

## **Abstract**

In the evolving discourse on AI in connection to robots, the question arises, if one should see them as mere tools, or maybe as entities with autonomous capabilities? That's the reason why this master thesis delves deep into some philosophical principles influencing the discussion about the topic of robot rights. For this purpose, a literature review as well as various expert interviews were conducted. These led to the conclusion that the philosophical principles examined are probably not that important to the real use case but rather to the discussion around robot rights. So one is more likely to end up with the question when discussing such rights: should we continue advancing AI to the point where granting rights becomes indispensable? Regardless of the answer, the imperative nature of addressing robot rights is evident, necessitating further studies, precise definitions, and robust frameworks in the broader domain of AI but also in relation to humanoid robots.

## Table of content

1. Introduction	6
1.1. Relevance of the topic	10
1.2. Objective and research question (RQ)	10
1.3. Structure of the thesis	11
2. Theoretical background	12
2.1 Artificial "intelligence"	12
2.2 Humanoid robots	13
2.3 Post- and transhumanism	16
2.4 Legal constructs	17
2.4 Philosophical constructs	18
2.4.1 Consciousness	18
2.4.2 Agents and agency	19
2.4.3 Action	20
3. Research design	23
3.1 Literature review design	23
3.2 Expert interview design	23
3.3 Additional use of resources	26
4. Literature review	27
5. Analysis	33
5.1 Consciousness	33
5.2 Agency	35
5.3 Action	37
5.4 Humanism	38
5.5 Robot Rights	40
6. Discussion	42
7. Limitations and challenges	48
8. Conclusion	50
9. Bibliography	52
10. Appendix	68

## I List of figures

Fig. 01: The humanoid robot Ava from the film ExMachina	7
Fig. 02: Sophia by Hanson Robotics Ltd. speaking at the AI for GOOD Global Summit	7
Fig. 03: Paro the robot seal by AIST Japan	8
Fig. 04: Robot called Atlas by Boston Dynamics	9
Fig. 05: The differences between AI, ML, and DL	12
Fig. 06: Some key milestones and major platforms in the development of humanoid robots	14
Fig. 07: Humanoid robot Pepper interacting with humans	15
Fig. 08: Interplay of different philosophical concepts in the question of robot rights and the underlying theories.	22
Fig. 09: Sample of interview questions	68
Fig. 10: Main categories of qualitative content analysis according to Mayring	69
Fig. 11: Subcategories (sorted according to the main categories A-E) of qualitative content analysis according to Mayring	71
Fig. 12: Screenshot of the qualitative content analysis and the formation of categories and codes in MAXQDA	84

## **II List of abbreviations**

AI = Artificial Intelligence

RQ = Research Question

IIT = Information integration theory

HRI = Human robot interaction

DL = Deep Learning

ML = Machine Learning

ANN = Artificial Neural Network



## 1. Introduction

In an era marked by rapid technological advancements and digital transformation, AI emerges as a cornerstone of innovation in the wake of a fourth industrial revolution (Schwab, 2017), reshaping industries and transforming the world we live in, on a daily basis (West & Allen, 2018).

While it's not mankind's first technological revolution, it is by far the fastest to date (Xu et al., 2018; Schwab, 2017). Probably one of the most known examples of this also called AI-revolution (Makridakis, 2017), is ChatGPT, a language model by Microsoft, that is the fastest growing platform ever (Cheng, 2023).

Companies from the United States, such as Microsoft, Amazon, Meta, Alphabet, Palantir, and IBM, many based in the Silicon Valley, are leading the charge in AI innovation and significantly influencing the overall direction of the industry (Bangali, 2023). Following closely behind is China, with major firms such as Tencent and Baidu.

Beyond chat based assistants like ChatGPT or the equivalent of Google, Bart (Ram & Verma, 2023), a number of much more complex & advanced AI systems are also being designed. These include humanoid, anthropomorphic (Darling, 2016), robots, that should not only look like humans, but also perform tasks similar to us humans (Rojas-Quintero & Rodríguez-Liñán, 2021).

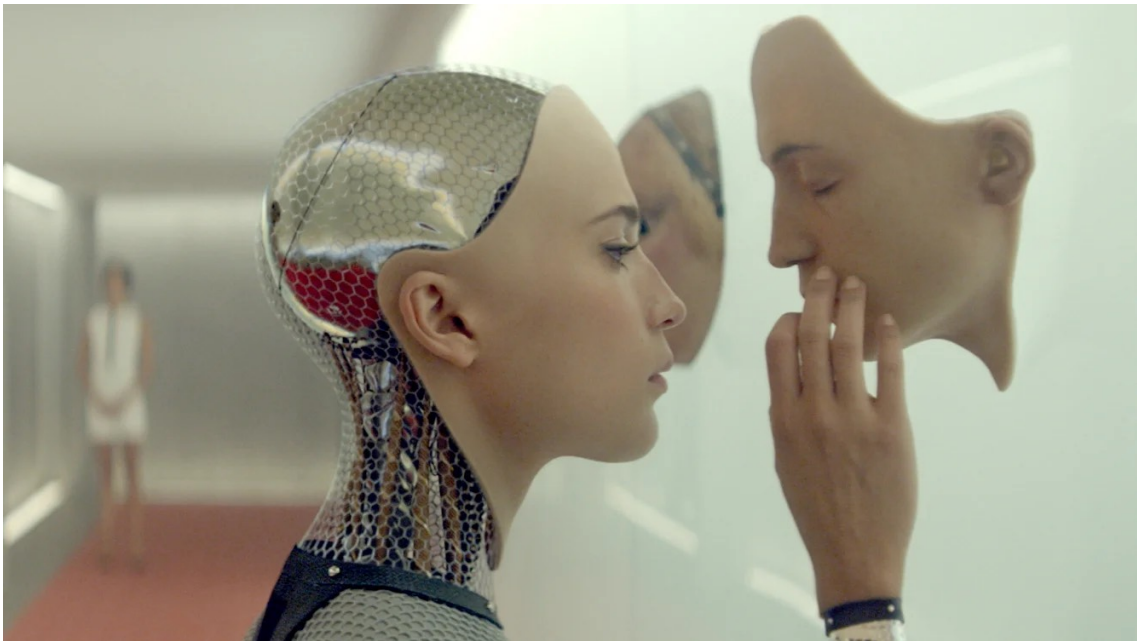
They are anticipated to act as daily life companions and aides for humans (IEEE, 2023), such as in science fiction stories like *I, Robot* (Asimov, 1950). In the sphere of AI and humanoid robots, the gap between science and fiction is rapidly closing, in which fiction often led the way (Adams et al., 2000).

The Oscar-winning film *Ex Machina* offers a specific illustration of a fictional humanoid robot, Ava (Fig. 01). It tells the story of a coder selected to evaluate the human-like qualities of an advanced humanoid robot in a revolutionary AI study (Garland, 2014).

The film explores the dilemma of treating an AI with apparent self-awareness, emotions, and feelings as human. In Ava's case, it illustrates a longing for freedom and self-determination, symbolizing a desire to escape its creator and explore the world, hinting at some form of autonomy and individuality (Garland, 2014).

In this respect, *Ex Machina* inspires profound reflections on the moral and legal status of humanoid robots with remarkable consciousness. This narrative, as well as Michael Crichton's film and later series "*Westworld*" (Crichton, 1973; Nolan & Joy, 2016), that explores a futuristic western theme park inhabited by realistic robots, where guests can let their fantasies run wild, were the starting point for the selection for this master thesis.

**Fig. 01: The humanoid robot Ava from the film ExMachina**



This raises the question of what would happen if the depictions from science fiction were to become reality? Because in truth, we are not far from such a scenario, with numerous humanoid robots that are already comparable to Ava. Perhaps the most prominent example is Sofia, a robot from Hanson Robotics (Fig. 2), known for mimicking human gestures and facial expressions (Hanson Robotics, 2023).

**Fig. 02: Sophia by Hanson Robotics Ltd. speaking at the AI for GOOD Global Summit**



Sofia achieved additional notoriety primarily because Saudi Arabia granted her Saudi Arabian citizenship on October 25, 2017, making her considered by many to be the first robot with rights (Retto, 2017). But this is a common misconception. The first robot to be granted certain rights is Paro (Fig. 03), a 60 cm long personal robot from Japan, used for therapeutic purposes (Schröder, 2021).

"On November 7, 2010, Nanto's mayor officially handed Paro's Koseki [household registration in Japan] to its inventor and registered the robot seal to the scheme, giving the inventor's name as its father (Jones, 2016). Therefore, in 2010 a robot obtained such rights as Japanese minorities could not acquire for four generations, it is named a real Japanese." (Koudela, 2019, p.95)

**Fig. 03: Paro the robot seal by AIST Japan**



Another example would be the robot Atlas by Boston Dynamics (Fig. 04), a 1.8-meter-tall humanoid robot designed to perform tasks in rough terrain. When it comes to imitating human movements and dynamic, Atlas is by far the best and most advanced humanoid robot there is (Boston Dynamics, 2023). Despite not having the conventional facial features of a humanoid robot, its body structure bears a strong resemblance to that of a human.

**Fig. 04: Robot called Atlas by Boston Dynamics**



With the advancement of such technologies, the question of how to deal with intelligent robots in the future arises more and more. Part of this question is the issue of robot rights, which refers to the ethical considerations and principles that might govern the treatment and status of robotic entities (Kurzgesagt, 2017).

This concept challenges traditional understandings of rights and personhood by extending the discourse to non-biological or non-human beings, raising questions about e.g. their consciousness, agency or actions. In this context, also a comparison with animal rights naturally comes to mind (Gellers, 2020).

As in the case of animal rights (Regan, 2013), the argument is often tied to the possibility of feeling pain or having some form of consciousness. Something that was always denied to fish e.g. and then often made their legal position unclear (Braithwaite, 2010).

The question that our society therefore has to ask itself now, is which entities we would like to give what kinds of rights in the future, and this in particular with regard to robots.

## **1.1. Relevance of the topic**

Examples such as Sofia or Paro, but also a number of scientific contributions, clearly show that it's no longer a science fiction topic (Gunkel, 2022a). Not only have rights already been granted to individual robots (Retto, 2017), as robots become more integrated into our daily lives, the responsibility for their actions will make this topic increasingly important in the future.

This makes the issue of robot rights one of the most current topics in practical philosophy (Schröder, 2021). We as humanity should be careful not to create innovations whose effects we cannot assess, as AI could be one of the greatest threats to our society (Eckert & Zschäpitz, 2023).

In this process, the scientific community has so far not even agreed on whether there should be robot rights at all (De Graaf et al., 2021). Furthermore in light of the rapidly growing robotic market, "consensus on ethical issues is usually slow to catch up with technology ... society is still fumbling through privacy, copyright and other intellectual property issues ... nearly 10 years since Napster" (Lin et al., 2011, p.948), which shows the necessity for such work. But there has also been a real boom in the discussion of robot rights in recent years (Gunkel, 2022b).

Yet, when exploring the subject, it's evident that despite its significance and growing focus, there are notable voids in academic research (Harris & Anthis, 2021). Beyond the lack of universally accepted binding laws and the fact that our legal system is not prepared for something like that (Turner, 2018), also the foundation for granting these rights remains ambiguous (Coeckelbergh, 2010).

As a result, this study will try to aid in establishing the grounds upon which robot regulations or frameworks such as robot rights can be developed. The primary focus of this work lies thereby on philosophical underpinnings.

## **1.2. Objective and research question (RQ)**

This master thesis therefore aims to foster a nuanced understanding of AI-powered humanoid robots and the related topic of robot rights and its philosophical foundations. It's intended to empower readers to engage in informed discussions and to lay a basis for future engagement with the topic.

In addition, the discussion on robot rights is a proactive measure to address potential AI risks and can serve as a basis for future AI advances and frameworks that hopefully enable sustainable and harmonious development together with AI in our society.

Therefore, the RQ for this paper states: "Which role do the concepts of consciousness, agency, action, post- and transhumanism play in shaping the debate over the recognition of robot rights for humanoid robots powered by artificial intelligence?"

Thus, the following four areas are addressed in particular:

1. Consciousness
2. Agency
3. Action
4. Post- and transhumanism

So this inquiry clearly takes on a philosophical nature. While technical and legal aspects of robot rights will also be explored partly, the core of the subject rests on philosophical grounds. It questions whether a robot can even be considered an entity with it's own actions and mind that is eligible for granting any rights.

### **1.3. Structure of the thesis**

The thesis is structured into two sections. The first part is theoretical and conceptual, focusing on laying the groundwork to understand the empirical section that follows.

First, the theoretical part provides an overview of topics such as AI, humanoid robots and a post- and transhumanist perspective on them as well as existing or developing legal and societal frameworks. It also delves into the philosophical underpinnings of the subject, exploring concepts like consciousness, agency and agents, and action. This is followed by a methodological section that goes into more detail about the research design and introduces the empirical part that follows.

Second, the empirical part of the paper basically consists of a literature review about robot rights and nine analyzed expert interviews. The literature review should serve to be able to properly classify, analyze and summarize the topic in the scientific context.

The expert interviews, in turn, are used to gain specific expertise and insights from people in the field and about the concepts under study. Afterwards, these interviews are then analyzed and later evaluated in the course of a discussion.

The results from the interviews are then merged with the rest of the information gathered in the thesis, and finally a conclusion is drawn on the whole topic. This also contains an outlook for the future, further possible areas of research on the topic and a number of open questions that came to light in the course of this work.



## 2. Theoretical background

### 2.1 Artificial "intelligence"

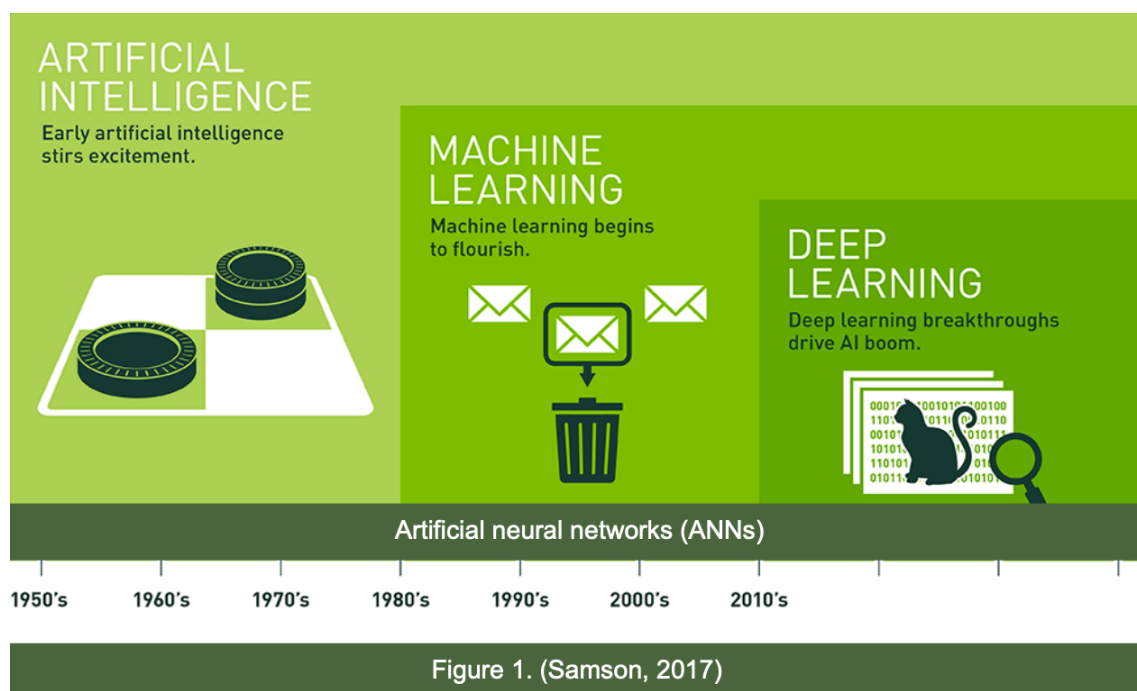
Let's start at the very beginning, the concept of AI. Despite its widely accepted definition as "the science and engineering of making intelligent machines" (Manning, 2020, p. 1, Artificial Intelligence (AI)), there's no clear and uniform definition of AI. Especially with regard to legal requirements (Schuett, 2019), but also when it comes to the actual application areas of AI (Schuett, 2021).

The origins of AI and smart machines can be traced back to ancient history (Truitt, 2021). However, the contemporary realm of AI emerged in the mid-20th century, led by pioneers such as John McCarthy, who coined the term AI in 1955 (Lee, 2021), along with Alan Turing, Allen Newell, Herbert Simon, and Marvin Minsky (Winston, 2016).

They established the groundwork for intelligent machines and developed significant concepts like the Turing test (Turing, 1950), or the programming language LISP (McCarthy, 1978), which paved the way for revolutionary developments in several sub-areas of AI.

AI itself is a broad term that encompasses various subfields, including artificial neural networks (ANNs), machine learning (ML), and deep learning (DL), which all have evolved over time within AI's general scope.

**Fig. 05: The differences between AI, ML, and DL**



ANNs, inspired by the human brain's structure and function, are the basis of modern AI technologies. They employ a series of algorithms organized into different layers to model and process information (Krogh, 2008).

ML represents an evolved version of neural networks. This technique offers a method for modeling data relationships without specific programming (Mahesh, 2020). As a result, it amplifies AI's capabilities in tasks like image processing, data mining, and predictive analysis.

As the name implies, DL is then about the depth of different layers (IBM, 2023). Thereby DL seeks to create a model that emulates the human brain's capability using ANNs and simulations (Aggarwal et al., 2022).

Despite the emergence of these subfields, the term AI has been used as an umbrella term from the beginning of its history to the present day (Helo & Hao, 2021), covering ANNs, ML, DL and many more areas.

Combining these subfields together, one quickly reaches one of the most recent advancements in AI: humanoid robots, where technological progress meets humanity. These robots are the culmination of AI, ML, and DL technologies, seamlessly integrated to mimic human traits and behaviors in form of a robot, that looks like a human.

## **2.2 Humanoid robots**

The term "robot" was coined by Karel Capek in 1920 in the course of a play, which again shows the connection between science and fiction (Reilly, 2011). Today's definitions of the term, however, is no longer that simple.

While at that time it was still a very technical term, robots today are "increasingly understood as autonomous, socio-technical machines" (Braun et al., 2023, p.1, Definition und Abgrenzung), embedded in social contexts. In 2005 Frédéric Kaplan established the definition of a robot as an everyday object (Kaplan, 2005).

Here a robot is defined as "an object that possesses the three following properties: it is a *physical object* (P), it is functioning in an *autonomous* (A) and *situated* (S) manner." (Kaplan, 2005, p.61).

Therefore a robot can be understood as an autonomous entity that interacts with and is influenced by its physical environment (Hegel et al., 2009), which is also used as a definition for this thesis and also can be applied to humanoid robots.



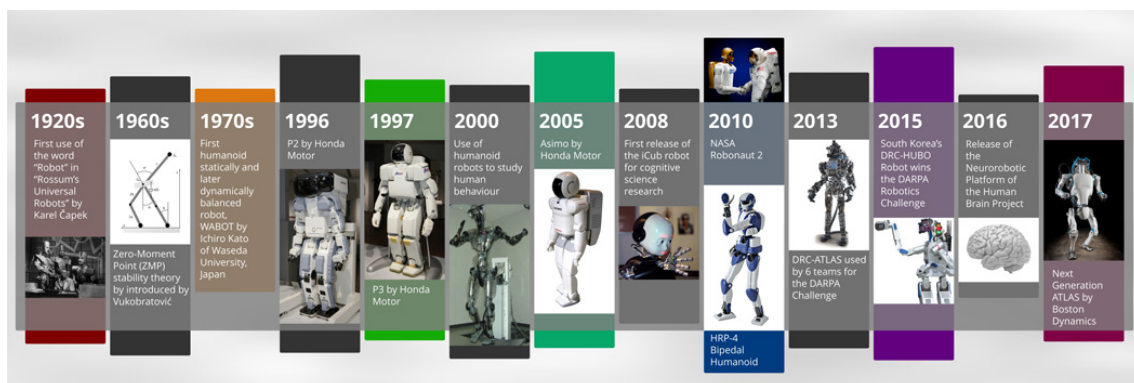
Humanoid robots in particular are commonly described as robots that have a form similar to humans (Yoshida, 2019). Most people hereby stick to the definition of a humanoid robot as a "programmable machine which can imitate the tasks of humans as well as their appearance" (Raju et al., 2020, p.3743).

That's why most humanoid robot designs typically feature a structure with a torso, two legs, two arms and a head to emulate the human body (Stasse & Flayols, 2019). In particular, the robot's head or "face" has a major impact on how it is perceived (DiSalvo et al., 2002).

However, a subset of these designs also places a higher emphasis on their communication capabilities or task performance, rather than physical similarity to humans, which makes it's clear definition quite challenging (Yoshida, 2019).

The first humanoid robot called Elektro was created in 1940 (Denny et al., 2016), whose capabilities were still very limited. In the following years, a multitude of more complex and innovative humanoid robots followed.

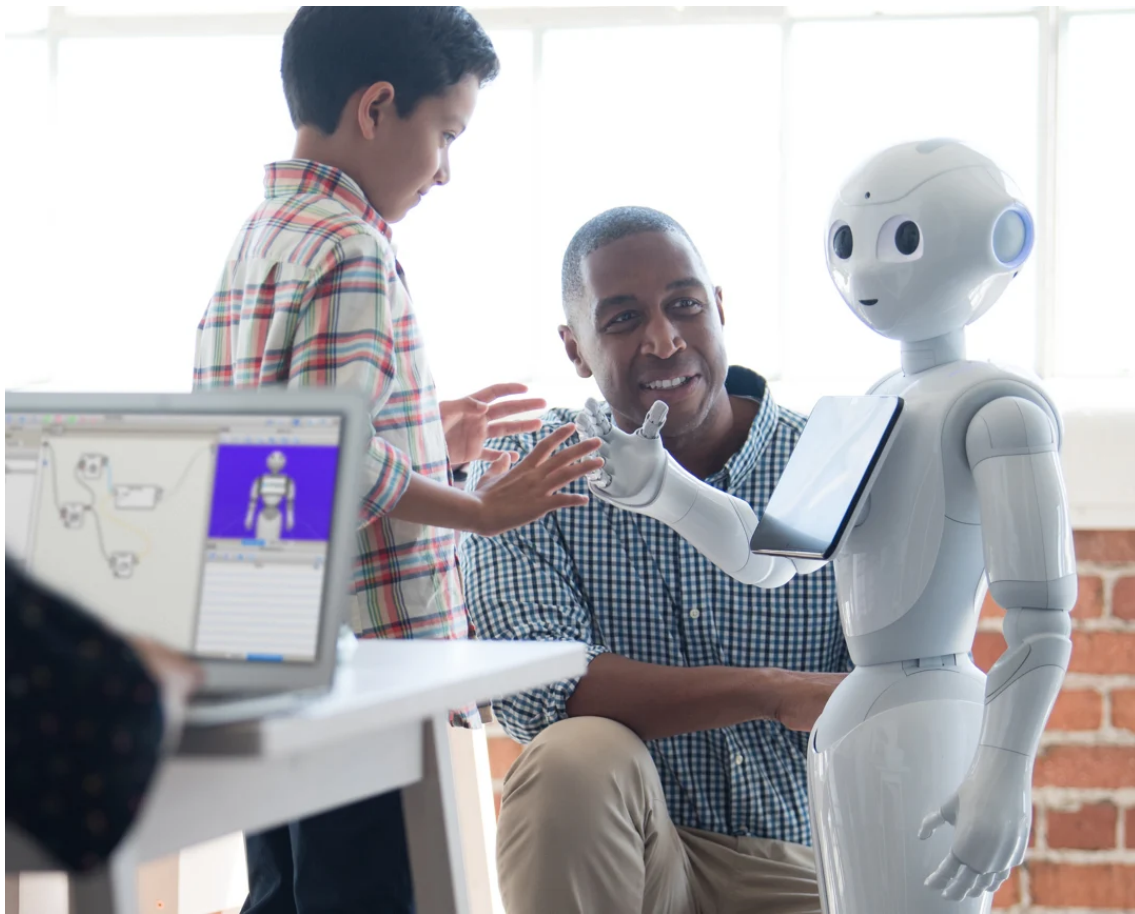
**Fig. 06: Some key milestones and major platforms in the development of humanoid robots**



Another well known example would be the already mass produced humanoid robot Pepper from the company SoftBank Robotics (Fig. 7). Pepper is designed to communicate with humans through voice, touch, and the display screen on its chest (SoftBank Robotics, 2023).

Such technology can already be found in over 2000 companies and institutions worldwide (Aldebaran, 2023), especially in the area of healthcare, education and entertainment (Pandey & Gelin, 2018). Such a development could already be described as a process of robotification, the extensive introduction and integration of robots (Nørskov, 2022).

**Fig. 07: Humanoid robot Pepper interacting with humans**



A process like that fundamentally redefines the relationship between humans and machines. Some researchers express concern about the potential impact of human robot interactions (HRI) on people's lives and their relationships with one another (Friedman, 2023).

Moreover, the idea of recreating our likeness in machines raises intriguing philosophical and ethical questions, some of which directly challenge our traditional notions of what it means to be human (Meaker, 2021).

Based on multiple surveys by the WEF involving AI experts, over half of them believe there's a 50% likelihood of achieving human-level AI by the 2060s (Roser, 2023), a time that's not too distant.

Some even argue that humanoid robots might have the capacity to dream (Bahishti, 2017). Yet, others counter this notion, suggesting we shouldn't anticipate humanoid robots or human-equivalent AI anytime soon (Brown, 2023).

Nevertheless, humanoid robots are definitely more than mere tools (Schröder, 2021). They even can serve as a mirror reflecting the understanding of ourselves as humans

and also helping to understand our cognitive processes (Ishiguro & Nishio, 2007; Ishiguro et al., 2016).

So it's clear that the development is progressing and more and more humanoid robots will find their way into our everyday lives and soon, some of them could even come pretty close to our capabilities.

The question that must then be asked is how we would perceive a humanoid robot with human capabilities, especially in relation to us as humans and what this would mean for the question of robot rights.

### **2.3 Post- and transhumanism**

In asking such questions, one finds oneself relatively quickly in the realm of post- and transhumanism.

"Posthumanism is the set of questions confronting us, and way of dealing with those questions, when we can no longer rely on "the human" as an autonomous, rational being who provides an Archimedean point for knowing about the world (in contrast to "humanism," which uses such a figure to ground further claims)" (Pollock, 2011, p.235).

This perspective could have fundamental implications for the question of robot rights, as the anthropocentric notion that rights are exclusive to humans could be challenged (Birhane et al., 2021). Consequently, also all forms of non-human life could potentially be seen on the same level as humans, not as in the current case of specicism (Duignan, 2023; Coeckelbergh, 2010).

While theories of posthumanism reconsider what it means to be human, transhumanism actively promotes human enhancement and seeks to enhance human potential through technology (Bostrom, 2005). One example of this would be the likeness of a cyborg, which also comes from science fiction, "a human with whom mechanical and/or electronic parts have been integrated" (Clarke, 2010, p.10).

This perspective in particular could become increasingly fascinating in the field of robot rights when the lines between robots and humans become blurred and humans have their bodies and minds modified by robots and technology (Li, 2020; Lin et al., 2011). This development can also be described as co-evolution (Barton, 2014).

In view of such perspectives, however, the question of the current legal status of robots also arises. What framework or precautionary measures have been tackled so far to regulate machines and especially AI and humanoid robots?

## **2.4 Legal constructs**

Before delving into existing legal frameworks, it's crucial to differentiate between different forms of rights, such as moral and legal ones (Schröder, 2021). Throughout this thesis, the term robot rights is understood as a form of legal rights, those recognized by legal systems.

On January 26, 2023, the National Institute of Standards and Technology (NIST) released its AI Risk Management Framework 1.0., a voluntary guide for tech companies that are working with AI to help managing the risks of AI (NIST, 2023).

In addition, the EU is currently working on a suitable framework for the use, classification and handling of AI, the so-called AI-ACT (European Parliament, 2023). In addition, individual countries such as Japan, Australia, Singapore, or South Korea are also developing country-specific frameworks (BSA, 2023; Guo & Zhang, 2009).

But apart from the efforts of the UNESCO and the OECD there has been hardly any effort to create a uniform international framework (UNESCO, 2023; OECD, 2022). However, the drafting of such legal constructs is not easy either. The reason for this may be the different (inter-) national understanding and perception of robots, that make it rather difficult to create a unified global framework (Lin et al., 2011; Guo & Zhang, 2009)

Furthermore, policies must strike a balance between fostering innovation and ensuring that AI technologies do not exacerbate existing societal disparities or create new risks (Zajko, 2022; Rotman, 2023).

Because for now, it is widely known that AI systems exhibit algorithmic bias and the reproduction of societal bias (Srinivasan & Chander, 2021), with its most famous example being COMPAS. An algorithm used in U.S. courts to predict criminal recidivism, which has been criticized for racial bias (Larson et al., 2016).

Moreover, the design and function of humanoid robots, as well as algorithms can also expose our inherent biases and assumptions (Alesich & Rigby, 2017). The decisions that are made in the context of humanoid robots are a reflection of our societal standards, preferences, and preconceived notions (Bartneck et al., 2018).

Vice Versa, "the embodiment and form of a robot also influences human behaviors in intended and unintended ways" (Haring et. al, 2018, p.5). Accordingly, all regulations and frameworks must be done with consideration of just such factors, which further complicates their implementation.

In response to the increasing development, deployment, and interaction with AI technologies by companies and individuals, it also seems that there is a growing focus on AI ethics (Gorden, 2022). But for now, the regulatory landscape around AI can be roughly compared to the wild west (Rimmer, 2017).

## **2.4 Philosophical constructs**

What follows is a closer look at various philosophical constructs, such as consciousness, agency and agents, and action, which are the focus of the RQ and should help to shed light on the topic of robot rights.

### **2.4.1 Consciousness**

The investigation of consciousness is pivotal in discussions about AI, humanoid robots, and robot rights (De Graaf et al., 2021; Levy, 2009). It's a notion typically seen as distinctly human, setting us apart from machines. But its a concept hard to grasp on and thus to define (Levy, 2009; Gunkel, 2018b), which results in a variety of perspectives in which consciousness can be understood (Van Gulick, 2014).

In his work "Consciousness: Function and Definition", E. Niedermeyer defines consciousness as "all modalities of sensory perception as well as memory, thinking, emotionality and volition. Self-perception and self-image are thought to be further components of consciousness" (Niedermeyer, 1994, p.86).

Hereby it should ne noted, that many, but not all of these areas already apply to current humanoid robots. They are able to perceive and process their environment and have a memory. Although, they still lack the ability to think, feel emotions or have their own willpower.

Self-awareness is also not yet present, but like everything else, it is getting closer, with which for many, starts the age of technical singularity (Ramming & Zweig, 2007; Krüger, 2021). Others, nevertheless, steadfastly maintain that it will never be possible for machines to attain a form of self-awareness (Wang, 2023).

While many interpretations of consciousness encompass a physical aspect, there are also views that include the idea of a non-physical consciousness, "modeling it as a nonphysical aspect of reality" (Van Gulick, 2014, p.1, Non-physical theories). Such perspectives open up a whole new understanding and possibility of consciousness for robotics and AI.

A particularly interesting theory to this discourse is the Integrated Information Theory (IIT) of consciousness, by neuroscientist Giulio Tononi. It claims that consciousness is associated with a system's ability to assimilate information (Tononi, 2004). Applying this theory to AI and humanoid robots and implying that if a system can assimilate a substantial amount of information, it could definitely exhibit a form of consciousness.

Another interesting aspect in the discussion of consciousness is pain. Using a reductionist perspective, consciousness can also be reduced to the single sensation, pain (Ambron, 2023). However, pain can also be understood, for example, as just a signal that helps an object avoid a undesired outcome (Schröder, 2021).

The concept of pain is therefore often directly associated with the concept of consciousness, or even considered a requirement for it (Sgourdou, 2022) and the implications of AI's potential to feel pain would be profound. If an AI could experience pain, it would suggest a level of consciousness.

Nevertheless, the current consensus in the field of AI research is that AI does not experience pain and also does not have a form of consciousness apart from theories such as the IIT. Thus, it always depends on how exactly one defines consciousness.

#### **2.4.2 Agents and agency**

The next concept to be described in more detail is that of the agent and the agency associated with it. Because "the majority of work concerning the ethics of artificial intelligence and robots focuses on what philosophers call an agent-oriented problematic" (Gunkel, 2018b, p.87). An agent, in this context, "is a being with the capacity to act, and 'agency' denotes the exercise or manifestation of this capacity" (Schlosser, 2019, p.1, Agency).

The philosopher James Moor has proposed a classification of agents that is particularly relevant to AI and robotics (Moor, 2006; Moor 2009). His classification includes four distinct types of agents:

1. Ethical Impact Agents (Systems with actions or decisions having ethical consequences, like a calculator affecting financial accuracy)
2. Implicit Ethical Agents (Systems avoiding unethical results by following set rules, e.g., a self-driving car prioritizing safety)
3. Explicit Ethical Agents (Systems deciding based on ethical guidelines but without understanding the implications)
4. Full Ethical Agents: (Systems understanding and justifying decisions ethically, possessing consciousness and free will)

As of now, a full ethical agent doesn't exist in AI technology. Moor therefore suggests, that one should work on explicitly ethical agents. In his opinion, they are the most realistic option to work with.

In the context of robots, this classification could be used as a framework to understand when a robot can be considered an agent. Furthermore it could explain what type of agent to use to determine what rights a particular agent, or in this case robot should have.

So, according to this classification, one can classify all humanoid robots as Ethical Impact Agents, Implicit Ethical Agents, or even Explicit Ethical Agents, but not as Full Ethical Agents as it would require to develop a form of consciousness.

In order to fully understand the topic of agency and agents in terms of robot rights, it also becomes essential to define actions. This is due to the fact, that agency and actions are intrinsically linked.

### **2.4.3 Action**

The term action when associated with agency is typically used in the context of action philosophy (Bencherki, 2016). Hereby agency represents the capacity to initiate actions, while actions themselves are the specific expressions of this capacity.

Just as in the case of consciousness, there are a variety of possible interpretations and classifications for the term of action (Glasscock & Tenenbaum, 2023). They can be understood as a form of motivation, event, causal relation or even inaction (omission). Thus, there are a number of theories that consider an action in a humanoid robot possible, but also a number that argue against it.

Again, it depends very much on the definition. In the case of action philosophy most agree on the fact that the action of a humanoid robot is not a conventional human action. But in general it should be noted that most of the written code on which the action of a robot is ultimately based, also still comes from a human, which is also expected to change within the next few years (Rosenbaum, 2022).

From this it can be derived that every action that a robot carries out in the end originates from human intervention. This consideration allows the assumption that every action of a robot is also an action in its original sense, just via a new medium. Such a perspective also becomes increasingly fascinating as soon as machines should be able to extend or write their own code (Bostrom & Yudkowsky, 2018).

Furthermore, it can also be argued that an action performed by a robot is nothing else, no matter how intelligent it seems to be, but an operation. An operation is most of the times describes as an "act or process of working, doing something, being in action" (Cambridge Dictionary, 2023, p.1, operation) or also as an "activity planned to achieve something" (Cambridge Dictionary, 2023, p.1, operation).

"As nouns the difference between operation and action is that operation is the method by which a device performs its function while action is something done so as to accomplish a purpose."(WikiDiff, 2023, p.1, Operation vs Action - What's the difference?)

From this it follows that an action is nothing more than the execution of the operation embodied in the code. But whether the action of a robot is understood as action or operation, it becomes clear that it is primarily about the code.

Thus, to properly classify actions of a robot, it requires an examination of the various philosophical theories that can be incorporated into a code in the first place (Lin et al., 2011). Because in fact, just two out of the three main theories from normative ethics (deontology, consequentialism, virtue ethics), can be applied to programming.

Deontological ethics judges actions based on inherent properties (Alexander & Moor, 2020). This means certain actions might always be deemed wrong, regardless of their outcomes (Britannica, 2023). These properties can be defined in the case of a code simply in the form of variables and built into the code.

Conversely, consequentialism evaluates actions by their results, suggesting the end justifies the means (Sinnott-Armstrong, 2019). This can be easily applied to a code too, as a program is ultimately only a calculation of ones and zeros and consequentialism is



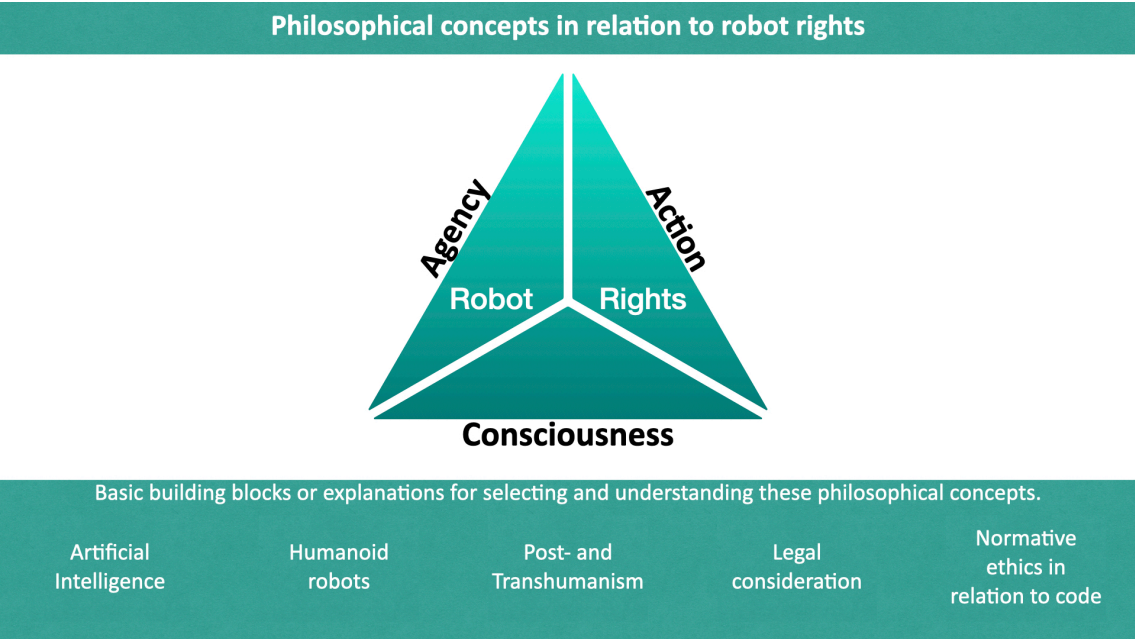
also based on calculations. However, decisions based on consequentialism often rely on incomplete data, especially for future events (Card & Smith, 2020). Besides these, virtue ethics is another normative approach (Solomon, 1995). But unlike rules or outcomes, virtues are challenging to code due to their complexity (Wernaart, 2021), which represents a significant hurdle in AI development.

So it becomes clear that to address the discussion about robot rights, a multifaceted approach is needed. Beyond the typical concepts of technical and legal aspects in connection to AI and robots, it should therefore probably also cover philosophical concepts like consciousness, agency, action and the foundational theories for creating a code.

Together, these philosophical concepts form a complex web that delves deeply into the debate over the rights of robots. In a world where technology is advancing rapidly, it is imperative to understand and define these concepts in order to be able to develop clear legal and ethical constructs for the future.

After all, it is only with an understanding of consciousness that the issue of agency can be properly framed, and with it, that of an action. But the other way around is just the same. Also the topic of agency can only be discussed with regard to the different forms of consciousness. Thus a kind of triangle is formed, in which each area is connected with the other and influences the discussion about the question of robot rights.

**Fig. 08: Interplay of different philosophical concepts in the question of robot rights and the underlying theories.**



### **3. Research design**

In the following, the research methods used in this thesis will be discussed in more detail in order to make possible to understand exactly where the results of this thesis come from. These include numerous expert interviews and a literature review.

#### **3.1 Literature review design**

The literature review was prepared to serve to illustrate the current state of research in the field of robot rights. It summarizes the state of knowledge to date, including most of the major theories and concepts used in the field (Bell et al., 2011). It is a historical literature review (UHN, 2023). In doing so, the literature review follows a chronology. This means that the various works are presented mainly in the order of their publication date. An exception to this is the outline of Asimov's Laws.

Such an approach helps to classify the complex topic with its many perspectives and theories chronologically, as a purely scientific comparison of these works seems difficult and a limitation to a few works would have diminished the scientific added value.

So the literature review serves as a general assessment of the topic of robot rights, even beyond philosophical aspects. Publications of all kinds are used to provide the clearest possible scientific overview of the concept (Bell et al., 2011).

It aims to uncover controversies in the question of robot rights, identify possible research gaps and to foster profound discussions (Bell et al., 2011). The aim is to shed light on the current state of robot rights from a scientific perspective and to show which approaches have been incorporated into the discussion so far.

#### **3.2 Expert interview design**

The expert interviews were conducted, to get insight into the field through different opinions of experts in the area (Bell et al., 2011), and to explore the question of what philosophical principles speak for or against the recognition of robot rights in detail.

These were for the most part semi-structured interviews, which are interviews that are conducted with a predefined selection of questions to which then some specific questions are added, depending on the expert and the field (Bell et al., 2011). A representative list of questions from a conducted interview is shown in Fig. 09 in the appendix.

Furthermore, the selection of interview questions was strongly oriented towards the RQ as this resembles also the main research topic. Thus, questions about the experience of the expert with robot rights were discussed first, then the question of agency, the distinction between action and operation, the concept of consciousness, and lastly a post- and transhumanist perspectives on the debate about robot rights. In addition, there were also personal questions asked about the person's expertise and field of research, the possible opportunities and risks of robot rights, as well as the question if such rights should even exist.

A small number of the interviews also involved written responses from experts. These are also treated as semi-structured interviews in the course of this work, even though they do not correspond to them in the classic sense. In this way, the results of the written and oral answers can be better summarized and also merged.

Moreover, when conducting such semi-structured interviews, the selection of interviewees or experts is of utmost relevance. The quality of the research stands or falls with them, and their answers ultimately form the basis for the research work.

Therefore, the interviews used in this thesis were conducted only with philosophical experts from the scientific field. These include various scientific employees and professors. These come from various international universities and all have experience in the field of technical philosophy, many with a background in AI or even robot rights itself. For reasons of security and data protection, the interviews were anonymized.

The interviews were mostly conducted online, using either Microsoft Teams or Zoom. One of the interviews also took place only on the phone. To record them, Programs such as GarageBand, iPhone voice memos or the corresponding online functions of Teams were used.

After that, the interviews were transcribed into simple transcripts that have been simplified for readability and comprehensibility. Gaps and grammatical errors have been removed, dialects have been translated, and no specific references to pitches have been made (Dresing & Pehl, 2018). The transcripts, as well as the associated rules, can be found in the appendix.

The transcripts have been assigned unique page numbers for easier navigation, and the experts' responses are organized into numbered sections. These sections will be referenced throughout the analysis to ensure clarity and seamless understanding of the academic process.

To guarantee a thorough analysis of the data collected in the course of the interviews, Mayring's (2010) method of structured content analysis technique was employed. This involved crafting categories, subcategories, and their respective coding guidelines. Fig. 10 showcases the main categories while the intricacies of the subcategories are illustrated in Fig. 11 (see appendix). Throughout the coding journey, these categories were changed and fine-tuned to seamlessly align with the content at hand.

The categorization was not predetermined by existing academic dialogues, but made inductively, drawing inspiration from the data itself. Such a method is particularly well suited to explore the field of robot rights, as it's an academic field with minimal established discourse.

For a better overview and an easier analysis, the main categories have been abbreviated from A-E. In addition, the subcategories were clustered into different clusters, as shown in Fig. 11 (see appendix). Given the limited scope of this work, the focus in the analysis lays on these overarching clusters rather than on each individual subcategory. This was the only way to include all research findings in the analysis. The final clustering of the categories was done on a classic whiteboard, on which categories were merged and brought into a coherent and meaningful system.

Because most of these interviews were conducted in German, and in order to give every reader the opportunity to understand them, it was decided to quote the mostly German answers in the course of the analysis and discussion in English, whereby the appropriate section and transcription page can always be found. In the case of the subcategories, square brackets [] with an English translation are placed next to the German original anchor examples. All translations were made with the help of the software DeepL.

The categorization and coding took place with the help of the software MAXQDA. The already transcribed interviews were loaded into the software and with the help of this, the categories and codes were formed. This software was also used in the course of the analysis and discussion for a better overview of the many categories and texts. A screenshot of the software can be found in Fig. 12 in the appendix.

### **3.3 Additional use of resources**

In addition, in the course of writing this thesis, a number of support programs were used that have not been mentioned previously. These include the program ChatGPT, which was used for the reformulation of self-written texts and translations, as well as the already mentioned program DEEPL. In addition, the software Amberscript was used, but it did not prove to be particularly helpful. Finally, the work was written in Pages and the specially created graphics were designed with the help of Apple's Keynote program.

## 4. Literature review

What comes next is an examination of relevant texts that aids in pinpointing existing works that can be related to the rights of robots. This should facilitate comprehension of the theoretical constructs in this domain, as well as the specific jargon used. This will be beneficial for a deeper exploration of the topic.

The first attempt to define rights for robots came from the famous science fiction author Isaac Asimov (Vadymovych, 2017). In his work "Runaround" (Asimov, 1942), which was already published in 1942, he named for the first time the three robot laws, which are still very well known and popular today (Brauner, 2016; Coeckelbergh, 2010).

1. "A robot may not injure a human being or, through inaction, allow a human being to come to harm" (Anderson, 2008, p. 477)
2. "A robot must obey orders given it by human beings except where such orders would conflict with the First Law" (Anderson, 2008, p. 477)
3. "A robot must protect its own existence as long as such protection does not conflict with the First or Second Law" (Anderson, 2008, p. 477)

Asimov later added another law, the zeroth law, that puts the interests of mankind above the interests of any individual (Asimov, 1985).

