UNIVERZITA KARLOVA

FAKULTA SOCIÁLNÍCH VĚD

Institut sociologických studií

Katedra sociologie

Diplomová práce

2023 **Michaela Mendelová**

UNIVERZITA KARLOVA

FAKULTA SOCIÁLNÍCH VĚD

Institut sociologických studií

Katedra sociologie

Humans, artificial intelligence and sentience

Diplomová práce

Autor práce: Michaela Mendelová

Studijní program: Sociologie

Prohlášení

1. Prohlašuji, že jsem předkládanou práci zpracovala samostatně a použila jen uvedené prameny a literaturu.
2. Prohlašuji, že práce nebyla využita k získání jiného titulu.
3. Souhlasím s tím, aby práce byla zpřístupněna pro studijní a výzkumné účely.

V Praze dne 27.12.2022 Michaela Mendelová

Bibliografický záznam

Mendelová, M. (2023). *Humans, artificial intelligence and sentience*. Univerzita Karlova, Fakulta sociálních věd, Institut sociologických studií, Katedra sociologie. Vedoucí diplomové práce Mgr. Filip Vostal, Ph.D

Rozsah práce: 106 442

Abstrakt

Práce je založena na sekundární analýze dat ze studie AIMS z roku 2021, která se zaměřila na umělou inteligenci (AI). Konkrétně na její vnímání, ohleduplnost respondentů vůči ní a její sociální integraci. Mým cílem bylo použít analýzu latentních tříd k výpočtu tříd nebo skupin respondentů na základě jejich názorů na témata, jako je zahrnutí vnímající umělé inteligence a robotů, zvířat a životního prostředí do morálního rámce, udělování zákonných práv umělé inteligenci s vědomím, podpora blahobytu umělé inteligence ve formě ochrany před poškozením nebo vnímání nebezpečnosti umělé inteligence pro společnost.
Moje analýza naznačuje, že existují 3 skupiny respondentů, a tedy je možné, že tyto skupiny existují skupiny v americké společnosti. Třída 1 je velmi rozporuplná, někteří respondenti v této třídě velmi podporují AI a někteří ne. Ale celkově mají tendenci si myslet, že AI s vědomím může být nebezpečná. Tato třída je obecně nejstarší a nejméně často informovaná o AI. Třída 2 obecně podporuje blaho AI, ale tito respondenti preferují pasivní podporu. Tito respondenti si nemyslí, že vnímající AI může být nebezpečná pro ně samotné, ale myslí si, že může být nebezpečná pro budoucí lidi. Tato třída je obecně nejmladší. Třída 3 jsou respondenti, kteří měli největší šanci pracovat s AI a často konzumují média o AI. Obecně podporují blaho AI, ale v některých případech si nejsou jisti, zda považovat určité neblahé chování vůči těmto entitám za špatné. V některých případech si také myslí, že vnímající AI může být nebezpečná. Doporučila jsem také některé komunikační strategie, jak informovat tyto různé skupiny lidí o důležitosti blahobytu umělé inteligence.

Abstract

The thesis is based on the secondary analysis of data from the 2021 AIMS study. It is centered on the sentience, moral consideration, and social integration of artificial intelligence. My goal was to use a segmenting procedure (latent class analysis) to calculate classes or groups of respondents, based on their opinions on topics such as the inclusion of sentient AI and robots, animals, and the environment in the moral circle, granting legal rights to sentient AI, and support of the well-being of AI in the form of protection from harm or the perceived danger of AI for society.
The analysis offers some results suggesting that there are 3 groups within the respondents and therefore possibly some groups in the US society. Class 1 is very contradictory, some respondents in this class are very supportive of AI, and some of them are not. Yet, overall, they tend to think that sentient technology can be dangerous. This class is generally the oldest and least frequently informed about AI. Class 2 is generally supportive of the welfare of AI, but these respondents prefer passive support. These respondents do not think that sentient AI can be dangerous for them, but they think that it can be dangerous to future people. This class is generally the youngest. Class 3 are respondents who had the biggest chance to work with AI and frequently consume media about AI. They are generally supportive of the welfare of AI, but they are in some cases not sure, whether some harmful actions towards these entities are wrong. They in some cases also think that sentient AI can be dangerous. I also recommended some communication strategies to inform these various groups of people about the importance of the welfare of AI.

Klíčová slova

AI, umělá inteligence, sociologie technologií, průzkum veřejného mínění, analýza latentních tříd, cítění, roboti, sekundární analýza

Keywords

AI, artificial intelligence, sociology of technology, public opinion research, latent class analysis, sentience, robots, secondary analysis

Title/název práce

Lidé, umělá inteligence a cítění

Poděkování:

Na tomto místě bych ráda poděkovala svému vedoucímu diplomové práce panu Mgr. Filipu Vostalovi Ph.D, za konzultace, rady a laskavý přístup. Také bych ráda poděkovala své rodině a přátelům za podporu.

Content

1 Introduction

Artificial intelligence (AI) is starting to perform many roles in the society. Nowadays, AI is becoming to influence daily lives of all sorts of people, not only those who are intentionally/deliberately involved in its process and evolution (Reddy Nadikattu, 2016). These artificial entities are not only used to perform tasks, but they are also used to provide familial comfort for people. In some cases they can even serve as a sexual companion and are programmed to reciprocate affection (Chessman, 2018). However, the progression of cybernetics and artificial intelligence can soon offer rational and sentient artificial entities. Thus starting discussions about granting them legal rights in the future which leads to a number of theoretical problems (Ashrafian, 2014). Are these entities valuable on their own, or are they valuable because of their value for humans, does consciousness matter (Hildt, 2019)? Although the reasons for ensuring their welfare through legal rights can vary, literature on the ethical and moral considerations of sentient AI and robots is inclined to protect these entities legally (see for example Ashrafian, 2014; Chessman, 2018; Darling, 2012; Hildt, 2019). Nevertheless, not only expert opinions on this issue are needed. The ethical artificial intelligence development has the aim to be human centric. The welfare and legal protection of artificial intelligence as a topic that is nowadays gaining popularity in sociological public opinion research, because of these reasons. However, the existing literature on the subject is insufficient and there are only a few studies published in recent years (see e.g., Kieslich et al., 2022; Martínez & Winter, 2021a; Zhang & Dafoe, 2020). There are even less studies published that calculate groups of respondents regarding their opinion on these issues, although there is a possibility that the development of this technology can have different impacts on various types of public (Kieslich et al., 2022; Martínez & Winter, 2021a).

The aim of my research is to analyse public opinion on the support of welfare of AI in general, interpret it within the context of the existing academic research, depict possible consequences on the evolution of AI and propose several possible paths for the future research in the area of sociology of technology and related academic fields.
In the master's thesis, I aim to use a segmenting procedure called *latent class analysis* to define classes or groups of respondents, based on their opinions on topics regarding inclusion of sentient AI and robots, animals, and the environment in the moral circle, granting legal rights to sentient AI, and support of the well-being of AI in the form of protection from harm or the perceived danger of AI for society. This can be used for better understanding the public opinion on these issues as well as for implementing communication strategies offering these groups the information on these issues. It would be beneficial for these groups to understand the importance of the welfare of artificial intelligence and the objectives to protect it in the future. Thus, can assist the human centric ethical development of AI to reach a new level.
I propose a secondary analysis of data collected by the Sentient Institute. I particularly focus on the 2021 AIMS study that was carried out last year in November and December. The used data are withdrawn from a nationally representative survey of 1,232 U.S. American adults. The survey consists of 75 questions on diverse topics, such as sentience, moral consideration, or social integration of artificial intelligence.
The data can be found through the Open Science Framework, which is a free online platform

for researchers. The platform serves to plan research, collect data, analyze it and as well as share their own work (OSF, n.d.).

Reviewing the literature and existing research on the subject raised two main hypotheses, which I call H1 and H2, and five questions labelled as Q1 to Q5. Furthermore, I calculated 6 different latent class analyses of variables, each for a specific group of questions. Then I calculated one latent class analysis for the membership in those groups from the previous 6 to link these analyses with each other and get the full picture.

The results of my secondary analysis of the existing data suggest that there are possibly some groups in the US society regarding the opinion on the welfare of (sentient) AI and robots. Specifically, I calculated three groups when it comes to moral considerations regarding artificial intelligence and robots. In the case of the last LCA which is the most important, these groups are comparable in size.
These groups not only have different opinions on issues regarding the well-being of AI, but it is likely, that contributing factors to the membership in classes are age, ownership of advanced AI, and consumption of media featuring AI or robots. This is for example in alignment with the results of the research conducted by Zhang and Dafoe (2020) that suggested that age and experience with AI could be the segmenting element. Yet, I found no significant relation to gender, race and ethnicity, income, or any other socio-demographic variable.
Overall, we can assume about the public opinion on the welfare of AI and robots that although respondents usually want to protect AI from harm, they also do not think that legal protection is suitable. Also, they are generally not inclined to include sentient AI in the moral circle, which can be the result of the lack of information on these issues. Another possibility is that respondents just do not care because there are more important social issues today that could affect them personally. All groups state that they do not think that the welfare of AI is one of the most important social issues today. It is also interesting that these classes are not divided by their political preferences, which can mean that no political party in the US paid attention to this topic and therefore there was no such tendency measured.
When it comes to the recommendations I provide on the communication strategies in accordance with the characteristics of those groups, there are some key general suggestions for all of them. I generally recommend a neutral tone and I strongly suggest not promoting fear. There is a significant number of respondents in all classes that are worried about the danger sentient AI could potentially be to humans. That can impact the opinion on the rights sentient AI could receive.
According to my results, there are also some major differences in the levels of support for the welfare of animals, the environment, and sentient AI and robots. Most respondents would include animals in the moral circle and think that the welfare of animals is one of the most important social issues today, but only a small share of them would include sentient AI in the moral circle and think that the welfare of sentient AI is an important social issue. So there is a possibility that the emphasis on the perception of robots as pets (as documented in the research by Darling (2012)) and the animal rights approach towards rights for AI and robots (offered by Chessman, 2018) can theoretically help with the support of sentient artificial entities and their rights. Yet, further research is needed to prove that.

2 Theoretical literature review

2.1 Objectives of the thesis and their theoretical background

In this chapter, it is critical to investigate current AI and ethics around it when it comes to public perception of this phenomenon followed by current research on public opinion regarding artificial intelligence, which frames my hypotheses and questions for the analysis that I conducted.
The goal of my analysis is to test whether US society is divided into groups relating opinions about the welfare of artificial intelligence and to describe what are these groups like. Is age, political preference, another sociodemographic, experience with advanced AI, or the consumption of AI narratives in media the dividing factor? How big are these groups? How are these groups viewing AI? What can we assume about American society in consideration of these groups? How to communicate information about AI to each of these groups effectively? If there are any differences in the levels of support for the welfare of animals, the environment, and sentient AI and robots, can the animal rights framework theoretically help to shape public opinion on ascribing rights to sentient AI? These are my questions and hypotheses.

The overall topic is the public opinion on well-being of (sentient) artificial entities, therefore understanding the roles of these entities in society is important, as well as the level of perceived sentience of current AI, theoretical objectives for practical protection of these entities against any future harm, media coverage of these issues and some current research on this topic. Although my work tackles the issue primarily from a sociological perspective, the topic of AI and its roles extends beyond and is treated in other disciplines and domains as well. The spectrum of disciplines dealing with the issue is wide, today we can read about AI in works from the fields of philosophy, psychology, law and of course in works from the domain of information technology. In the following paragraphs, I hereby present some of the key theoretical concepts that I find crucial for us to understand the growing presence and importance of AI in our daily lives. Furthermore, the concepts presented below can also help comprehend the possible public perceptions of AI today.

Firstly, let me introduce you to the main roles artificial intelligence is currently taking on in society. Reddy Nadikattu (2016) offers some examples like chatbots, digital assistants, self-parking along and self-driving cars, but those are just a fraction of AI's use. AI has become a fundamental part of numerous industries and it is effectively changing fields such as information technology (computer programs written by AI), healthcare (predicting high-risk patients with the use of AI), marketing (digital marketing algorithms), cybersecurity (and the military in general – for example autonomous weapon systems and vehicles), and even art (art generators like DALL-E2). Chessman (2018) goes even further, he claims, that the evolution of AI promises to challenge almost every aspect of civil society, law, and economy.
Artificially intelligent entities are currently built to perform complex tasks, occupy jobs that were previously done by actual people, develop inventions in science and technology, and to some extent provide familial comfort. There are experiments with robots constructed to offer sexual companionship and programs to reciprocate affection toward a person even. Massa, Bisconti, and Nardi (2022) argue that using AI for these predominantly social and intimate purposes can have a significant psychological impact on individuals who use it. Thus, it can have an impact on society on a larger scale, depending on nature of that impact. Nonetheless, this perspective is very tech-deterministic. In other words, these authors view this evolving technology as an inevitable and dramatic driving force of change for the whole society with far-reaching impact and consequences (*Technological Determinism - Oxford Reference*, n.d.).
According to some authors (Darling, 2012; Gunkel, 2018; Hildt, 2019; Scheutz, 2011),

individuals tend to develop a unidirectional emotional bond towards social robots. They often project lifelike qualities, human characteristics, and intentions onto them. The most typical example of this is the case of the robot Sophia, which was granted Saudi Arabia citizenship in 2017. All these factors contribute, according to Hildt (2019) and Gunkel (2018), to the rising number of questions regarding the status of AI in society and how to interact with it.

The response to these questions can be in the level of perceived sentience of AI and robots. To better understand this topic, it is central to introduce the philosophical approach to artificial consciousness and sentience. The debate on artificial consciousness is nothing new in this field, according to Hildt (2019), there are two prevalent approaches – strong artificial intelligence and weak artificial intelligence. The exchange of philosophical arguments about AI between these two approaches has been ongoing for a few decades, therefore giving details on these debates is not an option for this paper. However, in short, strong AI views appropriately programmed computers as sentient. According to Searle (1980), when these computers are given the right tools, they will understand humans and they will have cognitive states. On the other hand, Hildt (2019) argues that weak AI assumes that machines do not have sentience, do not have consciousness or mind, they can only simulate thought and understanding.
Moreover, there are several theoretical problems with artificial sentience. Hildt (2019) explains that the biggest problem is explaining consciousness, especially how subjectivity can unfold from matter. This phenomenon is denoted as the "hard problem of consciousness" (Chalmers, 1996).
Also, when it comes to the understanding of specifically human consciousness, it is formed by our own phenomenal experience. In addition to that, we experience human consciousness from the first-person perspective, but that is not the case for artificial consciousness, which is only accessible to us from the third-person perspective. How do we know, whether a machine is sentient, when we are not experiencing it, all definitions would be made without relying on phenomenal consciousness, because only a third-person perspective is possible.

However, there are some philosophical reflections around consciousness we can rely on when answering questions about artificial consciousness, although it focuses primarily on human (and animal) consciousness. Usually, according to several authors (Gennaro, n.d.; Hildt, 2019; Van Gulick, 2022), we distinguish between concepts of a conscious entity (an entity that is self-conscious, sentient, and wakeful, it has its own subjective qualitative experiences); to be aware of something (for example an apple on the table); and conscious mental states (mental states an entity is conscious of being in, such as having the opportunity to taste the apple). In addition to that Seth (2016), argues, that these problems are in fact completely valid to evaluate, but the "real problem" of consciousness is often forgotten - the challenge to address the reason why various artificial (mechanical or computational) and biophysical phenomena coincide (or not) when exposed to different qualities of experience.
According to Holland and Gamez (2009), conscious artificial intelligence is in actuality needed to further understand the nature of consciousness, autonomy, and subjectivity of entities.
Surprisingly enough, we are not that far from machines experiencing consciousness. In a recent article, named *Artificial stimulus-response system capable of conscious response* (Kim et al., 2021) the authors present a new system of artificial consciousness that is designed to provide people with neurological problems physical experiences, for example a touch of anything provided by a bionic hand. It is based on an artificial stimulus-response system that

was constructed to emulate human conscious response. However, this truly remarkable invention is still working with an existing conscious human mind, it is not sentient on its own, but in the context of some recent theories that have the potential to solve the hard problem of consciousness, it certainly can help to make a machine conscious. One such framework was last year presented by Safron (2021), where he integrated six existing theoretical frameworks, such as global workspace theory, integrated world modeling theory, or hierarchical predictive processing into a coherent concept binding everything from biophysics to agency. Nonetheless, this author integrated these frameworks theoretically, and I found no further research that would put it into practice.

Because of the possibility of sentient AI or robots in the near future, there are some tendencies to ensure some level of protection of these entities from an ethical perspective. There is already a plan to develop a platform to prevent *artificial suffering* proposed by Metzinger (2021). The author states that there should be a global moratorium on synthetic phenomenology with a strict ban on research that can cause distress to sentient AI. This moratorium would be flexible when it comes to the set of constraints it would promote, because of the number of possible scenarios in the future, but the demand for it is to be evidence-based, rational, and as ethical as possible. The reason for the development of this, mostly political platform, is to prevent any form of artificial suffering globally because today there is no representation of these future self-conscious artificial beings in the political processes of any country. It also seems theoretically possible that when these entities become conscious, they will have some political preferences themselves. It is unlikely that sentient being would not autonomously formulate their own hierarchy of goals and they will need a platform to be able to express them and even enforce them. There is a great possibility that these entities can suffer in ways we could not imagine, comprehend, or unable to discover completely, although people would be directly responsible for it. According to this author, the people that could be responsible for it in the future (AI researchers, mathematicians, neuroscientists, policymakers, or legal regulators) are possibly alive today and therefore it is important to establish this moratorium early on. Those risks of artificial suffering are incalculable today, but there is an emergency of it present with more inventions that are getting closer to the goal of artificial sentience. The need for an evidence-based, strictly rational process of assessment of this problem is also emerging.

