints. In this case, AI's miscalculation of carbon

emissions can mislead governments to create inaccurate policies about fighting climate change (Potts, 2023). In addition, if the data provided to AI in the process of its design and development is not accurate or complete, AI's decisions can end up spreading biases. If an AI system is coded with data that emphasizes the primacy of economic growth instead of protecting the environment, it can prioritize short-term economic gain over the sustainability of the environment (Kanungo, 2023). As Kak et al. (2023) state the ones who own artificially intelligent systems are Big Tech companies and there is no AI independent from these companies. Thereby, the users of AI such as a dairy factory that employs AI to monitor the behaviors of cows, are dependent on the policies of these Big Tech companies which are economically dominating capitalists. When this fact is taken into account, it is not difficult to assert that the value is given to financial gain instead of anything else, including environmental sustainability and the welfare of non-human animals.

Continuing with the ways in which the decisions made by AI technologies can harm the environment and non-human animals will further nurture the discussion on the need for ethically approaching this problem. For instance, a serious problem that could occur if anthropocentrically programmed AI is used in agriculture is the risk concerning the increase in monoculture. AI tools can be implemented in agriculture to produce crops with higher quality. However, growing one type of crop in a field at a time enables farmers to increase efficiency (Balogh, 2021). Accordingly, if AI is programmed to value efficiency instead of environmental sustainability, then it can make decisions for growing one type of crop. Such decisions end up increasing the risk for diseases related to plants (*Is AI Bad for the Environment?*, n.d.). The reason is that the lack of biodiversity also means the lack of species that prevent those diseases. This leads to an increase in the number of pesticides and herbicides used. The resulting effects are pollution of waters, reduction in the quantity of worms and insects that are available to birds, and loss of certain nutrients in the soil, which also cause soil exhaustion (Balogh, 2021). Another potential harm of AI's human-centered decisions on the environment comes from autonomous vehicles as well as drones that engage with environmental monitoring to explore resources. AI technologies with economic efficiency motivations may disrupt wildlife habitats, migration patterns, and cause ecosystem imbalances (Zhuk, 2023). To

benefit humans' motivation for financial gain, in this case, AI can make decisions without taking the interests of the environment as well as the non-humans into account.

2.4 The Negative Effects of AI on Non-human Animals

Importantly, the effects on the environment and non-human animals are interconnected, meaning that such results not only affect the sustainability of the environment but also the lives of non-human animals. Followingly, in this case, for instance, disruption of habitats and migration patterns leads to animal suffering and the loss of many animals' lives. This point on interconnectedness and how non-human animals are potentially harmed if AI is anthropocentrically biased needs further examples. Similar to the effects of AI on the environment, the ways in which AI can harm non-human animals are also overlooked in the literature, except for a very small number of scholars who direct our attention to algorithmic injustice and a need for including non-human animals in the agenda of AI ethics. For instance, Ziesche (2021) argues that not aligning the values of AI with the values of non-human animals besides humans would be speciesism. This indeed lies at the core of anthropocentrically biased AI and its decisions that only prioritize human values. Similarly, Bossert & Hagendorff (2023) claim that even studies advocating the development of sustainable AI remain looking at the issue from an anthropocentric framework, which states that technological applications should balance the needs of both present and future generations of people. There the interests of the ecosystem and non-human animals may only be considered if they benefit humans (p. 3460). On the other hand, other few studies that offer a need to include non-human animals within the scope of sustainable AI studies without remaining in anthropocentric perspectives are Owe & Baum (2021), van Wynsberghe (2021), and Bolte et al. (2022).

One of the important resources concerning the effects of AI on non-human animals belongs to Coghlan and Parker (2023) where they offer a systematic account of animal harm by AI. Throughout history technology has both benefitted and harmed non-human animals. In our century, especially together with recent developments, AI has a serious risk of harming animals (Coghlan & Parker, 2023). The table below provides various cases where non-human animals are negatively affected either intentionally or unintentionally by AI technologies.

Table 1 A practical framework for considering animal harm from AI

Anthropogenic harms to animals	AI harms to animals	Examples
<i>Intentional: socially condemned/illegal</i>	AI intentionally designed and used to harm animals in ways that contradict social values or are illegal AI designed to benefit animals, humans, or ecosystems is intentionally abused to harm animals in ways that contradict social values or are illegal	AI-enabled drones designed and used to locate target animals for illegal wildlife trade Poachers or illegal wildlife traders hack AI-enabled wildlife conservation drones to locate animals
<i>Intentional: socially accepted/legal</i>	AI designed to impact animals in harmful ways that reflect and amplify existing social values or are legal	AI-enabled precision livestock farming enables greater confinement and harmful treatment
<i>Unintentional: direct</i>	AI designed to benefit animals, humans, or ecosystems has unintended harmful impact on animals	<i>Ignorant, reckless, or prejudiced lack of consideration:</i> self-driving cars are not programmed to avoid collisions with small animals <i>Mistake or misadventure in operation:</i> precision livestock farming systems malfunction or operate in unintentional ways that harm animals
<i>Unintentional: indirect</i>	AI impacts human or ecological systems in ways that ultimately harm animals	<i>Material harms:</i> AI proliferation causes harm to the environment through energy use and e-waste thereby destroying animal habitat <i>Harms from estrangement:</i> replacement by AI of human observation and interaction leads to neglect of certain interests <i>Epistemic harms:</i> algorithmic recommender systems reinforce and amplify anthropocentric bias or desire of some people for animal cruelty as entertainment — leading to greater harm to animals through reinforcement of meat eating from factory farms, cruel uses of animals for entertainment, etc
<i>Foregone benefits</i>	AI is disused (not developed or deployed) in directions that would benefit animals (and instead developments that harm or do no benefit to animals are invested in)	Pharmaceutical companies do not invest in AI-enabled veterinary medicine for companion or wild animals because other areas are more profitable Environmental and animal groups fail to receive sufficient funding to develop and maintain AI to monitor and protect animals

Note. From Coghlan, S., & Parker, C. (2023). Harm to Nonhuman Animals from AI: a Systematic Account and Framework. *Philosophy & Technology*, 36(2).

<https://doi.org/10.1007/s13347-023-00627-6>

Although not all harms in those examples resulted from the decisions made by AI, there are still serious effects of AI's decision-making that show themselves, especially in livestock farming. AI that is not programmed to value the welfare of non-human animals and instead programmed to prioritize economic growth may end up crowding more animals in confined spaces and treating them harshly (Coghlan & Parker, 2023). They also draw our attention to the replacement of humans with AI, which can cause to overlook certain needs of non-human animals. This also means that animals' caretakers are becoming artificially intelligent machines. This lack of human-animal interaction leaves the lives of animals in the hands of machines that may place a higher value on efficiency. Even though AI tools are not programmed to merely prioritize efficiency, they may still fail to recognize the needs of non-human animals accurately due to a lack of related data about the signs of animal welfare when compared to humans. In their study, Singer and Tse (2022) also draw our attention to how integrating AI technologies in livestock farming can create harm for non-human animals. For instance, machine learning

models can be used to treat animals' diseases by feeding them extra or killing them. Additionally, sometimes these tools can physically engage with animals by making certain sounds and giving electric shocks to them if they cross the boundaries of the designated area (Singer & Tse, 2022, p. 542). As they also point out, a huge number of animals' lives will be impacted by these AI systems, which have the potential to lower the production costs of industrial farming.

Besides the application of AI in livestock farming, self-driving cars also carry the potential to harm non-human animals. Roadkill occurring frequently is already a huge problem especially when we become aware of the numbers of animals injured or dying due to being hit by a car. For instance, in Sweden, there occurs 20 car accidents everyday where a moose is hit by a car (Bland, 2015). When especially such statistics are taken into account, whether autonomous cars will spend effort to prevent hitting animals and how they are programmed to make those decisions is of high importance. Self-driving cars carry the potential to decrease the number of roadkills however, this is only possible if they are programmed accordingly (Singer et al., 2022). This is the very reason why scholars such as Bendel (2018) discuss animal-friendly machines from an ethical point of view.

2.5 The Need for Ethically Approaching the Anthropocentric Bias of AI

The potential negative effects point out that AI that is not programmed to value the environment, non-human animals, or sustainability in general and makes decisions according to the benefits of humans, especially of economically dominating capitalists that give importance to efficiency may result in suffering and harm¹. Nowadays, there is an effort to make certain changes in AI programming in order to decrease the number of biased decisions that concern human beings. Perhaps what led programmers to make such changes was the fact that the negative impacts became publicly known. One example is

¹ In addition to being anthropocentric, some AI applications also seem to employ a capital-centric approach which would be shortly called capitalocentric. Such systems often disregard the well-being of non-human animals and the environment in pursuit of economic efficiency. For example, agricultural AI may be used to optimize milk production by confining more cows in limited space, prioritizing yield over animal welfare. It is also crucial to note that capital-centric approaches do not necessarily benefit humans. These cases highlight that the ethical challenges surrounding AI have multiple layers, involving both anthropocentric and capitalocentric bias. However, this thesis also addresses AI systems that are not directly profit-driven such as a self-driving car programmed to avoid damage to the vehicle instead of avoid hitting an animal. Nevertheless, even when efficiency and profit-making are not the primary focus, ethical questions remain, particularly around whose interests are prioritized and why.

the gender-biased decisions about Amazon's artificially intelligent hiring tool being seen in the news and creating serious debates (Hamilton, 2018; Malik, 2023). The racial biases of algorithms also found a voice in a popular Netflix documentary called *Coded Bias*. It starts by illustrating how Joy Buolamwini, a researcher from MIT, discovered that the facial recognition technology she was testing could not recognize her black skin (Trenholm, 2021). Such depictions of biased approaches to AI technologies inform not just AI developers but also everyone who encountered this news in the media. Accordingly, newly occurring developments in AI programming started to aim for coping with biases that concern certain groups of people. However, as in the case of gender or racial biases, the problem of anthropocentric biases that concern non-human animals and the environment has not received adequate media attention, and it was not able to create a voice. As Baum & Owe (2023) state many modern societies already exhibit anthropocentrism which is why it is expected that algorithmic biases about the environment and non-human animals received zero attention (p. 140). For this reason, most people tend not to pay attention to, or are not even aware of, the harm that the environment and non-human animals are subjected to. "The AI industry will hire exactly zero employees who are nonhuman animals, so they will have no representatives if humans do not stand up to represent them" (Singer & Tse, 2023, "An Ethical Approach"). The same could also be said for the environment. That is why the discussion above points out a need for AI that is ethically responsible for the environment. Importantly, making AI green (lowering its energy consumption, water usage, carbon footprint, etc.) will not be sufficient rather what is needed is developing AI that is sustainable as well as just and ethical in its decisions. By recognizing this urgent need and taking the discussion one step further, my thesis asks: "What are some ethical ways to deal with the potential effects of AI on the environment and non-human animals?" and "Which ethical theory could be beneficial when programming AI that is not anthropocentrically biased and unjust?"

CHAPTER 3:

FEMINIST ETHICS OF CARE AS A SOLUTION TO THE ANTHROPOCENTRIC BIAS OF AI

This thesis aims to offer a theoretical lens to approach the problem concerning the potential negative impacts of AI on the environment and non-human animals if it is anthropocentrically biased. To investigate the answer to the question “Which ethical theory could be advantageous when programming AI that is not anthropocentrically biased and unjust?”, I delve into a discussion of feminist ethics of care and propose that the use of care ethics which takes interdependent relationships, vulnerabilities, contexts, and needs into account can be a potential solution to prevent AI’s negative impacts. For this aim, this third chapter is dedicated to explaining care ethics and the ways in which the application of this theory in issues concerning the environment and animals will prepare the ground for including AI within the scope of care ethics which is discussed in the next chapter.

