Emotional artificial intelligence and its social and ethical implications in Japan: A mixed method study

Author: Ho Manh Tung (ID: 61120002)

Supervisor: Peter Mantello

Doctoral Program

Graduate School of Asia Pacific Studies

Ritsumeikan Asia Pacific University

Beppu City, Oita Prefecture, Japan

February 20th , 2023

Contents

List of Tables

List of Figures

Acknowledgements

I would like to express my deepest gratitude to the people who have supported me in my journey toward the completion of my doctoral degree. To be able to study the impacts of artificial intelligence technologies for my Ph.D. thesis is really a dream come true for me. I thank Prof. Mantello Peter, my supervisor, for introducing me to the subject. Prof. Mantello's knowledge, generosity, enthusiasm, and curiosity have been an endless source of inspiration for me to go outside of my comfort zone and further develop my research skills. I would also like to thank Prof. Quan-Hoang Vuong, my mentor at the Centre for Interdisciplinary Social Research, Phenikaa University, without his support and guidance from the beginning, I would not be able to start, let alone, complete this journey.

I would like to thank my wife, Nguyen To Hong Kong, for supporting me all these years and helping me with editing my writing and sharpening my arguments in many research articles as well as my thesis. Being the first member of my family to achieve a doctoral degree produces a wonderful, proud feeling. For that, I would like to thank my parents for giving me endless support and encouragement. My colleagues and friends who cross my path during the last five years in graduate school also play a very important role in challenging and guiding me whenever I need them. For that, I thank Ho Manh Toan, Nguyen Minh Hoang, and Le Bui Ngoc Thang for their support.

I would also like to thank Prof. Andrew McStay and Prof. Vian Bakir and the UK team for creating the foundation for interdisciplinary social research regarding emotional AI technologies. Without them spearheading the way, I would not have been on this path. Finally, I thank Prof. Nader Ghotbi for his kindness and encouragement. Finally, I would like to thank the SGH Scholarship Foundation and APU Tuition Reduction Scholarship Fund for providing much-needed support during my study. In conclusion, I am grateful to everyone who has contributed to this thesis, directly or indirectly. Your support, encouragement, and guidance have been essential in helping me to complete this challenging and rewarding journey. Thank you.

Abstract

Emotional artificial intelligence (AI) is a narrow, weak form of AI systems that reads, classifies, and interacts with human emotions. This form of smart technology has started to become an integral layer of our digital and physical infrastructures. Thus, it will radically transform how we live, learn, and work in the years to come. This study is the first in the literature to bring the Technological Acceptance Model and the Moral Foundation Theory together under the analytical Three-pronged Approach (Contexts, Variables, and Statistical models) to study determinants of emotional artificial intelligence's acceptance in 10 different use cases in Japan. The statistical models in this study have successfully accounted for an average of 52.11% of the variation in the data (min = 38%; max = 67.8%). In the most successful case of statistical modeling, the case of Home Robots, our model accounts for 67.8% of the variation in the data, outperforming past models in the literature. Across all cases, we find women are more concerned about key ethical issues of emotional AI: algorithmic biases, data privacy, loss of autonomy, etc. Moreover, we find age is a negative correlate of attitude toward emotional AI applications, suggesting more public outreach efforts are needed to promote AI solutions for the elderly population—a major beneficiary of emotional AI technologies in the rapidly aging Japanese society. Interestingly and paradoxically, in many cases, accuracy concern, data management concern, and bias concern are found to be either non-significant or positively correlated with attitude toward emotional AI. These results suggest a willingness to adopt emotional AI applications despite its potential flaws and

muddy issues around data management. This attitude relaxes the concern that many technologists have raised over the hesitance of AI adoption due to its failure would be more psychologically jarring and salient. Yet, these results are worrying given the increasing number of immigrant workers and the lack of women in key decision-making positions in Japan. Based on the empirical findings, the final chapter provides seven lessons on algorithmic governance and AI ethics. Finally, the thesis calls for the development of theoretical frameworks that capture cross-cultural nuances in moral reasoning about effects of technologies on our daily lives for a better understanding of human-machines relationship in an era interactive AI.

Key words: emotional AI; technological acceptance model; Moral Foundation Theory; AI ethics

Chapter 1: Emotional AI in Society: Definitions, Applications, and Problems

1.1. Introduction

In the Hollywood blockbuster film, Ad Astra, celebrated astronaut Roy McBride (played by Brad Pitt) is charged with a mission to travel into deep space to find his long, lost father whose experiments to find intelligent life are causing a series of cataclysmic weather changes threatening to destroy Earth. Roy's fame aside from his relationship to his acclaimed yet negligent father lies with his stoic professionalism to remain calm under pressure. For the majority of the film, regardless of the mounting obstacles placed in his path, Roy remains a blank, emotionless canvas, cracking only slightly whenever his father is mentioned. Whether it is surviving a meteor shower, fending off murderous laboratory baboons, or surviving hand-to-hand combat with space pirates, Roy is always at the top of his game, in control.

Indeed, emotions take center stage in the story, not only in shaping conflict between one human to another but importantly, human to machine. But just before his final jump into deep space Roy's emotional body armor unravels during an AI-administered health check. When asked questions about his father, the AI sees what his human supervisors cannot - it can understand his emotional state by sensing his non-conscious body data. The AI's statistical reasoning concludes that in all probability his emotional state compromises the mission. And so, it disqualifies him from continuing on. Yet although AI can feel what Roy feels, it cannot sympathize or take exigent circumstances into account. It can only translate his non-conscious body data into an objective, quantifiable assessment.

While the scene in Ad Astra is set in the distant future, emotional artificial intelligence (EAI) is already here and changing how we live and work. Simply put, EAI is the ability of machines, sensors and devices to gauge, learn, interact, sense and simulate human emotions by reading a person’s biometric signs. While in the past, the focus is the reading of the exterior body, it now includes body language, gestures, skin conductance levels, blood pressure, eye movements, voice tone, respiration and heart rate (McStay, 2018).

According to Kate Crawford (2021), the industry centered around emotional AI is now worth around $22 billion, expected to double by 2024. The range of applications for the technology is expanding. For example, the music app, Spotify can suggest playlists by sensing a person’s mood based on previous playlist configurations. Amazon’s home assistant Alexa’s voice analytics can read the emotional state of its user and temper its responses (Richardson, 2020). Honda has created what they call the ‘Emotion Engine’ which embedded sensors in the car to detect whether a driver is angry or happy, alert or distracted, calm or stressed as well as offering personalized driving tips and health alerts. NEC, a Japanese security conglomerate, has developed software for McDonald's that measures customer sentiments as they are looking at digital menus to optimize the customer’s experience while at the same time increasing sales. In South Korea, nearly 25% of the top 131 corporations stated they were planning to facilitate their recruitment with emotional AI tools (Condie & Dayton, 2020). The Boston start-up Cogito and Japanese company Empath build voice recognition software for call center managers to monitor in real-time employees’ moods. The US company, Spot, markets an AI chatbot that uses natural language processing to identify patterns and problems associated with workplace harassment.

As the world is moving into a future of smart cities, internet of things, and ubiquitous computing, emotional AI is, thus, going to be an integral layer of our life, especially. How do we live well and ethically with machines that feels and feeds off our emotions? Answering this question depends on making sense of our perceptions of the technology and its impacts on our life. What are the benefits gained from interacting with machines that can sense our emotions? What are the risks involved? What does co-existing with emotional AI mean for values that we cherish: privacy, autonomy, fairness, trust, etc.? How do we begin to parameterize the ethics of emotional AI?

This study asks these questions in the context of Japan. While in the West, collective imagination of AI is often associated an innately dystopian form of synthetic intelligence such as Colossus: The Forbin Project (1970), War Games (1983), Terminator (1984), I, Robot (2004), and Ex Machina (2014), in Japan, emotional AI is associated in with beloved manga/animation characters such as Mighty Atom (Astro Boy) or Doraemon (Robertson, 2017). Robots that are designed to evoke feelings in humans and have some basic responses to human emotions have been introduced in Japan for decades. For example, in 1999, Sony introduced the world's first robot dog companion, AIBO, and recently in 2018, AIBO was reintroduced with an upgraded AI software that infused it with a "lovable quality". Or Fujisoft's humanoid robot, PALRO, which is capable of communicating with human through voice and can even remember faces of over 100 people, has been shown to reduce anxiety and stress in dementia patients and can encourage people with dementia to interact with others in seniors care facilities (Inoue et al., 2014).

As much as Japan is known for being home to futuristic, advanced technological innovations, when it comes to technological adoption, Japanese society is also saddled with many contradictions arising from the conflicts between two sides. One is the modern, industrialized way of life that prioritizes individualism, consumerism, flexibility, and productivity. The other is Japan's strong collectivist culture centered around the tradition of Bushidō (武士道)一way of the warrior, whose code of conducts demands meticulousness, unconditional loyalty, daily devotion, respect of hierarchy, and righteousness. For instance, even amid the height of the COVID-19 pandemic, when 80% of companies promote remote work, many were still reluctant to let go of the long-standing paper-based process with the stamping of the personal 'hanko' (seal) to signify approval and authority (Shoji, 2020). The contrast in Japan between its futuristic innovations and the greyest population holding on symbols of their conservative, traditional culture highlights the uniqueness of Japan as a context for studying technological adoption behaviors.

According to recent statistics, the AI market in Japan is estimated to be worth 3.7 trillion yen and to reach 87 trillion by 2030 (Ishii et al., 2020). As of 2020, Japan is home to roughly 200-300 AI-related companies and is the third major player in AI research and development after the US and China (Dirksen, 2020). In 2018, the number of domestic patent applications for AI-integrated inventions was approximately 4700, which was a massive increase of 54% from the previous year. For core AI technology, the number of applications was around 1500, an increase of 65% from 2017. These inventions have been surging since 2014, and half of them focused on deep learning (METI, 2020). Importantly, advanced deep learning systems are now beginning to be deployed in all areas of life in Japan from security cameras, job screening, workers monitoring, companion robots, student learning, interior sensing in cars, etc. In most cases, it is the current and forecasted shortage of labor force that provides the rationale for the adoption of AI and emotional AI in Japan.

A striking example is a case of security cameras with AI capabilities to detect emotions, movements, and facial identification, to help combat social problems arising from the rapid aging of the Japanese population. It is reported in a heart-wrenching New York Times article that around 17,000 people with dementia went missing in 2020, up from 9,600 in 2012, which makes dementia the leading cause of missing-person cases in Japan (Dooley & Ueno, 2022). Per the market research company Fuji Keizai's estimations, the Japanese domestic market for commercial security cameras is expected to grow from ¥56.3 billion in 2020 to ¥61.9 billion in 2024, and the above-mentioned AI capabilities are the main selling point for this product. AI-powered surveillance cameras, those that could detect real-time movements and emotions, are seen as an inevitable solution for the supervision of elderly people suffering from dementia.

In the education sector, given the reality of teachers being overworked and the labor market facing a shortage, AI technologies are considered vitally important in enhancing the effectiveness of teaching and learning. Ryo Uchida, professor of sociology of education at Nagoya University, and colleagues found over 70% of junior high school teachers in Japan have been overworked by 80 hours each month, which meets the technical threshold for determining death by overwork (過労死一karoshi) (Lee, 2022; Matsushita & Yamamura, 2022). Thus, a push toward AI adoption in schools is taking place in Japan with the hope that smart technologies will help by identifying more effective methods and areas to focus on for the students, thus shortening the time for training and quickly preparing the students to enter the workforce. Such logic has been epitomized in the 2019 "AI Quest" initiative, launched by the Ministry of Economy, Trade and Industry, to achieve goals such as “resolving the shortage of AI human resources” and “developing AI human resources (METI, 2022).”

These examples highlight just how important the role advanced technology will play in Japanese society. The Japanese national strategy of AI is centered around the concept of Society 5.0, which is defined as “A human-centered society that balances economic advancement with the resolution of social problems by a system that highly integrates cyberspace and physical space”(Cabinet Office Japan, 2019). In the national strategy of AI development, the Cabinet Office of Japan proposed seven social principles of human-centric AI: the human-centric AI principle the principle of education/literacy; the principle of privacy protection; the principle of ensuring security; the principle of fairness, accountability, and transparency; the principle of innovation. These principles are based on the basic philosophy that AI development and deployment must respect three values: 1) Human Dignity; 2) Diversity and Inclusion; 3) Sustainability (Cabinet Office of Japan, 2019). Indeed, more effective adoption of smart technologies that respect universal values seems to be the only way forward, not only for Japan but many societies.

Yet, from a social scientific perspective, the adoption of emotional AI in Japanese society is intriguing on many fronts. On the one hand, there is a clear rationale for the adoption of the technologies given the mounting pressures from the demographic conditions of Japan. On the other hand, while it is true that Japan has always been considered a technological powerhouse of the world’s economy when it comes to technological adoption, the country has always been known for many contradictions stemming from the conflicts between traditional culture, stiff in a male-dominant hierarchy and collectivism, with the modern, industrialized way of life that prioritizes individualism, consumerism, and productivity. For example, a classic text on contemporary Japanese culture, *The Japanese mind*, opened with the concept of Aimai (曖昧), or Ambiguity defined as “a state in which there is more than one intended meaning, resulting in obscurity, indistinctness, and uncertainty” (Osamu, 2002, p. 9). This cultural value is clearly at odds with the value presumed by emotional AI technologies, i.e., making emotions more transparent and visible to other people.

How do Japanese people view emotional AI, its utilities, and threats? How will these cultural tensions be resolved? Thus, this study sets out to explore various socio-demographic and behavioral determinants constitute the Japanese perceptions and acceptance behaviors regarding emotional AI applications.

This chapter introduces what constitutes emotional AI and how emotional AI technologies are being deployed in various sectors. Subsequently, it identifies concrete research problems related to emotional AI technologies. In Chapter 2, we will conduct a review of the current literature focusing on two key areas: 1) the social studies of emotional AI; 2) the studies of technological acceptance behaviors, concentrating on AI-driven technologies such as robotics, smartwatches, smart homes, etc. Chapter 3 will explain the mixed method approach deployed in this thesis to study the attitude toward emotional AI applications. The quantitative analysis and interpretation of data collected during this study draw from two prominent frameworks: Technological Acceptance Model (Davis, 1989) and the Moral Space of how humans judge machine (Hidalgo et al., 2021). The qualitative analysis of the data collected via interviews and focus groups is based on qualitative coding method (Braun & Clarke, 2006).

In the subsequent chapters (Chapter 4 to 9), insights from both the quantitative and qualitative analyses are presented. The results are organized into five contexts of uses: workplace, security, politics, education, and healthcare. Chapter 10 will focus on the theoretical and empirical contributions of the study as well as discussion of future research directions. The final chapter,

Chapter 11, will provide concrete policy recommendations of based on the empirical insights provided throughout the previous chapters as well as concluding thoughts on the philosophy and ethics of emotional AI.

1.2. What is emotional AI?

The state-of-the-art emotional AI technologies are grounded in the pioneering work of MIT Media Lab's Rosalind Picard, who coined the term 'affective computing,' which is now a growing multidisciplinary field that draws on computer sciences, engineering, psychology, physiology, philosophy, and even neuroscience. Originally, Rosalind Picard used the term to mean "computing that relates to, arises from, and deliberately influences emotion" (Calvo et al., 2015, p. 13). Now, the term 'emotional AI' designates a branch of affective computing that combines artificial intelligence, biosensors, and deep learning algorithms to sense, track, and classify human emotions and affective states (McStay, 2018). Figure 1.1 presents an emotional AI product, co-developed by Nippon Electric Corporation and RealEyes, an UK start-up (NEC Press Release, 2020). This AI system tracks in real-time emotional states of video conference participants, e.g., attention, relaxation, and confusion, and produce visualizations of their changes during different stages of a conference.

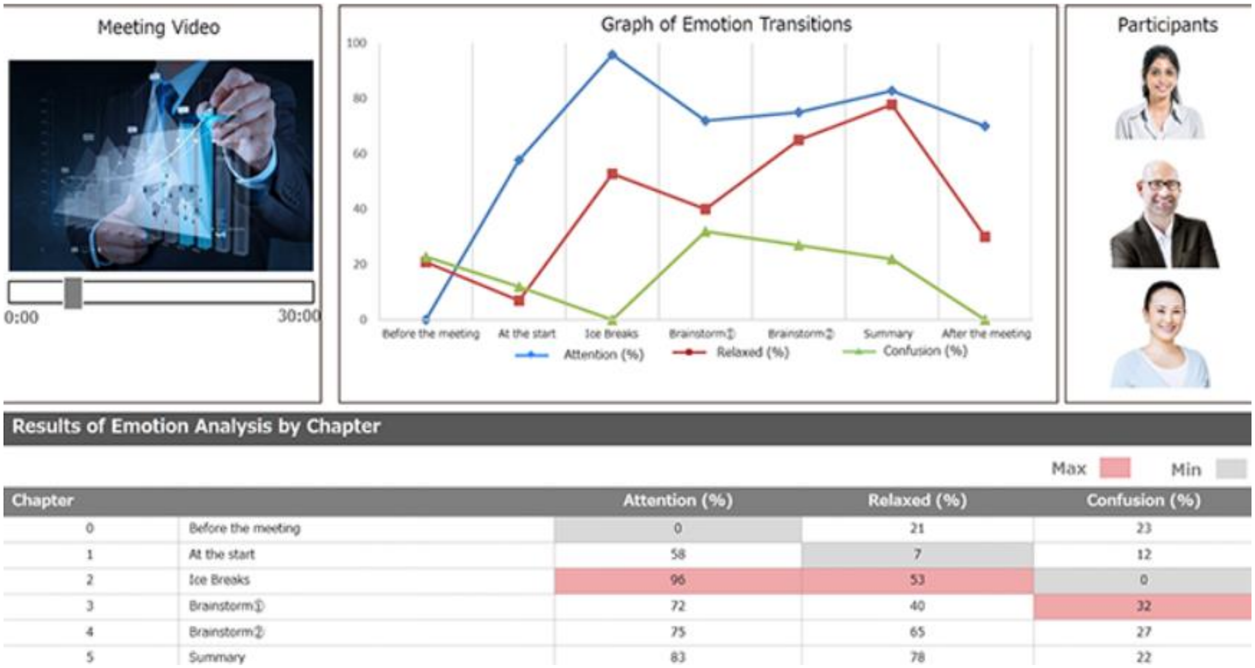


Figure 3.1. An example of data visualization by an emotional AI system co-developed by NEC and RealEyes. Source: NEC website.

Thus, in philosophical terms, emotional AI is a form of weak, narrow AI systems that use an ensemble of methods from machine learning, knowledge-based approach to teach machines to read, categorize, and react to human emotions. Figure 1.2 presents a categorization of different kinds of AI systems. Broadly speaking, AI systems can be categorized in terms of their capabilities and what techniques are being used to construct these systems.

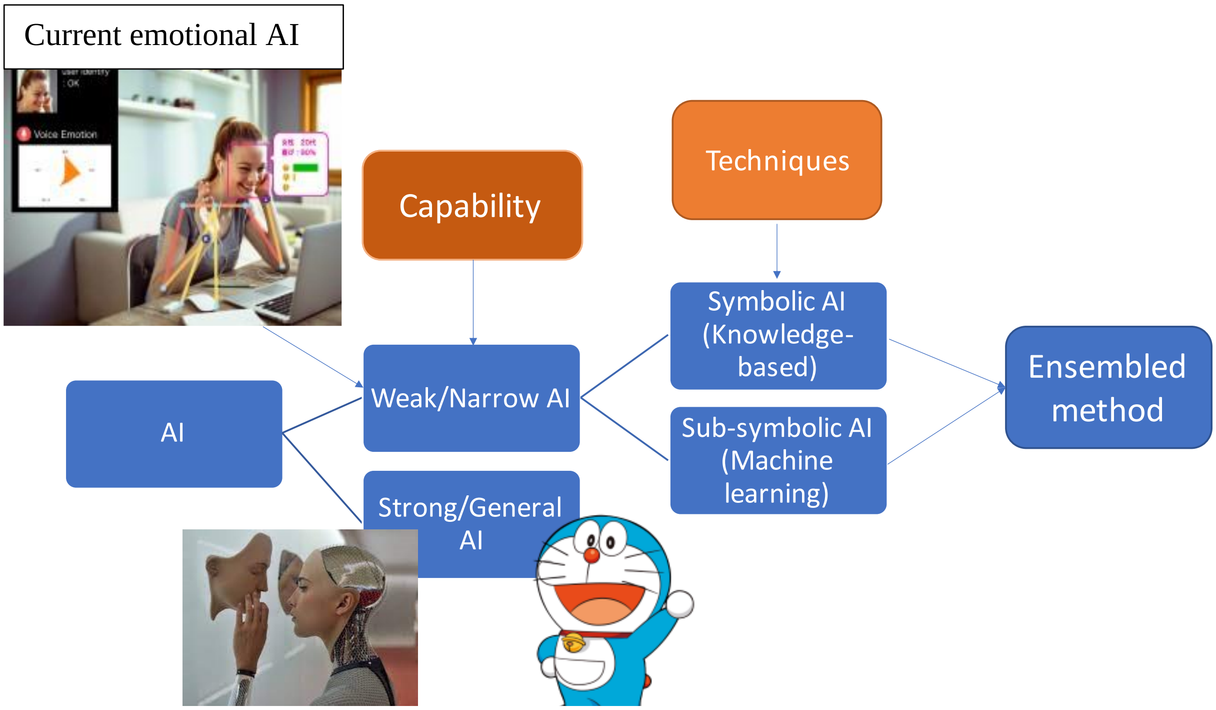


Figure 3.2. Categorizations of AI-systems and where current existing emotional AI systems are situated.

The General vs. Narrow AI distinction is popularized in the writing of Nick Bostrom (2017). Meanwhile, the distinction between Strong vs. Weak AI is first developed in the famous Chinese Room thought experiment of philosopher John Searle (Searle, 1980), where he made the case that a machine can only appear to have intelligence without having any real understanding of the world and its action. Strong and General AI systems do not exist in real life, and many experts predict they require unimaginable breakthroughs in theoretical and technical knowledge. In the fictional world, characters such as *Doraemon*, *Astro Boy*, *Terminator*, *Eva* in *Ex Machina*, *Ultron* in *Iron Man*, etc. can be considered as having achieved strong and general intelligence. These machines have consciousness and can apply their knowledge to new situations without being preprogrammed.

In contrast, existing emotional AI systems are narrow in the sense that it is limited by what has been programmed into them. For example, if an emotional AI system is trained to only identify anger, they will not be able to identify other emotions. In other words, its ability for recognizing anger cannot be generalized into the ability to recognize sadness or joy, for example. Furthermore, emotional AI is weak in the sense that it has no subjective awareness of the emotions it is trained to recognize. For example, the machine might correctly identify anger but has no self-awareness or subjective experience of anger itself, and their ability is limited to only what is programmed in them.

1.3. Current progress and applications of emotional AI systems

1.3.1. Applications: A growing global industry of more than USD21 billion

As emotional AI has huge potential for commercialization, and under the current ubiquity of a cheap data storing and computing solutions, emotional AI research has been progressing quickly, resulting in their commercial applications form a global industry worth USD21.6 billion and expected to double in value by 2024 (Crawford, 2021). In recent years, all these pioneers of affective computing have found ways to monetize their research on emotion-sensing AI systems. For example, Rosalind Picard, MIT Media Lab, founded *Affectiva* and *Empatica* to commercialize a range of emotional AI tools that can detect in real-time via data captured through wearable devices stress (Sano et al., 2018), frustration (Klein et al., 2002), even suicidal thoughts (Kleiman et al., 2018). Bjoern Schuller, from Imperial College London, co-founded *audEERING*, which manufactures emotion-sensing devices for audio media. Erik Cambria, Nanyang University of Technology, co-founded *SenticNet* which applies state-of-the-art sentiment analysis software for marketing. The following provides a description of current prominent emotional AI applications by sector. Below are some of the outstanding examples of emotional AI applications worldwide.

In the field of education, smartphone apps such as *ClassDojo* provide teachers with psychometric profiles of students, allowing them to score and reward positive behavior while giving a lesson (Williamson, 2017b, 2021). The smart toy Moxie, meanwhile, assists with a child's emotional, social and psychological development (Lyles, 2020). In the world of entertainment, Spotify's emotion recognition algorithms can suggest playlists by sensing a person's mood. And in response to an increase in elderly driver car accidents in Japan, Honda and Softbank's co-created a bio-sensing 'Emotion Engine' which detects if a driver is drowsy, distracted, or stressed (Dery, 2018).

In the workplace, emotional AI is an emerging layer in 'human-centric', automated management systems and data-driven corporate wellness programs (Brassart Olsen, 2020). For example, legacy companies such as IBM, Softbank, and Unilever are now using affect tools for recruitment, as well as monitoring the productivity of their workforce. Emotional AI companies such as *Empath* in Japan are marketing voice-analytic software that allows call center managers to monitor employee moods in real-time. For cash-strapped start-ups interested in low-cost wellness initiatives, Amazon provides a wrist wearable biosensor called 'Halo' that tracks a user's emotional state, detecting depression, anxiety, and even early signs of mental illness (Graziosi, 2020; Lecher, 2019). Similarly, the UK company, Moodbeam markets what they call a 'mood awareness' wearable that monitors an employee's 'happiness' level and then shares the daily findings with co-workers (Bearn, 2021). And in response to spiraling stress-related absenteeism at work caused by COVID, Microsoft has announced plans to use emotion-sensing devices in their worldwide offices to track wellbeing (Spataro, 2020).

As for security applications, the Russian security company ELSYS is making and selling a facial recognition camera system, called Vibraimage, to global sporting events that allegedly ‘predicts’ criminal intention and various emotional states such as neuroticism, depressions, etc. by analyzing a person’s gait, head and eye movements, and facial expressions (Wright, 2021). Vibraimage products have been used in Russian airports, Russian and Japanese nuclear powerplants as well as convenience and retail stores in Japan (Kobata, M., spokeperson of ELSYS Japan, personal communication, 2021). ELSYS’ technologies are said to be developed into lie-detecting devices by the South Korean Police and are currently being used to surveil the Uighur population in China (Wright, 2021).

In healthcare, emotional AI products are also considered to be a cheap, supplemental solution for mental health screening as well as digital medical counselling. For example, the UK’s National Health Service is investing in AI conversational agents such as *Wysa* to provide online medical counselling (Adikari et al., 2022). In Singapore, the social service agency Lion Befrienders has developed facial recognition software to provide early detection of depression, anxiety, cognitive decline in senior citizens (Menon, 2021).

1.3.2. Progress: Exponential growth toward multimodal processing of emotions

Extracting records of 3,386 publications with the key words ‘affective computing’ since 1995 in the Web-of-Science database, the study finds a high growth rate of 12% (Ho et al., 2021). The exponential growth of the field is evident in the fact that the number of publications within the recent 5 years outstrip the publications of the previous 20 years period between 1995 to 2015.

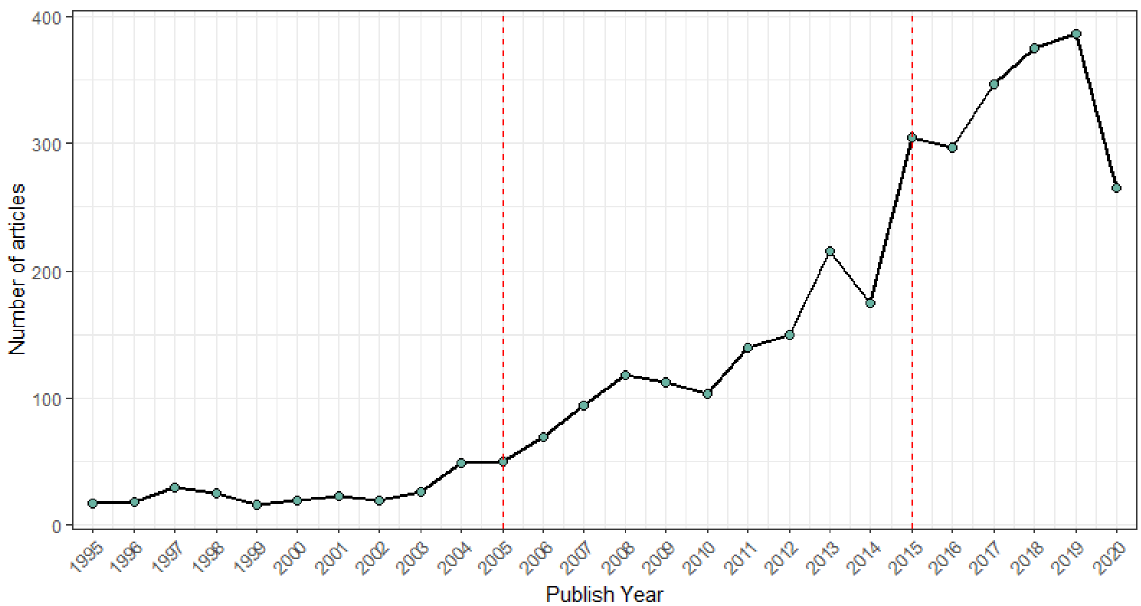


Figure 3.3. Annual scientific production on "affective computing", 1995-2020 (Source: Web of Science).