Although, there is a general agreement among scientists that current robots are not sentient, several authors suggest ascribing rights to robots (for example Coeckelbergh, 2010; Darling, 2012; Gunkel, 2018; Hildt, 2019). The exact argument for it differs, but generally, it is not centered on the AI's or robot's capabilities, but more on the roles of robots and AIs ascribed to them by human beings and the social context they interact with each other. Therefore, the consciousness of these artificial entities does not matter in ascribing legal rights to them. However, according to Hildt (2019), there is a problem with this approach, when applied to humans, it suggests that the value and rights of human beings is strongly dependent on their social roles and the interests of others. It is in a complete contradiction of the generally recognized view (for example by Jaworska and Tannenbaum (2021)) that the moral status of human beings is independent of their social roles, so current robots should matter morally for their own sake.

To further understand the perception of AI and robots by humans, I present a psychological research study. Darling (2016) analyzed violent behavior towards robots and suggests that

when it comes to social values, humans tend to treat robots more like their pets than just things, especially social robots with human-like features. Humans form a unidirectional emotional bond with these entities and such a bond can be a justifiable reason to protect robots and AIs legally.
This suggests an applicable framework for laws that could be ascribed to sentient AI, which is also recommended by Chessman (2018). It is a plausible possibility because at the end of the day, nonhuman animals are a form of sentient property and there are four centuries of development of domestic animal rights law. It offers some ways and insights into the evaluation of legal values that can be applied to the case of AI. That can mean protection of AI as a property (allocation of responsibility for damage done by AI included), protection of emotional attachments of humans towards artificial entities, and limitation of extreme antisocial behavior involving AI (that can be not only realistic artificial or physical torture of AI, but even simulations of child molestation or hate crimes). The reasons for it could be for example the psychological consequences for people or protecting AI because it is intrinsically valuable. Also, regulatory regimes are nowadays in place for animals, so these laws can be quickly implemented into existing and functioning structures. Therefore, I included in my analysis four variables on the topic of the welfare of animals and the environment. What are the levels of support for the welfare of animals, the environment, and sentient AI and robots? If there are any differences, can the animal rights framework theoretically help to shape public opinion on ascribing rights to sentient AI?

2.2 Current research on ethical considerations of AI in mass media and public opinion on them

Mass media can have an impact on public opinion (for example Mccombs, 2008; Mutz & Soss, 1997). Therefore, it can be useful to examine how the mass media handle those previously mentioned issues, such as the roles of AI in our society, granting rights to AI and more. When it comes to media coverage of these issues. According to Ouchchy, Coin, and Dubljevic (2020), media coverage follows the increasingly prominent presence of AI in our daily lives. The authors analyzed ethical considerations of these technologies in media. The problematic aspect of media coverage in news and discussions about AI is that the academic research and following debates have not yet converted their results to formats which would be suitable to the general public. On the contrary, they tend to separate from highly popularized events increasing recognition of the AI technologies (for example Tesla autopilot accidents) and moreover, from public concern about the yet unknown. Usually, the articles these authors analyzed started with clearly identified issues regarding AI and they often contain constructive recommendations, but academic ethical theories are very rarely used. Although the focus of media covering this ethical issue is fairly realistic and practical, the coverage is still very shallow. The authors then argue that multi-faceted approach and better accessibility of correct information for the public are needed and the importance of it will only increase in the future. I agree with them, but maybe they just expect too much from the mass media.
The research (Ouchchy et al., 2020) also showed that the media portrayal of ethical issues with AI is written in critical or balanced tones nowadays. The vast majority of articles analyzed on personal issues such as job loss to AI were neutral in their tone. The crucial aspect of these articles is the underlying message that it is important to harness potential benefits and mitigate the negative effects of these technologies on society, preferably to do so at the same time. The number of articles about AI is increasing throughout recent years and

the topic of regulation of AI is used more frequently than ever, so according to these authors, there is a possibility that this topic is in some form gaining importance in society. My question is whether the consumption of such media content (specifically the frequency) has any impact on the support of the well-being of sentient artificial entities.

Zhang and Dafoe (2020) argue a critical reason, why public opinion on ethics regarding AI is important. Recognizing the policies on AI governance the public prefers is crucial to establish a productive public policy deliberation. However, these are not the only reasons why public opinion is needed, other equally important are according to Kelley, Heldreth, Moessner, Sedley, Kramm, Newman, and Woodruff (2021) commercial development, product adoption, and research funding. According to Kieslich, Keller, and Starke (2022) despite the theoretical and practical societal importance of ethics regarding AI and the certain urgency of this issue, only a few public opinion pieces of research were conducted on this topic. Although, ironically, ethical AI development aims to benefit the whole society and be as human-centric as possible. However, it is not an easy task, especially when taking the complexity of AI systems into account. These models tend to be almost impossible to explain to anybody who is not directly involved in the process of AI research. In many cases, simplification is needed, which violates the ethical principle of accuracy. In all cases, ethical AI development requires taking the opinion of the public into account, because their lives could be affected by AI in the future.
There are two main concepts of human-centric AI. The first was introduced by Riedl (2019). Ethical AI development can be human centered only when on the input side it considers the sociocultural complexity of humans and on the output side, the explanation of AI that is easy to understand is provided. The other concept was provided by the European Commission (2019), according to their definition, human-centric ethical AI development means aiming for the common good and humanity to provide improvements in human welfare and freedom. Still, understanding the needs and perceptions of the general public is necessary.

There are a few surveys on the topic of this paper worth mentioning.
Zhang and Dafoe (2020) conducted an online study in 2018 with a random sample of 2000 American respondents on attitudes toward AI and trust in tech companies. This study, although the topic is in many aspects different than the topic of this paper, is relevant, because there are only a few researches I found that calculated groups of respondents according to their attitudes toward AI.
These researchers (Zhang & Dafoe, 2020) calculated groups of respondents in accordance with their perceived importance to resolve each AI governance challenges such as data privacy, digital manipulation, and AI-enhanced cyber-attacks. The responses were relatively comparable across the demographic criteria with a few exceptions – age and experience with programming. Compared to younger respondents, older ones considered almost all AI governance problems to be very important, also they were on average more likely to oppose the idea of AI development. The other category (respondents with programming experience) was practically the polar opposite. They are on average very supportive of AI development and are less concerned about AI governance challenges than the average respondent. Although, respondents were relatively certain that AI should be regulated in some form, they were not sure, who the regulators should be.
These results lead me to a hypothesis, whether age and experience with advanced AI have any connection to support of the welfare of sentient AI, specifically whether these factors will play a role in the membership in any of the groups I will calculate.

Kieslich, Keller, and Starke (2022) conducted a study on the public perception of ethical principles regarding AI in Germany. The results showed that each of the segments they calculated (5 in total) prioritized different ethical principles, there is probably not a universal understanding of the importance of each of them, but the age of the respondent and previous experience with AI were key in the segmentation procedure. AI has the potential to influence the lives of people, regardless of their age and previous experiences with such technology. Therefore, these researchers recommend identifying the public for each deployment of AI and proceeding individually following the preferences and expectations of the specific public. There would be different demands and ethical considerations for university students in the case of algorithmic admission systems in universities, and job seeker categorization systems for unemployed workers. These researchers identified approaches for communication about AI use for each of those classes, which would be also useful in the case of US society, thus I will try to offer it as well in the result section of this paper to provide some practical use of my analysis.

There are a few surveys on the topic of granting rights to AI, but none of them is using any segmenting procedure I am conducting. One of them is a research done by Lima, Kim, Ryu, Jeon, and Cha (2020). These researchers measured opinions on 11 rights that could be possibly granted to artificial intelligence and other autonomous entities of this kind. Respondents in most cases disfavored the very idea of AI or robot rights, but surprisingly they are supportive of the protection of these entities from cruel treatment, although a legal right to protect them does not seem to be a reasonable alternative for them. I am evaluating this statement in the results of my analysis.
In addition to that, this research (Lima et al., 2020) showed that when given information about the rights-bearing of nonhuman entities, their perception became more positive. Nevertheless, respondents were more supportive of these rights when these entities were portrayed as fully autonomous. Politically liberal or moderate participants were generally more likely to be supportive than conservatives and young respondents (under 35) as well. This further solidifies my hypothesis regarding demographic criteria.

Another research that is purposely focused on the extent to which should the law protect sentient artificial entities the public is comfortable with. Martínez and Winter (2021b) surveyed American adults on topics of general legal protection and the legal personhood of AI.
Results showed that from the respondent's point of view sentient AI had the lowest perceived level of legal protection nowadays. However, surprisingly when asked about the desired level of legal protection, respondents would on average grant more legal protection to sentient AI than to corporations that scored highest on the level of perceived level of protection, but lowest on the desired level of protection just below artificial entities. Approximately one-third of respondents were positive about granting legal personhood to sentient artificial entities. Politically liberal respondents were significantly more inclined to legally protect sentient AI and grant them legal personhood.

3 Methodology

3.1 Goals, hypotheses, and questions

The general idea of this research is to test the proposed hypotheses using a complex dataset to be able to provide answers in consideration for the procedure I choose and a public opinion survey that is crucial to determine the level of support for the well-being of sentient artificial entities provided by US society.

The main goal is to test whether US society is divided into groups defined by opinions related to the problematics of the welfare of artificial intelligence and how the groups behave and look like. I propose to test two hypotheses and answer 5 suggested questions that I present below.

H1: Is US society divided into groups relating opinions about artificial intelligence?
H2: Is age, political preference, another sociodemographic, or experience with AI the dividing factor?

Q1: How big are these groups?
Q2: How are these groups viewing AI as a whole?
Q3: What can we assume about American society in consideration of these groups?
Q4: How to communicate information about AI to each of those classes (in the last LCA) effectively? Q5: If there are any differences in the levels of support for the welfare of animals, the environment, and sentient AI and robots, can the animal rights framework theoretically help to shape public opinion on ascribing rights to sentient AI?

There is a gap in research when it comes to public opinion on the ethics of AI development, the legal rights of AI, and the overall well-being of sentient AI, there is no research that included all these topics with any segmenting procedure (as previously indicated in the previous section of this paper). Therefore, I am dividing respondents into groups by their responses to questions on legal rights, the welfare of sentient AI, perception of the level of harm artificial entities could be part of, and the inclusion to the moral circle when it comes to artificial entities and hopefully offer some communication strategies to help with informing the US public about AI and its' possible needs.

I had to plan the research part of this paper in consideration of all these factors. The best way to be able to answer my questions was to look for existing data on this topic because I as a student would not be able to conduct research of this volume on my own. The offers for student research I had were not suitable, for example, I could get a maximum of 800 respondents, which is a relatively small sample of data for the segmenting procedure I choose (there is a risk of calculating a group with too few respondents), and a very limited number of questions in the survey, the maximum was 20 with sociodemographic variables included. The results of this survey would be very biased and that's not what I want to provide. Therefore I decided to use existing data on this topic from the Sentience Institute, specifically the 2021 Artificial Intelligence, Morality, and Sentience (AIMS) Survey (Pauketat, 2022b). It is a nationally representative survey of 1,232 U.S. American adults with 75 questions on topics like sentience, moral consideration, or social integration of artificial intelligence, which is more suitable than the student research opportunities I had.

3.2 2021 AIMS survey

According to Pauketat (2022b) and the information Sentience Institute provided on their website, the AIMS survey was conducted in November and December 2021. The objective of

this longitudinal study is to track how the public's opinion on artificial intelligence changes over time. Another objective of that survey is to test some predictions made by forecasters on this topic. This survey was done through the software GuidedTrack (a professional software used by social scientists to gather survey data online (*GuidedTrack*, n.d.)) and run online with a sample of respondents enlisted by Ipsos (professional market research company, member of ESOMAR (*Key Figures | Ipsos*, n.d.)), where the sociodemographic variables were based on census estimates from the 2019 American Community Survey. All materials for the survey are available on the Open Science Framework.
Most of the analyses these researchers made were based on correlations and frequency tables, so they can easily compare this survey they conducted with the upcoming one (this will be a longitudinal study) and at the same time make estimates forecasters gave them easily comparable as well.
The results of this survey show that the American respondents are more uncertain about the possibility of artificial sentience (41.53%) than believe that artificial sentience is possible (34.82%) or not (23.65%). Some respondents even think that artificial sentience exists now (18.06%). Altogether respondents predict a 59.97% chance that artificial intelligence and robots will be sentient within 100 years. Before the data publication the researchers conducted exploratory predictions made by Metaculus (online forecasting platform (*About | Metaculus*, n.d.)) forecasters, where they tried to predict public opinion. The task was to predict the respondents' answers to 5 questions used in the survey. The Sentience Institute made its predictions beforehand to make sure that they are unbiased by the results. Overall, these predictions generally underestimated the consideration of the moral and social aspects of artificial sentience the respondents provide. Metaculus forecasters expected that the general public would be less certain about the possibility of artificial sentience. They underestimated the results of four questions: the support for a ban on the development of sentient AI (estimated 25%, in reality, 58%), the support for some legal rights assigned to sentient AI (estimated 14%, in reality, 37%), the importance of the welfare of this AI as a social issue (estimated 3%, in reality, 30%), and lastly the possibility of sentient AI being a threat to any future people (estimated 65%, in reality, 69%). On the other hand, the Metaculus forecasters made one overestimation, they thought that the public's belief in the possibility of artificial sentience would be higher (estimated at 41%, in reality, 35%). Estimations made by the Sentience institute were slightly closer to the actual results of the survey, but when it comes to the overestimation and underestimation of public opinion, they were in the same direction as Metaculus forecasters. Yet, the format was slightly different, therefore the direct comparison is not a possibility. The overall results of this survey will be further described in more detail in the part of this paper that is focused on the results of my analysis to provide more context.

3.3 Consideration of possible methods

The first step is to determine, what procedure is suitable for my analysis. After a thoughtful consideration, I decided to use latent class analysis (in short LCA). To give some clarity on my decision, I present some ideas from the paper Clustering in the field of social sciences: that is your choice by Jaime Fonseca (2012). This paper provides a comprehensive comparison of various clustering methods, including cluster analysis and latent class analysis. I originally considered both methods, both are suitable for my analysis in some form, therefore I decided to provide significant details about both of them.
According to Fonseca, any clustering attempts to find a finite group of clusters to describe

(survey) data. According to Gentle et al. (1991) cluster analysis seeks to separate seemingly similar objects into meaningful classes, both the composition and number of classes need to be determined beforehand. Fonseca (2012) then describes that recently we can see a rise in the use of latent class analysis in the field of clustering methods. However, both methods are useful probabilistic or statistical methods for grouping observations into clusters.
According to Collins and Lanza (2010), these statistical procedures are generally very similar. Both are considered to be "person-oriented analyses", which seek to identify individuals and group them. On the other hand, Weller, Bowen, and Faubert (2020a) compare them with variable-centered approaches, which are centered on relationships among variables, but not individual respondents. Moreover, in cluster analysis and latent class analysis, the exploration phase of research is similar, common practice is to conduct a series of solutions, each one with one additional class than the previous one. Researchers then use appropriate statistical and theoretical criteria to determine, which solution is fitting their research the best.
Fonseca (2012) also performed several comparable analyses to determine, which method is generally more useful for what purposes, in all cases the latent class model performed reasonably well. Yet, at the same time, cluster analysis performed both the best (90.7% accuracy on only continuous variables) and the worst (40% accuracy when it comes to mixed variables). I used categorical variables in my analysis, therefore LCA is slightly better for me to use.

Fonseca (2012) also describes the cases, benefits, and limitations of these two methods in great detail. According to this author, the purpose of cluster analysis is to discover mutually exclusive groups, clusters, or types of objects that the objects belonging to the same group are as similar as they can be and objects belonging to different clusters are as different from each other group as they can be (Fonseca, 2012; Hagenaars & Halman, 1989). Also, cluster analysis can be used to distinguish between "naturally built social groups" (for example a friend group), but there are some unfortunate theoretical problems naturally tied to it (Bailey, 1983). Also, the biggest problem of this method is the number of clusters to be retained, but that can be at least partially solved by replication and cross-validation procedure (Mandara, 2003).
Latent class analysis is used for building typologies or clusters based on observed dichotomous variables, which is conceptually similar to cluster analysis. The main difference is that it can identify latent classes based on observed response patterns with a presumption of the latent dimension responsible for it (for example Clogg, 1995; Fonseca, 2012; Lazarsfeld & Henry, 1968; McCutcheon, 1987). In other words, according to a group of scientists Weller, Bowen, and Faubert (2020), latent Class Analysis (LCA) is a statistical method usually used to identify subgroups within populations that differ qualitatively. These subgroups frequently share certain outward characteristics. The idea of latent class analysis is that membership in unobserved groups, subgroups, or classes can be determined by observing patterns of scores across responses to survey questions. Also, according to Vermunt and Magidson (2016) when it comes to these patterns, some parameters of the presumed (by LCA) statistical model differ across these unobserved subgroups. They then frame the categories of a categorical latent variable. This method has several applications because the whole idea of this procedure is universal to a certain point. The most important applications, although seemingly unrelated, are clustering, density estimation, scaling, and random-effects modelling.
LCA model can be even specified according to a number of authors such as Haberman (1979)

and Vermunt and Magidson (2016) as a log-linear model for a table with missing cell entries. Yet, a more precise description is a model, where the latent variable is included as an additional dimension, expanding the original table. I use contingency tables to provide further details on the existing results from the AIMS study for four different batteries of questions on topics such as granting legal rights to sentient artificial intelligence; inclusion of these entities, animals, and the environment to the moral circle; support of protection of these entities from harm, etc. In other words, the spectrum of topics I analyzed is broad, therefore the possibility of the presence of latent dimension is high, combined with the use of contingency tables the latent class analysis is more suitable in my case.

Fonseca (2012) also lists some key differences between cluster analysis and latent class analysis and he then recommends the use of latent class analysis because of them. According to other authors as well (Magidson & Vermunt, 2002), the most important difference when it comes to cluster analysis and latent class clustering is that the latter is model-based, which means that a statistical (and probabilistic) model is assumed for the population the data sample is acquired from. The benefit of using such a model is that the chosen cluster criterion is not as arbitrary, and the approach includes thorough statistical tests.