3.1 History of Care Ethics and Influential Care Ethicists

Care ethics is an ethical theory that is contrasted with principle-based ethical theories such as consequentialism and deontology. These theories suggest principles or rules about how to act morally right. For instance, utilitarianism, which is famously associated with John Stuart Mill and Jeremy Bentham, claims that you should act in a way that will produce the highest amount of happiness and the least amount of pain. However, care ethics does not require following such principles that may be based on utility or potential consequences. Rather, in moral action and decision-making, care ethics takes interdependent relationships, vulnerabilities, contexts, and needs into account and decides what is the right thing to do based on that.

One of the preliminary steps of developing care ethics comes from a psychologist called Carol Gilligan whose account is not normative yet gives rise to normative works on the development of care ethics. She focuses on the different approaches to morality based on gender. In “In a Different Voice”, Gilligan (2005) puts forward that Freud’s problem concerning the understanding of women’s development is due to the difference about how they experience relationships (p. 693). The difficult and mysterious looking difference of relationships becomes clearer when the way two eleven years old children,

Amy and Jake approach Kohlberg's moral dilemma depicting Heinz is compared. In this moral dilemma, Heinz finds himself in a situation where in order to save his wife's life, he can steal a drug that he cannot afford. Should he steal the drug? Jake, by using his logical reasoning comes to the solution that Heinz should steal the drug even if he does not love his wife. The reason is that the right to life weighs more than the right to property. In this case, stealing is the right thing to do, and a judge would understand Heinz's reasoning if he was caught. On the other hand, Amy's response indicates the failure of logic since she admits that Heinz should not steal the drug, but his wife should not die as well. In her reply, Amy focuses on the possible effect of stealing on the relationship between Heinz and his wife. Rather than looking at the moral dilemma as if it is a math problem, Amy considers it as a narrative of relationships over time (Gilligan, 2005, p. 68). Thus, it is described that women's approach to morality is a relational one whereas men tend to approach morality as a set of rules. Amy's understanding of morality includes the awareness of relationships between people which then "gives rise to a recognition of responsibility for one another, a perception of the need for response" (Gilligan, 2005, pp. 69-70). Such an approach brings us to what we call the ethics of care.

Besides Gilligan, the early 80s also introduced us to Nel Noddings, another influential name in the ethics of care tradition. In care ethics, first we encounter theories based on differences among genders however, later theories emphasize that caring is not specific to women. Noddings (2002) states that analyzing women's approach to morality does not aim to point out that women's morality is superior compared to men rather, the aim is to focus on an ethical perspective that is available through experiences of women which can make our lives better (p. 107). Traditionally in almost all societies, women are expected to care and nurture more than men (p. 19). However, Noddings (2002) adds that "caring is not just for women" (p. 20). In this case, not only women should be responsible for caring for others, but men should too. Therefore, the attitude of caring is universal, however, it is important to note that the element of universality in care ethics does not point to one specific type of moral education and a way of caring (p. 22). Here, being dependent on each other, finding yourself in a network of relationships is universal. This is the very reason why care ethics is contrasted with principled based ethical theories that perceive humans as independent and atomistic (Kwan, 2023).

In relation to the emphasis on the recognition of dependency on each other, people find themselves in a care relationship when they encounter a singular other. Importantly,

care ethics refuses to endorse the impartiality principle that traditional ethics require from moral agents, claiming that they should separate themselves from the context in making ethical decisions (Gruen, 2015a). That is why there is an emphasis on context and affective relationships that people find themselves in. The two parties in a caring relationship that are identified by Noddings are the “one-caring” and the “cared-for” and she argues that both sides are obligated in some way to care reciprocally and meet the other morally (Sander-Staudt, n.d.). Perhaps, an important question that will especially be relevant when we discuss AI in a caring relationship is whether reciprocity is possible, meaning that whether the cared-for can care for the one-caring. The question of reciprocity becomes even more complex when the caring relation is formed between beings that are distinct from each other. Such questions will be discussed in the next chapter which is allocated for the discussion of broadening care ethics to include AI.

Noddings (2002) argues that in a phenomenological analysis of caring relationship, two features characterize the consciousness of the carer. The first characteristic is the presence of a special kind of attentiveness which she explains as a relationship of engrossment. Here, engrossment which refers to empathy in a strong sense can be taken as another term for pregnancy. Therefore, it is as if you are under the skin of another person. In this relation, the one who cares for should “attend to or be engrossed in the cared-for” and there is a need for cared-for to receive the effort of the carer which is another form of attention (p. 16). The second feature that characterizes the consciousness of the carer is the presence of a motivational shift. Here the carer spends its motivational energy on the needs of the one cared-for (p. 28). Simply put, the purpose of a care relationship is to support the other, and in this caring relationship, the one who cares for makes the one who cared-for available to himself or herself. A crucial point that should be added here is that to care for, there is no requirement for an emotional component. Even though a loving relationship that already exists makes caring as well as reciprocity easier, care ethics as an ethical theory is not only possible to be applied in moral dilemmas where the two sides involved are in a loving relationship, such as mother to child, husband to wife, or sister to brother, and so on. Rather, for instance, it is possible for a nurse to care for a patient, a child to care for her dog, or a man to care for a random elderly person whom he saw on the street. Moreover, care relationships can extend to “professional relationships, political relationships with fellow citizens, and even global relationships with distant others across the world” (Kwan, 2023).

The last point above brings us to explain another important care ethicist, Joan Tronto who attempted to broaden the theory by arguing that caring relations do not merely occur between individuals but rather it can also be applied to societies. According to her account, care can contribute to developing a politically good society. However, Tronto (1993) also draws our attention to the fact that care is inseparable from gender, race, and class since in our culture, this determines “who cares and in what ways” (p. 112). Importantly, she states that the ones that are least well-off in society are doing a greater amount of caring while people who are best off make others do the work of care (p. 113). One example could be the fact that immigrant people or women in general are mostly hired for jobs such as housekeeper or as caregiver for the elderly and children. Perhaps, in the next chapter, this point can explain to us the reason why AI technologies are started to be used in healthcare for instance and a possible reasoning behind this is the fact that economically powerful ones make others (AI in this case) do the work of caring instead of doing it themselves.

Tronto (1993) argues for four elements of care: “caring about”, “taking care of”, “care-giving”, and “care-receiving”. She states that from these elements there arises four ethical elements within care: “attentiveness, responsibility, competence, and responsiveness” (p. 127). Attentiveness is an important dimension of care ethics since, in a caring relationship, there is a requirement to recognize a need. That is why, “by this standard, the ethics of care would treat ignoring others as a form of moral evil” (p. 127). In addition, the element of responsibility emphasizes responding to others’ needs. However, it should not be understood as conforming to obligations, and it does not arise from our promises (p. 131). The third element is competence, which refers to the actual work of caring. Sometimes, care can be inadequate due to resources related to providing care therefore, competently provided care is important for the need for care to be met. Lastly, the fourth element of responsiveness is explained as the response of the cared for to the one who cares for. Here we should be reminded that carer and the cared-for are in an unequal relationship since the cared-for is vulnerable which then poses challenges concerning dignity and autonomy. Imagine an elderly person being cared for by a nurse. The person is in need of eating, walking, getting dressed and he is dependent on his nurse and the care that is received. This creates a feeling of vulnerability and perhaps even neediness which might infringe the autonomy and dignity of the elderly person who used to be able to decide and act on his own. Therefore, it is important to emphasize that paying

close attention to potential abuse that results from such vulnerability is required.

3.2 Applied Care Ethics

Care ethics has been a popular theory that is applied to various ethical problems that occur in real life situations. Among the areas where care ethics is commonly applied, we can include medicine (Barns & Peachey, 2024), business (Hamington & Sander-Staudt, 2011), education (Diller, 2018) as well as environmental and animal ethics that I will discuss in more detail in the following subsections of this chapter.

Perhaps the reason why care ethics is preferred to be applied to such areas of life has to do with the fact that they involve the act of caring for others such as from a nurse to patients and a teacher to students. Care ethics has the potential to provide better understanding of the relationships that are present in these areas between different parties as well as the responsibility for recognizing a need, vulnerability, and caring as a result. For instance, Engster (2007), who conceptualizes care as a collection of practices guided by three key virtues of attentiveness, responsiveness, and respect explains how care practices can be applied in business (as cited in Friedman, 2009). Engster argues that governments and businesses are accountable for ensuring economic support during periods of illness, disability, old age, misfortune, or financial setbacks. Additionally, they are responsible for providing protection, healthcare, a clean environment, and safeguarding individuals' fundamental rights (Sander-Staudt, n.d.). This can increase the number of relationships that occur between beings that are seemingly disconnected from each other. Moreover, such a caring approach can allow the recognition of vulnerabilities and needs. Imagine a woman who is responsible for taking care of her sick son and she is having hard time at the office she works. An employer who is attentive to the needs of his/her employees should be able to respond the needs of this woman such as the time and understanding she needs for taking care of her son who is sick. Giving this woman a permission to leave the office early and even providing healthcare opportunities would be an instance of act of caring.

For an employer or a government to provide care for instance, there is no prerequisite for them to have a personal relationship with the woman mentioned above. Indeed, especially the government and the woman seem far distant from each other. This brings us to reemphasize an important point concerning the distance between beings that can take part in a caring relationship. Due to its focus on known individuals and

particular selves, care ethics initially appeared ill-suited as a moral framework for addressing relationships with distant others (Sander-Staudt, n.d.). However, accounts have been put forward to challenge this idea. For instance, Robinson (1999) argues for an account of care ethics that addresses dependency and vulnerability on a global level. In her feminist phenomenological version of care ethics account, responsiveness and attentiveness to the differences of others is crucial when compared to the universalist, rule-based ethical theories since it does not “approach difference as absolute but as existing only in relational terms” (p. 102). Such an approach allows the forming of a network of caring relationships between beings that are different and distant from each other without any form of hierarchy that occurs due to differences. Similarly, Held (2006) argues that incorporating the value of cooperation into society can change hierarchical structures and dominance rooted in gender, class, race, and ethnicity (p 152). Therefore, the relationships between different and distant others once again indicate that interdependence is valued and principle based ethical theories’ emphasis on independence is challenged.

3.2.1 Feminist care tradition in environmental and animal ethics

The discussion concerning caring relationships between distant others brings us to another common area where care ethics is applied: environmental ethics and animal ethics. Both areas focus on the questions of moral consideration of non-humans and how they take part in ethical relations as well as their relationships with humans.

How feminist care tradition is applied to issues that occur about the environment and animals? For the answer, we should first go back to 1974, the year in which the term *ecological feminism* coined by Françoise d'Eaubonne (Warren, 2015). The common characterization of ecofeminist philosophy is as follows:

- (1) it explores the nature of the connections between the unjustified dominations of women and nature; (2) critiques male-biased Western canonical philosophical views (assumptions, concepts, claims, distinctions, positions, theories) about women and nature; and (3) creates alternatives and solutions to such male-biased views. (Warren, 2015)

Ecofeminist ethics is a subfield of ecofeminist philosophy that challenges male-bias in existing ethical theories by proposing values such as care and empathy as well as developing theories of ethics as an alternative to male biased theories. Ecofeminism highlights how binary classifications divide groups, assign unequal value, and reinforce hierarchical perspectives (Bove, 2021). Those binary definitions used for differentiating male and female as well as human (or culture) and nature create oppositional dualism. However, one of the most significant themes within ecofeminist ethics is that their ontologies do not take humans, selves to be purely isolated, opposite and disconnected from others. Rather it is taken as “fundamentally relational, and therefore deeply social, historical, and ecological” (Cuomo, 2003, p. 203). If so, then it also means that while one is caring for himself/herself, not only rights that are associated with individuals are protected but also the others that he/she has relationships with are protected. This understanding allows for being attentive to complex and different forms of moral agency as well as relationships (Cuomo, 2003, p. 204). As one can imagine, those various kinds of relationships include the relationships formed with non-humans such as animals, and nature too.