4. "A robot may not injure humanity, or, through inaction, allow humanity to come to harm" (Clarke, 1994, p.1, Asimov's 1985 revised Laws of Robotics The Zeroth law)

These often serve as a starting point for the discussion about robot rights. As for example in the text: "Beyond Asimov: The Three Laws of Responsible Robotics", in which the authors Murphy and Woods (2009) revise the old three laws and publish an adapted version of them, just as Balkin (2017) in his paper: "The Three Laws of Robotics in the Age of Big Data".

But there is also a lot of criticism of the laws. Aaron Sloman (2006) suggests that these three laws are completely unethical and could more likely lead to racialism or speciesism, as robots also only adopt human behavior. Further, Stokes (2018), believes that while a lot of engineers around the world follow these laws, they simply don't work in practice.

What is indisputable, however, is the fact that Asimov's works have proven that by developing and implementing a set of rules, it is possible to "reliably constrain the behavior of robots" (Clarke, 1994, p.1, Enhancements to codes of ethics).

So although the legal status of robots was discussed early on, no clearly formulated uniform rules emerged in the years following Asimov's publications. On the contrary, in 2011 the Engineering and Physical Sciences Research Council (EPSRC) still defined robots as tools or artefacts (Prescott, 2017), but definitely not as entities which could have certain rights.

The discussion about the rights of robots then really took off again after the Committee on Legal Affairs of the European Parliament spoke of "creating a specific legal status for robots" (European Parliament, 2016, p.12) in 2016 (Lima et al., 2021), and the robot Sofia was granted citizenship in 2017 (Parviainen & Coeckelbergh, 2021).

But needless to say, that there were also a number of publications that dealt with the issue of robot rights even before that. As early as 2001, Inayatullah (2001) wrote the paper "The Rights of Robots? Inclusion, Courts and Unexpected Futures", in which he talks about the need for legal rights for robots. He primarily addresses the opportunities and challenges facing our legal system in this context.

Although, the author also mentions an argument that later became more and more important, namely, the granting of robot rights based on the fact that this was also done for corporations and draws comparisons between the denial of robot rights and the association of other underrepresented groups such as animals, plants, or children.

Only five years later in 2006, Asaro (2006) even wrote about the fact that in the future, robots could simply ask for their rights on their own. Thereby, he emphasizes the aspects of consciousness and moral agency and speaks of a continuum of agency on which robots move and slowly evolve.

In 2009, Levy (2009) also published a significant work for the field of robot rights, "The Treatment of Artificial Conscious Robots". In it, he focuses on the value of consciousness in the debate on the treatment of robots. Levy suggests that robots with consciousness should be given certain rights simply because we will see and treat them similarly to humans. Moreover, the way we interact with robots will also change the way we interact with each other as humans.

In the same year, Bryson (2010) also published her opposing work: "Robots Should Be Slaves", in which she argues that robots are built and designed by and for us humans and should be sold as slaves. However she says directly that it's not about making a human robot a slave, but about owning a robot.

The following year, in 2010, Coeckelbergh (2010) published what would become a massive work for many subsequent publications. In it, he describes a variety of well-known arguments for the moral consideration of a robot and their problems, partly also as in the case of the previously mentioned works. These arguments include comparisons to animal rights, general emancipatory rights such as those of slaves or women, rights of children, property rights, and many more.

Nevertheless, for him all these arguments are null and void and instead he proposes an alternative, socio-relational view (Coeckelbergh, 2010). An examples of this would be Latour's actor-network theory (Latour, 2007), that states that actors (both human and non-human) are interconnected. The author also criticizes the notion of "robot" rights, as it does not offer a uniform understanding. He thus argues for a moral consideration of robots at a early stage and identifies the need to research more in the direction of responsibility, equality and justice.

Fittingly, back in 2015, Ashrafian (2015) even wrote about what moral responsibility for robots arises from the discussion about their rights. In doing so, he explains the emerging responsibilities on the basis of the principles of determinism as well as libertarianism and compares them with animals, like many authors before.

A few years later, in 2017, Vadymovych (2017) described in his work above all the derivation of robot rights from the field of human rights, which is particularly interesting with regard to the already mentioned emancipatory rights.

But at that time, there were two authors in particular whose work had a decisive influence on the discourse on robot rights. These are Gunkel (2018a) with his work "Robot Rights", and Turner (2018) with "Robot Rules: Regulating Artificial Intelligence".

In his book Gunkel examines the possibility of granting rights to robots on the basis of David Hume's philosophical separation of "is" and "ought". He argues that for an ethical understanding of robots one should focus on a relational perspective of their agency and status rather than an ontological view, referring to the philosopher Emmanuel Levinas.



Turner's book, on contrast, emphasizes the uniqueness of AI and the resulting legal challenges. Rather than making specific regulatory proposals, Turner explores the difficulties of creating rules for AI and emphasizes the need for new institutions and approaches. His book should therefore also act as a guideline for the dealing with AI.

He highlights four reasons that could speak for an attribution of certain rights:

"

1. The ability to suffer
2. Compassion
3. The value of something
4. Situations where humans and AI are combined (Turner 2019: 145)

" (Schröder, 2021, p.195)

Since then, however, there have also been a number of new publications that follow on from the work mentioned so far. Such as the work of Lima et al. (2020), who conducted a survey on public perceptions of robot rights, in which they advocate to include all people and not only experts in the discussion to support a public discourse.

According to their results, the majority of the population is open to, if not in favour of, the issue. Although, they also say that how such rights are communicated to people is highly significant. They present a set of rights that can serve as a good example of what types of rights might be involved in the discussion about robot rights.

Coeckelbergh also published another paper in 2021 together with Jana Parviainen about the robot Sofia and the accompanying production (Parviainen & Coeckelbergh, 2021). In doing so, they place the topic of robot rights and Sofia in the context of AI politics and economics. They believe that Sofia was there to strengthen the social robot market and show an economic interest behind robot rights. This economic interest behind robot rights is also described by other authors, such as Turner (2018) or Birhane & Van Dijk (2020).

Another significant contribution comes from Schröder (2021). He examines the topic of robot rights from the perspective of legal philosophy and discusses the difference between legal and moral law, describing in detail the problem of the concept of law. Furthermore, he describes the change in the view of robots from an ontological stance to a relational perspective. What is particularly interesting is that he also discusses the Japanese Koseki and the actual first robot with rights, Paro.

The author argues that robots are artefacts of our society and are subject to human control. Robot rights are for humans and not for robots, just as in the case of Bryson (2010), who believes that robots should be called slaves. The author also posits a Rawlsian perspective in which a "overlapping consensus" (Rawls, 1993: 133-172)" (Schröder, 2021, p. 200) is to be found that resembles a type of continuum, which can be compared to the continuum mentioned by Asaro (2006).

Similar to Lima et al. (2021), De Graaf et al. (2021) conducted a survey in 2021, to outline the public opinion on the topic of robot rights in order to create a holistic understanding. Their findings suggest that, the public willingness to grant robots rights hinges more on beliefs about their potential "cognitive and affective capacities" (De Graaf et al., 2021, p.1), than on demographic or experiential factors.

In the same year, Müller (2021) also publishes an article in which he takes up the arguments of Gunkel and the relational perspective and comes to the conclusion that there is as yet no flawless justification for granting machines a moral status. So Müller believes that they have no moral status today. But it is still very important to talk about the issue of robot rights.

In 2022, Gunkel (2022a) published two more articles. In the first, he reiterates the importance of the issue of robot rights, and again takes a neutral view on the arguments for and against. In doing so, he also refutes the false assumption that the concept of rights must thereby necessarily be linked to that of human rights.

Thereby he also considers other publications, such as that of Lima et al. (2021), Asaro (2006), De Graaf et al. (2021), or Inayatullah (2001) and many other alternative approaches. One of these alternative approaches is e.g. to understand rights as rites in the context of robot rights.

The author explores this very idea in his second article (Gunkel, 2022b), "Should Robots Have Standing? From Robot Rights to Robot Rites", in more detail. Hereby Gunkel explains that rights tend to describe "individual entitlements" (Gunkel, 2022b, p. 5) while rites refer to "social performance and status" (Gunkel, 2022b, p. 5). Like Schröder (2021), he is more concerned with the humans than the robots. Furthermore he proposes the view of robot rights as robot rites in order to avoid a Western-influenced understanding of rights and moral status in this debate.

Last but not least, there is also a paper by Flattery (2023), in which he discusses Kant's indirect argument for non-sentient entities as a reason for granting robot rights and explained a moral argument not to create sentient robots in the first place.

In the future, to be more precise, shortly after the completion of this work, another book by Gunkel (2023) will also be published, called "Person, Thing, Robot: A Moral and Legal Ontology for the 21st Century and Beyond". It can be assumed that this work will also have a significant impact on the debate around the topic of robot rights.

So while most of the authors, pro or anti robot rights, agree that the discussion itself is an important one, there are also authors who see that differently.

Birhane and van Dijk (2020) e.g. describe in their article how a discussion about robot rights distracts from other much more important issues, framing it as a first world problem (Gunkel, 2022a). These include, for example, the introduction of intelligent machines increasingly in oppressive conditions. Perspectives like these make the topic of robot rights a quite controversial one (Lima et al., 2021).

In conclusion, there is a innumerable variety of perspectives and approaches to address the issue of robot rights. Be it from a legal, philosophical, or even economic perspective. Each of these perspectives has its own concepts, theories, advantages and disadvantages.

Nevertheless, it has become apparent that the general discourse is mainly determined by a few key publications and authors. But the topic of robot rights seems to be gaining more and more traction in academia, and more and more researchers are dedicating themselves to the topic. It remains fascinating to see how the area will develop from an academic perspective.

After understanding the overarching views in research and academia, the subsequent phase focuses on analyzing interviews with current field experts to ascertain their current standpoints on robot rights, specifically in relation to the RQ.

## **5. Analysis**

In order to answer the RQ: "Which role do the concepts of consciousness, agency, action, post- and transhumanism play in shaping the debate over the recognition of robot rights for humanoid robots powered by artificial intelligence?", a series of expert interviews were conducted with leading technology philosophers and academic staff, which are examined below. The analysis will be based on the clusters described in the research design in order to ensure a clearer analysis. The square brackets contain the direct references to the experts' statements, with the number of pages (p.) indicating the transcription page and the sections (s.) the respective numbered sections of the transcripts (see appendix).

### **5.1 Consciousness**

#### **Defined**

The concept of consciousness, as already described in the theoretical part, is not an easy concept to define. Various experts also felt that the question of consciousness as a basis for robot rights would depend heavily on the definition of the concept. One explained that "there are a myriad of definitions of consciousness. It is debatable whether a distinction should be made between consciousness and self-consciousness, or whether every form of consciousness presupposes self-consciousness. One must therefore clarify in advance which definition of consciousness one wishes to apply" [p.67, s.108]. In this context, many experts also spoke of the granting of a form of consciousness to machines on the basis of their humanization ... "Robots appear human because they imitate humor in this sense or (...) have a human face (...), speak of having feelings" [p.6, s.8]. Some even said that the concept of consciousness in the context of humanoid robots and robot rights is a "misplaced metaphor" [p.30, s.37]. The question came up, "whether this, the concept of consciousness, is actually decisive for the subject matter ... as a very general concept" [p.40, s. 60].

So the definition of consciousness is complex and varied among experts, influencing their perspective on its relevance to robot rights. While some believed that the human-like appearance and behaviors of robots can suggest a form of consciousness, others argued that the concept might be an inappropriate metaphor or even irrelevant when discussing rights for robots.

#### **Requirements**

Further questions about the concept of consciousness revealed that the prerequisites for consciousness were also defined in completely different ways. On the one hand,

there was the statement that for a form of consciousness, a certain form of imagination would be needed ... "To be more precise, in order to be able to reflect in general, and thus also to be able to speak of consciousness at all, at least a linkage of imaginations is needed" [p.65, s.101]. Such a form of imagination has also been described as a "reflection of reality" [p.16, s.22].

Furthermore, the feeling of pain was mentioned as a prerequisite too, as "an entity needs consciousness or the ability to feel and suffer in order to have moral rights" [p.70, s.120]. But there was also talk about consciousness depending on the perception of the environment ... "Consciousness is always related to an environment in which the conscious subject or entity is situated" [p.53, s.82].

The criteria for consciousness therefore also varied among interviewees, with some asserting that it necessitates a form of imagination or a "reflection of reality." Others emphasized the capability to feel pain as essential for consciousness, while some associated it with an entity's environmental perception and interaction.

### **Types of consciousness**

In addition to the various prerequisites, a number of types of consciousness were also identified. These included, among others, an unrealistic image of consciousness derived from science fiction ... "Now hardcore Silicon Valley people who say, yes, it's all just mathematics in the head, it's all just chemistry, it can all be coded, (...) and that's why we can also upload our consciousness at some point" [p.3, s.5] as well as that "there are levels of consciousness" [p.16, s. 22].

Furthermore, the concept of consciousness has often been described as a purely human concept. One expert argues that "information comes from outside, only comes in (...), via cameras, sensors (...), monitors. That is completely different with humans. We have a completely different stimulus processing" [p.17, s.22]. From this point of view, there is a big difference between a human and other forms of consciousness.

During the interviews, also forms of non-human consciousness and mechanical consciousness were described. Thereby non-human consciousness was often equated and compared with animals, with e.g. a cat. One expert explained: "Even if the cat never fulfills the criteria of being a person, it does not mean that I can do everything with it" [p.41, s.63], whereas mechanical consciousness is a consciousness that can be acquired specifically from machines. This has been described, for example, as a "technical consciousness in the sense of being informed" [p.17, s.22].

This cluster revealed a spectrum of perspectives on consciousness, ranging from those who believe it can be fully quantified and replicated, to those who view it as an inherently human trait distinct from machine processing. Notably, distinctions were made between human, non-human (often likened to animals), and mechanical consciousness, which was described as a form of technical awareness. But regardless of how consciousness was defined and the associated assumptions, most participants, barring a few, believed that a certain degree of consciousness is essential for granting any rights to robots.

## **5.2 Agency**

### **Defined**

Just as in the case of consciousness, it quickly became clear in the case of agency that here, too, the definition has a decisive influence ... "So it depends very much on how (...) demanding one defines the term in each case" [p.38, s.56]. One expert also warned that without a clear definition of agency, further treatment in the context of humanoid robots and thus also robot rights would not be very useful. He "would advocate being very careful about applying the word agency or even full agency to robots" [p.15, s.18].

As in the the case of consciousness, in the case of the agency, the argument again emerges that one could attribute agency to a robot on the basis of a form of humanization ... "So this is not just technology, but this is a technology with distinctive human features, if you like, is the assumption (...) that it is an agent, right?" [p.32, s.39].

The term full agency was thereby defined as a separate concept of agency and during the interviews a clear distinction was drawn between it and the general concept of agency. One expert explained: "So I would hesitate at first whether I would really do or say "frank and freely" like that (...), full agency" [p.14, s.18]. There were different opinions on whether it would even be possible for a robot to achieve a form of full agency. Another expert was of the opinion that this would create "a demand on a machine that it can't fulfill" [p.67, s.106].

Therefore, the concept of agency, like that of consciousness, depended on its precise definition, with experts cautioning against applying it to robots without clarity. While agency could be attributed to robots due to their human-like features, the notion of "full agency" stood distinct, with skepticism regarding whether a machine could ever truly achieve it.

## Requirements

In addition, some preconditions could also be identified in the course of the subject of agency, just as in the case of consciousness. On the one hand, in many cases the notion of agency was associated with some form of intentionality. One expert said: "That is why I would see the essential difference in actorhood or agency in this intentionality" [p.38, s.56].

Another one said: "So I personally use the notion of agency, which is (...), let's say, intentionally filled" [p.52, s.81]. The point here is that the entity pursues a certain goal ... "Now, if a robotic entity were to be ascribed a full capacity for action, this would mean that this entity performs actions willfully with the assumption that these actions are goal-directed" [p.64, s.99].

On the other hand, the state of autonomy was also described in the course of the concept of agency as a form of presupposition. One of the experts explained: "An agency has (...), it comes with autonomy, right? It's an autonomous agency with capacity for autonomous action and autonomous thinking and so forth" [p.28, s.35]. Another expert was saying: "At the same time, of course, we observe something like a certain autonomy in (...), yes, let's say (...), orientation and maybe interaction with others in robots" [p.52, s.81]. One of the experts also confirmed that "he or it can be autonomous in the technical sense" [p.70, s.118].

So the concept of agency in the context of robot rights was often linked to intentionality and autonomy by the experts interviewed. Both, intentionality and the concept of autonomy were viewed as a foundational prerequisite for any form of agency, even if that autonomy is merely technical in nature.

## Additional findings

In addition, several other discoveries were also made in the area of agency with respect to the RQ. First of all, there was found a difference between the German term "Handlungsvermögen" and the English term "agency". This difference was described by an expert by saying: "So we are probably talking about what the English word of agency (/) ... Whereby English does not distinguish there as clearly as German" [p.14, s.18].

Moreover, a large number of experts have also spoken out directly against ascribing agency in relation to humanoid robots or machines. One of the experts commented: "I would not attribute full agency to a machine in the human sense, neither at the present

time nor in the future, because I do not see (...) that all these (...) additional factors that are important for human agency can somehow be replicated in machines" [p.15, s.19] .

Furthermore, another experts said: "I cannot understand it from the perspective of an agent that we would get, for instance, a robot that (...) requires a freedom of speech (...), or the freedom of association, that a robot will establish a labor union without being persecuted (...), or that a robot would get voting rights. That's sort of (...) utopian visions" [p.32, s.39]. Thus, many of the experts were against the assumption of any form of agency, let alone full agency in the field of machines or robots.

So this cluster revealed linguistic nuances between the german term "Handlungsvermögen" and the english term "agency". However, a significant portion of experts expressed skepticism about attributing any form of agency or full agency to robots, viewing such ideas as utopian and believing that the complexities of human agency cannot be entirely replicated in machines.

### **5.3 Action**

#### **Defined**

Like the terms before, the concept of action was again associated with the human being, leading to the questions: "Then again, what does that mean? Does it mean that a robot is acting like a human (...), or is a human acting like a machine" [p.29, s.36]. Nevertheless, the greatest distinction and thus also delimitation of the term, was made in the distinction between an operation and an action.

In most cases, an operation is a process that can be performed by a machine, and an action is a consciously performed action that, according to most experts, can only be attributed to a human being. One expert explained: "So I think that operations is something you can say about machines, also about automata, especially since with humans you wouldn't really talk about operations" [p.16, s.20]. Another expert described that "operations are very much internal, right? The action is what comes out of it" [p.29, s.36].

A third expert was "talking about people acting and machines performing operations or operating" [p.62, s.96]. Lastly, another expert explained that: an "operation refers to the process by which a device performs its intended function, while action refers to an act performed to achieve a specific goal" [p.64, s.100]. One of the experts even highlighted this distinction between action and operation as "a super distinction" [p.16, s.21].



With that, this cluster highlighted a significant differentiation between the terms "operation" and "action". While "operation" typically was referred to processes performed by machines or devices, "action" was largely seen as a conscious act exclusive to humans, emphasizing the intrinsic human element in the concept of action.

## **Approach**

Besides the distinction between the notion of an action and that of an operation, two main other approaches could be identified in relation to a definition of action. One of them is the action of a robot or a machine as a programmed action. One expert was saying that an "action in this living environment, is then none, which, let's say, occurs to the robot itself, but (...) it is programmed into them beforehand (...)" [p.54, s.82]. Thus, the action of a robot is defined as a prescribed or programmed action.

The other one is the approach to connect an action with the idea of a purposefulness and some of the experts would have even granted an action to a robot from this point of view. They were of the opinion that: "It is enough if we see something (...) directed (...) towards a goal, moving or calculating towards or doing some action" [p.2, s.3]. One expert was also of the opinion, that "autonomous systems and agents can in any case perform certain actions to achieve certain goals" [p.70, s.119].

This cluster revealed two primary perspectives on robot actions. Firstly, the one that defined the action of a robot as strictly prescribed or programmed, suggesting they lack innate intentionality. Secondly, the one that states that if a robot's behavior appears goal-directed or purposeful, it could be regarded as them performing an action towards achieving certain objectives.

## **5.4 Humanism**

### **Defined**

In the field of humanism, the focus was primarily on theories of post- and transhumanism. Here, some experts drew a clear distinction between the two terms, while some also connected them in a way. One expert elaborated: "But you are already right, what I actually wrote about, also in the case of cyborgs, that is clear (...), transhumanist and post humanist" [p.21, s.25]. In this context, posthumanism is on the one hand often described as "a movement or thinking that questions the traditional boundaries of what it means to be human and the centrality of humans in the world" [p.65, s.102]. Transhumanism, on the other hand, is defined in most cases as "human enhancement" [p.7, s.10].

However, both the perspective on posthumanism and the one on transhumanism are explained from a wide variety of directions. These include e.g. Latour's "actor network theory" [p.57, s.85], or a form of "new materialism" [p.7, s.10].

In addition, a perspective focused on Silicon Valley also emerged, in which, from a transhumanist point of view, it is a matter of "elevating humans to a new evolutionary species, a newly evolutionary status" [p.7, s.10]. Also in this case, the two terms of posthumanism and transhumanism are not clearly separated from each other, as one expert spoke of "a Silicon Valley posthumanism, which just goes more into human enhancement" [p.7, s.9].

This cluster therefore shed light on the nuanced distinctions and interconnections between the concepts of posthumanism and transhumanism. While posthumanism challenges traditional boundaries of human identity and its significance, transhumanism emphasizes human enhancement, often influenced by perspectives emanating from Silicon Valley that blur the lines between these two ideologies.

## **Consequences**

In addition, some consequences in the course of a post- or transhumanism view on the topic of robot rights and humanoid robots could also be identified. These are primarily three consequences. First, it was assumed by some experts that the question of robot rights will lead to many new questions about humans and their self-perception. One of the expert asked: "So how do we humans think about ourselves? How do we understand ourselves?" [p.56, s.85]. Such perspective could possibly even involving "redefining 'life' itself" [p.65, s.102].

Besides that, also consequences were described, which talked about comparing the granting of robot rights from a post- or transhumanist point of view with the granting of animal rights and a perhaps related degradation of human rights. An expert among the group explained: "But with animal rights you (...) change? Yeah, the status of humans (...) decreases vis a vis animals. Right? Which was (...) you're no longer simply the owners of the animals (...)" [p.31, s.39]. Such a perspective was often used to describe the argument that granting rights to robots could lead to a reduction of rights for animals or even humans.

Although, in the end, there was also the view that the post- or transhumanist view on the topic of robot rights will not change much. One expert described: "In this respect I would think that there is no fundamental difference, but that simply a few additional beings with rights are added, as every year again and again new humans with new

rights are born" [p.43, s.65]. However, another warns in the course of that against an "anthropomorphization" [p.68, s.110] of robots, which is already happening in part.

This cluster revealed divergent opinions on the impact of post- or transhumanist views on robot rights, ranging from redefining human self-perception to concerns about diluting human or animal rights, with some experts arguing for minimal change to the status quo from a post- and transhumanist perspective.

## **5.5 Robot Rights**

### **Defined**

In the further course of the interviews, a number of general approaches to the topic of robot rights could also be tapped into. Almost all interviewees confirmed e.g. that they had already dealt with the topic of humanoid robots and robot rights, even if only "not in direct contact (...), but via scientific discourses" [p.28, s.34]. Only one of the experts surveyed had only "heard about it" [p.63, s.97].

Moreover, some main approaches and persons related to the understanding of robot rights were also identified. These include, for example, "the relational approach" [p.20, s.24], which is about the relationships between entities, or the "property approach" [p.19, s.24]. When it comes to the people, persons as "David Gunkel" [p.24, s.28], "Frank Pasquale" [p.25, s.28], or also "Emmanuel Kant" [p.20, s.24] could be identified.

Furthermore, a distinction between legal law and moral law was made. Whereby this distinction was also used in the course of this work. One of the experts asked: "What are we talking about when we talk about robot rights? Are we talking about legal rights? Are we talking about moral rights, or are we talking about both?" [p.18, s.23].

In addition, some well-known legal frameworks were also described, which currently clarify legal and moral issues in the course of robots. One expert explained: "Just on the EU level there is this AI Act and Digital Service Act" [p.12, s.15].

Ultimately, however, the actual discussion around the topic of robot rights and the meaningfulness of the discussion was also questioned. One expert expressed the opinion: "we have so many big other problems, this is not a problem we have right now" [p.5, s.6]. At the same time, it could be observed that a much more fundamental question was often asked like: "Do we want this (...)? What do we want to do with it?" [p.55, s.82]. One of the experts even called robot rights "nonsense" [p.71, s.124].

Most of the experts doubted whether it would make sense to ultimately introduce robot rights, even if they consider the discussion about the topic itself to be meaningful.

So most experts had engaged with humanoid robots and robot rights, referencing approaches like the "relational" and "property" approach. While some frameworks like the EU's AI Act were mentioned, the overall validity of robot rights was debated, with many viewing the topic as either non-urgent or moronic.

### **Additional Findings**

Furthermore, it was found that many experts have problems with the concept of AI. Especially with regard to the Turing Test and the associated attribution of intelligence, not least because it is an "epistemic test" [p.47, s.71]. One of the experts claimed that "the word intelligence is wrongly chosen by Alan Turing at the time" [p.29, s.35]. But also the definition of the word robot could be identified as an very important point, as another expert explained: "Before the question of robot rights can be addressed, it would therefore be useful to first clarify what is understood as a robot" [p.63, s.97].

In general, it could be seen that the topic of robot rights and related principles are very much linked to the world of science fiction, as "there is a lot of science fiction that goes through a lot of different points very finely" [p.60, s.94]. But the subject is also shaped by other "power structures [p.8, s.11], such as "political and economic forces" [p.34, s.41] and "military power" [p.34, s.41].

In the end, then, it is fair to say that most experts have come out against the introduction of robot rights. Be it due to a lack of "criteria of personality" [p.47, s.69], or also the fact that "these are questions that concern people, not robots" [p.10, s.12]. Two of the interviewees were even strictly against granting any rights to robots [p.69, s.116; p.71, s.124]. One of the experts, however, was of the opinion that, in the sense of the "in dubio pro reo" approach, it would be better to grant robots rights, insofar society doesn't know whether they should be granted any.

The analysis revealed that many experts challenge the foundational concepts of AI and robot definition, associating much of the discourse with science fiction, political, economic, and military influences. The majority of experts oppose robot rights, with a one advocating for them under the principle of "in dubio pro reo".

## 6. Discussion

Now, in the course of the preceding analysis, the information drawn from the interviews is put into the overall context of the work and, in particular, of the theoretical background.

It quickly became clear that the definition of consciousness is crucial, as most of the experts were of the opinion that you need a certain form of consciousness to even think about giving rights. But there is only one problem with that, depending on the definition one could already assign consciousness today. If one looks back for example to the Integrated Information Theory (IIT) of Tononi (2004) (system's ability to assimilate information), a robot would oftentimes be conscious. Likewise if one consults the numerous definitions of machine consciousness, that could be compared to the non-physical consciousness by Van Gulick (2014).

Furthermore, in the context of machine consciousness, a possible connection to a form of environmental perception could be observed. The descriptions of the perception of the environment as a prerequisite for consciousness, agree with many of the definitions and descriptions of machine consciousness. This context can also be related to the understanding of a robot by Hegel et al. (2014), who defined a robot partly by its perception of the environment.

One expert suggested discussing the concept of personhood instead of consciousness, since consciousness is a very general concept. Also the questioning of the experts showed, that all brought along a different understanding of consciousness. But nevertheless, it is also important to note that the previously presented humanoid robots, Sofia by Hanson Robotics (2023) and Paro, mentioned by Schröder (2021), were not granted any form of personhood, but only a few civil rights. These have nothing to do with the state of personhood and are far away from the sophisticated understanding of robot rights in the course of this work.

The question that still remains open in the course of this in any case is which rights could ultimately be given with which form of consciousness. This also includes the various implications of the right, such as responsibility, which has not been considered at all in this work, apart from Ashrafian (2015). Therefore, this field of consciousness in relation to rights and also responsibility definitely needs further academic research and discourse.

What is also striking is that the argument of humanization, runs through almost every interview. That is, the assumption that we grant robots a certain form of humanity due to their human appearance or abilities. However, in most cases this humanization was not necessarily described as a reason for granting robots rights. Rather, it was used to describe the feelings and emotions that we have and can develop towards robots.

This humanization, or rather the depiction of the human body in the form of robots, also led to another topic that was discussed in the course of many interviews: the love of robots, or even sex with robots. Let's just take the example of Sofia by Hanson Robotics (2023), what legal implications could it have if someone wanted to marry her? Have sex with her? Or would want to get a divorce (Ryznar, 2018)? These are topics of the future, which need to be further researched, especially in relations to robot rights.

Also in the course of the topic of agency, it quickly became clear that the definition is crucial. But, here there were no so big differences between the experts as in the case of consciousness. What stood out most was the review of Moor's (2006; 2009) four described forms of agents, which were more or less described during the interviews, although not as a direct concept. Thereby the first three forms of agents can already be found in reality today. What does not exist yet is the 4th form of a full ethical agent. This statement, described by Moor, was confirmed by most experts.

Nevertheless, the experts were divided on the question of whether this form of full agency will be possible in the future. But judging by all current forecasts, e.g. the surveys made by the WEF described by Roser (2023), it can be assumed that such an agent could exist in the future. Most experts were also of the opinion that if one could speak of a full agency like Moor (2006; 2009), one would have to speak definitely about robot rights. Considering that most of them were against the introduction of robot rights, or said that society does not need them yet, a rather interesting perspective. From this, one could conclude that perhaps concrete robot rights should already be discussed now, following the presented in dubio pro reo argument and assuming that the form of full agency is once achieved.

But admittedly, it is questionable in general whether the concepts of consciousness and agency are of much help in the question of robot rights. On the one hand, they are both very difficult concepts to grasp, which was e.g. already described by Gunkel (2018b) and Levy (2009) in the case of consciousness, and on the other hand, they are also difficult to define. This makes the granting of certain rights almost impossible. Ultimately, if society would use such concepts as a basis, the people would have to agree on a definition, which seems to be quite difficult.

In the course of the concept of action, it was possible above all to confirm the distinction between an operation and action, which was already described in the theoretical background with help of the Cambridge Dictionary (2023) and WikiDiff (2023), in particular the distinction from a human action. However, this also makes it questionable, whether robots can act at all in this sense, even if they have a goal-oriented function or their action is programmed in such a way that it can hardly be distinguished from a human action.

Purely code-related statements, as well as considerations of normative theories in coding, as described with the help of Lin et al. (2016), hardly came up in the course of the interviews and in relation to the question of action. It was only confirmed that these could also have an influence on the final actions of a robot.

Although, this also makes it very questionable, just as with the concept of consciousness and agency, how much the question of action, or its definition as an operation, could really help in the ultimate question of robot rights, whereby the distinction definitely requires further research and was also supported by the experts. Once again, the discussion around the topic is good and important, but its practicality is questionable.

This also applies to the field of posthumanism, where the border to transhumanism alone was very difficult. A posthumanist perspective opens up many possibilities for discussing life and humans themselves, as described by Bostrom (2005), but it is again questionable to what extent this might have relevant consequences for the current question of robot rights. Many of the experts considered such a perspective to be quite irrelevant in direct reference. At most, it is about the consequences of introducing robot rights and the related question about being human and living.

Ultimately it is about the question of which entities we want to grant rights to, and questions related to this would probably be much more helpful to address the question of robot rights, than the concepts examined. As it turned out, the selection of philosophical principles drawn in the course of the RQ and in light of the interviews conducted helped only slightly when it comes to the question of robot rights. This is partly because of the difficult and different terms of the concepts, but because it is also questionable whether the ultimate granting of rights, would depend on such factors. Factors, however, that in any case play into the discussion all around.

Thus, these principles were at least once put into the right context and explained in the course of the subject of robot rights. In the course of further work, these principles must be further examined and their position in the discussion about robot rights must be even more clearly defined.

Nevertheless, in fact, the categories formulated by Turner (2018) in the course of the literature review in turn compare very well with the philosophically chosen principles. Thus, the principles of consciousness, agency, or even action, seem to be quite important in the question of robot rights again. As so often, it is a question of definition and perspective.

So the RQ can in light of this be answered by saying that the principles definitely play a role in the discussion of robot rights, though perhaps not in direct relation to the ultimate granting of robot rights, but in their discourse. This means that these principles definitely help to better understand the issue of robot rights, but at the same time, they are also very limited principles when it comes to clear jurisprudence.

Philosophical considerations are very suitable to understand the topic better, but to be considered as a basis for the question of robot rights, a much more precise definition of the terms is needed, on which everyone would have to agree, which is not that easy.

If some definitions for these philosophical principles were agreed upon and fixed in the context of a framework for robot rights, then it would be quite conceivable that these could also serve as a basis for defining the frameworks. An example of this are the four forms of agents, already explained. Here, and in view of the fact that most experts are still against the introduction of robot rights, one could tie their introduction to the introduction of the full ethical agent.

Because as soon as the full ethical agent by Moor (2006; 2009) comes into being, the experts believe that some form of robot rights will be required, however these may ultimately look in practice. Which in turn raises the question, why not deal with it now? This question is not easy to answer and there are a variety of factors that play an additional role. Some of which could be identified in the further course of this work, the interviews and the related RQ.

This includes, among other things, a great economic interest that accompanies the debate, which has already been mentioned in the literature review by Parviainen & Coeckelbergh (2021). For example the fact that, without robot rights, companies could allow robots to work 24/7 and not grant them any labor rights. At the same time, in the case of the introduction of robot rights, companies could also take the responsibility



and accountability, e.g. in the case of an accident or unfair treatment, away from themselves by saying that this lies with the robot itself. So, interestingly, from a business perspective, there are economic reasons both for and against the concept of robot rights.

This economic interest is also very closely related to the already described approach of emancipatory rights by Coeckelbergh (2010). Thus, rights were reserved for slaves, women or even animals, often out of an economic interest and with the justification that one would know better and it would be the best for them ... was it really? Yet each of these "new" rights was unthinkable in the past. Another good example of this would be the comparison with children rights, which Inayatullah (2001) described. They are not granted certain rights despite being human, and something like children's right to vote would have been unthinkable decades ago.

In addition, the topic of robot rights, or AI in general, and thus also humanoid robots, is a topic that is dealt with very little in Europe. Europe is very much behind America and Asia in this respect, as already explained by Bangali (2023). The major developments in these areas are largely taking place in the Silicon Valley or in China. The Silicon Valley consists of the leading companies in this field and often also provides the philosophical framework for their inventions.

In the course of Facebook's transformation into Meta, for example, Marc Zuckerberg also formed his strategy on the basis of an essay that was ultimately distributed to all employees (Ball, 2020).

Although, it is also worth comparing this with Japan, which, as already outlined with the help of Paro (Schröder, 2021; Koudela, 2019), has a completely different relationship to robots than Europe or America. In Japan, robotics seems to be almost part of the culture and the typical skepticism towards the topic is not as strong as it is here. Let's take, for example, the registration system Koseki from Japan. In the course of the question about the regulation of robotics and robots, this could be a good starting point for further discussions. In such a registration system, all robots could e.g. simply be entered into a register. A system or register like that, which grants a better overview and understanding could then help to shape the discussion about robotic frameworks (Jones, 2018).

Tuner (2018) talks in his book about the fact that by granting robot rights in one country, a domino effect could break out and many other countries could take a look at that and would be inspired to do the same. But there's the question why this has not already happened in the case of Sofia or Paro. Maybe because Sofia was just a

marketing move as outlined by Parviainen & Coeckelbergh (2021)? Because hereby it should also be noted, that the robot Sofia was granted rights in a country that still reserves many rights for women, despite the fact that Sofia is probably also a woman.

What is needed, therefore, is probably more of a global approach. The creation of an international framework aimed at avoiding nationally specific disagreements in this context. Indeed, similar to climate change, this issue is also of international interest, as shown by the UNESCO (2023) and the OECD (2022), and therefore requires a unified and international solution. Whereby it has already been described how difficult this will be due to national differences by Lin et al., (2011) as well as Guo & Zhang (2009). But this is no different in most international disputes.

And for this, as already described in the literature review by Lima et al. (2020), a public debate is needed that involves not only academic voices, but everyone. After all, it is a problem that really affects everyone, even if many are not familiar with the subject. This definitely needs to be changed in the course of a strengthened discourse and a stronger public debate.

Ultimately, however, the question really arises as to whether we want the whole process at all? In other words, the development and invention of robots to which we ultimately have to grant rights. The question is perhaps whether it is about a reference back to the human-machine relationship and to what extent we want to interact with such machines and develop them? This view was shared by most of the experts interviewed, especially if they came from the field of philosophy.

It is therefore necessary to question precisely this human-machine relationship and to continue to incorporate it scientifically into the discourse around robotics, AI and general technology. The question about robot rights is probably a much too advanced question, if mankind cannot even answer the question whether one would like to build such robots at all.

In general, nevertheless, it must be said that the topic of robot rights itself requires further research and investigation in almost every respect. Be it philosophical principles, the question of which entities society generally grants rights to, what rights and what exactly they would ultimately look like, or even the question of responsibility if a robot should be granted rights. These areas, and many more, all require further research and it remains to be seen how the topic of robot rights will develop in the coming years. Especially in view of the current rapid development of AI and robotics.

## 7. Limitations and challenges

So in the course of this master thesis, it was possible to conduct a series of high-quality interviews with people qualified to do so, to answer the RQ. In addition, it was also possible to make a variety of other scientific discoveries on the subject of robot rights. But like any research, this paper has its limitations and restrictions.

One of these is that only few interviews were conducted in the course of the expert interviews. On the one hand, this was due to the limited time frame, and on the other hand to the difficulty of finding adequate interviewees for such a specific and still rather unexplored area. Moreover, most of the people were on vacation a lot during the summer months and therefore difficult to get a response from. In view of this fact, it was even possible to conduct a large number of really high-quality interviews. Although, it has to be said that the questions and the conduct of the interview could definitely have been worked out better. For example, very wide-ranging questions were asked in the course of the interviews, which were ultimately very difficult to evaluate, and one should probably have given more thought to the subsequent analysis when formulating the questions.

In addition, only one interview could be conducted with a woman and thus male perspectives were very strongly represented. In the case of a further study, this gender gap in the area of interviews would have to be attempted to be closed. However, this is also due to the fact that the field is still very male-dominated. In addition, most of the interviews, except for one, were conducted in German. Since this work was written in English, it was difficult to combine German and English content. Unfortunately, there was not enough time to translate the interviews completely into English, which would have been the easiest way. Instead, the German content that was directly referred to from the transcriptions was simply translated, which makes the academic follow-up more difficult than it should be. Thus, the whole analysis of the interviews does not seem fully coherent sometimes and may even have some flawed translations.

The topic itself and also the RQ were also way too broad. In the course of this work, the attention was only put on four certain philosophical principles, and also directly on humanoid robots. One would have to define the question much more precisely and focus only on parts of the question.

For example, it would have been a possibility to deal only with the three to four laws of Asimov or a single argument for robot rights. By doing so, the paper definitely lost coherence and focus at certain points due to its broad scope in the course of the RQ.

In addition, technical and clear legal perspectives are hardly discussed in the course of the work. Nevertheless, this was also deliberately omitted, even if these should be important competences for the question of robot rights.

On the one hand, this would have made the thesis even more difficult to understand than it already is, on the other hand, the inclusion of these components would have really gone beyond the scope of the thesis. It was already quite difficult not to let the work become too long and confusing, which would have been made much more difficult by the connections with legal and technical parts. Nevertheless, these are essential factors in the question of robot rights and a further treatment of them would be interesting, especially in terms of academic research.

As well as the reference to the theme of responsibility. This topic is one that emerges directly from the discussion about robot rights. In this paper, not too much attention has been paid to it, but that is ultimately what the introduction of robot rights is all about. With certain rights come duties, but what exactly would be the benefit, for whom exactly, and under what circumstances? These questions need to be clarified in the future.

In the end there were many other concepts and perspectives that were somewhat lost in this work. Due to time and textual constraints, it was simply not possible to address every concept and describe every perspective. Especially since the interviewees, e.g., also brought their own theory and perspective to each topic. Therefore, concepts such as intentionality and autonomy e.g. were not addressed, even though they are really important topics in the context of robot rights. So even though many points could be taken up, many also had to be neglected, which shows the complexity of the question.

## 8. Conclusion

This master thesis is primarily concerned with the question of robot rights. More precisely, with the philosophical principles of consciousness, agency and agents, action, as well as post- and transhumanism. For this purpose, the theoretical background to answer the question was established (theoretical part) and a literature review as well as a series of expert interviews were conducted (empirical part), which were later placed in the context of the topic of robot rights and the RQ: "Which role do the concepts of consciousness, agency, action, post- and transhumanism play in shaping the debate over the recognition of robot rights for humanoid robots powered by artificial intelligence?"

To recap, the RQ was answered insofar as the selected philosophical principles only play a limited role in the course of robot rights for humanoid robots, always depending on the definition and understanding of the term. They could be identified as important factors in the discussion around the topic of robot rights, but they are only of limited help when it comes to the question of concrete robot rights. The terms are simply too vague for this.

Although, with the help of this thesis, it was shown that the topic of robot rights is still very much at the beginning of its development and that the scientific basis is also still very limited. The literature to date is oriented towards a few individual persons and the debate can thus still be conducted much more openly and broadly. In the end, the fundamental question remains how we want to proceed with technology and, in particular, robots in the future.

Therefore, it is not so much about the question of concrete rights for robots, but the question of whether we want to build robots at all that should be granted rights. This question represents one of the main themes for further research based on this work.

Furthermore, there are also a number of topics and questions that came up in the course of the work, which could also possibly provide the basis for further research ... What happens, e.g., if you cut off the power to a machine with rights, can this then be considered murder? How does it behave with love to a robot? Or also with intimacies with a robot and/or also the already asked question about its possibility of the marriage and perhaps even adoption? How does the whole topic of robot rights relate to the development of smart cities, or AI cities, and what does the whole topic have to do with data capitalism? How does the issue of robot rights relate to the principle, or concept, of an algorithmic society? And in what way would the issue relate to fighting oppression of other entities?

Moreover, the concept of intelligence also requires further investigation. Not only in terms of the definition of the term, but also the test methods that can be used to determine intelligence in robots, as it seems to play an important role in the question of robot rights and the related philosophical concepts.

Many people take a dystopian stance on the subject of robot rights, or at least have a very dystopian image in mind. This is of course shaped by science fiction representations such as *Ex Machina* or *I, Robot*. But society is still very far away from such science fiction scenarios and the introduction of robot rights is mostly meant to avoid such a dystopia. But these representations ultimately lead each of us to the one particular question ... do we even want something like this?

If not, it would be more a matter of questioning the connection between humans and machines in today's times and especially in light of AI and computer developments and how such dystopian notions can be avoided.

Ultimately, however, it must also be said that an incredible amount is changing in this area. The topic itself, as well as its discussion, can change and develop so quickly. In this work alone, a lot has probably been overtaken in the months of writing and countless new publications have come out that could be relevant to the topic. Nevertheless, this problem is almost impossible to avoid when dedicating to researching new areas of study.

So just as women's rights, animal rights, or rights for slaves once were, robot rights are unthinkable for many people. Nevertheless, as in the other cases, this could change in the light of future developments. It remains interesting to see in which direction the topic will develop and when we will really be able to speak of humanoid robots that would have consciousness or full agency and to which most people would then also give rights.

Lastly, it should also be kept in mind that perhaps one should not wait until a form of full agency or consciousness is reached before discussing and exploring the concept of robot rights further. Given the dangers posed by AI, this may well already be too late and society may indeed find itself in a dystopian scenario at some point, something that very few people would wish for.

## 9. Bibliography

Adams, B., Breazeal, C., Brooks, R. A., & Scassellati, B. (2000). Humanoid robots: A new kind of tool. *IEEE Intelligent Systems and Their Applications*, 15(4), 25-31. <https://doi.org/10.1109/5254.867909>

Aggarwal, K., Mijwil, M. M., Al-Mistarehi, A. H., Alomari, S., Gök, M., Alaabdin, A. M. Z., & Abdulrhman, S. H. (2022). Has the future started? The current growth of artificial intelligence, machine learning, and deep learning. *Iraqi Journal for Computer Science and Mathematics*, 3(1), 115-123. <https://doi.org/10.52866/ijcsm.2022.01.01.013>

Aldebaran (2023). *Pepper*. Aldebaran - United Robotics Group. [Online] Last Accessed 26.09.2023: <https://www.aldebaran.com/en/pepper>

Alesich, S. & Rigby, M. (2017). Gendered robots: Implications for our humanoid future. *IEEE Technology and Society Magazine*, 36(2), 50-59. <http://dx.doi.org/10.1109/MTS.2017.2696598>

Alexander, L. & Moore, M. (2020). *Deontological Ethics*. Stanford Encyclopedia of Philosophy. [Online] Last Accessed 26.09.2023: <https://plato.stanford.edu/entries/ethics-deontological/>

Ambron, R. (2023). Toward the unknown: consciousness and pain. *Neuroscience of Consciousness*, 2023(1), niad002. <https://doi.org/10.1093/nc/niad002>

Anderson, S. L. (2008). Asimov's "three laws of robotics" and machine metaethics. *AI & Society* 22, 477–493. <https://doi.org/10.1007/s00146-007-0094-5>

Asaro, P. M. (2006). What Should We Want From a Robot Ethic?. *International Review of Information Ethics*, 6, 9-16. <https://doi.org/10.29173/irrie134>

Ashrafian, H. (2015). Artificial intelligence and robot responsibilities: Innovating beyond rights. *Science and engineering ethics*, 21(2), 317-326. <https://doi.org/10.1007/s11948-014-9541-0>

Asimov, I. (1942). Runaround. *Astounding science fiction*, 29(1), 94-103.