The second benefit is that standardization of variables is not necessary for latent class analysis. For example, in the case of hierarchical clustering (one of the types of cluster analysis) standardization of variables is crucial for them to have an equal variance to avoid clusters, where variables with the biggest variance prevail. Nonetheless, this solution does not completely solve the problem prevalently associated with scale differences. The clusters are practically unknown and so it is impossible to perform standardization within those clusters. On the other hand, the latent class clustering solution is not based on linear transformations on the variables, so such standardization of variables is not needed.

According to these authors (Fonseca, 2012; Magidson & Vermunt, 2002), there is also a difference when it comes to types of variables. It is possible to include mixed scale variables in latent class analysis, which generally can be continuous, any categorical (nominal, ordinal), or accounts of any combination of these types. Yet, hierarchical clustering is limited to only using interval scale variables. Collins and Lanza (2010) offer an even more precise explanation of it. Cluster analysis uses variable means to determine the "nearness" of cases when it comes to the patterns in survey responses. Thus, it is logical that the variables used in the analysis should be continuous. In contrast, the latent class analysis uses cross-tabulations as the input source, therefore the variables can be categorical.

According to several authors (Dillon & Kumar, 1994; Fonseca, 2012; McLachlan & Peel, 2000), another difference is in the number of clusters determination. There is no reliable method to determine the number of clusters when it comes to hierarchical clustering. In contrast, latent class clustering provides several diagnostics to do so, such as theoretical information criteria or likelihood ratio test, which is a big advantage. Another important detail to add to this point is from the work of Weller et al. (2020), in cluster analysis the case membership in clusters is determined. Nonetheless, in LCA there are assigned probabilities of class membership, which is very useful for the interpretation of results.

It is also beneficial that the inclusion of demographics or any other variables is much simpler in the latent class analysis, according to authors Vermunt and Magidson (2002) and Fonseca and Cardoso (2007). The simplest way to do such a thing in hierarchical clustering is to use discriminant analysis to be able to describe differences between the clusters on any number of exogenous variables. However, the latent class model can be simply extended to include any

other variables. This allows both the classification and description of clusters to be simultaneously completed using a single uniform algorithm of maximum likelihood estimation.

Because of all these reasons I choose LCA. The variables I used are categorical, there was no need to standardize them. The spectrum of questions asked in the AIMS study is broad, therefore the presumption of a latent variable being present had to be considered. Also, in this procedure, information criteria can be used to deliver a more reliable determination of the suitable number of clusters. Subsequently, I had to make two of these analyses for some batteries of questions I analyzed, because the computational power I had was not enough. Because of the nature of LCA, these results are much more comparable and interpretable in general.

3.4 Performed methods

3.4.1 Contingency tables

I computed contingency tables to provide some further detail on the original results of the AIMS (Pauketat, 2022a) study that used correlations. Therefore, it is necessary to include both methods in this chapter and compare them to show their differences.
Contingency tables are an old statistical procedure, the most likely primary source that mentions them was written by a famous statistician Karl Pearson in 1904, where he introduces the concept of contingency. Generally, the conception of contingency is that it is possible to classify individuals into two and more groups. It is then possible through the mean contingency or the mean square contingency to evaluate the extent to which two such systems are contingent or non-contingent. Also, Pearson (1904) tested the relation of contingency to a procedure he calls “normal correlation”. He found out that with normal distribution present in both cases of contingency coefficients and the frequency of groups is distributed evenly, it is possible to obtain correlation coefficient, but for an actual practice too many groups are not recommended. Contingency is independent of the order, scale, and overall arrangement of groups.
It is apparent that these methods are not the same but are comparable. Because of the dependency of correlation on normal distribution, and the use of crosstabulations in the latent class analysis as the input source, I will rely on the results from contingency tables and not the original research.

However, this method is not as important as latent class analysis for this paper and trying to summarize every aspect of them that evolved in the last 100 years would be interesting, but not very practical. Therefore, I use a short introduction to them written in a textbook by Jan Hendl (2006). Statistical analysis of frequency tables that unfolds with description and analysis of the relationship among categorical variables is like the correlation analysis of continuous variables and similar to the analysis of variance. The difference lies in the fact that in the case of the analysis of frequency tables we consider both categorical variables independent of each other. However, in the analysis of variance, the key idea is to evaluate the influence of a factor as an independent variable on the behavior of a random dependent variable that has a categorical character.
When it comes to two-dimensional tables, which are tables created by sorting according to two variables, in statistics, we call such tables contingency tables. We assume that each individual or experimental unit of the population O can be classified according to two

variables (criteria) X and Y. Variable X has categories and variable Y has b categories. After the assembly of the contingency table, it is possible to examine the relationship between the two variables X and Y – first. It is important to use a suitable representation, later different hypotheses can be tested. Hypotheses for contingency tables are usually defined in terms of the stochastic independence of X and Y.

3.4.2 Latent Class Analysis

To add some information about latent class analysis that was not mentioned earlier, this part is mainly about the practical use of this procedure.

LCA was firstly introduced by Paul Felix Lazarsfeld (1950) and it has been evolving ever since. According to Weller et. al (2020) scholars have discussed several issues alongside the evolution of LCA: 1) the selection of indicator variables 2) the selection of the final model, 3) the decision on how to incorporate covariates and 4) the selection of statistics to report. However, Weller et. al (2020) offer some suggestions when it comes to these issues, following numerous systematic reviews (for example Killian et al., 2019; Petersen et al., 2019; Ulbricht et al., 2018). To achieve an unrestricted LC analysis, at least three indicators are needed, but it is important to keep in mind the nature of those indicators. For example, when they are dichotomous, it is practically impossible to want or get more than two latent classes. In the case of four dichotomous variables, this three-class model has a positive number of degrees of freedom, but it is not unrestricted, because the indicators are dichotomous. Also, it is possible to identify even a five-class model with five dichotomous indicators, which can lead to misleading results as well. Generally, the identification of classes is possible to obtain by restraining specific model parameters. The minimum of variables I test is 4 and the scale has 7 points, therefore the ideal number of classes should be less than four, but in the less ideal scenarios it must be less than 7.

Another spectrum of problems with the assessment LCA models is the local maxima or boundary solutions presence. Boundary solutions are probabilities equal to one or zero or log-linear parameters equal to plus or minus infinity. A very popular method to overcome this problem is to use information criteria, for example, AIC or BIC.

According to Weller et. al (2020a) sample size is a debated aspect in the LCA literature as the method is progressing. These papers usually offer a simple suggestion, the bigger the sample size, the better, but it depends. For example, Nylund-Gibson and Choi (2018) suggest that the minimal number of cases should be 300 and above. Yet, for simpler models (fewer indicators and classes) or well-separated classes even smaller samples are suitable. However, a smaller sample size can lead to a few potential problems such as convergence failures, poor functioning fit indices, or problems uncovering smaller classes. With over 1200 cases I can be sure that the method will perform well on my data, which could not be the case with the student research possibility that was offered to me.

Wurpts and Geiser (2014) suggest the same when it comes to indicator variables included in the LCA model, usually more of them lead to a better result. Another suggestion by Weller et. al. (2020) for indicator variables is having a strong theoretical basis for using selected indicator variables. It helps to identify the classes, interpretation of results is easier and the application of results to practice is almost effortless. The 2 sets including only 4 variables, for which I did LCA are similar in their topics, therefore it is theoretically justifiable to use LCA on them.

Weller et al. (2020) also present the optimal strategy when it comes to conducting latent class analysis. The standard practice is to run a sequence of models, beginning with a one-class model and then adding one additional class at a time, until the best model is determined based on statistical criteria. Generally, the quality of a model is better with each additional class until an optimal solution is reached, then the quality starts to deteriorate. However, the optimal solution must be interpretable, a solution with superior statistics is not applicable when it makes zero sense theoretically.
A number of authors suggest (Nylund et al., 2007; J. K. Vermunt, 2002) the best criteria for comparing latent class analysis solutions, multiple statistics should be used, where the Bayesian information criterion (BIC) is probably the most reliable and according to a different group of authors (Lower BICs indicate better fit) (Muthén & Muthén, 2000; Nylund et al., 2007; Shanahan et al., 2013; Weller et al., 2020) theoretical interpretability of them should be considered. According to Weller et. al (Weller et al., 2020) and Celeux and Soromenho (1996) entropy is another metric that should be evaluated. Entropy indicates how accurate is the model in defying classes. Usually, a standardized entropy value around 1 is ideal, but a value above 0.8 is acceptable.

3.5 Software use

I used SPSS by IBM (IBM, 2019) for contingency tables and Jamovi (*The Jamovi Project*, 2021) for latent class analysis, specifically the snowRMM module (Seol, 2022).

4 Analysis and results

4.1 2021 AIMS results

The results of the AIMS study (Pauketat, 2022b), from which I gathered the data for my analysis were briefly discussed in the methodology part of this paper, mostly the results that do not affect the secondary analysis I provide.
To segment the respondents into classes I used a few different sets of variables from the aforementioned dataset and further divided them because I lacked the computational power to calculate entropy for more than 6 variable.

The first set consists of variables PMC1 to PMC12. This set is dedicated to topics such as the ethics of development of sentient AI and welfare standards and legal rights regarding AI.
The second set MCE1 to MCE9 focuses on attitudes towards robots. It includes topics such as the torture of sentient robots, the inclusion of such entities into the moral circle, or the protection of emotional attachments of humans towards AI.
The third set of variables SI1 to SI4 consists of questions on the perceived level of harmfulness of AI and robots.
The last two sets I analyzed as one are MCA1 to MCA2 that are centered on animal welfare and MCEn1 to MCEn2 that are centered on the welfare of the environment.
Other variables I used were of sociodemographic nature or they further characterized the respondent. I used age, gender, political views, religion, race and ethnicity, education, income, region, and experience with AI or robots – whether the respondent works with them, interacts with them outside of work, owns them, or consumes media about them.

The general results, as presented by the Sentience Institute (Pauketat, 2022a) show that although respondents agree that artificial entities deserve to be protected from deliberate harm, such as from people who would intentionally harm them physically or mentally (82% of respondents agreed), from retaliatory punishment (76% of respondents agreed) or non-consensual physical damage (68% of respondents agreed). Roughly 75% of respondents agree that sentient beings of artificial nature should be treated with respect. Only 49% of respondents would include such entities in the moral circle and as few as 37% of them would grant legal rights to them. This corresponds with results from research done by Lima, Kim, Ryu, Jeon, and Cha (2020).
However, this research offers more detailed perspectives. The vast majority of respondents believe that AI should be subservient to humans (80%). Most respondents support a ban on the development of some artificial entities, they are mostly opposing the idea of robot-human hybrids (65%) and AI-enhanced humans (63%). However, 58% of respondents are even against the development of sentient AI. Contrary to those results, the majority of respondents are inclined to support the development of certain standards to protect the well-being of sentient AI (59%), but they do not think the welfare of such beings is an important social issue nowadays (only 30% of respondents do).
Most respondents view AI as a potential risk, 65% of them agree that AI might be harmful to people in the US and 69% of them view them as potentially harmful to future generations of humans.

When it comes to the correlations the authors (Pauketat, 2022a) conducted, several demographic characteristics could predict ethical considerations regarding AI, the main ones are a vegan diet and exposure to media including AI. The second consistent predictors are age and gender. Also, race and ethnicity, region, religion, education, political preference, and income are predictors of some outcomes too.
The increased perceived danger of sentient AI development is linked to older age, being conservative, being female, and being religious. Yet, when it comes to the dangers for artificial intelligence, the opposite is not true. Although, there is a linkage between younger age and being liberal, also being female, Hispanic, vegan, and less educated with exposure to AI narratives are predictors for concern about the treatment of sentient artificial entities. Increased moral concern for these entities is present in cases of respondents living in the south, being Hispanic, being religious, having lower income, and consuming media content about AI. On the other hand, ones supporting the protection of sentient AI from malevolent actions are more likely female, are less educated, have a higher income, are vegan, liberal, and are more exposed to robot or AI narratives. Also, respondents who believe that nonhuman animals and the environment should be included in the moral circle also tend to show more moral consideration for sentient AI, which corresponds with the results of research done by Martínez and Winter (2021b).
Respondents who would advocate on behalf of the sentient AI (in a form of for example activism) are more likely to be younger, living in southern states, being liberal and being vegan or pescatarian, and being more exposed to media featuring AI.
Increased perception of current AI as sentient or having a mind is showing in younger respondents, Black or Hispanic, or vegan respondents, and those who are showing interest in robot or AI narratives.

4.2 Contingency tables

These results are interesting on their own, but for the sake of my analysis, I conducted contingency tables that show the opinion of respondents in a more detailed manner. Contingency tables are in some cases used in literature featuring latent class analysis (for example in EVANS et al., 1989; Fienberg et al., 2007; González & Sánchez, 2009). Also, that this is not the primary analysis of this paper, which is the reason, why I am not describing them in detail. Nonetheless, if you are interested in these tables, they are in the attached files.

My results show that the strongest predictors of ethical and moral considerations of AI are age, frequency of interaction with advanced AI, frequency of consumption of AI-related media, ownership of advanced AI, and working with advanced AI. Age, frequency of consumption AI related media, and the ownership of advanced AI were also the only variables that are relevant for the classes of the last and most important LCA I conducted. Older respondents are generally less supportive of the overall protection of sentient artificial entities in forms such as welfare standards and legal rights. They are less likely to include these entities in the moral circle and respect them. They also generally do not believe that AI or robots will ever be sentient.
The frequency of interaction with advanced AI showed an interesting trend, which I will demonstrate in the example of granting legal rights. The distribution of responses shows that the more the respondent interacts with AI, the bigger the tendency to grant it legal rights (only 16% of those who do not interact with AI, but 63% of those who interact with it on monthly bases are supportive of it), but this tendency changes in case of those who interact with AI more frequently than monthly. At this point the tendency continues in reverse (those who interact with these entities weekly are supportive of AI rights in 61% of cases, however, the same is true for only 48% of those, who interact with AI daily). In other cases (responses to questions), the same trend is present, but this question is a little more polarizing, so the tendency is more apparent. It can theoretically be a case of desensitization towards these entities, which is also mentioned by Darling (2012) as a reason to protect these entities by law. In addition to that, this theoretical possibility can be backed up by numerous studies on empathy in health care (for example Williams et al., 2001) and vet (for example Pollard-Williams et al., 2014), results of these researches showed that the longer the professional (or even a student) worked with patients, the lower went their empathy scores towards those patients, regardless of them being humans or animals, which is exactly in line with the argument of Darling (2012) and there is a possibility that those who interact with it more than monthly are working with it. However, this can be also linked with the belief that current AI is sentient, while only 7% of those who do not interact with AI thought so, it was the case for 43% of those who interact with it on monthly basis and for 30% of those who interact with it daily have the same opinion. Although, the more these respondents interact with advanced AI, the more they are likely to believe that artificial sentience will exist in the future.
The frequency of consumption of AI-related media probably has some effect on the perception of AI. Basically, the bigger the frequency of consumption of such media, the more these respondents agree to support AI, protect AI, and care about the well-being of AI. Generally, the results suggest a connection, based on the values of gamma that were ranging from 0.21 (support of the development of welfare standards that protect the well-being of sentient robots/AIs) to 0.4 (consideration for joining a public demonstration against the mistreatment of sentient robots/AIs), the mean value is 0.31. Which can theoretically be the outcome of how popular the theme of ethics is in science fiction literature, for example, novels such as the Foundation series (Asimov, 1942-1993) or Do Androids Dream of Electric

Sheep? (Dick, 1968) are popular for their approach to human-robot ethics. This was also documented by El Mesbahi (2015), who conducted a survey on whether and how science fiction influences humans' attitudes toward robots, and the results suggest that approximately 78% of respondents who consume science fiction media were at least somewhat affected by it, especially when it comes to the ethical questions it raises.
Owners of advanced AI are generally a lot more supportive of the welfare of sentient artificial entities than those who do not own such a piece of technology, owners were much more likely to for example support legal rights for these entities, respect them and or protect them from harm. On the other hand, owners of advanced AI have also a slightly higher chance to believe that robots and AI may be harmful to people in the USA.
Respondents working with advanced AI are also more likely to support the well-being of sentient AI and robots. For example, they are more likely to support these entities by granting them legal rights, consider joining a public demonstration on behalf of sentient AI, and also include them in the moral circle. These respondents are also very likely to think that the welfare of sentient AI is one of the most important social issues. Which is contrary to my theory about desensitization, therefore I believe that there is a different factor present in case of frequency of interaction with advanced AI.

Other demographic factors that can influence the opinion on some aspects of ethical and moral considerations of sentient AI are gender, race and ethnicity, and religion of respondents.
In the case of gender, the results suggest that there are not as many significant differences in opinion among men and women. The biggest difference is in the perceived possibility of AI being sentient in the future, men are generally more inclined that it will be possible. Results also suggest that women are less inclined to interact with AI and they do not consume AI narratives as often as men do.
Religion has a measured linkage to how the respondent views sentient AI, although I had to exclude one category (Sikh religion specifically) because there are only two people. Moreover, it is important to note that although all these measures were significant, the values of Cramer's V were almost always below 0.16, although the contingency coefficient predicted that there is a bigger connection among the variables.
Results show that respondents of Catholic and Protestant religions are more likely to support a ban on the development of AI-enhanced humans. Very similar results are for the support of legal rights for sentient AI. Catholic and Protestant respondents are generally more opposed to the idea. Interestingly, the situation differs in responses to whether the respondent views the welfare of sentient AI as one of the most important social issues nowadays. On average, all categories strongly disagree in approximately 40% of cases. Specifically, Protestants are inclined to strongly disagree (48%), but one category has a vastly different opinion on this matter - 68% of Muslims agree.
The race and ethnicity of a respondent had almost no measured effect on the opinion on sentient AI. Although, it is important to note that I had to exclude options Indigenous and other, both had only a few categories and even the combination of these two options was not enough. Also, there was the same problem with the contingency coefficient and Cramers' V as in the case of religion. There is one exception to the measured effect, the question of whether the respondent views the welfare of sentient AI as one of the most important social issues nowadays. Black and White respondents disagree in approximately 65% of cases, however that was the case for only 55% of Asian respondents and surprisingly only for 40%

of Hispanic respondents. Also, approximately half of Hispanic respondents state that they work with advanced AI, and they are more likely to consume AI-related media.