Relying on the connections and relationships, ecofeminists give various accounts of moral agency or, in the first place, moral consideration for non-human animals and the environment. One of the most commonly argued positions within ecofeminist ethics is care-focused ethics. In specific, feminist animal care theory emerged as a response to the dominance of animal rights and utilitarian theories, which had become the prevailing approach in animal ethics by the 1980s especially through the writings of Singer (1975) and Regan (1983). Both rights-based and utilitarian approaches disregard “sympathy, empathy, and compassion” as meaningful ethical and epistemological foundations for how humans treat nonhuman animals (Donovan, 2022, p. 48). Care-focused ecofeminists rely on the importance of such emotional responses as well as network of relationships, connections and interdependence when assessing ethical situations. Another highly important ecofeminist approach that is endorsed is the contextualized approach. Both rights theory and utilitarianism overlook the specific circumstances of ethical situations, along with their contextual and political complexities. As Emmerman (2022) also states, the traditional way to approach dilemmas that occur due to conflicting interests of humans and non-human animals often depict narrative-less, unknown beings and settings (p. 181). Without being provided with relevant information concerning the context, when we

attempt to make a decision by choosing one side over the other, we indeed engage in a “violence of abstraction” (Kheel, 1993, p. 253). As an example, Kheel (1993) gives us a case where we need to morally assess the use of animals in laboratory for researching a cure for humans who are ill. Here, in its simplest form, the task seems to be deciding whether to give more value to humans or non-humans. However, the problem is separated from the context that it occurs, and it is as if the story that is told to us is cut short and they want us to decide for an ending for the story. Without understanding the worldview that gave rise to the dilemma we are asked to assess; we can only evaluate the situation within its own framework (Kheel, 1993, p. 253). On the contrary, the conflicts between humans and non-humans have background story that is social, economic, political, and cultural (Emmerman, 2022, p. 182). Asking “should we prioritize human lives over animal suffering?” forces a binary choice between humans versus animals without considering the larger structures that created the problem. Given the importance of contexts in care ethics, ecofeminists who endorse care-focused ethics refuse to morally assess scenarios that are detached from the relevant narrative and features of the context. Instead of forcing a choice between the interests of two parties, if we ask, “Why do we depend on animal testing in the first place? What alternative research methods exist?”, we can better assess the situations. When making moral decisions, a situated self rather than a self that is detached from the context, also prevents them from falling into the trap of dualistic dilemmas, hierarchical structures, and exploitation as a result.

For most ecofeminists with a focus on care ethics, the environment and non-human animals with which we have interdependent relationships are conceived as part of vulnerable groups, and they draw our attention to the existing moral distance between humans and animals as well as the lack of responsibility that follows from that distance. As explained above, these studies aim to show that care ethics can provide a good basis for dealing with environmental issues as well as animal ethics. The works of Carol J. Adams and Josephine Donovan are perhaps the most prominent works on the application of care ethics to the problems concerning animal ethics. By expanding on Adams’s (2010) arguments in *The Sexual Politics of Meat* where she discusses the parallels between women’s oppression with the violence animals face with, like other feminist animal care theorists, they propose the ethics of care as “a superior foundation for animal ethics” when compared to rights theory. They challenge these theories due to their inclination to grant rights to animals based on human-like qualities, their rationalist foundations, non-

relational individualist ontology and their dismissal of emotions (Adams & Donovan, 1996; Donovan & Adams, 2007). They take care ethics as a better fit when discussing issues in animal ethics thanks to its relational ontology that recognizes love and empathy as fundamental to human-animal relationships, and also its contextual adaptability that enables a better understanding of animals along a spectrum of differences (Sander-Staudt, n.d.).

There are care ethicists who apply the important components of care ethics to discussions on our responsibility for the vulnerable groups of animals and the environment but there are also other care ethicists who see the responsibility for care towards these vulnerable groups as limited. For instance, Engster (2006) claims that humans should care for non-human animals only to the degree to which they depend on humans (as cited in Snader-Staudt). Similarly, according to Noddings, humans are only obliged towards animals “that are proximate, open to caring completion, and capable of reciprocity” (as cited in Sander-Staudt). On the same grounds, Manning (1996) also distinguishes domesticated animals from wild animals when discussing our obligation for care towards animals. Under such frameworks, it will be possible to argue for a caring relationship between a person and his dog. The person can care for the dog by petting, feeding him and playing with him. Reciprocally the dog can care for the person by showing his loyalty, cheering him up in times of emotional breakdown. But arguing for a relationship between a person living in New York and a chameleon that lives in the Madagascar Island seems not possible. The chameleon does not depend on humans as a pet does and again especially reciprocity from the chameleon’s side seems unattainable. Therefore, the second type of animals are not the types we as humans are obliged to care for according to some care ethicists like Engster and Noddings. However, such an approach that relies on proximity and dependency comes with various problems. It will require us to categorize non-human animals based on the degree to which they depend on us as well as based on their capacity for reciprocity. Such an attempt to categorize animals this way would be a highly complicated task since it is possible for different animals to form unexpected relationships with different people. Perhaps, intuitively, one would think that an octopus does not depend on a human being for its needs such as nutrition and other relevant survival needs. Thus, it makes no sense for a human being to be held responsible for taking care of an octopus that needs food for instance. However, the famous Academy Award winner documentary *My Octopus Teacher* depicts a real-life experience of a man

called Craig Foster who forms an unexpected friendship with an octopus living in South African kelp forest (*My Octopus Teacher* (2020) - IMDb, n.d.). Their intimate relationship allows them to form an emotional bond which then makes a caring relationship as well as responsibility attribution possible since an animal do not only need a human for the tangible things like food and shelter to be provided but they may also have non-physical emotional needs from humans. It also means that even Craig Foster had not dived into the South African kelp forest and had not even gotten in touch with the octopus, still a caring relationship would be possible. Indeed, this also explains why we can donate our money to save people, animals and forests affected by natural disasters like conflagration. We are able to care for them even though we have never met or got in touch with them before. These and many other similar examples indicate that even the seemingly distant beings are not disconnected from each other, and it is possible to form relationships that can also include care when we are reminded of the interconnectedness of beings in the world.

3.3 Overcoming the Moral Distance Between Humans and Non-Humans: Lori Gruen's Entangled Empathy

The last point that emphasizes the interconnectedness of beings brings us to a theory that will be one of the central points of this thesis. In this subsection, to strengthen the arguments above, I propose that the moral distance between humans and non-humans could be overcome by relying on the phenomenological concept of entanglement. In phenomenology, which is the study of first-person experience and consciousness, entanglement refers to the fact that beings and their experiences are interconnected and interdependent. Moreover, it emphasizes that our experiences and perceptions are not isolated, but we are constantly affected by our relationships with others and the world around us. More importantly, the fact that self and others are not the same does not mean that there should be dominance and subordination (Gruen, 2015, p. 96). Unfortunately, the binary classifications that divide us are used as a base to allow for subordination of nature and animals that live in it such as in a way that human-made objects viewed as completely distinct from the natural world. As humans pursue progress, natural environments are often perceived either as barriers to surpass or resources to exploit (Bove, 2021). This perspective frames human advancement as being in opposition to nature's preservation.

One of the reasons for subordination and hierarchy is that rationality is attributed to humans. In this view, morality is grounded in individuals' universal capacity to reason rather than in their relationships and connections with others (Gruen, 2015, p. 59). Following this understanding, there were those who did not give non-humans any moral consideration or right, because there seems to be an opposition between non-humans and humans based on rationality. As a reaction for this view, to extend the moral consideration from humans to non-human animals, traditionally, most animals ethicists started to focus on the similar traits of animals with humans. However, a problem within such accounts is that they mostly overlook the differences that are present. In the first place, it is obvious that there are many ways in which animals differ from humans. We are different in the sense that we are different individuals, and we belong to different biological groups. However, it is also important not to ignore that what we call animals indeed consists of a vast number of different beings. Perhaps, as Gruen (2015) also points out, the category "animal" is a peculiar name, and it is too vague (p. 62). Previously, feminists showed us the harmful results of overlooking the differences between women which ends up creating an image of women who are "white, heterosexual, cis-gendered, able body" (pp. 62-63). Overlooking how different women are from each other prevents us to understand the lived experiences of them which are also very different and unique in their own ways. The vulnerability of a white, middle-class, white-collar worker woman differs from the vulnerability of a black, homosexual immigrant woman. It may be true that both are women, both have vulnerabilities and both experience gender discrimination, but they do it in a way that is so different from each other.

From a similar way of thinking, in animal ethics, some ecofeminists who give accounts with care-focused perspectives attempted to explain how focusing on similarities and overlooking differences can create harmful results as in the case for women. From one perspective, discovering the ways in which animals are similar to humans allowed for the collapse of long-standing divisions between humans and other animals (Gruen, 2015, pp. 46-47). The similarities that are mostly emphasized are the fact that like humans, animals also feel pain, form social relationships, some even suffer when they lose their partners or friends. Pointing out such similarities ideally makes us rethink how we perceive and approach animals. Even though at first glance finding animals' traits that make them similar to us sounds positive, it still falls into the trap of anthropocentrism (p. 47). Gruen (2015) distinguishes two senses of anthropocentrism:

inevitable anthropocentrism and arrogant anthropocentrism. What makes the first one inevitable is the fact that we as humans obviously have human perspectives and we see things from our perspective. One person's perspective can differ from another but nevertheless both are human perspectives at the end of the day. Though this should not mean that an attempt to see things from a nonhuman perspective can never work. The second type is called arrogant anthropocentrism since it puts humans at the center and takes human perspective as superior to all other perspectives (p. 47). An emphasis on similarities of animals to humans makes us engage with arrogant anthropocentrism since it projects our conceptions which we created only by focusing on our perspective.

I believe that the inevitable type of anthropocentrism does no harm unless it is turned into an arrogant type of anthropocentrism. Being a human automatically brings a human perspective and most of the time, we look at the world from our own perspective. On the contrary, the arrogant anthropocentrism prevents us to endorse the perspective of non-humans in any sense and sees the non-human world as only a source that humans can benefit from. Whatever categorization and conceptualization that humans make concerning non-humans, it is accepted as it is, and non-humans are only considered to the extent that they benefit humans. Arrogant anthropocentrism would focus on similarities of non-humans to humans based on human conceptualizations about which features are crucial to gain moral consideration. This also means that non-human animals are reduced to the same perspective as humans. On the contrary, if we emphasize both the similarities and our differences as well as the specific contexts, this will provide us with a better account when we are trying to argue for animals' place in ethical relationships.

About the discussion on the two types of anthropocentrism, there is a need for an important clarification. By refusing the arrogant anthropocentrism and seeing it as harmful, I do not mean that there will never be any form of hierarchy between humans and non-humans in ethical decision-making. This point stands at the center of my thesis which argues for care ethics as a framework to be considered in AI design as a solution to the potential negative impacts on the environment and non-human animals if it is anthropocentrically biased. Therefore, I understand a non-anthropocentric AI as AI that does not prioritize humans over non-humans, nor does it prioritize non-humans over humans. Refusing arrogant anthropocentrism and endorsing care ethics with its relational ontology and context-dependency creates a framework in which we should not determine in advance as a set of rules and principles concerning whose interests should

be prioritized in ethical decision-making. Rather, we should look at the context to determine this. Perhaps animals should be prioritized in one context, while humans should be prioritized in another. This can help us avoid the dangerous pitfall of anthropocentrism and the creation of hierarchical systems that merely prioritize human needs and interests regardless of context.