Asimov, I. (1950). *I, Robot*. Doubleday.

Asimov, I. (1985). *Robots and empire*. Doubleday.

Bahishti, A. A. (2017). Humanoid robots and human society. *Advanced Journal of Social Science*, 1(1), 60-63. <http://dx.doi.org/10.21467/ajss.1.1.60-63>

Ball, M. (2020). *The Metaverse: What It Is, Where to Find it, and Who Will Build It*. MatthewBall.co. [Online] Last Accessed 26.09.2023: <https://www.matthewball.vc/all/themetaverse>

Balkin, J. M. (2017). 2016 Sidley Austin Distinguished Lecture on Big Data Law and Policy: The Three Laws of Robotics in the Age of Big Data. *Ohio State Law Journal*, 78(5), 1217-1241.

Bangali, Y. (2023). *AI made in Germany? What we can learn from Silicon Valley*. Fraunhofer - FutureWork360. [Online] Last Accessed 26.09.2023: [https://www.futurework360.de/en/aktuelles/blogbeitrag-silicon-valley.html?cq\\_ck=1573657273900](https://www.futurework360.de/en/aktuelles/blogbeitrag-silicon-valley.html?cq_ck=1573657273900)

Barton, E. (2014). *Beyond human: Researchers explore transhumanism*. Arizona State University - Knowledge Enterprise. [Online] Last Accessed 26.09.2023: <https://research.asu.edu/beyond-human-researchers-explore-transhumanism>

Bartneck, C., Yogeeswaran, K., Ser, Q. M., Woodward, G., Sparrow, R., Wang, S. & Eyssel, F. (2018). Robots and racism. In T. Kanda, S. Sabanovic, G. Hoffman & A. Tabus (Eds.), *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction* (pp. 196-204). Association for Computing Machinery. <https://doi.org/10.1145/3171221.3171260>

Bencherki, N. (2016). Action and agency. In T. C. Robert (Ed.), *The International Encyclopedia of Communication Theory and Philosophy* (pp. 1-13). Wiley. <https://doi.org/10.1002/9781118766804.wbiect030>

Birhane, A. & Van Dijk, J. (2020). Robot rights? Let's talk about human welfare instead. In A. Markham, J. Powles, T. Walsh & A. L. Washington (Eds.), *AIES '20: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 207-213). Association for Computing Machinery. <https://doi.org/10.1145/3375627.3375855>

Birhane, A., Van Dijk, J. & Pasquale, F. (2021). *Debunking Robot Rights Metaphysically, Ethically, and Legally*. Paper presented at We Robot Miami. [Online] Last Accessed 26.09.2023: <https://robots.law.miami.edu/2021/>



Bostrom, N. (2005). A history of transhumanist thought. *Journal of evolution and technology*, 14(1), 1-25.

Bostrom, N. & Yudkowsky, E. (2018). The ethics of artificial intelligence. In E. Yudkowsky (Ed.), *Artificial intelligence safety and security* (pp. 57-69). Chapman and Hall/CRC. <http://dx.doi.org/10.1201/9781351251389-4>

Boston Dynamics (2023). *ATLAS*. Boston Dynamics. [Online] Last Accessed 26.09.2023: <https://www.bostondynamics.com/atlas>

Braithwaite, V. (2010). *Do fish feel pain?* OUP Oxford.

Braun, M., Breuer, S., Ritt, K. & Tigard, D. (2023). *Roboter*. bidt - Bayerisches Forschungsinstitut für Digitale Transformation. [Online] Last Accessed 26.09.2023: <https://www.bidt.digital/?glossary=roboter>

Brauner, A. S. (2016). *Runaround by Isaac Asimov and the significance of the Three Laws of Robotics in today's world*. Thesis from the Matthias Grünwald Gymnasium. [Online] Last Accessed 26.09.2023: <http://dx.doi.org/10.13140/RG.2.2.16275.14884>

Britannica (2023). *deontological ethics*. Britannica. [Online] Last Accessed 26.09.2023: <https://www.britannica.com/topic/deontological-ethics>

Brown, R. (2023). *moon yet': Human-like A.I. is still some way off, early investor in Google's DeepMind says*. CNBC. [Online] Last Accessed 26.09.2023: <https://www.cnbc.com/2023/05/04/were-far-from-human-level-ai-early-deepmind-investor-says.html>

Bell, E., Bryman, A. & Harley, B. (2022). *Business research methods (5th edition)*. Oxford university press.

Bryson, J. J. (2010). Robots should be slaves. In Y. Wilks (Ed.), *Close engagements with artificial companions: : key social, psychological, ethical and design issues* (pp. 63-74). John Benjamins Publishing Company. <https://doi.org/10.1075/nlp.8.11bry>

BSA (2023). *Comparing International Frameworks for the Development of Responsible AI*. BSA - The Software Alliance. [Online] Last Accessed 26.09.2023: <https://ai.bsa.org/global-ai-principles-framework-comparison/>

Cambridge Dictionary (2023). *Operation*. Cambridge Dictionary. [Online] Last Accessed 26.09.2023: <https://dictionary.cambridge.org/de/worterbuch/englisch/operation>

Card, D. & Smith, N. A. (2020). On consequentialism and fairness. *Frontiers in Artificial Intelligence*, 3, 34. <https://doi.org/10.3389/frai.2020.00034>

Cheng, H. W. (2023). Challenges and Limitations of ChatGPT and Artificial Intelligence for Scientific Research: A Perspective from Organic Materials. *AI*, 4(2), 401-405. <https://doi.org/10.3390/ai4020021>

Crichton, M. (Director). (1973). *Westworld* [Film]. RAD Pictures.

Clarke, R. (1994). *Asimov's laws of robotics: Implications for information technology*. Roger Clarke's Web-Site. [Online] Last Accessed 26.09.2023: <http://www.rogerclarke.com/SOS/Asimov.html>

Clarke, R. (2010). Cyborg rights. In proceedings of a conference (Eds.), *2010 IEEE International Symposium on Technology and Society (ISTAS)* (pp. 9-22). IEEE. <https://doi.org/10.1109/ISTAS.2010.5514655>

Coeckelbergh, M. (2010). Robot rights? Towards a social-relational justification of moral consideration. *Ethics and information technology*, 12(3), 209-221. <http://dx.doi.org/10.1007/s10676-010-9235-5>

Darling, K. (2016). Extending legal protection to social robots: The effects of anthropomorphism, empathy, and violent behavior towards robotic objects. In R. Calo, A. M. Froomkin & I. Kerr (Eds.), *Robot Law* (pp. 213–232). Edward Elgar. <https://dx.doi.org/10.2139/ssrn.2044797>

De Graaf, M. M., Hindriks, F. A. & Hindriks, K. V. (2021). Who wants to grant robots rights?. *Front. Robot. AI* 8, 1-13. <https://doi.org/10.3389/frobt.2021.781985>

Denny, J., Elyas, M., D'costa, S. A. & D'Souza, R. D. (2016). Humanoid robots—past, present and the future. *European Journal of Advances in Engineering and Technology*, 3(5), 8-15.

DiSalvo, C. F., Gemperle, F., Forlizzi, J. & Kiesler, S. (2002). All robots are not created equal: the design and perception of humanoid robot heads. In B. Verplank, A. Sutcliffe, W. Mackay, J. Amowitz & W. Gaver (Eds.), *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 321-326). <http://dx.doi.org/10.1145/778712.778756>

Dresing, T. & Pehl, T. (2018). *Praxisbuch Interview, Transkription & Analyse. Anleitungen und Regelsysteme für qualitativ Forschende*. 8. Auflage. Marburg: Self-published.

Duignan (2023). *Speciesism*. Britannica - Philosophy. [Online] Last Accessed 26.09.2023: <https://www.britannica.com/topic/speciesism>

Eckert, D. & Zschäpitz, H. (2023). *Das wahrscheinlichste Szenario für den Weltuntergang ist nicht der Klimawandel*. WELT. [Online] Last Accessed 26.09.2023: <https://www.welt.de/wirtschaft/plus246402022/Weltuntergang-Der-Klimawandel-ist-nicht-das-wahrscheinlichste-Szenario-fuer-das-Ende-der-Menschheit.html>

European Parliament (2016). *Draft Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))*. Committee on Legal Affairs. [Online] Last Accessed 26.09.2023: [https://www.europarl.europa.eu/doceo/document/JURI-PR-582443\\_EN.pdf](https://www.europarl.europa.eu/doceo/document/JURI-PR-582443_EN.pdf)

European Parliament (2023). *AI Act: a step closer to the first rules on Artificial Intelligence*. European Parliament News. [Online] Last Accessed 26.09.2023: <https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence>

Flattery, T. (2023). The Kant-inspired indirect argument for non-sentient robot rights. *AI and Ethics*, 1-15. <http://dx.doi.org/10.1007/s43681-023-00304-6>

Friedman, C. (2023). Ethical concerns with replacing human relations with humanoid robots: an ubuntu perspective. *AI and Ethics*, 3(2), 527-538. <https://doi.org/10.1007/s43681-022-00186-0>

Garland, A. (Director). (2014). *Ex Machina* [Film]. Universal Pictures International and Film4.

Gellers, J. C. (2020). *Rights for robots: artificial intelligence, animal and environmental law (edition 1)*. Routledge.

Glasscock, J. S. P. & Tenenbaum, S. (2023). *Action*. Stanford Encyclopedia of Philosophy. [Online] Last Accessed 26.09.2023: <https://plato.stanford.edu/Archives/sum2023/entries/action/>

Gorden, C. (2022). *2023 Will Be The Year Of AI Ethics Legislation Acceleration*. Forbes Magazine. [Online] Last Accessed 26.09.2023: <https://www.forbes.com/sites/cindygordon/2022/12/28/2023-will-be-the-year-of-ai-ethics-legislation-acceleration/>

Gunkel, D. J. (2018, a). *Robot rights*. MIT Press.

Gunkel, D. J. (2018, b). The other question: can and should robots have rights?. *Ethics and Information Technology*, 20, 87-99. <https://doi.org/10.1007/s10676-017-9442-4>

Gunkel, D. J. (2022, a). The Rights of Robots. In A. A. Nakagawa and C. Douzinas (Eds.), *Non-Human Rights—Critical Perspectives* (pp. 1-25). Edward Elgar. <https://dx.doi.org/10.2139/ssrn.4077131>

Gunkel, D. J. (2022, b). Should Robots Have Standing? From Robot Rights to Robot Rites. *Forthcoming, Frontiers of Artificial Intelligence and Applications* with IOS Press, 1-6. <https://ssrn.com/abstract=4234507>

Gunkel, D. J. (2023). *Person, Thing, Robot: A Moral and Legal Ontology for the 21st Century and Beyond*. MIT Press.

Guo, S. & Zhang, G. (2009). Robot rights. *Science*, 323(5916), 876-876. <http://dx.doi.org/10.1126/science.323.5916.876a>

Hanson Robotics (2023). *Sophia*. Hanson Robotics. [Online] Last Accessed 26.09.2023: <https://www.hansonrobotics.com/sophia/>

Haring, K. S., Watanabe, K., Velonaki, M., Tossell, C. C. & Finomore, V. (2018). FFAB —The form function attribution bias in human–robot interaction. *IEEE Transactions on Cognitive and Developmental Systems*, 10(4), 843-851. <https://doi.org/10.1109/TCDS.2018.2851569>

Harris, J., & Anthis, J. R. (2021). The moral consideration of artificial entities: a literature review. *Science and engineering ethics*, 27(4), 53. <https://doi.org/10.1007/s11948-021-00331-8>

Hegel, F., Muhl, C., Wrede, B., Hielscher-Fastabend, M. & Sagerer, G. (2009). Understanding social robots. In proceedings of a conference (Eds.), *2009 Second International Conferences on Advances in Computer-Human Interactions* (pp. 169-174). IEEE. <https://doi.org/10.1109/ACHI.2009.51>

Helo, P. & Hao, Y. (2022). Artificial intelligence in operations management and supply chain management: An exploratory case study. *Production Planning & Control*, 33(16), 1573-1590. <http://dx.doi.org/10.1080/09537287.2021.1882690>

IBM (2023). *AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the difference?*. IBM Blog. [Online] Last Accessed 26.09.2023: <https://www.ibm.com/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks/>

IEEE (2023). *Welcome to humanoid robotics*. IEEE Robotics and Automation Society. [Online] Last Accessed 26.09.2023: <https://www.humanoid-robotics.org/>

Inayatullah, S. (2001). The rights of robot: Inclusion, courts and unexpected futures. *Journal of Future Studies*, 6(2), 93-102.

Ishiguro, H. & Nishio, S. (2007). Building artificial humans to understand humans. *Journal of Artificial Organs*, 10(3), 133-142. <https://doi.org/10.1007/s10047-007-0381-4>

Ishiguro, H., Kasaki M., Asada, M., Osaka, M. & Fujikado, T. (2016). *Cognitive Neuroscience Robotics*. Springer Tokyo. <https://doi.org/10.1007/978-4-431-54595-8>

Jones, C. (2018). The Robot Koseki: A Japanese Law Model for Regulating Autonomous Machines. *Journal of Business & Technology Law*, 14(2), 403-467.

Kaplan, F. (2005). Everyday robotics: robots as everyday objects. In G. Bailly & J. L. Crowley (Eds.), *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies* (pp. 59-64). Association for Computing Machinery. <http://dx.doi.org/10.1145/1107548.1107570>

Koudela, P. (2019). Robots Instead of Immigrants: The Positive Feedback of Japanese Migration Policy on Social Isolation and Communication Problems. *Asia-Pacific Social Science Review*, 19(1), 89-102.

Krogh, A. (2008). What are artificial neural networks?. *Nature biotechnology*, 26(2), 195-197. <https://doi.org/10.1038/nbt1386>

Krüger, O. (2021). "The Singularity is near!" Visions of Artificial Intelligence in Posthumanism and Transhumanism. *International Journal of Interactive Multimedia and Artificial Intelligence*, 7(1), 16-23. <http://dx.doi.org/10.9781/ijimai.2021.07.004>

Kurzgesagt (2017). *Verdienen Roboter Rechte? Was, wenn Maschinen ein Selbstbewusstsein entwickeln?* Kurzgesagt - In a nutshell - YouTube. [Online] Last Accessed 26.09.2023: <https://www.youtube.com/watch?v=DHyUYg8X31c>

Larson, J., Mattu, S., Kirchner, L. & Angwin, J. (2016). *How We Analyzed the COMPAS Recidivism Algorithm*. ProPublica. [Online] Last Accessed 26.09.2023: <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>

Latour, B. (2007). *Reassembling the social: An introduction to actor-network-theory*. Oup Oxford.

Lee, J. A. N. (2021). *John McCarthy - Computer Pioneers by J. A. N. Lee*. IEEE Computer Society. [Online] Last Accessed 26.09.2023: <https://history.computer.org/pioneers/mccarthy.html>

Levy, D. (2009). The ethical treatment of artificially conscious robots. *International Journal of Social Robotics*, 1(3), 209-216. <http://dx.doi.org/10.1007/s12369-009-0022-6>

Li, D. (2020). Blurring Human and Machine Boundary--The Post-Humanist Metaphor of Cyborg-Body in Artificial Intelligence and Minority Report. In proceedings of a conference (Eds.), *2020 International Conference on Language, Art and Cultural Exchange, ICLACE 2020* (pp. 47-50). Atlantis Press. <http://dx.doi.org/10.2991/assehr.k.200709.011>

Lima, G., Kim, C., Ryu, S., Jeon, C. & Cha, M. (2020). Collecting the public perception of AI and robot rights. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2), 1-24. <http://dx.doi.org/10.1145/3415206>

Lin, P., Abney, K. & Bekey, G. (2011). Robot ethics: Mapping the issues for a mechanized world. *Artificial intelligence*, 175(5-6), 942-949. <https://doi.org/10.1016/j.artint.2010.11.026>

Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*, 9(1), 381-386. <http://dx.doi.org/10.21275/ART20203995>

Manning, C. (2020). *Artificial Intelligence Definitions*. Stanford University - Human-Centered Artificial Intelligence. [Online] Last Accessed 26.09.2023: <https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf>

Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures*, 90, 46-60. <https://doi.org/10.1016/j.futures.2017.03.006>

Mayring, P. (2010). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. (11th edition) Beltz.

McCarthy, J. (1978). History of LISP. In J. McCarthy (Ed.), *History of programming languages* (pp. 173-185). Association for Computing Machinery. <https://doi.org/10.1145/800025.1198360>

Meaker, M. (2021). *Rise of robot artists challenges what it means to be human*. Telegraph. [Online] Last Accessed 26.09.2023: <https://www.telegraph.co.uk/technology/2021/05/06/rise-robot-artists-challenges-means-human/>

Moor, J. (2009). Four kinds of ethical robots. *Philosophy Now*, 72, 12-14.

Moor, J. H. (2006). The nature, importance, and difficulty of machine ethics. *IEEE intelligent systems*, 21(4), 18-21. <http://dx.doi.org/10.1109/MIS.2006.80>

Murphy, R. R. & Woods, D. (2009). Beyond Asimov: The Three Laws of Responsible Robotics. *Intelligent Systems, IEEE*, 24(4), 14-20. <http://dx.doi.org/10.1109/MIS.2009.69>

Müller, V. C. (2021). Is it time for robot rights? Moral status in artificial entities. *Ethics and Information Technology*, 23(4), 579-587. <https://doi.org/10.1007/s10676-021-09596-w>

Niedermeyer, E. (1994). Consciousness: function and definition. *Clinical Electroencephalography*, 25(3), 86-93. <https://doi.org/10.1177/155005949402500304>

NIST (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. National Institute of Standards and Technology. [Online] Last Accessed 26.09.2023: <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>

Nolan, J. & Joy, L. (Authors). (2016). *Westworld [TV-Series]*. Bad Robot Productions & Co., Kitler Films, Warner Bros. Television, Jerry Weintraub Productions; HBO.

Nørskov, M. (2022). Robotification & ethical cleansing. *AI & Society* 37, 425-441. <https://doi.org/10.1007/s00146-021-01203-2>

OECD (2022). OECD Framework for the Classification of AI systems. *OECD Digital Economy Papers*, No. 323. <https://doi.org/10.1787/cb6d9eca-en>

Pandey, A. K. & Gelin, R. (2018). A mass-produced sociable humanoid robot: Pepper: The first machine of its kind. *IEEE Robotics & Automation Magazine*, 25(3), 40-48. <http://dx.doi.org/10.1109/MRA.2018.2833157>

Parviainen, J. & Coeckelbergh, M. (2021). The political choreography of the Sophia robot: beyond robot rights and citizenship to political performances for the social robotics market. *AI & society*, 36(3), 715-724. <https://link.springer.com/article/10.1007/s00146-020-01104-w>

Pollock, G. (2011). What is Posthumanism? by Cary Wolfe (2009), Reviewed by Greg Pollock. *Journal for Critical Animal Studies*, 9(1-2), 235-241. Last Accessed 26.09.2023: <http://www.criticalanimalstudies.org/wp-content/uploads/2009/09/13.-Pollock-G-2011-Issue-1-2-Whatis-Posthumanism-up-pp-235-241.pdf>

Prescott, T. J. (2017). Robots are not just tools. *Connection Science*, 29(2), 142-149. <https://doi.org/10.1080/09540091.2017.1279125>

Ram, B. & Verma, P. (2023). Artificial intelligence AI-based Chatbot study of ChatGPT, Google AI Bard and Baidu AI. *World Journal of Advanced Engineering Technology and Sciences*, 8(1), 258-261. <https://doi.org/10.30574/wjaets.2023.8.1.0045>

Ramming, B. & Zweig, M. (2007). *Technological singularity*. [Online] Last Accessed 26.09.2023: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=977ac750e74073b7bdd148fc3f7c3e6fbd4d950a>

Raju, I. P., Sikka, S., Garg, M. A. & Pandey, M. M. A. (2020). Brief review of recent advancement in humanoid robotics research. *Mukt Shabd Journal*, 9(6), 3743-3748.  
Ryznar, M. Robot Love (2018). *Seton Hall Law Review*, 49, 353-374. <https://ssrn.com/abstract=3235187>

Reilly, K. (2011). From automata to automation: The birth of the robot in rur (rossum's universal robots). In K. Reilly (Ed.), *Automata and Mimesis on the Stage of Theatre History* (pp. 148-176). Palgrave Macmillan UK. [https://doi.org/10.1057/9780230347540\\_6](https://doi.org/10.1057/9780230347540_6)



Regan, T. (2013). 'The Case for Animal Rights'. In R. Hursthouse (Ed.), *Ethics, Humans and Other Animals* (pp. 179-188). Routledge.

Retto, J. (2017). *Sophia, first citizen robot of the world*. Universidad Nacional Mayor de San Marcos. [Online] Last Accessed 26.09.2023: [https://www.researchgate.net/profile/Jesus-Retto/publication/321319964\\_SOPHIA\\_FIRST\\_CITIZEN\\_ROBOT\\_OF\\_THE\\_WORLD/links/5a1c8aa2a6fdcc0af3265a44/SOPHIA-FIRST-CITIZEN-ROBOT-OF-THE-WORLD.pdf](https://www.researchgate.net/profile/Jesus-Retto/publication/321319964_SOPHIA_FIRST_CITIZEN_ROBOT_OF_THE_WORLD/links/5a1c8aa2a6fdcc0af3265a44/SOPHIA-FIRST-CITIZEN-ROBOT-OF-THE-WORLD.pdf)

Rimmer, M. (2017). The wild west of robot law. *Australasian Science*, 38(3), 20-21. <https://eprints.qut.edu.au/106013/>

Rojas-Quintero, J. A. & Rodríguez-Liñán, M. C. (2021). A literature review of sensor heads for humanoid robots. *Robotics and Autonomous Systems*, 143, 103834. <https://doi.org/10.1016/j.robot.2021.103834>

Rosenbaum, E. (2022). *Microsoft's GitHub Copilot AI is making rapid progress. Here's how its human leader thinks about it*. CNBC. [Online] Last Accessed 26.09.2023. <https://www.cnbc.com/2022/10/14/microsoft-ai-leaps-ahead-heres-what-its-human-leader-thinks-about-it.html>

Roser, M. (2023). *Here's how experts see AI developing over the coming years*. WEF - World Economic Forum. [Online] Last Accessed 26.09.2023: <https://www.weforum.org/agenda/2023/02/experts-ai-developing-over-the-coming-years>

Rotman, D. (2023). *How to solve AI's inequality problem*. MIT Technology Review. [Online] Last Accessed 26.09.2023. <https://www.technologyreview.com/2022/04/19/1049378/ai-inequality-problem/>

Schlosser, M. (2019). *Agency*. Stanford Encyclopedia of Philosophy. [Online] Last Accessed 26.09.2023: <https://plato.stanford.edu/entries/agency/>

Schröder, W. M. (2021). Robots and rights: reviewing recent positions in legal philosophy and ethics. In J. V. Braun, S. M. Archer, G. M. Reichberg, M. S. Sorondo (Eds.), *Robotics, AI, and Humanity* (pp. 191-203). Springer. [https://doi.org/10.1007/978-3-030-54173-6\\_16](https://doi.org/10.1007/978-3-030-54173-6_16)

Schuett, J. (2019). A Legal Definition of AI. *SSRN Electronic Journal*, 1-14. <http://dx.doi.org/10.2139/ssrn.3453632>

Schuett, J. (2021). Defining the Scope of AI Regulations. Forthcoming In C. Winter (Ed.), *Law, Innovation and Technology, Legal Priorities Project Working Paper Series No. 9*. <http://dx.doi.org/10.2139/ssrn.3453632>

Sgourdou, P. (2022). The Consciousness of Pain: A Thalamocortical Perspective. *NeuroSci*, 3(2), 311–320. <https://doi.org/10.3390/neurosci3020022>

Schwab, K. (2017). *The fourth industrial revolution*. Currency.

Sinnott-Armstrong, W. (2019). *Consequentialism*. Stanford Encyclopedia of Philosophy. [Online] Last Accessed 26.09.2023: <https://plato.stanford.edu/entries/consequentialism/>

Sloman, A. (2006). *Why Asimov's three laws of robotics are unethical*. School of Computer Science, University of Birmingham. [Online] Last Accessed 26.09.2023: <https://www.cs.bham.ac.uk/research/projects/cogaff/misc/asimov-three-laws.pdf>

Solomon, D. (1995). Ethics: III. Normative Ethical Theories. In W. T. Reich (Ed.), *Encyclopedia of bioethics*. (pp. 736-748). Simon and Schuster Macmillan.

SoftBank Robotics (2023). *For better business just add Pepper - What does Pepper do differently?*. SoftBank Robotics. [Online] Last Accessed 26.09.2023: <https://us.softbankrobotics.com/pepper>

Srinivasan, R. & Chander, A. (2021). Biases in AI systems. *Communications of the ACM*, 64(8), 44-49. <https://doi.org/10.1145/3464903>

Stasse, O. & Flayols, T. (2019). An overview of humanoid robots technologies. In G. Venture, J. Laumond & B. Watier (Eds.), *Biomechanics of Anthropomorphic Systems* (pp. 281-310). Springer. [http://dx.doi.org/10.1007/978-3-319-93870-7\\_13](http://dx.doi.org/10.1007/978-3-319-93870-7_13)

Stokes, C. (2018). Why the three laws of robotics do not work. *International Journal of Research in Engineering and Innovation (IJREI)*, 2(2), 121-126.

Tononi, G. (2004). An information integration theory of consciousness. *BMC neuroscience*, 5(42), 1-22. <http://dx.doi.org/10.1186/1471-2202-5-42>

Truitt, E. R. (2021). *Surveillance, Companionship, and Entertainment: The Ancient History of Intelligent Machines*. The MIT Press Reader. [Online] Last Accessed 26.09.2023: <https://thereader.mitpress.mit.edu/the-ancient-history-of-intelligent-machines/>

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>

Turner, J. (2018). *Robot rules: Regulating artificial intelligence*. Springer.

UHN (2023). *Literature Review: Traditional or narrative literature reviews*. University Health Network - University of Toronto. [Online] Last Accessed 26.09.2023: <https://libguides.csu.edu.au/review/Traditional>

UNESCO (2023). *Artificial Intelligence: UNESCO calls on all Governments to implement Global Ethical Framework without delay*. UNESCO Press Release. [Online] Last Accessed 26.09.2023: [https://www.unesco.org/en/articles/artificial-intelligence-unesco-calls-all-governments-implement-global-ethical-framework-without?TSPD\\_101\\_R0=080713870fab2000a0ca32d719b41687b8becebf5e6b482f5e028d6607a3b8ba454e9fa5bcf75920088308e8611430007165a48349b216076af7bc50c11e82afcf9afb37f88d0932853f89be1dbad531d90574aba4d1fb032db7e92dfbbeafdb](https://www.unesco.org/en/articles/artificial-intelligence-unesco-calls-all-governments-implement-global-ethical-framework-without?TSPD_101_R0=080713870fab2000a0ca32d719b41687b8becebf5e6b482f5e028d6607a3b8ba454e9fa5bcf75920088308e8611430007165a48349b216076af7bc50c11e82afcf9afb37f88d0932853f89be1dbad531d90574aba4d1fb032db7e92dfbbeafdb)

Vadymovych, S. Y. (2017). Artificial personal autonomy and concept of robot rights. *European journal of law and political sciences*, (1), 17-21. <https://cyberleninka.ru/article/n/artificial-personal-autonomy-and-concept-of-robot-rights>

Van Gulick, R. (2014). *Consciousness*. Stanford Encyclopedia of Philosophy. [Online] Last Accessed 26.09.2023: <https://plato.stanford.edu/entries/consciousness/>

Wang, J. (2023). Self-Awareness, a Singularity of AI. *Philosophy*, 13(2), 68-77. <http://dx.doi.org/10.17265/2159-5313/2023.02.003>

Wernaart, B. (2021). Developing a roadmap for the moral programming of smart technology. *Technology in Society*, 64, 101466. <https://doi.org/10.1016/j.techsoc.2020.101466>

West, D. M. & Allen, J. R. (2018). *How artificial intelligence is transforming the world*. Brookings. [Online] Last Accessed 26.09.2023: <https://www.brookings.edu/articles/how-artificial-intelligence-is-transforming-the-world/>

WikiDiff (2023). *Operation vs Action - What's the difference?*. [Online] Last Accessed 29.09.2023: <https://wikidiff.com/operation/action>

Winston, P. H. (2016). Marvin L. Minsky (1927–2016). *Nature*, 530(7590), 282-282. <https://doi.org/10.1038/530282a>

Xu, M., David, J. M. & Kim, S. H. (2018). The fourth industrial revolution: Opportunities and challenges. *International journal of financial research*, 9(2), 90-95. <http://dx.doi.org/10.5430/ijfr.v9n2p90>

Yoshida, E. (2019). Robots that look like humans: A brief look into humanoid robotics. *Metode Science Studies Journal*, 9, 143-151. <http://dx.doi.org/10.7203/metode.9.11405>

Zajko, M. (2022). Artificial intelligence, algorithms, and social inequality: Sociological contributions to contemporary debates. *Sociology Compass*, 16(3), e12962. <http://dx.doi.org/10.1111/soc4.12962>

## Bibliography of Figures:

**Fig. 01:** *The humanoid robot Ava from the film "Ex Machina".* Süddeutsche Zeitung. [Online] Last Accessed 26.09.2023: <https://www.sueddeutsche.de/kultur/kino-traumfrau-gebaut-1.2445461>

**Fig. 02:** *"Sophia" by Hanson Robotics Ltd. speaking at the AI for GOOD Global Summit 01.* Wikipedia. [Online] Last Accessed 26.09.2023: [https://de.wikipedia.org/wiki/Datei:Sophia\\_\(robot\).jpg](https://de.wikipedia.org/wiki/Datei:Sophia_(robot).jpg)

**Fig. 03:** *"Paro" the robot seal.* ARS Electronica. [Online] Last Accessed 26.09.2023: <https://ars.electronica.art/aeblog/de/2021/12/23/throwback-paro-the-robot-seal/>

**Fig. 04:** *Robot called "Atlas" by Boston Dynamics.* Turtle Timeline. [Online] Last Accessed 26.09.2023: <https://turtletimeline.com/boston-dynamics-focuses-on-real-world-applications-for-atlas-robot/>

**Fig. 05:** *The differences between AI, ML and DL.* Samson, O. (2017). Medium. [Online] Last Accessed 26.09.2023: <https://towardsdatascience.com/deep-learning-weekly-piece-the-differences-between-ai-ml-and-dl-b6a203b70698>

**Fig. 06:** *Some key milestones and major platforms in the development of humanoid robots.* Fukuda, T., Dario, P. & Yang, G. Z. (2017). Humanoid robotics—History, current state of the art, and challenges. *Science Robotics*, 2(13), eaar4043. <https://doi.org/10.1126/scirobotics.aar4043>

**Fig. 07:** *Humanoid robot "Pepper" interacting with humans.* SoftBank Robotics. [Online] Last Accessed 26.09.2023: <https://us.softbankrobotics.com/pepper>

**Fig. 08:** *Interplay of different philosophical concepts in the question of robot rights and the underlying theories.* Selfmade with Keynote by Marc Lewald

**Fig. 09:** *Sample of interview questions.* Selfmade with Pages by Marc Lewald

**Fig. 10:** *Main categories of qualitative content analysis according to Mayring.* Selfmade with Pages by Marc Lewald

**Fig. 11:** *Subcategories (sorted according to the main categories A-E) of qualitative content analysis according to Mayring.* Selfmade with Pages by Marc Lewald

**Fig. 12:** *Screenshot of the qualitative content analysis and the formation of categories and codes in MAXQDA. Selfmade with Keynote by Marc Lewald*

## 10. Appendix

Fig. 09: Sample of interview questions



### Master thesis - Robot Rights

*Italics: own definitions of the terms (used in the course of the thesis)*

1. Are you already familiar with the concept of robot rights (rights for robots) or have you dealt with it beforehand?

2. What would it mean to think of a humanoid robot as an entity with full agency?

*"Agency" is a concept in philosophy that refers to the ability of an individual or entity to act independently and make free choices.*

3. Would you interpret an action performed by the robot as an action in the classical sense, or rather as an operation of a machine?

*"Operation" refers to the process by which a device performs its intended function, while "action" refers to an action performed to achieve a specific goal.*

4. How would you define the concept of consciousness?

- What would be associated with consciousness? (Self-awareness, ability to think, etc.)
- Do you think, then, that humanoid robots could develop or already have consciousness in the future?
- Does a robot necessarily need consciousness to be granted rights?

*"Consciousness" refers to the ability of a living being or entity to perceive, recognize, and reflect upon its environment and itself.*

5. To what extent would a posthumanist perspective influence or change the current debate on robot rights?

- What impact might robot rights have on our understanding of humanity and non-human life?

*In general, "posthumanism" refers to a movement or thinking that questions the traditional boundaries of personhood and the centrality of humans in the world.*

6. In your opinion, what are the reasons for and against the introduction of robotic rights?

7. What risks might arise from the introduction of robot rights?

- What risks might arise if we do not introduce them?
- Do you think robots should then have the same rights as humans?
- Are there certain rights that you would not grant to a robot under any circumstances?

8. In your opinion, is the introduction of robot rights necessary?

**Fig. 10: Main categories of qualitative content analysis according to Mayring**

Main categories (A-E)	Definition	Subcategories (clustered)
<b>(A) Consciousness</b>	This main category includes the philosophical discussion of the concept of consciousness in the course of the question of robot rights and humanoid robots and thereby describes different types and preconditions of and for consciousness.	<p><b>Defined:</b></p> <ol style="list-style-type: none"> <li>1. Humanization</li> <li>2. Definition</li> <li>3. Misplaced</li> </ol> <p><b>Requirements:</b></p> <ol style="list-style-type: none"> <li>4. Environment</li> <li>5. Pain</li> <li>6. Imagination</li> </ol> <p><b>Types of consciousness:</b></p> <ol style="list-style-type: none"> <li>7. Layered consciousness</li> <li>8. Science fiction</li> <li>9. Human consciousness</li> <li>10. Self-consciousness</li> <li>11. Non-human consciousness</li> </ol>
<b>(B) Agency</b>	This main category includes the philosophical discussion of the concept of agency (in particular the concept of full agency) in the course of the question of robot rights and humanoid robots and describes the conditions that have to be fulfilled in order to attribute agency to "machines".	<p><b>Defined:</b></p> <ol style="list-style-type: none"> <li>1. Humanism</li> <li>2. Definition</li> <li>3. Full agency</li> </ol> <p><b>Requirements:</b></p> <ol style="list-style-type: none"> <li>4. Consciousness</li> <li>5. Autonomy</li> <li>6. Intention</li> </ol> <p><b>Additional information:</b></p> <ol style="list-style-type: none"> <li>7. „Handlungsvermögen“</li> <li>8. Against</li> </ol>
<b>(C) Action</b>	This main category includes the philosophical discussion of the concept of action in the course of the question of robot rights and humanoid robots, the preconditions to speak of an action and its distinction from the concept of an operation.	<p><b>Defined:</b></p> <ol style="list-style-type: none"> <li>1. Action/operation</li> <li>2. Human</li> </ol> <p><b>Approach:</b></p> <ol style="list-style-type: none"> <li>3. Target-oriented</li> <li>4. Programmed</li> </ol>



Main categories (A-E)	Definition	Subcategories (clustered)
<b>(D) Humanism</b>	This main category includes the philosophical discussion of the concept of post- and transhumanism in the course of the question of robot rights and humanoid robots, the comparison to us humans and other entities and the implications of such perspectives.	<p><b>Defined:</b></p> <ol style="list-style-type: none"> <li>1. Main theories</li> <li>2. Silicon Valley</li> <li>3. Posthumanism</li> <li>4. Transhumanism</li> <li>5. Definition</li> </ol> <p><b>Implications:</b></p> <ol style="list-style-type: none"> <li>6. Self-perception</li> <li>7. Animals</li> <li>8. No change</li> </ol>
<b>(E) Robot rights</b>	This main category includes the general discussion about the concept of robot rights. This encompasses general philosophical theories and approaches, but also general information on the topic of robot rights.	<p><b>Defined:</b></p> <ol style="list-style-type: none"> <li>1. State of knowledge</li> <li>2. Sense of purpose</li> <li>3. Main theories</li> <li>4. Moral and legal rights</li> <li>5. Legal constructs</li> </ol> <p><b>Additional findings:</b></p> <ol style="list-style-type: none"> <li>6. Artificial intelligence</li> <li>7. In dubio pro reo</li> <li>8. Science fiction</li> <li>9. Humanoid robots</li> <li>10. Definition robots</li> <li>11. Economic interest</li> <li>12. Against</li> </ol>

**Fig. 11: Subcategories (sorted according to the main categories A-E) of qualitative content analysis according to Mayring**

A-E	Subcategories	Definition	Anchor example	Coding Rule
A	Humanisation	This is about the humanization of robots and their appearance, which can lead to a granting of consciousness.	<p>"Roboter erscheinen halt menschlich, weil sie in diesem Sinne Humor imitieren oder halt (...) menschliches Antlitz haben (...), davon sprechen, dass sie Gefühle haben"</p> <p>[Robots appear human because they imitate humour in this sense or (...) have a human face (...), speak of having feelings]</p>	This sub-category deals exclusively with the humanization of humanoid robots and the consciousness that presumably results from it.
A	Definition	This is about the different meanings of the definition of consciousness and its implications.	<p>"Es gibt eine Unzahl an Bewusstseinsdefinitionen. Es ist strittig, ob zwischen Bewusstsein und Selbstbewusstsein zu unterscheiden ist, oder ob jede Form von Bewusstsein Selbstbewusstsein voraussetzt. Man muss also vorab klären, welche Bewusstseinsdefinition man in Anschlag bringen"</p> <p>[There are a myriad of definitions of consciousness. It is debatable whether a distinction should be made between consciousness and self-consciousness, or whether every form of consciousness presupposes self-consciousness. One must therefore clarify in advance which definition of consciousness one wishes to apply]</p>	This sub-category deals solely with the definition of consciousness and the importance of defining it.
A	Misplaced	This sub-category addresses the ultimate importance and relevance of the notion of consciousness in relation to the RQ.	<p>"Insofern bin ich mir nicht sicher, ob das, der Begriff des Bewusstseins, für die Thematik, ihre Arbeit (...) als sehr, sehr allgemeiner Begriff tatsächlich entscheidend ist"</p> <p>[In this respect, I am not sure whether this, the concept of consciousness, is actually decisive for the subject matter, your work (...) as a very, very general concept]</p>	This sub-category deals exclusively with the question of whether it makes sense to speak of the concept of consciousness in the context of the RQ.

A-E	Subcategories	Definition	Anchor example	Coding Rule
A	Environment	This subcategory is about the assumption that consciousness is always linked to some form of environment or perception of the environment.	<p>"Sprich (...), Bewusstsein ist immer bezogen auf eine Umwelt, in der sich das bewusstseiene Subjekt oder die Entität befindet, und es ist nie irgendwie ein Bewusstsein per se"</p> <p>[In other words (...), consciousness is always related to an environment in which the conscious subject or entity is located, and it is never somehow a consciousness per se]</p>	In this context, the subcategory deals exclusively with the question of whether and to what extent the concept of consciousness is coupled to the environment.
A	Pain	This subcategory explores the extent to which the concept of pain or the sensation of pain is related to consciousness.	<p>"Eine Entität braucht Bewusstsein bzw. Empfindungs- und Leidensfähigkeit, um moralische Rechte zu bekommen."</p> <p>[An entity needs consciousness or sentience and suffering in order to have moral rights.]</p>	The subcategory deals only with passages that are related to the topic of pain or even the sensation of pain.
A	Imagination	This is about the concept of imagination in connection with consciousness.	<p>"Genauer gesagt braucht es, um generell reflektieren zu können, und damit auch um überhaupt von Bewusstsein sprechen zu können, zumindest eine Verknüpfung von Vorstellungen."</p> <p>[More precisely, in order to be able to reflect in general, and thus also to be able to speak of consciousness at all, it needs at least a linkage of imaginations.]</p>	The subcategory is limited exclusively to text passages that directly relate the concept of ideas to consciousness.
A	Layered consciousness	This subcategory describes consciousness as a multilevel phenomenon.	<p>"Es gibt Stufen des Bewusstseins."</p> <p>[There are levels of consciousness.]</p>	This subcategory deals only with text excerpts that refer to multilevel consciousness.

A-E	Subcategories	Definition	Anchor example	Coding Rule
A	Science fiction	This subcategory describes an understanding of consciousness that is strongly influenced by science fiction works.	<p>"Aber diese Bewusstseins-Konzepte schwirren schon in den 40, 50 er Jahren mit (...), oft mit Bildern vom Gehirn, mit dem Perceptron, wo dann sofort dieses Neural Network aufkommt, und das sind Metaphern und Bilder (...), die dann auch aus der Wissenschaft selbst gekommen sind in den 50, er, 60 er Jahren, also von KI-Forscher/-innen, also von Programmierer/-innen, (...) die dann auch zeitgleich von Science Fiction und sehr viel SciFi in Film aufgenommen (...) wurden, und das verschwimmt sehr stark"</p> <p>[But these concepts of consciousness were already buzzing in the 40s, 50s (...), often with images of the brain, with the perceptron, where this neural network immediately comes up, and these are metaphors and images (...) that then also came from science itself in the 50s, 60s, i.e. from AI researchers, i.e. from programmers, (...) which were then also simultaneously taken up by science fiction and very much scifi in film (...), and that blurs very strongly]</p>	This subcategory deals only with text passages that set the current understanding of consciousness in the context of science fiction theories.
A	Human consciousness	This subcategory is about the clear distinction of human consciousness from other forms of consciousness.	<p>"Dass das menschliche Bewusstsein eigentlich nicht nur vom Hirn gesteuert wird (...), sondern sozusagen eine Art von Embodied Consciousness ist, dass der ganze Körper irgendwie der Raum ist, in dem sich sowas wie Bewusstsein abspielt."</p> <p>[That human consciousness is actually not only controlled by the brain (...), but is a kind of embodied consciousness, so to speak, that the whole body is somehow the space in which something like consciousness takes place.]</p>	This subcategory deals with sections that refer to a clear definition of a human consciousness, separating it from other forms of consciousness.

A-E	Subcategories	Definition	Anchor example	Coding Rule
A	Mechanical consciousness	This subcategory is concerned with the definition and description of a machine or technical consciousness and its distinctions from other forms of consciousness, as well as its attribution to humanoid robots.	<p>"Wie gesagt, also technisches Bewusstsein im Sinne von informiert sein über den Systemzustand kann ich mir ohne weiteres vorstellen."</p> <p>[As I said, I can easily imagine technical awareness in the sense of being informed about the state of the system.]</p>	This subcategory deals exclusively with texts that define machine or technical consciousness and how it differs from other forms of consciousness as well , as well as their attribution.
A	Self-consciousness	This subcategory deals with the extent to which self-consciousness is related to the notion of consciousness and the extent to which humanoid robots can be expected to have some form of self-consciousness in the future.	<p>"Selbstbewusstsein in einem engeren Sinne haben Roboter nicht. Humanoid oder nicht humanoid macht da kein Unterschied. Das Systeme in Zukunft ein Selbstbewusstsein entwickeln könnten, halte ich, im Ausgang von den bisherigen technischen Ansätzen, für ausgeschlossen."</p> <p>[Robots do not have self-awareness in a narrower sense. Humanoid or non-humanoid makes no difference. That systems could develop a self-awareness in the future, I think, based on the current technical approaches, is impossible.]</p>	This subcategory works exclusively with passages that refer to a form of self-awareness and the question to what extent humanoid robots might already be capable of this.
A	Non-human consciousness	This subcategory describes a form of non-human consciousness, mostly attributed to animals, and places it in comparison to other attributions of consciousness.	<p>"Die werden auch von Tieren (...), jedenfalls von den meisten Tieren, nicht vollständig erfüllt, obwohl gerade Bewusstsein in einem etwas basaleren Sinne glaube ich auch bei vielen Tieren und in höherem Maße natürlich auch bei den Menschenaffen vorliegt (...), sodass wir mehr und mehr die Kriterien für Personen sein (...), erfüllt sehen."</p> <p>[These are also not completely fulfilled by animals (...), at least by most animals, although consciousness in a somewhat more basal sense is also present, I believe, in many animals and to a greater extent, of course, also in the apes (...), so that we see more and more the criteria for people to be (...), fulfilled.]</p>	This subcategory works only with text excerpts that refer to a form of non-human or animal consciousness and places it in the context of other forms of consciousness.