There are a few notable developments in the field of affective computing and sentiment analysis in recent years. The latest trend, sentic computing, pioneered by Singapore-based SenticNet, features the ensemble of symbolic AI (knowledge-based methods which are strongly semantic) and sub-symbolic AI (machine-learning methods, neural networks and deep learning algorithms, which take advantage of the rise of computing power, availability of data and cloud storage possibilities) (Susanto et al., 2021). This hybrid style combines the top-down approach, which leverages symbolic models and knowledge bases (i.e., logical and semantic networks), and the bottom-up approach, which leverages advanced statistical NLP on large, labeled datasets. Cambria et al. (2020) found that this hybrid approach outperforms both symbolic representations and statistical methods alone, and Basiri et al. (2021) found that their proposed attention-based deep models for sentiment analysis of short and long texts reach state-of-the-art levels of performance on multiple benchmark datasets. The authors contend that the coupling of symbolic and sub-symbolic AI is a step further in the path from mere NLP to natural language *understanding* (Cambria et al., 2020).

Another important trend is that, rather than the discrete categorization of emotions, continuous modeling based on dimensions (valence, arousal, and dominance) is gaining (Schuller & Schuller, 2018). In this trend, the ensemble method is also utilized for predicting emotional intensity and ambivalence. Akhtar et al. (2019) proposed a stacked ensemble method to solve not only the problem of classification, i.e., the prediction of emotion/sentiment, but also the prediction of intensity in emotion/sentiment otherwise known as the regression problem. There have also been significant developments in ambivalence handling, which enhances the capacity of algorithms to classify sentiments into positive versus negative, but also recognize four other classes: neutral, mix-positive, mixed, negative, and mixed-neutral (Wang et al., 2020). Wang et al. (2020) proposed an algorithm that includes sentiment scales that can be adjusted for fine-tuned emotion-sensing, and strength-level tuning parameters to consider sentiment intensity and handling ambivalence. For example, words such as ‘extremely’ or ‘super’ are considered the highest enhancer parameter, while ‘minor’ or ‘mini’ are considered as reducer parameters. In the field of reading emotions and sentiments from images or videos, researchers also collate different neural networks for classifying different aspects relevant to reading emotions such as objects, scenes, and facial expressions for higher accuracy (Barros et al., 2020; Do et al., 2020).

Interestingly, notable differences in the algorithmic design of affective computing among different branches can be spotted. Text-based computational processing of emotions, pioneered by the Asia-Pacific cluster with the Singapore-based *SenticNet* as the focal industrial player, increasingly features the ensemble method which takes advantage of both deep learning (i.e., a statistical approach that uses deep neural networks) and symbolic AI systems (i.e., knowledge bases that contain semantic networks using linguistic rules, hand-coded by humans) (Susanto et al., 2020; Cambria, 2016). In contrast, voice-based computational processing, pioneered by the European collaborative network with the Germany-based *audEERin*g as the focal company, is moving towards end-to-end machine learning with AI systems being designed to learn, extract, and even synthesize emotions by themselves with lesser degrees of human-annotated emotion labels (Schuller & Schuller, 2018; Schuller & Schuller, 2020). Similarly, vision- and biosignal-based affective computing also features weakly supervised machine learning, where the human role is limited to providing emotion labels for images or videos (Ivanova & Borzunov, 2020; Ngai et al., 2022; Vuong & Parry, 2021).

Differences aside, various branches and modalities of affective computing have been shown to collaborate and combine their efforts and techniques. For example, Poria et al. (2017) highlighted the movement of the field from unimodal processing of human affect to multimodal fusion in which computing techniques from texts, audio, and visual signals are combined to generate higher accuracy. Schuller and Schuller (2020) also emphasized the exciting trend of 'transfer learning,' in which deep neural networks trained in one modality (for example, to read emotions from images) are then applied to a different modality (for example, speech emotion recognition). Such unifying movement toward affective multimodal computing is epitomized by the Multimodal Sentiment Analysis Challenge—the MuSe challenge, co-organized by top academic and industrial institutions in the field such as Imperial College London (UK), Nanyang Technological University (Singapore), University of Augsburg (Germany), BMW Group, audEEring, SenticNet, etc. This AI challenge aims "to provide a common benchmark test set for multimodal information processing and to bring together the Affective Computing, Sentiment

Analysis, and Health Informatics communities, to compare the merits of multimodal fusion for a large amount of modalities under well-defined conditions" (Stappen et al., 2021).

1.4. Social problems with emotional AI

With on-going investment fueling these exciting development trends, it is not a surprise that emotional AI technologies are moving fast across national borders while being advertised as objective and value-free 'AI-driven solutions' for a multitude of problems including workers' productivity, stress management, optimization of interpersonal relationship, loneliness, prevention of mental health problems, etc.

However, in recent years, multiple branches of literature have emerged to critique the uncritical integration of AI as well as emotional AI in our daily lives. These branches of the literature come by with different names including data feminism (D'Ignazio & Klein, 2020), AI discontents (Hanemaayer, 2022), and algorithmic colonialism (Birhane, 2020). Crucially, works in these areas increasingly challenge the underlying neoliberal narratives that drive the development and infusion of AI technologies in our daily lives. In these critical reflections, many dark aspects of AI technologies including the lack of explicability, algorithmic bias, privacy concerns, etc. are brought to light.

Regarding emotional AI, societal concerns for this technology are five-fold, which are visualized in Figure 1.3. First, is the unethical or malicious misuse due to emotional AI's stealth data tracking. Second, are cultural tensions arising from these emotional AI technologies crossing national and cultural borders. Third, is the lack of industry standard to govern its uses. Fourth, existing ethical frameworks for emotional AI are often vague and inflexible. Last, but not least, comes the shaky science of the emotion-recognition industry.

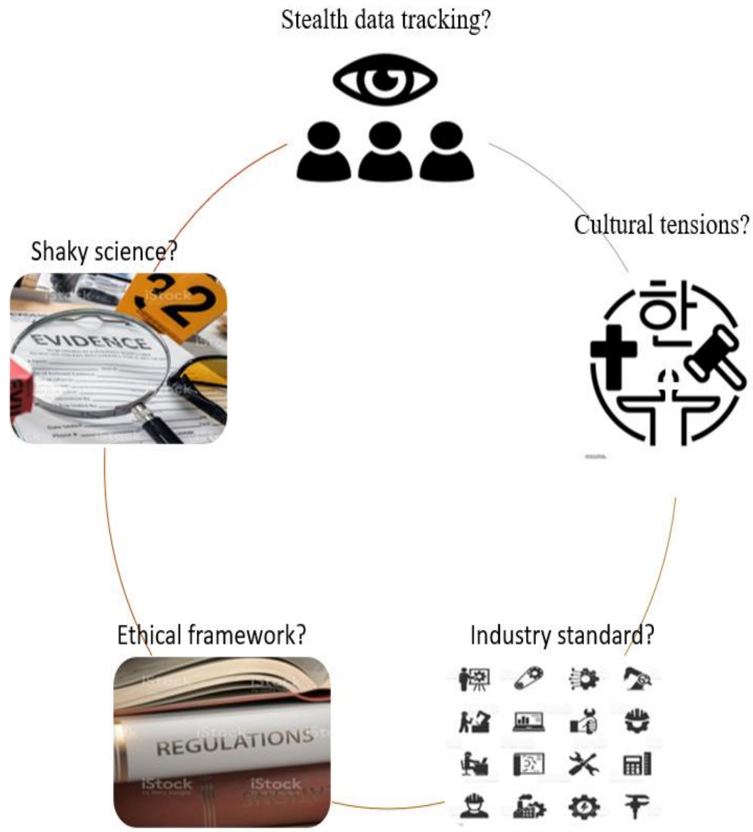


Figure 3.4. Five tensions impinging acceptance of emotional AI and non-conscious data harvesting.

First, affect tools are designed to harvest intimate data from an individual's subjective state without necessarily their awareness or permission. This creates multiple possibilities for its malicious or harmful misuse. For example, emotion-sensing devices in the workplace may lead to bias or discrimination against a worker for their lack of 'attitudinal conformity' (Mantello et al., 2021). In turn, affect tool may lead to emotional policing, creating a coercive pressure on the worker to always be happy, authentic, and positive. At the same time, diminishing their ability to backstage their feelings and in turn leading to anxiety, stress, and resentment. Similarly, affect tools in automobiles may lead to unfairly higher car or health insurance premiums. This is a concern that is examined deeply in Moore and Woodcock (2021)'s book titled "*Automated* *exploitation*." Concomitantly, in commercial settings, individuals may be exposed to empathic surveillance without their knowledge and depending on country, consent. For instance, AdMobilize markets AI software linked to public transit security cameras which monitor audience responses to interactive ads (AdMobilize, 2022). Besides analyzing gender, age, and dwell time, AdMobilize uses facial emotional analysis to detect micro-expressions of happiness, surprise, neutrality, and dissatisfaction. The intention is to track real-time ad performance and customer engagement. All captured analytics are then fed in real-time to a cloud-based dashboard that allows end-users to assess results, identify trends, and make optimize display content.

Second, similar to the hidden data gathering activities of many smart technologies, emotional AI will be far harder to collectively regulate as it is being developed as a proprietary layer in many products. A prime example is the automotive industry. Companies such as Ford, Porsche, Audi, Hyundai, Toyota, Honda, BMW, Volkswagen and Jaguar, in the name of safety and comfort enhancement, are developing in-cabin concierge systems that can track and respond to the emotional states of drivers (McStay & Urquhart, 2022). Yet, as researchers McStay and Urquhart (2022) observe, for the auto-industry, algorithmic secrecy is imperative for maintaining a competitive edge. This means that algorithmic transparency and collective standards for non-conscious biometric data collection will not occur for some time.

Third, although emotion-sensing technologies are predominantly designed in the West, they are being sold to a global marketplace. Problematically, although emotional AI is advertised as objective, value-free solutions that will bring many benefits for individuals and organizations such as increased well-being and productivity, it must be acknowledged that there are significant cross-cultural incongruences in interpreting the impacts of such smart technologies on core personal and traditional values (Ho et al., 2021; Mantello et al., 2021). As these devices cross international borders their algorithms are seldom tweaked for racial, cultural, ethnic, or gender differences (Buolamwini & Gebru, 2018a; Ho et al., 2021; Mitchell, 2019). A growing body of research shows that AI models that do not allow for difference or diversity can lead to unintentional bias or false positive identification, negatively impacting a target individual (N. T. Lee et al., 2019; Schelenz, 2022). This problem is further compounded by the lack of international consensus on the values and ethics that should be encoded into intelligent machines as well as cross-cultural incongruencies arising from a countries legal understanding of privacy (Mantello et al., 2021; Miyashita, 2021). For instance, while facial recognition and social credit systems are banned in many Western countries, China faces far less push-back because the notion of security is valued more than individual privacy (Roberts et al., 2021). Additionally, Chinese citizens are found to show greater trust in government-sponsored data collection than their Western counterparts (Aho & Duffield, 2020; Wang & Yu, 2015).

Although many countries across the globe have strengthened algorithmic regulations and data privacy laws, there are yet no uniform and consistent international agreements on artificial intelligence let alone AI systems that seek to harvest emotional data (ÓhÉigeartaigh et al., 2020). This is very concerning given the fact that AI technologies are being sold across national and cultural borders (Miyashita, 2021; Reddy et al., 2020). The absence of a global standard on data governance means a patchwork of legal and regulatory frameworks are being weighed against a range of neoliberal incentives and privacy considerations as well as elite stakeholder conversations which might elude rather than engage the public. For instance, although the United Nations Development Group published in 2017 a guidance for data privacy, ethics and protection in which it noted the need to engage stakeholders at all levels to ensure fundamental human rights are protected in the era of big data, the document remains a generic note without any means for enforcement or monitoring compliance in UN member states (United Nations Development Group, 2020). Similarly, the US Algorithmic Accountability Act, originally purposed in 2919, is still under review by Congress, with its original provisions watered down in order to placate the tech industry (Field, 2022).

Fourth, existing ethical frameworks for emotion-sensing technologies lack flexibility due to different businesses in various cultural settings having differing rationales or goals for adoption. For example, the Japanese voice analytics company, Empath, see the technology as a way for call-centers to optimize workplace productivity by providing supervisors with a panoptic window into the subjective state of each member of their customer service team. On the other hand, Moodbeam's emotion bracelet offer companies a neoliberal alternative to the far more management and expensive worker wellness programs. As the company suggests, a worker simply needs to wear the bracelet and it will automatically share data of his subjective state to both his managers and co-workers. This neoliberal approach to mindfulness is premised on the assumption that 'sharing is caring' (Mantello et al., 2021). Besides varying objectives to adoption comes the practical limitations of implementation and establishment of concrete metrics for measuring the technology's effectiveness. As the literature suggests, ensuring the efficacy of emotional AI technology requires having full-time staff skilled in data analytics and data management. However, many medium to large size companies employ automated managements systems without experts skilled in data analytics and data management experts (Bean, 2022).

Last, but not least, comes the shaky science of the emotion-recognition industry. A growing number of critics argue how can emotions be made computable when the science community cannot agree on exactly what emotions are, how they are formed or how they manifest themselves (Barrett, 2021; Birhane, 2020; Chen et al., 2018; Crawford, 2021a). Are emotions hard-wired into the psycho-physical makeup of an individual or per socially and culturally contingent? The science behind these affect-sensing algorithms has been heavily criticized for its lack of robust and solid evidence for the distinct biological blueprints of each human emotion. The fact is leading emotional AI companies are still relying on Paul Ekman (1999)'s now discredited theory of *eight basic universal emotions* (i.e., anger, fear, sadness, disgust, surprise, anticipation, trust, and joy) (Mitchell, 2019; Mohammad & Turney, 2013; Yue et al., 2019).

The pan-cultural, universal assumption of emotional expression is problematic because researchers such as decades of data from the science of emotion show the communication and inference of anger, fear, disgust, or any other Ekman's basic emotions have significant cultural and contextual variations (Barrett et al., 2019). Moreover, modes of emoting evolve since cultures are dynamic and unbounded, with the constant cultural transmission, learning, and unlearning (Boyd et al., 2011; Henrich, 2020; Vuong, Ho, et al., 2020; Q. H. Vuong, 2016). This truism about culture challenges the traditional and static ways of structuring emotion datasets into Ekman's eight basic types, the *valence* dimension (i.e., positive, neutral, negative sentiments), the *arousal* dimension (i.e., bored versus excited), favored by the tech companies (see McStay, 2018).

Computer science studies of algorithmic bias also converge on this concern about the accuracy of current emotional AI systems. For example, Rhue (2019) shows two facial recognition algorithms, *Microsoft AI* and *Face++*; both have a systematic bias about reading emotions such as anger and contempt, especially when interpreting emotions of different races. Timnit Gebru, the researcher who was fired from Google over a co-authored paper she wrote concerning the ethical risks of large language models, was also famously known for her discovery that that facial recognition systems are less accurate at identifying women of color (Buolamwini & Gebru, 2018b). Consequently, widespread use of emotional AI raises questions of fairness for marginalized groups such as women, ethnic minorities, people of colors, and people with disabilities, etc. (Barrett, 2021; Birhane, 2020; Chen et al., 2018; Crawford, 2021a). Pushing back against these arguments are affective computing engineers who insist emotions are in fact computable, that any limitations in diversity or cultural affordance will ultimately be solved by better algorithms (Barros et al., 2020; Schuller & Schuller, 2020).

As emotional AI emerges in cities, it will have profound impacts on the daily lives of citizens. By attempting to make internal emotional states visible it raises questions about data privacy in public spaces, empathic surveillance of everyday life and how governance mechanisms should best protect civic values and rights. Thus, the purpose of this study is to better understand what individuals think about a new era in human-machine relations, in which, intelligent machines *feel* but also *feed* off emotion as statistical fodder to reshape human behavior.

1.5. Research questions

Given the five-fold concerns for emotional AI uses in society, this thesis seeks to further explicate how emotional AI technologies are perceived by citizens. Two intuitions are presumably at play in determining the acceptance of smart affect-sensing technologies. The first intuition is, acceptance of a technology increases with its utilities: its ease of use, its help in managing stress, its help in improving security, its help with improving well-being, etc. The second is, acceptance of a new technology depends on a person's perception of its risks: it can be intrusive, it can diminish a sense of autonomy and freedom of expression, etc. These two intuitions are captured by two well-known frameworks of analysis: the Technological Acceptance Model (Davis, 1989), which postulates perceived utilities and perceived ease of use as two fundamental factors in technological acceptance; and the Moral Space of Hidalgo et al. (2021), in which how humans judge AI is a function of how the machine violates or conforms to the five moral dimensions:

Harm, Fairness, Loyalty, Authority, and Purity. This thesis will leverage these two frameworks and intuitions about our relationship with technology to offer a systematic and comprehensive analysis into the socio-demographic and behavioral determinants of attitude toward emotional AI application. Concretely, two research questions stand at the forefront of the project:

- How do sociodemographic factors (sex, income, educational qualifications, etc.) influence perceptions and acceptance of emotional AI applications?
- How do the concerns for fundamental values such as privacy, autonomy, safety, etc. of correlate with the attitude toward emotional AI applications?

Moreover, as the context where emotional AI is being deployed can change the user perception of AI systems, this thesis also seeks to understand the context sensitivity of attitude toward emotional AI. Thus, it will provide the answers to the following question:

- How do such correlations vary according to the domains of applications for example healthcare, education, security, politics, workplaces, etc.?

These questions are posed in the context of Japan. While studies on social perceptions of AI have been abundant in the West, studies on social aspects of AI in Japan are very few and far between. Yet, as introduced in section 1.1., Japan is home to many outstanding achievements in AI and robotics, such as the 2020 Guinness Record of the largest moving humanoid robot of 18-meter high Gundam, or the world's first companion robot in AIBO, launched in 1999. Given the multitude of social issues caused by a rapidly aging population, such as the lack of human resources, of caregivers, etc., the Japanese government views smart technologies, including emotional AI, as an integral, strategic solutions for said problems. Nevertheless, as much as Japan is known for being home to futuristic, advanced technological innovations, technological adoption in Japanese society is encumbered with many contradictions arising from the clash between values the technologies represent and prioritize (individualism, consumerism, flexibility, and productivity) and values of traditional Japanese culture (Bushido ethics of unconditional loyalty, daily devotion, respect of the hierarchy dominated by elderly men, and righteousness). Such contrast makes Japan a very interesting case for studying technological adoption behaviors.

This thesis, thus, provides a systematic and comprehensive investigation into the perception of emotional AI applications in various settings in Japan utilizing through conducting quantitative and quantitative analyses on original data sources: a national survey conducted in March 2022 (N = 2,000); a survey of Japanese clinic visitors in Beppu City, Japan (N= 245); a survey of foreign and Japanese students (N = 1,015); semi-structured interviews with stakeholders of emotion AI, e.g., vendors, union leaders, legal experts, working professionals in Japan (N = 31); group interviews via a citizen workshop setting (N = 24). Next, in Chapter 2, being informed by the above research questions, we will conduct a review of the current literature focusing on two key areas: 1) the social studies of emotional AI; 2) the studies of tech-acceptance and tech-adoption behaviors, concentrating on AI-driven technologies such as robotics, smartwatches, etc.

Chapter 2: Literature Review

Since this thesis focuses on the social and ethical implications of emotional AI technologies, in this chapter, we will review the current research landscape of studies on the perception and acceptance of emotional AI technologies. It must be stated studies that focus exclusively on emotional AI have been few and far between, as many problems with emotional AI are only first synthesized and introduced to the public in "*Emotional AI: The rise of empathic media*," a book published in 2018 by Andrew McStay, a UK-based scholar specialized in the ethics of digital technology. Since then, studies on this technology have gained momentum with works dealing with the social and ethical dimensions of emotional AI applications in various domains such as the workplace (Mantello et al., 2021), bioethics (Ghotbi & Ho, 2021; Ghotbi et al., 2022), cars (McStay & Urquhart, 2022), education (McStay, 2020a; McStay & Rosner, 2021), data governance (Ho et al., 2022; McStay, 2020b), security (Urquhart & Miranda, 2022; Urquhart, Miranda, et al., 2022b). Consequently, this chapter has three aims. First, it seeks to provide a review of the key findings in the latest works on the social and ethical dimensions of emotional AI technologies. Second, it seeks to explain and critically evaluate the prominent methodologies in studying the user perception of such technologies. Finally, it will provide a synthesis of key areas that are understudied in the literature.

2.1. Emotional AI: From sociological, ethical, and legal perspectives

Emotional AI is making its way into public and private uses. As demonstrated in Chapter 1, emotional AI products are growing in popularity and being used for many purposes. This section will focus on empirical findings on social and ethical issues related to the use of emotional AI in various commercial and public uses. Understanding the nuances in the acceptance and perception of this new technology is crucial for the future implementation of this technology.

2.1.1. Security

In the realm of security, existing research findings indicate that current emotional AI applications are questionable for their lack of transparency and suspect science (Wright, 2021), their inaccuracy, cost-effectiveness, and intrusiveness (Urquhart & Miranda, 2022). For example, in a sociological study of 26 frontline police officers' views of emotional AI applications for public security, Urquhart and Miranda (2022) find most of the UK officers interviewed expressed disbelief and skepticism regarding the use of emotional AI and facial recognition for public security. Investigating as to why the officers hold that view, the authors find there are four concerns: ineffectiveness, inaccuracy, distrust, and intrusiveness. The officers in the sample, speaking from direct experiences, state that the current level of accuracy of facial recognition technologies is "nowhere it needs to be" as the software easily makes wrong predictions when the quality of the footage is not high, as the software's accuracy is sensitive to even smallest natural actions such as squinting, smiling, of head-tilting occur. The sentiment of distrust is exceptionally high when it comes to decisions being made by the machines alone.

Moreover, the frontline officers also found the current technology too expensive for its level of accuracy and its limited relevance for most of the contexts in which the officers operate. In many cases, the officers tend to know the members of the public in areas where they are assigned, which brings into question the value of live facial recognition, and the fact that it records every interaction, and purports to know who everybody is at all times. The officers believe there must be transparency in the collection, storage, and analysis processes of such data for the public to deem the technology uses not intrusive.

Sharing similar concerns about the inaccuracy and intrusiveness of emotional AI technology, James Wright, the Alan Turing Institute, coins the term “suspect AI” to critique the lack of scientific legitimacy of the current generations of emotional AI applications. Applying the case study approach, Wright closely examined the algorithm behind Vibraimage and Mental Checker, two products by the Russian security company, ELSYS Corp, Wright (2021) argues such AI systems aim to algorithmically classify suspects/non-suspects, yet they themselves are suspect. Vibraimage and Mental Checker are algorithms that are trained to categorize mental states such as nervousness, happiness, arousal, etc. by reading the micromovements of the head.

Again, as explained in the first chapter, the scientific assumption of emotional AI technology is that emotion can be inferred from the machine reading of patterns in bio-signals such as facial expressions, gaits, tones, heart rates, skin conductance, etc. McStay (2018) refers to this view as a belief in ‘leaky emotions,’ where proponents of such a view think that an objective reading of emotions is possible given the fact that biological signals constitute the leak of our private emotions to the public.

Wright (2021), similar to earlier the argument put forth by McStay (2018), highlights the discrepancy between the priorities for simplistic and scalable machine learning models of industry of emotion recognition and the modern scientific understanding of the contingency and complexity of how emotion is constructed, expressed, and inferred. Citing the neuroscientific, anthropological, and psychological theories of emotions such as Barrett (2017)’s theory of constructed emotion, as well as closely examining the publications of ELSYS’s founder, Wright (2021) argues regarding the working of ELSYS’s products, “there is no coherent explanation of why certain intensities of head movements equate to a particular precise combination of emotions, behavior, intent or character.” Yet, the author is alarmed by the uncritical embrace of such technology, as they are now becoming a part of the surveillance assemblage in China, Japan, Russia, and South Korea, where they are increasingly incorporated in existing facial recognition systems used widely in convenient stores, ATMs, public events, etc.

Echoing similar worry, Urquhart and Miranda (2022) argue emotional AI will further entrench bio-deterministic framing of criminality, the automating suspicion based on the recognition of a person's identity, and his/her mental states poses a very high risk to the public. Thus, from a they urge cautions from public institutions, the use of less intrusive techniques such as fingerprint identification, and the adoption of participatory approaches in the deployment of new technologies. From a legal perspective, Urquhart and Miranda (2022) argue it is crucial for the UK law enforcement and legislation to take an proactive stance to create more precautionary and prescriptive guidelines for the use of facial recognition technologies rather than relying on existing innovative uses and resulting legal test cases. In other words, for high-risk technologies such as emotional AI security cameras, careful setting of standards before their deployment is a must to ensure the quality of law that can protect citizens from various moral harms.

2.1.2. Private spaces: Home and cars

Existing studies on emotional AI products in a private environment such as our homes and cars underline concerns about privacy loss, autonomy loss, and the lack of legal regulations for intelligent machines that constantly process our emotions.

Regarding the use of emotional AI in cars, McStay and Urquhart (2022) conduct 13 in-depth interviews with a range of experts on smart cities, the car industry, emotional AI, and data protection policy making. Analyzing the data with the sociological approach qualitative thematic coding, key themes about the concerns of the experts such as the questionable sensing-but-not-storing data claims of industry, manipulation of car users' behaviors (nudging vs. sludging distinction), etc. emerge out of the interview transcripts. The authors emphasize the challenges posed by the secretive and competitive nature of the car industry, which makes the consensus of a collective, publicly available industry standard for emotional AI in cars next to impossible. The authors also touch on multiple legal concerns that come with the use of emotional AI in cars. For example, given that emotional AI in cars will be part of the critical safety system, McStay and Urquhart (2022) argue such technologies should be deemed high-risk AI systems based on an analysis of the proposed European Commission AI Act.

Regarding the use of emotional AI in home setting, Urquhart, Miranda, et al. (2022b) combine the application of sociological approach and the review of recent legal cases involved smart homes to develop the concept of smart homes as 'invisible witnesses,' drawing from recent examples of data provided by Amazon Echo, Fitbit, Apple Health Data, etc. play a crucial role in solving crimes. Reflecting on the future of living with an 'invisible witness,' Urquhart et al. (2022) raise an important question about the loss of privacy implicated in the widespread adoption of IoT. Although users buy IoT products because they bring pleasure, safety, and many other utilities, in having a permanent witness, that constantly tracks and stores data on health, emotion, and daily activities, we risk losing a privacy space, quoting Goffman (1956), where we can enjoy *"backstage relaxation from playing social roles, having no fear of observation or judgment of others, and having utmost control over information flows." (p.634).* Moreover, besides individual concerns, there are also legal, forensic, and criminological concerns about how the technologies can be used and governed so as to protect the identity and agency of the users.

2.1.3. Education

Various aspects of the use of emotional AI in educational sectors have been explored in the works of McStay (2020a) on the rise of emotional AI in edtech (educational technology); Williamson

(2017a) and Williamson (2021) and Manolev et al. (2019) on ClassDojo, a popular app for students training; McStay and Rosner (2021) on children's toys. Similar to research studies in other areas, scholars in this area also emphasize the significant level of risk involved in the deployment of emotional AI in educational facilities and tools, despite the promises of advanced personalized learning, enhanced effectiveness in interventions when students are struggling with either with class materials or social, emotional challenges. McStay (2020a) identifies serious questions about the effectiveness, validity, and representativeness of training data in emotional AI edtech, the desirability of the chilling effects of the technologies (i.e., the feeling of inhibition and excessive self-consciousness) on students, the financial incentives of private companies might not align with the well-being of students, etc.