As previously discussed, considering both the similarities and differences between humans and non-humans, along with specific contexts, offers a more comprehensive approach when advocating for animals' inclusion in ethical relationships. Non-human animals should not be granted moral consideration solely based on their similarity to humans. Instead, recognizing their perspectives and our interdependence should lead to acknowledging their moral worth. Recognizing the perspective of non-human animals and bridging the moral gap between them and humans could be achievable through empathy—an essential element of the care tradition—alongside sympathy and compassion, which are regarded as different forms of attention. However, Gruen (2015) problematizes the standard accounts of empathy (putting oneself in another's shoes) since they overlook the entangled relationship between beings (p. 100). The right kind of empathy necessitates the ability to differentiate one's own perceptions, experiences, attitudes, and beliefs from those of others. When empathizing, distinguishing one and other's perspective can prevent a huge risk of projecting one's own perspective, desires, interests onto another (p. 95). When the ones that we want to empathize with are non-human animals, it is essential to attempt to understand the world as they do—considering how it appears, feels, smells, and sounds from their unique perspective. However, it is also crucial not to get lost in other's perspective. Gruen (2015) argues that the right kind of attentiveness can be achieved through entangled empathy which she defines as follows:

Entangled empathy is a type of caring perception focused on attending to another's experience of wellbeing. An experiential process involving a blend of emotion and cognition in which we recognize we are in relationships with others and are called upon to be responsive and responsible in these relationships by attending to another's needs, interests, desires, vulnerabilities, hopes, and sensitivities (p. 22).

In this way, when empathizing, one is attentive to similarities as well as differences

between herself, her situation and the other that she forms an empathic relation with. Thus, rather than reducing the other to our own perspective, or getting lost in other's perspective, through entangled empathy, we are able to recognize both our first-person perspective and looking at the other from the third-person perspective which is basically a movement between one's point of view and other's point of view. This helps us to understand that we are not part of the same perspective, but we are indeed in an entangled relationship (p. 101). Entangled empathy enables us to perceive and connect with a specific other within their unique context and this skill is fundamental to engaging in ethical relationships (p. 102). Such an understanding offers a great way to deal with moral distance and lack of responsibility between humans and non-humans. Gruen's study tells us that there is already an entangled relationship between us; therefore, we are intertwined. The concept of entangled empathy can change the way we perceive the world when we find ourselves empathetically entangled with vulnerable and the ones in need.

It is essential to recognize that entangled empathy is a process through which we cultivate the ability to be attentive to others' needs by accurately taking on the situation of another. As Jenni (2017) puts forward, in most cases, empathy is understood as the ability to share an affective state with others by attentively considering and imagining oneself in their situation which aligns with the first part of Gruen's definition of entangled empathy (p. 3). But her definition also includes the acknowledgement of relationships and responsibilities follow from that which makes her understanding of empathy more nuanced. The question is in this relationship, we are called upon to be responsive towards what? What are the needs, interests, desires, and vulnerabilities of the other? To avoid any form of error including the projection of our own feelings and needs onto the specific animal we are empathizing with, it is crucial to carefully approach the unique context of the animal so that given our responsibility, we would be responding to the relevant needs in an appropriate way. To succeed that, acknowledging the other's developmental, ecological, and evolutionary backgrounds is of high importance. Gruen (2015) states that this necessitates a willingness to learn, acquire knowledge, and "consultation with people who have experience with and knowledge of the life-worlds of specific others: for example, ethologists, ecologists, primatologists, and long-time caregivers" (p. 94). This point will especially be relevant when I discuss AI's ability for empathy and care in the next chapter.

When the skill for entangled empathy that I explained above is developed, it enables us to form an empathic relationship even with distant others. As one can observe, the concept of entangled empathy closely aligns with care ethics and should more precisely be regarded as a part of care tradition. Indeed, Gruen (2015) discusses care ethics to situate her concept of entangled empathy in its historical framework. Entangled empathy conforms to care ethics' emphasis on interdependence, context, recognition of vulnerabilities, attentiveness to the needs of others, and responding to those needs. At this point, we can say that what Gruen does in her study by using entangled empathy to argue for the place of non-human animals in ethical relationships also points out a revision concerning care ethics. The reason is that her inclusion of animals in ethical relationships (indeed we are already in an entangled relationship with them) by basing her argument on care tradition indicates that she also attempted to broaden care ethics to include non-human animals as well. This means that humans do not only form caring relationships with other humans but also with non-human animals. The whole idea of animals' place in ethical relationships suggests recognizing our responsibility towards them and responding to their needs by approaching them with care.

To sum up, this chapter begins with the introduction of care ethics and a discussion concerning important care ethics figures in history. The important features of care ethics when compared to other principle based ethical theories are its emphasis on context dependent, situated decision-making instead of impartial, abstract principles and rules, interdependent self and relational ontology instead of independent and disconnected self, as well as our responsibility towards vulnerable and the ones in need when making ethical decisions. Due to such features, care ethics became an advantageous theoretical framework to be applied for various areas ranging from business, education to environmental and animal ethics. Especially within ecofeminist ethics tradition, care is taken as fundamentally important in our relationships with distant others including non-humans without posing any hierarchical structures that put humans at the center of everything. It was argued that such hierarchies and the moral distance between humans and non-humans cannot be avoided if we focus on how similar non-human animals to humans when we want to morally consider their needs. The reason is that such approaches still cannot be saved from being anthropocentric. Alternatively, Gruen's concept of entangled empathy points out that we are already in an entangled, intertwined relationship with non-human animals. Recognizing our entanglements and accurately being attentive

to their needs in specific contexts, once again reminds us of our caring responsibility towards them and their place in ethical relationships. This chapter aims to lay the foundation for incorporating AI within the framework of care ethics, the topic explored in the following chapter.



CHAPTER 4:

REIMAGINING CARE ETHICS: THE ROLE OF AI IN ETHICAL RELATIONSHIPS

The last chapter demonstrates that what Gruen does in her work is crucial in the sense that the emphasis on entanglement between humans and non-human animals allows for arguing that animals also engage in ethical relations and accordingly, care ethics can be broadened to include animals. In my thesis, I aim to contribute to literature by proposing to rethink the scope of care ethics and entanglements by asking the kinds of beings that are included in ethical relations. Therefore, in this fourth chapter, I argue that a reimagining care ethics is needed since care ethics can be broadened to include AI which also takes part in ethical relationships and entanglements. First, I critically examine the potential role of AI within the ethical framework of care. Moreover, I consider key distinctions, such as what AI is versus how it appears and extension versus entanglement in relation to AI's role in ethical interactions. Drawing on the concept of social relationalism by David Gunkel and Mark Coeckelberg, this chapter discusses how AI might be integrated into ethical frameworks through its programmed capacities and interactions. Finally, it addresses a central objection: whether AI's simulation of moral actions could be seen as deception.

4.1 Is AI Capable of Caring?

Through care ethics and the concept of entangled empathy, the discussion in the previous chapter prepares the ground for arguing why the interests, vulnerabilities, and needs of non-humans should be morally considered if we aim to offer a solution to the potential negative impacts of anthropocentric bias of AI. Though the discussion concerning caring relations and the emphasis on entanglements involve care from humans to non-human animals and the entangled relationship between them. If the problem focused in this thesis were the negative effects of humans' decisions and actions on non-humans, then offering care ethics as a solution would almost be sufficient. Humans who approach animals with care and empathy would see the positive effects of their actions on animals and the environment. In our lives, we individually relate to others, we evaluate their vulnerabilities and needs. Care relations emerge in that moment. However, since our problem is the negative effects of AI's decision-making, now we need further arguments

concerning AI's ability to care for the environment and non-human animals. If care ethics and our entanglements offer a solution, and if we want to program AI in a way that recognizes entangled relationships, and responds to the needs of vulnerable ones, there are crucial questions that my thesis should try to answer: Can an artificially intelligent being care as human beings do? Can AI form an empathic relation with a non-human animal and accordingly become attentive to its needs so that it can care? Is AI even a moral agent that can approach non-human animals and the environment with moral concerns?²

Care ethics calls for moral agents to become attentive to the needs of others and then by taking responsibility, they should respond to those needs. As Gruen (2015) highlights, it is essential to accurately identify the needs of non-human animals without imposing a human-centered perspective. Additionally, she emphasizes that entangled empathy is a skill that develops over time. Considering these factors, even humans may sometimes fail to recognize their entanglements with others, attend to their needs, and act responsibly. This raises the question: how can AI possess these capabilities in the first place?

One of the preliminary problems that is associated with our question concerning AI's ability for care and empathy is whether AI can be held responsible for its decision-making or whether AI can act out of responsibility which is a topic that is discussed extensively in AI ethics and other relevant literature. If AI can act out of responsibility to recognize the needs of others, then there is also a possibility for it to care. However, even before the question on responsibility, what we need to focus on is the question concerning artificial moral agency since it brings up significant questions about whether it is possible to create artificial entities that are morally responsible for their acts (Himma, 2009). Indeed, it should be clear that AI technologies do not take on roles in areas that require responsibility

² The answers to these questions are closely tied to how we design AI and mitigate its negative impacts. In this sense, the question can also be reframed as: "How do we care for AI?" This is a crucial consideration because the more and better we care about AI (that is the more responsibly we design and guide its development) the better it will care for us as well as non-humans. Thus, the question of how AI cares for us, discussed in Chapter 4, is inseparably linked to how we care for and shape AI. For further exploration of caring for non-humans, including technologies, see Puig de la Bellacasa (2017) who argues for the idea that care is not something that merely humans do but is a speculative and political practice embedded in more-than-human entanglements. This idea suggests that technoscience and naturecultures are deeply intertwined, and both their sustainability and inherent characteristics rely significantly on how care is practiced and sustained within those entanglements.

and decision-making on their own. On the contrary, these roles are delegated to these technologies, suggesting that they have the status of a moral agent.

A moral agent can be defined as an entity that is capable of performing actions for good or evil (Floridi & Sanders, 2004, p. 349). Accordingly, it can be considered responsible for its actions including wrongdoing. Thus, as the idea suggests, artificial moral agents can engage in moral decision-making and act on them. In this sense, they can be “physically embodied robots as well as software agents or bots” (Misselhorn, 2022). Different arguments on whether AI is a moral agent have been put forward by different scholars. This question becomes especially relevant due to recent developments in AI technologies, their abilities for problem solving, offering smart solutions, and even engaging in decision-making that seems to have ethical implications. A common way to argue about the possibility of an artificial moral agency is to first determine the properties that are necessary and sufficient for an entity to be considered as a moral agent. Regarding the conceptions of moral agency, in AI ethics literature, we see an opposition between the standard and functionalist accounts in terms of such properties of moral agency. According to the standard view, moral agency is conceptually linked to being held responsible for one’s actions. Given that, the common conditions that must be met are “rationality, free will or autonomy, and phenomenal consciousness” (Behdadi & Munthe, 2020, p. 198). Johnson (2006), one of the main figures of the standard account of moral agency, provides strict conditions such that it necessitates artificial agents to have mental states, one of which is the intention to act, arising from the agent's freedom. These mental states are beliefs, desires, and other sorts of intentional states. However, artificial entities are not and can never be considered moral agents because they lack the internal mental states necessary to initiate actions that could be attributed to them (Johnson, 2006). Even if it is possible for them to act similar to humans, still they lack necessary internal mental states. Nevertheless, Johnson allows them to be components of moral agent alongside humans since they can be moral entities without being moral agents on their own. The fact that artificial artifacts are intentionally designed and deployed, and they are connected to human intentionality (designers put intentionality to machines) should make us recognize their intentionality which allows them to be important components of moral action. This means that when we are evaluating some action, decision, or situation from an ethical standpoint, the designer of the artificial artifact (say AI designer), artifact itself

(AI), and the user (human users and the ones affected including non-humans) should all be included (p. 204). Therefore, even though Johnson does not provide any possibility for AI to be a moral agent on its own, she nevertheless admits that AI is part of ethical relationships and should be ethically evaluated as an important component of moral actions.