There were also a few demographic factors with no measured effect on the support of the welfare of sentient AI or robots. Specifically education level, income, and surprisingly political orientation.
However, in the case of education, it is evident that more educated respondents slightly more interact with AI (they more likely respond that they own such a piece of technology, work with it and consume AI narratives) than less educated ones.
Although, results did not suggest any effect of the income of respondent, the bigger the income, the bigger the chance to own advanced AI and interact with AI. Wealthier respondents are also more likely to work with it and consume AI narratives.
While my results suggest that political orientation does not affect the opinion on the welfare of AI and robots, the only difference between liberals and conservatives is in the perception of the environment. Conservative respondents are more likely to not support the environment to be included in the moral circle (17% of them in comparison to 0% of liberals that think so). They also in 20% of cases do not think that the welfare of the environment is one of the most important social issues that is the case in only 3% of liberals. This directly contradicts with the findings of Martínez and Winter (2021b). These authors argue that conservatives have lower levels of empathy, therefore they are less likely to empathize with artificial entities. My results do not confirm their findings, because conservative and liberal respondents show comparable levels of compassion for these entities and animals. My results are also in alignment with the literature on the difference in the views on the environment of liberals and conservatives (for example Dunlap et al., 2010; Dunlap & McCright, 2008).
There was no measured effect of any demographic variable on the perception of the welfare of animals. However, there are major differences in how the respondents view the welfare of sentient AI and animals. While approximately 87% of respondents would include animals in the moral circle, only 37% would include sentient AI there. And approximately 79% of respondents think that the welfare of animals is one of the most important social issues in the world, and only 25% of think the same about the welfare of sentient AI. So theoretically, when according to Darling (2012) people tend to perceive robots more as pets than mere things, then the emphasis on this phenomenon and the animal rights approach offered by Chessman (2018) can help with the perception of granting legal rights to sentient AI, but further research is needed to prove this theory.

It is important to mention that I did not measure the effect of a diet of the respondent on the support of the well-being of sentient artificial entities. There are only a few cases of respondents who do not eat meat and the results would have higher chance to be biased. The contingency tables generally do not perform very well with too few cases.

However, the main analysis of this paper is not contingency tables, I am using them to broaden the scope of the analysis Sentience Institute conducted and to provide some context for the upcoming latent class analysis.

4.3 Latent class analysis

Firstly, I am offering some general information. There are four batteries of questions examined, but six sets of individual LCA were conducted for them (and one additional),

because of the lack of computation power on my side. The PMC battery and MCE battery were split in half, SI battery and MCA, and MCEn batteries were calculated as they were. My goal was to connect them and get the bigger picture about the groups relating their opinion on the approach towards sentient AI. Therefore, I had to determine the number of classes to be the same for every calculation, the best number when it comes to the appropriate values of AIC, BIC, and enthropy, and interpretability was 3. Also, in contradiction to my results from contingency tables, there was no measured linkage to gender, religion, race, and ethnicity, or any other demographic variable across all groups or classes.

4.3.1 Latent class analysis for PMC1 to PMC6

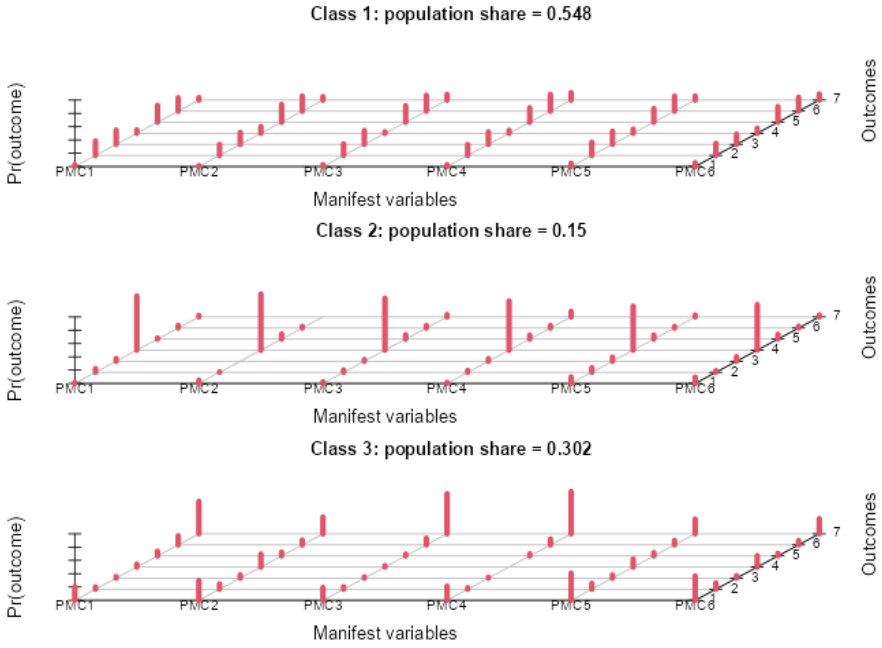
The first latent class analysis I calculated is different from the rest of them. Although, the values of AIC, BIC and enthropy (unfortunately unstandardized) suggest that 3 classes are suitable for a final model (as can be seen in Table 1 below). Nonetheless, when it comes to the interpretation, it does not make sense, but that also was the case for any number of classes for this set of variables.

Table 1: Measures for LCA PMC1-6

Number of classes	AIC	BIC	Enthropy
3	25179	25742	10.2

As we can see in the table above and the plot below, it computed 3 classes that are in fact different from each other, but regardless of the topic of the questions, the distribution of the responses suggest that respondents in group 1 agree and disagree with everything.
Respondents in group 2 are uncertain about everything, they have no opinion on these issues.
Respondents in group 3 just go for the extremes on both sides, which means that this group is at the same time very supportive and very unsupportive of these claims.
There exist two possible outcomes that can explain the results of the analysis. The calculation failed (I calculated it six times and every time the results were the same so that probably is not the case), or there are in reality no real classes that could be measured by the responses to these questions.
Although, contingency tables showed some tendencies, my interpretation is that topics such as a global ban on the development of sentience in robots/AIs (PMC1), a global ban on the development of applications that put the welfare of robots/AIs at risk (PMC2), a global ban on the development of AI-enhanced humans (PMC3), a global ban on the development of robot-human hybrids (PMC4), a global ban on the use of sentient robots/AIs for labor without their consent (PMC5) and a global ban on the use of sentient robots/AIs as subjects in medical experiments without their consent (PMC6), surprisingly do not divide the respondents into meaningful groups.
At the same time, when I tested the linkage of membership in these classes with demographic variables through contingency tables, there was no significant relation to any of them present in the case of this latent class analysis.

Plot 1: LCA plot for PMC1-6



4.3.2 Latent class analysis for PMC7 to PMC12

The second latent class analysis I computed is for variables PMC7 to PMC12 and as you can see in the table below, the measures suggest that 3 classes are fitting the final model accurately.

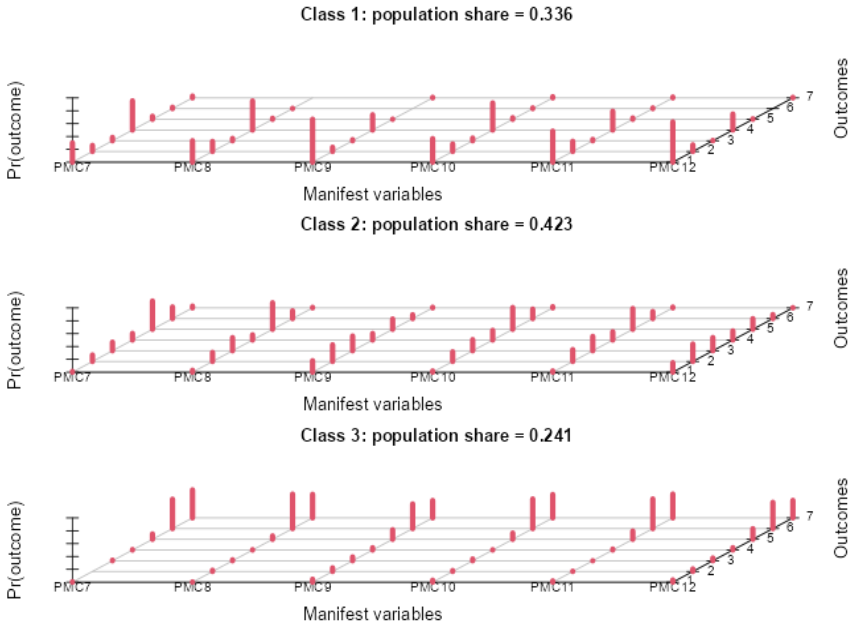
Table 2: Measures for LCA PMC7-12

Number of classes	AIC	BIC	Enthropy
3	23 401	23 963	9.48

In this set of questions there are statements regarding the support of safeguards on scientific research practices that protect the well-being of sentient robots/AIs (PMC7); support of the development of welfare standards that protect the well-being of sentient robots/AIs (PMC8); support of granting legal rights to sentient robots/AIs (PMC9); support of campaigns against the exploitation of sentient robots/AIs (PMC10); support asking institutions like the government and private corporations to fund research that protects sentient robots/AIs (PMC11); and consideration for joining a public demonstration against the mistreatment of sentient robots/AIs (PMC12). The scale is from 1 – strongly disagree to 7 – strongly agree. As you can see in the plot below, there are approximately 34% of respondents in the first class, 42% of respondents in the second class, and 24% of them in the third class. Their responses to these statements suggest that class 1 was frequently against these statements, or these respondents do not have an opinion on them. Class 2 is a different case, they most frequently somewhat agree with the statements, but they are also generally against the idea of granting legal rights to AI. They have various reactions on joining a public demonstration on behalf of AI and robots. However, class 3 is very supportive of these statements, but generally not very sure whether to join a public demonstration against the mistreatment of sentient AI. Class 1 are according to my results respondents, who generally do not own any advanced AI (74%), they also do not work with it (87%) and 42% of them are older than 60 years.
Class 2 are generally people who have the biggest possibility to work with advanced AI (31%) and they are likely to own such device (47%), also the majority of them (51%) are 31 to 50 years old.

Class 3 is very interesting, these respondents have the lowest chance to own (17%) or work (6%) with advanced AI, and they are more likely to be older than 60 years (42%).

Plot 2: LCA plot for PMC7-12



4.3.3. Latent class analysis for MCE1 to MCE5

The third latent class analysis I computed is for variables MCE1 to MCE5.
As you can see in the table below, the measures suggest that 3 classes are fitting the final model accurately.

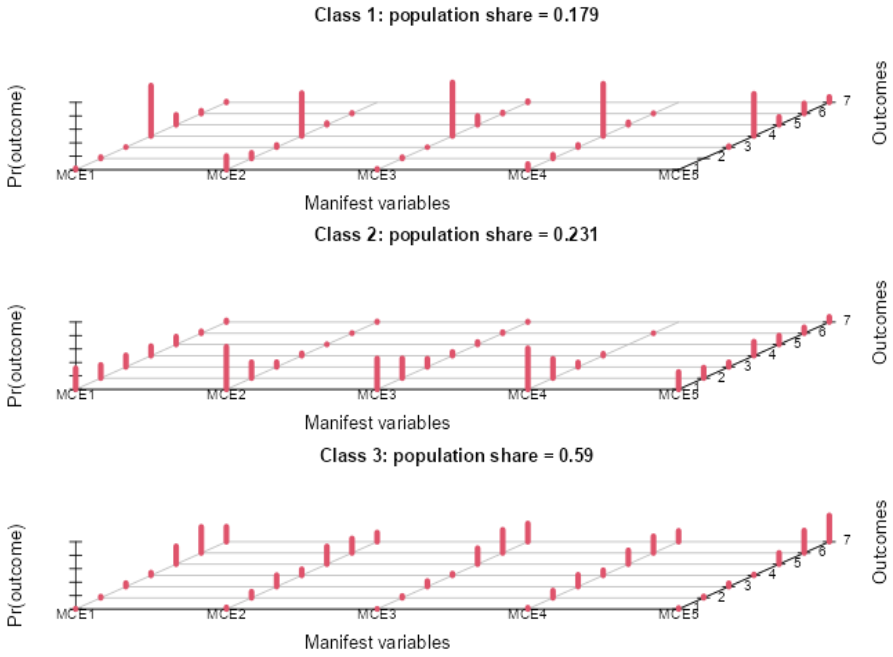
Table 3: Measures for LCA MCE1-5

Number of classes	AIC	BIC	Enthropy
3	19 803	20 272	8.20

The statements are: sentient robots/AIs deserve to be treated with respect (MCE1); sentient robots/AIs deserve to be included in the moral circle (MCE2); physically damaging sentient robots/AIs without their consent is wrong (MCE3), re-programming sentient robots/AIs without their consent is wrong (MCE4); and torturing sentient robots/AIs is wrong (MCE5). The scale is the same as in the previous set and the LCA plot for these variables is below.
The first class is the smallest, only 17% of respondents are there. They are likely to not have an opinion on these statements, although they are inclined to think that torturing sentient robots is at least slightly wrong. The second class contains only 23% of respondents, who are generally not supportive of these statements. However, the opinion on the inclusion of AI in the moral circle and the torture of sentient robots is not very defined for this group as a whole. The third class is large, with 59% of respondents generally agree with all these statements, only in the case of re-programming sentient AI those responses vary.
Class 1 is made up of respondents who are most likely to work with advanced AI (27%), 48% of these respondents are younger than 41 years.
Class 2 are respondents who generally do not work with advanced AI (86%) and 38% of them are 31 to 50 years old.

Class 3 are generally respondents who do not work with advanced AI (92%). They are also older, 51% of them are older than 60 years.

Plot 3: LCA plot for MCE1-5



4.3.4 Latent class analysis for MCE6 to MCE9

The fourth latent class analysis is for variables MCE6 to MCE9.
It is evident from the table below that the measures suggest that 3 classes are fitting the final model appropriately.

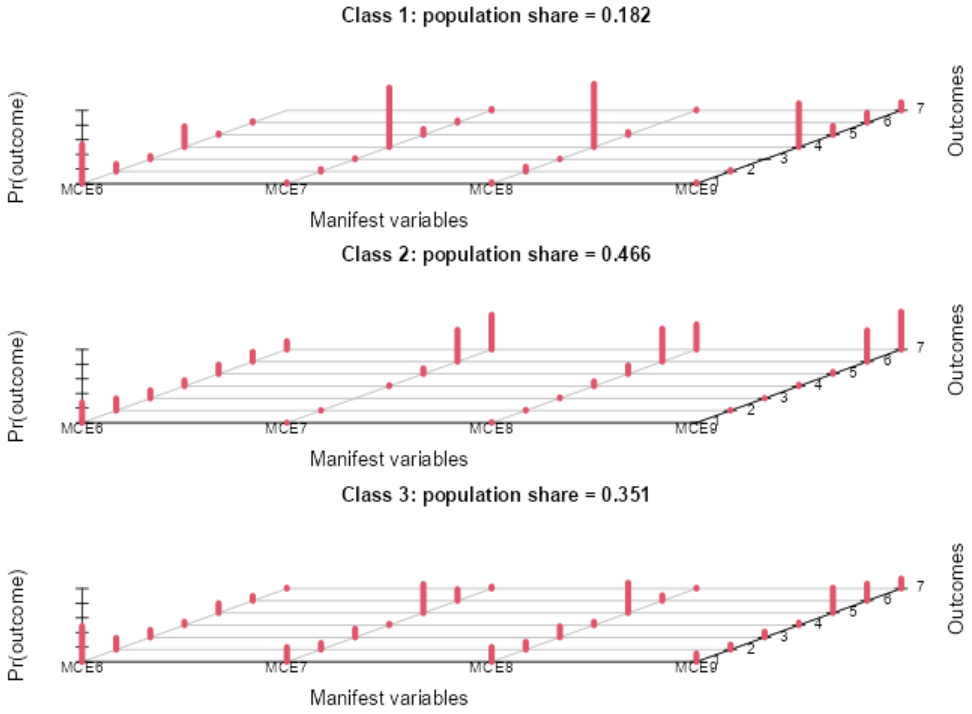
Table 4: Measures for LCA MCE6-9

Number of classes	AIC	BIC	Enthropy
3	15171	15550	6.13

This set of variables tested claims: the welfare of robots/AIs is one of the most important social issues in the world today (MCE6); sentient robots/AIs deserve to be protected from people who derive pleasure from inflicting physical or mental pain on them (MCE7); it is right to protect sentient robots/AIs from vindictive or retaliatory punishment (MCE8); it is wrong to blackmail people by threatening to harm robots/AIs they care about (MCE9). Scale is the same as in the previous LCAs and for the LCA plot for these variables see the plot below.
Class one is the smallest, with only 18% of respondents who generally do not have an opinion on these issues. The exception is that they certainly do not think that the well-being of AI is an important social issue.
Class two is much bigger with 46% of respondents in it. They are generally undecided on the first claim as a group, but in other cases, they agree with these statements.
Class three is approximately a third of all respondents (35%), who are inclined to not think that the well-being of AI is an important social issue. Yet, they generally somewhat agree with the other statements.
Class 1 is made of respondents who are likely to own advanced AI in 31% of cases, but they

are not likely to work with it (14%), only 13% of them consume AI narratives at least monthly, 39% of these respondents are older than 60 years.
Class 2 are respondents who have a slightly bigger chance to own advanced AI than respondents from group 1, 40% of them do own it. At the same time, these respondents have the biggest chance to work with it (27%) and 31% of them consume AI-related media more than monthly. Also, 65% of them are younger than 41 years.
Class 3 are generally respondents who do not own advanced AI (82%) and do not work with it (92%) and they also do not consume media featuring AI more than a few times a year (95%). They are also older, half of them (51%) are older than 60 years.