McStay (2020a) highlights the muddy legal and ethical concerns regarding deployment of emotional AI technologies in schools setting, as he shows it is clearly at odds with the UN's stance on the right of a child on a number of issues: the data minimization principle (i.e., collecting and processing data only toward necessary ends), the issues of maximizing the children flourishing, the child's right to liberty, to fullest development, to freedom of thought, etc. On first look, the deployment of emotional AI in education to purposefully modify students' behaviors via not only constant surveillance, but also reinforce 'desirable' behaviors and punish 'undesirable' behaviors, is at best questionable on these UN's moral considerations, at worst, outright dystopian. Such the modern phenomenon of modifying students' behaviors with technological tools is often legitimized on the basis of the belief that academic learning must be supplemented by social and emotional learning. As such, Williamson (2017) and others, who take on a more sociological approach, criticized that such operationalization of social and emotional learning through emotional AI epitomizes the problematic logic of technological solutism, i.e., the uncritically embrace of technologies to solve structural problems ranging from poorly funded schools to the unsound, outdated pedagogy.

Regarding emotional AI in children' toys, McStay and Rosner (2021), combining both qualitative interview data and quantitative survey data, highlight the unease around the issue of generational unfairness: Unlike adults, children have little control and ability to negotiate and challenge the uses of emotional AI technologies toward them. Consequently, the authors raise concerns about the manipulation of young children and their rights to have parts of their childhood forgotten. The presence of emotional AI toys in a home also raises concerns about parents' susceptibility since most parents lack the technical understanding regarding what data and how they are collected and processed in these toys.

2.1.4. Workplace

Scholarship that focuses on applications of emotional AI in the workplace emphasizes the neoliberal logic that drives the global workplace to adopt the emerging affect-sensing tools. On the one hand, intelligent emotion-sensing devices are supposed to help employees and companies manage stress, become more mindful and focused, and ultimately increase productivity and creativity at work. On the other, research studies on the effects of earlier versions of workplace surveillance technologies have highlighted serious drawbacks such as increased stress, and lower organizational commitment.

Even before the COVID-19 pandemic, self-tracking wearable devices have already become prevalent in the workplace. Early adopters of emotion-sensing wearables see this new technology as potential solutions for many problems in the modern workplace from interpersonal conflicts to the prevention of harassment. For example, Humanyze claims that its emotion analytics can promote harmonize interpersonal relationships at work via extracting contextual insights from the monitoring of internal communications, networks, relationships, location, and individual biometrics. HireVue provides AI products for recruitment purposes that can analyze personality and the likelihood of staying in the job via data from video interviews including spoken words, facial expressions, gestures, etc. In Japan, *Empath* in Japan markets voice-analytic software that allows call center managers to monitor employee moods in real-time, claiming it could reduce stress for the employees and turnover rate for the companies. Similarly, Behavox uses voice analytics of telephone conversations to track deviations from established patterns Emotional AI companies such as shouting at someone. With remote works becoming the norm in many companies due to COVID-19, employee-tracking technologies including emotion-tracking devices are on the rise (Crawford, 2021).

McStay (2018) raises three ethical concerns for the uses of emotional AI in the workplace. First, the default logic of emotional AI technologies privileges universal, basic emotions, thus when organizational decisions are made about employees on the basis of information provided by these technologies, there is a real risk of biases against minority groups such as migrant workers, women, LGBTQ, people of color, people with disabilities, etc. Second, there is a concern of coercion, as companies are in a stronger power position than most of their employees. And third, the logic of self-directed monitoring shifts the burden of solving social and emotional problems more toward the individuals rather than the structural and institutional setting, thus, it raises questions of fairness in organizational governance.

Added to these ethical concerns is the cultural incongruences when it comes to the interpretation of risks posed by the emerging technology. For example, Mantello et al. (2021), analyzing a multinational dataset of job seekers regarding their attitude toward various uses of emotional AI at work such as recruitment and monitoring, reveal that people with social privileges such as the male gender or a higher-income background tend to worry less about emotional AI at work. In addition, there is a higher acceptance among East Asian respondents of the technology compared to their Western counterparts, which might result from the differences in the collectivist and individualist cultural orientations. Such results, thus, invalidate the claim of universality of emotions made by existing emotional AI companies and raise an additional risk of inaccuracy and discrimination when these technologies are traveling across borders.

In a recent study, Urquhart, Laffer, et al. (2022) present scenarios of working with an emotional-AI-based supervisor in the workplace to 46 UK respondents from four distinct social groups, and find most respondents are ambivalent and negative about the prospect of having an AI that tracks their emotions at work. The participants concern the introduction of such technology entails a lack of trust in the employees. Moreover, these systems are seen as highly invasive, open to coded biases, and cannot consider variations among the individuals. The authors conclude the existing power asymmetries in the workplace necessitate cautions of the legislation, implementation, and development of affective computing for uses in the workplace. Unfortunately, as demonstrated by recent legal articles on existing laws that seek to regulate emotional AI, the existing laws are not fit for the job given the ‘black box’ nature of emotional AI algorithms and how emotional data are shared and stored by third parties (Bard, 2021; Bustamante et al., 2022). Bard (2021) and Melville et al. (2022) analyzing legal concerns that arise from the increasing use of emotional AI in the workplace context. Crucially, the authors highlight the shortcomings of current legal frameworks when it comes to concretely define the harms that would occur had emotional AI applications fulfill their purported claims of accurately reading emotions. For example, Bard (2021) asked if reading emotions by algorithms can be equated to reading thoughts and consequently, if it constitutes a violation of privacy. Problematically, the legal scholars point out the lack of legal protection in the US federal level against the use of emotional AI's generated information against individuals given that this type of information is currently possessing no concrete legal meaning. Bard (2021) points out that it is not clear emotional AI's generated information should be considered biometric data or should be treated as biological components such as blood or skin. Different ways of looking at emotional AI's generated information, i.e., the mental states, feelings, arousal, etc. can result in different legal statuses of such information, hence, influencing how it can be used in courts. These problems are indeed difficult and will require more careful thinking and legal analyses.

2.1.5. Emotional AI and the media

Mass media events have been found to shape public perception of AI. Neri and Cozman (2020) have shown that cautious opinions offered by technology pundits such as Stephen Hawking or Elon Musk could change AI risk evaluation by the public. Moreover, a study of media discussions on AI in the New York Times taken over 30 years showed a progressive increase in concern about the loss of control over AI, ethical concerns about the role of AI in society, and displacement of the human workforce (Fast & Horvitz, 2017). Using the NexisUni database, Ouchchy et al. (2020) found the tone of media coverage of AI's ethical issues was initially optimistic and enthusiastic in 2014 and became more critical and balanced until 2018, with the privacy issue being the most salient aspect of this debate.

Controversial political events such as the Cambridge Analytica case or the yet to be passed Algorithmic Accountability Act in the US can also shape the public discourse of the risk of AI misuse. The UK-based data broker company fell into disrupting when the public was made aware that its various parent companies, such as SCL Elections Ltd., had executed psychological operations (psy-ops), powered by harvesting massive social media data with algorithms to micro-target and allegedly change individual political beliefs and behaviors in more than 200 elections around the world, mostly in underdeveloped countries (Kaiser, 2019; Wylie, 2019). Bakir (2020) assessed the profiling offered by the company to the Leave.EU campaign in the 2016 Brexit Referendum and showed such a practice has both *deceptive* and *coercive* features.

Since then, in surveys around the world, where people are aware of digital micro-targeting practices, they have expressed a clear desire for action against technologies that exploit the emotionality of voters in political campaigns (Woolley & Howard, 2018). Yet, according to a YouGov survey in 2019, while 58% of the UK national sample were against tailoring political adverts, 31% of the UK sample were unaware of these problems (ORG, 2020). In response to the growing public concerns over the manipulativeness and intrusiveness of the AI-powered digital political and marketing campaign, politicians in advanced democracies have started to push for legislation that increases companies' transparency and accountability to build and deploy these AI systems (Badawy et al., 2018). Legislations such as the EU Digital Services Act, the Algorithmic Accountability Act, and the Filter Bubble Transparency Act in the US, the German *Medienstaatsvertrag* (State Media Treaty) have sparked heated public debates and received support from certain political factions and stakeholder groups (Rieder & Hofmann, 2020). However, it must be stated that crucial data are absent from public debates about AI governance other parts of the world such as the Global South and East Asia (Miyashita, 2021).

2.2. Determinants of user perception regarding emotional AI: Analysis of the empirical literature

Traditional theoretical models such as ‘Theory of Reasoned Action’ or ‘Social Cognitive model’ have provided a partial account for an individual’s reasoning process to explain the process of technological acceptance based on cost-and-benefit calculation, e.g., the perceived usefulness and perceived ease of use as the dominant predictors for tech-adoption behaviors (Okumus et al., 2018). However, these theories have struggled to account for cross-cultural differences in norms and values (Taherdoost, 2018). Davis's 'Technology Acceptance Model' (1989), the most cited model in this field, intentionally leaves out the cultural and subjective elements (e.g., norms, social roles, notion of self, and values) citing the difficulty to quantify as the reason for exclusion (Muk & Chung, 2015). Venkatesh and Davis (2000), however, expanded the original TAM model to include subjective norms. Yet even here, the authors' understanding of the term is based on whether most people who are close to a person think he or she should or should not adopt a technology (p.187). Such a narrow modulator for human behavior does not capture the complexity of cultural nuances in norms, social roles, the notion of self as well as personal values. For example, decades of psychological science research have shown people in collectivist cultures are more likely to conform to their group's expectations compared to individualist cultures (Henrich, 2020).

2.2.1 Cultural values as determinants of technological acceptance

An emerging body of literature has started to lend credence to the explanatory significance of cultural values in the behavioral mechanism of tech-adoption (Alina & Khalina, 2021; Dutot et al., 2019). Importantly, cultural values are found to underlie perceptions of risk and self-efficacy when dealing with new technologies such as carbon capture technologies (Hope & Jones, 2014), social media (Alsaleh et al., 2019), artificial intelligence (Vu & Lim, 2021), smartwatches (Dutot et al., 2019) or IoT based applications (Psychoula et al., 2018). Here, cultural values or cultural factors can be understood as broadly as "social learning, how we do and think about things, transmitted by non-genetic means" (Sapolsky, 2019, chapter 9). Past seminal studies that focus on cross-cultural differences tend to classify cultural characteristics into several dimensions: power distance, individualism/collectivism, masculinity/femininity and uncertainty avoidance (Henrich, 2020;

Hofstede, 1997). In the context of empirical studies of attitude toward smart technologies, cultural factors are often operationalized by comparing the differences between two populations, e.g., Japanese vs. American or Chinese vs. European, (Alina & Khalina, 2021; Alasaleh et al., 2019).

It is worth noting that the effects of cultural factors on attitude toward emotional AI and other smart technologies can vary case by case. For example, Psychoula et al. (2018) found for users of IoT-based applications, medical information is considered much less private among Asian respondents. Similarly, Mantello et al. (2021) found young East Asian, especially, the Japanese, respondents express much more trusting attitude than their Western counterparts when it comes to emotional AI applications in the workplace. The authors conjecture this is due to the dominant values in Confucianism, i.e., people in East Asian culture strongly value loyalty and respect to the workplace hierarchy, of which, AI applications can be thought of as an extension of the authority. Since the respondents are young job seekers, it is also possible that the embrace of AI applications in the workplace is a rection against the traditional hierarchies dominated by elderly men. Or Hope and Jones (2014)'s comparative study on the Muslim, Christian, and secular population demonstrated that perception of carbon capture technologies are guided by specific religious teachings such as stewardship and harmony values. Several American national surveys found, compared to highly religious people, non-religious and less religious people (measured by the number of times they attend religious services, for example (Brewer et al., 2020) held a more favorable view of AI (Northeastern University & Gallup, 2018; West, 2018).

It must also be said that there is a sharp difference between the public imagination of AI technology between the West and the East. Historically, in the West, public perception of AI has been greatly influenced by fictional representation in popular novels, film, and television, many of which represent AI as an innately dystopian form of synthetic intelligence such as *Colossus: The*

Forbin Project (1970), *Terminator* (1984), *I, Robot* (2004), *and Ex Machina* (2014). On the other hand, Asian people tend to associate AI in a more favorable capacity since they grow up with beloved manga/animation characters such as *Mighty Atom (Astro Boy)* and *Doraemon* (Robertson, 2017).

This section has highlighted the multitude of cultural factors that contribute to the shaping of public perception of AI technologies. Studies in this area have revealed socio-cultural backgrounds, fictional representation in popular culture, public positions of the experts in the field, political events and scandals, perception of economic and political rivals are also contributor of public perception toward AI.

2.2.2. Behavioral factors

One consistent finding in the literature is that people have little concern over job loss due to AI (Brougham & Haar, 2017; Pinto dos Santos et al., 2019). For example, a recent survey of 487 pathologists indicated that nearly 75% of the participants displayed excitement and interest in the prospect of AI integration in their work (Sarwar et al., 2019). Alternatively, there is also evidence that suggests greater anxiety related to the rise of AI applications in the workplace. Brougham and Haar (2017) found in a New Zealand study that the greater an employee's awareness of these technologies, the lower their organizational commitment and career satisfaction. These findings are concurrent with previous studies that have examined the relationship between biometric surveillance and employee trust in the workplace (Rosenblatt, 2018; Marciano, 2019; Mateescu and Nguyen, 2019; Manohka, 2020).

Of the few studies that looked at varying student attitudes toward AI from different university majors, mixed results were found. In terms of future sustainability, Gherheș and Obrad (2018) found Romanian students at technical universities held more positive views of AI than their humanities counterparts. Likewise, Chen & Lee (2019) found that Taiwanese students majoring in science and engineering are more positive about AI's social impacts than those in humanities, social science, management, education and arts. Noteworthily, it appears that curriculums of business schools with the Association to Advance Collegiate Schools of Business (AASCB) accreditation emphasize the importance and advantages of acquiring data analytics skills but little on data ethics to enter the increasingly AI-enabled business world (Clayton & Clopton, 2019). It is also common for business and marketing academic journals to emphasize the positive rather than negative aspects of AI in optimizing various operations and processes (Prentice et al., 2020). Consequently, one would expect business students to be more familiar with AI and have more positive attitude for AI in the workplace.

2.2.3. Socio-demographic factors

Besides cultural and behavioral factors, some socio-demographic factors are found to be important predictors for tech-acceptance. For example, male gender, higher income, and higher educational level are consistently found to predict a more tolerant attitude toward new technologies that run in the background, performing automatic data gathering and analysis (Ali, 2012; Hidalgo et al., 2021; McClure, 2017; Muriithi et al., 2016). Specifically, men are found to be more accepting of self-tracking apps (Okumus et al., 2018; Urueña et al., 2018) or AI/Robots (McClure, 2017) or drones (Aydin, 2019). Meanwhile, women are shown to exhibit more concern with third-party data usage beyond the original purpose and behavioral advertising techniques (Hoy & Milne, 2010), and were more likely to engage in privacy protection behaviors (Cecere et al., 2015) since they generally perceived more risks in the new technologies (Kasilingam, 2020).

Higher income is also a reliable predictor of willingness to adopt new technologies (Ali, 2012; McClure, 2017; Urueña et al., 2018). Higher level of education has also been shown to positively correlate with attitude toward automated decision-making and news recommendations by AI (Araujo et al., 2020; Thurman et al., 2019). McClure (2017)'s study of AI technophobia among the US population reveals people from non-dominant social classes such as lower income or non-white groups or female are far more likely to be threatened by new technologies. Batte and Arnholt (2003) argued people from dominant social classes tend to be early adopters of technology as they could afford the risks as well as they are often viewed as local opinion leaders.

Importantly, context also matters when it comes to the use of smart technologies as divergent attitudes occur when data-harvesting smart technologies are used by the government stakeholders rather than their private/commercial counterparts. For example, Wang and Yu (2015)'s study found female Chinese respondents were more likely to doubt the private sector's data practices, yet gender difference is negligible when it came to data governance by the public sector institutions (Wang & Yu, 2015). Damerji and Salimi (2021) found third and fourth-year students in university have higher perceived ease of use, perceived utility, and acceptance towards AI.

Although these socio-demographic factors are indeed useful in predicting AI perception, it is important to keep in mind that most of these studies are conducted from a single-country perspective (Ali, 2012; McClure, 2017; Batte & Arnholt, 2003; Damerji and Salimi, 2021, Araujo et al., 2020). And unfortunately, emerging scholarship of emotional AI has yet to produce conclusive evidence on technological acceptance of emotional AI. For example, McStay (2020b) found gender, social class, and region were not correlated with any clear differences in the UK's public attitude regarding emotional AI's data practices. In contrast, Mantello et al., (2021)'s seminal emotional AI in the workplace study showed that being male, higher income, and non-religious predicted a more trusting attitude toward the new technologies.

2.3. A synthesis of understudied areas

A review of the literature on emotional AI and determinants of tech-acceptance reveals an uneven and complex picture of human-machine relationship in the age of 'empathic media,' i.e., media technologies that can sense, read, and respond to human emotions (McStay, 2018). There are three key observations.

First, most of the current studies focus exclusively on emotional AI are based on qualitative thematic coding of data collected from the interviews and focus-group sessions (McStay, 2020a; Urquhart, Laffer, et al., 2022; Urquhart, Miranda, et al., 2022a). Consequently, it is necessary to expand the methodological toolkit via statistical analysis to more comprehensively study emotional AI's impacts in society. And this thesis seeks to contribute to this understudied area by drawing from the quantitative literature on technological acceptance behaviors (Ho et al., 2022; Kamal et al., 2020; Venkatesh & Davis, 2000) as well as quantitative moral psychology (Haidt, 2007; Hidalgo et al., 2021).

Second, the current literature shows there is a multitude of factors that underlies how an individual perceives a new technology. There are socio-demographic factors such as age, educational qualifications, gender, income level, etc. There are also key behavioral and perceptual variables such as perceived ease of use, perceived usefulness, privacy concerns, etc. How should these variables and their relations be mapped on to a theoretical framework? What are the strengths and weaknesses of current theoretical frameworks to study tech-acceptance behaviors given the ability of emotional AI technology to shape and nudge our behaviors and beliefs? These are the theoretical concerns that this thesis will seek to address.

Third, the current literature highlights the importance of contexts in studying the use of emotional AI. Clearly, different users of the technology (e.g., individuals or companies) and different contexts of using (e.g., whether the technology is used in the public or in the private) can vastly change the attitude towards emotional AI technology. Thus, this thesis aims to further explicate the context sensitivity of emotional AI uses via analyzing how correlations of several behavioral, socio-demographic variables with the overall attitude toward emotional AI change according to the context (i.e., policing use, workplace, media, etc.).

In the next chapter, we will further explore the theoretical frameworks and methodologies related to the social studies of emotional AI.

Chapter 3: Materials and Methods

This thesis draws from multiple sources of data and methodologies. Regarding the data sources, first, to understand determinants of attitude towards emotional AI in various sectors, three surveys have been conducted. The first survey collects viewpoints and demographic data from 1,015 international and domestic students in multicultural campus of a Japanese university. The second survey is conducted on Japanese residents living in Beppu City, Japan, which resulted in 245 responses. The third survey is national and representative, with 2000 responses from Japanese citizens aged 20 to 69. Second, for thematic qualitative analyses, more than 30 in-depth interviews with various stakeholders of emotional AI technologies (company representatives, legal experts, union leaders, working professionals) and four citizen workshops with more than 24 subjects are conducted. Finally, to understand the current research landscape, a bibliometric analysis of more than 3,300 articles indexed in the Web-of-Science database since 1995 is conducted, of which some insights from a preliminary analysis have been presented in the Introduction and Literature Review chapters.

Although the data come from various sources and with different collecting methods, the analysis and interpretation of research results are drawn from two theoretical frameworks: The Technological Acceptance Model of Fred Davis (1989), the Moral Space, a mathematical construct to quantify moral perception of machine actions (Hidalgo et al., 2021). Although the above theoretical frameworks have been well-tested and cited in the literature, each of them carries significant weaknesses when it comes to dealing with such a new technology as emotional AI. Yet there are elements from these models that can supplement each other and by putting them together in a sufficiently logical way, the new model will provide a better way of understanding user perception of emotional AI. So, this thesis argues.

3.1. Theoretical framework

3.1.1. Critiques of the Technological Acceptance Model (TAM)

The TAM is first proposed by Fred Davis (1989) in the prestigious journal Management Information System Quarterly, and it is one of the most well-cited models in the study of tech-adoption behaviors. In the original paper, In the original TAM model (1989), Davis hypothesized that the level of acceptance of a new technology is determined by two factors: perceived ease of use and perceived utility (See Figure 3.1). In the paper published in 2000, the model was extended to include subjective norms, but this factor is narrowly defined as whether most people who are close to (or familiar with) a person think he or she should or should not adopt a technology (p.187). In other words, it presumes a measurement of conformity due to social influence. Both the original (1989) and the extended TAM (2000) have enjoyed a high level of citation and empirical support. For example, a study found the extended model accounted for 61% of the variance in the behavioral intention (BI) to adopt mobile wallet technology (Lew et al., 2020). Another meta-analyses of digital technology adoption in education show the TAM models can account for up to 44% of variance in the BI (Scherer et al., 2019).

However, the rise of emotional AI and other forms of smart technologies such as ubiquitous computing or IoT embedded in our physical and digital infrastructures have exposed several limitations of the TAM: its linear, static assumptions about human-machine relationship as well as its lack of cultural sensitivity.

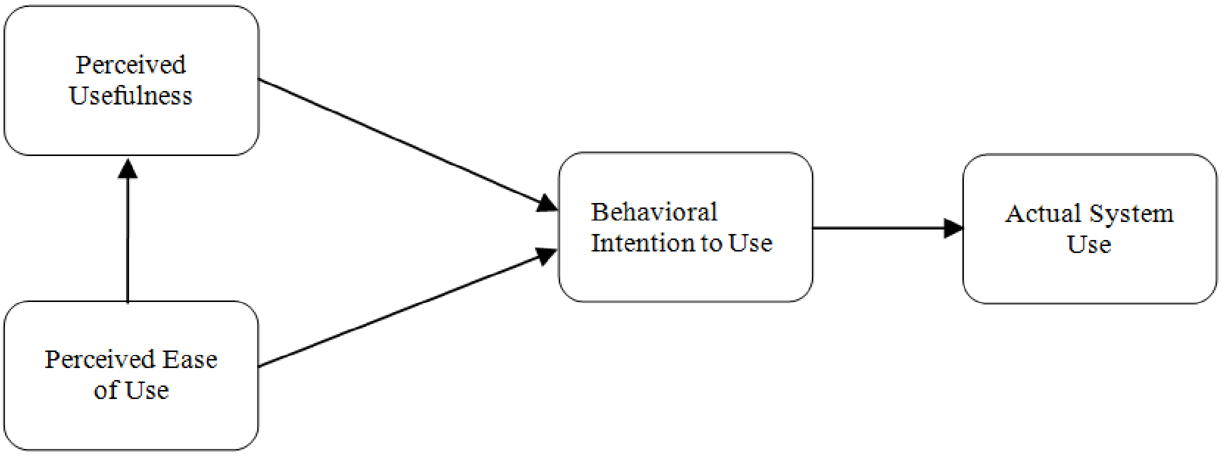


Figure 5.1. Technological Acceptance Model as originally conceived by Davis (1989).

First, TAM implies a linear subject-object relationship between a user and a technology. Indeed, there is an act of the user physically adopting the technology. Such a linear and tactile relationship is no longer a prerequisite with emerging smart technologies, which often runs in a ubiquitous, ambient fashion in the background of personal devices or in public spaces.

Second, affect-sensing algorithms are not static. Rather, besides being able to read and track our emotions, they can also respond to our various emotional states, and as demonstrated in the literature, they are often designed toward various ulterior aims such as mental health surveillance (Conway & O'Connor, 2016; Gruebner et al., 2016), or maximizing certain desired behaviors (Williamson, 2021; Zuboff, 2019), including purchases and engagement with contents and products in online platforms, or paying attention in classrooms, or managing stress in customer calls, etc. As the AI-powered physical and digital platforms constantly direct and nudge our behaviors and attentions, this new human-machine relationship dictates novel ways of conceptualizing models of technological acceptance.

Third and finally, there is also a crucial issue of the TAM not being culturally sensitive. A common critique that has been leveled at TAM is its lack of accounts for the cross-cultural variance (differences in core values, mindsets, etc.) in the way people form acceptance perceptions such as ease of use, of utility, of social influences) (Taherdoost, 2018). Indeed, there is an emerging literature that tests the TAM and TAM2 model in various countries and shows cultural values do indeed influence how people form the perceptions that are pertinent to the TAM (Dutot et al., 2019; Matson et al., 2012; Muk & Chung, 2015). Hence, to fully understand perception of emotional AI technologies, the TAM model must be supplemented by other theoretical frameworks.

3.1.2. The Moral Space: Applying the Moral Foundation Theory to study human-machine relationship

The Moral Space is a mathematical construct that Hidalgo et al. (2021) and colleagues use to unpack the perceived morality of a machine action quantitatively. The Moral Space originated from Jonathan Haidt's Moral Foundation Theory (2007), which posits five fundamental moral dimensions including fairness, loyalty, harm, purity, and authority (Figure 3.2). In the 2021 book, *How humans judge machines* published by the MIT press, Hidalgo et al. (2021) propose the morality of a machine's action can be captured by a function of how it has violated or validated the five moral norms in Haidt's Moral Foundation Theory.

The Moral Foundation Theory proposes there are five moral foundations. The first is the dimension of Harm/Care, which is the concern about and dislike for the suffering of others. The second is the dimension of Fairness, i.e., the concern about proportional versus egalitarian fairness (Fairness). The third is Loyalty, i.e., the concern for ingroup loyalty (Loyalty). The fourth is Authority, i.e., the concern about preserving the social structures, authority, and tradition. Finally, the fifth is Purity, the concern prompted by the feeling of disgust about physical or mental/spiritual contamination. The first two foundations about Harm/Care and Fairness are often regarded as the individualizing foundations, since they entail the concern about well-being of a person. Meanwhile, the later three foundations are considered the ‘binding foundation’, since they are about maintaining cohesion and order of the collective (Atari et al., 2020).

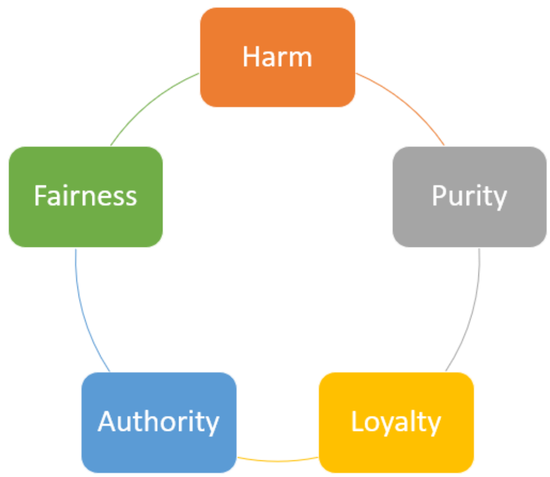


Figure 5.2. Five moral dimensions in the Moral Space (Hidalgo et al., 2021) and Haidt’s Moral Foundation Theory (2007).

In *How humans judge machines*, Hidalgo and colleagues presented to nearly 6,000 subjects hypothetical, but not far-fetched scenarios involving machines or humans making consequential decisions in different contexts, for example, a machine doing job screening vs a human doing job screening, a machine vs human security guard determining the legal status of immigrants in an airport, etc. The authors, using a 7-point Likert scale, ask the respondents to rate the moral wrongness of such situations first, then ask the respondents to rate how much a given action of a machine and a human has violated a moral norm in the Moral Foundation Theory.