Another proponent of the standard account is Himma (2009) who offers a relatively less stringent account compared to Johnson. An artificial ICT can only qualify as a moral agent if it possesses the ability to make free choices and grasp fundamental moral concepts and obligations—abilities that also assume consciousness (p. 28). That said, for AI technology to be considered as a moral agent, it has to be a highly sophisticated machine, something that is not currently available. The challenge concerning artificial free will that Himma points out is that we are not even sure about what free will is for humans and according to what sense our choices are considered to be free. If so, how can we attempt to develop such technologies? Regarding the consciousness of an artificial agent, there is no consensus among philosophers of mind on whether AI can be conscious. While some of them only accept biologically alive beings as beings that can be conscious, others point out that any brain as complex as the human brain will generate consciousness, no matter what material it is made from. Moreover, we should be reminded of one of the most popular problems within philosophy: the problem of whether other humans are also conscious or in other words, the problem of other minds. The same problem remains when it comes to determining whether a particular AI technology is conscious. Nevertheless, Himma (2009) admits that we may have a moral obligation to treat advanced ICTs as moral agents, even if we lack sufficient justification to believe they truly are—and cannot dismiss the possibility that they are not (p. 28). Indeed, this is what we have been doing in terms of recognizing another human being's moral agency. Even if I lack sufficient justification for whether there are other minds than my own, this does not mean we should not treat them as moral agents who are accountable for their actions. As Himma (2009) claims, if something acts like a human, I may not be justified in believing it has a mind, but I am still morally obliged to treat it as a moral agent. This indicates that even though the proponents of the standard account put high and strict criteria for an artificial agent to be a moral agent, they still recognize that based on how they act, we should indeed treat

them as morally relevant actors and even as moral agents that take part in ethical relationships and situations.

Another popular account regarding artificial moral agency is the functionalist account, an account that is extensively associated with Floridi and Sanders. The functionalist view contrasts with the standard account by relaxing the strict requirements for moral agency that the standard account upholds. They support the idea that there is a need for us to leave behind the traditional approach that focuses on determining whether agents possess mental states, emotions, or so on. Floridi & Sanders (2004) begin with the idea that whether an entity qualifies as a moral agent depends on the level of abstraction (LoA) used when deriving general criteria from typical examples of human moral agency (as cited in Behdadi & Munthe, 2020, p. 198). It is because abstraction significantly influences how we define an entity, meaning every definition is already shaped by an implicit LoA. “An *x* is defined or identified as *y* never absolutely (i.e. LoA-independently), as a Kantian ‘thing-in-itself’, but always contextually, as a function of a given LoA” (Floridi & Sanders, 2004). Each LoA enables a particular way of analyzing a system, and an entity can be interpreted across multiple LoAs.

Based on the arguments above, Floridi & Sanders (2004) put forward the idea that we should raise the level of abstraction when talking about artificial moral agency. The problem concerning the standard view is that it employs a very low level of abstraction which at the end enables only humans to be considered as moral agents. This is the reason why they blame the standard account for being anthropocentric when it focuses only on the human mind. Indeed, not even all human beings can be included in this case. This criticism towards the standard account also raised by Coeckelbergh (2009) who points out that we cannot be even sure whether young children or some psychologically disabled people are moral and responsible agents, making this account disadvantageous in the face of practical problems. (p. 182). Alternatively, Floridi and Sanders (2004) advocate for a “mind-less morality” by refusing criteria such as consciousness (p. 351). If we raise the level of abstraction, it moves away from human-specific traits, making the perspective less anthropocentric. I will return to this point on an account of moral agency being anthropocentric after I explain social-relational account, the account that I will also employ in my thesis.

Floridi & Sanders (2004) propose that moral agency can be approached from a more functional and system-oriented perspective. Most scholars take the question of moral agency and responsibility together. According to Floridi and Sanders (2004), it is possible for artificial agents to be accountable for their actions but not responsible. This means that they can be the cause of a problem, but this doesn't make them blame-worthy like a human. Such limitation does not disqualify the robot as a moral agent. The conditions they offer for moral agency are: "interactivity (response to stimulus by change of state), autonomy (ability to change state without stimulus) and adaptability (ability to change the 'transition rules' by which state is changed)" (p. 349). Without requiring human-like traits and being held responsible, this understanding allows for considering artificial agents as moral agents.

4.1.1 What is vs How It Appears

The opposition between the standard and functionalist accounts of moral agency has been a popular subject within literature but there also are new approaches that do not follow the same path as the advocates of standard or functionalist accounts. I believe a highly crucial argument was put forward by Behdadi & Munthe (2020). They note that the problems on which the philosophical debate on artificial moral agency focus are the problems we see in philosophy of mind and action, metaphysics, and epistemology. Their criticism towards such approaches, which I also agree with, is that they do not provide clear guidance for making practical decisions about the role of artificial entities in contexts that presume moral agency. "The central disagreement between a standard and a functionalist view of moral agency has limited importance for how we should approach artificial entities in practice" (Behdadi & Munthe, 2020, p. 197). In this sense, spending efforts to define properties needed for an entity to be considered moral agent is indeed practically useless. As Gunkel (2012) points out, one major reason is that for instance the concepts of consciousness (p. 55) and personhood (p. 90) come with problems that philosophers could not even have consensus when the agent is a human being. Once again, he reminds us of the problem of other minds and demonstrates the difficulty and complexity of the problem in terms of artificial entities. Besides, there are also epistemological problems associated with the approach that focuses on certain properties of agents that count as moral agents. As it is also stated by Dennett and Derrida, accounts discussed within the debate on artificial moral agency make a "leap from some externally

observable phenomenon to a presumption (whether negative or positive) concerning internal operations" (Gunkel, 2012, p. 64). Such an approach does not help us to understand the way we should approach artificial agencies that we interact with. What Behdadi & Munthe (2020) argues for is the need for shifting the debate to a normative place and giving up on asking the conditions for moral agency from a purely theoretical sense. Rather, the discussion should center on the ways and extent to which these entities ought to be integrated into human activities that typically rely on participants being moral agents and responsible.

The new direction that the debate should focus on as explained above indeed demonstrates an important point concerning the nature of an entity's moral status. If we are able to ask the question how and to what extent artificial agents should be included in certain practices that typically humans do, this suggests the idea that we, as humans assign these roles as well as moral status to these artificial agents. As I stated before in this chapter, it is indeed obvious that AI technologies do not take these roles on their own. Rather, humans delegated them to relevant technologies. This presupposes the idea that the answer to the question of how much AI should be involved in ethical relations is determined by humans rather than ontological characteristics.

Indeed, the answer to the question concerning how and to what extent AI should be included in roles that assumes moral agency and responsibility is already given in our interactions and relationships with AI. Social interactions determine how AI is approached, and how it is perceived as part of ethical relationships. This brings us to explain the social-relational account of moral status which I will endorse in this thesis. Coeckelbergh (2009), the prominent figure of this approach, also criticizes the standard and functionalist accounts of moral agency primarily because they hinge on the complex issue of deciding whether a particular natural or artificial agent possesses freedom and consciousness given when it is also difficult to determine for humans (p. 182). When we interact with other human beings, we do not interact based on whether they really have mental states, consciousness or any other property that is put forward by standard or functionalist accounts. Rather we approach, treat, and hold them responsible based on how they appear to us. That is why appearance is a very important component of morality. This is valid for our interactions with non-humans including AI too. Based on how they appear to us, we treat and interact with AI technologies as if they are human. It is not

necessary to be aware of their internal thoughts or mental states in order to hold them responsible, form a connection with them, or experience affection for them. This is what explains the existence of sex robots for instance. We do not need to know that they are conscious, have emotions and so on to interact with them. It is enough that they look like humans, that they have tastes, feelings, desires. As Coeckelbergh (2009) claims, “both the ascription of agency and of responsibility are, in practice, independent of the real” (p. 184). One of the concrete examples we can give is the humanoid robot Sophia who looks, talks, reacts, socially behave like a human. Sophia can even mimic certain gestures and facial expressions that can be regarded as emotions. People even conduct interviews with Sophia and treat her as a human without questioning whether she really has mental states, emotions, consciousness when interacting with humans. Interesting enough, Sophia was given citizenship in Saudi Arabia in 2017, making her the first robot that received legal personhood. Also, Sophia is the first robot Innovation Ambassador for the United Nations Development Programme (Sophia - Hanson Robotics, n.d.). Perhaps this is not that interesting given that appearance matters how we treat anyone. That citizenship is given to this social robot because of how it is perceived in social and ethical relationships. Appearance based approach is not only valid for robots. Think of your engagement with a conversational agent, a chatbot like ChatGPT or Gemini and how they appear as humans and are perceived as morally relevant actors that are able to engage with decision-making that even has ethical implications.

One can raise a criticism towards Coeckelbergh’s social-relational approach by challenging the exclusion of properties from relationalism if only appearances matter. However, he states that it remains possible to talk about properties, provided we recognize them as properties-as-they-appear to us within a social-relational and social-ecological context (Coeckelbergh, 2010b, p. 219). Gunkel (2012), who is also an advocator of the social relational view, points out that even before knowing the true nature of something, we are already called upon and morally obligated to respond to it in a meaningful way. The properties assigned to an entity are actually the result of how it relates to others, rather than having inherent qualities that exist beforehand (p. 13). Thus, this approach does not refuse the importance of properties but rather brings a new perspective that considers them as properties-as-they-appear to us.

The explanations until this point about social relational account may raise the question whether this account is solely a descriptive account, telling us how we treat moral status of artificial agents based on appearance. However, this would be a mistaken understanding since there is also the normative dimension of this approach given that the aesthetic and phenomenological aspect of assigning agency and responsibility already holds significant normative influence to support moral life (Coeckelbergh, 2009, p. 185). Rather than expanding how we define moral agency, he puts forward that we ought to expand how we define others by rethinking what is included in social and ethical relationships. Appearance and relationality should be the basis of moral status which enables us to broaden ‘others’ to include non-humans. We already treat others (humans and non-humans) based on how they are perceived in social relationships. Therefore, we would not be mistaken to state that Coeckelbergh somehow brings what is and what ought to be the case concerning moral status in line with each other.

Another important point needed to be discussed is whether the social-relational approach has any advantage when dealing with the criticism raised against other accounts for moral status: the fact that they are anthropocentric. The social-relational view argues that attributing consciousness, at least to non-human animals and artificial agents like AI, is shaped by social context rather than objective fact. Moreover, it refers to looking at moral agency in a more general way, rather than focusing on the specific traits of adult humans. If we define moral agency solely based on human traits, we require things like consciousness, free will, and emotions—which only humans can meet. This means AI, non-human animals, or other non-human entities would automatically be eliminated from being moral agents. Defining the properties needed for moral agency (based on human traits) and then evaluating non-humans accordingly can be an instance of arrogant anthropocentrism just like the problem of giving rights to animals based on their similar traits with humans. That is why, for instance, Floridi & Sanders (2004) criticize the standard account of moral agency for being anthropocentric and instead offer their non-anthropocentric account by examining entities from a perspective that does not focus on anthropocentric terms. However, Coeckelbergh (2009) problematizes their approach by arguing that it is not clear why their systemic view of moral agency is considered completely free from anthropocentric bias. The reason is that we create systems and assign the term ‘systems’ to artifacts or groupings of artifacts which makes it almost

impossible to think about them independent from humans. He is right to argue that technology is inherently anthropocentric and closely tied to human culture meaning that there is no artificial or cultural system that exists independently of the human practices that create, envision, and engage with it. (pp. 187-188). Perhaps, in this sense, non-human animals and AI differ from each other. We are not the creators of animals, but we are the creators of AI technologies. An attempt to determine the moral status of animals based on human traits can be considered an instance of arrogant anthropocentrism because as the previous chapter argues, animals should morally be considered regardless of our conceptualizations and decisions about them. Being part of entangled relationships already determines who is involved in ethical relationships, who is worthy of care and who can care. But perhaps the situation is different for AI since we cannot think AI independent from humans. In this sense, maybe assigning certain moral status to AI is not an instance of arrogant anthropocentrism. If anthropocentrism criticism still holds, then it is also possible to argue that the moral status attribution determined by social relations does not necessarily demonstrate the oppression of human conceptualization on non-humans. Perhaps, it is an instance of inevitable anthropocentrism since we normally evaluate those social and ethical relationships from a human perspective. But it is not a perspective that only accepts human categories and concepts regarding what it means to be a moral agent like we would see in instances of arrogant anthropocentrism. There is no checklist prepared based on human traits so that we evaluate whether an artificial agent can be considered moral agent. Rather, taking part in social relations, interacting with others is a prerequisite for an entity to be considered as a moral agent. As Torrance (2014) also explains, it is in these relationships that moral worth and other moral qualities emerge (p. 14). It should be clear that to be a part of social and ethical relations, intelligence and a living body are not required. Rather the appearance is sufficient.