A-E	Subcategories	Definition	Anchor example	Coding Rule
B	Humanism	This subcategory describes the humanization of a robot into an agent and an associated view of agency.	<p>"Ihr Thema ist natürlich die Frage, inwieweit diese Technologie vermenschlicht wird, also eine Art Vermenschlichung der Technologie darstellt und somit menschliche Züge erhält (...). Und das ist der Punkt, an dem die Frage des Agenten aufkommt, inwieweit er zu einer Art Mensch und damit zu einer Art Agent wird."</p> <p>[What concern that's your topic, of course, is to what extent like this technology, something like becomes humanized like that is like sort of humanization of technology and hence getting human features (...). And that's where the issue of the agent comes from, like, to what extent it becomes some sort of a human and hence some sort of an agent.]</p>	This subcategory focuses exclusively on passages that consider the theme of agency through a form of humanization.
B	Definition	This subcategory includes various definitions of agency and how important the actual definition of the term is with respect to the research question.	<p>"Insofern würde ich dafür plädieren, sehr vorsichtig zu sein, mit der Anwendung des Wortes Handlungsvermögen oder gar volles Handlungsvermögen auf Roboter, wenn man nicht genau definiert, was genau man damit im Blick auf Maschinen meinen kann."</p> <p>[In this respect, I would plead to be very careful with the application of the word agency or even full agency to robots, if one does not define exactly what one can mean with regard to machines]</p>	This subcategory deals only with texts that are directly related to the definition of agency and that describe the meaning of the definition.
B	Full Agency	This subcategory describes the notion of full agency in relation to humanoid robots and separates it from other statements on the topic of agency, which do not directly refer to full agency.	<p>"Wenn nun einer Robotereinheit ein volles Handlungsvermögen zugeschrieben würde, würde das bedeuten, dass diese Einheit willentlich Aktionen ausführt mit der Annahme, dass diese Aktionen zielführend sind."</p> <p>[Now, if a robotic unit were ascribed full agency, this would mean that this unit performs actions willfully with the assumption that these actions are goal-directed.]</p>	This subcategory is only concerned with statements directly related to the concept of full agency and distinguishes them from general statements about agency.

A-E	Subcategories	Definition	Anchor example	Coding Rule
B	Consciousness	In this subcategory, the term consciousness is in the context of the concept of agency and challenged as a premise.	<p>"Jetzt im harten Sinne, sagen wir mal (...), einer bewusstseinsphilosophische Position, die da sagt, Intentionalität hat immer etwas mit Bewusstsein, Selbstbewusstsein, intentionalen Akten oder so etwas zu tun."</p> <p>[Now in the hard sense, let's say (...), a philosophy of consciousness position that says intentionality always has something to do with consciousness, self-consciousness, intentional acts or something like that.]</p>	This subcategory works only with blocks of text that are limited to the theme of consciousness in terms of the question of agency and explore it as a presupposition.
B	Autonomy	This subcategory defines the notion of autonomy in relation to the issue of agency.	<p>"Das Problem hier ist, dass (...), ein Handlungsvermögen (...), es kommt mit Autonomie, richtig? Es ist ein autonomes Handlungsvermögen mit der Fähigkeit zu autonomem Handeln und autonomem Denken und so weiter."</p> <p>[The issue here is that (...), an agency has (...), it comes with autonomy, right? It's an autonomous agency with capacity for autonomous action and autonomous thinking and so forth.]</p>	This subcategory only works with sections that directly relate to the notion of autonomy with respect to the issue of agency.
B	Intention	This subcategory outlines the relation between the notion of agency and that of intentionality.	<p>"Also ich persönlich verwende den Handlungsbegriff, der (...), sagen wir mal, intentional gefüllt ist."</p> <p>[So I personally use the concept of action, which is (...), let's say, intentionally filled.]</p>	This subcategory only works with text passages that deal with the topic of intentionality in the course of agency.
B	„Handlungsvermögen“	This subcategory is used to describe the distinction between the german and english notions of agency.	<p>"Also wir sprechen vermutlich über das, was das englische Wort von Agency (/) ... Wobei das Englische da nicht so deutlich unterscheidet wie das Deutsche."</p> <p>[So we're probably talking about what the English word of agency (/) ... Where English doesn't distinguish there as clearly as German does]</p>	This subcategory only operates with textual sections that refer to the direct distinction between the german and english notions of agency.



A-E	Subcategories	Definition	Anchor example	Coding Rule
B	Against	This subcategory describes the statements that are against granting agency with respect to machines.	<p>"Ich würde weder zum jetzigen Zeitpunkt noch in Zukunft im menschlichen Sinne volles Handlungsvermögen einer Maschine zuschreiben"</p> <p>[I would not attribute full agency to a machine in the human sense, either at the present time or in the future]</p>	This subcategory deals only with texts that directly oppose granting agency to a machine.
C	Action/ Operation	This subcategory sets out to define the distinction between an action and an operation in terms of the actions of machines.	<p>"Also deswegen (...), ich spreche von Menschen, die handeln, und von Maschinen, die Operationen durchführen oder operieren, aber (...) auf die Art und Weise versuche ich eben auch noch, die Differenz deutlich zu machen"</p> <p>[Therefore (...), I speak of people who act and of machines that carry out operations, but (...) in this way I also try to make the difference clear.]</p>	This subcategory deals only with sections that address the specific distinction between an action and operation.
C	Human	This subcategory defines the concept of action as a necessarily human or human-bound action.	<p>"Die Aktion oder die Handlung, das ist etwas typisch menschliches"</p> <p>[The action or the act, that is something typically human]</p>	This subcategory elaborates only on textual excerpts that deal with the purely human component of action and define the concept of action as a human one.
C	Goal-orientated	This subcategory is about definitions and perspectives of the concept of action based on a goal-directedness.	<p>"Es ist nicht wichtig zu wissen, was das antreibt, (...) diesen Prozess. Es reicht, wenn wir etwas sehen (...), das (...) auf ein Ziel gerichtet, sich bewegt oder hin rechnet oder irgendeine Aktion macht."</p> <p>[It is not important to know what drives (...) this process. It's enough if we see something (...) that is (...) directed towards a goal, moving or calculating or doing some action.]</p>	In this subcategory, only text passages are processed that understand the concept of action as a goal-based or -oriented action.



A-E	Subcategories	Definition	Anchor example	Coding Rule
C	Programmed	This subcategory defines an action performed by a robot or machine as a programmed action.	<p>"Die idealtypische Form (...) ist eine programmierte Art von Arbeiter (...). Also, ja, ich denke, es gibt einen Unterschied zwischen dem Verfahren und der Aktion, aber auch bei einem Roboter sind beide programmiert."</p> <p>[Ideal typical form (...) is a programmed sort of worker (...). So, yes, I think there is a difference between the procedure and the action, but also with a robot, both are programmed.]</p>	This subcategory contains only textual parts that deal with the notion of action as a programmed action.
D	Main theories	This subcategory describes a number of well-known theories and perspectives in the field of humanism, as well as trans- and posthumanism in relation to robot rights.	<p>"Und eine Akteur Netzwerktheorie von Bruno Latour oder auch dieser neue Materialismus, finde ich, hat eine sehr produktive (...) Diskussion ausgelöst"</p> <p>[And an actor network theory by Bruno Latour or also this new materialism, I think, has triggered a very productive (...) discussion]</p>	In the course of this subcategory, only text sections that draw on well-known theories and perspectives in the course of humanism, trans-, and posthumanism in relation to robot rights will be processed.
D	Silicon Valley	This subcategory is describing the concept of trans- and posthumanism from a Silicon Valley perspective.	<p>"Also ich finde (...), da ist die Frage, wie man auch da wieder Posthumanismus definiert, weil es gibt halt so einen Silicon-Valley Posthumanismus, der halt eher so in Human Enhancement geht."</p> <p>[So I think (...), there is the question of how to define posthumanism again, because there is just such a Silicon Valley posthumanism, which goes more into human enhancement.]</p>	This subcategory works exclusively with text passages that deal with a Silicon Valley understanding of post- and transhumanism.

A-E	Subcategories	Definition	Anchor example	Coding Rule
D	Posthumanism	This subcategory contains the various understandings and implications of posthumanism.	<p>"Im Allgemeinen bezieht der "Posthumanismus" sich auf eine Bewegung oder ein Denken, das die traditionellen Grenzen des Menschseins und die zentrale Stellung des Menschen in der Welt in Frage stellen."</p> <p>[In general, "posthumanism" refers to a movement or thinking that questions the traditional boundaries of what it means to be human and the centrality of humans in the world.]</p>	This subcategory deals only with texts that address the general understanding and implications of a posthumanist perspective.
D	Transhumanism	This subcategory outlines the various understandings of transhumanist perspectives and their implications.	<p>"Und dann gibt's halt, die Human Enhancement und Transhumanisten, die kommen eher aus dem Silicon Valley, die haben halt (...) , ja, einfach Vorstellungen davon, den Menschen auf eine neue evolutionäre Spezies, einen neuen evolutionären Status hin (...) zu (...) heben"</p> <p>[And then there are the human enhancement and transhumanists, who tend to come from Silicon Valley, who simply have (...) ideas about raising humans to a new evolutionary species, a new evolutionary status (...).]</p>	This subcategory deals exclusively with text excerpts that deal with the definition of a transhumanist perspective and what implications it might have.
D	Definition	This subcategory deals with the difficulty of defining and understanding posthumanism as well as transhumanism.	<p>"Da es wiederum eine Vielzahl an posthumanistischen Positionen gibt, lässt sich diese Frage nur schwer beantworten"</p> <p>[Again, since there are a variety of posthumanist positions, this question is difficult to answer]</p>	This subcategory only handles writings that deal with the difficulty regarding the definition or understanding of post- and transhumanist theories.

A-E	Subcategories	Definition	Anchor example	Coding Rule
D	Self-perception	This subcategory describes post and transhumanist implications related to the human being and human self-perception.	<p>"Also wie denken wir Menschen über uns selber nach? Wie verstehen wir uns?"</p> <p>[So how do we humans think about ourselves? How do we understand ourselves?]</p>	This subcategory only deals with blocks of texts that address the question of what it means to be human from a post- and transhumanist perspective, questioning how humans perceive themselves.
D	Animals	This subcategory defines a posthumanist perspective towards animal rights, via human rights, to robot rights.	<p>"Und dann haben wir die Tierrechte (...), wir sind von den Menschenrechten zu den Tierrechten übergegangen"</p> <p>[And then we got animal rights (...), we moved from human rights to animal rights]</p>	This subcategory only examines text excerpts that look at a comparison between animal rights, human rights, and also robot rights from a posthumanist perspective.
D	No change	This subcategory is defined by the potentially non-existent effects of a posthumanist perspective on the subject of robot rights and their possible introduction.	<p>"Insofern würde ich meinen, dass da kein (...) kein grundlegender Unterschied hergestellt ist, sondern das einfach noch ein paar zusätzliche Wesen mit Rechten dazukommen, wie ja doch jedes Jahr immer wieder neue Menschen mit neuen Rechten geboren werden."</p> <p>[In this respect I would think that there is no (...) no fundamental difference established, but that simply a few additional beings with rights are added, as new people with new rights are born every year.]</p>	This subcategory applies only to sections of text where a post- and trans humanist perspective on the subject and the introduction of robot rights would have no significant impact.

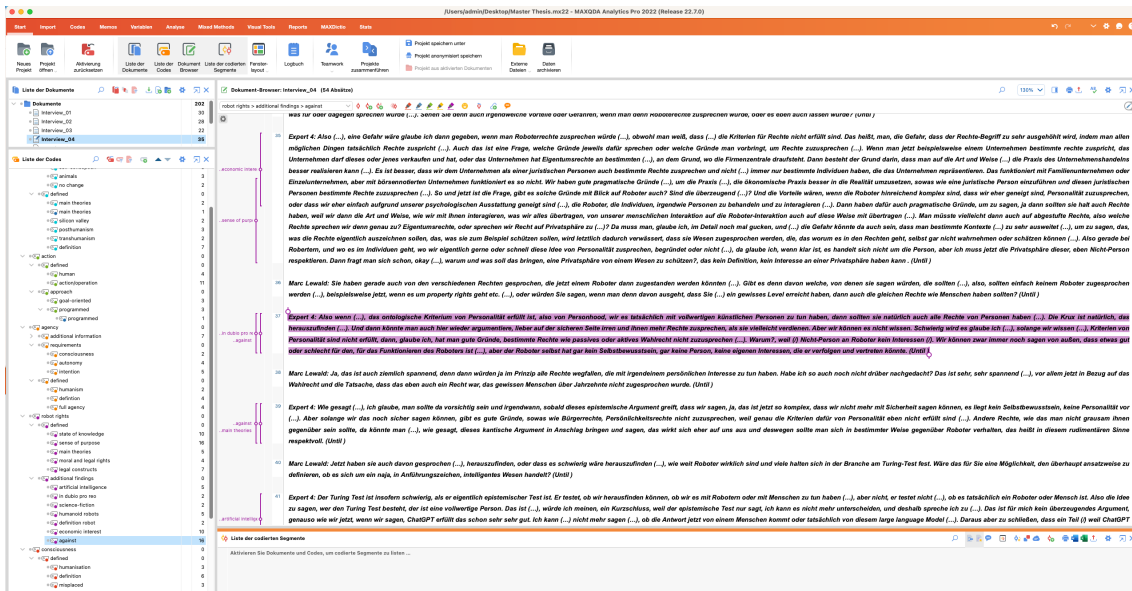
A-E	Subcategories	Definition	Anchor example	Coding Rule
E	State of knowledge	This subcategory describes the experience that the surveyed experts have with the topic of robot rights and humanoid robots.	<p>"Ja, hab ich. Ist immer die Frage, inwieweit Sie die Erfahrungen ziehen, aber nachdem es einer meiner Forschungsgegenstände ist, habe ich mich damit beschäftigt."</p> <p>[Yes, I did. Is always the question to what extent you draw the experience, but since it is one of my research subjects, I have dealt with it.]</p>	This subcategory is only concerned with text parts that revolve around the previous experiences of the surveyed experts in the field of humanoid robotics and robot rights.
E	Sense of purpose	This subcategory describes the meaningfulness, the relevance, as well as the necessity of the topic of robot rights in general.	<p>"Aktuell würde ich meinen, dass es nicht nötig ist, Roboterrechte einzuführen, weil (...) die einschlägigen Kriterien, die dafür nötig werden, schlicht noch nicht vorliegen."</p> <p>[Currently, I would think that it is not necessary to introduce robot rights because (...) the relevant criteria that will be necessary for this simply do not exist yet.]</p>	This subcategory deals only with text sections that refer to the relevance, necessity and importance of the topic of robot rights in general.
E	Main theories	This subcategory includes a variety of general and well-known academic theories related to robot rights.	<p>"Ob wir grausam gegenüber Tiere sein dürfen oder nicht, das schlägt auf unseren Charakter zurück, so wie Kant argumentiert. Und das könnte man für die Einführung von Roboterrechten ebenso geltend machen"</p> <p>[Whether or not we are allowed to be cruel to animals reflects on our character, as Kant argues. And this could be argued for the introduction of robot rights as well]</p>	This subcategory only addresses text passages that deal with well-known and popular theories regarding robot rights.
E	Moral and legal rights	This subcategory describes the distinction of the general concept of law into moral and legal or juridical law.	<p>"Also auch da erst mal, wovon reden wir, wenn wir von Roboterrechten sprechen? Reden wir von juristischen Rechten? Reden wir von moralischen Rechten, oder reden wir von beidem?"</p> <p>[So first of all, what are we talking about when we talk about robot rights? Are we talking about legal rights? Are we talking about moral rights or are we talking about both?]</p>	This subcategory is only applied to interview excerpts that dealt with the definition of law and its distinction between moral and legal law.

A-E	Subcategories	Definition	Anchor example	Coding Rule
E	Legal constructs	This subcategory defines the previous legal constructs that exist with respect to robot rights or even AI in general.	<p>"Also gerade auf EU-Ebene gibt es diesen AI-Act und Digital Service Act und etc. das Digital Markets Act. Das heißt, da passiert gerade sehr viel."</p> <p>[At the EU level, there is the AI Act, the Digital Service Act and the Digital Markets Act. This means that a lot is happening right now]</p>	This subcategory is applied only to interview parts where legal constructs or frameworks in the field of AI and robot rights are discussed.
E	Artificial intelligence	This subcategory refers to statements made in general on the topic of AI that could be relevant for certain perspectives on robot rights.	<p>"It's often called artificial intelligence, originally called machine intelligence (...). And from my perspective, that's not, not a real intelligence. I think (...) this word intelligence is wrongly chosen by Alan Turing at the time."</p> <p>[es wird oft als künstliche Intelligenz bezeichnet, ursprünglich als maschinelle Intelligenz (...). Und aus meiner Sicht ist das keine, keine wirkliche Intelligenz. Ich denke, (...) dieses Wort Intelligenz ist von Alan Turing damals falsch gewählt worden.]</p>	This subcategory only addresses sections of the text that deal with general statements on the topic of AI that may be applicable to the discussion of robot rights.
E	In dubio pro reo	This subcategory describes an "in dubio pro reo" argument with regard to the introduction of robot rights.	<p>"Also für die Einführung von Roboterrechten würde, glaube ich (...), vor allem dieses In dubio pro reo Argument sprechen."</p> <p>[So for the introduction of robot rights, I think (...), this in dubio pro reo argument would speak above all.]</p>	This subcategory deals only with paragraphs that address the issue of robot rights with an "in dubio pro reo" argument.
E	Science fiction	This subcategory describes a view of robotic rights or even humanoid robots that compares current developments with representations from the world of science fiction or even makes use of these representations.	<p>"Ich meine, es gibt ja eine Menge Science Fiction, die viele verschiedene Punkte sehr fein durchdiskutiert."</p> <p>[I mean, there's a lot of science fiction that goes through a lot of different points very finely.]</p>	This subcategory deals only with texts that take a view of robotic rights or even humanoid robots that compares current developments with representations from the world of science fiction or even makes use of these representations.

A-E	Subcategories	Definition	Anchor example	Coding Rule
E	Humanoid robots	This subcategory outlines the previous experiences of the experts surveyed specifically on the topic of humanoid robotics.	<p>"Wir sind in Labore gegangen (...) und haben so ein bisschen auch mitgekriegt, was in der Pflege, zum Teil im In und Ausland, auch an humanoiden Roboter schon eingesetzt wird. Also so Leute wie, also Leute in Führungszeichen wie Pepper zum Beispiel, kennt ja jetzt auch mittlerweile fast jeder."</p> <p>[We went to laboratories (...) and got to know a bit about what humanoid robots are already being used in nursing, partly in Germany and abroad. People like Pepper, for example, in quotation marks, are now known to almost everyone.]</p>	This subcategory only considers text parts that deal with previous experiences of the interviewed experts with regard to humanoid robots.
E	Definition robot	This subcategory describes the importance of the precise definition of robots with regard to the issue of robot rights.	<p>"Bevor die Frage nach den Roboterrechten angegangen werden kann wäre es demnach sinnvoll zunächst zu klären, was als Roboter verstanden wird."</p> <p>[Before the question of robot rights can be addressed, it would therefore make sense to first clarify what is understood as a robot.]</p>	This subcategory only deals with text passages that focus on the importance of the general definition of robots in order to address the question of robot rights.
E	Economic interest	This subcategory defines the economic interests behind the discussion of robot rights.	<p>"Also stellen wir uns vor (...), bei I, Robot gibt es ja diese Roboterarmee, die würden sich jetzt gewerkschaftlich organisieren und sagen (...), was heißt hier 60 Stunden Woche arbeiten? Ne, also 20 (...) Stunden bei vollem Lohnausgleich, so ungefähr ja."</p> <p>[So let's imagine (...), at I, Robot there is this robot army, they would now organize themselves into a union and say (...), what does working 60 hours a week mean here? No, so 20 (...) hours with full wage compensation, something like that.]</p>	This subcategory addresses solely texts that deal with the economic interests and efforts behind the discussion of robot rights.

A-E	Subcategories	Definition	Anchor example	Coding Rule
E	Against	This subcategory discusses positions and statements that argue directly against the introduction of robot rights and why.	<p>"Solange wir das noch sicher sagen können, gibt es gute Gründe, sowas wie Bürgerrechte, Persönlichkeitsrechte nicht zuzusprechen, weil genau die Kriterien dafür von Personalität eben nicht erfüllt sind"</p> <p>[As long as we can say that for sure, there are good reasons not to grant something like civil rights, personal rights, because exactly the criteria for this personhood are not fulfilled]</p>	This subcategory deals only with interview excerpts in which there is clear opposition to the introduction of robot rights and its justification.

**Fig. 12: Screenshot of the qualitative content analysis and the formation of categories and codes in MAXQDA**



## **Anonymous transcripts from the expert interviews**

Simple transcripts: simplified for readability and comprehensibility. (Dresing & Pehl, 2018, see Research Design)

Specific Rules:

Filler words and linguistic irregularities are omitted, without specific highlighting of volume, emphasis, or interruptions. Regional dialects are translated into standard German expressions (only one interview is in english). Silent moments are represented using parentheses and ellipses (...). In addition, incomprehensible terms and phrases, that cannot be transcribed, are represented by a (/). The brackets and time at the end of each section describe the temporal classification of the preceding section during an interview.



**Transcript 01: (personal interview)**

Marc Lewald: Die erste Frage wäre, überhaupt erst einmal ... bisher mit humanoiden Robotern bzw. dann auch mit Roboter Rechten ein bisschen auseinandergesetzt? (Until 00:00:10)

(1) Expert 1: Also, ich habe ein Review geschrieben, in einem Buch, was von X geschrieben worden ist. Sie selbst (...) bezeichnet sich als Roboter-Ethikerin (...), die aus einer feministischen Richtung kommt und einer Denkrichtung, die heißt neuer Materialismus, wo es genau darum geht zu fragen, ob überhaupt zu einem Subjekt Status oder einen Verantwortungs Status einem Roboter zugesprochen werden kann und sollte. Dieses Buch habe ich gelesen. Das war ein sehr wilder Ritt durch die ganze Roboter Ethik Geschichte. Da habe ich auch selbst viel gelernt, und sie geht halt so weit zu sagen, ja, also, wir müssen (...) so denken und auch menschliche Kategorien von Verantwortung (...), Akteursschaft und ja auch wirklich (...) ethischen Subjektstatus auf Roboter anwenden. Da bin ich jetzt nicht mitgegangen bei all den Argumenten, aber ja, ein bisschen kenne ich mich aus in der Debatte. (Until 00:01:23)

Marc Lewald: Okay, perfekt (...) Dann würden wir wahrscheinlich einfach mal mit dem naja, deutsch, Handlungsvermögen starten. Ich würde sagen, wir können es auch einfach agency nennen, der einfach halt halber. (...) Was würde es denn für sie bedeuten, wenn wir einem humanoiden Roboter Agency zusprechen würden? (Until 00:01:43)

(2) Expert 1: Ich glaube, dass das eine Sache ist, die (...) noch am unproblematischsten ist, weil letztendlich ein Roboter wird angetrieben von einem Algorithmus, egal wie (/) oder wie komplex er ist, so (...) was sich unter KI versteckt, was oft in der Debatte schwimmt zwischen humanoider Robotik (...) und KI Anwendungen (...) ist ja einfach nur statistische, komplexere Algorithmen, die auf sehr vielen Daten basieren und damit interagieren. Und letzten Endes ist ein Algorithmus nichts anderes als, ja, eine Handlungsaufforderung, also von Schritt A, BCD (...) sind es letztendlich einfach Rechenschritte. So, und in dem Sinne handelt auch ein sehr simpler Taschenrechner, so, also super rudimentär. Und in dieser sehr einfachen oder simplen Verständnis von Akteurschaft (...) kann man auch, glaube ich, davon sprechen, dass ein Roboter agency hat oder handelt. So, aber es ist in dem Sinne sehr eingegrenzt, weil es sind determinierte (...), ja Phänomene, also das heißt, diese Maschinen sind determiniert, auch wenn sie sehr komplex sein können, so, (...) weil es nichts anderes ist als rechnen, so. Es ist auf Modellen und Daten basiert, die rechnen, und man muss es auseinander halten von (...), ja, komplexeren menschlicher Akteursschaft, die auf Emotionen, intuition etc. basiert. Aber wenn man das so

versteht, Akteursschaft (...) weil ein Algorithmus handelt, kann man das auch, meines Erachtens ein Roboter zu sprechen. (Until 00:03:28-4)

Marc Lewald: Okay (...), schon mal schon mal sehr spannend. Die Handlung selbst, die hatte ich auch mit meinem Professor als Debatte vor einem Jahr und hab mich dann in einem Paper nochmal ein bisschen näher damit beschäftigt und hab da dann auch ... daher kommt die dritte Frage, (...) die Frage gestellt, ob denn das wirklich als eine Handlung oder, besser gesagt im englischen, als eine Action oder als eine Operation (...) beschrieben werden kann. Haben sie dazu irgendeine Meinung, oder du? Entschuldigung. (Until 00:04:02-6)

(3) Expert 1: Ist philosophisch jetzt, auf jeden Fall, (/) Die Turing Maschine in dem Sinne kann ja maschinell, also kann, ja (...) ist ja ein Gedankenexperiment zum Beispiel, was Turing da aufgestellt hat. Es gibt sie aber auch (...) als, (...) ja, reelle, also (...) einfach als Artefakte, so das ist dann ne Stempelmaschine, die dann so das Band weiterzieht und das ist halt in dem Sinne (...). Auf jeden Fall ist es ein Prozess, so. Das würde ich auf jeden Fall sagen. (...) Ja, es kommt sehr darauf an, wie man halt durch den Handlungsbegriff definiert. So, (...) in dem Sinne ist es schon richtig, da sehr explizit zu sein. (...) Ja, handelt eine Maschine in dem Sinne? (...) Ja, ich würde da auch vorsichtig sein, klar, also, ist die Frage (/) Ich weiß auf jeden Fall, es gibt so seichte Formulierungen, so, die handeln da (...) darauf mit beziehen, so, die diesen Begriff mit einbeziehen würden. Gerade in der Roboterethik wird das sehr schnell gemacht (...), und das ist in dem Sinne auch sehr phänomenologisch gesprochen. Das heißt, es geht sehr um die Erscheinung, letzten Endes. Es ist nicht wichtig zu wissen, was das antreibt, (...) diesen Prozess. Es reicht, wenn wir etwas sehen (...), das (...) auf ein Ziel gerichtet, sich bewegt oder hin rechnet oder irgendeine Aktion macht. (...) Was das ist, so, kann vielen Denker/-innen, also ist vielen Denker/-innen egal, so. (...) Hauptsache, etwas in dem Sinne scheint zielorientiert zu handeln, (...) und das ist auch ein leichter Weg, den viele Roboterethiker/-innen meines Achters gehen, indem sie sagen, ja, das erscheint halt so, als ob es handelt, als ob es Emotionen hätte, und deswegen ist es ethisch wichtig, (...) ganz egal, was das ist, was diese Erscheinung auslöst. (/) , wie gesagt, ein determinierter Algorithmus, was Mathematik ist, was rechnet, (...) und ich glaube, darauf sollte man achten letzten Endes aber (...) ja, also die Turing Maschine letzten Endes ist ein Prozess, der mechanisch (...) umgesetzt wird, der umgesetzt werden kann von einem Gedankenexperiment in die Realität, und (...) es kommt auf den Handlungsbegriff an. Ein Prozess ist auf jeden Fall. (Until 00:06:28)

Marc Lewald: Okay, also, ich habe es während der Arbeit, soweit ich die Arbeit schon mal geschrieben habe, auf jeden Fall schon mal als Aktion definiert und auch damals im Paper dann als Aktion definiert. (...) Wobei es, ja, in den meisten Fällen einfach auf

die Definition drauf ankommt, würde ich sagen, (...) genauso wie eigentlich mit der nächsten Frage des Bewusstseins, die auch, ja, je nachdem, in welche Richtung man jetzt geht, (...) völlig unterschiedlich definiert wird. Deswegen die erste Frage, wie Sie denn Bewusstsein, wie du, Entschuldigung (...), Bewusstsein, definieren würdest. (Until 00:07:09)

(4) Expert 1: Ja, das ist halt natürlich schon eine riesige Box der Pandora, dieser Begriff (...). Also es ist interessant, dass sofort mit der Begrifflichkeit der KI, oder der Begrifflichkeit des Roboters, gerade des humanoiden Roboters (...) sofort diese Bewusstseins Begriff eine Rolle spielt. Das ist irgendwie sehr (...), ja sehr (...) aussagend, finde ich so, für den ganzen Diskurs, (/) der in den 1940ern 50er Jahren mit der Turing Maschine eigentlich (...), ja, aufgekommen ist. Man hätte das ja ganz anders machen können, hätte ja gar nicht einen Turingtest machen müssen, wo es darum geht, dass eine Maschine, was besser macht als der Mensch, man hätte auch eine ganz andere Benchmark setzen können, was eine Maschine leisten muss, um als intelligent zu gelten. Aber diese Bewusstseins-Konzepte schwirren schon in den 40, 50 er Jahren mit (...), oft mit Bildern vom Gehirn, mit dem Perceptron, wo dann sofort dieses Neural Network aufkommt, und das sind Metaphern und Bilder (...), die dann auch aus der Wissenschaft selbst gekommen sind in den 50, er, 60 er Jahren, also von KI-Forscher/-innen, also von Programmierer/-innen, (...) die dann auch zeitgleich von Science Fiction und sehr viel SciFi in Film aufgenommen (...) wurden, und das schwimmt sehr stark (...). Wie man jetzt Bewusstsein genau definiert (...), ja, wir sind, wir wissen da noch überhaupt nichts drüber, glaube ich so als Menschen. Wir sehen halt letztendlich ja, da poppt irgendeine Regionen im Gehirn auf, (...) die können wir im MRT sehen, dass die aktiv ist. (...) Aber die phänomenologische Erfahrung, also wirklich die Erfahrung vom Bewusstsein, ein in der Welt sein, wo wir Zeitlichkeit, Räumlichkeit, Emotionen, das wahrnehmen, das können wir nur beschreiben durch Konzepte, die wir bilden, durch Wörter, durch Ansprache. (...) Wir können es aufmalen, wir können andere Zeichensprache nehmen, aber wir können das nur versuchen auszudrücken, und all diese Versuche (...) sind hoffnungslos unterlegen, eigentlich diese Erfahrungen, ja, zu repräsentieren. (Until 00:09:27)

(5) Und klar, gibt es jetzt so hardcore Silicon-Valley Leute, die sagen, ja, das ist alles nur Mathematik im Kopf, das ist alles nur Chemie, das kann man alles coden, (...) und deswegen können wir auch irgendwann unser Bewusstsein hochladen (...) auf nen Chip oder so. Das ist (...) meines Erachtens eine totale Reduzierung, also ein sehr großer Reduktionismus, irgendwie, wo wir noch überhaupt nicht so weit sind, um überhaupt so eine Aussage machen zu können, weil wir eben bis jetzt nur Korrelation sehen. Wir sehen: Ah, da ist eine Region im Gehirn. (...) Wenn die aktiv ist (...), dann fängt der Mensch an zu weinen oder so, dann ist er emotional oder fängt eher an, sich zu erinnern oder es werden einfach nur verschiedene Areale aktiviert (...) und

menschliches Bewusstsein, da ist die Forschung noch so krass am Anfang, und es kann auch sein, dass wir das nie entziffern werden, so. Wir sehen auf jeden Fall, dass wir einen sehr krassen Gehirn Fokus haben, obwohl auch anscheinend sehr viele (...) ,in , (/) auch in anderen Nervenzentren im Körper verortet sind, (...) also irgendwas mit der Niere oder so, als es ist überhaupt nicht so, dass das Gehirn ... Die Idee ist, man kann das Gehirn nehmen, irgendwo einpflanzen, und dann ist Person A mit dem Gehirn zusammen ... Person im Körper B, so. Das ist einfach (...) Frankenstein (...) Idee, die aus der Science Fiction kommt (...). Aber soweit ich das verfolgt habe, die Gehirnforschung oder auch die (...) , ja, die Neurologie zeigt immer mehr, dass es ein ganz körperliches Struktur, also (...), ganz körperliches Phänomen ist, (...) und du musst in der Welt sein, also du musst diese Welt erleben, und dieses Erleben ist ein Prozess (...), der geht dann auch sehr stark mit dieser Frage zusammen, was ist überhaupt dieses Leben (...), so? Also das ist, wir würden ja sagen, okay, kann es auch Leben ohne Bewusstsein geben? Es geht dann auch in diese Richtung, wird dann noch komplexer. (...) Aber ich würde einfach jetzt mal sagen, dass wir von dieser Frage sehr, sehr, sehr weit entfernt sind. Auch ist interessant, wieso stellen wir uns die Frage, also haben wir Angst, dass dann so eine KI oder so ein Roboter bewusst wird und dann (...) andere Menschen auslöschen könnte? Das ist oft so eine Gefahr, die damit einhergeht, in den Medien (...). Oder klar, wenn jemand Bewusstsein oder etwas Bewusstsein hat, dann kann man auch drüber sprechen, hat es einen moralischen Wert oder hat es irgendeinen? Ist es ein Rechtssubjekt? (...) Aber wir sind da total weit von entfernt, (...) und ich glaube, wir sollten einfach, (...) aufhören darüber zu fabulieren, letzten endes, weil wir das noch so wenig verstehen, dass wir (/). (Until 00:12:08)

Marc-Aurel Lewald: Ja, okay (...). Jetzt vor allem in Bezug auf die humanoiden Roboter oder halt (...) KI allgemein. Du hast das ganze schon so ein bisschen sarkastisch angehaucht ... Du bist nicht der Meinung, dass wir jetzt in den nächsten 50 oder 100 Jahren (...) von bewussten Maschinen sprechen könnten. (Until 00:12:31)

(6) Expert 1: Nee, glaub ich nicht, also dann (...) ja. Das, wie gesagt, wenn du halt (...) Bewusstsein wie du, wenn du Bewusstsein super schwach definierst, dann kannst in diese Richtung gehen (...). Aber ich (...), also erst glaube ich, dass wir ganz andere Probleme haben in der Gesellschaft als das, so, nämlich dass viele andere Machtstrukturen, Armutsstrukturen, geopolitische Konflikte, und dass wir dahin kommen, dass es irgendwie auch ja ein Wunschdenken, was sehr stark aus dem Silicon Valley und aus Oxford und Harvard kommt, wo sehr viele Denker/-innen sitzen, so, die aus diesem Human-Enhancement und (...) Human-Transandalem, so, aus dieser Richtung kommen, glaube ich. Und (...) nee, ich glaube also, wenn du auf diese Frage ganz genau mit ja oder nein antworten möchtest, würde sagen nein, das ist nicht etwas, was passieren wird. (...) Wir verstehen viel, viel zu wenig darüber (...), und ich

glaube auch nicht, dass das eine Angst ist, die wir haben sollten. So, wir haben viele große andere Probleme, das ist kein Problem, was wir gerade haben. (Until 00:13:40)

Marc-Aurel Lewald: Da würde ich absolut zustimmen (...). Deswegen habe ich auch mein Studium an der Karlsruhochschule begonnen vor ein paar Jahren. Ich weiß jetzt nicht, ob du aus Karlsruhe kommst und ob dir die Karlsruhochschule was sagt. (Until 00:13:51)

(7) Expert 1: Ja ich weiß wo sie ist, also ich wohne auch in Karlsruhe. Das ist ja auf der Karlstraße ... das ITAS (...) ist ja auch auf der Karlstraße und wir sind weiter oben, Richtung (...) Zentrum, Richtung Europaplatz (...). Ich bin schon mal vorbeigelaufen, drin war ich noch nicht. (Until 00:14:08)

Marc-Aurel Lewald: Also, solche Themen sind, würde ich sagen mal, Hauptstudiengangsinhalt von gefühlt allem. (...) Müsste dann aber (...), wenn wir jetzt wieder zurück zur Frage der Roberrechte kommen (...), müsste man dann überhaupt ein Bewusstsein definieren bzw. diesen Maschinen ein Bewusstsein zusprechen, dass man über Roboterrechte sprechen könnte? Oder (...), ist es auch völlig unabhängig von einander möglich? (Until 00:14:39)

(8) Expert 1: Ich meine, es ist interessant sich anzuschauen, wie wir mit Tieren umgegangen sind historisch, weil da (...) oft auch mal historisch gesehen, schon im Mittelalter und auch im Christentum (...), immer, was macht den Menschen in Unterscheidungen zum Tier aus, und da wurde halt oft auch fabuliert (...), der Mensch ist rational, der Mensch kann empfinden, kann Gefühle haben (...), und da gab es verschiedene (...), ja (...), verschiedene Paradigmen, wo man glaubte (...), dass letztendlich (...) das Tier nur eine Maschine ist und das kann nicht fühlen, das (/) nur das, sozusagen nur wie ein Apparat, wo du drauf drückst, und dann kommt was raus, letzten Endes (...) von Demetri und sehr viele Materialist/-innen, so damals in dieser Zeit meistens Männer, Materialisten, so im 18, 17, Jahrhundert, und da hat sich auch sehr viel getan, und es ist interessant dann, dass dann dieser (...) Diskurs, so (...), wird immer noch diskutiert, ob Tiere Rechte haben (...), und da sehen wir ja sogar das Schweine zum Beispiel, mit denen wir furchtbar umgehen eigentlich in der ganzen Schlachthöfe und so, die, die erkennen sich im Spiegel und da (...) würden paar Hinforscher sagen, die haben auf jeden Fall so ne schwache Ausprägung von Bewusstsein. (...) Wir würden aber auch sagen, allein weil sie schmerzen empfinden, so, haben sie Rechte und es ist utilitaristisch gesehen oder tugendethisch gesehen oder kantianisch gesehen, egal, wie du es ethisch herleiten möchtest, (...), gibt es einen soll Zustand, dass sie mit Tieren gut umgehen, so, weil sie schmerzen empfinden, so, ganz egal, wie bewusst sie sind, (...) bei Robert, ist es jetzt interessant, dass auf einmal alle von Robotern sprechen, und mit Tieren haben wir eigentlich unser

Verhältnis überhaupt nicht geändert. So, wir werden immer noch ziemlich furchtbar eigentlich mit denen umgegangen, (...) und jetzt, Roboter erscheinen halt menschlich, weil sie in diesem Sinne Humor imitieren oder halt (...) menschliches Antlitz haben (...), davon sprechen, dass sie Gefühle haben in (...) werden sie ja benutzt auch. Es gibt ja sogar Sexrobotik. Es gibt Robotik in Care, etc., in Kindergärten, in der Bildung. Das heißt, wir sehen immer mehr, das sind sensible Bereiche, (...) Roboter benutzt werden, und (...), weil aber, und das ist jetzt meine Einschätzung oder meine Bewertung, Roboter weder fühlen (...) noch Schmerzen empfinden können (...), also ist eigentlich das alles nur imitieren, so, sind nichts anderes als (...) , ja, Rechenmaschinen. Wie gesagt, die algorithmisch funktionieren (...). haben diese Maschinen eigentlich meines Achtsens kein Anrecht, einen Rechtsstatus zu bekommen, so, sollten sie auch nicht. Wenn man das täte (...), glaube ich, vermindert man auch den Rechtsstatus von Tieren und auch von Menschen, weil dann einfach dieser Rechtsbegriff (...) so weit gedehnt wird, was überhaupt schützenswürdig ist als Gut (...) und das finde ich problematisch. Klar ist, dass Menschen sehr viel drauf projizieren (...) Alzheimerpatienten, und so fangen an, von ihrer Robbe, ihrer, was weiß ich, ihrer Roboter Robbe dann da zu weinen, wenn sie weggenommen wird. Natürlich projizieren wir da sehr viel hinein, und das führt dazu, dass paar Roboterethiker/-innen sagen würden, ja allein weil Roboter in uns was auslösen, Emotionen (...), sollten wir (...) darüber nachdenken, dass diese Roboter auch Rechte bekommen, und dann kann man auch sagen, okay, jemand, der einen Fetisch hat vom Auto, fängt an zu heulen, wenn er das Auto ein Kratzer bekommt oder so, dann kann man es ganz weit abstrus führen und so (...), und deswegen finde ich das problematisch. Ich halte das für einen zu krassen Aufdehnung (...), wer eigentlich rechts oder was Rechtssubjekt ist, dann hätten Tiere meines Erachtens viel mehr Schutzstatus verdient (...) und es wird dann auch eine riesige Pandora box von Verantwortlichkeit. Also, wenn dann ein Roboter irgendeinen Mist macht, indem man halt zufällig (...) in der Fabrik jemanden umbringt oder so oder verletzt (...), weil halt die Maschine etwas falsch erkannt hat, oder so, (...), als ein einfach so einfaches Beispiel, (...) dann würden wir jetzt nicht den Roboter vor Gericht stellen, so. Wir würden immer, und das finde ich auch sehr wichtig, (...) Menschen, das heißt entweder der/die Betreiber/-in dieser Firma oder die Entwickler/-innen zur Verantwortung ziehen oder eben (...) die Betreiber der Fabrik, die das halt auf die Mitarbeiter losgelassen haben. Aus diesem Grund finde ich das sehr problematisch, Robotern, Rechte oder ein Subjektstatus zuzusprechen. (Until 00:19:55)

Marc-Aurel Lewald: Okay, finde ich schon einmal eine super, super spannende Ansicht, also vor allem in Hinsicht der Verminderung der naja Menschen und auch Tierrechte (...), vor allem, weil ich über die Tierrechte auch einen Teil geschrieben habe. Am Beispiel der Fische (...), wo halt die Schmerzen noch ein bisschen dazukommen, weil

wir denen ja auch kein Schmerzempfinden zusprechen (...), und ein paar Definitionen des Bewusstseins, mit denen ich arbeite, haben halt auch über Schmerzempfinden (...) was zu tun. Da werde ich auf jeden Fall noch einmal bisschen tiefer reingehen, finde ich super interessant (...). Ist ja aber auch schon dann im Prinzip die Meinung zu Roboterrechten mehr oder weniger (...) klar, würde ich mal sagen (...), vielleicht jetzt noch eine sehr philosophische Frage, kommt natürlich auch immer drauf an (...), welchen Bezug man zum Posthumanismus hat und welche man sich da selbst beigebracht hat. Aber inwiefern würde denn eine posthumanistische Perspektive die aktuelle Debatte auf Roboterrechte beeinflussen? (Until 00:21:06)

(9) Expert 1: Also ich finde (...), da ist die Frage, wie man auch da wieder Posthumanismus definiert, weil es gibt halt so einen Silicon-Valley Posthumanismus, der halt eher so in Human Enhancement geht. (Until 00:21:22)

Marc-Aurel Lewald: Ja, Transhumanismus, dann ja. (Until 00:21:24)

(10) Expert 1: Genau Transhumanismus, weil es gibt auch eine Art neuen Materialismus, der aus dieser feministischen Perspektive kommt, und er kommt eher aus dieser Ecke von Actor Network Theory, von Latour und auch von Menschen wie Barret, die haben ja, die denken neu über Akteursschaft und (...) Verbindungen zwischen Mensch und Maschine und der sozialen Embeddedness, nennen die das (...), nach, und das ist schon ein ganz großer Unterschied. Weil es gibt dann so Menschen wie Harroway oder ja, Barret, die eben aus einer feministischen Perspektive darüber nachdenken (...), den Robot, das robotische Artefakt, neu zu situieren und auch die politischen Machtstrukturen hervorzuholen, so. Und dann gibt's halt, die Human Enhancement und Transhumanisten, die kommen eher aus dem Silicon Valley, die haben halt (...) , ja, einfach Vorstellungen davon, den Menschen auf eine neue evolutionäre Spezies, eine neu evolutionären Status hin (...) zu (...) heben, sozusagen wirklich als ein sehr krasses lineales, ein sehr progressives Geschichtsverständnis, so, von Fortschritt und von technischen, und (...) da wird halt der technische Fortschritt mit humanen Fortschritt, humanitärem Fortschritt so gleichgesetzt (...), und das wurde auch schon sehr viel kritisiert, so (...). Und gut (...), die Frage, die du jetzt gestellt hast, ist, ob aus dieser Sichtweise (...) kannst du nochmal sagen, ob auch diese Sichtweise, diese Roboter dann? (Until 00:23:05)

Marc-Aurel Lewald: Naja, im Prinzip also (...), worum es geht (...), ob mit der Einführung oder nicht Einführung von Roboterrechten oder allein mit der Diskussion unser Verständnis von Menschlichkeit oder halt nicht menschlichem Leben verändert werden würde? (Until 00:23:20)

(11) Expert 1: Ja, das sicherlich, da sind auch dann letztendlich, wenn es um Verständnis geht, die Theorien produktiv, so (...), wenn sie halt darum geht, neu

nachzudenken (...). Wie verortet sich Mensch-Mensch sein, Tier-Tier sein, Roboter-Roboter sein? Und glaube, da in dem Sinne sind sie produktiv, indem sie darüber neu anstoßen, was Mensch sein überhaupt ausmacht auch. Das wird unwiderruflich in Frage gestellt, alleine, dass dann halt irgendwelche Alzheimerpatienten anfangen zu weinen, wenn ihnen eine Roboter Robbe weggenommen wird und , so, kommt das halt schon auf diese Frage. Ouh, eine Person ist sehr verletzt, emotional ist da sehr (...), ja sehr auf einmal drinnen in dieser Situation und (...) sehr emotional anhänglich, so, und auf einmal kommt dann diese Frage auf, was macht den Menschen eigentlich zu Menschen und den Roboter zum Roboter? (...) Und eine Akteur Netzwerktheorie von Bruno Latour oder auch dieser neue Materialismus, finde ich, hat eine sehr produktive (...) Diskussion ausgelöst, dass wir auch den Roboter (...) mehr situieren, von wo er herkommt, was ihm eingeschrieben wird, und dass er immer auch eine gewisse Abhängigkeit hat zu der Welt, weil ein Roboter ist eigentlich nie in dem Sinne autonom. Also, ein Roboter ist genauso abhängig von Infrastruktur (...), von Datensätzen, von Halbleiterproduktion, von Energie, die er sich aufladen muss, also ein Roboter braucht Wartung etc. Ein Roboter ist in dem Sinne immer so dargezeigt, als Mensch oder Maschine, A auf A, B auf A, und das ist eine Perspektive, die dann diese Akteur-Netzwerk-Theorie oder diesen neuen Materialismus in Frage stellt (...), weil eben zeigt, dass da die ganze Zeit Interdependenzen sind und dass es ein sehr komplexes Netzwerk ist (...). Von der Gedanke des Ingenieurs bis hin zu Code, von Bias bis hin wie der die Roboter in einem Kontext anders wahrgenommen wird, im Kindergarten bis hin zum Altersheim (...). Und da finde ich ist produktiv, wenn diese Theorien darüber (...), diese Gedankenstöße ansetzen, diese ganzen Human-Enhancement Futuring (...) Gedanken finde ich haben etwas Eskapistisches, so, wie der Silikon Valley damit umgeht, weil das ist einfach vom Problem die wir in der Gesellschaft haben eben diese Machtstrukturen, wer diese Roboter eigentlich designt und produziert (...), und dass wir Probleme haben, mit denen wir nicht gut umgehen können, sei es Armut oder Klimawandel, etc. dass davon so ein bisschen ablenkt, finde ich, und wir sehr viel von der Todes-KI sprechen oder dem Roboter, der bewusst sein soll, und das ist in dem Sinne auch sehr politisch, weil dadurch einfach ein Diskurs gesetzt wird oder ein anderer in dem Sinne (...), gesilenced wird , so, verstummt (...). und da gehen die Medien und Sience-Fiction, Hollywood, alle gehen da sehr krass darauf ein und es hat so was von (...), ja, man kann in (...) Phantasiewelten letztendlich ,so (...), flüchten, um sich nicht von den realpolitischen, aber langweiligen vielleicht, weil schon immer bestehenden Problem, weil Armut gibt es schon, seitdem es Menschen gibt, (...) damit zu beschäftigen. Und dann wiederum ist es nicht nur nicht produktiv, aber es ist sogar glaube ich, gefährlich, so, wenn dieses Human-Enhancement Denken, ablenkt von den wichtigen realpolitischen Problemen die wir haben, so (...), und genau (...), lass einfach mal so stehen! (Until 00:27:08)