This approach is indeed suitable for emotional AI social scientific research because it enables researchers to unpack various ethical dimensions and concerns related to the technology. For example, applying the Moral Foundation Theory, we can survey various moral perceptions of the technology: emotional AI as a threat to privacy, emotional AI as a threat to autonomy, or its utility such as increased safety or intimacy. It also highlights the contingency of human evaluation toward machines, as Hidalgo et al. (2021) found we judge machines more harshly based on the outcomes rather than intention, and uses of AI by the government are judged differently than uses of AI by private sectors.

3.1.3. Combining the TAM and the Moral Space: A three-pronged approach

Clearly, understanding which factors determine the attitude toward a new technology such as emotional AI is a nuanced act. Drawing insights from the TAM (Davis, 1989) and the Moral Space of how humans judge machines (Hidalgo et al., 2021), as well as various qualitative studies that have been reviewed in previous chapters, we can begin to synthesize and advance our understanding of determinants of emotional AI user perception.

As such, I propose a three-layered approach toward the unpacking of ethical and social dimensions in our attitude toward emotional AI applications: *Contexts, Variables, and Statistical Models* (Figure 3.3).

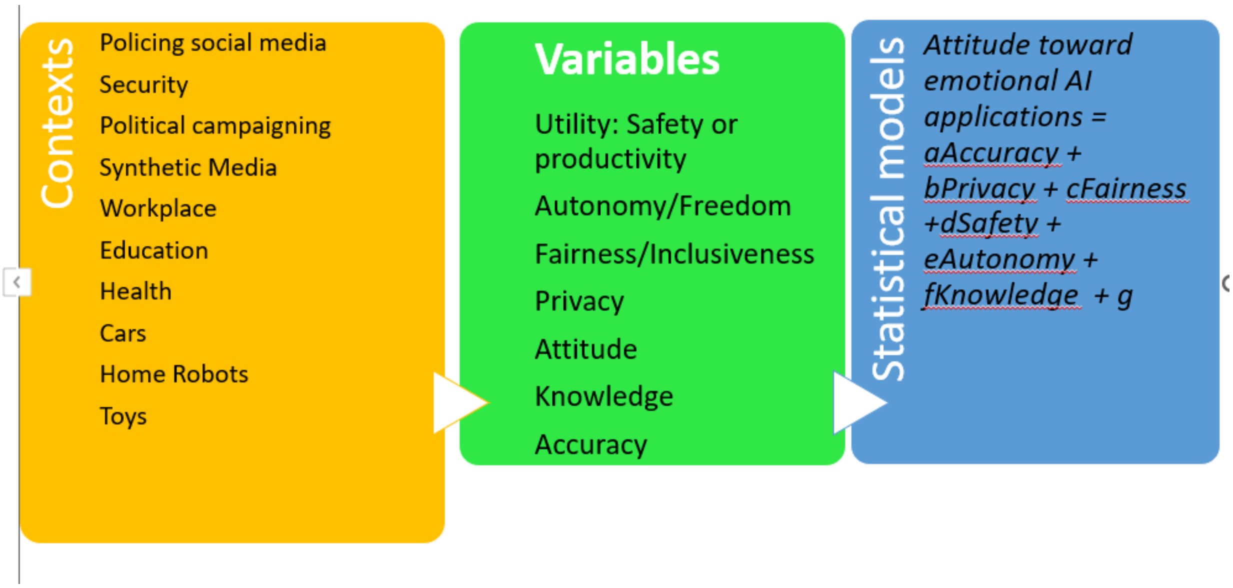


Figure 5.3. A three-pronged approach toward synthesizing our understanding of emotional AI user perception.

Regarding the contexts, as pointed out in the literature review, each different use cases will bring about a different set of concerns or making certain concerns more pronounced than in other cases. Thus, the context sensitivity of emotional AI user perception must be studied.

Regarding the variables, first, based on the TAM, utilities such as how easy it is to use/understand a new technology or perceived increased safety or productivity, etc. will play a role in our attitude towards the technology. Thus, we will include these variables in our survey design. Second, based on the Moral Foundation Theory, harms conceived as a violation of moral values such as privacy, fairness/inclusiveness, and autonomy are hypothesized to influence our acceptance of the technology. Finally, as shown previous studies on emotional AI in cars, toys, and security (McStay & Rosner, 2021; McStay & Urquhart, 2022; Urquhart & Miranda, 2022), background factors such as the level of trust in government regulation of new technologies and the level of trust in the private sector's self-regulation of new technologies are also factors worth examining.

These key considerations will help us explore the terrain of ethical concerns for emotional AI applications and its context sensitivity. Table 5.1 summarizes the hypotheses which will be tested in this study.

Table 5.1. A three-pronged approach toward synthesizing our understanding of emotional AI user perception.

No.	Hypotheses	Literature/ Theories	Research questions
1	H1: **Being male** is positively correlated with attitude toward emotional AI. While the opposite is true for **female**.	Empirical findings on attitude toward AI applications (Ali, 2012; McClure, 2017; Urueña et al., 2018)/ Sex differences in Moral Foundation Theory (Atari et al., 2020; Graham et al., 2011; Hidalgo et al., 2021)	RQ1
2	H2: **Female** express more concerns about emotional AI's implications for moral harms such as privacy violation, autonomy loss, biased algorithms.	Sex differences in Moral Foundation Theory (Atari et al., 2020; Graham et al., 2011)	RQ1/ RQ2
3	H3**: Income** is positively correlated with the attitude toward emotional AI.	(Brewer et al., 2020, McClure, 2017; Cai et al., 2017; Huffman et al., 2013)	RQ1
4	H4: **Higher educational qualification** positively correlated with attitude toward emotional AI.	Empirical findings on attitude toward AI applications (Ali, 2012; Chen & Lee, 2019; McClure, 2017; Urueña et al., 2018)	RQ1
5	H5: **Perceived utilities** of emotional AI technologies	Predictions from Technological Acceptance Model (Alina & Khalina, 2021; Davis, 1989; Kamal et al., 2020; Taherdoost, 2018)	RQ2

	positively correlate with attitude toward them		
6	H6: **Self-rated knowledge** with emotional AI technologies is positively correlated with attitude toward the emerging technologies.	Predictions from Technological Acceptance Model's (Alina & Khalina, 2021; Davis, 1989; Kamal et al., 2020; Taherdoost, 2018)	RQ2
7	H7: Concern about emotional AI's **negative impacts on moral values** (privacy, autonomy, biases, respect for authority/tradition, etc.) is negatively correlated with attitude toward emotional AI technologies.	Predictions from Moral Foundation Theory as adapted in the book How humans judge machines (Atari et al., 2020; Graham et al., 2011; Hidalgo et al., 2021)	RQ2
8	H8: Concern about **accuracy of the technology** is negatively correlated with attitude toward emotional AI technologies.	Predictions from Moral Foundation Theory as adapted in the book How humans judge machines (Atari et al., 2020; Graham et al., 2011; Hidalgo et al., 2021)	RQ2
9	H9: Transparency on how emotional data is managed, stored, processed positively correlated with attitude toward emotional AI. The opposite is true when no transparency is provided.	Qualitative research results from various use cases including cars (McStay & Urquhart, 2022), toys (McReynolds et al., 2017; McStay & Rosner, 2021), data management (McStay, 2020b), education (McStay, 2020a), smart homes , security (Urquhart & Miranda, 2022); workplace (Mantello et al., 2021; Urquhart, Laffer, et al., 2022), etc.	RQ2
10	H9: **Trust toward the government's ability to regulate** the technology is positively correlated with attitude toward emotional AI technologies.	Empirical findings from attitude toward AI/Robots and government effectiveness index (Vu & Lim, 2021)	RQ2
11	H10: **Trust toward the private sector's ability to regulate** the technology is positively correlated with attitude toward emotional AI technologies.	Empirical findings from attitude toward AI/Robots and techno-social environment (Vu & Lim, 2021)	RQ2

12	H11 (**The context sensitivity hypothesis**): Determinants of attitude toward emotional AI varied in according to different contexts.	Qualitative research results from various use cases including cars (McStay & Urquhart, 2022), toys (McReynolds et al., 2017; McStay & Rosner, 2021), data management (McStay, 2020b), education (McStay, 2020a), smart homes , security (Urquhart & Miranda, 2022); workplace (Mantello et al., 2021; Urquhart, Laffer, et al., 2022), etc.	RQ3

3.2. Materials

The empirical results are based on three main sources. The first is statistical analysis results of a cross-sectional data in two national, representative surveys conducted in March 2022, a municipal survey conducted in August to October 2021, and a convenient survey sample of international students in a Japanese campus conducted between June 2020 to April 2021.

The second is 31 in-depth interviews conducted with various stakeholders of the technologies, including producers and sellers of the technologies, legal experts, and companies' employees from various sectors, in the period 2020-2021.

The third is qualitative data from four citizen workshops, each lasted three (3) hours, organized in March and April 2022 to collect the viewpoints from four distinct social groups: Japanese nationals of the two age groups: 18-35 years old (n=6) and over 65 years old (n=4), foreigners working and studying in Japan (n=10), Japanese people who have some forms of mental and physical handicaps (n=4).

All the data analyses, both quantitative and qualitative, will be driven by the theoretical framework presented in the previous section. In other words, insights from the TAM and the Moral

Space model of how humans judge machines will be used to construct statistical models as well as generating themes for the qualitative coding of interviews and focus-group discussions data.

Below are the full descriptions of the data collection method.

3.2.1. Surveys

Table 5.2 and Table 5.3 present the socio-demographic breakdown of the two surveys collected in this study. The third survey collected from 245 clinic visitors in Beppu City, Oita Prefecture, in the South of Japan will be presented in chapter 5, which explores the perception of Japanese people regarding emotional AI in healthcare setting.

Table 5.2. Socio-demographic breakdown of the first survey from international and domestic APU students. The titled of the survey is: "Students' perception of emotional AI in smart cities."

Variables	**Category**	**Male (N = 437)**		**Female (N = 578)**	
		Frequency	**Percentage**	**Frequency**	**Percentage**
Region	Africa	5	1.14%	6	1.04%
	Central Asia	11	2.52%	5	0.87%
	Eastern Asia	224	51.26%	262	45.33%
	Europe	9	2.06%	11	1.90%
	Northern America	7	1.60%	10	1.73%
	South-Eastern Asia	137	31.35%	226	39.10%
	Southern Asia	41	9.38%	48	8.30%
	Oceania	2	0.46%	8	1.38%
Income	Low	39	8.92%	43	7.44%
	Medium	327	74.83%	483	83.56%
	High	71	16.25%	52	9.00%
School year	First year	63	14.42%	66	11.42%
	Second year	118	27.00%	198	34.26%
	Third year	128	29.29%	186	32.18%
	Fourth year	111	25.40%	109	18.86%
	Fifth year or more	11	2.52%	9	1.56%
Major	Business management/Economics	233	53.32%	185	32.01%
	Social science/Humanities	204	46.68%	392	67.82%

Religions	Atheism	132	30.21%	157	27.16%
	Buddhism	64	14.65%	129	22.32%
	Christianity	59	13.50%	66	11.42%
	Islam	52	11.90%	58	10.03%
	Others or Unidentified	130	29.75%	168	29.07%
Religiosity	Not/little religious	372	85.13%	494	85.47%
	Very religious	36	8.24%	45	7.79%

Table 5.3. Descriptive statistics from a national, representative survey on the Japanese population. The title of the survey is: "General Japanese citizens' perception of emotional AI technologies."

		Frequency	**%**
AGE AND GENDER		2000	100.0
1	Male／20s	200	10.0
2	Male／30s	200	10.0
3	Male／40s	200	10.0
4	Male／50s	200	10.0
5	Male／60s	200	10.0
6	Female／20s	200	10.0
7	Female／30s	200	10.0
8	Female／40s	200	10.0
9	Female／50s	200	10.0
10	Female／60s	200	10.0
	EDUCATIONAL LEVEL	N=2000	100%
1	Middle School	42	2.1
2	High School	530	26.5
3	Colleges of technology (高等専門学校)	37	1.9
4	Vocational School (専門学校・専修学校)	235	11.8
5	Junior college (短期大学)	186	9.3
6	Bachelor's	852	42.6
7	Master's	90	4.5
8	PhD	19	1.0

9	Others	9	0.5
	INCOME LEVEL	N=2000	100%
1	Under 3,300,000 JPY	450	**22.5**
2	Between 3,300,000 – 9,000,000 JPY	865	43.3
3	Between 9,000,000 – 18,000,000 JPY	228	11.4
4	Over 18,000,000 JPY	43	2.2
5	Do want to answer	414	20.7

3.2.2. Qualitative thematic coding of the interviews and citizen workshops

Qualitative thematic coding was deployed for the encoding of the transcripts of the interviews and the workshops. Thematic analysis is defined as a systematic method of "identifying, analyzing, and reporting patterns (themes) within data" (Braun & Clarke, 2006). In a methodological review, Castleberry and Nolen (2018) recommend five steps to make thematic analysis thorough and systematic, and reduce subjectivity: compiling, disassembling, reassembling, interpreting, and concluding. Similar processes for themes generating, coding, and organizing for qualitative research data were also recommended by Creswell (1994) and Clarke and Hoggett (2019).

We first transcribed all the interviews and citizen workshops (compiling). Then for disassembling, the coding strategy was based on the initial structure of the interviews and the focus group discussions: (1) General explanations of technologies and their applications; (2) Data management and algorithmic transparency; (3) Rules and trusts; (4) Minimizing harms and misuse.

During the reassembling process, based on insights from the TAM and the Moral Foundation Theory presented above, we further generated themes including:

1. guarding against biases
2. algorithmic opaqueness
3. protecting personal data

4. cultural conflicts
5. legal obstacles
6. misuse concerns
7. mitigating harms

After all transcripts of the interviews and citizen workshops were codified accordingly, we looked across all interviews to identify commonalities. In the next two sections, details on the demographics of the interviewees and the citizen workshops are presented.

3.2.3. Interviews

The first set of the interviews were conducted between January 15th to August 2nd, 2021, via an online conference platform, each lasted up to one hour. Here, interview subjects are representatives of global market players in AI business solutions based in Japan such as IBM Global Business Services, EMPATH, ELSYS, Preferred Networks, and HireVue. We also interviewed a leading data privacy expert and two labor union leaders involved in a series of ongoing lawsuits against Amazon Japan. Due to the COVID-19 pandemic, we have sent our interview requests to these stakeholders via emails, only nearly one-third of the contacted stakeholders reply and agreed to have an interview online. Verbal consent was obtained for all the interviews to be recorded and transcribed. Most interviewees were happy to be identified by name and their organizations, only two interviewees did not give consent to be mentioned by their names or their companies' names. Relevant information on the interview subjects is given in Table 5.4.

Table 5.4. Characteristics of the interviewed companies (organizations)'s representatives.

	Company name	**Interview subject's position**	**Company introduction**

Business/Industries			
1.	TalentA	SVP, Chief Financial Officer	The exclusive distributor of HireVue's AI-driven Video Interviewing and pre-hiring system.
2.	IBM Global Business Services	Talent and Engagement Associate Partner	Marketing algorithms powered by IBM Watson to optimize the flow of human resource management.
3.	Preferred Networks	PFN Fellow	Using cutting-edge deep learning technologies to create real-time sensing of the physical world, making devices intelligent and computable.
4.	WACUL	CEO	Using machine learning to optimize websites based on analytics of consumer behaviors.
5.	Empath	Co-founder and CSO	Empath uses emotional AI to detect 4 emotions which are joy, calm, anger and sorrow, besides an energy point. Their products have been used by more than 500 customers over 40 countries.
6.	Talented People Laboratory Inc.	CEO	Providing consulting services to companies and job seekers about important trends in human resource management in Japan.
7.	Business Research Lab	CEO	Providing consulting services to companies and job seekers about important trends in human resource management in Japan.
8.	ELSYS JAPAN	CEO	A partner of ELSYS Corp, the Russian company that sell technologies that detect emotions from head movement and facial expressions.
9	Anonymous	Senior engineer in emotion analytics team	A giant security company in Japan with advanced biometric identification as well as video analytics technologies.
10	Anonymous	Representative of the Japanese branch	A global leader in emotion analysis technology for ad testing.
Union members			

11	Amazon Japan Union	Mr. Masafumi Ito; Union leader	Mr. Masafumi Ito, who worked in Amazon Japan for more than seven years, sued Amazon Japan for wrongful contract termination.
12	Tokyo Managers's Union	Mr. Takeshi Suzuki; Chairman	Mr. Takeshi Suzuki have organized union activities for over to provide consultation and support for employees who feel their rights are violated by companies' practices, especially with the increased use of the performance improvement plan (PIP).
Legal expert			
13	Legal scholar	Hiroshi Miyashita	Miyashita lectures law in Chuo University and he is a leading legal scholar in Japan in Constitutional Law and Information Law.

The second set of interviews focuses on professionals working in Japan. During October and November 2021, using the snowball sampling method, we interviewed 18 working professionals from various sectors including high-tech, retailing platforms, hospitality, transportation and logistics, NGOs., etc. The interviewees were selected on the basis that they must be currently working for, or they have at least three years of working experiences in a Japanese company either in Japan or overseas.

The interviewees gave consent to the recording and interviews to be transcribed. All interviewees were explained that their names will be anonymized. Moreover, they were asked if their company's names can be identified, and most agree to their organizations or business areas to be identified given what they state are their own opinions and *do not represent the companies' views*. More importantly, all interviewees were explained when their opinions are quoted, names of companies are not mentioned. Key information regarding the working professionals are summarized in Table 5.5.

Table 5.5. Characteristics of the working professionals' groups.

Interview Subject No.	Companies/ Fields	Sex	Age	Working experience in Japan	Nationality	Working position/ rank
1	Oracle Netsuite Japan	Male	40s	6	Rwanda	Software Engineer
2	Glico Inc.; Misumi	Female	30s	10	Thailand	Business development/ Manager
3	World Family Co.	Male	30s	9	Japan	Sales/ Staff
4	Japan Post Office	Male	30s	10	Japan	Delivery employee
5	Transportation company	Female	20s	7	Japan	Global Business Department/ Staff
6	Tech companies	Male	50s	25	Vietnam	Management; Business development/CEO
7	FPT Holding	Male	30s	7	Japan	Consulting/ Staff
8	Hilton Group	Female	30s	10	Japan	Reception/Senior staff
9	PWC Japan	Female	30s	9	South Korea	Consulting/ Manager
10	Rakuten; GMO Cloud; Pipeline	Male	30s	9	Germany	Business developer focused on IT products/ Staff
11	Amazon Japan; Rakuten	Female	30s	10	Vietnam	Sales; Consulting
12	Hitachi Group; DSsama (an IT Start-ups)	Female	30s	9	Vietnam	Sales and Consulting/ Staffs

13	Treasure Data; Rakuten;	Male	30s	14	Vietnam	Software engineer/ Staff and self-starter
14	A labour import firm; IT Security firm	Male	30s	9	Japan	Sales and consulting/ Staff
15	Highschool	Female	30s	8	Japan	Teacher
16	Yamato Holding; Shiseido; Deloitte; Daison	Male	30s	13	Vietnam	Consulting for logistics services/Staff
17	Lawson; A consulting company	Female	30s	12	Vietnam	Sales and business development/ Staff
18	Manufacturing	Male	30s	8	Vietnam	Sales and business development/Staff

3.2.4. Citizen workshops

The third is qualitative data from four citizen workshops, each lasted three (3) hours, organized in March 2022 to collect the viewpoints from four distinct social groups: Japanese nationals of the two age groups: 18-35 years old (n=6) and over 65 years old (n=4), foreigners working and studying in Japan (n=10), Japanese people who have some forms of mental and physical handicaps (n=4).

The participants are presented with a short introduction of the current application of the technology via concrete examples of current use cases such as chapter 1's NEC and Realeyes AI for analyzing emotions during video calls. Then the participants are presented with a fictional story of a day-in-life of a person who uses and is subjected a range of emotional AI products embedded in daily activities: home hub smart assistant, targeted advertising, security cameras, music playlist, workplace, political news, toys, and cars (See Table 3.4.). For each use case, the participants are asked to provide their reactions for the use of the technology as well as their thoughts on its ethics and regulations.

Table 5.6. List of emotional AI use cases presented to during the citizen workshops.

Form of Emotional AI	**Context of uses**	**Narrative premise**
Home-hub smart assistant (Voice analysis)	Home Health Commercial	Home hub monitors user's voice and makes recommendations (e.g., daily schedule, dietary, doctor visits, local pharmacy products, etc.) based on analysis of mood and health conditions.
Bus station surveillance sensor (Biometric sensor)	Security	Citizens surveilled at transport hub and provided assistance or inconvenienced based on EAI.
Fake news/ Disinformation (Emotional profiling and trigger)	Civic discussion Social Media	Participants react to a deep-fake video and experience the social media platform's profiling response.
Spotify music recommendations (Voice and background analysis)	Commercial Social setting	Participants presented with new Spotify terms and conditions asking if they are happy with emotion and social setting data being collected to improve music recommendations.
Sales call evaluation and prompt tool (Voice and facial expression)	Workplace	Introduction of a system that monitors employees' expressions and tone of voice to evaluate job performance.
Emotoy (Voice and movement)	Children entertainment and education	Purchase of a toy that has emotional AI tool to collect and respond to children's emotional data, building a profile to make learning and marketing recommendations.
Rental car (Biometric, voice and movement analysis)	Commercial/ Road safety	EAI for in-cabin customization and driving recommendations. Identification when a driver is tired or distracted.

The data from these focus groups are transcribed and coded according to the qualitative thematic coding method presented previously.

Chapter 4: Workplace

It is no longer a surprise that Japan's workforce suffers considerably from the country's decades of declining birth rates and subsequent aging population. A report by the World Economic Forum projected in 2019, prior to the COVID-19 pandemic, that approximately 20% of the workforce, or up to 12 million Japanese people, will be out of the labor market by 2040 (Fleming, 2019). Before the pandemic broke out in end-2019, about 27% of existing work tasks in Japan were expected to be automated (Horii & Sakurai, 2020). This process has been accelerated over the past two years due to an extensive shift toward remote working and digitization to reduce the spread of the virus. For instance, a number of conveyor belt sushi chains around Japan have advanced moves to automate customer services, including the use of AI software to check in customers, to count the number and type of dishes consumed, and to pay the bill (Kamo, 2020). In other cases, besides installing self-checkout registers, convenience stores such as FamilyMart have also rolled out plans to use robot workers to restock refrigerated beverage shelves at 300 outlets by 2025 to ease the labor shortage (Yoshida, 2022). contribute to reshaping the Japanese work and consumer culture. Thus, a focal point of research on the future of work in Japan is on the transformative effects of digitization, automation, and new AI-power technologies on the traditional workplace (Schneider et al., 2019). This chapter explores the attitude of Japanese people toward the use of emotional AI technologies in the workplace.

4.1. The rise of emotional AI in Japanese workplace

Emotional AI is the more specific form of AI that identifies, tracks, interacts with, and/or reacts to human emotions through text, voice, computer vision and biometric sensing. In the workplace, emotional AI can be integrated into various processes ranging from recruitment, human resource management, to real-time monitoring of staff's emotional states, to name a few. AI software is recognized as useful in automating and streamlining parts of the recruiting workflow (Odyssei, 2019).

In the initial stage, companies might deploy recruiter chatbots or pre-hiring software to interview, screen, and shortlist applicants. One popular function is the utilization of AI-powered video analysis software in the interview and assessment of candidates. Some prevalent recruitment platforms in Japan include the U.S.-developed HireVue Hiring Platform and IBM Global Business Services, or other local services such as Preferred Networks, or Talent and Assessment Inc. (T&A). Companies have also developed in-house AI-powered candidate screening and hiring systems, as in the cases of telecommunication giant Softbank Corp. and brewer Kirin Holdings Co. (Horiuchi, 2020).

In other parts of the workplace, companies are also turning to AI to keep track of their employees' mental state of being and to micro-assess employee performances. For instance, staff recruiting group Recruit Holdings in 2018 implemented an internal program that utilizes AI to detect employees who might want to quit based on a host of data, including a past database of departed employees, current employee data and performance evaluation (Nikkei staff writers, 2018). The goal, the company said, was to let managers intervene early when such signs emerge.

The need to micro-manage employees looms larger because of telework during the pandemic. To help companies keep track of the health status of their employees, Futjisu Group has joined hands with the University of Tokyo to develop algorithms capable of giving the optimal advice based on individual health status (Okada & Iwatsu, 2021). The algorithms draw on a diverse set of personal health data, such as employees' physical and mental responses to stress, factors affecting stress responses, and various causes of stress (Fujitsu Website, 2022). Similarly, researchers from the University of Tsukuba in June 2021 suggested that companies could deploy an AI system using machine learning to predict psychological distress among workers, which is a risk factor for depression (Doki et al., 2021).

There is clearly an expanding utilization of AI-powered software in the workplace to monitor employees' mental being and provide employers with real-time workspace updates. For example, Preferred Networks, one of the most valuable Japanese AI start-ups, has developed a video camera system using deep learning algorithms to make assembly lines and factory floors safer (Interview data, 2021). Meanwhile, Empath's real-time detection of emotions from voices technology is currently being used via a Web API by over 2,000 companies and call centers over 50 countries. Nakamura Toru, CFO of TalentA, Hirevue's Japanese exclusive distributor concurred: "We are already asked by the hiring managers if the AI could detect honesty or some kind of that or mental illness in job candidates," while Hazumu Yamazaki, the co-founder and co-CEO of *Empath*, a voice analytic company, stressed: "So far we attract people from government sectors, especially, in the legal department, who want to use our technology especially for checking whether criminals are lying" (Interview data, 2021).

Emotional AI is inevitably part of the future of work, yet there remain serious concerns about its utilization. First, is the problem with algorithmic bias. As with any data-driven systems running on past data, EAI software is at risk of unconscious profiling bias. This is considering even programs that use training datasets avoiding the use of sensitive features like race, gender or age. Far from being statistically 'objective', the datasets themselves can replicate a programmer or even a society's innate preconceptions of race, gender or ethnicity (D'ignazio & Klein, 2020). Concurringly, Nakamura Toru, TalentA said: "In Japan, one of the key problems with existing hiring processes is how to empower women more. If the datasets are simply curated from existing workforce population, the same gender bias against women will emerge" (Interview data, 2021)

Second, while companies always stress data protection as a priority, there have been cases in which companies design workarounds to circumvent data protection obligations. One well-known controversy in Japan is the Rikunabi data scandal in which the Japanese job information portal Recruit Career Co. was found to have sold users' and students' data, such as the algorithmic scores of candidates, to client companies in 2019 without consent (Fumiko et al., 2020).

Third, unions in Japan and across the world have raised serious questions over the implications of AI-based management for the worsening condition of worker precarity. In a recent controversial court case, Amazon Japan is sued for wrongfully terminating a company employee contract. Moreover, according to Tokyo Union, Amazon Japan warehouse workers are current being monitored by AI that tracks various physical and mental states, these systems are allegedly allowed to terminate contracts of warehouse workers.

Next, this chapter will present quantitative and qualitative findings from a national survey, a students' survey and 31 interviews that have been thematically coded.

4.2. Study 1: A national survey of Japanese perception regarding the use of emotional AI in the workplace.

4.2.1. Descriptive statistics

Below is the explanation on how emotional AI might be used in the workplace that is given to the respondents.

> "Employers are interested in using technology to understand the emotional behaviour of their employees. For instance, some call centres are analysing emotion in their employees' voices to make sure they have the right tone of voice when speaking with customers. Other

businesses, such as retailers, are interested in using cameras to make sure that shop staff meet their standards of 'appropriate' facial expressions and behaviour towards customers."

Then, the respondents are asked to rate how likely they agree to a statement concerning the utilities, self-rated knowledge about emotional AI (derived from the TAM), and ethical implications of emotional AI applications (derived from the Moral Foundation Theory) in the workplace on a scale of 1 to 5 (1 means strongly disagree, and 5 means strongly agree).