4.1.2 AI as an Extension of Humans vs entangled

There is no problem in remaining agnostic about what AI technologies really are. Perhaps, rather than having its own agency, technology is the extension of human agency. This reminds us of the extended mind hypothesis of Clark & Chalmers (1998). Even before AI or our cell phones, our minds are articulated via writing for instance. Mind has always been something that is articulated towards the outside. In order not to forget (to overcome one such vulnerability) we take notes, write our thoughts down. Then we

started taking notes on our cell phones. Now there emerges another type of vulnerability in which we become dependent on technology and lose the ability to do work on our own. This also allows for the extension of our agency towards AI which can even be considered as extended moral agent with whom we may share responsibility. Since AI takes part in ethical relationships, it may also be part of entanglements. But there is also another option. Perhaps, AI technologies do not share moral responsibility and agency, but they are mere tools. In this they are no different than a pencil. Other accounts can also be put forwards but what I am trying to argue is that reality does not matter. Regardless of whether AI is an extension of humans (that is capable of simulating moral action) or entangled with others, we indeed perceive it as a being like us. At this point, this view is also in line with Behdadi & Munthe (2020). The question of whether AI has moral agency is practically useless and irrelevant. In this regard, the normative approach would be more useful which may open the question of how much AI should be involved. Indeed, AI already takes part in ethical relationships.

Social-relationalist account has the advantage of showing us the dangers associated with AI decision-making. Regardless of what they really are, we treat them as moral agents. As Behdadi & Munthe (2020) argue, we can now ask whether we really want this. But the question is not whether we should treat them as moral agents. Rather the first question is whether we should delegate such roles (that have ethically crucial implications) and responsibilities to these technologies and let designers and developers of AI work on making AI appear as moral agents. AI technologies do not have the potential to harm non-human animals and the environment because we perceive them as moral agents. Rather they carry this potential to harm because we assign these roles to these machines. It is because of how we program, design and deploy these machines for which purposes. In this sense, Behdadi & Munthe (2020) make an accurate point when they relocate the discussion on artificial moral agents to a normative place that questions whether we should assign them this role. But this should not mean that whenever AI has potential to harm, we should get rid of it. It also should not mean that we need to give up using all these technologies (or at least some) in relation to the environment and non-human animals. This was never my point in this thesis. What I am trying to argue here is programming AI with care ethics can help better, more equal, just and sustainable use of AI. One of the ideas presupposed behind the starting question of this thesis was that the

development of AI is inevitable. Indeed, this is not only about AI but about everything else we consider technology. Unless something extreme happens (war, pandemics, natural disasters?), technology and humankind will keep developing as history suggests. I believe the same is true for AI use and I highly suspect that we can stop using AI. Besides, we should not forget the benefits these technologies provide. Chapter 2 begins with explaining how AI technologies promise advantages when they are used in relation to the environment and non-human animals. Nevertheless, its potential harms should not be overlooked. Therefore, once again the social-relational approach emphasizes the place and role of AI in our society and reminds us of the roles assigned to these technologies.

We have discussed the social-relational approach towards moral status and now I will argue that we can also benefit from this approach when discussing whether an entity is able to care. For the sake of simplicity, let us call this care status. Just like moral status, care status also depends on how it appears and is perceived. Care can happen in various ways in the form of, for instance, healthcare from a nurse to patients, parental care from a mother to children, or other forms like from a human to his car, pet, plant and so on. There are many other forms of caring not only due to different agents involved but also due to intensity, duration, motivation and other relevant factors. It may be said that the nurse cares for the patient since it is her job while this is not the case for a mother. Also, one's care for her child can be a lot more intense and for the long term when compared to caring for an object like a car, a house which also needs cleaning, for instance. These differ from each other but at the end of the day, we can all consider them as caring.

In social relations, we stated that to treat another human being as moral agent and morally responsible, we do not question the presence of consciousness, mental states, autonomy, etc. Rather how the other appears already tell us how we should treat him/her. Following the same line of thought, when appreciating a nurse's care for the patient for instance, we do not question whether she conforms to the four ethical elements of care that Tronto specifies (attentiveness, responsibility, competence responsiveness). To be able to say that what she is doing is caring, questioning whether she first recognized a need and then took responsibility to meet that need is practically useless and irrelevant. It is sufficient for her to appear as caring to treat her as a caring agent. The same goes for the care-receiver whom we may call a care patient.

At this point, once again the social-relational approach also holds for non-human animals. A mother lion's care for her offsprings may not in reality be completely the same as a mother's (human) care for her daughter for instance. But when lions teach their offsprings how to hunt, when penguins go for a long journey to find food for their babies or even when your pet dog comforts you and gives emotional support anytime you cry, we also call them caring. We do not even question the presence of Tronto's four elements for human beings, why do we have to do it to consider an animal as caring agent or even care patient. Without really possessing the four elements, appearing as caring is sufficient. Now the same understanding should also be valid for artificial agents like AI. In order to consider AI as being able to care, questioning its moral status, ability for taking responsibility and the presence of four phases of care ethics is practically irrelevant. In reality, AI's care may not be the same as a human being's care just like different humans care differently. However, we should still consider AI as a caring agent if we perceive it as caring in a social relational sense. This also means that AI's appearance as a care patient should remind us of our responsibility towards AI. Given reciprocity, we may owe them care in relation to their needs (for instance, this could be in the form of a physical repairment, new, relevant data input or providing comfort when they are harmed). Social relational understanding of care is also valid for AI's ability for empathy. Based on appearance, we can treat AI that provide nutrients for soil as caring, and AI that recognizes a cow is stressed due to confined spaces and then provides it more space as empathizing and caring. What this discussion also suggests is that if care and empathy abilities are understood in a social relational sense then the burden is on the developers and designers of AI to make it appear as caring and empathizing.

How could such an appearance be possible? AI could be coded to recognize the needs of animals and the environment such as recognizing when cows on a farm need more space and proper food or recognize when a tree needs watering and certain nutrients, etc. Specific facial reactions and behaviors of animals, other related signs from the environment that point out their needs could be learned by AI. This learning can occur when AI is put in an environment and physically and socially interacts with its surroundings via a body. Therefore, even though AI may not be able to empathize in a real sense as humans do, thanks to proper coding and learning resulting from interaction, it can appear as if it can empathize and become attentive to the needs of vulnerable ones.

These points contribute to supporting my thesis for incorporating AI into the ethics of care as a solution for its potential harm.

4.2 Objection: Is AI's simulating moral acts a form of deception?

The arguments in this chapter might receive an objection that need to be addressed. The objection might be given as a follow up to the point above concerning how AI appears. One might object that the fact that AI simulates moral action is a form of deception. It deceives us to believe that it can form empathic relations, care for the environment and also deserve to be cared for in return. However, as a reply, it could be put forward that “to appear in a certain way” does not necessarily refer to a form of deception or a morally problematic situation. As Coeckelbergh (2010a) claims “appearance is the glue of social life and involves learned, often involuntary, salient deception” (p. 10). In their social lives, to play their social roles people also appear in a way that is different from what they really are. In social relations, people hide certain emotions and show expected behaviors, and this is something other than deception. In that sense, AI's appearance as caring is similar to people's appearances in social settings and should not be necessarily considered as deception.

This chapter by first discussing different accounts of moral agency in the literature provides the ground for claiming that care and empathy can also be thought of in line with social-relational approach. This approach does not come up with ontological properties needed for an entity to be considered as a moral agent. Rather, in social relations, moral status is attributed to others (even to humans). In this sense, it is sufficient for an entity to appear as a moral agent. Similarly, if AI appears and is perceived as caring and forming empathic relations with the environment and non-human animals then we should also consider it as caring and empathizing. This point makes the following argument possible: we can approach the problem concerning AI's negative impact on the environment and non-human animals from the perspective of care ethics. If through certain programming, AI cares for non-humans by recognizing their needs and through entangled empathy (appearance is sufficient) then this promises a solution to its potential harm.

CHAPTER 5: CONCLUSION

In this thesis, in order for us to understand the scope of the problem, I first focused on explaining the specific ways in which AI technologies are applied to domains that are related to the environment and non-human animals. The reason why AI applications are becoming widespread is due to the promises AI use makes in related industries: efficiency in terms of time, labor, and resources. I gave examples of AI use such as automated watering systems, animal behavior monitoring, wildlife tracking drones, automated feeding systems and so on.

Even though AI technologies provide certain benefits, the discussion based on the literature indicates that they also carry a huge potential to harm animal welfare and environmental sustainability. There were significant questions that need to be addressed: How ethical is AI in these practices? Who uses it? Who benefits as a result? Who is involved in the development of AI? Whose interests are prioritized? Whose values are transferred to these intelligent machines? Does AI serve the wellbeing of the planet as a whole or solely focus on benefiting some humans? This brought us to consider how the fact that AI engages in moral decision-making, in other words, decisions that have moral consequences, affect the environment and non-human animals. As I already explained in the introduction, the working process of AI tools includes making decisions based on the data provided by programmers. If AI developers while programming AI mostly give value to economic growth rather than environmental protection, that means for the sake of profit-making, sustainability might be overlooked, and we may end up facing severe consequences.

For instance, the increased use of AI in agriculture with economic and efficiency concerns in mind may lead to “the overuse of pesticides and fertilizers, damaging the soil and water, which accordingly harms biodiversity.” (Kanungo, 2023). Animal’s contact with machines instead of their human caretakers, roadkill due to self-driving cars and AI-enabled harmful treatment and confinement in livestock farming also point to negative effects of AI applications.

These were some examples of AI’s negative impact. The existence of such negative effects points out that AI is not programmed to care for the environment, animals, or sustainability in general; rather, it makes decisions based on economic

efficiency and perhaps gives importance to economic growth. Nowadays, I stated that there is an effort to make certain changes in AI programming in order to decrease the number of biased decisions that concern human beings. Perhaps what led programmers to make such changes was the fact that the negative impacts became publicly known thanks to the media. Again, gender-biased decisions about Amazon's artificially intelligent hiring tool being seen in the news and creating serious debates can be an example here. Despite such newly occurring developments in terms of biases that concern humans, the problem of anthropocentric biases was not able to create a voice. That is why it was argued that we are in need of potential solutions and ways to approach the anthropocentrism of the development of AI and its decision-making from an ethical point of view.

In my thesis, I take the overlooking of sustainability and animal welfare concerns for AI's decision-making as my starting point. The research question I focused on is the following: Which theoretical framework could be beneficial when approaching AI that does not have anthropocentric tendencies? This raised further questions: Should AI operate solely according to human interests, even when those interests harm non-human animals and the environment? Can AI be designed in a way that acknowledges the moral significance of non-human entities? What are some ethical ways to deal with the problem of AI's negative impact on the environment and non-human animals?

One point that is of high importance was about what we mean by anthropocentric AI. AI that is developed based on purely anthropocentric (arrogantly) views, will only put humans at the center and disregard the interests of non-humans. As an alternative, in my thesis, I do not argue that animal and environmental interests should be prioritized before humans. Perhaps we should not determine in advance as a set of rules and principles concerning whose interests should be prioritized. Rather, we should look at the context to determine this. Thus, maybe animals should be prioritized in one context, while humans should be prioritized in another.