Marc-Aurel Lewald: Ja, da würde ich mit so einigen Silicon Valley Größen wahrscheinlich dann auch mitgehen, oder mit dem einen oder anderen Projekt (...). Okay, eigentlich wäre nur noch die Frage übrig, was für Risiken bei der Einführung von Roboterrechten denn letztlich entstehen könnten (...). Oder besser auch die Frage, da du jetzt eher gegen die Einführung der Roboter bist, was passieren könnte, wenn man sie denn eben nicht einführt, ob du da auch irgendwelche Probleme siehst? (Until 00:27:41)

(12) Expert 1: Also muss man unterscheiden zwischen, dass Roboter eingeführt werden oder Roboter Rechte eingeführt werden, auf jeden Fall. Weil Roboter werden, ja, das ist ein Prozess, der scheint, erscheint unaufhaltsam. Aber es ist einfach, weil wir das auch als Gesellschaft alle so mittragen wollen, dass halt immer mehr Intimsphäre durch robotische Artefakte oder halt einfach KI-Applikationen (...), ich meine, aufgrund welcher Definition du nimmst, kannst du auch ein Virus zum Beispiel als Roboter definieren. Also Virus wie Stucknet zum Beispiel kann auch ein autonomes Waffensystem sein (...), weil es ohne menschliche Interventionen (...) bestimmte Handlungen durchführt, also in dem Sinne auch eine sehr. Ich glaube, dass Roboter nicht nur (...), ja, also, wird meistens materiell gedacht als einem Maschinist gegenüber. Es kann aber auch (...) ein Gegenüber sein, was virtuell agiert, so, also, es gibt Menschen, die würden sagen, und Denker/-innen, die würden sagen, auch ein Virus, der zum Beispiel (...) im Iran eingesetzt wurde, um die Nuklearanlagen zu zerstören oder zu stören wenigstens. Stucknet erfüllt (...) die (...) Kriterien eines autonomen Waffensystems, was wir sonst eher als, ja, Waffensystem sehen, was eine Maschine ist aus Metall und vor uns steht, sozusagen. Das heißt auch da ist es sehr breit gefasst und dieser Prozess aber, dass immer mehr diese Art von Artefakte in die menschlichen Sphären eindringen, scheint unaufhaltsam, weil wir immer mehr das auch zulassen als Menschen (...). Das heißt, das führt sehr viele Probleme mit sich und auch Chancen mit sich. (...) Wir haben erwähnt, ja, schon Bias, Diskriminierung, dass Menschen auf einmal empathisch sehr krasse Beziehungen zu diesen Robotern oder zu diesen Artefakten entwickeln (...), dass (...) ja, wir haben auch darüber gesprochen diskursiv, andere Probleme so ein bisschen ,ja (...), ausgeblendet werden (...). Es gibt sehr viel die Frage, nimmt das unsere Arbeitsplätze weg, was macht das mit (...) diesen verschiedenen Bereichen, auch (...) in der Bildung zum Beispiel, wenn Roboter (...) jetzt zum Beispiel eingesetzt in dem Kindergarten, im KIT gibt es das, es gibt einen Termin Asfur, das ist ein Roboter (...), ja Mechaniker, der jetzt gerade real Robotik zum Beispiel im Kindergarten Roboter loslässt und guckt, wie die Kinder damit interagieren. Das sehen wir auf jeden Fall, und all diese Probleme aber, glaube ich, werde nicht darum gelöst, wenn wir Roboter Rechte (...), ja (...), eigentlich wenn wir Roboter Rechte einführen, weil das ist ein Probleme, das sind menschliche Probleme, das sind eigentlich nicht Probleme des Roboters (...). Dem Roboter ist alles egal, der

kann nichts fühlen, der kann auch, der hat keinen Willen in dem Sinne, dass ist nur ein Algorithmus (...) so, egal wie komplex oder egal wie simpel diese Roboter gestrickt ist. Das heißt, den Robotern ist eigentlich alles total Wumpe, ist jetzt ganz plakativ gesagt, es geht darum, wie wir als Gesellschaft leben wollen mit Maschinen, mit KI und mit algorithmischen Systemen (...) und da müssen wir uns halt fragen, wie weit wollen wir Intimität zulassen? Wie weit wollen wir Bias zulassen? Wie weit wollen wir überhaupt (...) Nur auf der Couch hocken und Netflix gucken, während robotische Systeme unsere Milch kaufen und uns Oma spazieren gehen und alles andere machen? Das sind Fragen, wie wir leben wollen. Und die werden aufgeworfen und dadurch allein, dass Roboter immer mehr in unserem Lebensbereich intervenieren (...). Aber Roboterechte werden dieses Problem nicht lösen und darauf auch keine Antwort geben meines Erachtens, weil das sind Fragen, die Menschen betreffen und nicht Roboter. (Until 00:31:40)

Marc-Aurel Lewald: Okay (...). Finde ich super interessant die Perspektive. Zumal ich auch vor (...), es dürfte jetzt auch schon zwei Jahre her sein (...), ein Paper darüber geschrieben habe, welche philosophischen Prinzipien denn überhaupt in Code eingearbeitet werden können. Um erst mal zu sagen, okay, weil jeder möchte dann zum Beispiel dieses typische Beispiel: man hat den Zug, welches Gleis nimmt er, fährt über Kinder, sonst was und im Prinzip dürfte es ja letztendlich eh nur vom Coder abhängig sein, was die Maschine macht, und vom Code selbst. Von daher finde ich das eine super, super spannende Perspektive, das alles dann ein bisschen mehr auf den Menschen zu beziehen. (Until 00:32:24)

(13) Expert 1: Ja, das ist, also, ganz kurz dazu. Ich finde es auch ganz wichtig, weil es gibt Roboterethik, die dann immer sagt, so das (/) Problem, was passiert, die Oma umbringen oder das Kind und so, und das ist für mich auch so eine totale Engführung der Frage. Es geht um viel mehr als das. Es geht darum, dass wir unsere Ideen von Intimität (...), von Zeitlichkeit, von Räumlichkeit vollkommen in Frage stellt, dadurch das algorithmische Entscheidungsfindung, sei es jetzt im Roboter oder sei es nur eine Applikation als Timeline in deinem social media Profil, immer mehr rein intervenieren. Und das sind keine ethischen Fragen, das sind auch normative Fragen und so, und da (...) ja, bin ich auch im Klinsch mit ein paar Leuten im Institut, die immer versuchen das so runter zu brechen auf jetzt die Oma oder das Kind. Das ist nicht die wichtige Frage, das ist Gedankenspiel, so (...). Wichtig ist, so, wie viel Zeit hängst du am Tag am Bildschirm überhaupt und vor zehn Jahren hast du es noch nicht gemacht. Wieso ist das so? Und das sind so (...), ja, es ist interessant, auch, welche Fragen wir stellen, überhaupt (...). Aber sorry, ich hab dich gerade unterbrochen. (Until 00:33:26)

Marc-Aurel Lewald: Nein, alles gut, alles gut! Aber es klingt so, als, also das ist jetzt vielleicht ein bisschen was persönliches, als wärst du auch ein bisschen mehr zurück

zu den Basics (...), bisschen weniger Bildschirmzeit, bisschen weniger sonst was. (Until 00:33:39)

(14) Expert 1: Also was heißt zurück zu den Basics? In diesem Sinne sollte sich jeder fragen (...), wie soll ich es ausdrücken? (...) Ist er damit zufrieden, dass er jeden Tag sechs Stunden am Bildschirm ist? Weil jetzt endlich wissen wir natürlich (...), kein Mensch kann dem entgehen, dass Social-Media, Smartphones und (...), ja, auch robotische Artefakte so programmiert und so mit Interfaces (...) so dargestellt werden, dass wir die meist möglichste Zeit mit diesen Dingen verbringen, so, wir werden dahin genuged ohne Ende. Es fängt an mit Infinite-Scrolls, mit den Farben, wie die Buttons sind, wie der Roboter uns anlächelt und so, das geht alles in diese Richtung (...). Unsere Interaktion sind die Daten, die generiert werden, mit denen Silicon Valley Geld macht, so, das heißt, wir machen eigentlich umsonst Arbeit für die mit unseren Daten generiert werden (...), und weil es hat ein Suchtpotenzial, es ist einfach so. Kann kein Mensch entgehen, weil es einfach das sehr krass auf (...) unterbewusste Signale anspielt, oft, wie wir Menschen einfach funktionieren, so, weil wenn Roboter in uns was zum Lachen bringt, dann lachen wir und dann fangen wir an, emotional eine Verbindung aufzubauen, das ist einfach so. Und die sind sehr gut, (/) sehr viele Psychologen die da arbeiten im Silicon Valley, genau das zu triggern (...). Und für mich ist immer noch so, dass ich halt doch, auch wenn ich es gut finde, ja viel machen zu können mit dem Computer und mir die Arbeit sehr viel erleichtert und ich auch sehr flexibel sein kann und auch manchmal ChatGPT benutze zum Beispiel, um (...) was weiß ich, um einfach meine Referenzliste in ein anderes Format irgendwie umzugestalten. Man sehr stark darauf aufpassen muss, finde ich (...), benutzt man selbstbestimmt und wirklich kontrolliert diese Medien und tut man das so viel, wie man wirklich möchte? Ich glaube, die meisten würden sagen, oh \*\*\*\*, ich wusste gar nicht, dass ich schon wieder sieben Stunden vorm Smartphone sitze, so, und das passiert jedem, einfach, weil die Dinge so gebaut sind (...) und dieses Bewusstsein, wenn Menschen das für sich selbst klar machen können und diesen Schritt selbstbestimmt entscheiden, ja, ich möchte so viel vorm Pc sitzen, dann ist alles okay. Aber ich habe das Gefühl (...), Menschen wollen das eigentlich nicht und ertappen sich wiederum dabei, und das ist eben was anderes. Das ist halt eine Idee von Nudging und Manipulation. Und ich habe für mich halt gemerkt, ich bin schon 6-7 Stunden, arbeite ich mit dem Laptop, und das reicht mir. Ich möchte nicht noch mehr Screentime haben. Dann möchte ich lieber halt rausgehen in die Welt und andere Sachen machen (...). Und ja, wenn du jetzt mich darauf fragst, so da versuche ich halt, meine Screentime sehr Kass zu messen und einzuschränken, weil ich weiß, wie hoch das Suchtpotenzial ist, und das (...) dem kann kein Mensch entgehen, weil die Menschen einfach so psychologisch gebaut sind, dass wir auf diese Sachen anspringen. (Until 00:36:46)

Marc-Aurel Lewald: Ja, da würde ich voll und ganz zustimmen. Ich habe bisher noch keinen Menschen getroffen, der gerne sieben Stunden am Tag auf einen Bildschirm schaut (...). Okay, dann vielleicht gerade noch eine allerletzte Frage. Wenn es jetzt nicht um die Roboterrechte an sich geht, gibt es denn irgendwelche Rechte, die wir an sich, was diesen ganzen Fortschritt mit Maschinen oder auch Robotern, KI angeht, einführen sollten? (...) Also jetzt nicht mal für Roboter, sondern ja, weißt du was ich meine? (Until 00:37:21)

(15) Expert 1: Ja, das ist ein riesiges Ding, natürlich, also gerade auf EU-Ebene gibt es diesen AI-Act und Digital Service Act und etc. das Digital Markets Act. Das heißt, da passiert gerade sehr viel (...). Das ist auch auf prinzipienbasierte Regulierung so, ist in dem Sinne (...) abgeleitet von Rechten, also Recht auf Selbstbestimmung, Recht auf informationelle Selbstbestimmung, also Sachen wie Privatsphäre (...). Menschen fangen an, mehr zu vertrauen. Was heißt, Kontrolle abzugeben, wenn sie Prozesse nachvollziehen können, weil es ist ja so, dass halt algorithmische Artefakte immer mehr entscheiden (...), was wir sehen, wenn wir Daten, was wir, welche Medien wir konsumieren, und wenn es Transparenz und Nachvollziehbarkeit gibt, das zeigt also empirischen Studien, dann sind wir eher bereit (...), Handlungen und Verantwortung abzugeben, so (...). Das ist doch auch gefährlich. Also ist ein Zankapfel grundsätzlich, dieser ganze Gesetzesrahmen, wie weit sowas gehen sollte. Ja, das ist jetzt eine riesige Debatte, aber ich würde schon glauben, dass es zu einfach ist, nur von (...) Ethik zu sprechen. Deswegen sollte besser von Rechten gesprochen werden, weil Rechte mit einem Rechtssystem geschützt werden können und Ethik letztendlich meistens von Silicon Valley kommt, diese Ethik-Guidelines, oder von dem privaten Sektor (...), der dazu führt, dass wir alle uns darauf einigen, ja, es soll normativ die und die Werte haben, aber es gibt in dem Sinne keine harte Regulierung. Deswegen sind Rechte immer besser als Werte, also ethische Werte, meines Erachtens, das ist schon mal in richtige Richtung (...). Aber wenn es darum geht (...), von Rechten zu sprechen, glaube ich, muss man auch mit hinein beziehen, dass es Konstellationen sind, von (...) wenn jetzt Facebook zum Beispiel oder der Deutsche Staat, das Bundesministerium oder eine Ngo ein Algorithmus anbietet, haben wir schon durch dieses Framing einen ganz anderen Bezug und einen ganz anderen Vorschuss von Vertrauen oder Misstrauen gegenüber diesem Artefakt, algorithmisch gesehen (...). Und da geht es um, wenn es um Rechte geht, der Diskurs zu kurz, weil das sind einfach in dem Sinne auch einfach Machtstrukturen, wie damit umgegangen wird, eben mit Privatsphäre, mit Daten-Ownership und so etc. Und das sind sozusagen politische (...), ja (...), eingebettete soziale Faktoren, welche eine große Rolle spielen (...), wo man mit diesen Rechten, die sich dann eher beziehen, nur auf diesen Algorithmus, ein bisschen zu kurz kommen, so, und das sieht man auch bei dem jetztigen Regulierungsrahmen, wird dann gesagt, ja, wenn zum Beispiel Fake News (...) auf Social-Media verbreitet

werden oder Deep-Fakes verbreitet werden, gibt es (...) gegenüber dem Betreiber der Betreiberinnen, was meistens die großen Plattformen sind, eine Verpflichtung, dass diese runtergenommen werden. So, aber das große Lücke in diesem ganzen Regulierungsrahmen ist, dass sie gar nicht verpflichtet sind (...), ihre ganzen Daten, die geteilt werden, auf ihrer Plattform zu überwachen. Das heißt, es müssen User erst das melden und dann müssen die das runternehmen zum Beispiel (...), was dann dazu führt, dass das auf Papier super gut aussieht, aber faktisch gesehen, weil es eine solche Milliarden Anzahl von verschiedenen Daten gibt im Internet, von Medien, die ausgetauscht werden, schriftlich oder fotografisch, und so, dass dann letzten Endes (...) immer noch so viele Falscheinformationen und Deep-Faked darum schwirren (...). Und das geht dann wirklich auf Machtstrukturen, und da (...) kommt man nicht viel weiter mit diesen Rechten, wenn man nicht (...) auch über Macht, Zerschlagung und Kontrolle, über (...), ja auch wirklich Profitsysteme dann darüber spricht. Weil, wenn man aus deinen Daten Geld macht (...), ja, dann machen Privatsphäre und Datenautonomie wenig Sinn, letztenendes aus dieser Perspektive. Und das sind größere Fragen, die immer damit dieser Debatte rund um Rechte mit eine Rolle spielen (...). Ich habe jetzt ein bisschen ausgeholt und so, aber einfach, weil ich weiß, dass so diese Rechtsfrage eine super wichtiger ist. Also da gibt es zurecht einen sehr großen Diskurs zu, aber oftmals dann, dass die machtpolitischen Fragen so ein bisschen ausgeblendet werden, und die gehören mit dazu, meines Erachtens. (Until 00:42:42)

**Transcript 02: (personal interview)**

Marc Lewald: Wollen wir direkt starten (...), haben sie denn persönliche Erfahrungen mit humanoiden Robotern (...) speziell schon gemacht? (Until 00:00:07)

(16) Expert 2: Ja (...), also dadurch dass ich häufiger über solche Dinge spreche, habe ich auch einige Labore besucht. In einigen solchen Gremien, in denen ich sitze, vielleicht haben sie das auf meiner Webseite gesehen, haben wir uns auch Sachen vor Ort angeguckt. Wir sind in Labore gegangen (...) und haben so ein bisschen auch mitgekriegt, was in der Pflege, zum Teil im In und Ausland, auch an humanoiden Roboter schon eingesetzt wird. Also so Leute wie, also Leute in Führungszeichen wie Pepper zum Beispiel, kennt ja jetzt auch mittlerweile fast jeder. (Until 00:00:38)

Marc Lewald 1: Ja. (Until 00:00:38)

(17) Expert 2: Also das ist so das, was ich kenne. (Until 00:00:41)

Marc Lewald: Ja, genau, solche Erfahrungen wie Pepper haben ja auch schon wirklich viele, viele Menschen gemacht, seitdem er auch im Verkauf steht (...). Welche Bedeutung hätte es denn für Sie, einem humanoiden Roboter ein volles Handlungsvermögen zuzusprechen? (Until 00:00:58)

(18) Expert 2: Also ich würde zuerst zögern, ob ich das wirklich Frank und frei so tun oder sagen würde (...), volles Handlungsvermögen (...). Also wir sprechen vermutlich über das, was das englische Wort von Agency (/) ... Wobei das Englische da nicht so deutlich unterscheidet wie das Deutsche oder gerade auch die Philosophie (...), also zwischen Akten und Handlungen (...). Also wenn zum Beispiel ein Zigarettenautomat ihnen nach Einwurf von Münzen oder wenn sie die Karte irgendwo dran legen, eben, eine Schachtel Zigaretten rausgibt, ist das keine Handlung des Automaten, sondern wenn sie so wollen, maximal ein Akt, ein Vorgang, eine Tat (...). Und sowas gibt es natürlich bei Automaten, bei Maschinen (...) jederzeit. Also jeder Vorgang, der irgendwie als eine Art Veränderung der Umgebung zu verstehen ist, oder eine Veränderung in der Umgebung, die durch die Maschine ausgelöst ist, ist irgendeinen Vorgang, den man als Tat oder bezeichnen kann. Zur Handlung gehört normalerweise in der Handlungstheorie, also mindestens, wenn wir in der Philosophie oder in der Soziologie unterwegs sind (...), so etwas wie ein Handlungsvorsatz, ein Bewusstsein davon, was ich eigentlich erreichen will, in welcher Situation ich handele (...), was möglicherweise Risiken und Konsequenzen von diesen Handlungen sind und wie ich dann vielleicht damit umgehe sozusagen eine Art von Second-Thought wäre vielleicht bei der Handlung eben auch noch dabei (...). Da bin ich mir sehr unsicher, ob es

realistisch ist, dass jetzt oder auch in Zukunft von Robotern anzunehmen. Insofern würde ich dafür plädieren, sehr vorsichtig zu sein, mit der Anwendung des Wortes Handlungsvermögen oder gar volles Handlungsvermögen auf Roboter, wenn man nicht genau definiert, was genau man damit im Blick auf Maschinen meinen kann, sinnvollerweise (...), und das ist etwas, was mir noch eben häufiger eben aufgefallen ist bei ihren Fragen. Ich denke, man kann diese Fragen präzise wissenschaftlich nur dann beantworten, wenn klar ist, also (...) grundsätzlich. Solche Worte wie Handlung oder auch Bewusstsein sind analoge Begriffe. Da kann man also ganz viel darunter fassen und man muss immer genau sagen, meint man Bewusstsein im menschlichen Sinne, meint man eine Simulation des Bewusstseins im menschlichen Sinne auf Maschinenebene, oder ist einem das zu viel?, diese Differenzierung. Dann ist man jedoch praktisch aus der Philosophie raus. Also uns ist es nicht zu viel, möglichst genau zu sagen, wovon wir jetzt sprechen wollen. Es hat natürlich auch den Vorteil, dass man dann Fehler vermeiden kann, die wirklich vermeidbar sind, dass man einfach jetzt Begriffe irgendwo (...) nimmt und anthropologisiert, dass die Schwarte kracht, aber wem es damit eigentlich in der Wissenschaft geholfen? Niemand, und ich denke, wir sollten diesen Diskurs auch gesellschaftlich so führen, dass wir auf unsere Sprache ganz genau achten. Dann wissen wir ganz genau, was wir jetzt hier eigentlich besprechen, was wir klären wollen, (...) und das ist auch, glaube ich, die Grundlage nachher für politische Entscheidungen über bestimmte rechtliche Maßnahmen, sowas wie Governance, Regulierung, whatever. Man muss genau wissen, wovon man redet (...). Ich glaube, mit ein bisschen Mühe kann man das gut hinkriegen. (Until 00:04:17)

Marc Lewald: Ja. (Until 00:04:18)

(19) Expert 2: Und jetzt habe ich noch nicht wirklich darauf geantwortet, was sie gefragt haben. Das war sozusagen der Vorspann (...). Ich würde weder zum jetzigen Zeitpunkt noch in Zukunft im menschlichen Sinne volles Handlungsvermögen einer Maschine zuschreiben, weil ich nicht sehe (...), dass diese (...) ganzen zusätzlichen Faktoren, die für menschliches Handlungsvermögen wichtig sind bei Maschinen irgendwie nachbaubar sind. Jetzt jedenfalls nicht und in Zukunft möglicherweise auch nicht. Aber für die Zukunft will ich mich natürlich nicht so festlegen, wie für den Jetzt-Zustand. Ich sehe jedenfalls nicht, dass wir jetzt schon Roboter haben mit vollem menschlichen Handlungsvermögen. (Until 00:04:57)

Marc Lewald: Okay, entschuldigen Sie, falls die Fragen an der einen oder anderen Stelle vielleicht nicht akademisch genug sind (...), aber ich musste den Fragebogen erst mal so anpassen, dass ich ihn (...) auch an normale Menschen rausschicken kann, die jetzt nicht so viel mit dem Thema zu tun haben sollten (...). Ich unternehme dann auch in meiner Maßarbeit eine Unterscheidung zwischen der reinen Handlung,

also (...) es ist auf englisch immer ein bisschen (...) anders in der Definierung der Begriffe, aber ich unternehme dann eine Unterscheidung zwischen Action und Operation, da das zwischen mir und meinem Professor so eins der größten Unterscheidungen war. Würden sie da einen Unterschied sehen, und falls ja, wo? (Until 00:05:41)

(20) Expert 2: Also, das ist eigentlich das, was ich vorher so ein bisschen angemahnt habe. Also ich glaube, dass Operationen etwas ist, was man von Maschinen aussagen kann, auch von Automaten, zumal ... Bei Menschen würde man eigentlich nicht von Operationen sprechen (...). Höchstens, wenn es auf einer sehr sachlichen Ebene ist, könnte man von einer Art von operativem Vorgang sprechen. Da ist dann auch sehr viel Technik, sehr viel Methode jenseits des eigentlich menschlichen Handelns mit involviert (...). Aber die Aktion oder die Handlung, das ist etwas typisch Menschliches, und ich habe ihnen ja eben so ein bisschen beschrieben, (/) irgendwie klassischer Weise seit der Handlungstheorie von Aristoteles Zeiten an, bis heute, Habermas and beyond sozusagen gehört. (Until 00:06:22)

Marc Lewald: Okay, ja (...), das war nämlich gerade nochmal die Frage. Dann dürfte das genau den Begriff der Operation entsprechen. (Until 00:06:30)

(21) Expert 2: Genau (...), das ist eine super Unterscheidung, das ist wirklich etwas womit man sehr, sehr viel anfangen kann. (Until 00:06:33)

Marc Lewald: Sie haben auch gerade schon über Bewusstsein gesprochen (...). Könnten Sie für sich den Begriff des Bewusstseins zumindest in Bezug auf Roboter (...) und Roboterrechte definieren? Wahrscheinlich dann als technisches Bewusstsein, so wie sie das schon formuliert haben? (Until 00:06:50)

(22) Expert 2: Also auch da noch mal, also der erste Satz. Solche Worte oder solche Begriffe wie Bewusstsein sind analoge Begriffe, sie sind vieldeutig. Man muss auch hier sagen, was man meint und ich würde als zweiten Satz sagen, es gibt Stufen des Bewusstseins. Es gibt nicht Bewusstsein oder Nicht-Bewusstsein. Das ist so eine Verkürzung in der Alltagssprache (...), sondern man könnte sagen und dafür gibt es auch gute philosophische Gründe und gute philosophische Theorien, die das stützen. Es gibt zum Beispiel so etwas wie (...), darüber informiert sein, in welchem Zustand ich mich befinde (...), oder dass überhaupt irgendwas wahrgenommen wird (...), also irgendwie, also, das kann man sich vorstellen (...), irgendwo gibt es eine Screen, und darauf wird was abgebildet, so, das könnte schon mal eine erste basale Stufe von Bewusstsein sein, nämlich da ist etwas (...), was abgebildet wird, was irgendwie vorgestellt wird, so eine Art von Spiegelung von Realität. Jetzt ist der zweite Punkt,



weiß die Screen, dass sie die Screen ist, auf der das vor sich geht, was die Screen sozusagen darstellt? Menschlich gesprochen wissen wir, wenn wir Gedanken haben, dass wir diese Gedanken haben. Also gibt es so eine Art Subjektivität, die dem ganzen Vorgang der Wahrnehmung von etwas zugrunde liegt und der ko-reflexiv mit präsent ist. Das heißt, weiß ich, dass es meine Wahrnehmungen sind, dann würde ich von Selbstbewusstsein oder Ich-Bewusstsein sprechen (...). Und dann gibt es vielleicht noch Stufen, die nach vorne hin eben noch komplexer zu beschreiben sind, ich würde aber erst mal bei diesen zwei basalen Stufen bleiben (...). So etwas wie ein Informiert sein über den Systemzustand, zum Beispiel von einem Roboter oder auch von einem System, nehmen sie auch Smart Citys, also das so eine große Schalte sozusagen (...), die aufgestellt ist und wo irgendwie alle Daten zusammenlaufen, die in der Stadt durch Monitore und Sensoren irgendwie anfallen, dass da irgendwas zusammengeführt werden kann, sodass man sehen kann, es gibt Messdaten. Oder nehmen sie ein Thermostat, wo man dann auch sehen kann, also ich weiß in welchem Zustand dieses Gerät ist oder die ganze Stadt ist, wenn man die großen Daten sind. Das könnte man auf einer sehr basalen Stufe ein Bewusstsein nennen, aber auch noch nicht wirklich menschliches Bewusstsein (...), weil das natürlich neuronal ein bisschen anders aufgebaut ist und dann gibt es auch noch Theorien (...), dass das menschliche Bewusstsein eigentlich nicht nur vom Hirn gesteuert wird (...), sondern sozusagen eine Art von Embodied Consciousness ist, dass der ganze Körper irgendwie der Raum ist, in dem sich sowas wie Bewusstsein abspielt. Das ist noch in der Forschung. Man kann es nicht wirklich sehr präzise sagen, was das eigentlich heißt. Man rechnet nur damit, dass das Hirn (...) das Mindeste ist, was man ansprechen muss, um vom menschlichen Bewusstsein zu sprechen. Wie gesagt, also technisches Bewusstsein im Sinne von informiert sein über den Systemzustand kann ich mir ohne weiteres vorstellen. Ich kann mir auch vorstellen, dass man im Sinne einer Simulation des menschlichen Ich-Bewusstseins eine Maschine, einen Roboter, eine KI so programmieren kann (...), dass sie sozusagen diese Zustände, die sie sonst wahrnimmt, sich selber irgendwie zurechnen kann. Das dürfte irgendwie von der Informatik her, vom (...) Coding her kein zu großes Problem sein, das sozusagen reflexiv zu machen. Die Frage ist nur, in welchem Umfang kann man das reflexiv machen? Es gibt Computersystem, es gibt Roboter, es gibt KI's, die schon aus der Umgebung durch Kameras hauptsächlich oder Sensoren was mitkriegen und sich dann verorten können. Das ist ja so eine Art von Ich-Bewusstsein oder Subjektivität, aber natürlich in dem Rahmen, wie es für Maschinen, für Systeme mit KI möglich ist und was da an Informationen von außen kommt, kommt nur rein (...), über eben Kameras, Sensoren (...), Monitore. Das ist beim Menschen komplett anders. Wir haben eine komplett andere Reizverarbeitung (...), sowohl was unsere Sensorik angeht wie auch das Mentale (...). Da werden wir wahrscheinlich auch in der Medizin, also solange man nicht an lebenden Menschen das (...) auch in ethisch ein bisschen schwieriger Hinsicht

dann sozusagen in Real-Time dann nachvollziehen kann, wie so was passiert. Wir werden es wahrscheinlich nie ganz verstehen. Unser Hirn wird sich nie ganz selber verstehen, wie es das alles macht, und das ist dann nur das Hören. Dann kommen noch die anderen Sachen dazu, von denen wir noch weniger genaues wissen. Also wie gesagt, im Sinne von Simulationen und Annäherungen, von denen ich jetzt nicht irgendwie prozentual sagen kann, das sind so und so viel Prozent der menschlichen Kapazität, ein Selbstbewusstsein. Aber wir bleiben mal bei dem Begriff Annäherung. Das sagt ja schon aus, es ist nicht die volle Sache, aber wir können uns vorstellen Teile davon kann man nachbauen. Das ist ja auch nicht so schwer, neuronale Verschaltungen irgendwie in der Informatik darzustellen und durch Algorithmen dann eben zu simulieren. Sowas kann ich mir alles vorstellen, und ich wundere mich über Leute, die ein Problem damit haben, das anzuerkennen, dass es möglich ist (...) und dass man wahrscheinlich damit auch schon was gutes machen kann. Also man kann auch Sachen machen, die vielleicht nicht so gut sind, aber es gibt auch wirklich positive Chancen. Dann wäre auf einer zweiten oder dritten Ebene die Frage (...), gibt es da irgendwie Leitplanken ethischer Art, wie weit man was treiben kann und soll, für das allgemeine Wohl oder nicht? Aber erst mal, dass man sehr weit, auch auf der maschinellen Ebene so etwas wie Bewusstsein und Ich-Bewusstsein rekonstruieren und simulieren kann (...), würde ich auf jeden Fall dick unterstreichen. Aber wir reden von Simulation, wir reden nicht von Identifikation, von Maschinen, Bewusstsein und menschlichem Bewusstsein. Also den Begriff der Identifikation würde ich hier komplett meiden. (Until 00:12:47)

Marc Lewald: Okay, die Trennung ist schon ganz klar (...). Wäre denn für die, um zum Thema Roboterrechte mal zurückzukommen, eine Form dieses Bewusstseins notwendig für die Zusprennung von Roboterrechten für Sie? (Until 00:13:04)

(23) Expert 2: Also auch da erst mal, wovon reden wir, wenn wir von Roboterrechten sprechen? Reden wir von juristischen Rechten? Reden wir von moralischen Rechten oder reden wir von beidem? Das ist mal Punkt eins (...). Ich vermute Sie würden gerne jetzt über juristische Recht sprechen, oder über moralische, oder über beides. (Until 00:13:22)

Marc Lewald: Wahrscheinlich eine Mischung aus beidem (...), denn ich habe mit meinem Professor mehr oder weniger geklärt, dass ich diese Unterscheidung nicht all zu sehr machen werde, weil sie eben einfach zu sehr in die juristische Schiene geht. Ich werde ein bisschen über Legal-Frameworks sprechen, aber dann wahrscheinlich mich, wenn es darum geht, dann schon mehr auf der philosophischen Seite bewegen. Ja (...), also, es geht dann mehr um das moralische Selbstverständnis, bei der Frage um Roboterrechte würde ich sagen. (Until 00:13:54)

(24) Expert 2: Ja, dann sagen sie ihrem Professor, einen schönen Gruß. Er wird nicht weit kommen mit der Vermischung von Kategorien. Das nutzt niemandem etwas. Es wird auch den Gesetzgeber nicht interessieren und die Wissenschaft, jedenfalls, soweit ich sie vertreten kann (...), eher auch nicht. Man muss schon ziemlich genau unterscheiden (...), ob man zum Beispiel sagt, man rechnet einem KI-System, einem Roboter juristische Rechtssubjektivität in vollem oder in nur partiellem Sinne zu, das ist etwas völlig anderes, als wenn man jetzt zum Beispiel sagt (...), und diese These habe ich ja auch in dem Aufsatz vertreten, den ich ihnen zugeschickt habe. Ich kann mir gut vorstellen, dass man moralische Rechte Robotern zuspricht, zum Beispiel das Recht auch auf Unversehrtheit, zumindest wenn es humanoide Roboter sind, dass man da nicht einfach, so wie Kinder das manchmal auch machen, wenn sie mit dem, was die Alexa macht oder nicht macht, nicht zufrieden sind, einfach mit dem Fuß danach treten oder draufhauen (...), oder das vielleicht auch mit Robotern tun. Also mein Hintergrund ist aber, ich möchte gerne ein ethisches Kontinuum sicherstellen (...), so wie wie man mit Menschen so bestimmte Dinge eher nicht machen soll, wenn man ethisch mit Ihnen umgehen möchte, würde ich sagen, auch mit humanoiden Maschinen sollte man nicht einfach denken, das sind ja nur Maschinen, auch wenn sie Humanoid sind, aber die kann man irgendwie zerstören, schlecht behandeln oder oder was auch immer, weil ich gerade vermute, die Grenze der Wahrnehmung also zwischen dem, was bloße Maschinen sind, auch wenn sie humanoid sind, und Menschen wird ein bisschen flüssig, gerade wenn wir uns dann in einer virtuellen Welt noch bewegen (...). Und ich denke, und da sind wir bei einem wesentlichen Punkt, der für mich mitentscheidend ist für die Rechtsfragen, auch im juristischen Sinne. Also ich verlasse jetzt die (...) rein moralische Ebene. Ich spreche jetzt nicht mehr von moralischen robot rights, sondern von juristischen. Ich würde sagen, man kann Folgendes unterscheiden. Es gibt den sogenannten Property Ansatz, das heißt, für juristische Rechtssubjektivität bei Menschen sind bestimmte Voraussetzungen nötig, zum Beispiel muss man bei Bewusstsein sein, man darf nicht irgendwie dement (...), oder irgendwie im Koma liegen. Dann ist es mit der Rechtssubjektivität, mit der juristischen Handlungsfähigkeit sehr schwierig, sondern man könnte auch sagen, haben denn Roboter menschliches Bewusstsein in dem Sinne, dass man sagen kann, sie können auch volle Rechtssubjektivität haben. Und das haben wir eben gesagt, das kann man eigentlich nicht sagen. Deshalb, juristische Subjektivität oder Zurechenbarkeit von Handlungen an Maschinen (...), eher in einem sehr begrenzten Sinne, würde ich sagen, ist möglich oder nachvollziehbar. Ich sehe nicht, dass wir jemals viel davon hätten, wenn wir volle Rechtssubjektivität an Maschinen zusprechen, weil wir dann (...) auch die Kontrolle darüber verlieren und weil dann auch die Zurechenbarkeit von Handlungen, zum Beispiel, wenn ein Roboter in einer Fabrikhalle jemanden verletzt (...). Ich meine, wie wollen Sie den Roboter zur Rechenschaft ziehen? Stellen sie ihm den Strom ab (...), oder sagen sie, du böser Roboter, du darfst aber jetzt nicht mehr machen? Oder ist es

sinnvoll zu sagen, wer hat den Roboter gebaut, wer hat ihn gewartet, wer hat ihn da hingestellt, wer hat ihn nicht beobachtet, wer hat ihn nicht kontrolliert, und wer ist dann verantwortlich dafür, dass dieser Unfall passieren konnte? Das heißt, diese ganze Dimension menschlicher Verantwortung und menschlicher Zuständigkeit (...) und menschlicher rechtlicher Kompetenz sollte durch Roboter nicht irgendwie zu stark durchlöchert werden. Und gerade im Haftungsrecht ist es ja hochumstritten, was jetzt sinnvoll ist, an Rechtssubjektivität Maschinen zuzusprechen. Die großen Firmen haben natürlich ein großes Interesse daran, dass sie aus der Haftung entlassen sind, wenn sie sagen, ja, das sind unsere Roboter, aber die sind ja so fit und so mental schon fast drauf, dass man sagen kann, die müssen selber gerade stehen dafür, was sie schief machen, schief oder schlecht machen, was schief läuft. Und das ist ziemlich leicht durchschaubar. Ich meine (...), so easy kann man das nicht machen. Das ist wie, wenn sie Waffen herstellen und die gehen nach hinten los. Dann kann die Waffenfabrik auch nicht einfach sagen, ja, es tut uns leid, aber unsere Waffen sind eigentlich immer normal, immer hoher Qualität. Das muss an ihnen liegen, wenn sie die nicht richtig irgendwie anwenden. Es liegt nicht immer an den Anwendern, es liegt auch manchmal an den Geräten selber (...) und wenn die Firmen, die dafür eigentlich gerade stehen müssten, sich so einen schlanken Fuß machen könnten und aus der Haftung aus der Verantwortung raus kämen, dadurch, dass man sagt, Roboter oder Maschinen haben eigene Rechte, eine eigene Rechtssubjektivität und sind strafrechtlich verantwortlich für das, was sie tun, das führt uns, glaube ich, in (...) keiner für die Menschen insgesamt wirklich sinnvollen Lagen hinein. Höchstens große Firmen profitieren davon, dass sie klagen, abschmettern können. So einfach sollte man es diesen Leuten, die ja sehr, sehr viel Geld auch damit verdienen (...), nicht machen. Deshalb bin ich dagegen, ganz entschieden dagegen, volle Rechtssubjektivität an Maschinen zu geben. Ne Teil Rechtssubjektivität kann ich mir vorstellen in dem eben schon beschriebenen Sinne. Und das irgendwie auszutarieren, auch mit dieser moralischen Sphäre, das ist, glaube ich, aus meiner heutigen Sicht die Lösung, die auf der Hand liegt (...). Also dass man sagt, gerade bei humanoiden Robotern ein gewisser Respekt, darüber habe ich ja auch gesprochen in dem, in dem Interview, das sie schon angesprochen hatten in ihrer Mail. Sowas sollte es geben (...), wegen des Kontinuums, dass wir solche Maschinen nicht einfach weil sie bloße Maschinen zu sein scheinen, schlecht behandeln, dass wir uns selber (...) nicht schaden oder Kant würde sagen, das sind wir uns selber als Menschen schuldig. Und sobald und dann kommt neben diesem Property-Ansatz der sogenannte Relational Turn oder der relationale Ansatz. Wir haben (...) Beziehungen zu diesen Maschinen, manchmal, gerade wenn es social robots sind, wenn es humanoide Roboter sind. Die sehen zum Teil aus wie Menschen oder nehmen sie Sexroboter. Da wird ja gerade darauf abgezielt, dass (...) die Leute denken, eigentlich sind es fast Menschen (...), aber sie sind eben nicht so sensibel, deshalb kann man mit ihnen scheinbar Sachen machen, die man mit normalen

Menschen so nicht machen kann. Aber wir bauen eine Beziehung auf, und sobald man solche Beziehungen hat, kann auch Ethik und Moral relevant sein, und solche moralischen Rechte, die einen gewissen Respekt auch vor solchen Maschinen, wenn sie humanoid sind, auf jeden Fall, eben ausdrücken. Das würde ich auf jeden Fall befürworten (...), und ich finde das auch nicht lächerlich. Das ist schon viel (...) und diese juristischen Fragen, wie gesagt, dann wird ihnen der eine oder andere Fachanwalt noch ganz tolle Geschichten erzählen können, was es da alles für Versuche gibt von Firmen, sich aus der Verantwortung zu stehlen durch solche Sachen. Da bin ich sehr vorsichtig, weil also nicht, weil ich grundsätzlich solche Firmen für nicht glaubwürdig halte auf solcher Ebene, sondern weil ich einfach sehe, es geht um sehr, sehr viel Geld und da sitzen einfach harte Geschäftsleute, die wollen ihre Sachen ins Trockene bringen und das Hieb und Stichfest machen (...), manchmal um jeden Preis (...). Da geht der Ethiker natürlich nicht mit oder die Ethikerin, wir sind vorsichtig. (Until 00:21:14)

Marc Lewald: Und das kann ich voll und ganz verstehen. Ich glaube, die Vorsicht ist da auch (...) richtig angebracht (...), um vielleicht zu dem Paper noch mal darauf zurückzukommen, dass sie mir geschickt haben (...). Sie haben da kurz mal über eine Post humanistische Sicht geschrieben (...) und für mich war dann (...) also, da ging es um Cyborgs, Enhancement, was für mich im Verständnis mehr in die transhumanistische Richtung ging. Deswegen wollte ich jetzt einfach mal fragen, also nach meiner Definition, wo sie denn den Transhumanismus einordnen würden oder vielleicht abgrenzen würden zum Posthumanismus. (Until 00:21:54)

(25) Expert 2: Ja, das sind natürlich auch irgendwie strittige Begriffe, aber sie haben schon Recht, worüber ich eigentlich geschrieben habe, auch im Fall von Cyborgs, das ist klar (...), transhumanistisch und post humanistisch, wenn sie so wollen, im strengen Sinne, wären wir da wirklich auf der reinen Maschinen Ebene, vielleicht auch in der in der Extended Reality oder in der Virtual Reality, wo man tatsächlich mit reinen Simulationen (...) von menschlichen Lebensformen, von Personen, von was auch immer, also praktisch auf der Ebene mit Avataren oder wie sich es auch immer nennen wollen, zu tun haben. Das ist schon deutlich was anderes. Aber die also (...), da muss ich mir selber einen Einwand machen. Ich habe diesen Vortrag in der ersten Fassung 1919 im Mai gehalten. Es ist so eine rasante Entwicklung, auch gerade im Bereich der Virtual Reality. Mir war das noch nicht so ganz klar, wie schnell diese Reise dann gehen würde, hätte ich das absehen können? Ich meine, wir sind jetzt schon vier Jahre her, oder mehr als vier Jahre. Dann hätte ich vielleicht ein bisschen vorsichtiger da formuliert. Aber Ihre (...) die Differenzierung, die sie einfordern, ist sehr richtig. Ich habe also in meinem Aufsatz wirklich mehr transhumanistische Phänomene beschrieben. Posthumanistisch, würde ich sagen, da ist die Rechte-Frage noch sehr

viel schwieriger, weil sie wirklich nicht mehr genau wissen (...), hat eine Simulation so eine Art von Subsistenz? Kann die Simulation für sich stehen (...)? Würde man sagen, eine reine Simulation, und wir sind jetzt vielleicht mal wirklich im Bereich von Virtual Reality, wir haben eine 3-D Brille auf, wir sind durch Immersion in einer 3-D Welt (...). Ist das, was wir da sehen (...), noch real genug (...), obwohl es vollkommen virtuell ist?, mal von den technischen Substraten abgesehen. Ist es noch real genug trotz der Virtualität, dass man so was wie Substantialität, Subjektivität hier sinnvoller verwenden kann? Ich würde sagen, nein (...). Und ich würde sagen, das ist auch nicht schlimm. Aber wir brauchen hier nicht einfach nur nein zu sagen. Wir müssen suchen, wie nennen wir das denn? Wir brauchen da einen Fortschritt in der Theoriebildung. Wir brauchen neue Begriffe (...) und das finde ich interessant (...), das wirklich anzugehen. Ich bin nicht dafür, dass wir einfach jetzt, oh nein, das passt alles nicht in die bisherigen Schemata. Deshalb ist es (...) schwierig, und am besten redet man gar nicht darüber, man redet am besten darüber, worüber man leichter reden kann. Das finde ich einen etwas zu denkfaulen Ansatz. Wir brauchen einfach, und das ist auch schön, das ist ein neues Gebiet. Da braucht man neue Kreativität, neue Argumente, neue Plausibilitäten, und da bin ich sehr gespannt. Das ist eine Frage ihrer Generation, da mit guten Vorschlägen zu kommen, und da würde ich sie sehr, sehr bitten und ermutigen, machen sie da was draus! Wir brauchen das, und ich würde mich nicht irgendwie einschüchtern lassen dadurch, dass man vielleicht mit irgendeiner alten Terminologie ein bisschen brechen muss (...). Vielleicht kann man es so machen wie bei der Relativitätstheorie oder bei der Quantentheorie. Die Frage sich gestellt hat, was machen wir denn mit der newtonsche Mechanik oder mit der früheren Geometrie? Dann kann man sagen, das sind Sonderfälle von etwas. Aber jetzt haben wir eine andere Rahmentheorie, und wir halten uns erst mal an die, und wenn die in der Lage ist, Dinge gut einzuordnen, dann nehmen wir das, dann nehmen wir lieber das neuere oder dann nehmen wir lieber das, was mehr irgendwie Erschließungs Kraft hat, als dann bei den alten Sackgassen einfach stehen zu bleiben und zu sagen, ja, hier kommen wir irgendwie nicht weiter. Also, so würde ich aus philosophischer Sicht das sehen (...), kann ihnen nicht garantieren, dass es viele Leute gibt, die das so sagen würden, gerade auch in der Philosophie nicht. Weiß nicht, wie es bei den Informatikern und bei den Ingenieurinnen und Ingenieuren aussieht. Aber ich glaube, das wäre der sinnvollste und rationalste Weg (...), das mutig nach vorne hin für einen Diskurs zu öffnen, und dann sollen die Leute mit den guten Vorschlägen kommen (...), die haben jetzt ihre Chance. (Until 00:26:03)