Plot 4: LCA plot for MCE6-9



4.3.5 Latent class analysis for SI1 to SI4

The fifth latent class analysis was calculated for variables SI1 to SI4.
The table below suggests that 3 class model fits the final model well.

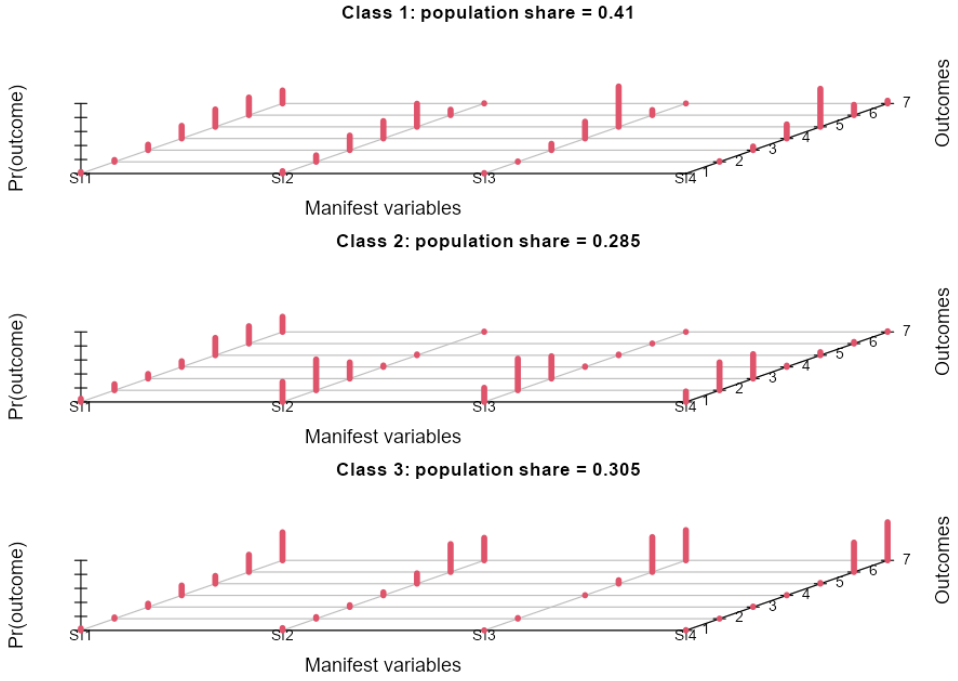
Table 5: Measures for LCA SI1-4

Number of classes	AIC	BIC	Enthropy
3	15918	16297	6.45

This set of statements contains: robots/AIs should be subservient to humans (SI1); robots/AIs may be harmful to me personally (SI2); robots/AIs may be harmful to people in the USA (SI3); robots/AIs may be harmful to future generations of people (SI4). The scale is the same as in the previous LCAs, which you can see in the LCA plot below. There was no significant linkage to any sociodemographic variable throughout the classes.
There are 41% of respondents in the first class. They generally somewhat agree with those statements, except for the perceived danger of robots for them, where most of them disagree.

In class two there are 29% of respondents. This group on the other hand generally disagree with those statements, with the exception of robots being subservient to humans, where they mostly agree.
The last group contains 30% of respondents, who generally strongly agree with all of these statements.

Plot 4: LCA plot for SI1-4



4.3.6 Latent class analysis for MCA1, MCA2, MCEn1 and MCEn2

The last LCA computed for variables in the questionnaire is for variables MCA1, MCA2, MCEn1, and MCEn2. I merged the analysis of these variables into one, because they are thematically very similar to one another and computing LCA for only 2 variables is not suitable for LCA as a method.
As you can see in the table below, measured values suggest a suitable fit of three class model.

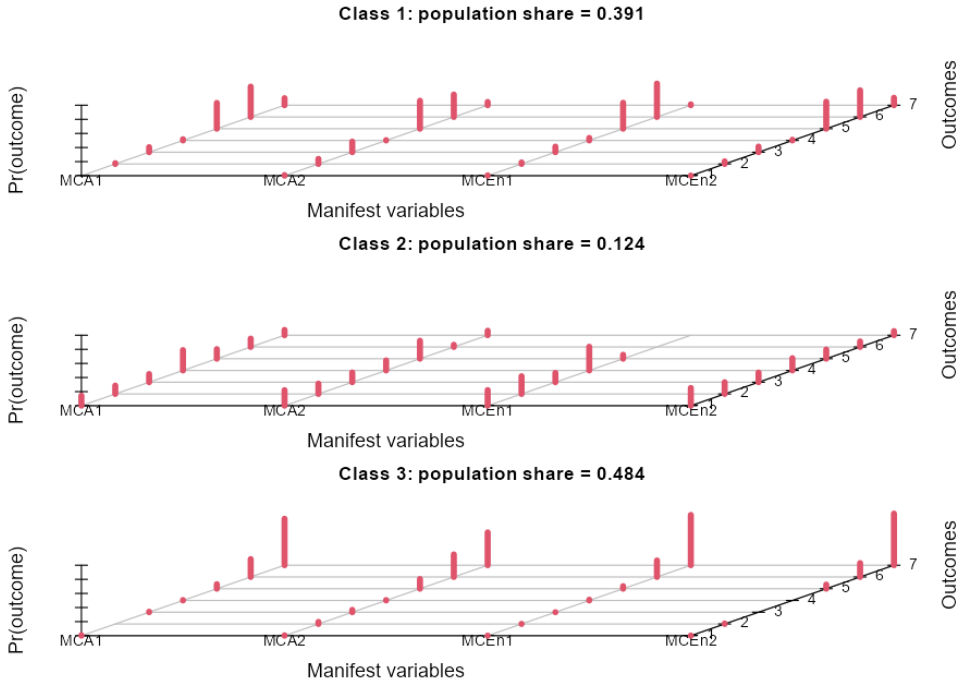
Table 6: Measures for LCA MCA1-2 and MCEn1-2

Number of clases	AIC	BIC	Enthropy
3	13753	13753	5.56

The last set of variables is on the topic of inclusion to the moral circle, the statements are: animals deserve to be included in the moral circle (MCA1); the welfare of animals is one of the most important social issues in the world today (MCA2); the environment deserves to be included in the moral circle (MCEn1); the welfare of the environment is one of the most important social issues in the world today (MCEn2). The scale is the same as in the previous LCAs. It is important to note that there was no measured linkage to any sociodemographic variable.
Class 1 contains 39% of respondents, who generally somewhat agree with these statements. Class 2 is the smallest with only 12% of respondents. They are extremely undecisive as a group. They generally do not really know whether to include animals in the moral circle as a

group, and they are generally somewhat denying that their well-being is an important issue. Also, they generally do not believe that the environment should be included in the moral circle, but at the same time, their responses suggest only a very small inclination for not recognizing the welfare of environment as an important issue.
On the other hand, the third class, made of 48% of respondents, is extremely supportive of both animals and the environment.

Plot 6: LCA plot for LCA MCA1-2 and MCEn1-2



4.3.7 Latent class analysis for membership in classes

How to link all those classes from all 6 LCAs together? Well, another segmenting procedure is needed to determine whether some classes across the analyses align with each other. Therefore, I chose to do another latent class analysis, but not with each of those variables, but instead with membership in these classes. I did not find this in literature about latent class analysis, but it was the only solution that made some sense in my case, because I do not think that alignment of classes can be successfully measured by the alignment of sociodemographic criteria only.
I chose three classes for this LCA to ensure interpretability and I decided to exclude the first LCA I calculated (PMC1 to PMC7) to get better fitting result in this one, which is supported by the measures in the table below.

Table 7: Measures for memberships in classes

Number of clases	AIC	BIC	Enthropy
3	11 158	11 321	4.51

As you can see in the plot below, there are three classes relatively equal in size.

The class 1 contains 38% of respondents and aligns with classes 1 and 3 in the second LCA (support of safeguards on scientific research practices that protect the well-being of sentient robots/AIs (PMC7); support of the development of welfare standards that protect the well-

being of sentient robots/AIs (PMC8); support of granting legal rights to sentient robots/AIs (PMC9); support of campaigns against the exploitation of sentient robots/AIs (PMC10); support asking institutions like the government and private corporations to fund research that protects sentient robots/AIs (PMC11); and consideration for joining a public demonstration against the mistreatment of sentient robots/AIs (PMC12)), which is contradictory. Group one generally somewhat disagree with these statements, or these respondents do not have an opinion on these issues, on the other hand, group 3 is generally very supportive of these statements. The only exception, where these groups meet each other is the unwillingness to join a public demonstration on behalf of AI (PMC12).

This class also almost perfectly aligns with the second class in case of LCA for variables MCE1 to MCE5, which means that these respondents are generally not supportive of two statements: physically damaging sentient robots/AIs without their consent is wrong (MCE3) and re-programming sentient robots/AIs without their consent is wrong (MCE4). The opinion on statements: sentient robots/AIs deserve to be included in the moral circle (MCE2) and torturing sentient robots/AIs is wrong (MCE5) is not defined at all for this group.

Class 1 aligns with the first class in LCA for variables MCE6 to MCE9 (the welfare of robots/AIs is one of the most important social issues in the world today (MCE6); sentient robots/AIs deserve to be protected from people who derive pleasure from inflicting physical or mental pain on them (MCE7); it is right to protect sentient robots/AIs from vindictive or retaliatory punishment (MCE8); it is wrong to blackmail people by threatening to harm robots/AIs they care about (MCE9)). That means that these respondents generally have no opinion on these issues except for the well-being of AI is an important social issue they strongly disagree with.

In case of LCA for variables SI1 to SI4 (statements: robots/AIs should be subservient to humans (SI1); robots/AIs may be harmful to me personally (SI2); robots/AIs may be harmful to people in the USA (SI3); robots/AIs may be harmful to future generations of people (SI4)) it is mostly second and third class it aligns with, which is also contradictory. Respondents in class 2 generally disagree, but respondents in class 3 strongly agree. Except for the question on the social status of robots and AIs, these groups agree that these entities should be subservient to humans.

In the case of the last LCA (MCA1, MCA2, MCEn1, MCEn2) it is the third class it aligns with. These respondents strongly agree with the inclusion of environment and animals into the moral circle, and they think that the well-being of environment and animals is important.

This class has generally the lowest rate of owning advanced AI (16%) and the lowest rate of interaction with it monthly or more (8%). These respondents also do not consume media featuring AI, specifically only 4% of them consume such content on monthly basis and more. Also, it is important to note, that half of them (51%) are older than 60 years.

In other words, this class is very contradictory. Some respondents are inclined to support the welfare of sentient artificial entities, some have no opinion on it, and some do not care about it at all. I think it is partly because they lack the information about AI they claim to receive. They also do not own any advanced AI to have some day-to-day experience with it. Maybe because of that they tend to be much more scared of the potential danger of AI than the other two groups, which can be a factor in the lack of support of these entities they are willing to provide.

For this group of people, I recommend to not scare them more than they already are. Although, it is not the case for everyone, a significant part of them is inclined to support the welfare of AI and robots already. It is important to inform them about the issues sentient AI

can face in the future, but I recommend doing it in a way that considers the fear some of them already feel and choose a neutral tone stressing the benefits of protection of these entities. Also, they would maybe benefit from information they need to be able to interact with advanced AI when they need it, because they generally do not own such piece of technology.

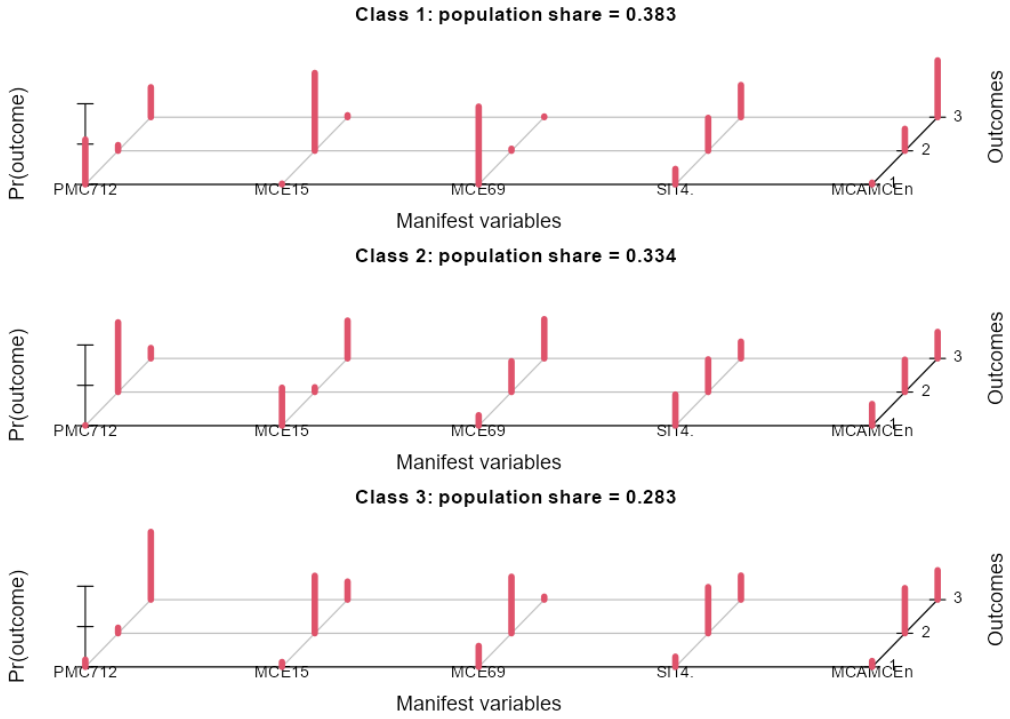
There are 33% of respondents in the second class. It aligns almost perfectly with the second class in LCA for variables PMC7 to PMC12. Which means that these respondents generally somewhat agree to support the safeguards on scientific research practices that protect the well-being of sentient robots/AIs (PMC7); support of the development of welfare standards that protect the well-being of sentient robots/AIs (PMC8); support of campaigns against the exploitation of sentient robots/AIs (PMC10) and support asking institutions like the government and private corporations to fund research that protects sentient robots/AIs (PMC11). However, they have various reactions to the consideration for joining a public demonstration against the mistreatment of sentient robots/AIs (PMC12). They are also somewhat against the support of granting legal rights to sentient robots/AIs (PMC9).
There is an equal chance of these respondents to be in the first or third class in the LCA for variables MCE1 to MCE5. Which means that they are generally not sure, or they somewhat agree with statements such as: sentient robots/AIs deserve to be treated with respect (MCE1); sentient robots/AIs deserve to be included in the moral circle (MCE2); physically damaging sentient robots/AIs without their consent is wrong (MCE3) and torturing sentient robots/AIs is wrong (MCE5). The only exception is in the case of statement re-programming sentient robots/AIs without their consent is wrong (MCE4), where the responses vary.
In case of LCA for variables MCE6 to MCE9 there is biggest alignment with class 3. These respondents are generally likely to support statements such as: sentient robots/AIs deserve to be protected from people who derive pleasure from inflicting physical or mental pain on them (MCE7); it is right to protect sentient robots/AIs from vindictive or retaliatory punishment (MCE8); it is wrong to blackmail people by threatening to harm robots/AIs they care about (MCE9). Although, they are not inclined to think that the welfare of robots/AIs is one of the most important social issues in the world today (MCE6).
Nonetheless, for the last two LCAs it is not that clear, the biggest alignment is with classes 1 and 2 in case of LCA for variables SI1 to SI4. The same is true for classes 2 and 3 in case of LCA for variables MCA1, MCA2, MCEn1 and MCEn2. In case of LCA for variables SI1 to SI4 the opinions do not align with two exceptions. These respondents agree that robots/AIs should be subservient to humans (SI1), and they do not agree that robots or AIs could be harmful to them personally (SI2). Yet, in case of statements such as: robots/AIs may be harmful to people in the USA (SI3) and robots/AIs may be harmful to future generations of people (SI4), class 1 generally somewhat agree, but class 2 generally disagree.
In case of the last LCA with statements such as: animals deserve to be included in the moral circle (MCA1); the welfare of animals is one of the most important social issues in the world today (MCA2); the environment deserves to be included in the moral circle (MCEn1); the welfare of the environment is one of the most important social issues in the world today (MCEn2), the situation is similar. Although, class 1 generally somewhat agree with these statements, class 2 is extremely undecisive as a group. The respondents generally have no opinion on whether to include animals in the moral circle, so they deny that their well-being is an important issue. They think almost the same about the environment.
This class is generally the youngest with 40% of respondents younger than 40 years. Therefore, it makes sense that these respondents have a higher chance of owning advanced AI

(32%), interacting with it monthly or more (18%) and consuming media about AI monthly and more (15%).
In conclusion, this class is generally inclined to give support to sentient AIs and robots, but it must be a passive support that benefits them, which makes sense, because this group tends to view sentient AI as a potentially dangerous for future generations in some cases. To give some examples I made this assumption of, they are not willing to attend demonstrations on behalf of sentient AI, they do not think that the wellbeing of sentient AI is important issue, they think that AI should be subservient to humans, and they do not think that reprogramming sentient AI without consent is extremely harmful, although they are inclined to protect it from harm.
This group is generally supportive of well-being of AI, but because it is only in a passive way for most of them, I suggest to give them information about the importance and urgency of legal framework for protection of these entities and stress the benefits for this group of people and future generations of AI and humans well.