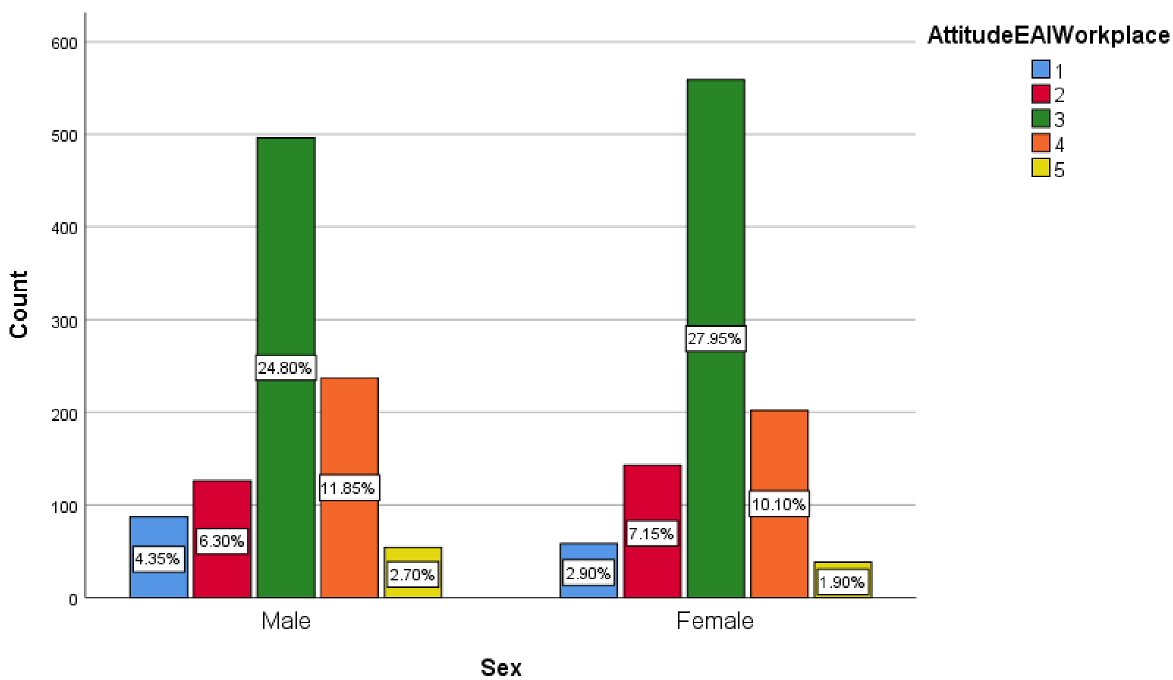


Figure 6.1. Attitude toward Emotional AI in the workplace.

Overall, based on the distribution of the questionnaire answer, regarding whether EAI in the workplace will be beneficial for society overall (i.e., the variable of attitude toward the technology in the work setting), there are slightly more people who report feeling positive about the technology. Figure 6.1 shows that 26.6% report feeling positive about the benefits of the technology, while 20.7% report feeling negative, the majority of 52.8% report feeling neutral.

4.2.2. Sex differences

Conducting the Chi-square tests, we found that the following variables exhibit statistically significant differences between the sexes: AtttitudeEAIWorkplace (*p=0.003*), AccuracyConcern (*p<0.001*), PrivacyConcern (p=0.002), ManagerNoAccess (*p=0.003*), FairnessConcern (*p<0.001*), FreedomConcern (*p=0.011*).

Specifically, we find that Japanese female respondents are on average express more worries regarding EAI's implications for freedom, privacy, accuracy, and fairness in the workplace. Moreover, female respondents are on average less positive about the benefits that EAI in the workplace will bring and feel more comfortable if the manager has no access to their emotional data.

Table 6.1. Distribution of concerns regarding emotional AI applications in the workplace.

Sex		FreedomConcern	FairnessConcern	ManagerNoAccess	UnionAccess	TrustPrivate
Male	Mean	3.35	3.48	3.25	2.99	2.88
	N	1000	1000	1000	1000	1000
	Std. Deviation	1.053	.951	.976	.991	.975
Female	Mean	3.45	3.64	3.36	3.00	2.92
	N	1000	1000	1000	1000	1000
	Std. Deviation	.960	.849	.887	.903	.897
Total	Mean	3.40	3.56	3.30	3.00	2.90
	N	2000	2000	2000	2000	2000
	Std. Deviation	1.009	.905	.934	.948	.937

Sex		PrivacyConcern	Knowledge	AccuracyConcern	AttitudeEAIWorkplace	TrustGov
Male	Mean	3.44	3.01	3.46	3.05	2.83
	N	1000	1000	1000	1000	1000
	Std. Deviation	1.019	.945	.945	.962	1.009
Female	Mean	3.58	2.97	3.60	3.02	2.83
	N	1000	1000	1000	1000	1000

	Std. Deviation	.939	.886	.839	.854	.928
Total	Mean	3.51	2.99	3.53	3.03	2.83
	N	2000	2000	2000	2000	2000
	Std. Deviation	.982	.916	.896	.910	.969

4.2.3. Socio-demographic factors

Table 6.2 presents the regression results in terms of socio-demographic factors.

Table 6.2. Regression results for socio-demographic factors in attitude toward emotional AI in the workplace.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	3.219	.111		29.122	.000
	Age	-.007	.002	-.107	-4.276	.000
	Income	.014	.033	.011	.415	.678
	Education	.026	.013	.051	1.947	.052

a. Dependent Variable: AttitudeEAIWorkplace; R square = 0.015. Note: * means p ≤ 0.05; ** means p≤ 0.01 means*** p≤ 0.001; **** means p ≤ 0.0001

For how socio-demographic factors influence the attitude toward applications of emotional AI in the workplace, we have two statistically significant results. First, age is a *negative* significant predictor of attitude toward the use of Emotional AI in the workplace (β_{age}= - 0.007, p<0.001). Second, education is a *positive* significant predictor of attitude toward EAI's application in the workplace ($\beta_{Education}$=0.26, p=0.052). Meanwhile, income has no statistical significant relationship with the dependent variable, which is contradictory to recent studies' results (Mantello et al., 2021).

4.2.4. Utilities, values, and concerns

Table 6.3. Regression results for behavioral determinants of emotional AI in the workplace.

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.742	.088		8.435	.000
	FreedomConcern	-.049	.020	-.055	-2.518	.012
	FairnessConcern	.017	.023	.017	.738	.461
	ManagerNoAccess	.175	.018	.180	9.926	.000
	UnionAccess	.275	.021	.287	13.403	.000
	PrivacyConcern	-.087	.020	-.094	-4.337	.000
	Knowledge	.127	.019	.128	6.684	.000
	AccuracyConcern	.014	.022	.013	.608	.543
	TrustGov	.230	.020	.245	11.443	.000
	TrustPrivate	.076	.022	.078	3.441	.001

a. Dependent Variable: AttitudeEAIWorkplace; R square = 0.554; Note: * means p ≤ 0.05; ** means p≤ 0.01 means*** p≤ 0.001; **** means p ≤ 0.0001

Regarding concerns about fundamental values, concerns about freedom and privacy are statistically significant negative predictors of attitudes toward the use of emotional AI in the workplace ($\beta_{FreedomConcern} = -.055^{*}$; $\beta_{PrivacyConcern} = -.094^{***}$). In other words, when people hold the concern about the intrusiveness of the emotional AI and its resulting loss in autonomy and freedom in the workplace, they tend to reject the technology. These results are aligned with the implications of the Moral Foundation Theory, i.e., the rejection of a new technology is a function of its violation of fundamental values (Hidalgo et al., 2020).

Interestingly, two other results that diverge from the Moral Foundation theory: the concern about biases against disadvantaged group in emotional AI systems and the concern for inaccurate emotional AI have no statistically significant association with acceptance of emotional AI in the workplace.

4.2.5. The importance of data governance

We also find our Japanese respondents are sensitive to the issues of emotional data governance. As those who want to deny manager access to emotional data and those who want emotional data to be shared with the unions that properly represent workers' right are more likely to trust the benefit of the technologies ($\beta_{ManagerNoAccess}$ = 0.18***; $\beta_{UnionAccess}$ = 0.287***). Perception of regulatory framework as provided by the government and the private sector also play a key role in people's overall trust for societal benefits of the technologies since we found people who express more trust in the government and private sectors' ability to regulate the technologies are more likely to report trust in the benefit of emotional AI use at work ($\beta_{TrustGov}$ = .245***; $\beta_{TrustPrivate}$ = .078***).

4.3. Study 2: A survey of international and Japanese students regarding applications of emotional AI in the workplace.

4.3.1. Being managed by AI is the greatest concern

Analyzing a dataset of 1,015 students' perception of emotional AI, first, this study discovers that being managed by AI is the greatest AI risk perceived by the international future job seekers.

We presented students with a list of nine ethical problems with AI proposed by the World Economic Forum (Bossman, 2016) and asked them to choose the top three. Interestingly, Figure 4.2.1 shows the top concern for international students' body is essentially about human-machine interaction, i.e., "Humanity. How do machines affect our behavior and interaction?" with 561 responses (55.3%). The second greatest concern, at 488 responses or 48.1%, is about the security of these smart systems, i.e., "how do we keep AI safe from adversaries?". The third-place is about unemployment with 467 responses or 46%, and the fourth-place is about unintended consequences of deploying AI with 445 responses or 43.8%. Although previous studies on AI integration at work have pointed out people are not concerned about AI replacement, at least in the short-terms (Pinto dos Santos et al., 2019; Sarwar et al., 2019), our survey results provide a more nuanced understanding of people's perception of various risks regarding automated management systems.

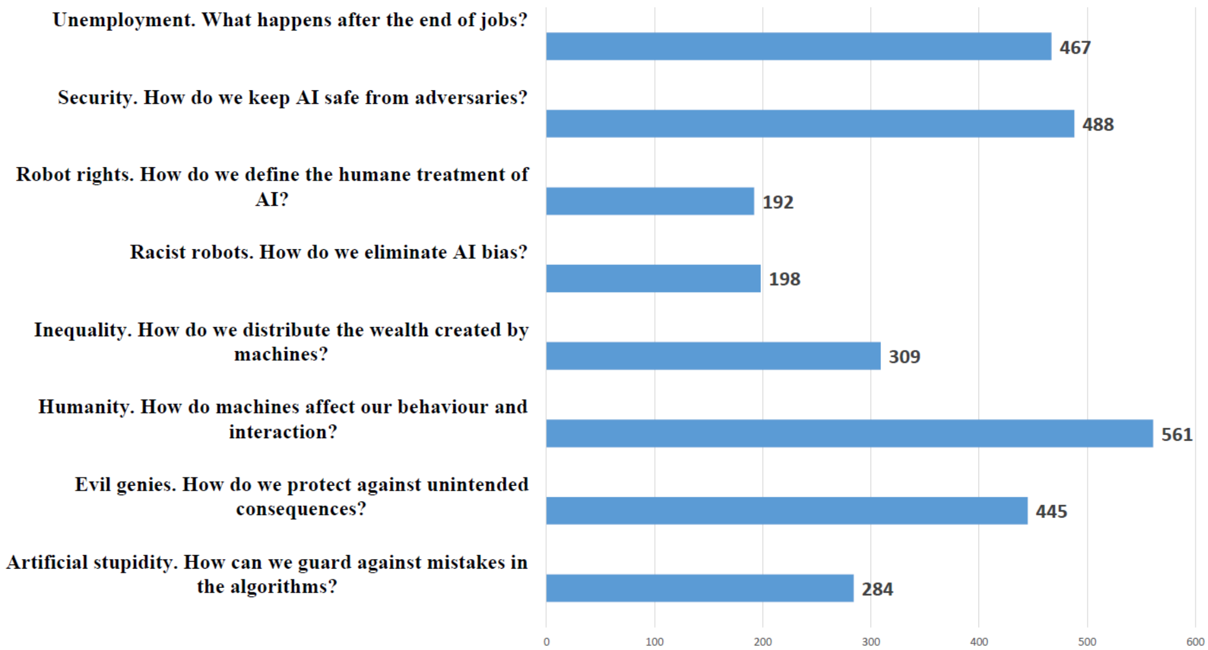


Figure 6.2. WEF's nine ethical concerns regarding AI ranked by the students.

Table 6.4. Distribution of familiarity with AI and attitude toward EAI-based HR management.

Variables	**Category/Group**	**Male (N = 437)**		**Female (N = 578)**	
		Frequency	**Percentage**	**Frequency**	**Percentage**
Familiarity with AI (1: Not familiar; 5: Very familiar)	1 to less than 2	42	9.61%	101	17.47%
	2 to less than 3	137	31.35%	239	41.35%
	3 to less than 4	207	47.37%	202	34.95%
	4 to 5	51	11.67%	36	6.23%
Attitude toward automated management (1: Very worried; 5: Not worried)	1 to less than 2	45	10.30%	74	12.80%
	2 to less than 3	149	34.10%	260	44.98%
	3 to less than 4	200	45.77%	214	37.02%
	4 to 5	43	9.84%	30	5.19%

Table 6.4. also shows 52% of the future job seekers express negative concern about the EAI-enabled HR management, 51% rated themselves below average regarding AI knowledge.

Figure 6.2. shows that human-AI interaction is the top ethical concern with nearly 55% of the total responses, while job loss to AI only ranks third with 48%. These insights will prove crucial when communicating in educational settings about the risks of AI. As the workplace moves toward a more invasive form of neo-Taylorism where AI tools seek to go beyond the exterior of the physical body and datafy our emotional lives (Marciano, 2019; Richardson, 2020), our results suggest young jobseekers have started express a greater level of concern regarding AI supervising and making decisions about their performance and career advancement, rather than AI replacing their jobs.

Moreover, the analytical insights highlight the urgent needs for better education and science communication concerning the risks of AI in the workplace. As the data on the level of awareness of EAI among the future job seekers, although nearly 80% picked a very close definition of EAI (Figure 4.2.2A), when students are asked to rate their level of familiarity with EAI, roughly 40% rate themselves as unfamiliar or very unfamiliar and 36.7% of the respondents are unsure of their level of knowledge (Figure 4.2.2B).

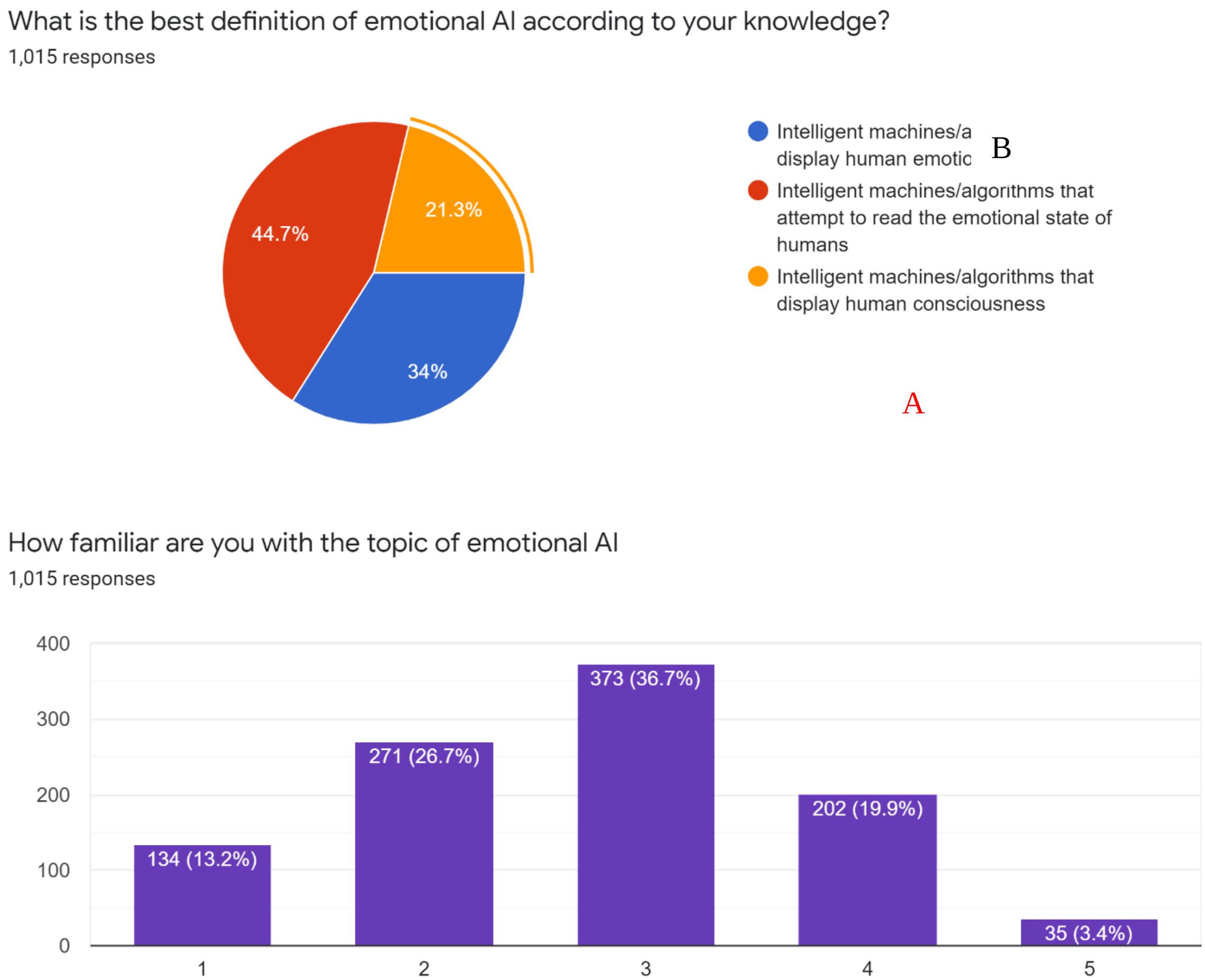


Figure 6.3. Familiarity of the respondents with EAI. A) Students choose among three definitions of EAI. B) Students rate their familiarity with the topic. The scale is from 0 (not familiar) to 5 (very familiar)

4.3.2. Cross-cultural differences in attitude toward emotional AI for workplace management

In exploring the effects of various factors on the attitude toward emotional AI-based human resource management via the Bayesian MCMC approach, this study also highlights various cross-cultural and socio-demographic discrepancies in concern and ignorance about the EAI-enabled management of the workplace that must be bridged to bring more equalities to the AI-augmented workplace. The analyses show people from different socio-cultural, economic backgrounds do tend to form different perceptions of emerging technologies: We consistently people from a more dominant social class (being male and being from a higher-income background) are likely to have less anxiety toward EAI-based HR management.

Figure 6.4. presents the correlations between socio-demographic factors and attitude between emotional-AI-based HR management. It shows students with higher income, the male gender, business major, and being in senior are likely to have a less-worried outlook toward EAI-enabled HR management. Regarding income, an explanation might be the students with higher income are likely to have higher educational attainment (Aakvik et al., 2005; Blanden & Gregg, 2004) and end up in high-status occupations (Macmillan et al., 2015); thus, in all likelihood, they are more likely to become future managers who will use those AI tools to recruit and monitor their employees. Regarding the sex variables, validating H2 and H9, our result is aligned with the literature showing male-ness is correlated with higher perceived technological self-efficacy (Cai et al., 2017; Huffman et al., 2013). Being a business major is correlated with less anxiety for EAI-enabled HR management might be a product of the lack of emphasis on AI's ethical and social implications in business education. Another reason may be that hoping to become a manager would incline a person to adopt the company position, thus seeing management supervision only in terms of productivity and performance results. Future studies are required to understand the underlying cause.

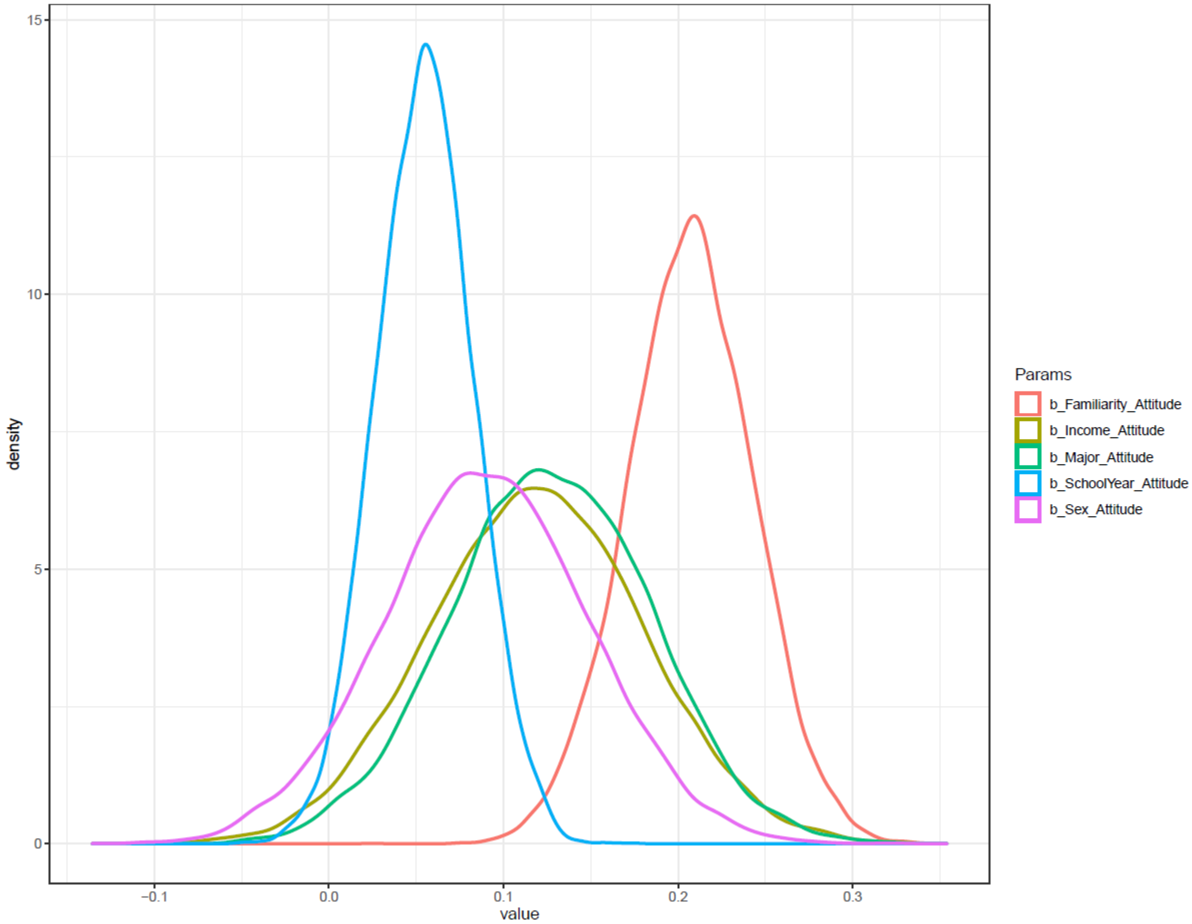


Figure 6.4. Density plot from Model 10 for five variables: familiarity, income, major, school year, and sex.

These factors are also correlated with a higher self-rated knowledge for AI and as demonstrated in Figure 6.4, self-rated familiarity with AI has a positive correlation with the attitude toward AI's use in HR setting (***β_Familiarity_Attitude's*** mean = 0.21, sd = 0.04). This finding implies that students who rated themselves to have more knowledge of AI might be unaware of the biases in and inaccuracy of emerging technologies. Taken together, these facts indicate many students might be ignorant of the ways in which social biases and privileges can lead to harmful EAI's use in the workplace, as shown in various studies on algorithmic biases (Rhue, 2019; Crawford, 2021, Moore and Woodcock, 2021; Buolamwini & Gebru, 2018). Even though the problem of algorithmic bias has now moved to the center of public discourse in Western media (Singh, 2020), when it comes to a multi-national sample this study indicates a clear lack of knowledge as 51% of the respondents rated themselves below average in AI knowledge (Table 6.4.).

Past studies have shown student engagement with ethics is contingent on several factors: first, the type of curriculum adopted by higher education institutions (Culver et al., 2013); second, how the concept of bias is communicated and understood through the course literature. As such, our study indicates that university curriculum would strongly benefit from inclusion of courses on social and ethical implications of AI in the workplace, especially in the business major, which has been shown to correlate with less concern about AI in HR management in this paper (see Figure 7 and H3). This is to correct any students' misconceptions and enrich their understanding of the positive and negative potential of such technologies. Given the strong emphasis on the importance and advantages of acquiring data analytics skills in current curriculums of AASCB-accredited business schools (Clayton & Clopton, 2019), ethical training and critical thinking about the ethics of these technologies should be integral to institutional higher learning epistemology that prepares younger generations for the quantified workforce.

Here, it is worth mentioning previous studies show that an employees' awareness of the presence of the smart surveillance technologies negatively correlates with organizational commitment (Ball, 2010; Brougham & Haar, 2017). These two tendencies combined with the risk of AI being misunderstood (Wilkens, 2020) are important obstacles to overcome before such technologies can be harnessed in ways that safeguard the worker's best interests.

4.3.3. Regional differences

Our analysis also shows people from economically less developed regions (Africa, Oceania, Central Asia) exhibit less concern for EAI-enabled management, while people from more prosperous regions (Europe, Northern America) tend to be more cautious. Interestingly, however, an economically prosperous region such as East Asia correlates with less anxiety toward the EAI-enabled HR management. Our data in Figure 6.5. show that, for East Asian, 63.62 % of the Japanese, 56.32% of the South Korean, and 41.77% of the Chinese respondents express a more accepting attitude (averaging the score of equal or more than 3 in the attitude scale). While for European and Northern Americans, an overwhelming majority of 75% possess the worried attitude toward being managed by AI. Since these East Asian countries have different political systems, the consistency of accepting attitudes for EAI across these countries could be explained by a common factor—Confucianism. Specifically, there might be antipathy toward individual rights in Confucian culture (Weatherley, 2002), as well as stronger emphasis on harmony, duty, and loyalty to the collective will (Vuong et al., 2020; Whitman, 1985). Finally, in Confucian culture, there is much more acceptance of intervention by higher authority as it is thought of as a source of moral guidance (Roberts et al., 2020). As for why there is the largest percentage of Japanese young job-seekers who express a positive attitude toward emotional AI applications in the workplace, it is possible that this embrace of technologies in the decision-making process of the workplace is young Japanese' reaction against the traditional social hierarchies dominated by elderly men in Japan.

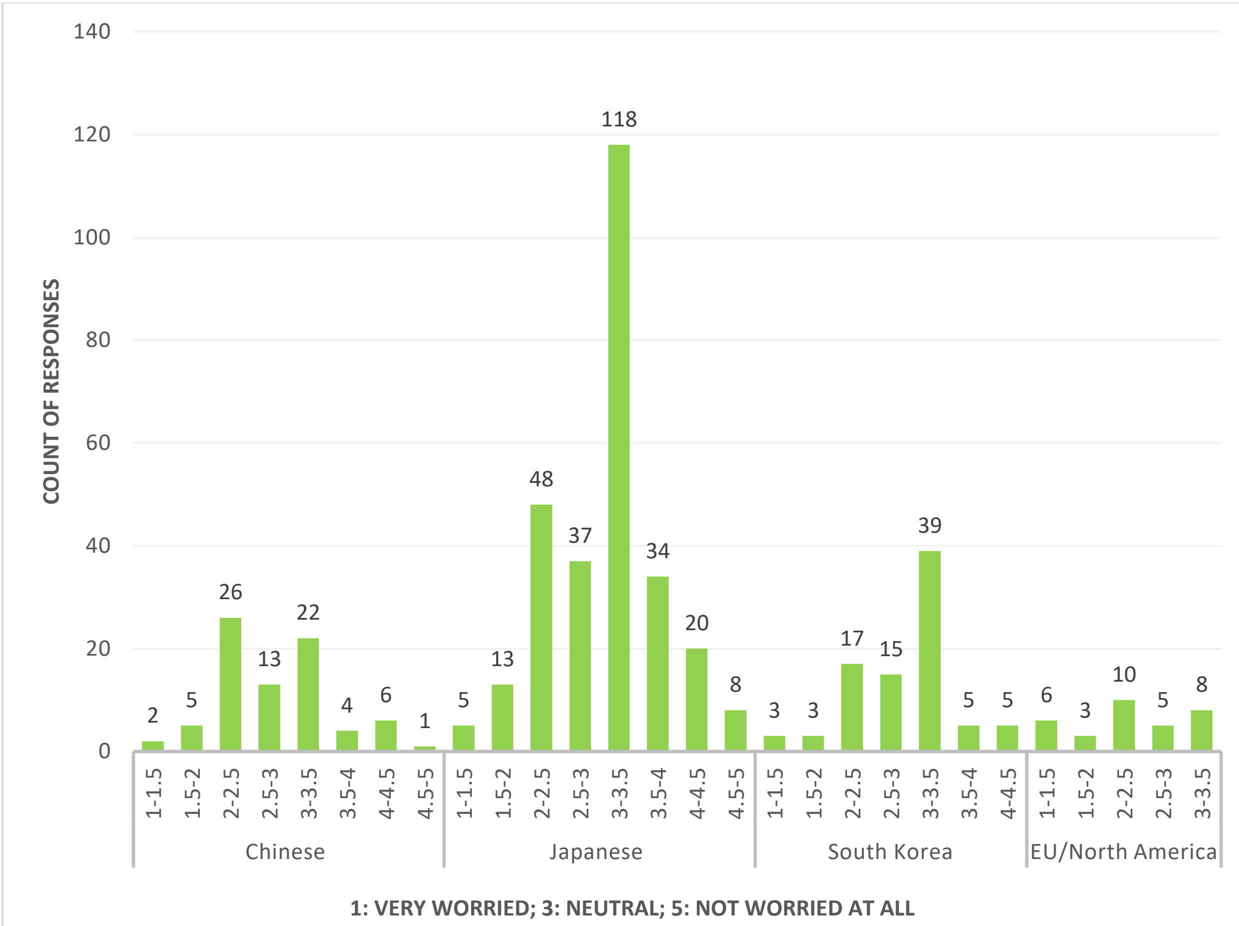


Figure 6.5. Comparing the distribution of different attitudes toward EAI-enabled management by three major East Asian countries (China, Japan, Korea) and Europe/North America.