What I want to draw attention to in this thesis is that many interests of animals and the environment that are overlooked due to people's economic interests carry a huge potential to harm. These should be seriously considered in AI development. In doing this, I ask which ethical framework can be helpful?

In this thesis, I proposed approaching the problem from the theoretical framework

of feminist care ethics and examined whether it offers a solution for potential harms. One of the preliminary steps on developing care ethics comes from a psychologist called Carol Gilligan whose account is not normative yet gives rise to normative works on the development of care ethics. Gilligan focuses on the different approaches to morality based on gender and describes women's approach to morality as a relational one whereas men tend to approach morality as a set of rules. As it is also pointed out by one of the early care ethicists, Nel Noddings, traditionally in almost all societies, women are expected to care and nurture more than men (p. 19). However, she adds that "caring is not just for women" (p. 20). In this case, not only women should be responsible for caring for others, but men should too. Therefore, the attitude of caring is universal although different people can care differently. People find themselves in a care relationship when they encounter a singular other. Importantly, care ethics refuses to endorse the impartiality principle that traditional ethics require from moral agents, claiming that they should separate themselves from the contexts in making ethical decisions. That is why there is an emphasis on context and affective relationships that people find themselves in. Noddings explains the caring relationship as a relationship of engrossment, which is another term for pregnancy. It is as if you are under the skin of another person. In this relation, the one who cares for should "attend to be engrossed in the cared for" (p. 16). The term engrossment refers to empathy in a strong sense. Simply put, the purpose of a care relationship is to support the other, and in this caring relationship, the one who cares for makes himself available to the one who is cared for.

Followingly, a care ethicist, Joan Tronto (1993) attempted to broaden the theory by arguing that caring relations do not only occur between individuals rather it can also be applied to societies. Accordingly, care can contribute to developing a politically good society. Tronto argues for four ethical elements within care: "attentiveness, responsibility, competence, and responsiveness" (p. 127). Attentiveness is an important dimension of care ethics since, in a caring relationship, there is a requirement to recognize a need. That is why, "by this standard, the ethics of care would treat ignoring others as a form of moral evil" (p. 127). In addition, the element of responsibility emphasizes responding to others' needs. Whereas competence refers to the actual work of caring, responsiveness refers to the response of the cared for to the one who cares for.

Care ethics has been applied to issues in environmental and animal ethics,

particularly within ecofeminism which critiques the binary definitions and hierarchical structures that place men over women, humans over non-humans or culture over nature. Ecofeminists highlight that those binary definitions create oppositional dualism. However, one of the most significant themes within ecofeminist ethics is that their ontologies do not take humans, selves to be purely isolated, opposite and disconnected from others. Rather it is taken as relational as well as social. Care-focused ecofeminism also challenges traditional rights-based and utilitarian animal ethics for overlooking the relational and contextual dimensions. Care ethics is the rejection of moral abstraction, the tendency to evaluate ethical dilemmas in simplified, binary terms (e.g., human vs. animal interests).

For most ecofeminists with a focus on care ethics, the environment and non-human animals with which we have interdependent relationships are conceived as part of vulnerable groups, and they draw our attention to the existing moral distance between humans and animals as well as the lack of responsibility that follows from that distance. Such a moral distance could be overcome by entangled empathy, a concept coined by Lori Gruen (2015). Like other care ethicists, Gruen criticizes animal ethicists who defend animal rights by emphasizing their similar traits with humans for allowing animals to gain moral consideration as humans. But this reduces them to the same perspective as ours and again points out a hierarchy of beings. Such understanding also overlooks the differences that are present. A similar problem occurs if we overlook the differences among women which ends up creating an image of women who are “white, heterosexual, cis-gendered, able body” (pp. 62-63). Overlooking how different women are from each other prevents us to understand the lived experiences of them which are also very different and unique in their own ways.

Instead of focusing on similarities between humans and non-humans to allow them to gain moral consideration, Gruen draws our attention to an already existing entangled relationship between humans and non-human entities by suggesting that the appropriate way to be attentive to the needs of others is possible through a phenomenological concept called entangled empathy (p. 54). Under this view, our experiences and perceptions are not isolated, but we are constantly affected by our relationships with others and the world around us.

If we recognize that, when empathizing, we do not merely put ourselves in

another's shoes relying on our similar traits, rather we are attentive to similarities as well as differences between us and others. Thus, rather than reducing the other to our own perspective, we are able to recognize both our first-person perspective without getting lost in other's perspective and looking at the other from the third-person perspective. This helps us to understand that we are not part of the same perspective, but we are indeed in an entangled relationship. We are called upon to be responsive and responsible in these relationships by attending to another's needs, interests, and vulnerabilities. In short, the concept of entangled empathy can change the way we perceive the vulnerable and the ones in need, and it suggests approaching animals with care.

In her work, what Gruen does is crucial in the sense that she advocates broadening care ethics so that it includes animals because thanks to our entanglement, besides humans, animals also engage in ethical relations. With my thesis, I aimed to contribute to the literature by proposing to rethink the scope of care ethics and entanglements by asking the kinds of beings that are included in ethical relations. I argued that a revision concerning care ethics is needed since care ethics can be broadened to include AI which also takes part in ethical relationships and entanglements.

There occurred crucial questions that my thesis should try to answer: Can an artificially intelligent being care as human beings do? Is AI even a moral agent that can approach non-human animals and the environment with moral concerns? Can AI form an empathic relation with a non-human animal and accordingly become attentive to its needs so that it can care? These questions brought me to first discuss whether AI is a moral agent that can act out of responsibility. I have discussed standard and functionalist approaches to moral agency. The normative side of this thesis lies in the argument that moral agency should not be limited to beings with consciousness or emotions, nor should moral consideration be granted only to those who function as humans in relevant ways. Instead, following the insights of social- relational theorists such as Coeckelbergh and Gunkel, I have suggested that moral agency and moral status in general can emerge from appearance and interaction within social and ethical relations. Moral consideration should not depend on what something is (ontological properties) but on how it appears and interacts within a relationship. This also prevents social-relational approach for being arrogantly anthropocentric (although it can be considered as inevitable anthropocentrism). Questions concerning whether AI is conscious, whether it is a sentient

being, that can care, that can empathize, are practically irrelevant and useless. Indeed, these properties are assigned based on how it appears and is perceived in relations because in the first place, we delegated these roles to AI technologies. This means that AI systems need not "feel" empathy or "understand" morality in the human sense in order to participate meaningfully in ethical relationships. What matters is whether they can appear as caring agents, embedded in entangled relationships with others, and whether their actions contribute to the well-being of those others.

Following these arguments, I claimed that the burden is on AI programmers and designers to make AI appear as a caring agent. In this sense, even though AI lacks genuine empathy and care in the human sense, it could be designed to acknowledge these entangled relationships and consider the vulnerabilities of non-human animals and the environment in its decision-making processes and not solely human interests, especially profit related ones. For instance, with proper coding and the opportunity to interact with its surroundings, AI could recognize that cows in a farm need more space and proper food or recognize when a tree needs watering and certain nutrients, etc. Specific facial reactions and behaviors of animals, other related signs from the environment that point out their needs could be learned by AI.

This would involve recognition that different entities may have distinct needs that should be considered when making decisions, even if those interests do not align with human interests in every context. So, a care-focused AI might be understood as AI that should not only be aligned with the interests of humans but also take the environment and non-human animals into consideration by means of care.

In conclusion, offering care ethics as a framework to approach AI's potential harm on the environment and non-human animals brings new understanding in AI design. Through programming and a certain appearance, it assigns an important role in AI to recognize the vulnerabilities, to become attentive to the needs of non-human entities and not solely human interests, especially the ones that harms others. However, one would be right to ask the question of whether programming AI in a way that cares by focusing on vulnerabilities and taking needs into account in each context is even technically possible or not. This is a crucial question that requires an answer. Such a technical question should be addressed by relevant scholars if we want to make AI contribute to sustainability before it is too late.

An important point that may refer to a limitation regarding my thesis is the absence of comprehensive discussion of the dangers of assigning a certain moral status (i.e., moral agency and roles that have significant ethical implications). One major concern regarding AI technologies having moral agency would be their potential for evolving into evil tools that attack humanity. Indeed, that fear was already a very common theme in various movies that one day, robots will supersede and control us. This raises the following question: While giving AI roles that assumes moral agency and enabling it to care, are we risking it turning into something that will get out of our control and turn against us in the future? This problem is not limited to AI's ability to care alone but also applies to any AI use. If AI technologies start taking roles on many different fields that used to be taken up by humans, this may result in human's overdependency on technologies that will turn humans into highly vulnerable beings at the end. Future works on AI ethics need to discuss the relationship between the results of assigning roles on AI and technological determinism. While designing AI to provide benefits and minimize harm, we would not want to face unexpected results. If AI will have a prominent place in our lives and if it carries the potential for becoming superintelligent, then what kind of changes we should expect in social, political, and cultural domains.

Having explained the limitations of my work and suggestions for further studies, it is important to reemphasize that my research aims to offer a starting point on AI's environmental impacts and its role in ethical discussions. If we need to deal with the challenges of the future from an ethical standpoint, we should start by asking not only what AI is capable of doing, but what kind of world we want to live in and how AI can help us achieve it by means of care.

BIBLIOGRAPHY

- Adams, C. J. (2010). *The Sexual Politics of Meat: A Feminist-Vegetarian Critical Theory*. Continuum.
- Adams, C. J., & Donovan, J. (1996). *Beyond Animal Rights: A feminist Caring Ethic for the Treatment of Animals*. Continuum.
- Balogh, A. (2021). *The rise and fall of monoculture farming* | *Horizon Magazine*.
<https://projects.research-and-innovation.ec.europa.eu/en/horizon-magazine/rise-and-fall-monoculture-farming>
- Barns, K. J. E., & Peachey, L. (2024). Seeking a voice: the inadequacy of the “four principles” and the need for care ethics in the provision of health care to vulnerable populations. *The Medical Journal of Australia*, 221(1), 25–28.
<https://doi.org/10.5694/MJA2.52349>
- Baum, S. D., & Owe, A. (2023). Artificial Intelligence Needs Environmental Ethics. *ETHICS, POLICY & ENVIRONMENT*, 26(1), 139–143.
- Bendel, O. (2018). Towards animal-friendly machines. *Paladyn*, 9(1), 204–213.
<https://doi.org/10.1515/PJBR-2018-0019/MACHINEREADABLECITATION/RIS>
- Bland, A. (2015). *Will Driverless Cars Mean Less Roadkill?* | *Smithsonian*.
<https://www.smithsonianmag.com/innovation/will-driverless-cars-mean-less-roadkill-180957103/>
- Bolte, L., Vandemeulebroucke, T., & van Wynsberghe, A. (2022). From an Ethics of Carefulness to an Ethics of Desirability: Going Beyond Current Ethics Approaches to Sustainable AI. *Sustainability 2022, Vol. 14, Page 4472*, 14(8), 4472.
<https://doi.org/10.3390/SU14084472>
- Bossert, L. N., & Hagendorff, T. (2023). The ethics of sustainable AI: Why animals (should) matter for a sustainable use of AI. *Sustainable Development*, 31(5), 3459–3467. <https://doi.org/10.1002/SD.2596>
- Bove, T. (2021). *Ecofeminism: Where Gender and Climate Change Intersect*.
<https://earth.org/ecofeminism/>
- Cattiau, J. (2018). *A tale of a whale song*. <https://blog.google/technology/ai/tale-whale-song/>
- CattleEye | *Autonomous Livestock Monitoring - CattleEye*. (n.d.). Retrieved September 25, 2024, from <https://cattleeye.com/>