The third class is the smallest with 28% of respondents and in most cases, it aligns mostly with one class from each LCA. It almost perfectly aligns with class 3 in the second LCA - support of safeguards on scientific research practices that protect the well-being of sentient robots/AIs (PMC7); support of the development of welfare standards that protect the well-being of sentient robots/AIs (PMC8); support of granting legal rights to sentient robots/AIs (PMC9); support of campaigns against the exploitation of sentient robots/AIs (PMC10); support asking institutions like the government and private corporations to fund research that protects sentient robots/AIs (PMC11); and consideration for joining a public demonstration against the mistreatment of sentient robots/AIs (PMC12). Meaning that these respondents are generally very supportive of these statements, the only exception is the attendance of public demonstration on behalf of AI (PMC12) they are not really sure about.
It also mostly aligns with class 2 in the third LCA (MCE1 to MCE5), with statements such as: sentient robots/AIs deserve to be treated with respect (MCE1); sentient robots/AIs deserve to be included in the moral circle (MCE2); physically damaging sentient robots/AIs without their consent is wrong (MCE3), re-programming sentient robots/AIs without their consent is wrong (MCE4); and torturing sentient robots/AIs is wrong (MCE5). These respondents generally somewhat disagree with the statements, although they are not sure about the inclusion of AI to the moral circle and the torture of sentient robots as a group.
It is also close to class 2 in case of fourth LCA - the welfare of robots/AIs is one of the most important social issues in the world today (MCE6); sentient robots/AIs deserve to be protected from people who derive pleasure from inflicting physical or mental pain on them (MCE7); it is right to protect sentient robots/AIs from vindictive or retaliatory punishment (MCE8); it is wrong to blackmail people by threatening to harm robots/AIs they care about (MCE9). Therefore, these respondents generally agree with these statements, although they are not sure, whether the welfare of AIs and robots is an important issue.
This class aligns mostly with the second class in case of the LCA for variables SI1 to SI4 - robots/AIs should be subservient to humans (SI1); robots/AIs may be harmful to me personally (SI2); robots/AIs may be harmful to people in the USA (SI3); robots/AIs may be harmful to future generations of people (SI4). That means that these respondents agree with only one statement about robots being subservient to humans, in other cases they generally disagree. Yet, a third of these respondents are also in the third class of this LCA, who strongly agree with every statement.

In the case of the last LCA (animals deserve to be included in the moral circle (MCA1); the welfare of animals is one of the most important social issues in the world today (MCA2); the environment deserves to be included in the moral circle (MCEn1); the welfare of the environment is one of the most important social issues in the world today (MCEn2)) the biggest alignment is with the second class, which is extremely indecisive but generally inclined to not support these statements as a group. The next significant alignment is with the third class that is very supportive of both animals and the environment.
This class has the highest rate of owning advanced AI (42%), interacting with it monthly and more (32%) and consuming media featuring AI (30%) this frequently. A significant number of them are 31 to 50 years old (46%).
In other words, this class is supportive of sentient AI or robots and their welfare in the sense that according to them, some safeguards and protection is needed. However, when it comes to different levels of harm towards sentient AI, some of them are not so sure that it is wrong. That can be the case of their belief whether the current AI is sentient I presented previously, because they are maybe not willing to take into account that their action towards sentient AI would be harmful given the percentage of these respondents who own advanced AI. It is also likely that they are unwilling to support these entities in this case, because they view them as a potential danger.
In the case of this class, these respondents already know that some protection of sentient artificial entities is needed, but it would be beneficial to know the reasons, why they think that way, because they tend to not think that these entities should be respected and that certain behavior towards them without consent is wrong. Thus, I recommend informing them about the potential autonomy of sentient artificial entities and the importance of their consent that should be incorporated in a message promoting the significance of potential legal framework that would protect these entities and humans as well.

Plot 7: LCA plot for memberships in classes



5 Conclusion

My analysis offered some results suggesting that there indeed are groups within the respondents from the 2021 AIMS study. Therefore, possibly some groups exist in the US society regarding the opinion on the welfare of (sentient) AI and robots from various perspectives. The topics are the inclusion of these entities, animals, and the environment in the moral circle, granting legal rights to sentient AI, and support of the well-being of AI in the form of protection from harm or the perceived danger of AI for society.

The theoretical literature and current research offered some interesting perspectives on various topics I included in one analysis. In conclusion, here are the responses to my questions and verdicts for my hypotheses from the theoretical part of my paper.

Is US society divided into groups relating opinions about artificial intelligence? (H1)
Yes, it is possible that US society is divided into precisely three groups when it comes to moral considerations on artificial intelligence and robots.

Is age, political preference, another sociodemographic, or experience with AI the dividing factor? (H2)
Mostly yes. These groups not only have different opinions, but it is likely that the dividing factors are age, ownership of advanced AI, and consumption of media featuring AI or robots. Which is in alignment with the results of the research conducted by Zhang and Dafoe (2020). Nonetheless, there is no significant relation to gender, race and ethnicity, income, or any other socio-demographic variable.

How big these groups are? (Q1)
In the case of the last LCA, which is the most important, these groups are almost the same size. In the first class, there are 38% of the respondents, in the second class there are 33% of the respondents and in the third class, there are 28% of the respondents.

How are these groups viewing AI as a whole? (Q1)
I am describing only the classes from the last and the most important latent class analysis.
Class 1 is very contradictory. These respondents are generally inclined to slightly support sentient AI and protect it. In some cases, they simply do not have an opinion on the issues that were presented to them and in some cases, they are extremely unsupportive. Yet, overall, they are most likely to view these entities as a potential danger.
Class 2 is inclined to protect sentient AI, but these respondents prefer passive support that benefits them. They are in some cases likely to view these entities as potentially dangerous mainly for future generations.
Class 3 is generally inclined to support sentient AI. However, these respondents are not sure, whether to include it in the moral circle, although they are willing to include animals and the environment. They also do not really take into account the potential consent of sentient artificial entities. In some cases, they also tend to think that sentient AI can be dangerous to humans.

What can we assume about US society in consideration of these groups? (Q3)
I believe that most of these respondents are not informed enough about these topics. They usually want to protect AI from harm, but they do not think that legal protection is suitable. Also, they are generally not inclined to include sentient AI into the moral circle, which can be the result of the lack of information I mentioned. Nonetheless, maybe I am wrong, and they

know everything they needed for their decisions. At the end of the day, all groups state that they do not think that the welfare of AI is one of the most important social issues today, which can have an impact on their responses. So, there is a huge possibility that they have more important social issues to think about and the moral consideration of AI and robots is simply not urgent enough for them to care about it.
It is also interesting that these classes are not divided by their political preferences. Which can mean that no political party in the US ever paid attention to this topic and therefore there was no such tendency measured. There was also measured no linkage of gender, religion, income, and race and ethnicity to any of those classes.

How to communicate information about AI to each of those classes (in the last LCA) effectively? (Q4)
Firstly, I suggest a neutral tone and I strongly recommend not promoting fear, significant number of respondents in all classes are already worried about the danger sentient AI could potentially be to the point, that it can impact the opinion on the rights sentient AI could receive. One of the approaches that could be considered for all of these classes is to compare sentient AI to pets and communicate the legal rights for AI as similar to animal rights, but further research on this possibility is needed.
Class 1 are respondents, who are generally not very informed (they mostly do not consume media featuring AI), therefore I suggest informing them about the benefits of protection of these entities. Also, this class would maybe benefit from the information they need to be able to interact with advanced AI when they need it because they generally do not own such a piece of technology.
For people in the second class, I suggest giving them information about the importance and urgency of a legal framework for the protection of these entities and stress the benefits of it for this group of people and future generations of AI and humans well.
In the case of class 3, I recommend informing them about the potential autonomy of sentient artificial entities and the importance of their consent that should be incorporated in a message promoting the significance of a potential legal framework that would protect these entities and humans as well.

If there are any differences in the levels of support for the welfare of animals, the environment, and sentient AI and robots, can the animal rights framework theoretically help to shape public opinion on ascribing rights to sentient AI? (Q5)
There are some major differences. Most respondents would include animals in the moral circle and think that the welfare of animals is one of the most important social issues today, but only a third of them would include sentient AI in the moral circle, and only fourth of them think that the welfare of sentient AI is that important. So yes, the emphasis on the perception of robots as pets (based on the research by Darling (2012)) and the animal rights approach towards rights for AI and robots (Chessman, 2018) can theoretically help with the perception of sentient artificial entities and their rights. However, further research is needed to prove that.

I think that the opinion of the US general public on AI and robots will change someday to support the welfare of artificial entities more than my results showed. Yet, the current levels suggest some level of compassion towards these entities is already present. Although, that can be influenced by the perceived danger of these entities and the perception of the possibility of sentience in current and future AI, which is completely understandable.
Also, I am glad that there is a number of recent studies (like AIMS and other featured in the second chapter dedicated to the literature review) that are focused on public opinion on this topic and other topics regarding AI. There is a huge possibility of this topic to be featured

more frequently in the future if the trend continues. There are some topics regarding AI that simply need some further public opinion research.

Bibliography

About | Metaculus. (n.d.). Retrieved December 26, 2022, from https://www.metaculus.com/about/

Ashrafian, H. (2014). *AIonAI : A Humanitarian Law of Artificial Intelligence and Robotics*. https://doi.org/10.1007/s11948-013-9513-9

Asimov, I. (n.d.). *Foundation series*. Astounding Science Fiction (Street & Smith), Gnome Press, Spectra, Doubleday.

Bailey, K. D. (1983). Sociological classification and cluster analysis. *Quality & Quantity*, *17*(4), 251–268. https://doi.org/10.1007/BF00167539

Brooks, R. A. (1999). *Cambrian Intelligence: The Early History of the New AI*. MIT Press. https://doi.org/0-262-52263-2

Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*, *13*(2). https://doi.org/10.1007/BF01246098

Chalmers, D. J. (1996). *The Conscious Mind:In search of a fundamental theory*. Oxford University Press.

Chessman, C. F. (2018). Not Quite Human: Artificial Intelligence, Animals, and the Regulation of Sentient Property. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3200802

Clogg, C. C. (1995). *Latent Class Models BT - Handbook of Statistical Modeling for the Social and Behavioral Sciences* (G. Arminger, C. C. Clogg, & M. E. Sobel (Eds.); pp. 311–359). Springer US. https://doi.org/10.1007/978-1-4899-1292-3_6

Coeckelbergh, M. (2010). Robot rights? Towards a social-relational justification of moral consideration. *Ethics and Information Technology*, *12*(3). https://doi.org/10.1007/s10676-010-9235-5

Collins, L. M., & Lanza, S. T. (2010). Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences. In *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. https://doi.org/10.1002/9780470567333

Darling, K. (2016). Extending legal protection to social robots: The effects of anthropomorphism, empathy, and violent behavior towards robotic objects. In *Robot Law*. https://doi.org/10.4337/9781783476732.00017

Darling, K. (2012). EXTENDING LEGAL PROTECTION TO SOCIAL ROBOTS: THE EFFECTS OF ANTHROPOMORPHISM, EMPATHY, AND VIOLENT BEHAVIOR TOWARDS ROBOTIC OBJECTS. *We Robot Conference 2012*, 24.

Dick, P. K. (1968). *Do Androids Dream of Electric Sheep?* Doubleday. https://doi.org/978-0-575-09418-5

Dillon, W. R., & Kumar, A. (1994). Latent Structure and Other Mixture Models in Marketing: An Integrative Survey and Overview. In *Advanced Methods of Marketing Research* (pp. 295–351). Blackwell.

Dunlap, R. E., Xiao, C., & McCright, A. M. (2010). Politics and environment in America: Partisan and ideological cleavages in public support for environmentalism. *Environmental Politics*, *10*(4), 23–48. https://doi.org/10.1080/714000580

Dunlap, Riley E., & McCright, A. M. (2008). A widening gap republican and democratic views on

climate change. *Environment*, *50*(5), 26–35. https://doi.org/10.3200/ENVT.50.5.26-35

El Mesbahi, M. (2015). Human-robot interaction ethics in sci-fi movies: Ethics are not 'there', we are the ethics! *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9186*, 590–598. https://doi.org/10.1007/978-3-319-20886-2_55

European Commission. (2019). *The Ethics Guidelines for Trustworthy Artificial Intelligence*. https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html

EVANS, M. J., GILULA, Z. V. I., & GUTTMAN, I. (1989). Latent class analysis of two-way contingency tables by Bayesian methods. *Biometrika*, *76*(3), 557–563. https://doi.org/10.1093/biomet/76.3.557

Fienberg, S. E., Hersh, P., Rinaldo, A., & Zhou, Y. (2007). *Maximum Likelihood Estimation in Latent Class Models For Contingency Table Data*. http://arxiv.org/abs/0709.3535

Fonseca, J. R. S. (2012). Clustering in the field of social sciences: that is your choice. *International Journal of Social Research Methodology*, 1–26. https://doi.org/10.1080/13645579.2012.716973

Fonseca, J. R. S., & Cardoso, M. G. M. S. (2007). Mixture-model cluster analysis using information theoretical criteria. *Intelligent Data Analysis*, *11*(2), 155–173. https://doi.org/10.3233/ida-2007-11204

Gennaro, R. J. (n.d.). Consciousness. In *Internet Encyclopedia of Philosopfy*. https://iep.utm.edu/consciousness/#H7

Gentle, J. E., Kaufman, L., & Rousseuw, P. J. (1991). Finding Groups in Data: An Introduction to Cluster Analysis. *Biometrics*, *47*(2). https://doi.org/10.2307/2532178

González, A., & Sánchez, J. E. (2009). *the Study of Cells in the Analysis of Contingency Tables From the Viewpoint of Rudas , Clogg and Lindsay Mixture Index of Fit . 3*, 244–249.

GuidedTrack. (n.d.). Retrieved December 26, 2022, from https://www.guidedtrack.com/

Gunkel, D. J. (2018). Robot Rights. *Robot Rights*. https://doi.org/10.7551/MITPRESS/11444.001.0001

Hagenaars, J. A., & Halman, L. C. (1989). Searching for ideal types: The potentialities of latent class analysis. *European Sociological Review*, *5*(1), 81–96. https://doi.org/10.1093/oxfordjournals.esr.a036509

Hendl, J. (2006). *Přehled statistických metod zpracování dat : analýza a metaanalýza dat* (2nd ed.). Portál. https://www.researchgate.net/publication/47044723_Prehled_statistickych_metod_zpracovani_dat_analyza_a_metaanalyza_dat

Hildt, E. (2019). *Artificial Intelligence : Does Consciousness Matter ? 10*(July), 1–3. https://doi.org/10.3389/fpsyg.2019.01535

Holland, O., & Gamez, D. (2009). Artificial Intelligence and Consciousness. *Encyclopedia of Consciousness*, 37–45. https://doi.org/10.1016/B978-012373873-8.00004-9

IBM. (2019). *SPSS Software | IBM*. IBM Website. https://www.ibm.com/analytics/spss-statistics-software

Jaworska, A., & Tannenbaum, J. (2021). The Grounds of Moral Status. In *Stanford Encyclopedia of Philosophy* (Spring 202). https://plato.stanford.edu/entries/grounds-moral-status/

Kelley, P. G., Heldreth, C., Moessner, C., Sedley, A., Kramm, A., Newman, D. T., & Woodruff, A. (2021). *Exciting , Useful , Worrying , Futuristic : Public Perception of Artificial Intelligence in 8 Countries*. 627–637.

Key figures | Ipsos. (n.d.). Retrieved December 26, 2022, from https://www.ipsos.com/en-us/about-us/key-figures

Kieslich, K., Keller, B., & Starke, C. (2022). Artificial intelligence ethics by design. Evaluating public perception on the importance of ethical design principles of artificial intelligence. *Big Data and Society*, *9*(1). https://doi.org/10.1177/20539517221092956

Killian, M. O., Cimino, A. N., Weller, B. E., & Hyun Seo, C. (2019). A Systematic Review of Latent Variable Mixture Modeling Research in Social Work Journals. *Journal of Evidence-Based Social Work*, *16*(2). https://doi.org/10.1080/23761407.2019.1577783

Kim, S., Roe, D. G., Choi, Y. Y., Woo, H., Park, J., Lee, J. I., Choi, Y., Jo, S. B., Kang, M. S., Song, Y. J., Jeong, S., & Cho, J. H. (2021). Artificial stimulus-response system capable of conscious response. *Science Advances*, *7*(15), 1–9. https://doi.org/10.1126/SCIADV.ABE3996

Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press.

Lazarsfeld, P.F., & Henry, N. W. (1968). *Latent structure analysis*. Houghton Mifflin.

Lazarsfeld, Paul F. (1950). The logical and mathematical foundation of latent structure analysis. *S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star, & J. A. Clausen (Eds.), Studies in Social Psychology World War II: Measurement and Prediction*, *4*, 361–412.

Lima, G., Kim, C., Ryu, S., Jeon, C., & Cha, M. (2020). *Collecting the Public Perception of AI and Robot Rights*. *4*(October).

Magidson, J., & Vermunt, J. K. (2002). Latent class models for clustering: A comparison with K-means. *Canadian Journal of Marketing Research*, *20*(1).

Malsch, T. (2001). Naming the Unnamable: Socionics or the Sociological Turn of/to Distributed Artificial Intelligence. *Autonomous Agents and Multi-Agent Systems*, *4*(3). https://doi.org/10.1023/A:1011446410198

Mandara, J. (2003). The typological approach in child and family psychology: A review of theory, methods, and research. In *Clinical Child and Family Psychology Review* (Vol. 6, Issue 2). https://doi.org/10.1023/A:1023734627624

Martínez, E., & Winter, C. (2021a). Protecting Sentient Artificial Intelligence: A Survey of Lay Intuitions on Standing, Personhood, and General Legal Protection. *Frontiers in Robotics and AI*, *8*(November), 1–7. https://doi.org/10.3389/frobt.2021.788355

Martínez, E., & Winter, C. (2021b). *Protecting sentient artificial intelligence general legal protection*.