Such cross-regional and cross-cultural differences prompt us to further investigate the differences among the top 10 countries represented in our sample size. Controlling all other socio-demographic and behavioral variables, the Japanese have the strongest correlation with an accepting attitude toward EAI in HR management, followed by the Vietnamese, Chinese, and Korean (See the Supplementary file). Indians, on the other hand, correlate with the highest level of anxiety toward automated management followed by their Bangladeshi and Indonesian counterparts. The Japanese participants' lack of reservation for EAI-based management is perhaps unsurprising given the extent to which workplace norms and conventions dictate unquestioning obedience, loyalty and mandatory volunteerism (Stukas et al., 1999), especially in relation to managerial superiors (Meek, 2004; Rear, 2020). For example, it is an unspoken convention in Japanese corporate culture that no one leaves the office before the *katcho* (office head) does. Our findings suggest that as a more invasive form of automated management, EAI may exacerbate anxiety amongst foreign workers in Japan, opening up the possibility of conflict with Japanese managers who are culturally conditioned to value conformity, loyalty and to punish 'attitudinal diversity'. As the Japanese saying goes, "出る杭は打たれる*",* (*deru kugi wa utareru*—the nail that sticks up must be hammered down) (Sana, 1991; Luck, 2019).

Overall, the study 2 demonstrates cross-cultural factors are indeed important in predicting the attitude toward emotional AI-based HR management. This result contradicts theories such as Technology Acceptance Model or Theory of Planned Behaviors or Theory of Reasoned Action, that only prioritize the cost and benefit calculation in predicting human behaviors (Davis, 1989; Taherdoost, 2018). The empirical findings on such stark cross-cultural and cross-regional differences could help educators, businesses, and policymakers to shape their action programs to address any stakeholder's concern or lack thereof for the future of AI-driven work.

4.4. Study 3: Qualitative thematic analysis of interviews and citizen workshops

4.4.1. Cultural conflicts: trust, privacy and biases

The first level of cultural discord is at the data level, it arises from cross-cultural differences and legal distinctions over a worker's privacy, and what constitutes sensitive data. For example, even though current and proposed AI regulations in Europe and North America are, on the surface, compatible with existing data privacy laws in Japan, in the Japanese workplace, national labor law says almost nothing about health data privacy in the context of the workplace.

IBM Global Business Services' executive, Christian Vlad, with more than 25 years of work experience in Japan, stated: "Not a single European organization would ever consider requesting their employees to give them access to their health data, as this is a matter for lawsuit. But in Japan, it is quite commonplace" (Vlad C, 2021, Interview data). Concurring, the majority of our interviewed working professionals confessed that it is the norm for employees to 'volunteer' their data, e.g., their video conferences or calls being recorded in Japan.

Many interviewed employees have explained the rationale as follows. A Japanese employee (Male, 30s) working for a national transportation company state: "Currently, my company uses the tablet-based tracking system and the managers can see all of my on-the-road activities: how much time on the road, how much time in each stop, where the person is in real-time, etc. So far, I have no negative experiences. I personally think it is a good thing as it improves alertness, efficiency, and reduce road accidents." Another interviewee (Male, 30s, Japanese, Education sector) express the pressure of being monitored makes him perform better. He further elaborates that his company is using a tablet-based system that records all locations and times when he meets with clients. Concomitantly, a female interviewee (Female, 30s, Vietnamese expat, E-commerce) also thinks: "I think managers lack tools to visualize the efforts or stress levels of employees. A lot of the time, in my experiences, I think managers are oblivious to whether an employee is experiencing overcapacity. Thus, AI tools can be very useful if there are clear guidelines for how and when it is being used."

Yet, as Japanese corporations have unchecked access to health data, there is a risk of workers being put into an ever weakened and precarious power relations. The staff of the transportation company concurs: "AI monitoring of delivery activities will make it easy for anyone to replace my job." An interviewee (Male, Japanese, 30s) working in a start-up company also express his worries: “I would resist automated tracking of all the details of my work activities as well as mental states if I know there is a huge wage gap between the management and the employee. The situation seems very unfair.”

Our interviews with working professionals in Japan also reveal that, except for two people who work for global companies and have real experiences with obtaining data licenses, the current commonsensical understanding in Japan of what it means to protect personal data is surrounding the practice of anonymization and obtaining consent. The majority of interviewees who work for smaller firms admit their companies/managers have very weak training in data privacy and protection. An interviewee (Male, 30s) states: “I don’t feel like the personal data protection practices exist at all. We just have to make sure as long as information such as names, addresses, or birthday are not revealed. Beyond that, there is not much else for the companies and us to do.” However, these practices are ill-prepared for the rise of AI systems that create people’s behavioral profiles to manipulate users’ moods and behaviors. For example, AI-driven management system might see correlations between certain personal factors such as age, gender, personality traits, etc. and behaviors of employees, and making recommendations for the management to optimize performance without the employees knowing it. Thus, again, AI-driven management practices such as these can put employees in a further weaker position.

Generally, the notion of privacy in Japan as with other Confucian/East Asian societies is convoluted, especially in a communal context such as the workplace. Miyashita (2011) argues the notion of privacy in Japan is undergoing a process of acculturation whereby the traditional notions of privacy as symbol of trust in the relationship of self and the collective is interacting with the Western privacy notion rooted in the respect for individual liberty (Miyashita, 2011). In our interview with Miyashita (2022), the data privacy expert worries that if companies in Japan decide to use algorithms to detect lies or mental illness of their employees, there is little or no regulation against such use. There are also serious questions over the ability to reinforce data protection laws because even though there have been many public data breaches scandals, the national agency of Personal Data Protection Commission (PPC) has not issued any fine since its establishment of 2003. It has been noted that staffs of PPC are from the ministries and the commissioners are from various industries, which raise questions over conflicts of interest.

4.4.2. The lack cultural sensitivity in algorithmic design

Many of the current AI tools used to monitor and assess Japanese workers, having been designed in the West, operate with a one-size-fits-all approach to their design. Not only do the algorithms driving these technologies come embedded with pre-conceived, Western-centric ideas of the ideal worker, the datasets used to measure performance and set benchmarks for productivity are often taken from North American and European sample populations. Although vendors and their company clients we interviewed are aware of such limitations in their product design, they still cling to the belief that the issue will resolve itself by future development of a more inclusive, ‘perfect’ algorithm.

For example, all of the companies in our sample, whether local or global, expressed an awareness of the need to accommodate cross-cultural differences in building emotional AI algorithms. A representative from ELSYS Japan, the distributor of ELSYS—a controversial Russian company that produces a software that purportedly detects ‘suspicious behaviors’ from the head’s micromovement, acknowledged that their technology does not know the difference between races (Interview Data, 2021). Meanwhile, a spokesman of a Japanese giant facial recognition company’s emotional analytics team admitted the need to have a better geographically specialized algorithm to deal with race, gender and cultural diversity (Interview Data, 2021).

Critically, none of the stakeholders from the companies interviewed, however, were able to provide any confirmed, detailed plans to move away from Paul Ekman's six basic emotions model, a theory that has fallen into disrepute (Crawford, 2021; Barrett, 2017) as recent literature on algorithmic bias suggests that machines express systematic biases in reading the emotions of people of color and minority groups (Buolamwini & Gebru, 2018; Purdy et al., 2019; Rhue, 2019).

The interviewed companies allude to making the efforts to rectify the current disjuncture between discourse and practice by using a kind of hybrid model and collecting more localized datasets. For example, Empath's co-founder, Hazumu Yamazaki confirmed the use a model of four emotions (sadness, calm, joy, and anger), which is a reduced version of Ekman's six model (Interview Data, 2021). Other interviewed companies confirmed the use of local data subjects and local in-house annotators for training datasets in the hope that the obtained data reflect accurately the local nuances in emotional expressions. Although all interviewed companies express an interest in gaining access to more localized datasets, the problem at the algorithmic design level, i.e., the problem with the scientifically controversial theoretical model of emotions, remained unresolved. For instance, TalentA (HireVue's exclusive distributor) extracted data from 1,000 model answers curated from Japanese college students who obtained jobs from large corporations, nonetheless, the questions given to the students were originally designed by US psychologists. TalentA's Toru Nakamura admitted the difficulty of making the translated versions of the questions an appropriate and polite fit for local and cultural context.

In the case of AI-driven management practice, the majority of our interviewed employees express serious doubts of the accuracy of the technology and whether it could fit into Japanese team-work culture. An interviewee (Male, 30s) professed: "In performance tracking, a value of a task is judged by the team rather than a manager. Thus, it raises difficult questions for a blanket application of AI-based performance monitoring.” Another (Male, 30s, European, IT industry) concurs: “"I think there must be a better culture of problem-solving to understand why a tech-solution is necessary. In my experiences until now working for Japanese companies, tech solutions are often introduced for the sake of face-saving: To look as if management is doing something. They don't often solve the problem they are supposed to solve but create more unforeseen problems." (Interview data, 2021).

The efficacy of AI products used for management practices is contingent on many cultural changes in many interviewees’ opinions: “If performance tracking is focused on each individual, for example, one person is in charge of one client, rather than a group being in charge like my companies are doing, then I can see AI will make performance tracking and review more reliable.” (Female, 30s, Electronics and IT industry). Meanwhile, a Rwandan software engineer who has experiences using an AI system to manage his staffs say: “In my experiences, for Japanese companies, interactions between managers and staffs does not change in the present of an AI-driven platform. The numbers (produced by these AI systems) are used best as tools of reminders and reflections, rather than for performance evaluation” (Male, 40s). It should be noted that reflection meetings (反省会), i.e., often several hours long meetings that reflect on the process of past performances, are very important part of the Japanese working culture. Thus, the informant alludes to the fact that the use of AI in the workplace should conform to the native cultures and meet the goals and purposes of each department. Concurringly, IBM Global Services, Cristian Vlad states: “Any advanced technologies must be presented in a culturally intelligent manner if they want to be adopted by Japanese companies.”

4.4.3. AI undermines trust in the workplace

Numerous studies have shown that workers placed under constant monitoring experience lower degrees of motivation but also, significantly, higher degrees of stress and anxiety (Bondanini et al., 2020; Brougham & Haar, 2018). In a Japanese context, automated management translates into a ubiquitous form of surveillance that violates time-worn, unspoken bonds of trust between employer and employee.

For example, Amazon Japan is now embroiled in a series of labor disputes largely due to its culturally insensitive performance improvement plan for errant staff and general hostility toward collective bargaining (Ishibushi & Matsakis, 2021). Takeshi Suzuki, the chairman of the Tokyo Managers' Union, stated "from a purely Japanese perspective, the use of AI to monitor workers signifies there is no trust from the corporation towards its workers" (Suzuki T., Interview data, 2021). Moreover, the fact that Amazon uses temp agencies as intermediaries to hire warehouse workers, means they can cancel the contracts of those they deem unfit for their culture. In Amazon, 6-10 % of low performance people will be fired or put into a performance improvement plan, that is hard to get out. Yet this challenges the traditional work culture in Japan which values loyalty over productivity and instead focuses on solidarity, long-term trust, and human growth. According to Suzuki, while labor protection law in Japan is very strong, rarely do workers engage in lawsuits or grievances. Moreover, he also stated that very few workers join unions to exercise their rights against unfair practices at work. This reluctance to assert themselves in the case of wrongful dismissal can partly be explained by the influences of Confucianism in Japanese culture, an unwavering compliance and submissiveness to authority but also, the uncontested bond of trust that exists between employer and employee (Miyashita, Interview Data, 2021).

This observation is not far from the employee's perspectives. An interviewee (Male, 50s). claims: "It is often the case that Japanese companies, even if they are global companies, they want to stay Japanese. In Japan, reading between the lines and paying attention to what the other people thinks are very important in the workplace. The prioritization of numbers provided by the AI over the human relationships would signify a serious lack of trust." He also mentions his direct experience in a case where his Japanese headquarters rejected the proposal of foreign branches to use AI systems for performance review.

A female Japanese employee in her 30s of a transportation company concurs: "In Japanese culture, at least in my experiences, people have an old-fashioned way of thinking. They are not trusting the technologies. It is not about money; it is more about trust…They trust people more than the AI although they understand using AI might be more cost-effective." This attitude is the polar opposite of the attitude of Amazon Japan, in which, as mentioned above, 6-10 % of low performance people will be fired or put into an allegedly arbitrary performance improvement plan (Ishibushi & Matsakis, 2021).

4.4.4. Gendered bias: Can AI help improve traditional gendered bias in the Japanese workplace?

Most of our interviewed employees tend to also express worries over gendered bias in Japan. A Japanese female informant (30s, Logistics) stated "Women in the Japanese companies are not afforded opportunities to showcase their talents such as going to business meetings, travels, finding more customers, etc.". Another agrees with this point: "In practice, I think there are many unspoken disadvantages for women in Japanese companies. It happens to me once when I was declined a chance of a promotion due to pregnancy" (Vietnamese, Female, 30s). However, as to whether AI-driven management would improve or exacerbate the situation, it is a split between optimism and pessimism.

According to the legal scholar at Chuo University Miyashita, in some ways, Japanese young employees, especially the women, would tend to embrace the automated management system as a reaction against the traditional dominant social hierarchies. They are rather being judged by a machine than being judged by an elderly, male superior (Interview Data, 2021). An employee of an e-commerce company expresses her desire for more AI applications: “In my experiences, managers do not have tools to visualize the workloads and stress levels of their employees. Therefore, a tool such as AI systems that help in that aspects can be a plus for the modern Japanese workplace (Female, Interview Data, 2021). Nakamura Toru, from TalentA and York Date, from Business Research Lab all agreed that it is especially important in Japan to make sure the training datasets for AI do not replicate the current inequalities in the managerial ranks of the modern Japanese workplace. They confirm that their companies have taken steps to ensure equal representation of male and female in their training data.

However, many other interviews remain from agnostic to highly critical as to whether AI systems will be useful in combatting gendered bias in Japan. As for why, they cite two reasons. One is a deep reluctance for adoption of new technologies in traditional workplace. A male employee at a manufacturing company states: “I always feel that Japanese elderly managers understand that machines are better than humans in many aspects. But until the next generation of managers, who might be more familiar with the technology, I don’t see its adoption anytime soon” (Interview data, 2021). The second is that many aspects of modern Japanese companies are still paper-based, and digitalization will not happen as fast as it should. This is confirmed in the interviews with Business Research Lab, TalentA, and IBM Global Services.

4.5. Chapter summary

Chapter 4 has presented a 360 view of social perceptions of emotional AI in modern Japanese workplace via four empirical sources: a national survey, a survey of young jobseekers in Japan, 31 interviews with various stakeholders, and 4 citizen workshops. The quantitative analyses of the surveys reveal significant sex differences. In the national survey, we find that Japanese female respondents are on average express more worries regarding EAI's implications for freedom, privacy, accuracy, and fairness in the workplace. Moreover, similar to the jobseekers' survey, female respondents are on average less positive about the benefits that EAI in the workplace will bring. Interestingly, as a reaction against the traditional male-dominant hierarchy, many interviewees see AI-based HR management as their allies in improving the situation of gendered bias in the Japanese workplace. These findings imply just how important it is to have more women in decision-making position to ensure a fairer and more humane applications of emotional AI.

Interestingly, the analysis of the national survey reveals concern about the accuracy of emotional AI and concern about social biases in these automated systems do not have a statistically significant relationship acceptance of the technology. This suggests people would tend to accept the technology despite of its flaws. In addition, it reflects a cultural belief in homogeneity among the Japanese respondents.

Many cultural tensions that can arise from a more widespread adoption of emotional AI in the workplace are also uncovered. First, our interviews and citizen workshop participants reveal that the use of AI for management of workforce is likely to be seen as a betrayal to the traditional value of employer-employee trust that has been an essential feature in Japanese society. Second, serious questions have been raised by participants in our interviews on the cultural sensitivity of AI systems that will be used in the Japanese workplace. In responding to how to resolve these tensions, our interviewees argue it is necessary to articulate various aspects of Japanese workplace culture that might be impacted by the use of AI, and thus, creatively using the emerging technologies in ways that respect the culture.

Chapter 5: Education and Toys

This chapter systematically explores citizen perceptions of the applications of emotional AI in education and toys. As mentioned in the previous literature review as well as the introduction, emotional AI technologies are introduced to schools and children's toys under the rhetoric that they can deliver personalized learning, enhanced effectiveness in interventions when children are struggling emotionally. Kate Crawford reports in a Nature article that 4 Little Trees, an AI system developed in Hong Kong has been introduced in schools to monitor children's emotions (happiness, sadness, anger, disgust, surprise and fear) in the classrooms to gauge 'motivation' and even forecast grades (Crawford, 2021b).

The logic behind this modern phenomenon of modifying children's behaviors in educational setting is that academic learning needs to be supplemented by social and emotional learning. Here, emotional AI is often seen as a solution to operationalize such social and emotional learning (Williamson, 2017). Earlier works in this area have pointed out a multitude of problems with emotional AI in schools: effectiveness of the technology in supporting learning (Williamson, 2021), accuracy of reading emotions, and inclusiveness of training data in emotional AI edtech, the desirability of the feeling of inhibition and excessive self-consciousness on the students' parts (McStay, 2020a). It is important to critically evaluate whether the financial incentives of private companies are aligned with the well-being of our next generations.

In this chapter, the focus is on the Japanese perspective, particularly with regard to how the Japanese public perceives the adoption and integration of emotional AI technologies in schools and toys. The national survey and transcripts from the citizen workshops serve to help answer the research questions in two educational contexts of schools and toys:

- How do sociodemographic factors (sex, income, educational qualifications, etc.) influence perceptions of emotional AI applications?
- How do the concerns for fundamental values such as privacy, autonomy, safety, etc. of correlate with the attitude toward emotional AI applications?

5.1. Emotional AI in educational facilities

5.1.1. The Japanese context of AI in educational setting

In Japan, utilizing AI in education has been considered one of the solutions for the shortage in domestic human resources. Here, the crucial utility of AI in education is to enhance the effectiveness of learning by identifying better methods and areas to focus on for the students, thus shortening the time for training and quickly preparing the students to entering the workforce. Such logic has been epitomized in the 2019 "AI Quest" initiative, launched by the Ministry of Economy, Trade and Industry (METI, 2019); as well as the 2022 "Digital Agency's Roadmap on utilization of data in education" initiative by Ministry of Internal Affairs and Communications; Ministry of Education, Culture, Sports, Science, and Technology; and Ministry of Economy, Trade and Industry (Digital Agency JP, 2022). The AI Quest initiative sets out to achieve goals such as "resolving the shortage of AI human resources" and "developing AI human resources" (METI, 2022), while the Digital Agency initiative sets out a long-term goal to create a digital environment where learners can store and utilize their own data over the course of their life (Digital Agency JP, 2022).

Noteworthily, AI technologies powered with affective computing capabilities are also considered as a potential solution for the arduous, long hour working hours of Japanese teachers. According to a national survey conducted by Uchida, Nagoya University and colleagues found over 70% of junior high school teachers in Japan have overworked by 80 hours each month, which meets the technical threshold for determining death by overwork (過労死—karoshi) (Lee, 2022; Matsushita & Yamamura, 2022).

In a recent article, Yamada Seiji, Professor of Information Science at the National Institute of Informatics has provided an overview of the current trend in adoption of AI technologies in Japan. Seiji (2018) identifies emotional AI as an important tool for supporting the learning process and enhancing the pedagogical effectiveness. Citing the famous example of the *duolingo* app, a global EAI-enabled service that is popular among foreign language learners in Japan, Seiji notes how role-playing game in the app has generated more fun in learning a new language. Seiji thus urges it is important to leverage the strong interest of youngsters and young adults in smartphones game for educational purposes. It is reported that in Japan, there are more than 28 million Japanese smartphone game users, mainly men in their teens, and elementary, junior high and high school students.

AI startups for educational purposes have sprung up in Japan in recent years. For example, NEC, a security conglomerate in Japan, launched in 2020 an AI tool that analyzes emotions of participants in video conferences or zoom classes (Abe & Iwata, 2022). Or Qubena, an AI tool provided by the Tokyo-based COMPASS Inc., is promoted as capable of providing the best questions that match a student's learning level (AI Smiley, 2022). The AI-powered teaching material is reportedly used by about 500,000 people in more than 1,800 elementary and junior high schools nationwide. As Qubena is integrated into the online learning system MEXCBT[1] of the Ministry of Education, Culture, Sports, Science and Technology starting September 2022, the

[1] MEXCBT is the combination of MEXT (Ministry of Education, Culture, Sports, Science and Technology) and CBT (Computer Based Testing). MEXCBT, whose first prototype was introduced in 2020 and became official in November 2021, is envisioned as an online test bank that students, with one terminal per user, can use them to study. See https://g-apps.jp/ict-education/about-mexcbt-and-e-learning-portal/

public's familiarity with adaptive learning systems will increase (Compass, 2022). Another famous case is *atama +*, a tablet-based AI tool that can assess students' level of comprehension, mistakes, learning history, concentration, etc. in real time to propose the shortest curriculum with the highest learning effect. The company advertises that its product has been adopted by more than 3,200 educational facilities across Japan and in one case, it has been shown to provide nearly a 6 times reduction in the learning time (AI Smiley, 2022).

Digital learning is not new in Japan, given that private establishments such as the Uchida Yoko Institute for Education Research, founded in 1998, have long strived to develop online learning or computer-based testing systems. The Japanese government even launched the JPY460-billion (USD3.12 billion) Global and Innovation Gateway for All (GIGA) School Program at the end of 2019 with the aim of enhancing digital learning experience nationwide (Ishizaki, 2021). Yet, it is worth noting that major developments in education technology in Japan took place after a number of global reports, including the OECD's TALIS 2013 (OECD TALIS Report, 2013) and PISA 2015 (OECD, 2017), revealed a below-average use of computers and Internets among Japanese students. What followed was a surge of public spending on improving the hardware infrastructure for education in the new era. The five-year infrastructure plan for ICT in education (2018–2022) allocated local governments with an annual budget of JPY180.5 billion (USD1.7 billion) from 2018 to 2022 (MEXT, 2018), which means each elementary school receives about JPY6.22 million yen (almost USD60,000) (Gazzano, 2021).

In examining Japan's vision of Society 5.0 in the education sector, Holroyd (2020) suggests that connecting education to the national priorities should not be an issue considering Japan's unitary state and historical pattern of letting its strong national government control the centralized education system (Holroyd, 2022). However, the disconnect between the central policy and the local implementation remains a substantial point of concern. Even with the push of the COVID-19 pandemic toward remote learning, the rate of smart technologies adoption, including AI services, in Japanese schools is very slow. One of our informants, who is a teacher at a high school in Kyushu Japan, states: “There has been a push toward using more tablets and smartphones to facilitate the students with doing class works during the pandemic. However, it has never come into fruition” (Interview Data, 2021). This is in line with the literature on the surprising lack of readiness for remote learning in Japan (Masami, 2021; Sato, 2020). Thus, it is important to investigate what are the reactions and perceived threats of emotional AI in the educational setting.

5.1.2. Descriptive statistics

Below is an explanation given to the survey respondents. Then, the respondents are asked to respond to a series of Likert-scale questions on the topic of emotional AI in school.

> “Schools in some countries are employing companies to install cameras and artificial intelligence in classrooms to track students’ facial expressions to try to work out their emotional states and attention levels. This aims to tailor teaching approaches by understanding if some students are struggling with class material or if other students need to be challenged more. It also aims to identify students’ attention levels, to help teachers to monitor and record in-class attention levels.”

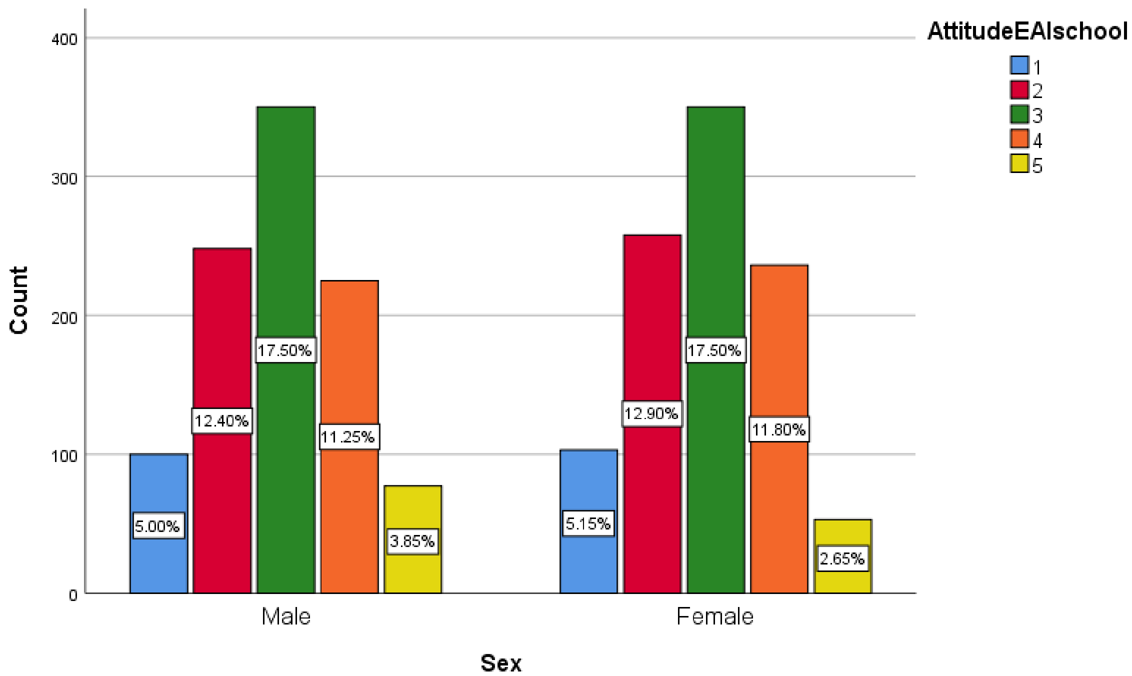


Figure 7.1. Distribution of attitude toward emotional AI in schools by sex. 1 means strongly disagree, 5 means strongly agree.

The descriptive statistics in Figure 7.1. show that there are slightly more people who report feeling *negative* about emotional AI in schools (35.4%) than those who report feeling *positive* (29.6%). Meanwhile, 35% of the respondents stay neutral on the topic. The mean score for the attitude toward emotional AI in schools is 2.9 (sd = 1.068), suggesting Japanese people are on average more negative about the use of emotional AI in schools.