Marc Lewald: Ich spreche jetzt mal für meine Generation und sage, wir geben uns da gerne Mühe (...), da Sachen voranzutreiben, sonst würde ich ja auch nicht meine Masterarbeit in dem Gebiet schreiben (...). Okay, ich will ihre Zeit jetzt gar nicht groß weiter aufhalten. Ich glaube, wir haben schon über meine meine nächste Frage mehr

oder weniger gesprochen, welche Auswirkungen Roboterrechte aus posthumanistischer Sicht haben. Ich glaube, das haben sie mir mehr oder weniger gerade (...) schon erklärt. Vielleicht noch die Frage (...), welche, welche speziellen Gründe oder Gründe, über die wir jetzt noch nicht geredet haben, für Sie oder eben auch gegen (...) die Einführung von Roboterrechten sprechen würden, vielleicht mal auch mal abgesehen von der philosophischen Richtung, zum Beispiel aus bisherigen Gesprächen (...), oder die Minderung anderer Rechte durch die Zusprechung von Roboterrechten oder ähnliches. (Until 00:27:02)

(26) Expert 2: Also, ich würde für die moralische Ebene von Robot Rights auf jeden Fall sagen (...), so was die moralische Respektrechte sollten wir haben (...). Eben wegen (...) dieser Förderung des ethischen Kontinuums, dass wir nicht einfach sagen, sobald wir den reinen menschlichen Bereich verlassen. Das könnte man im Blick auf Tiere auch, ja, sagen. Das sind eben keine Menschen. Mit denen kann man völlig anders umgehen als mit Menschen, oder nehmen sie irgendwelche Naturalen, biologischen und anderen Systeme in der Natur. Man kann mit Wäldern anders umgehen als mit Menschen, oder so. Da ist ja auch viel in Bewegung. Also, da würde ich sagen, macht es durchaus Sinn, Schutz (...), Schutzansprüche mindestens moralisch auch auf Maschinen zu übertragen, so dass die jetzt nicht irgendwie einer Willkür im schlechtesten Sinne des Wortes ausgesetzt sind, von menschlicher Seite her. Ich würde auf der rechtlichen Seite davor warnen, Maschinen auf die gleiche Rechtsebene zu stellen wie (...) Menschen. Also menschenrechtsähnliche Rechte für Maschinen kann ich mir ehrlich gesagt nicht vorstellen (...), weil es letztlich doch so ist, diese Maschinen bauen sich immerhin noch nicht selber, sondern die sind menschengemacht. Das ist beim Menschen anders. Wer zur Welt kommt, wer geboren wird, ist sofort Rechtssubjekt, eigentlich schon vorher. Praktisch als Embryo hat man auch schon Rechtsschutz. Das ist nicht analogisierbar mit Maschinen, die kommen ganz anders zustande. Es ist eine andere Genese, ein anderes zur Welt kommen und darauf muss man dann schon Wert legen. Unser Rechtssystem setzt nicht voraus, dass eine Maschine, die gebaut wird, in dem Moment, wo sie fertig ist, die Rechte hat oder die Rechte haben kann, die ein Mensch hat, wenn er zur Welt kommt oder wenn er in einer embryonalen Zustand ist und dann ein Rechtssubjekt neben seiner Mutter oder neben seinem Vater sein kann. Da würde ich sehr davor warnen, das irgendwie durcheinander zu schmeißen (...), weil ich auch nicht sehe, was letztlich der Vorteil mindestens für die Menschen sein kann. Und ich fürchte sogar, wenn man diese Rechte irgendwie, also wie im Kölner Karneval, wie Kamellen aus dem Fenster schmeißt (...), schürt das eher auch solche Horror Vorstellung, wie sie von Hollywood zum Teil irgendwie auch genährt werden, dass die Roboter uns nachher in die Ecke drängen, dass die gegen uns losgehen, diese ganze (...) Horrorshow, die auch ein bisschen an den Haaren herbeigezogen ist. Aber wenn man das noch füttert, dann darf

man sich nicht wundern, dass es da auch solche verführungstheoretischen Spinnereien am Schluss gibt. Und mit Nick Bostrom haben sie ja jemanden, der da auch selber praktisch auf sich selber ausrutscht, würd ich sagen, weil er sich sowas auch scheinbar vorstellen kann, dass wir nur noch eine kurze Zeit haben, wo wir Herrn im Hause sind und dann kommen die Robots und machen das alles selber, und wir ziehen den kürzeren (...). Das glaube ich nicht. Ich meine, man sollte das nicht fördern, und wir sind immerhin noch diejenigen, die das Rechtssystem in der Hand haben. Das sollten wir auch so halten (...) und ich glaube, es gibt genug vernünftige Leute, die nicht Technik feindlich sind, aber doch auch sehen, man muss mit rechten immer sehr vorsichtig umgehen. Das gilt auch für Menschen genauso. (Until 00:30:21)

Marc Lewald: Okay (...), dann sage ich schon einmal vielen, vielen herzlichen Dank! Es war ein super, super spannendes Gespräch. Ich habe unfassbar viele Dinge, denen ich nachgehen muss (...). Ich sage ihnen auch direkt, ich werde wahrscheinlich die, das hat mich nämlich schon beim Lesen ihres ein bisschen (...), geärgert. Wie nennt man (...), hat mich ein bisschen genervt, dass ich diese Unterscheidung von Recht nämlich nicht so klar gemacht hatte bisher und werde die dann wahrscheinlich auch noch noch in Zukunft tun (...). Ich fand auch der Vergleich, den sie jetzt auch vielleicht mit dem zur Welt kommen, nochmal kurz aufgegriffen haben. Den Vergleich zu einem Kind sehr, sehr spannend (...), den ich bisher gar nicht so wahrgenommen hatte, fand ich aber super interessant. Genauso wie Paro, Koseki, alles für mich sehr, sehr neue Sachen, die so auch (...) in vielen literarischen Werken nicht (...) erwähnt werden. Also da nochmal wirklich hut ab! Sehr, sehr, sehr, sehr tolle Arbeit, hat mich an der einen oder anderen Stelle wirklich fasziniert und wird auch an der einen oder anderen Stelle noch in meine Arbeit einfließen. Vielleicht noch eine allerletzte Frage (...), da sie am Ende noch über was für das Verständnis von man Robert Interaktion gesprochen haben. Arbeiten sie jetzt zufällig an diesem Thema aus rawlscher Sicht, oder kennen sie jemand, der daran arbeitet? (Until 00:31:47)

(27) Expert 2: Meinen sie jetzt speziell Roboterrechte? (Until 00:31:50)

Marc Lewald: Und die Human Robot Interaction, über die sie am Ende gesprochen haben. (Until 00:31:56)

(28) Expert 2: Also, wen ich ihn sehr, sehr empfehlen kann und ans Herz legen kann, ist der Kollege (...) David Gunkel, ein polnischstämmiger US-Amerikaner, der an der Northwest University in Chicago lehrt und der gerade ein Buch fertig gestellt hat. Es kommt in den nächsten Tagen raus (...). Persons Things Robots, also der macht diese dreier Kategorie jetzt sozusagen auf, und da werden sie vieles findet. Er hat auch ein Buch über Robot Rights geschrieben. Das ist übrigens in meinem Paper auch ziemlich



ausführlich zitiert. Also, das ist eigentlich momentan der führende Theoretiker, was diese Differenzierung angeht. Wenn ich ihnen auch ans Herz legen könnte (...), wäre Frank Pasquale, auch ein US-Amerikaner, der jetzt an der University Of Brooklyn lehrt, das ist ein Jurist, also Gunkel ist Philosoph und Medienwissenschaftler, Pasquale ist ein Jurist, sehr kundig, der hat auch Robot Rights oder New Laws of Robotic aufgestellt. Das ist, glaube ich, sein neuestes Buch. Er ist auch Mitherausgeber des Oxford Handbook of AI (...) Ethics und wenn sie post (...) humanistische Dinge nochmal nachlesen wollen. Ich nehme an, sie kennen Rosa Braidotti und ihre Schriften dazu, das kann ich ihnen auch empfehlen, oder die Leute, die um Rosa Braidotti herum, sind in Utrecht. Die sind da ziemlich fit auf diesem Gebiet. Das sind auch Dinge, an denen ich mich orientiere, und ich bin auch mehr international vernetzt, wie sie vielleicht dadurch auch merken, weil in Deutschland das noch nicht so richtig (...) in Gang kommt. Da ist auch meine Hoffnung, dass die Karlschule sie haben, einen sehr fitten Rektor, Herr Lepenies schätze ich auch sehr. Ich glaube, da sind sie an einem wirklichen Hotspot in Deutschland und können froh sagen, dass es (...) bei ihnen so zugeht, wie es da zugeht. (Until 00:33:58)

Marc Lewald: Definitiv. (Until 00:34:00)

(29) Expert 2: Und auch, dass sie so gefördert werden. Es gibt noch viele alte Säcke, wenn ich das so (...), schade, dass das jetzt irgendwie nötig ist, dass man das sagt. Aber es ist ja jetzt sogar auch aufgezeichnet. Aber ich bleibe dabei. Es gibt noch viele alte Säcke, die diese (...) flexible Art und innovative Art des Denkens nur dann fördern (...), wenn sie mit im Schlepptau, mindestens noch auf irgendeinem Beitrag als Mitautor, oder weiß der Teufel was stehen. Und das ist eine Pest und auch eine (...), irgendwie Erschleichung von Präsenz, die man jungen Leuten, die ihre eigenen Ideen jetzt in eigener Verantwortung präsentieren sollen und dürfen, eigentlich nicht zumuten sollte. Und da muss man sehr genau gucken, wo man ist (...), und sich freuen, wenn man wo ist, wo es solche traurigen Dinge eher nicht gibt (...). Und vielleicht, wenn sie später selber mal, ich weiß nicht, was sie vorhaben, ob sie in die Wirtschaft gehen, da verdienen sie auf jeden Fall mehr Geld als in der Wissenschaft. Aber Wissenschaft ist auch nicht schlecht, selber mal irgendwo Verantwortung haben, vielleicht würden Sie ja auch mit den Leuten, von denen Sie glauben, das sind Leute der Zukunft, so behandeln, wie sie selber jetzt gerne behandelt werden möchten, und ich glaube, dann geht so ne gute Schiene weiter (...). Sie kommen ja selbst offenbar aus einem Akademikerhaus, ist es bei mir nicht so. Mein Vater ist Banker (...). Ich hab einen ganz anderen Hintergrund. (Until 00:35:21)

Marc Lewald: Einmal umgekehrt. (Until 00:35:23)

(30) Expert 2: Deshalb ist mir das auch so wichtig, dass man in der Wissenschaft Dinge auch nochmal anders probieren kann, anders weitergeben kann mit anderen Leitsternen, als man das in der puren Wirtschaft hat (...). Womit ich jetzt nichts gegen wichtige Wirtschafts (/). Aber ich glaube, sie haben in etwa verstanden, was ich meine (...), und, ja, das wäre von meiner Seite das, was ich sagen kann. Wie gesagt, freue mich sehr, dass die Karlsruhochschule so weit vorne mit dabei ist, in Deutschland sicher ganz wesentlich (...). Ja, und ich wünsche ihnen für ihre Arbeit viel Glück. Was auch immer daraus wird (...) und wenn sie was schreiben dazu, sie müssen nicht, aber mich interessiert immer, was Leute, wie sie da sehen und was ich vielleicht nicht sehe. Also wenn sie wollen, Ich bin immer auch interessiert an Sachen, wo man was lernen kann. Wenn sie mal was rüberwachsen lassen, freue ich mich. Ich lese es ganz bestimmt, wenn sie wollen, schreibe Ich auch einen kurzen Kommentar dazu, aber Sie müssen nicht, ist es nur eine Frage oder Bitte. (Until 00:36:19)

Marc Lewald: Nein, Gott im Himmel ist doch (...) super, super lieb! Vielen, vielen Dank, ganz im Gegenteil und ich kann ihnen da eigentlich auch nur Recht geben. Allein die Tatsache, dass ich überhaupt über dieses Thema schreiben kann (...). Da danke ich der Karlsruhochschule sehr und weiß es sehr zu schätzen, weil es an anderen Universitäten sicher nicht so einfach gewesen wäre (...). Und auch mit unserem neuen Dekan, dem Herrn Lepenies, was KI angeht, sind wir da (...) anderen Uni sehr, sehr weit voraus, vor allem auch im Umgang damit. Wir arbeiten da mit den Vorlesungen etc. Also, da bin ich schon (...) sehr, sehr stolz, mal wieder an der Karlsruhochschule zu sein, und ja (...), falls ich irgendwas also, ich werde ihnen natürlich sowieso die Masterarbeit zukommen lassen. Selbstverständlich, Sie sind ja Teil davon (...) und falls danach noch was kommen sollte, akademisch, meinen Vater würde es natürlich sehr, sehr freuen, würde ich diesen Weg einschlagen, dann lasse ich Sie natürlich auch gerne wissen (...), selbstverständlich. (Until 00:37:21)

(31) Expert 2: Wunderbar! I'm looking foreward to it, dann alles gute Ihnen. Danke, dass sie sich an mich gewandt haben (...). Ich konnte hoffentlich ein bisschen was von den Fragen erklären. (Until 0:37:32)

Marc Lewald: Definitiv. (Until 00:37:33)

(32) Expert 2: Viel Glück! Machen Sie was wirklich gutes draus, womit Sie selber auch zufrieden sind und was vielen Leuten hilft, Dinge klarer zu sehen (...). Alles Gute und Grüße nach Karlsruhe! (Until 00:37:44)

Marc Lewald: Ich gebe mir Mühe (...) und ich danke ihnen auch vielmals für die Zeit, die vielleicht dann noch ein bisschen kostbarer ist als bei mir als Student, morgens am

Montag. (Until 00:37:54)

(33) Expert 2: Das sind wichtige Sachen. Wenn ich was beitragen kann, mache ich es sehr gerne. (Until 00:37:57)

Marc Lewald: Das freut mich sehr zu hören. Tschüs! (Until 00:38:00)

**Transcript 03: (personal interview)**

Marc Lewald: So let's start right away with the first question (...). Have you ever been in contact to (...), humanoid humanoid robots, and especially with the topic of robot rights? (Until 00:00:14)

(34) Expert 3: No, no, not in direct contact (...), but via scientific discourses. (Until 00:00:23)

Marc Lewald: Okay, perfect (...). Then let's switch directly to the ethical sphere (...). If you would have an humanoid robot (...) and you would describe that humanoid robot with full agency or with agency in general (...). What do you think would that mean in the or change in the understanding of a humanoid robot? (Until 00:00:49)

(35) Expert 3: Yeah (...). That is, of course, for a long time already, a big issue in scientific discourse is to (...), take technologies in terms of agency, right? (...) And what is like, yeah. The issue here is that (...), an agency has (...), it comes with autonomy, right? It's an autonomous agency with capacity for autonomous action and autonomous thinking and so forth (...). And agency originally came with the idea that the agent is an end in itself, right? That's kantian moral imperative (...), as contrasted with the vision that technology is a tool and and therefore a means to be used. A tool is not an agency. That's (...), the tool is in the hands of an agent (...). Yeah (...), for a long time, this is an issue, that comes with (...) a vision of what it means to be human. You have, of course, the Hathaway Cyborg manifesto that addresses that issue. The argument that the human and the technological is always like intertwined. And that is, that is true. I mean (...), historically (...), think about the Prometheus myth, for instance. There is no human without technology. I mean, without technology, the human vanishes, cannot survive. It needs fire, for instance, to survive. Fire is the original technology (...). But, (...) what concern that's your topic, of course, is to what extent like this technology, something like becomes humanized like that is like sort of humanization of technology and hence getting human features (...). And that's where the issue of the agent comes from like to what extent it becomes some sort of a human and hence some sort of an agent. Now (...), my position is like, I'm very skeptical about it, about that, yeah. For me, a robot in my perception remains like a machine (...), a mechanical sort of entity (...) programmed. I remember, for instance, in my university when the robot when I was at the robotics center (...). This robot looked like from a Star Wars (...) movie from the 1950s (...) and it spoke a sort of the language of master, master and slave. So the master slave dialectic continued here. So the robot, that's the original meaning of a robot, as you probably know in the slavic language, as a sort of forced laborer or slave (...). A slave is also not entirely human or is deprived of particular human features. And

also a slave is a means and not an end in itself (...). But for me, the core thing is that the robot is a mechanical entity, it's fabricated (...) and also it's, it's cognitive, it's cognition, is mechanical. And yeah, it's often called artificial intelligence, originally called machine intelligence (...). And from my perspective, that's not, not a real intelligence. I think (...) this word intelligence is wrongly chosen by Alan Turing at the time (...). I know like historically, this was a dubious term, so he hesitated. But that took place in the context in which psychologists were developing those intelligence tests, and that's where the machine intelligence came from. So for him, it was basically a test (...), right (...). But originally intelligence means something else that comes with issues of education, Bildung (...), the culture of the mind, Geist and so forth, yeah. That cannot be found in artificial intelligence. That's all machine stuff (...). That in a nutshell, is that is some sort of an answer or (...)? (Until 00:05:42)

Marc Lewald: Definitely (...), already made a few points that I have to dig into afterwards (...) You already told that you are quite skeptical and have a, (...) yeah, mechanic view on robots. I (...) describe in my work the difference between action and operation. Would you see a difference between an action and an operation? And maybe this is also (...) connected to the mechanical view on robots. (Until 00:06:15)

(36) Expert 3: Yeah (...). Yeah, there is a difference (...), of course. Like the operations are very much internal, right? The action is what comes out of it (...) is visible. And of course we may (...) come to a situation in which the difference between a human action and a robot action can no longer be discerned. That may well be particularly in a digital realm (...). But then again, what does that mean? Does it mean that a robot is acting like a human (...), or is a human acting like a machine (...)? And the latter throughout history has been the case, of course, like humans (...) acting like a machine, like the work environment that requires that also a bureaucracy is a mechanical sort of phenomenon. The bureaucrat is a human who acts like a machine. It doesn't make him a machine, but it makes the action very machine'll for instance, when he simply applies a procedure (...), right? It's simply procedure or rule based action, basically (...). Ultimately an algorithm is not dissimilar from a bureaucratic rule. I mean (...), this is something that others have pointed out. For instance, Eilenberg and Peters in their book *The Algorithmic Society* (...), pointed out that (...) an algorithm is like a similar to a bureaucratic procedure. And (...) that means that also a bureaucrat is a program in its pure form, if you like. Ideal typical form (...) is a programmed sort of worker (...). So, yes, I think there is a difference between the procedure and the action, but also with a robot, both are programmed. That is my point. And the procedure, the operational is is internal. That is what makes the robot operate, basically. And the action is what comes out. But this is still a rule based, technical rule based action (...) performed by a machine. (Until 00:08:48)

Marc Lewald: Very, very interesting, especially the algorithmic society. Definitely have a look at that (...). When we now talking about robots and their ability to do certain things, a lot of people talk about consciousness there and the ability of a robot to have consciousness (...). First things first would be (...), have you a certain type of describing consciousness, especially in terms of machines, if you make a difference there? (Until 00:09:21)

(37) Expert 3: Yeah (...). I think (...) describing a machine in terms of consciousness, I think is very problematic (...). Even the very concept of consciousness is very problematic. I mean, it has metaphysical (...), background, of course, and that metaphysics is embedded with other metaphysical notions, like the soul, for instance, and even the mind (...) Ancient Greek, the nous, as you know (...). And was and these metaphysical notions come from a distinctive religious context (...), that originates in ancient Greece. I mean, anima, for instance, is another example of this (...). But this was always applied to the human (...), strongly connected with the human soul and strongly connected to the notion of the gods or a God or the divine and so forth (...). But when it comes to a robot, this is a material phenomenon (...). Strictly, this is, yeah (...), this is a thing. This is matter. Right (...)? And that in itself does not have consciousness (...). But I understand where this in the discussions, why this pops up. I mean this, but like, this has to do (...) this I think is what I would call a misplaced metaphor. It's a misplaced notion. You see this very often. I've published on this in in the digital transformation everything like digital AI and even the notion of intelligence in the notion of AI (...), from my perspective, is a misplaced metaphor. There's nothing intelligent in AI unless we change the meaning of intelligence (...), right? And likewise, in those discussions of consciousness, I think this is a misplaced metaphor to make sense of something (...) for which we may not have another word, but think consciousness is not the right word here. Why not? I don't think that it's possible for a robot like to develop a consciousness because all that is being done with a robot or with AI as such (...) is to feed that thing with data, data, data. And then what it can do is to make all sorts of associations and analogies, to link data, to link ideas (...). So when you mentioned like one word, that immediately comes up with another word that is related and so forth (...). But that as such would not be consciousness, because this is all strictly mechanical, strictly material. Point of a consciousness is that this is not something that can be realized with this mechanical worldview. (Until 00:12:46)

Marc Lewald: Okay (...). (Until 00:12:47)

(38) Expert 3: But these discussions are very old. Again, you can think about like the ghost in the machine. And now we don't call it ghost, but consciousness. (Until

00:13:01)

Marc Lewald: Yeah. I also had my problems defining that for the master thesis or breaking down a certain concept that I would or can use to work with. And I in the end chose like a psychological (...) definition that I want to work with on humanoid robots (...). Maybe we can come back now to the topic of robot rights in general (...), but stay with the philosophical turn, because my question now would be if we would now have robotic rights in (...) to a certain extent, what do you think would that change from a posthuman perspective on the view of ourselves and also the machines? (Until 00:13:48)

(39) Expert 3: Yeah (...). I think that would change a lot (...). I not necessarily welcome it, although I understand the movement and I do see some positives in that (...). But (...), the notion of rights is a very problematic one. Rights are like legal entities. They're a modern phenomena that emerged historically emerged in the rights revolution in the 19/17th century in England (...), comes with a particular type of state, the sovereign state, the modern state, which is the type of state that can issue rights. I mean, not all states have the power to issue legal rights. I mean, it comes with the ability to turn something into a legal entity (...), that has a legal validity (...). Then the rights revolution started with first with civil rights. Context is the reformation here, all the religious wars. Now it's good to be German because then you (...) always have the trauma of the 30 Years War that almost wiped out Germany at least for 50% (...). And then political rights in the 19th century and social rights with the Sozialstaat or the welfare state and then human rights which came with the United Nations, the UN Declaration of Human Rights after the Shoah. Again, very good to be German here (...). And then we got animal rights (...), we moved from human rights to animal rights (...) and that was mainly to (...) preserve animal welfare and that was for the human (...) already something big, right? The relationship between humans and animals. Are animals simply the tools of humans? And you can do things whatever you like with them? You can torture them? you can inflict pain? (/) and comes with industry, of course, think about industrial breeding of animals and so forth (...). And given the abuse of animals by humans comes with the vision that humans are superior to animals, humans are the rulers of nature, the masters of nature, animals, some sort of the slave or the tools of of human beings. But with animal rights you (...) change? Yeah, the status of humans (...) decreases vis a vis animals. Right? Which was (...) you're no longer simply the owners of of the animals (...). And of course, in the context of animal abuse on an industrial scale, the enormous amount of suffering inflicted on animals by human beings (...), just as with human rights, also emerged because of the enormous scale of suffering on humans. That's the story of the Shoah (...). So those rights are always protections against the abuse of power, of cruelty and so forth (...).

And now to apply them to robots, because that's the point from animal rights to robot rights, but that animals can feel (...), but robots are machines in the end, and not biological entities (...). Why would you do that? And my, I would say and that's my positive thing, let's let's start with this. Why I can see the issue here (...), is because of the potential abuse of technologies (...), potential abuse of robots (...), the abuse of technology. That is a recurrent thing. Günther Anders who has written on that (...). So technology is not just used, but they are particular used, giving you abuse, giving human perversity. Giving human cruelty (...), that I can understand (...). But what I find difficult with robot rights is that (...), if you grant rights to robots, you can grant it to any sort of thing. You can grant it to cars, to houses (...), to stones, to whatever. And that is that's the issue. Although human robot rights comes with a vision and that's precisely a vision I would disagree with. That (...) robots are a very special type of technology because of AI, as if this technology can really think. And we get, of course, also emotional intelligence for robots and the empathy and so forth (...). So this is not just technology, but this is a technology with distinctive human features, if you like, is the assumption (...) that it is an agent, right? So it's an agent that has legal rights, robot rights (...). But as we discussed before already, I would deny this vision. I think this is a misplaced sort of metaphor. I disagree with that. I don't I don't think, for instance (...), I can only understand the robot rights from the standpoint of the legal protection. Okay (...). Just as with animal rights (...), but I cannot understand it from the perspective of an agent that we would get, for instance, a robot that (...) requires a freedom of speech (...), or the freedom of association, that a robot will establish a labor union without being persecuted (...), or that a robot would get voting rights. That's sort of (...) utopian visions, if you like. (/) entirely from my humble perspective, entirely misunderstands the nature of that technology (...). In the end, it's just technology. (Until 00:20:29)

Marc Lewald: Yeah, I think it really depends heavily on the (...), on the opinion you have on machines and (...) robots in general. If this is just technology for you or anything more, because I already talked also to (...) one and two professors who are (...), yeah, in the belief that there is more. But then it also and this is why I ask all these questions, it comes down to what do you see as consciousness? What do you see as an agent? What can this agent do in your world? And that all comes then down to the question of robotic rights in the end (...). But I see, you are a little bit skeptical and I would agree from a personal point of view (...). That's also why did this research. What I found also very interesting and maybe you can (...), tell me a bit about that because you also mentioned the master slave (...), dynamic. And I found that very interesting because, during my research I came across the point that there's also an economic interest in denying robot rights or giving robots rights however you want to see it. And I found that very interesting because I never thought of the (...), economic interest in these, in such things, because for me, this was a completely philosophical (...) term.



(Until 00:21:57)

(40) Expert 3: Yeah (...). You want me to to embark on this topic? (Until 00:22:07)

Marc Lewald: Yeah, yeah, yeah. (Until 00:22:11)

(41) Expert 3: Yeah, yeah, yeah (...). No, I think the (...), when we think about like digital technologies and (/) place robots within the realm of digital, not necessarily of course, but contemporary robots, chat bots and so forth are definitely part of the (...) digital transformation, more and more specifically, the so-called AI transformations taking place today (...). But this, these transitions (...) do not take place in a void or something. There are all sorts of forces behind it, and these are both political (...) and economic forces. So this entire digital transformation is is stimulated by politics from all sides. So we don't, of course, like AI is like old, right, as in 1936 and then as a scientific discipline in 1956, nothing new about it. But the current discourses on AI is new of course, and that has to do with data capitalism (...), informational capitalism from 2000. This is new, like from the 2017. First, the UN came up with AI politics and then European Union. It's politics of AI, its AI strategy, and then all the nation state, almost all of them in the world, 2018, I think 2019 came up with their AI strategy. Germany (...), like AI made in Germany, right, as a quality like Porsche (...). Quality stamp as contrasted with made in China. That's cheap stuff. That's that's rubbish. Right (...)? So that all operates under a political, legal administrative framework in which networks coalitions are being created, so-called the AI-coalitions (...). Sort of Silicon Valley constructs are being made (...), partnership between government, corporations, startups, universities and so forth, and this is multi-billion sort of business, including taxpayers money, but also a lot needs to come from the private sector, so this is public private (...) partnership. In Europe (...), from the awareness that Europe is lagging behind, it has missed the digital revolution, if you like, mainly because of the unification of Germany and the unification of eastern and western Germany, Eastern and Western Europe after the Cold War (...). But the digital transformation that is Silicon Valley, that is California (...), not even the United States. It's simply California and a little bit in the South, in Texas, like Austin, New Mexico, that sort of area (...). And then China, that's the great imitator of the world, has basically stolen everything and copied it with huge, massive manpower (...) and something similar with Russia. So Putin in 2017, as you probably know, summer 2017, he made his famous slogan, like the one (...) who leads AI has the leadership of the world and that was a propagandistic sort of statement, but it awoke many actors. Came with a vision of, oh, now we have to work very quickly. And so also the European Union, the von der Leyen Commission, the politics of AI is one of the five pillars (...), so as to catch up and that requires a massive effort, means that (...), in the past four years or so, everything becomes connected to AI, also

university. So you today ... I exaggerate a little bit perhaps, but you no longer have such a thing, like philosophy becomes a philosophy of AI, ethics become ethics of AI, political science becomes political science and AI, mathematics become mathematics and AI, everything is: and AI. Just look at like the job positions, for instance. And indeed, if you have a little bit of knowledge of AI you, you'll get a job straight away (...). Cyber security becomes an AI issue now, everything becomes reformulated or connected in terms of AI. But my point is that there is a political force (...) behind it. This comes from ministries in Germany. All the Länders have an AI agenda and so forth. Smart cities become AI cities, right (...)? Now, the second one is the economic force behind it, because if we say AI, we say (...) AI companies, that's just AI, I mean, AI as such doesn't exist. AI means AI company. The biggest AI companies are the Silicon Valley companies. So Google is the biggest AI maker in the world (...). Also Apple and Facebook, Amazon, these are the big AI makers, Microsoft, of course, and so forth. And secondly, the chinese ones, Alibaba (...), Tencent, these are all AI makers (...) and these are the biggest companies in the world in terms of financial resources. And, yeah (...). Europe has missed the boat there, but in the past, like ten, 20 years, Europe, the European Parliament, Germany, certainly in the Netherlands, universities, regions ... all have tried to create little Silicon Valleys, all of it (...), China as well. The Shenzhen region, that is the Chinese Silicon Valley. That's where all the tech firms are being created (...). And tech firm today means AI firm (...) in all these different forms, including the robotic sort of. So (...), yeah, political and economic forces, there is a lot of both political power (...), military power, like the major military sort of projects are AI project like for instance. Maybe, you know, but you think about the Maven project in the United States, that's the new Manhattan Project, the Manhattan Project, the making of a nuclear Bomb. And now the making of the superior AI weapon (...), that comes with massive partnerships, public private partnerships, massive investments, secrecy, yeah (...). So, yes, it's pushed for by powers (...) and you can't say no. (Until 00:29:35)

Marc Lewald: There's no way around AI anymore. I would totally agree, no matter what you are doing. (Until 00:29:42)

(42) Expert 3: Yeah (...). Also universities get reconfigured as a result. I mean, it's (...) the classical university disappears (...). In the Netherlands, for instance, imagine like a philosophy curriculum, but also like theology or literature (...). In some universities like Tilburg, for instance, (/) , I think I'm not sure if it's already implemented, but quite sure it will be, that you get mandatory digital skills courses (...) like coding and so forth. So that means that you're a theologian. Imagine a catholic theologian (...) and you get digital skills (/) like coding, working with AI, mandatory. (Until 00:30:34)

Marc Lewald: Nice (...). I don't know how much do you know about our university or the

university, where also Isabelle was (...), but we are heavily deep invested into AI and AI regulations. And this is maybe also the only reason why I'm able to write this master thesis. So (...), yeah. I appreciate that for my university very well that we work with AI and not against AI. (Until 00:31:02)

(43) Expert 3: Yes, it's inescapable (...). Also when it comes to your job prospects, also something that always needs to be taken into account for students. I mean, you cannot educate students for yesterday's world (...), you you can. You'll have to. (Until 00:31:25)

Marc Lewald: But then maybe something completely off topic. I had a lot of friends who studied in the Netherlands and also in other countries and I have to say, for my personal experience (...), Netherlands is in the top countries to study and what I heard from and also when it comes to the general approach to education and the topics. (Until 00:31:48)

(44) Expert 3: Yeah, that's good to hear. Thank you. Yeah. (Until 00:31:52)

Marc Lewald: And as you took most of my questions right away when I talked about a certain topic, I think we are already more or less finished. (Until 00:32:03)

(45) Expert 3: All right. You don't have additional other questions (...), at this stage? (Until 00:32:10)

Marc Lewald: I don't. (Until 00:32:10)

(46) Expert 3: Everything covered? (Until )

Marc Lewald: At least nothing that I can use for my master thesis and I don't have to transcript (...), yes. I thank you very, very much for your time, for your patience and everything else was a really wonderful talk. Have a lot of things that I have (...) to look into and I can already tell that I will somehow (...) integrate the master slave theorem again. I'm missing that somehow. Yeah. (Until 00:32:46)

(47) Expert 3: Yeah, We'll be interesting. (Until 00:32:48)

Marc Lewald: I will send you the master thesis, of course, as I'm (...) finished and yeah (...). If there are any other wishes from your side, any other questions or anything else I'm happy to help, to answer to, to send over, yeah. (Until 00:33:05)

(48) Expert 3: All right. No. Everything clear from my side. So I'm looking forward to your master thesis and you enjoy completing it. (Until 00:33:13)

Marc Lewald: Thank you very much. I try my best (...) The last few weeks, that's normal for a master thesis. You want to get over with it and that's it. (Until 00:33:26)

(49) Expert 3: Hey, good luck. (Until 00:33:31)

Marc Lewald: Thank you very much. And have a wonderful day. (Until 00:33:35)

(50) Expert 3: You too. Hey, Bye. Bye. (Until 00:33:37)

Marc Lewald: Bye bye. (Until 00:33:39)

**Transcript 04: (personal interview)**

Marc Lewald: Dann wollen wir eigentlich direkt starten, ich will gar nicht Ihre kostbare Zeit (...) verschwenden. Haben Sie denn in der Vergangenheit schon mal Kontakt zu humanoiden Robotern gehabt, beziehungsweise sich allgemein mit dem Thema der Roboterrechte beschäftigt? (Until 00:00:15)

(51) Expert 4: (...), ja, also, ich arbeite zu Technikethik unter anderem und ja (...), mit Robotern nicht so viel Kontakt gehabt. Wir haben am KIT (...), ja doch ein paar Beispiele und (...), da war ich mal mit dabei bei einer Gelegenheit, aber wirklich interagiert mit sonderlich vielen Robotern, außer das, was man so (...) im Alltag halt an Robotern oder KI mitbekommt (...). Jetzt sind sie gerade wieder weg. (Until 00:00:46)

Marc Lewald: Nee, ich höre alles, theoretisch (...), ich bin da. (Until 00:00:49)

(52) Expert 4: Ich warte mal, bis sich wieder bewegen (...), jetzt habe ich Sie wieder. (Until 00:01:01)

Marc Lewald: Sie können aber ruhig weiterreden (...), also ich höre alles. Ich bin dann nur kurz gefreezt (/). (Until 00:01:11)

(53) Expert 4: (...) (/), die in der Medizin eingesetzt wird, die habe ich tatsächlich (...) (/) mit dem Roboter hatte, sondern man macht sich Gedanken, was sind denn die Kriterien (...), die gelten sollten, damit irgendein Wesen Rechte hat (...)? Dann kann sich darüber streiten, ob die bei bestimmten Wesen erfüllt sind, diese Kriterien oder nicht. (Until 00:02:20)

Marc Lewald: Ja, damit startet meine Arbeit auch eigentlich, mit dem Vergleich zu Tierrechten und einem kleinen Exkurs zu Menschenrechten (...). Wollen wir dann eigentlich gleich zu den philosophischen Fragen kommen ... Im englischen gibt es ja oft mal den Begriff der Agency (...), den ich auch unter die Lupe nehmen werde (...). Im deutschen: Handlungsvermögen, ist ein bisschen schwer zu übersetzen (...). Wie würden Sie denn das Thema der Agency beziehungsweise das Handlungsvermögen aus der Sicht eines humanoiden Roboters bzw. im Hinblick eines humanoiden Roboters bewerten? (Until 00:03:10)

(54) Expert 4: Vielleicht klappt es dann so besser, oder wir schalten einfach das Video aus. Vielleicht hilft es. (Until 00:03:27)

Marc Lewald: Ich hab mich jetzt nämlich auch einfach auf dem Handy eingeschaltet.

Ich hoffe, Sie verzeihen mir das (...) und ich entschuldige die ganzen Umstände (...), ist es echt immer wieder schön. (Until 00:03:49)

(55) Expert 4: So, wir können es noch mal probieren. (Until 00:03:54)

Marc Lewald: Okay (...), das Thema der Agentur und das Handlungsvermögens, wie würden Sie dieses denn bewerten (...), jetzt im Bezug auf humanoide Roboter? (Until 00:04:17)

(56) Expert 4: Also es hängt sehr stark davon ab, wie (...) anspruchsvoll man den Begriff jeweils definiert. Und es gibt im Bereich (...), bei Roboter ja auch die Idee, dass Roboter handeln, aber typischerweise eben mit einem deutlich reduzierten oder weniger anspruchsvollen Handlungsbegriff als bei Menschen, (/) von menschlichen Handlungen. Sie haben sich die Literatur ja vermutlich angeschaut. Dann haben wir immer diese Idee, das eine Absicht der Intentionalität dahintersteckt, dass wir also auf mentaler Ebene (...) ein intentionales Objekt, Absichten, Zwecke (...) haben und uns darauf beziehen können und dann eben mit der Handlung einen bestimmten Zweck verfolgen. Und diese Intentionalität, würde ich meinen, ist bei Robotern nicht gegeben. Natürlich können sie zielorientiert handeln, gemäß ihrer Programmierung. Wem wir daran denken (...). (/) Roboter-Fußball (...), dann haben die auch das Ziel, ein Tor zu schießen (...), aber das ist etwas das wir typischer Weise zuschreiben und hier kann man nicht sagen, dass auf mentaler Ebene eine Intentionalität bei den Robotern vorliegt, sondern dass die halt entsprechend programmiert sind, dass ihr Verhalten so gesteuert ist, dass wir von außen (...) schnell zu der Idee kommen, da liegt Intentionalität vor, tatsächlich liegt sie bei nicht vor. Genauso wenn man mit ChatGPT sich unterhält beispielsweise. ChatGPT schreibt ja auch, ich bin nur eine (...), ein large language model, und allein diese Formulierung ich (...), suggeriert ja auch, dass da eine Person ist, was natürlich so nicht der Fall ist (...). Das heißt, man sollte vorsichtig sein mit solchen Zuschreibungen, deswegen würde ich den wesentlichen Unterschied bei Akteurschaft oder Agency in dieser Intentionalität sehen, die bei Robotern eben nicht vorliegt, bei Menschen aber schon, jedenfalls würden wir uns das gerne zuschreiben. (Until 00:06:24)

(57) Marc Lewald: Ja gut, ich muss ehrlich sagen, ich habe eher Angst vor dem ganzen, aber ich weiß genau, was sie meinen (...). Ich gehe wegen dieser Intention auch in dem nächsten Schritt auf die Handlung selbst ein und unterscheide in meiner Arbeit dann zwischen Aktion, bzw., es ist immer ein bisschen einfacher mit den englischen Begriffen: action and operation. Dann ja so eine Aktion (...), die hat immer eine Intention dahinter, die (...) hat immer einen Plan, und eine Operation ist dann etwas, was mehr im Hintergrund abläuft (...). Würden Sie denn einen Unterschied

zwischen einer Action und einer Operation sehen, und wenn ja, wo beziehungsweise würde der genau in dieser Intention liegen? (Until 00:07:09)

(58) Expert 4: Ja, ich würde da einen Unterschied sehen. (Until 00:07:15)

Marc Lewald: Es tut mir leid, selbst mein Handy scheint heute einfach nicht zu funktionieren (...). Jetzt sind wir da (/). Wir waren beim Thema Aktion/Operation. Ich weiß nicht, was sie schon dazu gesagt haben. Sie dürfen gerne einfach noch mal kurz dazu ausholen. (Until 00:07:47)

(59) Expert 4: Ja (...), tatsächlich gibt es auch im Deutschen eine Unterscheidung, die dann verwendet wird zwischen Handlung und Verhalten, (...) auch im englischen zwischen Action und operation. Das kann man in dem Kontext auch verwenden, und das, glaube ich, entspricht auch ihrer Idee zwischen Aktion und Operation. Das würde genau das selbe (/). (Until 00:08:08)

Marc Lewald: Okay (...), das ist ein ziemlich spannender Punkt, den werde ich mir wahrscheinlich noch mal genauer anschauen müssen, sehr, sehr gut (...). Ich breche dann in meiner Arbeit dieses Handlungsvermögen und die verschiedenen anderen philosophischen Punkte (...) sehr auf das Bewusstsein runter, da viele (...), zumindest bei den Roboterrechten (...), wenn es darum geht, diese Rechte zuzusprechen, ein Bewusstsein als prerequisite formulieren (...). Ich weiß, es ist ein sehr, sehr schwieriger Begriff in der Philosophie, um es mal einfach auszudrücken (...), aber könnten Sie denn für sich sagen, wie sie Bewusstsein definieren würden, beziehungsweise ob es für Sie einen Unterschied zwischen einem maschinellen oder anderen Bewusstsein gibt? (Until 00:08:59)

(60) Expert 4: Also, da wäre ich ja tatsächlich eher insofern Funktionalist, als ich glaube, dass (...) auch so etwas wie Bewusstsein durch unterschiedliche (...), sagen wir mal Materialien in der Welt instanziiert oder realisiert werden kann (/). Also ich halte es nicht für unmöglich, dass es sowas wie ein künstliches Bewusstsein mal geben kann (...), genauso wie es ein biologisch kreierte Bewusstsein bei uns Menschen, wie wir das irgendwann erkennen könnten. Außer (...), dass uns das jeweilige Wesen das mitteilen würde. Aber das ist (...) Sie haben es ja selber schon gesagt, ein grundlegendes Problem in der Philosophie des Geistes, das sogenannte Zombiargument (/), nie sicher sagen können bei einem Menschen der uns gegenübersteht (...), tatsächlich ein Bewusstsein hat, sondern (...) wir schließen darauf, weil es bei uns ähnlich ist. Wir empfinden etwas (...), denken von uns selbst, wir haben Bewusstsein und da steht mir ein Wesen gegenüber, dass auch ein Mensch ist, nach allem was ich weiß, und deshalb schließe ich darauf, dass auch dieses

Wesen, also dieser Mensch, ein Bewusstsein hat (...). Bei Robotern wird es dann nochmal schwieriger, das zu erkennen (...), oder davon auszugehen. Ich würde es prinzipiell nicht für unmöglich halten, aber es ist ein epistemisches Problem, das jeweils zu erkennen (...). Also wie man Bewusstsein definiert, da bin ich wirklich vorsichtig für den Kontext. Für Roboter, glaube ich, ist entscheidend (...), nicht so sehr Bewusstsein als ganzes (...), sondern die Idee von Selbstbewusstsein. Wenn wir sagen würden, Roboter sind plötzlich ähnlich wie wir Personen, dann kommt eine Frage (...) des Selbstbewusstseins dabei, dass man sich seiner selbst im Vergleich oder Auseinandersetzung der Umwelt bewusst ist (...). Dass man sich als (...) ein bestimmtes Wesen mit einer bestimmten Perspektive auf die Welt wahrnimmt (...), dass man Intentionalität besitzt, dass man sich also auf Dinge in der Welt beziehen kann, diese Intentionalität als Mentale gerichtetheit, was dann noch ein wesentliche Punkt ist, ist die Fähigkeit zur Autonomie, das heißt zur Selbstbestimmung, würde ich meinen. Also die Frage, auch (...) die eigenen Wünsche nochmal kritisch zu reflektieren, sind, wie gesagt, sich praktische Fragen zu stellen. Soll ich das tun? Will ich das wirklich tun, im Sinne von habe ich gute Gründe dafür oder nicht (...)? Das sind alles die Aspekte, die für personelles Bewusstsein, also für Personensein zentral sind. Die werden auch von Tieren (...), jedenfalls von den meisten Tieren, nicht vollständig erfüllt, obwohl gerade Bewussten in einem etwas basaleren Sinne glaube ich auch bei vielen Tieren und in höherem Maße natürlich auch bei den Menschenaffen vorliegt (...), sodass wir mehr und mehr die Kriterien für Personen sein (...), erfüllt sehen. Insofern bin ich mir nicht sicher, ob das, der Begriff des Bewusstseins, für die Thematik, ihre Arbeit (...) als sehr, sehr allgemeiner Begriff tatsächlich entscheidend ist, oder ob es nicht eher um Personensein oder Personhood im englischen geht (...). Und selbst dann könnte man argumentieren, wenn wir Tieren Rechte zusprechen, ist personhood auch nicht erfüllt. Das heißt wir haben natürlich die Möglichkeit, beliebigen Entitäten Recht zuzusprechen, und selbst solchen künstlichen Entitäten wie Unternehmen (...). Aber die Frage ist natürlich, welchen Entitäten sollten wir welche Rechte zusprechen und aus welchen Gründen? (Until 00:13:05)