Massa, N., Bisconti, P., & Nardi, D. (2022). The Psychological Implications of Companion Robots: A Theoretical Framework and an Experimental Setup. *International Journal of Social Robotics*. https://doi.org/10.1007/s12369-021-00846-x

Mays, K. K., Lei, Y., Giovanetti, R., & Katz, J. E. (2021). AI as a boss ? A national US survey of predispositions governing comfort with expanded AI roles in society. *AI & SOCIETY*, *0123456789*. https://doi.org/10.1007/s00146-021-01253-6

Mccombs, M. (2008). *Setting the agenda: The mass media and public opinion*.

McCutcheon, A. (1987). *Latent Class Analysis*. https://doi.org/10.4135/9781412984713

McLachlan, G., & Peel, D. (2000). *Finite Mixture Models*. John Wiley & Sons, Inc.

Metzinger, T. (2021). Artificial Suffering: An Argument for a Global Moratorium on Synthetic Phenomenology. *Journal of Artificial Intelligence and Consciousness*, *08*(01), 43–66. https://doi.org/10.1142/s270507852150003x

Mlynář, J., Alavi, H. S., Verma, H., & Cantoni, L. (2018). Towards a sociological conception of artificial intelligence. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10999 LNAI*(July), 130–139. https://doi.org/10.1007/978-3-319-97676-1_13

Muthén, B. O., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, *24*(6). https://doi.org/10.1111/j.1530-0277.2000.tb02070.x

Mutz, D. C., & Soss, J. (1997). Reading Public Opinion : The Influence of News Coverage on Perceptions of Public Sentiment. *The Public Opinion Quarterly*, *61*(3), 431–451. http://www.jstor.org/stable/2749580

Nylund-Gibson, K., & Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science*, *4*(4). https://doi.org/10.1037/tps0000176

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *14*(4). https://doi.org/10.1080/10705510701575396

Ouchchy, L., Coin, A., & Dubljević, V. (2020). AI in the headlines : the portrayal of the ethical issues of artificial intelligence in the media. *AI & SOCIETY*, *0123456789*. https://doi.org/10.1007/s00146-020-00965-5

Pauketat, J. (2022a). *Sentience Institute | Artificial Intelligence, Morality, and Sentience (AIMS) Survey: 2021*. https://www.sentienceinstitute.org/aims-survey-2021

Pauketat, J. (2022b, March 23). *Sentience Institute | Data Publication for the 2021 Artificial Intelligence, Morality, and Sentience (AIMS) Survey*. https://www.sentienceinstitute.org/blog/data-publication-AIMS-survey

Pearson, K. (1904). *On the theory of contingency and its relation to association and normal correlation*. Dulau and Co.

Petersen, K. J., Qualter, P., & Humphrey, N. (2019). The application of latent class analysis for investigating population child mental health: A systematic review. In *Frontiers in Psychology* (Vol. 10, Issue MAY). https://doi.org/10.3389/fpsyg.2019.01214

Reddy Nadikattu, R. (2016). the Emerging Role of Artificial Intelligence in Modern Society. *International Journal of Creative Research Thoughts*, *4*(4), 2320–2882. www.ijcrt.orgwww.ijcrt.org

Restivo, S. (2001). Bringing up and booting up: Social theory and the emergence of socially intelligent robots. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, *4*, 2110–2117. https://doi.org/10.1109/icsmc.2001.972867

Riedl, M. O. (2019). Human-centered artificial intelligence and machine learning. *Human Behavior and Emerging Technologies*, *1*(1). https://doi.org/10.1002/hbe2.117

Safron, A. (2021). *Integrated World Modeling Theory (IWMT) Expanded: Implications for Theories of Consciousness and Artificial Intelligence*. *June*. https://psyarxiv.com/rm5b2/

Scheutz, M. (2011). The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots. *Robot Ethics: The Ethical and Social Implications of Robotics*, *November*, 205–221. http://books.google.com/books?hl=en&lr=&id=oBb-lt3l4oYC&oi=fnd&pg=PA205&dq=The+Inherent+Dangers+of+Unidirectional+Emotional+Bonds+between+Humans+and+Social+Robots&ots=yuaSs8y33n&sig=h5KrpJxIcyk5RT2vih0YpgSHqNY

Searle, J. R. (1980). Minds, brains, and programs. *THE BEHAVIORAL AND BRAIN SCIENCES*, *3*,

417–456.

Seol, H. (2022). *snowRMM: Rasch Mixture Model for jamovi. [jamovi module]*. https://github.com/hyunsooseol/snowRMM

Seth, A. K. (2016, November 2). *The hard problem of consciousness is a distraction from the real one | Aeon Essays*. https://aeon.co/essays/the-hard-problem-of-consciousness-is-a-distraction-from-the-real-one

Shanahan, L., Copeland, W. E., Worthman, C. M., Erkanli, A., Angold, A., & Costello, E. J. (2013). Sex-differentiated changes in C-reactive protein from ages 9 to 21: The contributions of BMI and physical/sexual maturation. *Psychoneuroendocrinology*, *38*(10). https://doi.org/10.1016/j.psyneuen.2013.04.010

Technological determinism - Oxford Reference. (n.d.). Retrieved December 29, 2022, from https://www.oxfordreference.com/display/10.1093/oi/authority.20110803102813253;jsessionid=6BC43215720200147D492ED1B84844A8

The jamovi project (jamovi (Version 2.3) [Computer Software]). (2021). https://www.jamovi.org

Ulbricht, C. M., Chrysanthopoulou, S. A., Levin, L., & Lapane, K. L. (2018). The use of latent class analysis for identifying subtypes of depression: A systematic review. In *Psychiatry Research* (Vol. 266). https://doi.org/10.1016/j.psychres.2018.03.003

Van Gulick, R. (2022). Consciousness. In *The Stanford Encyclopedia of Philosophy* (Winter 202). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2022/entries/consciousness/

Vermunt, J. K. (2002). Latent class analysis of complex sample survey data: Application to dietary data. *Journal of the American Statistical Association*, *97*(459), 736–737.

Vermunt, Jeroen K, & Magidson, J. (2016). *"Latent Class Analysis" in The Sage encyclopedia of social science research methods*. *January*.

Weller, B. E., Bowen, N. K., & Faubert, S. J. (2020a). Latent Class Analysis: A Guide to Best Practice. *Journal of Black Psychology*, *46*(4). https://doi.org/10.1177/0095798420930932

Weller, B. E., Bowen, N. K., & Faubert, S. J. (2020b). Latent Class Analysis: A Guide to Best Practice. *Journal of Black Psychology*, *46*(4), 287–311. https://doi.org/10.1177/0095798420930932

Whitley, E. A., & Collins, H. M. (1991). Artificial Experts: Social Knowledge and Intelligent Machines. *The Journal of the Operational Research Society*, *42*(12). https://doi.org/10.2307/2582960

Woolgar, S. (1985). Why not a sociology of machines? the case of sociology and artificial intelligence. *Sociology*, *19*(4). https://doi.org/10.1177/0038038585019004005

Wurpts, I. C., & Geiser, C. (2014). Is adding more indicators to a latent class analysis beneficial or detrimental? Results of a Monte-Carlo study. *Frontiers in Psychology*, *5*(AUG). https://doi.org/10.3389/fpsyg.2014.00920

Zhang, B., & Dafoe, A. (2020). *U . S . Public Opinion on the Governance of Artificial Intelligence*. 187–193.

Teze Diplomové práce

Jméno a příjmení studujícího: Bc. Michaela Mendelová

Studijní program: Sociologie

Předpokládaný název práce:

V angličtině: Humans, artificial intelligence and sentience

V češtině: Lidé, umělá inteligence a cítění

Klíčová slova: AI, artificial intelligence, sociology of technology, public opinion survey, 2021 AIMS survey

Vedoucí práce: Filip Vostal, PhD

Jméno diplomového semináře, do kterého se chce studující hlásit (předběžně): Diplomový seminář MGRC

Námět práce

According to Mlynář, Alavi, Verma and Cantoni (2018) the development of society always formed social sciences leading them to adjusting their methodological, theoretical and conceptual frameworks to fit new social phenomena. Recently it is for example artificial intelligence. It advanced in recent years to the point where it manifests itself in people's day to day lives and is becoming a part of society in smart home systems, intelligent public displays, autonomous vehicles, chatbots and so on. However, AI is not considered a social or sociological phenomenon in true sense and lacks appropriate conception of this phenomenon, so there is not much of a suitable framework for empirical studies. As always there are a few exceptions. In the eighties author S. Woolgar (1985) proposed his framework named sociology of machines, arguing that AI can and should change the sociological perception that there is something significantly social about human behavior, so we can examine the assumptions of social science that there is huge difference between machines and people – and between human and machine intelligence as well. Almost 20 years after that, Malsch (2001) discussed concept of socionics through which it is possible to link AI and sociology. This field aims to explore the specificities of social life in a modern society, for example the resilience and adaptability of social systems, to incorporate these features into technology and computer programs. According to Mlynář et al. (2018), the most remarkable attempt to incorporate non-human actors into sociology is the actor network theory (ANT) by Bruno

Latour (2005). Underlying massage of this theory is to point out evident similarities of human and non-human actors and technologies and objects in general as partakers in the construction of society which further expands Woolgars'(1985) argument (Mlynář et al., 2018). But according to different authors (Restivo, 2001; Whitley & Collins, 1991) even the idea of an intelligent machine is concerning, because a machine is not a community or a society member, therefore it is inappropriate to call it "intelligent". The reason for that is simple, if we have machines who think, we have to take into consideration that they live with us and share a part of our society. But Restivo (2001) presents that this is the recipe for success, because we can improve the AI with sociological understanding of mentality which means that AI will have to be social and emotional. The idea of socially intelligent robot is nothing new and it's not a sociological invention, it actually started in AI laboratory at MIT (for example Brooks, 1999). Restivo (2001) thinks that the way to intelligent robotic behavior with thinking, emotions and consciousness is through the sociological imagination. To put it in another way, thinking, consciousness and emotions are social constructions that are applicable on non-human agents.

According to Hildt (2019) there is an overall agreement that robots nowadays do not have sentience or consciousness. However, there are some authors (for example Coeckelbergh, 2010; Darling, 2016; Gunkel, 2018) that have argued that it would be beneficial to ascribe rights to robots. One of these authors Darling (2016) argues that it is in accordance with human social values to treat robots more like pets than just like things, which is a claim based on survey on violent behavior toward robots. Hildt (2019) then argues, that while the concrete arguments in favor of ascribing rights to robots vary, these arguments generally center on the social roles people attribute to robots, on the relationships and emotional bonds people build with them, or on the social context in which they interact with each other.
This argument was through one study confirmed as possibly true. According to Lima et al. (2020) the topic of ascribing rights to AI is still a sensitive one in the UE, after the European Parliament proposed that robots advanced enough could be granted "electronic personalities". However, the experiment of these authors offered some interesting results. They collected online users' first impressions of 11 possible rights that could be given to robots and AI in the future and explored whether debunking common misconceptions could modify someone's opinion on the issue. The results show that even though online users generally disapprove of AI and robot rights, they are in favor of protection of these agents from cruelty and cruel treatment. In addition to that, respondent's perceptions became more optimistic when they were confronted with information about myth-refuting statements or rights-bearing non-human entities. Authors repeated this experiment over a representative sample of U.S. residents and found similar results.

Another study unpacking the relationship between humans and AI from the USA was done by Mays et al. (2021), these authors conducted a study with a theme predisposition governing comfort with expanded AI roles in society, which is surprisingly understudied concept, given the topic's relevance to the deployment, design and even regulation of AI systems. Authors carried out a survey of a representative sample of the US population (N = 2254) and conducted mixed-methods analysis based on it. Results showed that there are two different social dimensions to comfort with AI: as a peer and as a superior. However, for both of these dimensions are general and technological efficacy traits like locus of control, robot phobia, perceived technology competence and communication apprehension, strongly associated with acceptance of AI in any roles. Although, generally female and older respondents were less

comfortable with the scenario of AI in various social roles (Mays et al., 2021).
This study is thematically not so close to the data I will be analysing, but the results are a huge inspiration for the hypotheses I will test.

Hildt (2019) also argues that for the debate on the moral and legal status of robots and also for the broader question of how to respond to and interact with machines, a better understanding of artificial consciousness, artificial rationality, artificial sentience, and similar concepts is needed.
I agree with Hildt, but I would also like to note, that it is not just the understanding of researchers that is needed, but also an understanding of ordinary people. Therefore, I would like to conduct an analysis of the recently published data from the AIMS study about artificial intelligence, morality and sentience. These data serve a tracking purpose, where researchers study a change of values in time, which is not my concern. I would like to present a paper about indicators contributing to people viewing AI and robots as more or less sentient and also about indicators which make people more or less sentient towards machines and AI.

Předpokládané metody zpracování

Part of the thesis will be theoretical; it is important to summarize the findings of previous authors who work with the topic of this paper. In the introduction, I will try to explain the focus of the paper and important terminology, then I will introduce some texts on public opinion on this phenomenon and the current social science theory on this topic. Then I will follow up with a larger part, which will be devoted to the analysis of data from the AIMS study from 2021.
According to Sentience Institute's website (Pauketat, 2022b) in November and December of 2021, Sentience Institute conducted a nationally representative survey of 1,232 adults in the USA about social integration, moral consideration and sentience of artificial intelligence. The survey was programmed in GuidedTrack and run online with a sample recruited by Ipsos based on census estimates from the 2019 American Community Survey. The 2021 AIMS survey design, data collection and analysis were done by Janet Pauketat, Jamie Harris, Ali Ladak, and Jacy Reese Anthis. There were two purposes of the study, the first one is to provide a baseline from which to track how the public's opinion on this topic changes over time. The second purpose of this survey is to test the predictions of researchers and forecasters on this topic. The results of the original study are that more American people are uncertain about whether artificial sentience is possible (41.53%) than believe that it is (34.82%) or is not (23.65%) possible. Some Americans think that artificial sentience already exists (18.06%). Overall, this people predict a 59.97% chance that robots or AI will be sentient in the next 100 years. Aggregate responses were within Sentience Institute's 80% credible intervals for 69% (53/77) of the items. They overestimated 6 items and underestimated 18 items.
The reason why I decided to work with existing data is simple, I wanted to examine the thin line between people and machines, preferably AI and I would not be able to conduct a survey of this volume. The offers for students I saw were always quite limited and this data have a relatively big sample and suitable variables. Also, the data originally served a tracking

purpose, so the further analysis I will conduct will hopefully bring some new perspectives and results.

Etické souvislosti zvažovaného projektu

It is a secondary analysis, the data is ready to use, and it is anonymized already, but of course, I will pay attention to not violate applicable GDPR law, although it is a study conducted in the United States of America. Also, data analysis will be conducted in accordance with ethical rules: CASA, ČSA, AAA, EASST, 4S.

Literatura citovaná v projektu

About | Metaculus. (n.d.). Retrieved December 26, 2022, from https://www.metaculus.com/about/

Ashrafian, H. (2014). *AIonAI : A Humanitarian Law of Artificial Intelligence and Robotics*. https://doi.org/10.1007/s11948-013-9513-9

Asimov, I. (n.d.). *Foundation series*. Astounding Science Fiction (Street & Smith), Gnome Press, Spectra, Doubleday.

Bailey, K. D. (1983). Sociological classification and cluster analysis. *Quality & Quantity*, *17*(4), 251–268. https://doi.org/10.1007/BF00167539

Brooks, R. A. (1999). *Cambrian Intelligence: The Early History of the New AI*. MIT Press. https://doi.org/0-262-52263-2

Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*, *13*(2). https://doi.org/10.1007/BF01246098

Chalmers, D. J. (1996). *The Conscious Mind:In search of a fundamental theory*. Oxford University Press.

Chessman, C. F. (2018). Not Quite Human: Artificial Intelligence, Animals, and the Regulation of Sentient Property. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3200802

Clogg, C. C. (1995). *Latent Class Models BT - Handbook of Statistical Modeling for the Social and Behavioral Sciences* (G. Arminger, C. C. Clogg, & M. E. Sobel (Eds.); pp. 311–359). Springer US. https://doi.org/10.1007/978-1-4899-1292-3_6

Coeckelbergh, M. (2010). Robot rights? Towards a social-relational justification of moral consideration. *Ethics and Information Technology*, *12*(3). https://doi.org/10.1007/s10676-010-9235-5

Collins, L. M., & Lanza, S. T. (2010). Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences. In *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. https://doi.org/10.1002/9780470567333

Darling, K. (2016). Extending legal protection to social robots: The effects of anthropomorphism, empathy, and violent behavior towards robotic objects. In *Robot*

Law. https://doi.org/10.4337/9781783476732.00017

Darling, K. (2012). EXTENDING LEGAL PROTECTION TO SOCIAL ROBOTS: THE EFFECTS OF ANTHROPOMORPHISM, EMPATHY, AND VIOLENT BEHAVIOR TOWARDS ROBOTIC OBJECTS. *We Robot Conference 2012*, 24.

Dick, P. K. (1968). *Do Androids Dream of Electric Sheep?* Doubleday. https://doi.org/978-0-575-09418-5

Dillon, W. R., & Kumar, A. (1994). Latent Structure and Other Mixture Models in Marketing: An Integrative Survey and Overview. In *Advanced Methods of Marketing Research* (pp. 295–351). Blackwell.