5.1.3. Sex differences

Table 7.1. Sex differences regarding attitude toward and concerns about emotional AI in schools.

Sex	AttitudeEAI school	Bias Concern	DataMisuse Concern	Dystopian Concern

Male	Mean	2.93	3.49	3.41	3.39
	N	1000	1000	1000	1000
	Std. Deviation	1.085	.946	.969	.987
Female	Mean	2.88	3.67	3.58	3.56
	N	1000	1000	1000	1000
	Std. Deviation	1.051	.839	.855	.858
Total	Mean	2.90	3.58	3.49	3.47
	N	2000	2000	2000	2000
	Std. Deviation	1.068	.898	.917	.928

Sex		Knowledge	Safety Utility	Accuracy Concern	TrustGov	TrustPrivate
Male	Mean	2.99	3.09	3.46	2.83	2.88
	N	1000	1000	1000	1000	1000
	Std. Deviation	.969	.951	.951	1.004	.960
Female	Mean	2.95	3.07	3.57	2.83	2.93
	N	1000	1000	1000	1000	1000
	Std. Deviation	.875	.855	.865	.932	.903
Total	Mean	2.97	3.08	3.51	2.83	2.91
	N	2000	2000	2000	2000	2000
	Std. Deviation	.924	.904	.911	.968	.932
RANGE:	1 (Strongly disagree) to 5 (strongly agree)					

Running the Chi-square test, we find statistically significant sex differences among the following variables: the concern that EAI used in school to monitor emotion and attention can be biased against certain disadvantaged groups (BiasConcern, $p < 0.001$); the concern that emotional data of children collected by EAI might be used against them now or in the future (DataMisuseConcern, $p < 0.001$); the concern that EAI in school is dystopian as emotional expressions of children are constantly monitored (DystopianConcern, $p < 0.001$); self-rated knowledge about EAI use in school (Knowledge, $p = 0.017$); the recognition of increased safety due to EAI use in school (SafetyUtility, $p = 0.003$); the concern about accuracy of EAI systems (AccuracyConcern, $p = 0.002$). Meanwhile, there is no meaningful sex differences in the variables of trust in the government or the private sector's ability to regulate the technology (TrustGov and TrustPrivate) as well as the variable of attitude toward EAI use in school (AttitudeEAIschool).

Thus, we find female respondents are, on average, more concerned about the potential biases in EAI systems being used in school, the potential for data misuse, the dystopian feature of constantly monitoring children's emotions, and the potential for inaccurate reading of emotions compared to their male counterparts. Female respondents are also less positive about the increased safety utility in school as the result of using EAI systems and rate themselves as having less understanding about the technology.

5.2. Regression analysis

5.2.1. Socio-demographic factors

Table 7.2. Regression results for socio-demographic factors and attitude toward emotional AI in schools.

c		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.323	.130		25.560	.000
	Age	-.010	.002	-.129	-5.161	.000
	Income	.023	.039	.015	.596	.551
	Education	.001	.015	.001	.055	.956

a. Dependent Variable: AttitudeEAIschool; R square = 0.017

Regarding the socio-demographic determinants of attitude toward EAI in school (Table 7.2.), only age exhibits a statistically significant relationship with the attitude toward EAI in school, which is negative (β_{Age} = - 0.129***).

5.2.2. Utility, Values and Concerns

Table 7.3. Regression results for behavioral determinants of attitude toward emotional AI in schools.

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.564	.111		5.060	.000
	BiasConcern	.077	.029	.065	2.643	.008
	DataMisuseConcern	-.025	.030	-.021	-.835	.404
	DystopianConcern	-.109	.030	-.094	-3.650	.000
	Knowledge	.190	.025	.164	7.643	.000
	SafetyUtility	.421	.026	.356	16.108	.000
	AccuracyConcern	-.051	.028	-.043	-1.830	.067
	TrustGov	.100	.026	.091	3.796	.000
	TrustPrivate	.195	.028	.170	6.926	.000

a. Dependent Variable: AttitudeEAIschool; R square = 0.445

This model explains 44.5% of the variation in the data. As Table 7.3 shows, positive correlates of attitude toward emotional AI in education include the recognition of increased safety as a consequence of using EAI in school ($\beta_{SafetyUtility} = 0.356^{***}$); concerns for biases toward disadvantaged groups ($\beta_{BiasConcern}=0.065^{**}$); self-rated knowledge of the technology ($\beta_{Knowledge}= 0.164^{***}$); having trust in the government's regulation ($\beta_{TrustGov}= 0.091^{***}$); having trust in the private sector to regulate the technology ($\beta_{TrustPrivate}= 0.17^{***}$). With safety utility being the strongest correlate, the surveyed population considers increased safety at school, which includes but is not limited to smart camera surveillance, intelligent tutoring systems, secured computer-based testing, and anxiety monitoring function, as a major advantage the technology will offer for education. Here, the results resonate with the TAM's predictions on perceived utility and perceived ease of use, as well as, the importance of techno-social environment and government effectiveness in enhancing tech-adoption (Vu & Lim, 2021).

Notably, one paradoxical result is the positive correlation between the bias concern and the attitude toward EAI in school. In other words, even though this study finds Japanese participants to be concerned about the biased treatment of disadvantaged groups, their attitude toward emotional AI in school remains positive. This implies that people are willing to accept the technology regardless of the biases latent in the technology. This positive correlation perhaps speaks to the long-standing cultural belief in a homogenous Japanese society (Woo, 2022).

In terms of negative correlates, people who agree that emotional AI in school which constantly monitors children's emotions is too dystopian are more likely to disagree that such use of the technology will be beneficial for society ($\beta_{DystopianConcern}$= - 0.094***). The result agrees with the prediction from the Moral Foundation Theory, that a violation of privacy via the constant monitoring of emotions should increase unease toward the technology. This confirms the finding by Kucirkova et al. (2021) that privacy concern is among the key considerations of Japanese parents and teachers regarding personalized digital learning devices. Kucirkova et al. (2021) also highlight the risks posed to children's safety by the disclosure of personal information and the difficulty of ensuring personal data security in the Japanese educational setting. More importantly, the result stresses the need for a transparency of how an emotion-sensing technology will be used in the classroom. Clearly, in this context, it is useful to consider the data minimization principle expressed in Article 5(1) of the GDPR, which limits the collecting and processing data only toward necessary ends.

Interestingly, concerns for data misuse (i.e., answer to the question "I would be concerned about what happens to the emotional data about the child, and whether it might be used against the child in some way (now or in the future).") and concerns about accuracy of the technology have no statistically significant relationship with the attitude toward EAI in school. This result somewhat contradicts Kucirkova et al. (2021)'s finding that concerns for data misuse and how the technology might influence students being the key concerns of Japanese teachers.

In the context of the literature, the findings from the regression analysis highlight the ambivalent attitude toward the smart technology in the Japanese classroom, which has been documented in the study on perspectives regarding the use of personalized digital devices by Kucirkova et al. (2021). Here, the authors find while the teachers and parents in the study welcome the new technologies' benefits in personalized learning, they also feel that the technology must be closely monitored by responsible adults.

Given the fact that there is still a raging debate on the nature of human emotions, whether it is wired in our biology or it is socially constructed, it is clear that we need to be cautious with the use of emotional AI in school setting. Before considering a tech-solution, structural causes for students' low motivation and performance need to be considered. Within this debate, it is important to remember that a national survey found that overall, over 70% of junior high school teachers in Japan have overworked by 80 hours each month, which meets the technical threshold for determining death by overwork (Lee, 2022; Matsushita & Yamamura, 2022). This statistics highlights the importance of addressing these structural issues before thinking of using emotional AI technology to monitor and modify students' behaviors, concurring with the critique by Williamson (2021). Next, we will turn to the case of toys.

5.3. Children toys

5.3.1. Descriptive statistics

Below is the explanation to survey respondents regarding the use of emotional AI toys. Then, the respondents are asked to give their response on the scale of 1 (strongly disagree) to 5 (strongly agree) to various statements about the utilities and concerns implicated in the use of emotional AI in children toys.

> "This question is about interactive toys for children up to 12 years old. Toymakers are interested in building toys with capabilities for basic conversations, meaning they can increasingly understand and derive meaning from children's speech. These toys would also try to interpret emotion in child speech, through tone of voice, so that the toy can respond appropriately by adapting play activities or trying to cheer them up if they are sad."

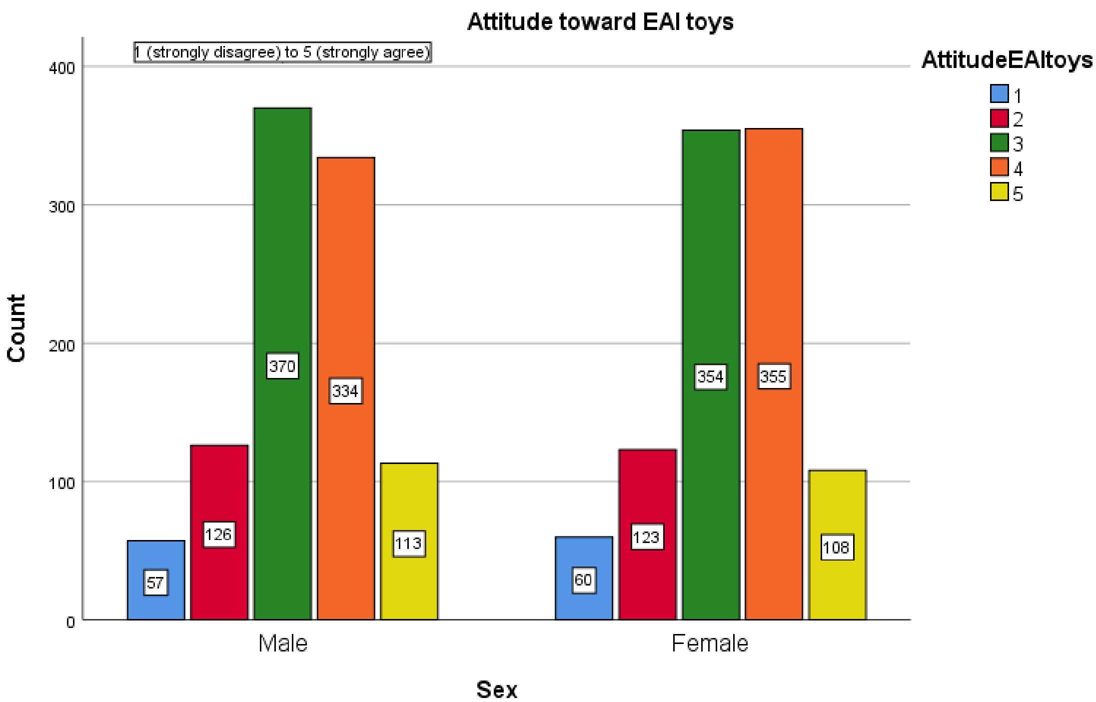


Figure 7.2. Distribution of attitude toward emotional AI in children's toys by sex. 1 means strongly disagree, 5 means strongly agree.

Figure 7.2. shows the distribution of answers regarding the attitude toward emotional AI in toys. Compared to the case of school, emotional AI in toys receives more acceptance from the Japanese population. Overall, 36.2% report feeling neutral about such toys, while 34.5% and 11.1% report

feeling somewhat positive and very positive, respectively, regarding emotional AI in toys. Thus, about 46% report being positive and accepting of the emerging technology, while only 17% report feeling negative about the technology. It is worth noting that only 5.9% report a strong disagreement regarding the technology. The willingness to embrace this technology in children's toys is in line with a previous study on young children's use of personalized technologies. Kucirkova et al. (2021) note the growing interactions of such technologies with children is almost unavoidable but there is a need to ensure children's agency in using the smart personalized toys.

5.3.2. Sex differences

Table 7.4. Sex differences regarding attitude toward and concerns about emotional AI in toys.

Sex		AttitudeEAI Toys	UndueInfluence	DataManage Concern	OK AliveIllusion	TrustPrivate
Male	1000	3.32	3.31	3.40	3.09	2.96
	.959	1000	1000	1000	1000	1000
	3.02	1.019	.919	.939	.924	.959
Female	1000	3.33	3.41	3.54	3.09	3.02
	.851	1000	1000	1000	1000	1000
	2.99	1.021	.851	.896	.903	.851
Total	2000	3.32	3.36	3.47	3.09	2.99
	.907	2000	2000	2000	2000	2000
	TrustPrivate	1.020	.887	.921	.913	.907

Sex		Privacy Concern	Knowledge	Accuracy Concern	Bias Concern	TrustGov
Male	Mean	3.13	3.06	3.36	3.37	2.89
	N	1000	1000	1000	1000	1000
	Std. Deviation	.960	.898	.897	.897	.992
Female	Mean	3.24	3.05	3.45	3.47	2.91
	N	1000	1000	1000	1000	1000
	Std. Deviation	.888	.866	.850	.817	.881
Total	Mean	3.19	3.05	3.40	3.42	2.90

N	2000	2000	2000	2000	2000
Std. Deviation	.926	.882	.875	.859	.938

Regarding sex differences, running the Chi-square test, we find statistically meaningful differences between the sexes in the following variables: the concern about undue influence of the emotional AI toys on children (UndueInfluence, $p = 0.026$); the concern about how emotional data of the children are managed (DataManagementConcern, $p < 0.001$); the concern about loss of privacy or too much monitoring of children emotions (PrivacyConcern, $p = 0.01$); the concern about overall accuracy of the technology (AccuracyConcern, $p = 0.039$); the concern about social biases embedded in emotional AI toys (BiasConcern, $p = 0.011$); the trust in government's regulation (TrustGov, $p = 0.006$); the trust in the private sector's regulation (TrustPrivate, $p = 0.003$).

Thus, we find that women express more privacy concern, more accuracy and bias concern, worry more about the interaction of the EAI toys with the children. These heightened worries of women make sense within the Moral Foundation Theory. It is found that women care more about the moral dimensions of Harm, Fairness, and Purity than men (Atari et al., 2020). This is similar to the finding by Kucirkova et al. (2021), in which the teachers, primarily female, are concerned about the teacher-children relationships if a smart teddy bear can hold AI-powered conversations with the children. The reasons for this worry are attributed to the teachers' desire to cultivate agency and autonomy in children as well as to the deep-seated fear of personal data breach. Overall, the literature on digitalized, smart, Internet-connected toys is more focused on the issues of data privacy and human interactions with the emerging technologies than the more fine-tuned aspects of technological accuracy and inherent social biases (Martín-Ruíz et al., 2018; McReynolds et al., 2017; Yankson et al., 2017).

5.4. Regression analysis

5.4.1. Socio-demographic factors

Table 7.5. Correlations of socio-demographic factors and attitude toward emotional AI toys

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.933	.105		37.477	.000
	Age	-.010	.002	-.133	-5.984	.000
	Income	-.007	.016	-.009	-.413	.680
	Education	-.034	.012	-.062	-2.793	.005

a. Dependent Variable: AttitudeEAItoys; R square = 0.02

Running a regression analysis on socio-demographic factors and attitude toward emotional AI toys, we find that both age and education are negative correlates of the dependent variable (β_{Age} = -0.133***; $\beta_{Education}$ = -0.062 **). Here, as elderly people tend to reject new emerging technologies, it is expected that age would negatively correlate with attitude toward EAI toys. However, it is unexpected that education is a negative correlate since the results in the literature have indicated that people would higher educational qualification tend to view new, emerging technologies such as AI or robots favorably. This result might be due to the subject being children toys. There might be an uncomfortable feeling among more educated parents regarding how the EAI toys might interact or influence the children.

Indeed, it is likely that the more educated the parents, the more they worry about the impact of technologies on children's development. There is a growing psychological scholarship which indicates that compared with previous generations, there exists a higher correlation between Gen Z 's technological dependency and mental health issues such as loneliness, depression, and anxiety (Anderson & Jiang, 2018), higher levels of individualism in learning and teamwork (Dombrosky et al., 2018), as well as higher levels of reliance of digital devices for interpersonal communication (Chicca & Shellenbarger, 2018).

5.4.2. Values and concerns

Table 7.6. Correlations of behavioral factors and attitude toward emotional AI toys

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Significance
1	(Constant)	1.117	.115		9.748	.000
	UndueInfluence	-.020	.028	-.018	-.730	.466
	DataManagementConcern	.131	.027	.118	4.789	.000
	OKAliveIllusion	.415	.025	.372	16.779	.000
	PrivacyConcern	-.219	.025	-.199	-8.749	.000
	Knowledge	.100	.025	.087	3.962	.000
	AccuracyConcern	-.039	.030	-.034	-1.290	.197
	BiasConcern	.080	.029	.067	2.807	.005
	TrustGov	.099	.028	.091	3.560	.000
	TrustPrivate	.167	.030	.149	5.571	.000

a. Dependent Variable: AttitudeEAItoys; R square = 0.38

This model explains 38% of the variation in the data, which is quite low compared to other studies that use the extended TAM model (Lew et al., 2020; Scherer et al., 2019). Positive correlates of attitude toward emotional AI in toys include being OK with a child having an illusion that the toys might be alive and having a personality ($\beta_{OKAliveIllusion}$= 0.372***); concerns regarding the management of emotional data collected by the toys ($\beta_{DataManagementConcern}$ =0.118***); concerns for biases toward disadvantaged groups ($\beta_{BiasConcern}$=0.067**); self-rated knowledge of the technology ($\beta_{Knowledge}$= 0.087***); having trust in the government's regulation ($\beta_{TrustGov}$= 0.091***); having trust in the private sectors to regulate the technology ($\beta_{TrustPrivate}$= 0.149***).

The strongest positive correlate is between being OK with the alive toy illusion, which suggests the surveyed population considers the feeling that the toy has a personality and its being alive is a major advantage, when choosing whether to buy an emotional AI toy. This finding However, the second strongest correlation is a *negative* one, which is between privacy concern and attitude toward emotional AI in toys. Here, people who agree that emotional AI in school that constantly monitors children's emotions is too intrusive are more likely to disagree that they want the emotional AI in toys ($\beta_{PrivacyConcern}$= - 0.199***). Thus, there is tug of war between the utility of the toys presenting an illusion of being alive with the privacy concern.

It is interesting and seemingly paradoxical that concern about data management and concern about embedded biases are positive correlates with attitude toward EAI toys and while accuracy concern is not a statistically significant predictor. This suggests the surveyed population, when it comes to acceptance of the EAI toys, might have little concern about biases embedded in the toys, which might be a product of the homogenous nature of Japanese society. Also, the result regarding data management concern suggests the respondents might accept the toys regardless of their concerns about who might access the emotional data.

Another notable result is the concerns about data misuse and concerns about accuracy of the technology have no statistically significant relationship with the attitude toward EAI in toys. When we take a normative stance, this result is somewhat worrying. As McStay and Rosner (2021) pointed out in their seminal work on emotional AI toys: children, unlike adults, have little control and ability to negotiate the uses of emotional AI technologies toward them. The authors name this problem 'generational unfairness.' Since, adults make decisions for the children about which toys they can play with, the fact that our analysis shows concerns about accuracy and data misuse do not figure into their attitude toward emotional AI toys is worrying. The results support the concern that McStay and Rosner (2021) raised about parents' susceptibility and naivety since most parents lack the technical understanding regarding what data and how they are collected and processed in these toys.

Consequently, the rise of emotional AI toys and their increasing presence in home requires an urgent need to educate parents and responsible parties to understand better the working of new technologies as well as their social and ethical implications to protect vulnerable children from its negative effects.

In a citizen workshop, a male participant (30s) raises his concerns: "Although I would like to buy these smart toys for my children, I would like to know more how the data the toys collected are stored and processed. I certainly do not want the data to be used in ways that lead to more and more power for the companies. We can imagine these companies have data on behavioral patterns and characteristics of our children and come up with more ways to manipulate us and sell their products." Here, the participant essentially raises his concerns about how to maintain the data minimization principle, which is vulnerable to the presence of emotional AI toys in home.

5.5. Chapter summary

In this chapter, we have explored perceptions of implications of emotional AI in the context of educating and developing children: in schools and in toys. In both cases, there are more respondents that express a positive attitude toward such use of the technology than those two express negative feeling. We also find the older people are less receptive of the technology, as age is a negative correlate of attitude toward emotional AI in both cases.

In terms of sex differences, aligning with predictions from the Moral Foundation Theory, we find that women express more privacy concern, more accuracy and bias concern of emotional

AI in schools and in toys. As for the regression analysis to understand how social and ethical perceptions of emotional AI influence acceptance of the technology, we find that self-rated knowledge and perceived utilities of the emotional AI (i.e., improving safety at school or making toys more interactive) are positive correlates of its acceptance. These findings agree with the predictions of the TAM. We also find that in both cases, trust in regulation of the government and trust in the private sector positively correlate with the acceptance of the technology.

Importantly, we find that two seemingly paradoxical results. The first is a positive statistically significant correlation between bias concern and attitude toward EAI in school. The second is concerning emotional AI toys: the concern about data management and the concern about embedded biases are positive correlates with attitude toward EAI toys. We can interpret the results as people accept the technology even when they acknowledge its shortcomings. The results might also reflect a cultural attitude regarding values such as privacy and homogeneity in Japan.

Chapter 6: Private Space: Home Robots and Cars

Emotional AI might be a new term, yet robots that are designed to evoke feelings in humans and have some basic responses to human emotions have been introduced in Japan for decades. It is widely accepted that Japan is the home to many of the first sophisticated companion robots in the world. In 1999, Sony introduced its first robot dog companion AIBO, which has been hugely popular. In 2018, AIBO was reintroduced as being updated with AI software that infused it with a "lovable quality" according to assertions from Sony. Along with AIBO, comes other companion robots such as the humanoid robot, Pepper by Softbanks; the conservational robot, Palro by Fujisoft (White & Galbraith, 2019); the vulnerable, cuddly, even moody robot NICOBO, by Panasonic, etc. (Nikkei Staff Writers, 2021).

Another prominent example of emerging emotional AI technology in Japan is the character Azuma Akari, which is often likened to Amazon's Alexa or Google's Assistant. However, according to leading scholars in the anthropology of Japanese human-robot relationships, Daniel White and Galbraith (2019), Azuma is vastly different due to its deep roots in Japanese anime and manga culture. Being represented as an image of a 'cute girl character" (bishojo-美少女) to deliberately make 'her' "far more deliberately affectionate and alive." Gatebox Inc., Azuma's company, unapologetically promotes Azuma as a niche product for male users who are "otaku," i.e., committed anime and manga fans.

In Japan, home care/companion robots are not only seen as mere commercial products but they are also considered real, potential solutions for a rapidly aging society (Wright, 2019). According to the latest annual report on aging society, there are 36.19 million people aged 65 or older, which takes up to 28.8% of the total population in Japan (Cabinet Office Japan, 2021). More importantly, compared to other countries, Japanese elderly have the lowest number of interactions with their neighbors in terms of consulting or helping when another person is ill (Cabinet Office Japan, 2021).

In this chapter, we seek to explore various concerns and determinants of that influence the acceptance of regarding the presence of emotional AI in private spaces including home robots and cars.

6.1. Home robots

Below is the explanation given to the respondents, prior to them answering a series of Likert-scale questions on their perception of the technology.

> "Robots that assess people's emotion expressions and behaviour have potential uses in the home. This can include performing basic domestic chores and acting as personal assistants and companions (e.g., scheduling the week's events, making reservations, and helping with home security). Home robots can also assess people's emotions to have basic conversations, and engage in interactive activities (such as play, education, and companionship). In the case of silicon robots that respond to touch, they can even have sex with humans."

Then the respondents are asked to give their answer to 10 different Likert-scale questions (range 1 to 5; 1 being "strongly disagree" and 5 being "strongly agree").

6.1.1. Descriptive statistics

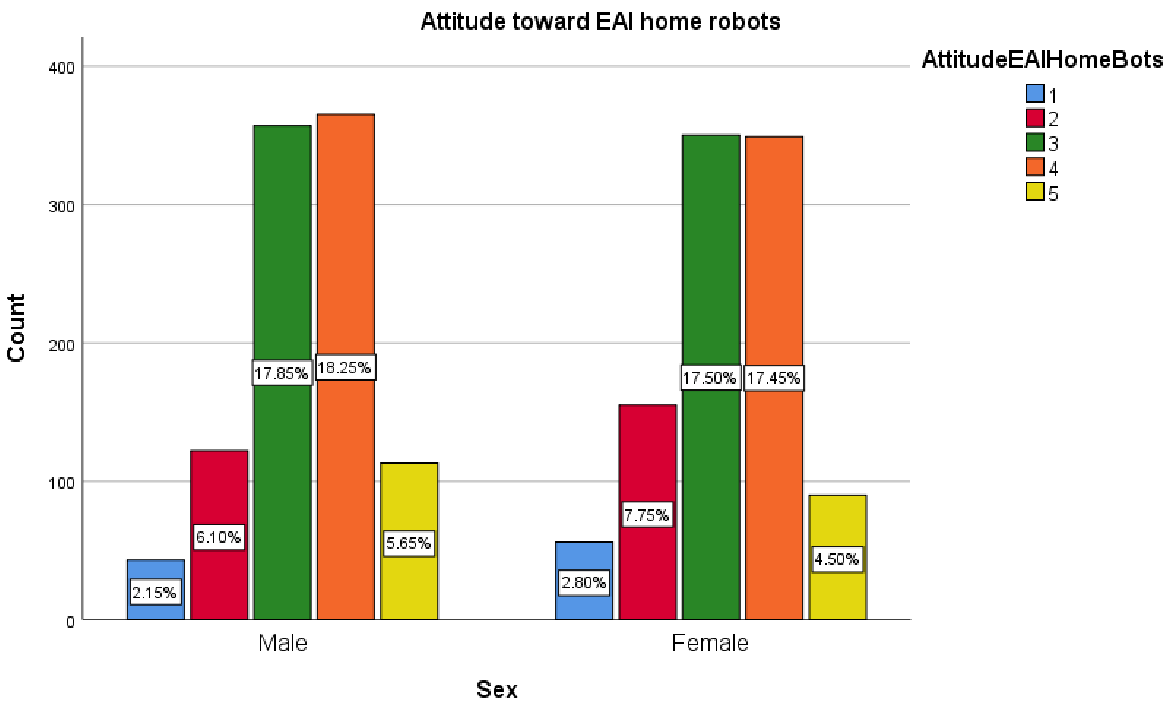


Figure 8.1. On the acceptance of an EAI home robot that assesses emotions and behaviors. 1 strongly disagree to 5 strongly agree.

Looking at the descriptive statistics across all variables in Figure 8.1., the following results stand out. First, there are more people that are willing to accept the home robots that assess their emotions and behaviors that those who are not. Specifically, while 45.9% (24.1% male and 21.95% female) report an acceptance of the home robots, only 18.8% report a negative feeling toward the home emotional AI robots. Only 35.4% of the people report feeling neutral about the robots.

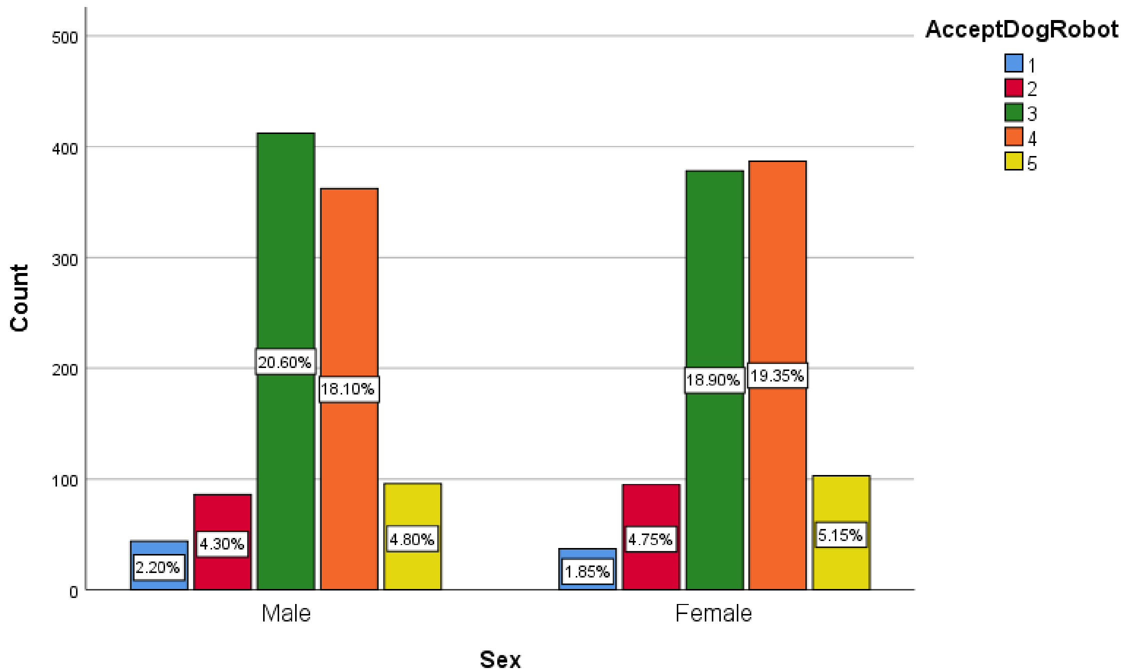


Figure 8.2. On the acceptance of a dog robot gaining the cherished status of being a household member. (1 strongly disagree to 5 strongly agree).