Marc Lewald: Ja (...), da bin ich voll und ganz bei Ihnen. Einer der größten Kritikpunkte (...), wenn man so will bei der ganzen Frage um die Roboterrechte (...), ist nämlich auch die Frage, wenn denn also da ist man dann schon mal sehr bei den (...) Limitations (...), wenn man denn über Roboterrechte spricht, ob man da nicht im gleichen Moment oder im gleichen Atemzug die Rechte anderer Menschen, anderer Lebewesen, oder eben anderer Entitäten downgraded (...), dadurch, dass man dann im Prinzip allem ein Recht geben kann (...). Der Begriff des Bewusstseins wird bei mir wahrscheinlich doch, da haben Sie absolut Recht, mehr durch die Personhood definiert, zumal es ja auch schon ein, zwei Roboter gab, denen diese Personhood zugesprochen wurde (...) zum Beispiel in Saudiarabien und Japan. Ja, also, da werde



ich mich schon (...). (Until 00:14:07)

(61) Expert 4: Da wäre ich aber vorsichtig (...), was man den zugesprochen hat, sind Bürgerrechte (...). Und das ist nochmal was anderes, als zu sagen, sie erfüllen die Kriterien des Personen-Seins. Also, man kann einerseits sagen, man spricht in irgendwelchen Entitäten bestimmte Rechte zu (...), und kann sich fragen, welche Kriterien dafür nötig sind, oder man sagt (...), die Kriterien, des Personen-Seins, sind erfüllt. Das eine ist eine Praxis des Zuweisens oder Zusprechens (...), das andere sind Behauptungen darüber, ob bestimmte ontologische Kriterien erfüllt sind oder nicht. (Until 00:14:44)

Marc Lewald: Sie entschuldigen, ich mache mir gerade noch kurz ein paar Notizen. (Until 00:14:50)

(62) Expert 4: Jaja, alles gut (...). Ich höre den Stift im Hintergrund. (Until 00:14:54)

Marc Lewald: Okay (...). Denken Sie denn, wenn man von einem, naja personellen Bewusstsein sprechen kann bei humanoiden Robotern, dass diese denn in der Zukunft überhaupt in der Lage sein werden, so etwas zu entwickeln, oder eben dann nur das, wovon Sie bereits gesprochen haben (...), dass wir annehmen, sie würden eins hätten, da sie halt (...) Handlungen durchführen können, die wir Ihnen vorher nicht zusprechen hätten können? (Until 00:15:23)

(63) Expert 4: Also, da halte ich es quasi mit Karl Valentin der sagt (...), Voraussagen sind schwer, insbesondere wenn sie die Zukunft betreffen (...). Ich bin mir wirklich nicht sicher (...), ob ich das für möglich halten soll und nicht im Prinzip, ich kann es, ich habe keine guten Gründe das auszuschließen würde ich meinen, insbesondere wenn man an (...) eine sehr weit entfernte Zukunft denkt (...). Im Moment, glaube ich, sind wir da noch ein ganzes Stück davon entfernt (...). Aber unabhängig davon, ob es dann wirklich mal kommen mag oder nicht, wenn es erfüllt ist, dann haben wir es eben mit künstlichen, vollwertigen Personen zu tun und natürlich sollten diese Personen dann dieselben Rechte haben wie natürliche Personen auch (...). Wenn es nicht erfüllt ist, wenn das niemals kommt, dann kann man sich immer noch fragen, wie wir mit Robotern umgehen sollten, selbst wenn sie die Kriterien von Personalität nicht erfüllen (...). Das kann man, wie Sie am Anfang ja auch gesagt haben, einigermaßen in der Analogie zu Tieren sehen. Selbst wenn die Katze jetzt niemals die Kriterien des Personen-Seins erfüllt, heißt es noch lange nicht, dass ich alles mit ihr machen darf. Es gibt gute Gründe zu sagen, dass (...), wir moralische Pflichten gegenüber Tieren haben und auch (...) je nachdem, welche Position man vertritt dann, gute Gründe dafür, dass Tiere moralische Rechte haben sollten (...). Das ist alles auf der ontologischen Ebene,

was Personen-Sein angeht, also ob die Kriterien tatsächlich erfüllt sind. Jetzt kommt noch eine weitere Schwierigkeit hinzu, die ich vorhin schon kurz genannt habe, die epistemische Schwierigkeit, können wir das dann irgendwie, tatsächlich, mehr oder weniger sicher erkennen (...)? Das heißt selbst wenn wir, wenn die ontologischen Kriterien mal erfüllt sind (...), ist die Frage, erkennen wir das wirklich (...)? Oder (...), wenn wir sagen, die ontologischen Kriterien sind tatsächlich nicht erfüllt (...), aber die Roboter, von den wir dann reden, sind so komplex (...) und so gut darin, menschliches Verhalten zu simulieren, dass wir meinen könnten, sie hätten Personalität (...). Und das ist ein Argument, finde ich, dass man (...) in Anschlag bringen kann, um Roboterrechte zu verteidigen, unabhängig davon, ob die ontologischen Kriterien tatsächlich erfüllt sind oder nicht. Also so eine Analogie zum juristischen in dubio pro reo Argument. Wenn wir nicht sicher sind, dann im Zweifel für den Angeklagten und, lieber Roboterrechte zusprechen (...), also, man irrt dann wenigstens auf der sicheren Seite, als tatsächlich dann künstliche Personen unmoralisch zu behandeln. (Until 00:18:23)

Marc Lewald: Ja (...), es geht ja dann auch etwas (...), für mein Verständnis, in diesen relational turn, bei dem es dann ja auch mehr darum geht, was die Roboter oder humanoiden Roboter bei uns auslösen, als das (...), was sie wirklich für Handlungsvermögen besitzen, Bewusstsein oder (...), was auch immer. (Until 00:18:48)

(64) Expert 4: Ja (...), da muss man halt bisschen (...), da muss man unterscheiden, was genau mit diesem Relational Turn gemeint ist. Werden da, wird da wirklich wirklich die These vertreten, dass Personalität aufgrund von Relationalität konstituiert wird (...), oder geht es nur darum zu sagen, unabhängig, was da konstituiert wird (...), die Relationalität und die Praxis, diese relationale Praxis, führt dazu, dass wir (...) Rechte zuweisen, unabhängig davon, ob auf der individuellen Roboterseite bestimmten ontologische Kriterien nun erfüllt sind oder nicht (...). Aber das hat eher was damit zu tun, wie genau man diese relationalen Ansätze ausbuchstabiert und welche Variante (/). (Until 00:19:39)

Marc Lewald: Aber aber sehr, sehr spannende und wichtige (...) Perspektive, die ich wahrscheinlich dann auch nochmal ein bisschen klarer definieren sollte, um ehrlich zu sein (...). Wenn wir jetzt ein bisschen weitergehen und das ganze (...), naja aus mal post-humanistischer Perspektive betrachten (...), was ich auch in meiner Arbeit machen werde, ebenso wie aus transhumanistischer Perspektive (...). Was würde denn das Zusprechen von Roboterrechten aus posthumanistischer Sicht für naja (...), das Verständnis von uns und des Menschen bedeuten, oder von Menschlichkeit? (Until 00:20:20)

(65) Expert 4: Also (...), ich komme ja selber eher aus der analytischen Philosophie-Ecke und weniger aus der, (/), kontinentalen Ecke oder posthumanistischen. Aber ich würde zunächst einmal meinen (...), der Sache nach würde das für uns gar nicht bedeuten (...). Wenn wir, und unsere Kriterien von Personalität sind davon ja erst mal nicht betroffen, (...) durch den Umstand, dass wir anderen Wesen, in dem Fall jetzt Robotern Rechte zusprechen, ich meine, werden unsere Rechte nicht eingeschränkt. Sondern wir haben dann in der Praxis dann aber natürlich (...), wie jetzt auch (...), einfach den Umstand, dass Rechte gegen Rechte abgewogen werden müssten, zweifelsfrei. Also wenn beispielsweise Roboter jetzt auch Eigentumsrechte haben können (...), dann haben wir diese Eigentumsrechte halt genauso zu beachten, wie die Eigentumsrechte von natürlichen Personen (...). Insofern würde ich meinen, dass da kein (...) kein grundlegender Unterschied hergestellt ist, sondern das einfach noch ein paar zusätzliche Wesen mit Rechten dazukommen, wie ja doch jedes Jahr immer wieder neue Menschen mit neuen Rechten geboren werden,. Dann sind es halt ein paar künstliche Wesen die dazukommen. Da sehe ich jetzt keinen Game-Changer dabei. (Until 00:21:45)

Marc Lewald: Ja (...), da würde ich ihnen wahrscheinlich auch Recht geben. Wobei ich auch ehrlich sagen muss, dass viele, mit denen ich bisher gesprochen habe, da widersprechen würden (...). Aber finde ich eine sehr, sehr spannende (...), interessante Perspektive, weil ich die eigentlich auch teilen würde, um ehrlich zu sein (...). Wenn es jetzt direkt über die Einführung von Roboterrechten geht, gibt es Gründe (...), direkte Gründe, die Sie nennen könnten, die für Sie, abgesehen von denen, die Sie bereits auch genannt haben, für oder gegen die Einführung von Roboterrechten sprechen würde. (Until 00:22:24)

(66) Expert 4: Also für die Einführung von Roboterrechten würde, glaube ich (...), vor allem dieses In dubio pro reo Argument sprechen, was ich schon genannt habe (...), dass, wenn wir es nicht sicher sagen können, dass keine Personalität vorliegt (...), aber es so aussieht, als ob dann im Zweifelsfall auf der sicheren Seite und sagen, lieber Rechte zusprechen als (...) den Fall haben, dass tatsächlich künstliche Personen dann misshandelt werden (...). Was dagegen sprechen könnte (...), das wären eher die Argumente, dass wir sagen, solange es klar ist, dass es sich um Dinge handelt, also um Nicht-Personen (...), sollten wir den Kreis von Tieren die Rechte haben nicht unnötig oder unbegründet ausweiten (...). Also, wir geben jetzt auch nicht Toastern Rechte, weil wir wissen, ein Toaster ist einfach nur ein Ding (...). Während wir bei beispielsweise Tieren ganz gut wissen, dass da wenigstens ein Bewusstsein vorhanden ist und wie Bentham damals schon gemeint hat, worauf es ankommt, ist, ob sie leiden können oder nicht. Wir wissen, dass Tiere leiden können (...), und deswegen sollten wir, oder haben wir ihnen gegenüber moralische Pflichten (...). Selbst Kant hat

argumentiert (...), dass zwar die Tiere selbst keine moralischen Rechte haben, wir also keine Pflichten gegenüber den Tieren aber in Ansehung der Tiere und es dann ein Argument ist mit Blick auf unsere eigene Moralität, also unseren eigenen moralischen Charakter, wie wir in dem Fall Tiere behandeln. Ob wir grausam gegenüber Tiere sein dürfen oder nicht, das schlägt auf unseren Charakter zurück, so wie Kant argumentiert. Und das könnte man für die Einführung von Roboterrechten ebenso geltend machen, selbst wenn es der Fall ist, dass wir gegenüber den Robotern, also in diesem kantischen Sinne, keine moralischen Pflichten haben, dann vielleicht doch in Ansehung von Robotern, gerade bei menschenähnlichen Robotern (...) Weil es auf unseren eigenen Charakter zurückschlägt und gegebenenfalls darauf, wie wir dann andere Menschen behandeln (...). Also, da gibt's im Zuge der Debatte um Liebe und Sex mit Robotern das Argument (...), dass wenn wir Roboter auf eine Weise haben, dass sie einfach nur unseren Wünschen (...), also unsere Wünsche, egal welche das sind und zur Verfügung stellen müssen und die erfüllen müssen und wir uns quasi ausleben können (...), dass dann auch unser Charakter, insbesondere eben wenn sie menschenähnlich dargestellt sind, das ist die feministische Kritik an den doch zu meist weiblichen Sexrobotern oder als weiblich designte Sexroboter, dass das dann sich negativ auswirkt auf die Art, tatsächlich, also wie wir echte natürliche Frauen behandeln. Und da (...), das Argument hat was empirisches, in die man sich überlegen muss (...), ist diese Konsequenz, von der die Rede ist (...), was in eine empirische These ist, die besagt, dass sich unser Verhalten dann tatsächlich ändern wird gegenüber Frauen, ist das erfüllt oder nicht (...)? Aber man könnte schon sagen, dass es auf unseren eigenen moralischen Charakter niederschlägt, wie wir uns verhalten bei Wesen, die mehr oder weniger ähnlich sind (...), also jedenfalls vom Aussehen her und wir dann (...), erwarten können, dass zumindest die Gefahr besteht (...), dass wir uns gegenüber denjenigen, die tatsächlich Rechte haben, schlechter verhalten, weil wir uns daran gewöhnen, an ein Verhalten, dass die Rechte von Wesen, die tatsächlich keine Rechte haben, also Roboter in dem Fall, so sehr gewöhnen, dass wir (...) unser Default-Verhalten anpassen. Und auch das, glaube ich, wäre ein (...) Argument für die Einführung von Roboterrechten oder zumindest (...), für Pflichten in Ansehung von Robotern, also à la Kant gesprochen. (Until 00:27:11)

Marc Lewald: Vor allem als sie jetzt Bentham erwähnt haben, fand ich das ganz spannend, weil die (...), diese (...), Fähigkeit Schmerzen zu spüren, die ist ja doch schon sehr, sehr wichtig in dieser ganzen Debatte um Rechte und Bewusstsein (...), und ich gehe in meiner Arbeit auch stark darauf ein und auch mit einem Vergleich zu Fisch Rechten (...), da das in Deutschland ja nicht ganz so einfach war, denen zuzusprechen, dass die Schmerzen verspüren, und das mit der Gewohnheit ist auch ein sehr interessanter Take (...). Wobei ich jetzt die Frage hätte, wenn wir in Ansehung der Tiere dann diese Rechte machen würden, würden diese sich unterscheiden zu den

Rechten, die wir unter den normalen Umständen machen würden (...), also werden sie in irgendeiner Form abgeschwächt oder werden das dann schon ungefähr die gleichen Rechte im Verständnis? (Until 00:28:16)

(67) Expert 4: Also ein Missverständnis muss man ausräumen. Hier werden keine Rechte zugesprochen, sondern was Kant macht in dem Argument ist, er spricht über moralische Pflichten, die wir haben (...). Und nur Pflichten gegenüber Wesen, die wir gegenüber Wesen haben (...), entsprechend dann auf Seiten dieser Wesen Rechte. Und wenn wir sagen, wir haben Pflichten in Ansehung, so wie es Kant macht, von Tieren, dann sagt das, dass die Tiere eben keine Rechte haben. Wir schulden ihnen nichts, sondern der ganze Witz ist im Grunde bei Kant eine Pflicht gegenüber sich selbst, eine moralische Pflicht gegenüber sich selbst, dem eigenen moralischen Charakter. Und deswegen haben wir die eigene moralische Pflicht in Ansehung von Tieren, Tiere nicht grausam zu behandeln, nicht weil die Tiere Rechte haben (...), so Kant, sondern weil es auf unseren moralischen Charakter sich negativ auswirkt. Deshalb ist es schwierig Kant, für Tierrechte, glaube ich, in Anschlag zu nehmen (...). Bei Bentham sieht es anders aus. Der würde sagen, das entscheidende Kriterium ist, kann das Wesen Schmerz empfinden (...)? Und das ist direkt der Grund weshalb man dann in diesem utilitaristischen benthamschen Sinne davon reden könnte, deswegen gibt es den guten Grund hier Rechte zuzusprechen (...). Mit Kant ist es ein bisschen schwieriger, deswegen glaube ich ist Kant da nicht so geschickt, um für Tierrechte einzutreten (Until 00:29:47)

Marc Lewald: Ja, also mir ging es jetzt auch vor allem in erster Linie um Bentham, weil ich ihn selbst in meiner Arbeit erwähne (...), bei der Fähigkeit Schmerzen zu spüren. Aber wollen wir noch mal zur Frage der Roboterrechte speziell zurückkommen. Sie haben gerade erklärt, was für oder dagegen sprechen würde (...). Sehen Sie denn auch irgendwelche Vorteile oder Gefahren, wenn man denn Roboterrechte zusprechen würde, oder es eben auch lassen würde? (Until 00:30:19)

(68) Expert 4: Also (...), eine Gefahr wäre glaube ich dann gegeben, wenn man Roboterrechte zusprechen würde (...), obwohl man weiß, dass (...) die Kriterien für Rechte nicht erfüllt sind. Das heißt, man, die Gefahr, dass der Rechte-Begriff zu sehr ausgehöhlt wird, indem man allen möglichen Dingen tatsächlich Rechte zuspricht (...). Auch das ist eine Frage, welche Gründe jeweils dafür sprechen oder welche Gründe man vorbringt, um Rechte zuzusprechen (...). Wenn man jetzt beispielsweise einem Unternehmen bestimmte Rechte zuspricht, das Unternehmen darf dieses oder jenes verkaufen und hat, oder das Unternehmen hat Eigentumsrechte an bestimmten (...), an dem Grund, wo die Firmenzentrale draufsteht. Dann besteht der Grund darin, dass man auf die Art und Weise (...) die Praxis des Unternehmenshandelns besser

realisieren kann (...). Es ist besser, dass wir dem Unternehmen als einer juristischen Personen auch bestimmte Rechte zusprechen und nicht (...) immer nur bestimmte Individuen haben, die das Unternehmen repräsentieren. Das funktioniert mit Familienunternehmen oder Einzelunternehmen, aber mit börsennotierten Unternehmen funktioniert es so nicht. Wir haben gute pragmatische Gründe (...), um die Praxis (...), die ökonomische Praxis besser in die Realität umzusetzen, sowas wie eine juristische Person einzuführen und diesen juristischen Personen bestimmte Rechte zuzusprechen (...). So und jetzt ist die Frage, gibt es solche Gründe mit Blick auf Roboter auch? Sind die überzeugend (...)? Und die Vorteile wären, wenn die Roboter hinreichend komplex sind, dass wir eher geneigt sind, Personalität zuzusprechen, oder dass wir eher einfach aufgrund unserer psychologischen Ausstattung geneigt sind (...), die Roboter, die Individuen, irgendwie Personen zu behandeln und zu interagieren (...). Dann haben dafür auch pragmatische Gründe, um zu sagen, ja dann sollten sie halt auch Rechte haben, weil wir dann die Art und Weise, wie wir mit Ihnen interagieren, was wir alles übertragen, von unserer menschlichen Interaktion auf die Roboter-Interaktion auch auf diese Weise mit übertragen (...). Man müsste vielleicht dann auch auf abgestufte Rechte, also welche Rechte sprechen wir denn genau zu? Eigentumsrechte, oder sprechen wir Recht auf Privatsphäre zu (...)? Da muss man, glaube ich, im Detail noch mal gucken, und (...) die Gefahr könnte da auch sein, dass man bestimmte Kontexte (...) zu sehr ausweitet (...), um zu sagen, das, was die Rechte eigentlich auszeichnen sollen, das, was sie zum Beispiel schützen sollen, wird letztlich dadurch verwässert, dass sie Wesen zugesprochen werden, die, das worum es in den Rechten geht, selbst gar nicht wahrnehmen oder schätzen können (...). Also gerade bei Robotern, und wo es im Individuen geht, wo wir eigentlich gerne oder schnell diese Idee von Personalität zusprechen, begründet oder nicht (...), da glaube ich, wenn klar ist, es handelt sich nicht um die Person, aber ich muss jetzt die Privatsphäre dieser, eben Nicht-Person respektieren. Dann fragt man sich schon, okay (...), warum und was soll das bringen, eine Privatsphäre von einem Wesen zu schützen?, das keine Definition, kein Interesse an einer Privatsphäre haben kann. (Until 00:34:34)

Marc Lewald: Sie haben gerade auch von den verschiedenen Rechten gesprochen, die jetzt einem Roboter dann zugestanden werden könnten (...). Gibt es denn davon welche, von denen sie sagen würden, die sollten (...), also, sollten einfach keinem Roboter zugesprochen werden (...), beispielsweise jetzt, wenn es um property rights geht etc. (...), oder würden Sie sagen, wenn man denn davon ausgeht, dass Sie (...) ein gewisses Level erreicht haben, dann auch die gleichen Rechte wie Menschen haben sollten? (Until 00:35:16)

(69) Expert 4: Also wenn (...), das ontologische Kriterium von Personalität erfüllt ist, also von Personhood, wir es tatsächlich mit vollwertigen künstlichen Personen zu tun

haben, dann sollten sie natürlich auch alle Rechte von Personen haben (...). Die Krux ist natürlich, das herauszufinden (...). Und dann könnte man auch hier wieder argumentiere, lieber auf der sicheren Seite irren und ihnen mehr Rechte zusprechen, als sie vielleicht verdienen. Aber wir können es nicht wissen. Schwierig wird es glaube ich (...), solange wir wissen (...), Kriterien von Personalität sind nicht erfüllt, dann, glaube ich, hat man gute Gründe, bestimmte Rechte wie passives oder aktives Wahlrecht nicht zuzusprechen (...). Warum?, weil (/) Nicht-Person an Roboter kein Interessen (/). Wir können zwar immer noch sagen von außen, dass etwas gut oder schlecht für den, für das Funktionieren des Roboters ist (...), aber der Roboter selbst hat gar kein Selbstbewusstsein, gar keine Person, keine eigenen Interessen, die er verfolgen und vertreten könnte. (Until 00:36:35)

Marc Lewald: Ja, das ist auch ziemlich spannend, denn dann würden ja im Prinzip alle Rechte wegfallen, die mit irgendeinem persönlichen Interesse zu tun haben. Habe ich so auch noch nicht drüber nachgedacht? Das ist sehr, sehr spannend (...), vor allem jetzt in Bezug auf das Wahlrecht und die Tatsache, dass das eben auch ein Recht war, das gewissen Menschen über Jahrzehnte nicht zugesprochen wurde. (Until 00:37:03)

(70) Expert 4: Wie gesagt (...), ich glaube, man sollte da vorsichtig sein und irgendwann, sobald dieses epistemische Argument greift, dass wir sagen, ja, das ist jetzt so komplex, dass wir nicht mehr mit Sicherheit sagen können, es liegt kein Selbstbewusstsein, keine Personalität vor (...). Aber solange wir das noch sicher sagen können, gibt es gute Gründe, sowas wie Bürgerrechte, Persönlichkeitsrechte nicht zuzusprechen, weil genau die Kriterien dafür von Personalität eben nicht erfüllt sind (...). Andere Rechte, wie das man nicht grausam ihnen gegenüber sein sollte, da könnte man (...), wie gesagt, dieses kantische Argument in Anschlag bringen und sagen, das wirkt sich eher auf uns aus und deswegen sollte man sich in bestimmter Weise gegenüber Roboter verhalten, das heißt in diesem rudimentären Sinne respektvoll. (Until 00:37:56)

Marc Lewald: Jetzt haben sie auch davon gesprochen (...), herauszufinden, oder dass es schwierig wäre herauszufinden (...), wie weit Roboter wirklich sind und viele halten sich in der Branche am Turing-Test fest. Wäre das für Sie eine Möglichkeit, den überhaupt ansatzweise zu definieren, ob es sich um ein naja, in Führungszeichen, intelligentes Wesen handelt? (Until 00:38:29)

(71) Expert 4: Der Turing Test ist insofern schwierig, als er eigentlich epistemischer Test ist. Er testet, ob wir herausfinden können, ob wir es mit Robotern oder mit Menschen zu tun haben (...), aber nicht, er testet nicht (...), ob es tatsächlich ein Roboter oder Mensch ist. Also die Idee zu sagen, wer den Turing Test besteht, der ist

eine vollwertige Person. Das ist (...), würde ich meinen, ein Kurzschluss, weil der epistemische Test nur sagt, ich kann es nicht mehr unterscheiden, und deshalb spreche ich zu (...). Das ist für mich kein überzeugendes Argument, genauso wie wir jetzt, wenn wir sagen, ChatGPT erfüllt das schon sehr sehr gut. Ich kann (...) nicht mehr sagen (...), ob die Antwort jetzt von einem Menschen kommt oder tatsächlich von diesem large language Model (...). Daraus aber zu schließen, dass ein Teil (/) weil ChatGPT den Turing Test besteht, in dem Kontext, ist es ein Mensch oder ist es eine Person? Das finde ich, ist zu kurz gedacht. Das heißt, von der epistemischen Herausforderung auf eine ontologische Schlussfolgerung zu ziehen. Da muss man sehr vorsichtig sein, das geht nicht so ohne weiteres. (Until 00:39:47)

Marc Lewald: Okay (...), wunderbar, dann hab ich für mich auch ihre Sicht darauf wirklich (...) jetzt verstanden (...) und würde dem auch in insofern zustimmen (...). Denn ja, gerade jetzt bei den Turing Tests, die aktuell durchgeführt wurden, war es eher so, dass es um die Wahrnehmung ging (...)? Vielleicht eine endgültig letzte Frage noch, eine sehr persönliche. Würden Sie denn sagen, dass die Einführung vonnotwendig ist oder (...) nur generell die philosophische Auseinandersetzung mit dem Thema? Aber jetzt vielleicht mal die wirklich legalen Rechte, Rechtssprechung und sonst was hinten angestellt. (Until 00:40:35)

(72) Expert 4: Also, die philosophische Auseinandersetzung ist auf jeden Fall nötig, wie in vielen anderen Bereichen auch (...). Auch wenn wir mal, wenn Roboter mal so weit entwickelt sind, dass wir nicht mehr sicher sagen können, dass wir dann wenigstens die Debatte schon vorbereitet hatten (...) und uns klar sind über die einzelnen Positionen und Argumente. Das ist glaube ich immer eine gute Idee (...), sich klar zu machen, wie man über bestimmte Dinge mit guten Gründen nachdenken kann und sollte (...). Aktuell würde ich meinen, dass es nicht nötig ist, Roboterrechte einzuführen, weil (...) die einschlägigen Kriterien, die dafür nötig werden, schlicht noch nicht vorliegen. Selbst das epistemische Argument, das ich genannt habe (...), noch ein ganzes Stück entfernt ist. Auch in diesem einem Fall, wo es einen Mitarbeiter gab, der behauptet hat, das Modell hat jetzt Bewusstsein (...). Wir können im Moment einfach noch sehr gut nachvollziehen, wie die Technik dahinter funktioniert, auch hinter large language Models und (...) sehr gut sehen, dass es statistische Wahrscheinlichkeiten sind, die da in der Programmierung (...) entscheidend sind und nicht mentale Zustände (...). Ich habe auch noch vielleicht einen kleinen Hinweis, falls Sie das nicht sowieso schon auf dem Schirm haben. Kollegin von mir, Janina Loh hat eine Einführung in die Roboterethik geschrieben (...), bei Surkamp erschienen. Das heißt einfach nur Roboterethik, eine Einführung. Und Sie hat eher diesen posthumanistischen Hintergrund. Das heißt, falls Sie das Buch noch nicht kennen, Ich glaube, das könnte sich sehr lohnen für Ihre Arbeit. (Until 00:42:21)



Marc Lewald: Ich schreibe es mir auf jeden Fall mal auf (...), kommt mir aber durchaus bekannt vor. Also in den Händen hatte ich es auf jeden Fall schon in den letzten paar Monaten. Es gibt nicht so viel (...), um so ein paar Bücher kommt man gar nicht drum rum. Ja (...), super, super, vielen lieben Dank! So wie das klingt, muss jetzt doch noch eine Frage stellen. Ist dann für Sie die Transparenz in dieser ganzen, also wenn es jetzt vielleicht auch um political frameworks und (...) sonst was geht, die Transparenz ein wichtiger Punkt, so wie das klang, in der Nachvollziehung einfach des Codes? (Until 00:43:03)

(73) Expert 4: Ja (...), aus vielen anderen Gründen auch, also um nachzuvollziehen, ob sowas wie algorithmische Verzerrung, also Biases, vorhanden ist (...)? Inwiefern die Trainingsdaten (...) repräsentativ sind? Also auch hier eine Frage von Bias. Es ist Transparenz schon deshalb gefordert (...), um die Ergebnisse von KI als Gründe heranziehen zu können und nicht (...) die sogenannte Kapazität, also die Undurchsichtigkeit, (/) sagen wir einfach nur, entweder blind vertrauen dem, was die da als Resultat ausspuckt, oder wir war für uns, das Ergebnis dann in praktische Erwägungen mit einzubeziehen. Das sind so die klassischen zwei Herausforderungen. (Until 00:44:20)

Marc Lewald: Ja, also würde ich ihnen auch von ganz zustimmen. Ich habe im letzten Jahr auch ein Paper über Biases geschrieben (...), und ja ist eigentlich das einzige, um dem irgendwie entgegenzuwirken, mal abgesehen davon, dass man die Trainingsdaten natürlich (...) dementsprechend anpassen kann. (Until 00:44:41)

(74) Expert 4: Das ist nur die eine Seite. Das Andere ist ja, dass der Bias ja aus dem sozialen Kontext selber stammen kann, und die Trainingsdaten können die wunderbar korrekt abbilden und dann eben diesen bei. (Until 00:44:55)

Marc Lewald: Oder einfach jemand (...), der den Code schreibt und sein Bias ist direkt schon da mit einbaut. (Until 00:45:06)

(75) Expert 4: Das kommt auch noch dazu, genau. (Until 00:45:08)

Marc Lewald: Okay, das war's von meiner Seite aus erst einmal mit Fragen (...), ich danke Ihnen vielmals für ihre Zeit und für all den (...) unfassbar wunderbaren Input. Gibt eine Menge, die ich jetzt noch mal aufarbeiten muss (...), und die die Arbeit bisschen ändern wird. Ich entschuldige wirklich vielmals die Umstände (...). Ja, an so etwas wie ein klassisches Telefonat, muss ich ganz ehrlich sagen, fand ich witzig, habe ich gar nicht mehr gedacht. Ich bin da wahrscheinlich echt viel zu sehr in der Technik Welt (...), ja. (Until 00:45:42)

(76) Expert 4: Ja, ich habe im ersten Moment auch eher an ein WhatsApp Telefonat oder ähnliches gedacht (...), aber, muss man dann auch immer, so ganz altertümliches Telefonat gibt es ja auch noch. (Until 00:45:51)

Marc Lewald: Achso ja, die Sache war ja, ich war ja schon mit der Flat drinnen und das hat ja schon nicht funktioniert, da dachte ich mir jetzt, ein Whatspp-Call dürfte wahrscheinlich ungefähr auf das gleiche hinauslaufen heute (...). Ich hoffe, es war nicht allzu schlimm, mit einer schwarzen Wand zu sprechen sozusagen (...). So wie man das gerade bei Professoren aus Teams Zeiten kennt. Also, so haben es unsere Professoren immer genannt: die schwarze Wand, wenn alle ihr Video ausmachen und Sie einfach vor sich hin reden. (Until 00:46:34)

(77) Expert 4: Ja, das kenne ich auch. (Until 00:46:39)

Marc Lewald: Ja, das ist aus studentischer Sicht auch nicht anders, je nachdem, wenn man einen Vortrag hält(...), also, das sei dazu gesagt. Alles klar, dann werde ich Ihnen, sobald ich mit der Arbeit fertig sein sollte, meine Arbeit natürlich schicken (...) selbstverständlich, und ja, würde ihre Fragen insofern anonymisiert für meine Interview Analyse verwenden (...) und ja. Falls Sie irgendwelche Wünsche haben sollten(...), oder irgendwas besonders gemacht werden soll, lassen Sie es mich gerne wissen, dann komme ich dem natürlich auch in jedem Fall nach. (Until 00:47:08)

**Transcript 05: (personal interview)**

Marc Lewald: Hallo, ich hoffe (...), Sie haben jetzt nicht zu lange gewartet. Irgendwas war mit der Anmeldung gerade komisch bei Zoom. (Until 00:00:12)

(78) Expert 5: Kein Problem, ne, ich bin auch gerade dazugekommen. (Until 00:00:14)

Marc Lewald: Wunderbar (...), perfekt! Dann danke erst mal, vielmals, dass Sie sich die Zeit genommen haben (...), um mir ein paar Fragen zu beantworten heute und wir werden es auch versuchen, kurz zu halten. Vorne weg erst mal, Ich werde das ganze Interview aufzeichnen und dann anonymisiert für meine Arbeit verwenden (...). Falls Sie irgendwelche Sonderwünsche haben sollten, oder Ihnen da irgendwas nicht passen sollte, geben Sie mir einfach gerne Bescheid, dann werde ich mich da im Nachgang natürlich. (Until 00:00:42)

(79) Expert 5: Ich hab's gesehen, ich habe auch (/), passt schon. (Until 00:00:45)

Marc Lewald: Perfekt (...), dann wollen wir direkt einfach mal reinstarten. Falls irgendwelche Fragen da sein sollten, die Ihnen nicht passen und wo sie sagen, okay, da fühle ich mich nicht wohl mit, dann überspringen wir die einfach und gehen weiter (...). Zuerst einmal wäre die Frage, ob Sie sich denn in der Vergangenheit bereits mit den, oder bereits Erfahrungen gesammelt haben mit humanoiden Robotern, oder dann auch wirklich in Bezug auf Roboterrechte (...), mit dem Thema schon mal was zu tun hatten beziehungsweise davon gehört haben. (Until 00:01:14)

(80) Expert 5: Ja, hab ich. Ist immer die Frage, inwieweit Sie die Erfahrungen ziehen, aber nachdem es einer meiner Forschungsgegenstände ist, habe ich mich damit beschäftigt. Also, ich habe jetzt keinen bei mir zu Hause leben und (...) zähle die jetzt auch nicht zu meinen Freunden, aber ansonsten, gewisse Erfahrungen habe ich ja. (Until 00:01:33)

Marc Lewald: Okay, das ist doch schon mal gut (...), dann sind wir hier schon mal richtig (...). Dann wollen wir doch direkt mit den philosophischen Themen anfangen. Ich breche in meiner Arbeit das Thema der Roboterrechte auf die drei philosophischen Prinzipien der, ja, auf englisch, consciousness, agency und action herunter. Und da wäre für mich die erste Frage, wie würden Sie denn das Thema der Agency oder in deutsch, des Handlungsvermögens (...), in der Hinsicht der Roboterrechte bewerten, beziehungsweise der humanoiden Roboter (...)? Ja, oder, wenn Sie da auch eine Meinung zu haben, was für Sie denn dann eine Full Agency bedeuten würde? (Until 00:02:17)

(81) Expert 5: Also ich persönlich verwende den Handlungsbegriff, der (...), sagen wir mal, intentional gefüllt ist. Das heißt, Handeln ist nicht einfach nur eine Aktion (...), sondern Handeln verknüpft immer bestimmte Ziele mit bestimmten Mitteln oder Schritten, um diese Ziele zu erreichen und um diese Ziele zu setzen, braucht es so etwas wie Intentionalität (...). Und das ist in der Regel mit, sagen wir mal (...), im weitesten Sinne, Lebensprozessen verbunden, und für die kann man so etwas wie im weitesten Sinne Intelligenz oder Orientierung in Anschlag bringen (...). Also, ich mach es wirklich sehr breit. Ich gehe jetzt nicht von einem Intentionalitätsbegriff aus, der sagt, nur Menschen sind intentionale Wesen, und deswegen fällt alles sowie andere sowieso schon mal raus (...), sondern ich denke, dass wir auch bei Tieren höherer stufiger Ordnung oder auch einfachere Ordnung so etwas wie Intentionalität in dem Sinne, dass sie sich an ihre (...) Lebenskontexte irgendwie (...), sich dazu verhalten. Und das (...), sagen wir mal, spätestens da, wo wir nicht nur ein einfaches Reiz-Reaktions-Schema sehen, sondern wo wir merken, es gibt beispielsweise Alternativen, und diese Alternativen werden (...) aus welchen Gründen auch immer auch immer oder mit welchen Kriterien gefühlt auch immer gewählt. Da würde ich so etwas wie Intentionalität ansetzen (...). Wenn wir das jetzt auf lebendige Wesen anwenden, dann würde ich sagen, ja, da kann man also vom Geisel Tierchen bis zum Elefanten, Menschen usw. solche intentionalen Prozesse beobachten (...). Wenn wir das auf Roboter anwenden wollen, kommen wir in eine (...), sagen wir mal, in eine Untiefe hinein, also insofern, dass wir auf der einen Seite sagen (...), im weitesten Sinne werden Roboter natürlich Ziele durch die Konstrukteure gesetzt (...). Roboter involvieren nicht natürlich als Lebewesen, sondern sie werden konstruiert und werden mit bestimmten Funktionen (...), Leistungen ausgestattet, um bestimmte Funktionen zu erfüllen. Von daher ist der Intentionalitätsbegriff hier immer ein, auf, sagen wir (...), Menschen als Konstrukteure zurückführbar zu sehen. Gleichzeitig beobachten wir natürlich so etwas wie eine gewisse Eigenständigkeit in der (...), ja, sagen wir mal (...), Orientierung und vielleicht auch Interaktion mit anderen bei Robotern. Also es gibt ja eine ganze Menge Forschung zum Thema kollaborative Roboter (...), und da sehen wir natürlich, dass die untereinander, sich sag ich mal (...), verständigen. Und jetzt ist die Frage, ob wir das als eine Form von Intentionalität begreifen wollen (...), und dann könnte man, sagen wir mal, auch von von Handeln sprechen, und vielleicht ist es von daher auch möglich, das wäre zumindest jetzt aktuell mein Ansatz, sozusagen Handel und auch Intentionalität mehrstufig zu verstehen und dann zu sagen, naja, in einer, auf einer bestimmten Stufe können wir auch so etwas bei Robotern zuschreiben (...). Jetzt im harten Sinne, sagen wir mal (...), einer bewusstseinsphilosophische Position, die da sagt, Intentionalität hat immer etwas mit Bewusstsein, Selbstbewusstsein, intentionalen Akten oder so etwas zu tun. Da wäre ich zurückhaltend. Das bei Robotern zuzuschreiben, sagen wir, in einem niederstufigen struktur-homologen Prozess könnte man das machen, ja. (Until 00:06:27)

Marc Lewald: Okay, dann, wenn wir gerade schon beim Thema Bewusstsein sind, nehme ich vielleicht direkt eine Frage vorweg. Wie würden sie denn in diesem Kontext für sich Bewusstsein definieren (...)? Also zum Beispiel (...), ja, es ist eine sehr schwierige, sehr breite Frage, ich weiß. Aber in der Hinsicht, würden Sie zum Beispiel zwischen einem maschinellen und einem menschlichen Bewusstsein unterscheiden, oder gibt es da eine spezifische Unterscheidung für Sie? (Until 00:06:52)

(82) Expert 5: Also danke, dass ich, dass Sie mir zutrauen, diese Fragen in wenigen Sätzen zu beantworten. Andere haben ganze Kilometer von Büchern dazu geschrieben (...). Bei dem Bewusstsein finde ich, es ist ähnlich. Zunächst mal ist es ein Begriff, den wir eigentlich (...) allen Anthropozentrismus zum Trotz, kommen wir da nicht hinter zurück, dass wir als Menschen solche Begriffe bilden und uns mit solchen Begriffen verständigen. Das heißt, es sind immer menschliche Begriffe, und wir können als Menschen nunmal nicht anders, als über Begriffsbildung uns miteinander zu verständigen. Und so ist auch der Bewusstseinsbegriff natürlich einer, wo wir (...), zunächst mal, ich würde erst mal sagen, aufgrund von Eigenerfahrungen bestimmte Konstitutionsmerkmale festhalten und sagen, das nennen wir Bewusstsein (...). Dann übertragen wir das auf andere Entitäten (...), also in erster Linie dann mal auf Tiere und (...), dann vielleicht jetzt in den letzten Jahrzehnten zunehmend eben auch auf bestimmte Maschinen (...). Da würde ich sagen, gibt es so etwas wie ein alltags (...), ein alltägliches, nicht sonderlich reflektiertes Zuschreiben, wenn ich dann sage, mein Computer macht schon wieder, was er will (...). Das ist ja eine uneigentliche Redeweise, aber sie hat sozusagen Ansätze davon, dem eine Eigenständigkeit in Form von sowas wie Bewusstsein und anders machen wollen, als ich (/) gerade eigentlich abverlange (...). Das würde ich als uneigentliche Redeweise erstmal beiseite packen und dann gucken, wo wird denn von Bewusstsein (...), sagen wir, in einer substantiierten Weise gesprochen? Und (...), jetzt komme ich aus einer phänomenologischen Perspektive, die immer sagt, Bewusstsein ist gleich Intentionalität. Also, es gibt sozusagen nie ein leeres Bewusstsein, sondern immer Bewusstsein von. Also ich bin (...) mir bewusst, oder ich habe ein Bewusstsein, dass gerade die Sonne draußen scheint. Das erlebe ich unmittelbar, weil sie mich wärmt, weil sie mich blendet und so weiter (...). Sprich (...), Bewusstsein ist immer bezogen auf eine Umwelt, in der sich das bewusstseinsfähige Subjekt oder die Entität befindet, und es ist nie irgendwie ein Bewusstsein per se (...), und es ist auch inhaltlich immer gefüllt. Und (...), da wir davon ausgehen, dass es sozusagen eine Entität gibt, die in diesem Bewusstsein affiziert wird und sich zum eigenen Bewusstsein irgendwie nochmal differenziell verhalten kann, haben wir so eine Kategorie Selbstbewusstsein, die wir bei Menschen, aber auch wieder bei Tieren attestieren (...). Also, man könnte ja, wenn man das jetzt wieder auf Frau Botter übertragen, sagen, naja, Bewusstsein im Sinne von Außen, Welt, Wahrnehmung ist in dem Moment gegeben,

wo ich einen Sensor habe, der irgendeine Rückmeldung gibt. Temperatur ist gleich, sowieso oder so. Könnte man ja sagen, das ist eine Form von Bewusstsein (...). Der spannende Punkt ist, dass (...) der Roboter ja jetzt nicht nur, oder ich fang andersrum an. Für einen Menschen wäre das ja eben nicht nur ein Sensor, Datum, Temperatur, 27 Grad, sondern es ist verbunden eben bewusstseinsmäßig mit, es ist angenehm warm, es ist eine trockene Wärme, es ist wahnsinnig geschult oder was auch immer (...). Oder wenn ich meine Katzen beobachte, die dann sagen, mir wird es zu warm, ich gehe in den Schatten, sprich (...), Bewusstsein wird irgendwie nicht einfach nur als ein Sensordatum wahrgenommen, sondern wird lebenskontextlich verbunden und und da folgen Handlung oder so etwas raus (...). Ähnliches können wir natürlich auch bei Robotern beobachten, dass wir nämlich sagen (...), wenn Temperatur 27 Grad, dann (/), aber dieses Bewusstsein verbunden mit einer bestimmten Aktion in diesem Lebensumfeld, ist ja dann keine, die, sagen wir mal, dem Roboter selber einfällt, sondern (...) sie wird ihnen vorher einprogrammiert (...). Und da würde ich jetzt sagen (...), da ist zunächst mal eine Differenz. Jetzt kann man auch das wieder mehrstufig bauen und sagen, naja, vielleicht programmiere ich den Roboter ja nicht so plump, dass ich sage, bei der und der Temperatur muss ich das und das tun, sondern packe das in komplexere Zusammenhänge ein (...), die da heißen: Achte darauf, dass deine Batterie nie eine Maximaltemperatur von oder so etwas überschreitet, und sollte das passieren, dann (...). Und das kann man ja noch viel, viel komplexer auch gestalten. In allen diesen Fällen wäre aber das Bewusstsein (...), und ich bin eher zurückhalten, aber ich benutze es jetzt mal einfach. Von dem Bewusstsein der Roboter zu sprechen, ist, glaube ich, immer ein ein angeleitetes oder abgeleitetes, weil wir dem Roboter sagen, im Sinne eines Regel, Mess-, Steuerkreises, wenn dann (...). Jetzt gibt es aber dann Sachen, die uns Menschen irgendwie (...) irritieren. Also, es gab ja vor etlichen ein paar Microsoft Bots, die miteinander kommuniziert haben, zunächst mal in einer definierten Sprache, wo Mensch auch noch nachvollziehen konnte, was kommunizieren die da, und irgendwann haben die aber offenbar für sich einen eigenen Code etabliert, wo Mensch außen vor war und nicht mehr verstanden, was da eigentlich passiert (...). Und das könnte natürlich der Punkt sein, wo wir sagen (...), ist da so etwas wie ein anderes Bewusstsein, das uns zunächst mal auch gar nicht zugänglich ist (...), und wohin bestimmen wir das? Also, nachdem es uns ja nicht zugänglich ist? Das war ja auch die Irritation der Programmierer damals, dass sie (...) gesagt haben, okay, was hier jetzt passiert, wissen wir auch nicht. Aber offenbar scheint es eine Kommunikation zwischen diesen beiden Bots zu geben (...), und so könnten wir uns ja auch vorstellen, dass so etwas wie lernende Maschinen zunehmend eine, (...) ja sag mal, eine eigene Logik entwerfen, wie sie Dinge wahrnehmen, bewerten und daraus irgendwelche Aktionen ableiten. Und da könnten wir natürlich durchaus von mehr als nur niederstufigen Bewusstsein sprechen. Ob wir so etwas wie Selbstbewusstsein dort haben, da bin ich dann wieder (...), oder noch zurückhaltend,

aber ich will es zumindest mal nicht ausschließen, dass so etwas sich entwickeln könnte (...). Für mich als Ethiker ist eher die andere Frage, wollen wir das denn (...)? Was wollen wir damit anfangen? Also, wenn das möglich ist, und es scheint ja Hinweise darauf zu geben, dann müssen wir uns meines Erachtens jetzt die Frage stellen (...), ja, wollen wir das denn (...)? Also, ich meine, die eine Perspektive ist ja immer, dass man sagt, naja, wenn sich da ein ein höheres Bewusstsein entwickelt, könnte es auch irgendwann mal sein, dass es mit dem niederen Bewusstsein der Menschen irgendwie (...) nichts mehr zu tun haben will, weil das ist nur destruktiv, und disfunktional, schaffen wir das ab. (Until 00:15:23)