Dunlap, R. E., Xiao, C., & McCright, A. M. (2010). Politics and environment in America: Partisan and ideological cleavages in public support for environmentalism. *Environmental Politics*, *10*(4), 23–48. https://doi.org/10.1080/714000580

Dunlap, Riley E., & McCright, A. M. (2008). A widening gap republican and democratic views on climate change. *Environment*, *50*(5), 26–35. https://doi.org/10.3200/ENVT.50.5.26-35

El Mesbahi, M. (2015). Human-robot interaction ethics in sci-fi movies: Ethics are not 'there', we are the ethics! *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9186*, 590–598. https://doi.org/10.1007/978-3-319-20886-2_55

European Commission. (2019). *The Ethics Guidelines for Trustworthy Artificial Intelligence*. https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html

EVANS, M. J., GILULA, Z. V. I., & GUTTMAN, I. (1989). Latent class analysis of two-way contingency tables by Bayesian methods. *Biometrika*, *76*(3), 557–563. https://doi.org/10.1093/biomet/76.3.557

Fienberg, S. E., Hersh, P., Rinaldo, A., & Zhou, Y. (2007). *Maximum Likelihood Estimation in Latent Class Models For Contingency Table Data*. http://arxiv.org/abs/0709.3535

Fonseca, J. R. S. (2012). Clustering in the field of social sciences: that is your choice. *International Journal of Social Research Methodology*, 1–26. https://doi.org/10.1080/13645579.2012.716973

Fonseca, J. R. S., & Cardoso, M. G. M. S. (2007). Mixture-model cluster analysis using information theoretical criteria. *Intelligent Data Analysis*, *11*(2), 155–173. https://doi.org/10.3233/ida-2007-11204

Gennaro, R. J. (n.d.). Consciousness. In *Internet Encyclopedia of Philosopfy*. https://iep.utm.edu/consciousness/#H7

Gentle, J. E., Kaufman, L., & Rousseuw, P. J. (1991). Finding Groups in Data: An Introduction to Cluster Analysis. *Biometrics*, *47*(2). https://doi.org/10.2307/2532178

González, A., & Sánchez, J. E. (2009). *the Study of Cells in the Analysis of Contingency Tables From the Viewpoint of Rudas , Clogg and Lindsay Mixture Index of Fit . 3*, 244–249.

GuidedTrack. (n.d.). Retrieved December 26, 2022, from https://www.guidedtrack.com/

Gunkel, D. J. (2018). Robot Rights. *Robot Rights*.

https://doi.org/10.7551/MITPRESS/11444.001.0001

Hagenaars, J. A., & Halman, L. C. (1989). Searching for ideal types: The potentialities of latent class analysis. *European Sociological Review*, *5*(1), 81–96. https://doi.org/10.1093/oxfordjournals.esr.a036509

Hendl, J. (2006). *Přehled statistických metod zpracování dat : analýza a metaanalýza dat* (2nd ed.). Portál. https://www.researchgate.net/publication/47044723_Prehled_statistickych_metod_zpracovani_dat_analyza_a_metaanalyza_dat

Hildt, E. (2019). *Artificial Intelligence : Does Consciousness Matter ? 10*(July), 1–3. https://doi.org/10.3389/fpsyg.2019.01535

Holland, O., & Gamez, D. (2009). Artificial Intelligence and Consciousness. *Encyclopedia of Consciousness*, 37–45. https://doi.org/10.1016/B978-012373873-8.00004-9

IBM. (2019). *SPSS Software | IBM*. IBM Website. https://www.ibm.com/analytics/spss-statistics-software

Jaworska, A., & Tannenbaum, J. (2021). The Grounds of Moral Status. In *Stanford Encyclopedia of Philosophy* (Spring 202). https://plato.stanford.edu/entries/grounds-moral-status/

Kelley, P. G., Heldreth, C., Moessner, C., Sedley, A., Kramm, A., Newman, D. T., & Woodruff, A. (2021). *Exciting , Useful , Worrying , Futuristic : Public Perception of Artificial Intelligence in 8 Countries*. 627–637.

Key figures | Ipsos. (n.d.). Retrieved December 26, 2022, from https://www.ipsos.com/en-us/about-us/key-figures

Kieslich, K., Keller, B., & Starke, C. (2022). Artificial intelligence ethics by design. Evaluating public perception on the importance of ethical design principles of artificial intelligence. *Big Data and Society*, *9*(1). https://doi.org/10.1177/20539517221092956

Killian, M. O., Cimino, A. N., Weller, B. E., & Hyun Seo, C. (2019). A Systematic Review of Latent Variable Mixture Modeling Research in Social Work Journals. *Journal of Evidence-Based Social Work*, *16*(2). https://doi.org/10.1080/23761407.2019.1577783

Kim, S., Roe, D. G., Choi, Y. Y., Woo, H., Park, J., Lee, J. I., Choi, Y., Jo, S. B., Kang, M. S., Song, Y. J., Jeong, S., & Cho, J. H. (2021). Artificial stimulus-response system capable of conscious response. *Science Advances*, *7*(15), 1–9. https://doi.org/10.1126/SCIADV.ABE3996

Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press.

Lazarsfeld, P.F., & Henry, N. W. (1968). *Latent structure analysis*. Houghton Mifflin.

Lazarsfeld, Paul F. (1950). The logical and mathematical foundation of latent structure analysis. *S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star, & J. A. Clausen (Eds.), Studies in Social Psychology World War II: Measurement and Prediction*, *4*, 361–412.

Lima, G., Kim, C., Ryu, S., Jeon, C., & Cha, M. (2020). *Collecting the Public Perception of AI and Robot Rights*. *4*(October).

Magidson, J., & Vermunt, J. K. (2002). Latent class models for clustering: A comparison with K-means. *Canadian Journal of Marketing Research*, *20*(1).

Malsch, T. (2001). Naming the Unnamable: Socionics or the Sociological Turn of/to Distributed Artificial Intelligence. *Autonomous Agents and Multi-Agent Systems*, *4*(3). https://doi.org/10.1023/A:1011446410198

Mandara, J. (2003). The typological approach in child and family psychology: A review of theory, methods, and research. In *Clinical Child and Family Psychology Review* (Vol. 6, Issue 2). https://doi.org/10.1023/A:1023734627624

Martínez, E., & Winter, C. (2021a). Protecting Sentient Artificial Intelligence: A Survey of Lay Intuitions on Standing, Personhood, and General Legal Protection. *Frontiers in Robotics and AI*, *8*(November), 1–7. https://doi.org/10.3389/frobt.2021.788355

Martínez, E., & Winter, C. (2021b). *Protecting sentient artificial intelligence general legal protection*.

Massa, N., Bisconti, P., & Nardi, D. (2022). The Psychological Implications of Companion Robots: A Theoretical Framework and an Experimental Setup. *International Journal of Social Robotics*. https://doi.org/10.1007/s12369-021-00846-x

Mays, K. K., Lei, Y., Giovanetti, R., & Katz, J. E. (2021). AI as a boss ? A national US survey of predispositions governing comfort with expanded AI roles in society. *AI & SOCIETY*, *0123456789*. https://doi.org/10.1007/s00146-021-01253-6

Mccombs, M. (2008). *Setting the agenda: The mass media and public opinion*.

McCutcheon, A. (1987). *Latent Class Analysis*. https://doi.org/10.4135/9781412984713

McLachlan, G., & Peel, D. (2000). *Finite Mixture Models*. John Wiley & Sons, Inc.

Metzinger, T. (2021). Artificial Suffering: An Argument for a Global Moratorium on Synthetic Phenomenology. *Journal of Artificial Intelligence and Consciousness*, *08*(01), 43–66. https://doi.org/10.1142/s270507852150003x

Mlynář, J., Alavi, H. S., Verma, H., & Cantoni, L. (2018). Towards a sociological conception of artificial intelligence. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10999 LNAI*(July), 130–139. https://doi.org/10.1007/978-3-319-97676-1_13

Muthén, B. O., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, *24*(6). https://doi.org/10.1111/j.1530-0277.2000.tb02070.x

Mutz, D. C., & Soss, J. (1997). Reading Public Opinion : The Influence of News Coverage on Perceptions of Public Sentiment. *The Public Opinion Quarterly*, *61*(3), 431–451. http://www.jstor.org/stable/2749580

Nylund-Gibson, K., & Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science*, *4*(4). https://doi.org/10.1037/tps0000176

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *14*(4). https://doi.org/10.1080/10705510701575396

Ouchchy, L., Coin, A., & Dubljević, V. (2020). AI in the headlines : the portrayal of the ethical issues of artificial intelligence in the media. *AI & SOCIETY*, *0123456789*. https://doi.org/10.1007/s00146-020-00965-5

Pauketat, J. (2022a). *Sentience Institute | Artificial Intelligence, Morality, and Sentience (AIMS) Survey: 2021*. https://www.sentienceinstitute.org/aims-survey-2021

Pauketat, J. (2022b, March 23). *Sentience Institute | Data Publication for the 2021 Artificial Intelligence, Morality, and Sentience (AIMS) Survey*. https://www.sentienceinstitute.org/blog/data-publication-AIMS-survey

Pearson, K. (1904). *On the theory of contingency and its relation to association and normal correlation*. Dulau and Co.

Petersen, K. J., Qualter, P., & Humphrey, N. (2019). The application of latent class analysis for investigating population child mental health: A systematic review. In *Frontiers in Psychology* (Vol. 10, Issue MAY). https://doi.org/10.3389/fpsyg.2019.01214

Reddy Nadikattu, R. (2016). the Emerging Role of Artificial Intelligence in Modern Society. *International Journal of Creative Research Thoughts*, *4*(4), 2320–2882. www.ijcrt.orgwww.ijcrt.org

Restivo, S. (2001). Bringing up and booting up: Social theory and the emergence of socially intelligent robots. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, *4*, 2110–2117. https://doi.org/10.1109/icsmc.2001.972867

Riedl, M. O. (2019). Human-centered artificial intelligence and machine learning. *Human Behavior and Emerging Technologies*, *1*(1). https://doi.org/10.1002/hbe2.117

Safron, A. (2021). *Integrated World Modeling Theory (IWMT) Expanded: Implications for Theories of Consciousness and Artificial Intelligence*. *June*. https://psyarxiv.com/rm5b2/

Scheutz, M. (2011). The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots. *Robot Ethics: The Ethical and Social Implications of Robotics*, *November*, 205–221. http://books.google.com/books?hl=en&lr=&id=oBb-lt3l4oYC&oi=fnd&pg=PA205&dq=The+Inherent+Dangers+of+Unidirectional+Emotional+Bonds+between+Humans+and+Social+Robots&ots=yuaSs8y33n&sig=h5KrpJxIcyk5RT2vih0YpgSHqNY

Searle, J. R. (1980). Minds, brains, and programs. *THE BEHAVIORAL AND BRAIN SCIENCES*, *3*, 417–456.

Seol, H. (2022). *snowRMM: Rasch Mixture Model for jamovi. [jamovi module]*. https://github.com/hyunsooseol/snowRMM

Seth, A. K. (2016, November 2). *The hard problem of consciousness is a distraction from the real one | Aeon Essays*. https://aeon.co/essays/the-hard-problem-of-consciousness-is-a-distraction-from-the-real-one

Shanahan, L., Copeland, W. E., Worthman, C. M., Erkanli, A., Angold, A., & Costello, E. J. (2013). Sex-differentiated changes in C-reactive protein from ages 9 to 21: The contributions of BMI and physical/sexual maturation. *Psychoneuroendocrinology*, *38*(10). https://doi.org/10.1016/j.psyneuen.2013.04.010

Technological determinism - Oxford Reference. (n.d.). Retrieved December 29, 2022, from https://www.oxfordreference.com/display/10.1093/oi/authority.20110803102813253;jses

sionid=6BC43215720200147D492ED1B84844A8

The jamovi project (jamovi (Version 2.3) [Computer Software]). (2021). https://www.jamovi.org

Ulbricht, C. M., Chrysanthopoulou, S. A., Levin, L., & Lapane, K. L. (2018). The use of latent class analysis for identifying subtypes of depression: A systematic review. In *Psychiatry Research* (Vol. 266). https://doi.org/10.1016/j.psychres.2018.03.003

Van Gulick, R. (2022). Consciousness. In *The Stanford Encyclopedia of Philosophy* (Winter 202). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2022/entries/consciousness/

Vermunt, J. K. (2002). Latent class analysis of complex sample survey data: Application to dietary data. *Journal of the American Statistical Association*, *97*(459), 736–737.

Vermunt, Jeroen K, & Magidson, J. (2016). *"Latent Class Analysis" in The Sage encyclopedia of social science research methods*. *January*.

Weller, B. E., Bowen, N. K., & Faubert, S. J. (2020a). Latent Class Analysis: A Guide to Best Practice. *Journal of Black Psychology*, *46*(4). https://doi.org/10.1177/0095798420930932

Weller, B. E., Bowen, N. K., & Faubert, S. J. (2020b). Latent Class Analysis: A Guide to Best Practice. *Journal of Black Psychology*, *46*(4), 287–311. https://doi.org/10.1177/0095798420930932

Whitley, E. A., & Collins, H. M. (1991). Artificial Experts: Social Knowledge and Intelligent Machines. *The Journal of the Operational Research Society*, *42*(12). https://doi.org/10.2307/2582960

Woolgar, S. (1985). Why not a sociology of machines? the case of sociology and artificial intelligence. *Sociology*, *19*(4). https://doi.org/10.1177/0038038585019004005

Wurpts, I. C., & Geiser, C. (2014). Is adding more indicators to a latent class analysis beneficial or detrimental? Results of a Monte-Carlo study. *Frontiers in Psychology*, *5*(AUG). https://doi.org/10.3389/fpsyg.2014.00920

Zhang, B., & Dafoe, A. (2020). *U . S . Public Opinion on the Governance of Artificial Intelligence*. 187–193.

Orientační seznam literatury

1. Ashrafian, H. (2014). *Artificial Intelligence and Robot Responsibilities :* https://doi.org/10.1007/s11948-014-9541-0
2. Ashrafian, H. (2014). *AIonAI : A Humanitarian Law of Artificial Intelligence and Robotics*. https://doi.org/10.1007/s11948-013-9513-9
3. Baum, S. D. (2017). Social choice ethics in artificial intelligence. *AI & SOCIETY*. https://doi.org/10.1007/s00146-017-0760-1
4. Brooks, R. A. (1999). Cambrian Intelligence: The Early History of the New AI. MIT Press. https://doi.org/0-262-52263-2

5. Coeckelbergh, M. (2010). Robot rights? Towards a social-relational justification of moral consideration. Ethics and Information Technology, 12(3). https://doi.org/10.1007/s10676-010-9235-5
6. Darling, K. (2016). Extending legal protection to social robots: The effects of anthropomorphism, empathy, and violent behavior towards robotic objects. In Robot Law. https://doi.org/10.4337/9781783476732.00017
7. Gunkel, D. J. (2018). Robot Rights. Robot Rights. https://doi.org/10.7551/MITPRESS/11444.001.0001
8. Hildt, E. (2019). *Artificial Intelligence : Does Consciousness Matter ? 10*(July), 1–3. https://doi.org/10.3389/fpsyg.2019.01535
9. Chessman, C. F. (2018). Not Quite Human: Artificial Intelligence, Animals, and the Regulation of Sentient Property. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3200802
10. Kelley, P. G., Heldreth, C., Moessner, C., Sedley, A., Kramm, A., Newman, D. T., & Woodruff, A. (2021). *Exciting , Useful , Worrying , Futuristic : Public Perception of Artificial Intelligence in 8 Countries*. 627–637.
11. Latour, B. (2005). Reassembling the Social: An Introduction to Actor-Network-Theory. Oxford University Press.
12. Lima, G., Kim, C., Ryu, S., Jeon, C., & Cha, M. (2020). *Collecting the Public Perception of AI and Robot Rights*. *4*(October).
13. Martínez, E., & Winter, C. (2021). *Protecting sentient artificial intelligence general legal protection*.
14. Malsch, T. (2001). Naming the Unnamable: Socionics or the Sociological Turn of/to Distributed Artificial Intelligence. Autonomous Agents and Multi-Agent Systems, 4(3). https://doi.org/10.1023/A:1011446410198
15. Mays, K. K., Lei, Y., Giovanetti, R., & Katz, J. E. (2021). AI as a boss ? A national US survey of predispositions governing comfort with expanded AI roles in society. *AI & SOCIETY*, *0123456789*. https://doi.org/10.1007/s00146-021-01253-6
16. Mlynář, J., Alavi, H. S., Verma, H., & Cantoni, L. (2018). Towards a sociological conception of artificial intelligence. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10999 LNAI*(July), 130–139. https://doi.org/10.1007/978-3-319-97676-1_13
17. Ouchchy, L., Coin, A., & Dubljević, V. (2020). AI in the headlines : the portrayal of the ethical issues of artificial intelligence in the media. *AI & SOCIETY*, *0123456789*. https://doi.org/10.1007/s00146-020-00965-5
18. Restivo, S. (2001). Bringing up and booting up: Social theory and the emergence of socially intelligent robots. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, *4*, 2110–2117. https://doi.org/10.1109/icsmc.2001.972867
19. Roese, N. J., & Amir, E. (2009). *Human- Android Interaction in the Near and Distant Future*. *4*(4), 429–434.
20. Whitley, E. A., & Collins, H. M. (1991). Artificial Experts: Social Knowledge and Intelligent Machines. The Journal of the Operational Research Society, 42(12). https://doi.org/10.2307/2582960
21. Woolgar, S. (1985). Why not a sociology of machines? the case of sociology and artificial intelligence. Sociology, 19(4). https://doi.org/10.1177/0038038585019004005
22. Zhang, B. (2021). *Public Opinion Toward Artificial Intelligence*. 1–21.

23. Zhang, B., Anderljung, M., Kahn, L., Dreksler, N., Horowitz, M. C., & Dafoe, A. (2021). *Ethics and Governance of Artificial Intelligence : Evidence from a Survey of Machine Learning Researchers*. *71*, 591–666.
24. Zhang, B., & Dafoe, A. (2020). *U . S . Public Opinion on the Governance of Artificial Intelligence*. 187–193.

Data

Pauketat, Janet; Ladak, Ali; Harris, Jamie; Anthis, Jacy (2022), "Artificial Intelligence, Morality, and Sentience (AIMS) 2021", Mendeley Data, V1, doi: 10.17632/x5689yhv2n.1

Seznam příloh

Příloha č.1: Tabulky pro analýzu (tabulky v souboru Excel)