- Chen, C., Zhu, W., Steibel, J., Siegford, J., Han, J., & Norton, T. (2020). Recognition of feeding behaviour of pigs and determination of feeding time of each pig by a video-based deep learning method. *Computers and Electronics in Agriculture*, 176, 105642. <https://doi.org/10.1016/J.COMPAG.2020.105642>
- Choi, C., Kim, J., Kim, J., Kim, D., Bae, Y., & Kim, H. S. (2018). Development of Heavy Rain Damage Prediction Model Using Machine Learning Based on Big Data. *Advances in Meteorology*, 2018(1), 5024930. <https://doi.org/10.1155/2018/5024930>
- Cifuentes, J., Marulanda, G., Bello, A., & Reneses, J. (2020). Air temperature forecasting using machine learning techniques: a review. *Energies*, 13(16), 4215. <https://doi.org/10.3390/en13164215>
- Coghlan, S., & Parker, C. (2023). Harm to Nonhuman Animals from AI: a Systematic Account and Framework. *Philosophy and Technology*, 36(2), 1–34. <https://doi.org/10.1007/S13347-023-00627-6/TABLES/1>
- Cowls, J., Tsamados, A., Taddeo, M., & Floridi, L. (2023). The AI gambit: leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations. *AI and Society*, 38(1), 283–307. <https://doi.org/10.1007/S00146-021-01294-X/FIGURES/11>
- Cuomo, C. (2003). Ethics and the Ecofeminist Self. “*Ethics and the Eco/Feminist Self*” (2003) Published in *Environmental Philosophy: From Animal Rights to Radical Ecology (4th Edition)*, Eds. Michael E. Zimmerman, J. Baird Callicott, et al, Prentice Hall, 2003. https://www.academia.edu/4731796/Ethics_and_the_Ecofeminist_Self
- Diller, A. (2018). The Ethics of Care and Education: A New Paradigm, Its Critics, and Its Educational Significance. *The Gender Question in Education*, 89–104. <https://doi.org/10.4324/9780429496530-10>
- Donovan, J. (2022). *Caring to Dialogue: Feminism and the Treatment of Animals* (C. J. Adams & L. Gruen, Eds.; 2nd edition, pp. 47–67). Bloomsbury Academic.
- Donovan, Josephine., & Adams, C. J. . (2007). *The feminist care tradition in animal ethics* . Columbia University Press.
- Emmerman, K. S. (2022). Inter-Animal Moral Conflict and Moral Repair: A Contextualized Ecofeminist Approach in Action. In C. J. Adams & L. Gruen

- (Eds.), *Ecofeminism: Feminist Intersections with Other Animals and the Earth* (second edition, pp. 181–196). Bloomsbury Academic.
- Environmental Sustainability Through AI: 8 Conservation Strategies*. (n.d.). Retrieved September 22, 2024, from <https://2030.builders/8-ways-ai-can-contribute-to-environmental-conservation/>
- Friedman, M. (2009). Reviewed Work: *The Heart of Justice: Care Ethics and Political Theory* by Daniel Engster. 256–258. <https://www.jstor.org/stable/41441884?seq=1>
- Gilligan, C. (2005). Images of Relationship. *North Dakota Law Review*, 81(4).
- Green, G. (2022). *Five ways AI is saving wildlife – from counting chimps to locating whales | Artificial intelligence (AI) | The Guardian*.
<https://www.theguardian.com/environment/2022/feb/21/five-ways-ai-is-saving-wildlife-from-counting-chimps-to-locating-whales-aoe>
- Gruen, L. (2015a). *Entangled Empathy: An Alternative Ethic for Our Relationships with Animals*. Lantern Books.
- Gruen, L. (2015b). *Entangled Empathy: An Alternative Ethic for Our Relationships with Animals*. Lantern Books.
- Hamilton, I. A. (2018). *No Surprise Amazon's AI Was Biased Against Women, Says Sandra Wachter - Business Insider*. <https://www.businessinsider.com/amazon-ai-biased-against-women-no-surprise-sandra-wachter-2018-10>
- Hamington, M., & Sander-Staudt, M. (2011). *Applying Care Ethics to Business* (M. Hamington & M. Sander-Staudt, Eds.). Springer.
- Hao, K. (2019). *Training a single AI model can emit as much carbon as five cars in their lifetimes | MIT Technology Review*.
<https://www.technologyreview.com/2019/06/06/239031/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes/>
- Held, V. (2006). *The Ethics of Care : Personal, Political, and Global*. Oxford University Press.
- Is AI Bad for the Environment?* (n.d.). Retrieved October 21, 2024, from <https://insights.taylorandfrancis.com/sustainability/ai-bad-environment/#section-9xCmaI2cX4>
- Jenni, K. (2017). Review of “Entangled Empathy: An Alternative Ethic For Our Relationships with Animals.” *Essays in Philosophy*, 18(1).
<https://doi.org/10.7710/1526-0569.1577>

- Kak, A., West, S. M., & Whittaker, M. (2023). *Make no mistake—AI is owned by Big Tech* | MIT Technology Review.
<https://www.technologyreview.com/2023/12/05/1084393/make-no-mistake-ai-is-owned-by-big-tech/>
- Kanungo, A. (2023). *The Real Environmental Impact of AI* | Earth.Org.
<https://earth.org/the-green-dilemma-can-ai-fulfil-its-potential-without-harming-the-environment/>
- Kheel, M. (1993). From Heroic to Holistic Ethics: The Ecofeminist Challenge. In G. Gaard (Ed.), *Ecofeminism: Women, Animals, Nature* (pp. 243–271). Temple University Press.
- Kwan, J. (2023). *Care Ethics*. Markkula Center for Applied Ethics at Santa Clara University. <https://www.scu.edu/ethics/ethics-resources/ethical-decision-making/care-ethics/care-ethics.html>
- Li, P., Yang, J., Islam, M. A., & Ren, S. (2023). *Making AI Less “Thirsty”: Uncovering and Addressing the Secret Water Footprint of AI Models*.
- Malik, A. (2023). *AI Bias In Recruitment: Ethical Implications And Transparency*. Forbes Technology Council.
<https://www.forbes.com/councils/forbestechcouncil/2023/09/25/ai-bias-in-recruitment-ethical-implications-and-transparency/>
- Manning, R. (1996). *Caring for animals*. <https://philpapers.org/rec/MANCF>
- Moyano-Fernández, C., & Rueda, J. (2023). AI, Sustainability, and Environmental Ethics. *International Library of Ethics, Law and Technology*, 41, 219–236.
https://doi.org/10.1007/978-3-031-48135-2_11
- Mulligan, C., & Elaluf-Calderwood, S. (2021). AI ethics: A framework for measuring embodied carbon in AI systems. *AI and Ethics* 2021 2:3, 2(3), 363–375.
<https://doi.org/10.1007/S43681-021-00071-2>
- My Octopus Teacher* (2020) - IMDb. (n.d.). Retrieved February 24, 2025, from https://www.imdb.com/title/tt12888462/?ref_=ttawd_ov
- Noddings, N. (2002). *Educating Moral People: A Caring Alternative to Character Education*. Teachers College Press.
- Owe, A., & Baum, S. D. (2021). *The Ethics of Sustainability for Artificial Intelligence*.
<https://doi.org/10.4108/EAI.20-11-2021.2314105>

- Patterson, D. A., Gonzalez, J., Le, Q. V., Liang, C., Munguía, L.-M., Rothchild, D., So, D. R., Texier, M., & Dean, J. (2021). Carbon Emissions and Large Neural Network Training. *ArXiv.Org*.
- Potts, E. (2023). *Climate action requires less-biased AI, Cambridge researchers say*.
<https://www.innovationnewsnetwork.com/climate-action-requires-less-biased-ai-cambridge-researchers-say/36343/>
- Price, K. (2024). *How AI is helping scientists protect birds*.
<https://www.nationalgeographic.com/premium/article/birds-apps-ai-citizen-science-conservation>
- Puig de la Bellacasa, M. (2017). *Matters of care: Speculative Ethics in More than Human Worlds*. University of Minnesota Press.
- Regan, T. (1983). *The Case for Animal Rights*. University of California Press.
- Rizzoli, A. (2021). *AI in Agriculture: 8 Practical Applications [2024 Update]*.
<https://www.v7labs.com/blog/ai-in-agriculture#h2>
- Robinson, Fiona. (1999). *Globalizing care : ethics, feminist theory, and international relations*. Westview Press. <https://www.routledge.com/Globalizing-Care-Ethics-Feminist-Theory-And-International-Relations/Robinson/p/book/9780813333571>
- Rusilowati, U., Ngemba, H. R., Anugrah, R. W., Fitriani, A., & Astuti, E. D. (2024). Leveraging AI for Superior Efficiency in Energy Use and Development of Renewable Resources such as Solar Energy, Wind, and Bioenergy. *International Transactions on Artificial Intelligence*, 2(2), 114–120.
<https://doi.org/10.33050/ITALIC.V2I2.537>
- Sander-Staudt, M. (n.d.). *Care Ethics | Internet Encyclopedia of Philosophy*. Internet Encyclopedia of Philosophy. Retrieved February 5, 2025, from
<https://iep.utm.edu/care-ethics/>
- Singer, P. (1975). *Animal Liberation*. New York Review Books.
- Singer, P., & Tse, Y. F. (2023). *What AI Means For Animals - NOEMA*.
<https://www.noemamag.com/what-ai-means-for-animals/>
- Singer, P., Yip, , Tse, F., & Tse, Y. F. (2022). AI ethics: the case for including animals. *AI and Ethics* 2022 3:2, 3(2), 539–551. <https://doi.org/10.1007/S43681-022-00187-Z>
- Technology - Waste Robotics*. (n.d.). Retrieved September 23, 2024, from
<https://wasterobotic.com/technology/>

- Trenholm, R. (2021). *Coded Bias review: Eye-opening Netflix doc faces racist technology* - CNET. <https://www.cnet.com/culture/entertainment/coded-bias-review-eye-opening-netflix-documentary-faces-up-to-racist-tech/>
- Tronto, J. C. (1993). *Moral Boundaries: A Political Argument for an Ethic of Care*. Routledge.
- van Wynsberghe, A. (2021). Sustainable AI: AI for sustainability and the sustainability of AI. *AI and Ethics* 2021 1:3, 1(3), 213–218. <https://doi.org/10.1007/S43681-021-00043-6>
- Wang, G., Sun, Y., & Wang, J. (2017). Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning. *Computational Intelligence and Neuroscience*, 2017(1), 2917536. <https://doi.org/10.1155/2017/2917536>
- Warren, K. J. (2015). Feminist Environmental Philosophy. *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/feminism-environmental/#SecKinPosFemEnvEco>
- Wu, T. L. (2023). *What Is E-Waste Recycling and How Is it Done?* | Earth.Org. <https://earth.org/what-is-e-waste-recycling/>
- Yun, C. (n.d.). *7 Applications of AI in Agriculture | 2024 Updated | BasicAI's Blog*. Retrieved September 23, 2024, from <https://www.basic.ai/blog-post/7-applications-of-ai-in-agriculture>
- Zheng, G., Li, X., Zhang, R. H., & Liu, B. (2020). Purely satellite data-driven deep learning forecast of complicated tropical instability waves. *Science Advances*, 6(29). https://doi.org/10.1126/SCIADV.ABA1482/SUPPL_FILE/ABA1482_SM.PDF
- Zhuk, A. (2023). Artificial Intelligence Impact on the Environment: Hidden Ecological Costs and Ethical-Legal Issues. *Journal of Digital Technologies and Law*, 1(4), 932–954. <https://doi.org/10.21202/JDTL.2023.40>
- Ziesche, S. (2021). (PDF) *AI Ethics and Value Alignment for Nonhuman Animals*. https://www.researchgate.net/publication/350868232_AI_Ethics_and_Value_Alignment_for_Nonhuman_Animals