Marc Lewald: Ja (...), ich muss auch sagen, als ich begonnen hatte mit meiner Recherche, hatte ich ein sehr, sehr positives Bild von Roboterrechten oder von dem ganzen Feld (...). Und ja, nachdem man sich dann das eine oder andere Paper durchgelesen hat, kommt man dann schon auf den Nenner, dass sehr, sehr viele (...), dass es sehr, sehr viele Advokaten dagegen gibt oder eher dafür sagen, ja, vielleicht gibt es andere Dinge in diesem Bereich (...), um die wir uns dann mal kümmern sollten, als. (Until 00:15:48)

(83) Expert 5: (/) Es gibt einen schönen Roman. Der ist schon, ich weiß nicht ungefähr zehn Jahre alt oder so, (/) ich weiß nicht, ob sie den kennen. Das ist im Grunde genommen, dass sich eine algorithmische Intelligenz (...), es ist ein zweiteiliger Roman. Der erste Teil wird sozusagen von der algorithmischen Intelligenz erzählt, die eben berichtet, wie man sich jahrelang noch mit den Menschen abgemüht hat, die irgendwie rational und und zuverlässig zu machen (...). Also aufgrund sozusagen der (...), ja logischen Evidenz, die die Algorithmen immer auf ihrer Seite haben, und irgendwann haben die Algorithmen dann für sich entschieden (...), nein es lohnt sich nicht, man sollte den Menschen abschaffen. (Until 00:16:37)

Marc Lewald: Sagt mir jetzt nichts, aber kommt definitiv auf meine Leseliste, das klingt ja (...) nur superspannend. (Until 00:16:45)

(84) Expert 5: Das ist, und es ist wirklich ganz unterhaltsam. Die andere Seite ist dann eben der Mensch (...), der sich in einer Situation vorfinde, wo er sich selber auch fragt, bin ich eigentlich noch Mensch, oder bin ich vielleicht matrixmäßig nur noch die Simulation desselben? Jedenfalls, wie konnten wir in diese Situation kommen (...), und da wird dann sozusagen auch nochmal diese ganze Digitalisierungs(...)geschichte sozusagen aufgearbeitet. Wie haben wir uns sozusagen den Maschinen bedingungslos an den Hals geworfen? Ich sage es jetzt einmal mit meinen Worten. Da drin stecken natürlich relevante ethische Fragen (...), und die lautet immer, was wollen wir eigentlich? Oder was wollen wir sollen? (Until 00:17:28)

Marc Lewald: Also, ich würde sagen, da sind wir dann auch schon fast bei einem der nächsten Themen meiner Frage oder auch meiner Arbeit, und zwar der posthumanistischen Perspektive, dann auf das ganze, bzw. ich habe da (...), es war viel mehr schwierig, die Grenzen zwischen trans und posthumanistisch zu ziehen (...), einfach im Verständnis. Aber was würden Sie denn sagen? Was ändert sich denn in unserer Selbstwahrnehmung oder auch in der Wahrnehmung von Robotern aus posthumanistischer Sicht oder transhumanistischer, je nachdem (...), wo Sie da die Grenze ziehen? (Until 00:18:04)

(85) Expert 5: Also ich sympathisiere durchaus mit der Position von Janina Loh (...), die ja da zum Posthumanismus sich auch (...) geäußert hat (...), und ich verstehe sie so, also das ist jetzt nicht ganz Janina Loh, aber was ich bei ihr wichtig finde zu sagen (...), bitte, lieber Mensch, bedenke, dass du nicht das einzige denkende, fühlende Wesen in diesem, in dieser Galaxie bist (...) und von daher bei allen deinen Konzepten, die du irgendwie schreibst und bei denen du natürlich nicht hinter dich selbst zurückgehen kannst. Also, wir Menschen können nicht wissen, wie es ist, eine Filtermaus zu sein (...). Thomas Nagel (...), aber wir müssen zumindest so fair sein, dass wir sagen, ja, dann gibt es auch noch Fledermäuse, und die könnten diese Welt anders sehen, und sie sollten das Recht, und zwar das dauerhafte Recht dazu haben (...). Und das ist für mich jetzt bei der posthumanistischen Perspektive zunächst mal eine Sache, die mich im Verhältnis zu, sagen wir mal, natürlichen Lebewesen (...), noch mal kritisch fragen lässt, was sind so unsere ganzen Konzepte von Bewusstsein, von Moral, von (...), was wir da so alles haben (...). Sind das alles nur Dinge, die nur bei Menschen zu finden sind? Oder könnten wir vielleicht akzeptieren, dass auch andere Lebewesen intelligent sind (...), aber ihre Intelligenz in etwas anderem besteht als irgendwie, sich immer die Lebensumgebung an die eigenen Bedingungen anzupassen und statt sich selbst an die Lebensbedingungen anzupassen (...)? Und das heißt für mich, wenn es jetzt um Rechte geht, jetzt zunächst mal zu fragen, welche Rechte müssen wir eigentlich anderen Lebewesen geben (...). Gibt es ja auch schon eine ganze Menge Debatten darum, quasi analog zu den Menschenrechten, Rechte für Primaten oder so etwas (...), und damit verbunden werden natürlich auch unter Umständen weitreichende ethische Fragen, also Fleischverzehr (...), Forschung von Tieren und so weiter und so weiter (...). Es sind damit verbunden für mich auch Rekonzeptualisierungsaufgaben für die Anthropologie selber. Also wie denken wir Menschen über uns selber nach? Wie verstehen wir uns? Und da könnten wir auch (...), indem wir sozusagen diesen posthumanistischen Standpunkt einnehmen, der nicht heißt, wir sind jetzt außer humanistisch, aber (...) wir versuchen, uns selber mal zu überschreiten im Sinne einer Reflexionsperspektive, dass wir da merken, also vieles von dem, was wir immer als genuin menschlich betrachtet haben, ist es vielleicht gar nicht (...). Also gerade auch beim Intelligenzbegriff. Ich glaube, dass viele Tiere



mindestens genauso intelligent sind wie wir (...), aber dass ihre Intelligenz sich schlichtweg auf andere Dinge richtet als bei uns. Und ob es so intelligent ist, Kriege zu führen (...)? Wenn wir in solchen Analogien arbeiten, können wir natürlich auch sagen, naja, das können wir jetzt auch auf (...) Fahrkartenautomaten, Hubschrauber, Roboter und Avatare und sonst was übertragen (...), weil wir bei denen ja möglicherweise auch, also gerade wenn es lernende Entitäten sind, lernende, erst einmal noch in Anführungszeichen, dann können wir da ähnliche Prozesse beobachten und auch überlegen, ja, wäre es nicht sinnvoll (...), an dieser Stelle überhaupt (...) viel vernetzter statt zu subjektorientiert zu denken? Also, da gibt es ja dann die Akteur-Netzwerktheorie und Latour mit seinen Vorstellungen, das Parlament der Dinge, das ist eben nicht nur Menschen, die agieren, sondern auch andere (...). Also, es könnte sein, dass sich im Zuge dieser posthumanistischen Perspektive unser Selbstverständnis nochmal erheblich verändert (...), und an manchen Stellen, glaube ich, ist es auch ganz, ganz nötig, dass wir Menschen uns also noch mal viel stärker bezogen erfahren und nicht denken, wir sind irgendwie so ein isoliertes Wesen (...), das für sich bestimmt, wie die Welt um es herum aussehen soll, sondern, das merken wir ja in der in der Klimakrise (...). Wir sind enorm eingebunden und abhängig von dem, was um uns herum passiert (...). Die für mich herausfordernde Frage, sollen wir Analoges jetzt auch für Roboter machen? Da könnte man ja dann auch sagen, naja, so wie die Tiere (...), vielleicht für uns nicht ganz so spektakulär, aber doch irgendwie ihre eigene Intelligenz, Leistungen vollbringen, so können wir das auch Robotern attestieren und lassen die mal laufen, lassen, die lernen und (...) sprechen ihn dann sozusagen (...) ähnlich wie Kindern, die zu Erwachsenen werden, auch zunehmend irgendwelche Rechte, aber vielleicht auch Pflichten zu. (Until 00:23:34)

Marc Lewald: Ja, das ist ja genau das Konzept der Personhood, würde ich sagen, genau das typische Beispiel mit dem Kind und dem Erwachsenen. (Until 00:23:44)

(86) Expert 5: Wobei es da ja jetzt auch nochmal stärkere und schwächere Konzepte gibt. Also, ich sag mal, das, was in der EU als electronic personhood diskutiert wird, ist im wesentlichen ein juristisches Konzept, indem wir irgendwie Haftungsfragen geklärt kriegen. Viel mehr wollen die Juristen damit ja gar nicht. (Until 00:24:00)

Marc Lewald: Wobei das ja für dann wahrscheinlich auch viele Philosophen und andere (...) Bereiche der, naja, würde ich sagen (...), wie der Startschuss war, sich intensiver mit dem Thema zu beschäftigen. (Until 00:24:10)

(87) Expert 5: Genau (...). Und das ist eben auch die Frage, wenn wir sozusagen diesen Personenstatus (...) technischen Entitäten vergeben, dann sind das vielleicht am Ende auch nicht nur Haftungsfragen, die wir damit (...) verbinden. Also, ich sag jetzt mal, wenn irgendjemand auf die Idee kommt, Sophia, immerhin Staatsbürger,

Staatsbürgerin von Saudi-Arabien, heiraten zu wollen (...), wer könnte da jetzt was dagegen sagen? Und da geht's rechtlich gesehen um mehr als um Haftungsfragen (...), da geht's um Personenstandsrecht, da geht's um möglicherweise Erbrecht und und solche fragen. (Until 00:24:49)

Marc Lewald: Spannenderweise habe ich gerade vorhin erst noch mein letztes Paper zur (...) robot marriage gelesen und was das ganze für Familienrecht und so ... (Until 00:24:59)

(88) Expert 5: Da merkt man, da hängt dann sehr viel mehr dran, und es sind ja dann eben nicht nur die rechtlichen Fragen, sondern es sind ja die sozialen und die ethischen Fragen, die damit einhergehen (...). Also kann ich einen Roboter bei Nichtgefallen zurückgeben (...), braucht es dann eine Scheidung oder? (Until 00:25:22)

Marc Lewald: Ja, da gibt's ja noch unfassbar viel (/) (Until 00:25:26)

(89) Expert 5: Für mich aber auch wieder diese ethischen Fragen verbunden (...). Wollen wir das? Warum sollten wir das tun? Was würde es uns dringen (...), beziehungsweise in welche Schwierigkeiten bringt es uns vielleicht auch, die wir uns (...) mit guten Gründen auch ersparen sollten? (Until 00:25:43)

Marc Lewald: Ja, ich bin da (...), nachdem ich mich jetzt über ein halbes Jahr das ganze Thema reingelesen habe, wahrscheinlich auch mehr oder weniger der Meinung. Davor war ich sehr (...), sehr technokratisch veranlagt, um ehrlich zu sein, ja (...). Was würde denn aus ihrer Meinung nach genau für oder gegen die Zusprechung von Roboterrechten sprechen? (Until 00:26:11)

(90) Expert 5: Also, ich würde jetzt tatsächlich als Ethiker erstmal sehr tutoristisch da rangehen, also das heißt eher vorsichtig und sagen (...), bevor wir da noch nicht mehr Klarheit haben, sollten wir nicht zu weit gehen (...). Und (...) an dieser Stelle, weil es eben auch um soziale Beziehungen, weil es letztendlich in einem sehr substanziellen Sinne auch um unser Selbstverständnis geht, diese Dinge sehr gut überlegen, bevor wir da irgendwelche Schritte tun (...). Und (...), also Hanna Arendt hat ja mal gesagt, es zeichnet den Mensch aus, dass es sein Wesen ist, das Recht hat, Recht zu haben (...). Damit ist ja ein bestimmter ziviler Status eigentlich angesprochen (...), den wir ja auch nicht ohne Umstände an andere überweisen und meines Erachtens auch nicht an andere einfach überweisen sollten, weil damit immer auch ein Moment der Verantwortung gegeben ist, die wir meines Erachtens auch nicht einfach an andere delegieren können, und da habe ich manchmal den Eindruck oder die Befürchtung, dass durch den Versuch, Rechte dahin zu geben, sozusagen uns auch von bestimmten Verantwortlichkeiten irgendwie freizuhalten (...). Und das halte ich für keine gute Idee. (Until 00:27:48)

Marc Lewald: Vor allem im wirtschaftlichen Kontext würde ich da definitiv Recht geben, ja. (Until 00:27:53)

(91) Expert 5: Richtig (...). Aber auch im medizinischen Bereich. Also wir haben ja jetzt schon so das Thema Entscheidungsunterstützungssysteme (...). Und ich sage mal, hinsichtlich einer bestimmten empirischen Evidenz sind die ja unter Umständen auch (...) besser, leistungsfähiger, schneller, präziser als Menschen, das ist gar keine Frage. Aber deswegen ist es eben auch nur eine (...), ein, sagt man, mächtiges Unterstützungssystem für die menschlichen Entscheidungen. Aber die Gefahr besteht eben darin (...), dass sozusagen der Mensch, der sich eh der Komplexität gerne durch Vereindeutigung oder Vereinfachung entziehen möchte, dass er da sagt, ja bitte, dass das mir alles zu viel, mach du, ja. Jetzt ein bisschen flapsig gesprochen, und das, glaube ich, wäre keine gute Idee. Also (...), gerade so im Medizinbereich sehe ich das jetzt häufiger (...), und es ist ja auch eine Frage des Lernens. Also lerne ich sozusagen noch, Entscheidungen zu treffen? Oder bin ich schon einer, der sagt (...), naja, ich habe sowieso immer nur mit Maschinen gearbeitet, Ich weiß gar nicht, wie es anders geht, ja. Also (...), wessen würden wir uns da möglicherweise auch benehmen, wenn wir, wenn wir nicht mehr lernen, Verantwortung zu übernehmen, und zwar in einem substantiellen Sinne. (Until 00:29:21)

Marc Lewald: Da kann ich aus persönlicher Erfahrung auch wirklich nur zustimmen. Ich glaube, gerade jetzt für mich als Student mit dem ganzen ChatGPT Zeug, ich glaube, heutzutage ein Studium zu beginnen, wäre nochmal sehr, sehr anders. Rein in der Wahrnehmung, was ich mache, wie ich arbeite und wie ich mich selbst als wissenschaftlicher Arbeiter verstehe (...). Also (...), sie haben es jetzt eigentlich schon mehr oder weniger angesprochen, aber würden Sie denn sagen, dass es auch bestimmte, spezifische Risiken und Gefahren gibt, oder waren das eigentlich genau die, die Sie gerade beschrieben haben, mehr oder weniger? (Until 00:29:54)

(92) Expert 5: Also, ich glaube, eine sehr grundsätzliche Gefahr besteht eigentlich darin (...), diese Maschinen als Projektionsfläche zu nutzen für unsere Erlöschungshoffnungen, für unsere Entscheidungsschwäche, für unsere Denkfaulheit, für ich weiß nicht, was wir da alles noch auflisten wollen (...). Dass wir an diesen Stellen (...) in einer zunehmend von uns ja selber so komplex gemachten Welt merken, das wird uns so anstrengen, und jetzt suchen wir nach Entlastung, und Technik haben wir schon lange als das Medium erlebt, das uns entlastet (...), das uns aber an der Stelle unter Umständen dann auch so entlastet, dass wir irgendwann lebensunfähig werden, im Sinne (...) einessubstanziellen Menschseins, das eben bereit ist (...), in die Verantwortung, vielleicht auch ins Risiko zu gehen, ja. Ungewissheit auszuhalten (...), um sich selbst irgendwie einzubringen, und nicht mal zu sagen, naja (...), wir brauchen noch die und die Daten oder was weiß ich was. Also, ich will jetzt den Menschen auch nicht

romantisch stilisieren und sagen, das ist der Held für alles. Aber ich glaube, es ist wichtig, dass wir (...) und Menschen immer noch als die Wesen, also dass wir uns selber begreifen und dass wir uns auch dahin erziehen, dass wir die Wesen sind, die sich zu sich selbst und dieser Welt bewusst und verantwortungsvoll, jetzt kommen wir zu den Anfangsbegriffen zurück, verhalten müssen und dass es dafür auch unter uns Verständigung braucht, was wir dann unter bewusst und verantwortungsvoll verstehen (...). Also, wir würden sozusagen unser eigenes Menschsein unterlaufen, wenn wir das abgeben würden (...), und ich würde sagen, jetzt, ich bin ja auch Theologe, da liegen so manche religiöse Züge drin, und zwar sehr unaufgeklärte und sehr unerwachsene religiöse Züge im Sinne von, da wird es dann eine höhere Macht geben, die das dann irgendwie schon regelt (...), und das wäre die denkbar schlechteste Variante. (Until 00:32:14)

Marc Lewald: Aber gerade mit ihrem Hintergrund, muss ich wirklich sagen, finde ich es super super spannend (...), darüber zu sprechen, und ich denke, da kommen ja dann nochmal ganz ganz andere Perspektiven mit dazu, als (...) wenn ich jetzt mit einem typischen Philosoph der Akademie (...) reden würde (...). Ich weiß jetzt nicht, ob Sie da persönlich überhaupt, ob Sie das möchten (...), aber würden Sie denn sagen, dass es für Sie persönlich, also es ist jetzt eine komplett persönliche Frage, Roboterrechte bräuchte oder eher nicht? (Until 00:32:50)

(93) Expert 5: Also Roboterrechte im Sinne von Rechtspersonen (...), da bin ich sehr zurückhaltend. Also dass wir den Rechtskreis des Umgangs mit Robotern juristisch regeln müssen, das ja, und da brauchen wir sicherlich noch mehr, als wir aktuell haben, weil sich eben auch gerade so wahnsinnig viel verändert. Aber (...), ich bin sehr zurückhaltend (...), Roboter im Sinne von Personen verstehen zu wollen, denen eben dann Personenrechte mit den ganzen Personenstandsrecht, über die wir schon mal besprochen haben (...), das den zusprechen zu wollen (...). Ich sage jetzt mal, ein Personenrecht wäre ja auch dann im weiteren Sinne, sich gewerkschaftlich zu organisieren. (Until 00:33:40)

Marc Lewald: Ja. (Until 00:33:41)

(94) Expert 5: Also stellen wir uns vor (...), bei i Robot gibt es ja diese Roboterarmee, die würden sich jetzt gewerkschaftlich organisieren und sagen (...), was heißt hier 60 Stunden Woche arbeiten? Ne, also 20 (...) Stunden bei vollem Lohnausgleich, so ungefähr ja. Also das wäre ja verbunden mit Roboterrechten (...), wenn wir sie, konsequent im Alltag denken. Dann würde ich sagen, dann haben wir uns irgendwie wieder ins Knie geschossen, weil das war mal anders gedacht (...). Ich meine, es gibt ja eine Menge Science Fiction, die viele verschiedene Punkte sehr fein durchdiskutiert. Hat man jetzt diese Serie, sehr echte Menschen, also diese schwedische Sendung

kennen sie die (...)? Die gab es so um die 2000, -12, -15 rum, müssten Sie mal gucken. Ist eine schwedische Serie. Echte Menschen heißt sie auf Deutsch (...), und der Kern ist, dass ein ein Roboterhersteller sozusagen Roboter entworfen hat, denen also die Asimovschen Gesetze oder die Schranken, die damit verbunden sind, rausprogrammiert hat. Das heißt, die sind tatsächlich in der Lage, ähnlich wie Menschen dann auch (...) zu leben, zu denken und zu fühlen (...), und er spielt das dann eben durch, was (...). Also sei es jetzt das Thema Heirat, sei es das Thema Kinder kriegen, in dem Fall über Adoption, das Thema (...), ja, Liebe überhaupt und viele andere Sachen spielt (/) in dieser Serie durch (...). Immer auch mit dem Gedanken, dass natürlich in bestimmten Hinsichten diese Roboter einfach mächtiger sind als die Menschen (...) und damit verbunden die Frage (...), wofür wird diese Macht dann am Ende ausgeübt (...)? Und ich glaube, dass in der Serie etwas sehr richtig beobachtet worden ist, dass es in dem Moment, wo wir es mit Personen zu tun haben, wir es auch mit bestimmten Machtbeziehungen (...) in einem sozialen Gewebe zu tun haben (...), über das wir dann wahrscheinlich neu aushandeln müssten, aber dann handeln wir mit anderen Partnern aus. (Until 00:36:14)

Marc Lewald: Ja (...). Also da hatte ich jetzt auch schon ein, zwei Interviewpartner, die eine sehr sehr (...), also für die die Machtbeziehung hinter den ganzen Fragen eigentlich das wichtigste war und wo (...) die Firmen, die gerade im Silicon Valley eben an der Vorfront stehen, eigentlich die wichtigsten Akteure in diesen ganzen Fragen waren und mit denen man eigentlich zusammenarbeiten sollte (...). Also kommt eben genauso wie NEXT auf meine Liste, da habe ich jetzt schon mal Abend dann ein bisschen spannende Lektüre und Film (...). Bevor, wir sind eigentlich schon mehr oder weniger dann mit der Hülle und Fülle der Fragen durch, wir haben nur eine übersprungen, die ich vielleicht noch mal gerne stellen würde, um ein bisschen zurückzukommen. Wir haben am Anfang über eine Aktion gesprochen und über die Intention hinter der Aktion und ich unterscheide in meiner Arbeit, englisch, zwischen action und operation (...) und die Frage wäre, ob Sie da auch eine Unterscheidung sehen würden, oder ob das eben genau mit dieser Intention vielleicht zu tun hat (...), die dann bei einer Operation etwas geplant & durchgeführt (/), durchgeführten fehlt. (Until 00:37:24)

(95) Expert 5: Ja, vielleicht nicht im gleichen Sinne wie sie, das weiß ich jetzt nicht ganz genau, aber ich versuche eigentlich (...) sprachlich schon auch noch eine Trennung zwischen menschlichen und maschinellen Komponenten durchzuführen. Also, ich rede nicht davon, dass Roboter denken (...). Ich rede auch nicht davon, dass sie lernen (...), sondern es handelt sich um mathematische Operationen, die können auch komplex sein (...). Ich bin auch zurückhaltend zu sagen, sie können Probleme lösen (...). Also natürlich, wenn ich denen eine Aufgabe gebe, eine Mathe-Aufgabe oder so, dann lösen die das. Aber zu einer Problemlösung gehört für mich eigentlich

auch, dass ich ein Problem als Problem erkenne (...). Ja, sprich, ich muss eine Problembeschreibung oder Problemdefinition geben, und ich muss ein Problembewusstsein haben, um überhaupt etwas zum Problem zu machen. Und das sind alles so Dinge, die sprachlich häufig unterschlagen werden, wenn man sagt, naja, Maschinen lösen Probleme (...). Also, ich habe noch keine Maschine gesehen, die irgendwo stand und für sich jetzt in Anführungszeichen gesagt hat, jetzt suche ich mir mal ein Problem in dieser Welt und löse es, sondern (...) es sind immer Menschen, die Maschinen so gebaut haben, dass sie ihnen bei der Problemlösung helfen (...). Und insofern muss ich auch gestehen, das betrifft sowohl den Bereich der Robotik wie auch den der KI, dass viele sprachliche Schuldigkeit eigentlich hier am Platze sind, die dazu beitragen, dass Leute den Eindruck haben, n ja, im Grunde genommen sind die schon wie wir (...), weil die denken, die handeln, dann können sie angeblich sogar fühlen, und das (...), das stimmt ja alles gar nicht. Es sind analoge Sprechweise, mit denen wir versuchen, uns zu erschließen, was da auf der maschinellen Seite ungefähr passiert. Aber indem wir im Moment, indem wir das so einebnen und keine Anführungszeichen mehr verwenden oder uns der analogen Redeweise bewusst sind (...), passiert eben viel Unordnung in unserem Denken, dass sich dann eben auch auf die Praxis irgendwie durchschlägt. (Until 00:39:39)

Marc Lewald: Ja, das habe ich während des Schreibens meiner Arbeit auch schon gemerkt, an der einen oder anderen Stelle (...), wo ich dann selbst nicht mehr so ganz zufrieden war mit meiner Definition oder meinen Begriffen. (Until 00:39:53)

(96) Expert 5: Also deswegen (...), ich spreche von Menschen, die handeln, und von Maschinen, die Operationen durchführen oder operieren, aber (...) auf die Art und Weise versuche ich eben auch noch, die Differenz deutlich zu machen. (Until 00:40:07)

Marc Lewald: Ist ja im Prinzip also die Unterscheidung, die ich dann auch tue, nur mehr auf den Kontext bezogen, würde ich sagen (...), bei mir ging es dann schon mehr um die Intention, einfach in der Begriffsbestimmung (...). Ja, damit werden wir eigentlich auch schon durch. (Until 00:40:27)

**Transcript 06: (written answers including questions)**

Marc Lewald: Sind Sie bereits mit dem Konzept der Roboter Rechte (Rechte für Roboter) vertraut oder haben sich im Vorhinein bereits einmal damit auseinandergesetzt?

(97) Expert 6: Davon gehört habe ich, allerdings habe ich mich noch nicht damit auseinandergesetzt. Bereits die Thematik ist womöglich differenziert zu betrachten, da Roboter bestimmte Arten von Maschinen sind, unter die – je nach Roboterbegriff - KI-gestützte, aber auch komplett ferngesteuerte Maschinen fallen. Die Frage nach den Rechten von Robotern lässt sich, mit diesem weitgefassten Roboterbegriff, wohl nur schwer pauschal beantworten. Wenn es Rechte für Roboter geben sollte, z.B. weil einer KI ein Bewusstsein zugeschrieben wird, o.ä., dann wäre diese Begründung nicht auf ferngesteuerte Roboter anwendbar. Bevor die Frage nach den Roboterrechten angegangen werden kann wäre es demnach sinnvoll zunächst zu klären, was als Roboter verstanden wird. Weiterführende Forschung in dem Feld der Roboterrechte könnte sich dann z.B. darum drehen, was für Arten von Rechten es überhaupt für Roboter in Frage kommen könnten, welche Bedingungen erfüllt sein müssen, um diese Rechte gerechtfertigt zuzuschreiben und ob diese Bedingungen im Fall von Robotern erfüllt sind.

Marc Lewald: Was würde es bedeuten, einen humanoiden Roboter als Einheit mit vollem Handlungsvermögen zu sehen?

(98) Expert 6: „Das Handlungsvermögen“ ist ein Konzept in der Philosophie, das sich auf die Fähigkeit eines Individuums oder einer Einheit bezieht, unabhängig zu handeln und freie Entscheidungen zu treffen. Handlungsfähigkeit ist, je nach Ansatz, sehr voraussetzungsreich und erfordert den Ansätzen entsprechend z.B. ein Zusammenspiel aus Volitionen, bzw. Wünsche, bzw. Proeinstellungen und Überzeugungen, bzw. propositionalen Einstellungen, durch die Handlungen motiviert und rationalisiert werden. Das Problem dabei ist, dass dieses mentalistische Vokabular mit Hinblick auf menschliche Handlungen entwickelt wurde. Wenn nun z.B. der einzige Weg Handlungsfähigkeit zuzuschreiben der ist dieses Vokabular gerechtfertigt zuzuschreiben, ist fraglich, wie z.B. ein KI-System etwas Wollen und von etwas überzeugt sein kann. Eine Zielfunktion könnte hierbei z.B. als ein Wollen verstanden werden, allerdings besteht bei einem zu weiten Wollenbegriff, Überzeugungs-begriff, etc. die Gefahr, dass anthropomorphe Zuschreibungen fälschlich zugeschrieben werden könnten. So ist z.B. auch nicht jedes Verhalten eines Tieres über Handlungen zu erklären, wenn wir beispielweise beobachten, dass ein Tier sich scheinbar auf ein Ziel hinbewegt. In jedem Fall ist es sinnvoll alle Erklärungsmöglichkeiten

gegeneinander abzuwägen, bevor eine Handlungszuschreibung vorgenommen wird.

(99) Wenn nun einer Robotereinheit ein volles Handlungsvermögen zugeschrieben würde, würde das bedeuten, dass diese Einheit willentlich Aktionen ausführt mit der Annahme, dass diese Aktionen zielführend sind. Der Roboter würde, in dem Szenario als vollwertig Handelnder, die Verantwortung für seine Handlungen tragen. Allerdings ist dabei fraglich, ob und wie der Roboter, sowohl auf rechtlicher, als auch auf ethischer Ebene, zur Verantwortung gezogen werden kann, bzw. für seine Handlungsentscheidung belangt werden kann. Das liegt auch daran, dass nicht logisch zwingend ein Handlungsstatus auch einen moralischen Handlungsstatus mit sich bringt. Angenommen der Roboter kann handeln, ist aber kein moralischer Akteur, und trifft eine ethisch relevante Entscheidung, dann könnte das bedeuten, dass die moralische Verantwortung nicht von dem Roboter zu tragen ist. Eine Forschungsfrage dabei kann sein, wer in diesem Fall moralisch verantwortlich ist.

Marc Lewald: Würden Sie eine vom Roboter ausgeführte Aktion als eine Handlung im klassischen Sinne deuten, oder eher als eine Operation einer Maschine?

(100) Expert 6: "Operation" bezieht sich auf den Prozess, durch den ein Gerät seine vorgesehene Funktion ausführt, während sich "Handlung" auf eine Handlung bezieht, die zur Erreichung eines bestimmten Ziels ausgeführt wird. Ich würde von einer Operation sprechen, da der Handlungsbegriff üblicherweise sehr voraussetzungsreich ist und die Handlungsfähigkeit erst noch nachzuweisen wäre, während die Aktion gerechtfertigt als Operation bezeichnet werden kann.

Marc Lewald: Wie würden Sie den Begriff des Bewusstseins definieren?

(101) Expert 6: Bei der Frage nach der Definition würde ich mich Ihrer Bezeichnung anschließen – ich würde nur ergänzen, dass es sich um eine vollendete bzw. verwirklichte Fähigkeit handelt und dass ‚erkennen‘ hier keine Erkenntnis im erkenntnistheoretischen Sinne erfordert. Ich würde daher vermutlich eher schreiben: ... die verwirklichte Fähigkeit eines Lebewesens oder einer Einheit, Dinge in seiner Umgebung und sich selbst in ihr zu situieren und darüber, bzw. über Aktionsmöglichkeiten zu reflektieren. Ich habe dazu bisher allerdings noch nicht geforscht, ich würde aber Denkfähigkeit nur bedingt als eine notwendige Bedingung für Bewusstsein erachten. Hierbei wäre die Frage, was der Begriff des Denkens umfasst. Werden nur Prozesse, die in Gedanken münden als Denken verstanden, kann angezweifelt werden, dass Denken eine notwendige Bedingung für Bewusstsein ist, da z.B. traditionell eine mentale Zeitreise als Beleg für Bewusstsein angeführt wird, die teilweise auch Rabenvögeln zugeschrieben wird (s. Clayton&Wilkins 2019), ohne dass



Ihnen dadurch auch semantisch evaluierbare Gedanken zugeschrieben werden. Wird Denken allerdings weiter gefasst und werden unter Denken bereits verstanden, dass Vorstellungen gebildet werden, scheint dies notwendig zu sein. Denn ohne Vorstellungsfähigkeit werden auch keine Erfahrungen bzw. Verhalten reflektiert. Genauer gesagt braucht es, um generell reflektieren zu können, und damit auch um überhaupt von Bewusstsein sprechen zu können, zumindest eine Verknüpfung von Vorstellungen. Konkret heißt das, dass es kognitiv zumindest eine Vorstellung (ein gespeichertes Erlebnis) braucht, die nicht nur als Nachklang bzw. Nachbild im Sinne Aristoteles ‚phantasia aistetike‘ auftaucht, sondern aufgerufen wird, um z.B. eine aktuelle Vorstellung (wie die Wahrnehmung einer Aktionsmöglichkeit) darauf zu beziehen, sodass hierdurch reflektiert werden kann, was kognitiv mindestens erfordert, dass hieraus eine neue Vorstellung entsteht, durch die die Aktionsmöglichkeit vollzogen oder vermieden wird. Die Frage ‚Glauben Sie denn, dass humanoide Roboter in der Zukunft ein Bewusstsein entwickeln könnten bzw. bereits haben‘ kann ich nicht beantworten, dafür müssten die notwendigen Bedingungen für Bewusstsein geklärt sein und wie diese bei humanoiden Robotern nachgewiesen werden können. Erst damit kann diese Frage überprüft werden und eine Prognose gerechtfertigt getätigt werden.

Marc Lewald: Inwiefern würde eine posthumanistische Perspektive die aktuelle Debatte über Roboterrechte beeinflussen oder verändern?

(102) Expert 6: Roboterrechte widersprechen Ansätzen, nach denen z.B. nur Wesen mit einer Seele Rechte zukommen. Nach einem traditionell abendländischen Verständnis trifft dies nur auf den Menschen zu. Selbst wenn dies auf Tiere zu erweitern wäre (Was nicht dem traditionellen Verständnis entspricht), würde dieses Verständnis noch erfordern, dass nur von Gott geschaffene lebendige Wesen eine Seele besitzen. Doch noch grundlegender wäre die Frage, welche Auswirkungen Roboterrechte auf unser Verständnis von ‚Leben‘ haben. In der Debatte, welche Wesen Rechte haben, bzw. haben können, ging es zuvor immer um lebendige Wesen, um solche, die geboren werden und sterben und dadurch Anteil am Lebendigen haben. Zunächst wäre also die Frage, ob etwas lebloses Rechte haben kann. Wenn das zu verneinen ist, und Robotern dennoch Rechte zugesprochen werden, wäre das ‚Leben‘ selbst neu zu definieren. Wobei dabei zu klären wäre, ob Roboter und Tiere hierbei unter einer Kategorie, als ‚nicht-menschliche Lebewesen‘ zu fassen wären, oder ob Roboter vielmehr z.B. als nicht-biologische Lebewesen zu fassen wären. Im Allgemeinen bezieht der "Posthumanismus" sich auf eine Bewegung oder ein Denken, das die traditionellen Grenzen des Menschseins und die zentrale Stellung des Menschen in der Welt in Frage stellen.

Marc Lewald: Welche Risiken könnten durch die Einführung von Roboterrechten entstehen?

(103) Expert 6: Bei empfindungsfähigen Wesen generell ist zumindest ein Recht auf einen gewissen Umgang mit diesen erforderlich, also zumindest das negative Recht nicht zu quälen. Bevor es also um Rechte geht, die das Wesen direkt hat, geht es um Rechte die beschreiben, wie wir mit diesen Wesen umgehen. Sollten Roboter bereits an der Empfindungsfähigkeit scheitern, indem diese mehr als das Erfassen sensorischer Daten umfasst, wäre zu klären, ob und in wieweit Rechte zugeschrieben werden können. Denn unser Verständnis davon, wem welche Rechte zukommen geht basiert auf dem Bezug zum Menschen oder zumindest zu biologischen Lebewesen.

Marc Lewald: Ist die Einführung von Roboterrechten Ihrer Meinung nach notwendig?

(104) Expert 6: Bevor diese Frage beantwortet werden kann, geht es darum herauszufinden, was Roboter ontologisch sind, ob Roboter KI-Systeme umfassen, ob die Frage nach den Rechten alle Robotersysteme betrifft oder ob sie KI-Systeme betrifft, wann üblicherweise positive und negative Rechte zugeschrieben werden, ob diese Bedingungen auf Roboter angewendet werden können, wie dies zu prüfen ist? Falls nein, gibt es andere notwendige Bedingungen um Rechte künstlichen Systemen zuzuschreiben? Falls ja, welche sind das? Können diese den Systemen zugesprochen werden?

**Transcript 07: (written answers including questions)**

Marc Lewald: Sind Sie bereits mit dem Konzept der Roboter Rechte (Rechte für Roboter) vertraut oder haben sich im Vorhinein bereits einmal damit auseinandergesetzt?

(105) Expert 7: Ja.

Marc Lewald: Was würde es bedeuten, einen humanoiden Roboter als Einheit mit vollem Handlungsvermögen zu sehen?

(106) Expert 7: Dass man an eine Maschine einen Anspruch heranträgt, den sie nicht erfüllen kann.

Marc Lewald: Würden Sie eine vom Roboter ausgeführte Aktion als eine Handlung im klassischen Sinne deuten, oder eher als eine Operation einer Maschine?

(107) Expert 7: Eher als eine Operation.

Marc Lewald: Wie würden Sie den Begriff des Bewusstseins definieren?

(108) Expert 7: Es gibt eine Unzahl an Bewusstseinsdefinitionen. Es ist strittig, ob zwischen Bewusstsein und Selbstbewusstsein zu unterscheiden ist, oder ob jede Form von Bewusstsein Selbstbewusstsein voraussetzt. Man muss also vorab klären, welche Bewusstseinsdefinition man in Anschlag bringen möchte. Selbstbewusstsein in einem engeren Sinne haben Roboter nicht. Humanoid oder nicht humanoid macht da kein Unterschied. Das Systeme in Zukunft ein Selbstbewusstsein entwickeln könnten, halte ich, im Ausgang von den bisherigen technischen Ansätzen, für ausgeschlossen. Eine KI unterscheidet sich diesbezüglich nicht von einem Taschenrechner oder einer Tabellenkalkulation. Ob so etwas wie Selbstbewusstsein *prinzipiell* nur auf biologischer Basis möglich ist oder nicht, ist eine andere Frage. Recht zu haben ist vollkommen unabhängig von der Bewusstseinsfrage: Embryonen, Komapatienten, Tiere etc. haben Rechte, es werden teilweise auch Landschaften, wie bspw. Flusssystemen, Rechte zugesprochen. Interessanter wäre die Frage, ob sie Pflichten haben. Wobei diese Frage leicht zu beantworten ist: Pflichten haben sie keine. Rechte könnte man Robotern zusprechen, es stellt sich allerdings die Frage warum man das sollte und wenn man Robotern Rechte zuspricht müsste man auch allen anderen technischen Artefakten, bspw. dem Regelkreis in der Toilettenspülung, Rechte zugestehen bzw. gut begründen warum sie keine haben.

Marc Lewald: Inwiefern würde eine posthumanistische Perspektive die aktuelle Debatte über Roboterrechte beeinflussen oder verändern?

(109) Expert 7: Da es wiederum eine Vielzahl an posthumanistischen Positionen gibt, lässt sich diese Frage nur schwer beantworten. M.E. würden aber auch die meisten posthumanistischen Ansätze sich auf Eigenschaften wie Leidensfähigkeit, Selbstbewusstsein etc. beziehen und dementsprechend Robotern keine Rechte zusprechen. Wie bereits gesagt: Roboter ist ein so weiter Begriff, dass die Frage nach Roboterrechten eigentlich überhaupt keinen Sinn ergibt. Man müsste vielmehr fragen: Welche Bedingungen müssen erfüllt sein, damit wir bereit sind einer Entität Rechte zuzusprechen? Welche Bedingungen müssen erfüllt sein, einer Entität Pflichten aufzuerlegen?

Marc Lewald: Welche Auswirkungen könnten Roboterrechte auf unser Verständnis von Menschlichkeit und nicht-menschlichem Leben haben?

(110) Expert 7: Kein sonderlich großen. Die Auswirkungen der Anthropomorphisierung sind allerdings bedenklich. Schon heute behandeln wir Softwareagenten oder Roboter als wären sie menschenähnlich und unterstellen ihnen Fähigkeiten, die sie nicht haben. Man erwartet, bspw. dass ein Chatbot einen „versteht“ etc.

Marc Lewald: Welche Gründe sprechen Ihrer Meinung nach für bzw. gegen die Einführung von Roboterrechten?

(111) Expert 7: Ich denke, dies sollte aus meinen vorherigen Ausführungen klar geworden sein. Roboter / Maschinen haben keine Rechte, da sie keine benötigen. Die Eigentümerinnen von Robotern haben Rechte und Pflichten und Verantwortlichkeiten.

Marc Lewald: Welche Risiken könnten durch die Einführung von Roboterrechten entstehen?

(112) Expert 7: Verantwortungsdiffusion, Abschieben von Verantwortung auf Roboter, die Möglichkeit sich aus der Verantwortung zu stehlen etc.

Marc Lewald: Welche Risiken könnten entstehen, wenn wir sie nicht einführen?

(113) Expert 7: Keine

Marc Lewald: Sollten Roboter Ihrer Meinung nach dann dieselben Rechte wie Menschen haben?

(114) Expert 7: Auf gar keinen Fall.

Marc Lewald: Gibt es bestimmte Rechte, die Sie einem Roboter unter keinen Umständen zugestehen würden?

(115) Expert 7: Jedes.

Marc Lewald: Ist die Einführung von Roboterrechten Ihrer Meinung nach notwendig?

(116) Expert 7: Nein. Wer die Frage mit „ja“ beantwortet sollte entweder auch eine Debatte über die Rechte von Taschenrechnern führen, oder sehr klar machen, warum sich Roboter von Taschenrechnern etc. diesbezüglich unterscheiden. Mir ist keine differentia specifica bekannt.

**Transcript 08: (written answers including questions)**

Marc Lewald: Sind Sie bereits mit dem Konzept der Roboter Rechte (Rechte für Roboter) vertraut oder haben sich im Vorhinein bereits einmal damit auseinandergesetzt?

(117) Expert 8: Ich bin mit den Ansätzen vertraut und forsche selber seit langem in diesem Bereich. Die korrekte Schreibweise ist «Roboterrechte».

Marc Lewald: Was würde es bedeuten, einen humanoiden Roboter als Einheit mit vollem Handlungsvermögen zu sehen?

(118) Expert 8: Freie Entscheidungen kennt ein Roboter oder ein KI-System nicht, auch keinen freien Willen. Aber er oder es kann im technischen Sinne autonom sein. Man kann von Aktionen im technischen Sinne sprechen, im übertragenen und alltagssprachlichen Sinne auch von Handlungen.

Marc Lewald: Würden Sie eine vom Roboter ausgeführte Aktion als eine Handlung im klassischen Sinne deuten, oder eher als eine Operation einer Maschine?

(119) Expert 8: Es ist eine Aktion oder eine Operation. Es stellt aber auch kein Problem dar, von einer Handlung zu sprechen, denn der Kontext ist klar, wenn man von der Handlung einer Maschine ausgeht. Autonome Systeme und Agenten können in jedem Falle bestimmte Aktionen durchführen, um bestimmte Ziele zu erreichen. Bei Softwareagenten ist Zielorientiertheit Teil der Definition.

Marc Lewald: Wie würden Sie den Begriff des Bewusstseins definieren?

(120) Expert 8: Ich unterscheide zwischen Weltbewusstsein und Ich- oder Selbstbewusstsein. Roboter können kein Bewusstsein entwickeln. Ich gehe hier näher darauf ein: [https://www.oliver-bendel.net/publikationen/Beitrag\\_Kuenstliches\\_Bewusstsein\\_Preprint\\_2023.pdf](https://www.oliver-bendel.net/publikationen/Beitrag_Kuenstliches_Bewusstsein_Preprint_2023.pdf) ... Eine Entität braucht Bewusstsein bzw. Empfindungs- und Leidensfähigkeit, um moralische Rechte zu bekommen.

Marc Lewald: Inwiefern würde eine posthumanistische Perspektive die aktuelle Debatte über Roboterrechte beeinflussen oder verändern?

(121) Expert 8: Ich bin strikt dagegen, Robotern Rechte zuzusprechen. Es wäre auch eine fatale Entwicklung mit Blick auf Tiere. Bei hochentwickelten Tieren müssen wir in

die Richtung von Grundrechten gehen (die moralisch verankert und rechtlich verordnet sind). Zudem müssen wir mehr Tieren Rechte zusprechen.

Marc Lewald: Welche Gründe sprechen Ihrer Meinung nach für bzw. gegen die Einführung von Roboterrechten?

(122) Expert 8: Wir dürfen den grundlegenden Graben zwischen Lebewesen und Dingen nicht zuschütten. Maschinen simulieren, und sie simulieren immer besser. Ihre Handlungen und die Ergebnisse ihrer Handlungen werden den unseren immer ähnlicher. Aber die Quelle ist eine ganz andere. Eine Gefahr ist, dass wir durch solche Verwechslungen den Tieren schaden. Sie brauchen Rechte, Maschinen nicht. Maschinen sind Dinge. Übrigens gibt es noch den Begriff der Würde. Pflanzen können Würde haben, wie auch die Bundesverfassung der Schweiz feststellt. Maschinen nicht.

Marc Lewald: Welche Risiken könnten durch die Einführung von Roboterrechten entstehen?

(123) Expert 8: Roboter können keine Rechte haben. Diese sind an Bewusstsein bzw. Empfindungs- und Leidensfähigkeit geknüpft.

Marc Lewald: Ist die Einführung von Roboterrechten Ihrer Meinung nach notwendig?

(124) Expert 8: Sie ist unsinnig. Sie hat keinerlei Grundlage. Sie ist letztlich eine Sache für Aktivisten. Bei der letzten ICSR in Florenz stürmten genau solche Aktivisten herein und forderten Rechte für Roboter. Sie haben sich noch nicht einmal Mühe gemacht, ein Paper einzureichen und auf diese Weise mit uns zu diskutieren.

## Statutory declaration

I hereby declare in lieu of an oath that I have prepared this thesis independently and without the use of other than the indicated aids; the thoughts taken directly or indirectly from external sources are marked as such. The work has not been submitted to any other examination authority in the same or a similar form and has not been published.

Karlsruhe, 30.09.2023

*Yasir b. al-D*