In the case of whether people would accept a dog robot gaining the cherished status of a household member like a real dog, the number is also equally positive in favor of the dog robot (Figure 8.2). Specifically, 47.5% (10% strongly agree, 37.5% tend to agree) report they would be comfortable with the dog robot achieving such status. In contrast, only 13.1% report non-acceptance (strongly disagree or tend to disagree) of the companion dog robot.

These two items are combined to create a new variable that measure the acceptance of EAI home robots (AttitudeHomebotCombined), which will be used as the dependent variable for later regression analyses. A reliability test was conducted and the Cronchach alpha is 0.784, suggesting the two items track each other well.

6.1.2. Sex differences

Table 8.1. Sex differences in attitude toward and concerns about emotional AI in home robots

Sex		AttitudeEAI HomeBots	Decrease Loneliness	AcceptDog Robot	HumansReplace	TrustPrivate
Male	Mean	3.38	3.45	3.38	3.22	2.94
	N	1000	1000	1000	1000	1000
	Std. Deviation	.982	.939	.930	.910	.966
Female	Mean	3.26	3.44	3.42	3.37	2.96
	N	1000	1000	1000	1000	1000
	Std. Deviation	1.010	.937	.929	.913	.874
Total	Mean	3.32	3.45	3.40	3.30	2.95
	N	2000	2000	2000	2000	2000
	Std. Deviation	.998	.938	.929	.914	.921

Sex		SexBotsChange Connections	Privacy Concern	AutonomyLoss	Knowledge	TrustGov
Male	Mean	3.20	3.40	3.31	3.08	2.88
	N	1000	1000	1000	1000	1000
	Std. Deviation	.961	.913	.887	.902	1.004
Female	Mean	3.34	3.60	3.45	3.02	2.81
	N	1000	1000	1000	1000	1000
	Std. Deviation	.907	.900	.848	.887	.921
Total	Mean	3.27	3.50	3.38	3.05	2.84
	N	2000	2000	2000	2000	2000
	Std. Deviation	.937	.911	.870	.895	.964

Applying the Chi-square test, we find statistically significant sex differences in the following variables: the utility of companion robot to decrease loneliness (DecreaseLoneliness, $p = 0.046$); the concerns that companion robots would replace human visitors, friends, or families (HumansReplace, $p = 0.008$); the worry that sex robots would change expectations of human sexual connections (SexBotsChangeConnections; $p = 0.005$); the privacy concern over management of emotional and behavioral data (PrivacyConcern, $p < 0.001$); the concern that emotionally intelligent robots can influence their owners' thinking and feeling (AutonomyLoss , p = 0.002); trust in the government's regulation (TrustGov, p = 0.013), trust in the private sector (TrustPrivate, p = 0.021).

To interpret the above results, we find male respondents are more positive that home robots would reduce loneliness, which means male Japanese respondents recognize the utility of reducing loneliness more than their female counterparts.

Meanwhile, female respondents are more concerned that robots would replace human visitors (friends or families) (i.e., the dimension of Purity), that robots would change expectations of human sexual relations (i.e., the dimension of Purity), that data collected by the home robots would not be securely managed and stored (i.e., the dimension of Harm), that presence of emotionally intelligent robots would have undue influence in one's thinking and feeling (i.e., the dimension of Harm). These results are well within the prediction of the Moral Foundation Theory.

Regarding trust in the regulation, while men express more trust in the government's ability to regulate the technology, women express more trust in the private sector's ability to regulate the technology.

6.1.3. Socio-demographic factors

Table 8.2. Regression results for socio-demographic determinants of emotional AI in home robots.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.405	.106		32.105	.000

	Age	-.003	.002	-.047	-1.862	.063
	Income	.043	.032	.035	1.361	.174
	Education	.008	.013	.017	.634	.526

a. Dependent Variable: AttitudeHomeBotCombined

We find no statistically significant relationship between socio-demographic factors of age, income, education, and acceptance of home robots (Table 8.2).

6.1.4. Values and concerns

Table 8.3. Regression results for behavioral determinants of emotional AI in home robots.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.393	.070		5.589	.000
	DecreasedLoneliness	.657	.014	.705	48.414	.000
	HumansReplace	-.030	.016	-.031	-1.891	.059
	SexBotsChangeConnections	.007	.016	.007	.427	.669
	PrivacyConcern	.013	.017	.014	.781	.435
	AutonomyLoss	.007	.019	.007	.391	.696
	Knowledge	.093	.015	.095	6.072	.000
	TrustGov	.057	.016	.063	3.480	.001
	TrustPrivate	.090	.017	.095	5.173	.000

a. Dependent Variable: AttitudeHomebotCombined; R square = 0.678

Running a linear regression analysis on behavioral determinants of attitude toward emotional AI home robots, we find the model accounts for nearly 68% of the variation in the data, which is the highest among all models. Here, per Table 8.3., the statistically significant positive correlates include the utility of decreased loneliness as the result of adopting these robots ($\beta_{DecreaseLoneliness}$ = 0.705***); self-rated knowledge of the technology ($\beta_{Knowledge}$= 0.095***);

having trust in the government's regulation ($\beta_{TrustGov}= 0.063^{***}$); having trust in the private sectors to regulate the technology ($\beta_{TrustPrivate}= 0.095^{***}$).

Surveying the numbers in-depth shows a very strong correlation between the utility of decreased loneliness and acceptance of home emotionally intelligent robots, which can be interpreted that the surveyed Japanese population considers reducing the feeling of loneliness the most important determinants of acceptance toward home robots While other concerns for values such as privacy, autonomy, and human replacement are of little concerns. Such a strong correlation makes sense in the context of Japan still during the COVID-19 pandemic, when the survey was conducted, and many studies have reported a sharp increase in social isolations in Japan during the pandemic (Yamada et al., 2021).

These significant findings on decreased loneliness and self-rated knowledge also make sense under the TAM framework since the perceived utility of decreased loneliness and perceived ease of use (i.e., self-rated knowledge of the technology) are two strongest correlates of the acceptance of home robots.

On the contrary, concerns for privacy violation, autonomy loss, human replacement, changing expectations of human connection are found to be non-significant. This is unexpected by the theory of Moral Foundations. Applying the Moral Foundation Theory, things that violate moral norms such as harms (such as privacy violation or autonomy loss) and purity (i.e., changing expectations of human connection or replacing human visitors of friends) should undermine the acceptance of the technology. Yet, here, in the case of the Japanese population, these concerns are not statistically related to acceptance of home robots. This suggests in the case of home robots, the Japanese population care first and foremost about its utility to reduce loneliness, and accepting the concerns about privacy, autonomy loss, and human replacement as part of the transaction.

The result on privacy concern agrees with comparative empirical findings in the literature. For example, a study that compares the attitude toward home care robots among Japanese, Irish, and Finish found that Japanese has the highest acceptance rate of agreeing to home robots taking pictures and record videos that can identify the user with permission (55% compared to around 45% in the case of Ireland and Finland) (Suwa et al., 2020).

6.2. Cars

Interior sensing in car is an area where emotional AI applications are hugely potential. In recent years, Japanese companies have started to launch emotional AI products for cars. The rationale for these companies is to reduce chances of road accidents due to an aging population. For example, Honda and Softbank have co-created the 'Emotion Engine', which detects if a driver is drowsy, distracted, or stressed as a response to the spike in elderly drivers' accidents (Dery, 2018). In the course of this research, in both interviews with working professionals and the citizen workshops, the use of emotional AI in cars seems to receive the least push-back.

In the citizen workshops, all Japanese participants are welcome emotional AI in cars. They see little or no privacy violation because it seems that the master value they see is to make sure travelling is safe. Another participant concurs: "I see the future belongs to autonomous vehicles; thus, it is only natural that emotion-sensing in cars will be utilized. Personally, I don't see it as violating my privacy in cars" (Interview data, 2021). A few participants in the workshops stated that if the smart cars wouldn't interfere with their driving and making commands, but keep a cooperative relationship, they don't too many problems with interior sensing with cars. Next, we will turn to the quantitative results. In an interview with an officer of Japan Post (Male, 30s): "It is inevitable that cars with emotion-sensing technologies will be common given the dramatic growth of products transportation during and after the pandemic" (Interview data, 2021).

Below is the explanation given to the respondents prior to them answering 10 different Likert-scale questions (range 1 to 5; 1 being "strongly disagree" and 5 being "strongly agree").

> "Car manufacturers are interested in understanding how drivers and passengers feel in cars. Unobtrusive in-car cameras would monitor driver tiredness, distraction, emotion expressions (such as stress, anger, or frustration) and behaviour of passengers. A driver might have the option to allow insurance companies to use this data about behaviour and emotion expressions to lower car insurance costs."

6.2.1. Descriptive statistics

Regarding how people feel about emotional AI in cars, while 47.6% of the respondents report a neutral stand regarding this application of emotional AI, there are nearly three (3) times more people who report feeling positive than those who report feeling negative about the technology: 38.3% versus 13.6%. The average score of attitudes toward EAI in car is 3.3 with a standard deviation of 0.886, quite similar to the case of home robots. The descriptive statistics suggest that emotional AI car is view quite positively by the Japanese population.

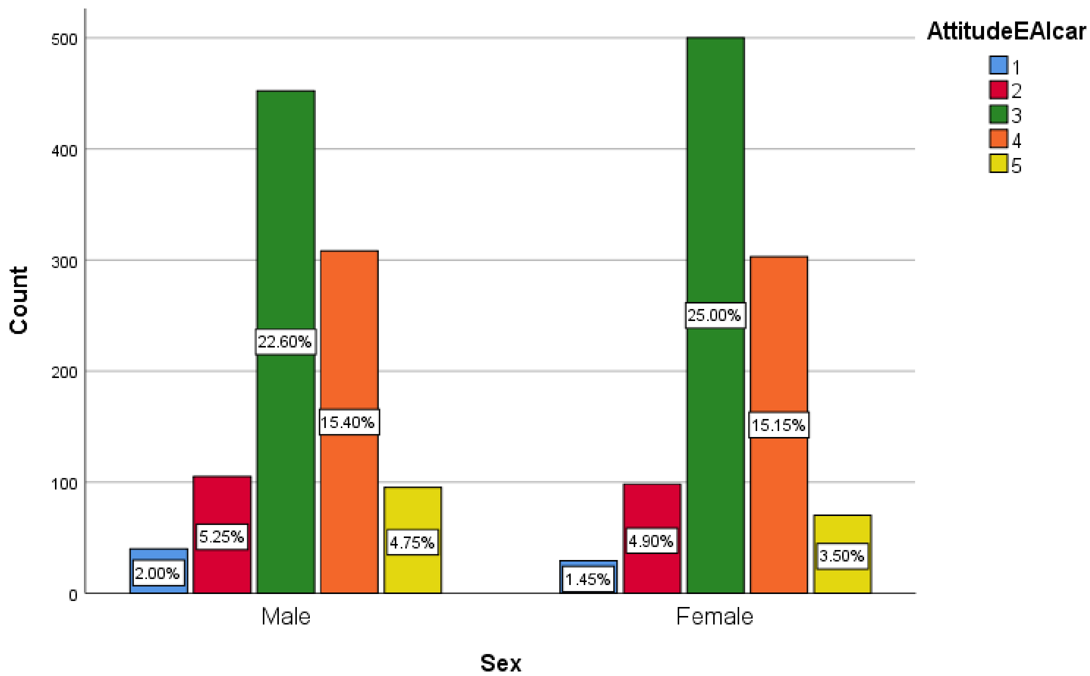


Figure 8.3. Distribution of answers to the question on attitude toward emotional AI in cars: Range: 1 strongly disagree to 5 strongly agree.

6.2.2. Sex differences

Using the Chi-square test, statistically significant differences between the sexes have been found in the following variables: InsuranceAccess ($p<0.001$); Dystopian Concern ($p<0.001$); Privacy Concern ($p<0.001$); Knowledge ($p=0.02$); Accuracy Concern ($p<0.001$). In other words, women tend to display more worries regarding insurance companies accessing their emotional data in cars (i.e., the dimension of Harm). They also tend to feel emotional AI cars pose a dystopian and privacy concern about its constant monitoring of emotion (i.e., the dimension of Harm/Care). Moreover, they also tend have higher concern about the accuracy level of the technology (i.e., the dimension of Harm). These results are aligned with the Moral Foundation Theory's literature on sex differences (Graham et al., 2011; Atari et al., 2021). We also find women rate themselves as having less knowledge about the technology, which aligns with the literature on technological self-efficacy, where women are found to believe they have less self-efficacy related to technologies (Huffman et al., 2013).

Table 8.4. Sex differences in key concerns about emotional AI in cars.

Sex		Safety Utility	Comfort	Insurance Access	Dystopian Concern	Privacy Concern
Male	Mean	3.71	3.38	3.32	3.15	3.15
	N	1000	1000	1000	1000	1000
	Std. Deviation	.953	.983	1.022	1.019	1.022
Female	Mean	3.74	3.38	3.24	3.28	3.37
	N	1000	1000	1000	1000	1000
	Std. Deviation	.870	.933	.927	.902	.928
Total	Mean	3.72	3.38	3.28	3.22	3.26
	N	2000	2000	2000	2000	2000
	Std. Deviation	.912	.958	.976	.964	.983

Sex		Knowledge	Accuracy	AttitudeEAIcar	TrustGov	TrustPrivate
Male	Mean	3.23	3.33	3.31	2.92	3.02
	N	1000	1000	1000	1000	1000
	Std. Deviation	.941	.945	.925	.999	.970
Female	Mean	3.18	3.47	3.29	2.94	3.05
	N	1000	1000	1000	1000	1000
	Std. Deviation	.865	.831	.846	.919	.876
Total	Mean	3.21	3.40	3.30	2.93	3.03
	N	2000	2000	2000	2000	2000
	Std. Deviation	.904	.892	.886	.960	.924

6.3. Regression analysis

6.3.1. Socio-demographic factors

Table 8.5. Socio-demographic determinants of attitude toward emotional AI in cars.

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	3.306	.108		30.524	.000
	Age	-.004	.002	-.069	-2.729	.006
	Income	.043	.032	.034	1.323	.186
	Education	.032	.013	.064	2.443	.015

a. Dependent Variable: AttitudeEAIcar; R square = 0.012

Regarding how socio-demographic factors influence the perception of emotional AI in cars, we find two results that are diverging compared to the case of home robots. First, age exhibits a negative correlation with attitude (β_{age}= -0.069**). And second, education has a positive correlation ($\beta_{Education}$= 0.064*). That means older people are less likely to be in favor of emotional AI in cars, while people with higher education are more likely to think emotional AI in cars are beneficial for society.

6.3.2. Values and Concerns

Table 8.6. Behavioral determinants of attitude toward emotional AI in cars.

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.483	.088		5.502	.000
	Knowledge	.192	.017	.196	10.971	.000
	SafetyUtility	.205	.018	.211	11.131	.000
	Comfort	.125	.018	.135	6.735	.000
	InsuranceAccess	.149	.018	.164	8.442	.000
	TrustPrivate	.130	.021	.136	6.344	.000
	TrustGov	.118	.019	.128	6.374	.000

DystopianConcern	-.003	.019	-.003	-.135	.892
PrivacyConcern	-.138	.020	-.154	-6.996	.000
Accuracy	.072	.018	.072	3.979	.000

a. Dependent Variable: AttitudeEAIcar; R square = 0.583

This model explains 58% of the variation in the data. Positive correlates of attitude toward emotional AI in cars include: increased safety (β_{Safety} = 0.211***); increased comfort ($\beta_{Comfort}$ =0.135***); self-rated knowledge of the technology ($\beta_{Knowledge}$= 0.196***); concern about emotional AI's accuracy ($\beta_{AccuracyConcern}$= 0.072***); being ok with insurance companies having access to emotional data in cars ($\beta_{InsuranceAccess}$= 0.164***); having trust in the government's regulation ($\beta_{TrustGov}$= 0.128***); having trust in the private sectors to regulate the technology ($\beta_{TrustPrivate}$= 0.136***). With safety utility being the strongest correlate, it suggests that the surveyed population cares the most about safety utility when it comes to their judgment of the overall benefit of emotional AI cars. Such a finding resonates with the workshop, as most workshop participants state that if they would feel comfortable with emotional AI in cars as it is a good way to safeguard their safety.

In terms of negative correlates, people who agree that emotional AI in cars would create too much scrutiny of their emotional lives are more likely to disagree that such use of the technology will be beneficial for society ($\beta_{PrivacyConcern}$= - 0.154***).

Regarding the dystopian concern, i.e., the fear the constantly monitoring emotions in the interior of a car represents a shift toward a dystopian society, is found to bear no statistically significant relationship with the dependent variable AttitudeEAIcars. This result suggests that people do not think of a new feature of interior sensing of mental states of car drivers and passengers, however intrusive, represent that society is becoming more dystopian.

Interestingly, concern for accuracy of emotional AI is positively correlated with attitude toward the technology. This suggests that people are willing to accept the technology despite its flaws. This result somewhat contradict the worry voiced by many technologists that due to the lack of perfect performance, people will not accept AI technologies even though in utilitarian terms, an imperfect, 95% correct AI would deliver huge benefits: reducing an absolute number of deaths by road accidents in the case of self-driving cars or

As mentioned in the beginning of the section, the citizen workshop participants are aligned with the results we find here. During the citizen workshop, all of the Japanese participants think it is inevitable that future smart cars as well as self-driving cars will need to have this feature. They welcome this feature, citing two reasons: it will help them drive safer as well as have a better experience in cars. Only among some foreigners who work in Japan raise the concern that they do not like it when the cars started to make its own decisions, for example, demanding the drivers to stop because the car senses the driver is tired. They think it is still important that the drivers can make their own decisions.

6.4. Chapter summary

Chapter 6 has provided a close examination of the perceptions toward two uses of emotional AI in our private space: home robots and cars. In both cases, we find that there are around three times more people who accept the technology than people who are not, suggesting a high willingness for adoption of these technologies in society.

In the case of home robots, we find a strong correlation between the utility of reduced loneliness and attitude toward these companion robots, which perhaps reflects the increase of social isolations that has happened since the beginning of the COVID-19 pandemic.

In cars, attitude toward emotional AI is strongly correlated with the utilities of increased safety and increased comfort. While in both cases, privacy concern and data management concern are not significant predictors of attitude toward the technology, which highlights the acceptance of the privacy risk associated with installment of emotional AI in these private spaces. Concern about violation of the purity norms (such as sex robots can change human expectations of sexual relations) or concern about the undue influence of home robots on how a person thinks and feels are found to be non-significant. Meanwhile, the acceptance of emotional AI cars is not influenced by whether a car constantly sensing the drivers' emotions is too dystopian. These quantitative results resonate with the qualitative insights from the citizen workshops, where most participants are not opposed to emotional AI in these private environments.

Chapter 7: Healthcare

Artificial intelligence is going to transform the healthcare sector worldwide. In this transformation, emotional AI is expected to play an important role in self-care and preventive medicine for mental health (McStay, 2018). Recognizing the potential for mental health support of emotional AI, companies and governments around the world are exploring a wide array of algorithms that can detect early signs of mental illnesses or help patients and the common consumer become more mindful and less stressful. Moreover, the most advanced generation of conversational bots, endowed with natural language processing and affective computing ability, could respond to a patient's emotions in their voice tones or facial expressions, etc. For example, the UK's National Health Service is investing in AI conversational agents such as *Wysa* to provide online medical counselling (Adikari et al., 2022). In Singapore, the social service agency Lion Befrienders has developed facial recognition software to provide early detection of depression, anxiety, cognitive decline in senior citizens (Menon, 2021). Meanwhile, EU nations are trialing affective robots such as JustCat and Hobbit as surrogate caretakers to assist elderly facing cognitive or physical disabilities (Johnson et al. 2020).

In this chapter, we are going to study to perceptions of emotional AI for healthcare purposes in the context of Japan. With the rapid pace of population aging, Japan's medical workforce and healthcare system are facing a huge challenge in maintaining healthcare quality for their increasing elderly patients while suffering from a shortage in medical staff (Marć et al., 2019). Consequently, in recent years, Japan has put significant efforts in improving its technological capacity, investing heavily to develop artificial intelligence (AI), and utilizing its application to not only maintain but also revolutionize its healthcare system.

In the Japanese healthcare sector, it is estimated that the market for AI-related medical sector will reach 15 billion yen in 2025 (JETRO Australia, 2020). In terms of image diagnostics, Olympus Corporation, Fujifilm Corporation, Astellas Pharma, Toray Engineering, Envoy Ai, Cyberdyne, and Lpixel are the biggest contributors, interestingly all being domestic companies. Cooperation in AI investment is also relatively common. In 2015, Softbank allied with IBM, a multinational technology and consulting corporation, to introduce IBM Watson, a tool for facilitating medical research, clinical research, and healthcare management health to the public. The integration of AI into the healthcare system offers many prominent solutions and even innovations to the field (Jiang et al., 2017; Yu et al., 2018). For instance, Ubie, one of the top AI start-up companies in Japan (Onikle Inc., 2021) offers a variety of AI products that enable physicians to effectively communicate with patients: online medical records, AI-based preliminary examination, and assisting chatbots and questionnaires (Ubie, 2022). Looking at other AI products in healthcare startups in Japan, services provided vary from cancer diagnosis, and image-based cell sorting to clinical decision management, and patient monitoring (Tracxn, 2022).

Next, we will explore various determinants and concerns that influence the acceptance of emotional AI in the healthcare setting. The first part of the chapter focuses on the results from a Japanese national, representative survey (N = 2000) conducted in March 2022. The second part of the chapter presents the data and results from a smaller sample, which collects the viewpoints of 245 clinic visitors in Beppu City, Oita Prefecture Japan (N=245) during August to October, 2021.

7.1. Study 1: The national survey of attitude toward emotional AI used in mental healthcare and screening

Below is the explanation given to the survey respondents before they were asked to give answers on how much they agree or disagree with a series of Likert-scale questions (Range 1 means strongly disagree to 5 means strongly agree) concerning the use of emotional AI in mental healthcare.

> "Dementia is a syndrome associated with an ongoing decline of brain functioning. Companies are using tablet computers in professional elderly care-home settings to provide therapy to elderly dementia sufferers to improve their stress levels and well-being. In these cases, the carer shows the dementia patient images designed to trigger memories. The patient's facial expressions (e.g., a big smile, or a very small change of expression) are captured by the tablet's camera. This aims to infer the patient's emotions when they look at the images. This data about the patient's emotional experience is said to help the carer and care home quantify if the therapy is working."

7.1.1. Descriptive statistics

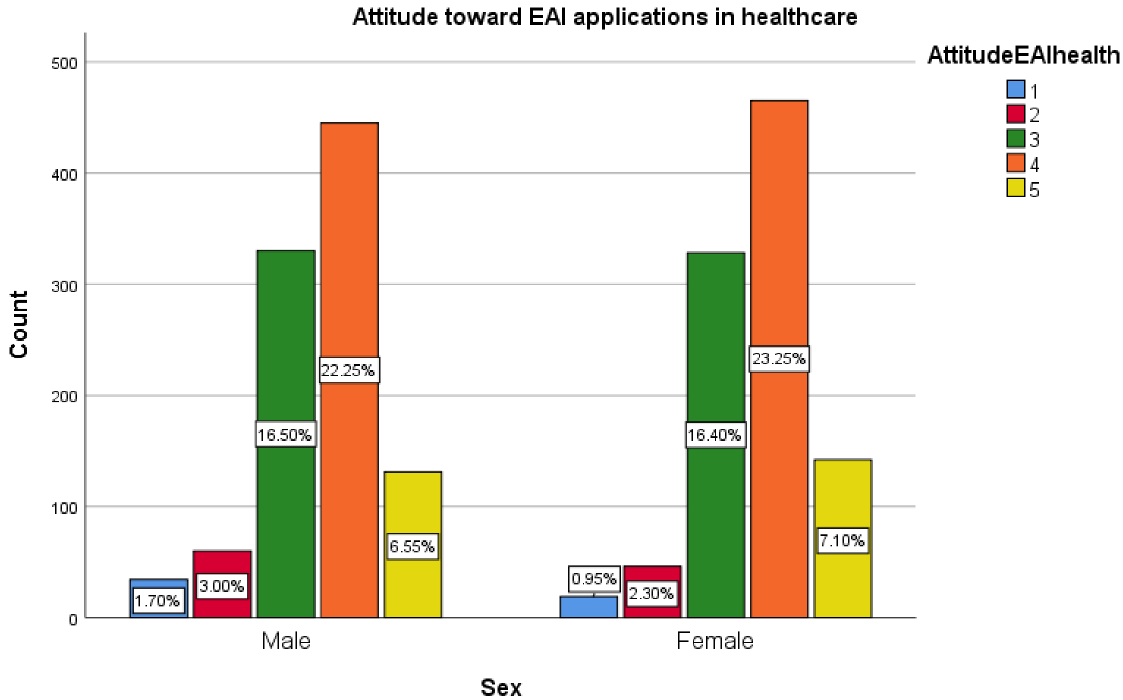


Figure 9.1. Distribution of attitude toward EAI applications in healthcare: Range 1: Strongly disagree to 5: Strongly agree.

Regarding the attitude toward EAI applications in healthcare, 59.2% of the respondents report agreeing with the statement that they are comfortable with using an EAI application and would like to see it more widely used in healthcare for elderly people. It is worth noting that while 13.7% people report "strongly agree," only 2.7% report strongly disagree. Only 32.9% report feeling neutral.

7.1.2. Sex differences

Table 9.1. Sex differences in the attitude toward and concern about emotional AI in mental healthcare.

Sex		Attitude EAIhealth	Undue Influence	Replace Humans	DataManage Concern
Male	Mean	3.58	3.08	3.23	3.26
	N	1000	1000	1000	1000

	Std. Deviation	.911	.937	.878	.916
Female	Mean	3.66	3.14	3.40	3.39
	N	1000	1000	1000	1000
	Std. Deviation	.845	.842	.826	.843
Total	Mean	3.62	3.11	3.31	3.32
	N	2000	2000	2000	2000
	Std. Deviation	.879	.891	.857	.882

Sex		Knowledge	AutonomyLoss	SafetyUtility	TrustGov	TrustPrivate
Male	Mean	3.18	3.12	3.39	3.02	3.05
	N	1000	1000	1000	1000	1000
	Std. Deviation	.907	.875	.856	.935	.945
Female	Mean	3.19	3.23	3.44	3.05	3.14
	N	1000	1000	1000	1000	1000
	Std. Deviation	.902	.835	.777	.849	.840
Total	Mean	3.19	3.17	3.41	3.03	3.09
	N	2000	2000	2000	2000	2000
	Std. Deviation	.904	.857	.818	.893	.895

Applying the Chi-square test, statistically significant differences between the sexes have been found in the following variables: the worry that EAI applications are allowed to make judgments on sensitive matters such as emotions and psychology of at-risk populations (UndueInfluence, $p=0.002$); the worry that EAI applications are considered better than human care workers in emotion recognition skills (ReplaceHumans, $p < 0.001$); the concern about how data on patients are stored and managed (DataManageConcern, $p = 0.003$); the concern that using EAI can undermine the autonomy of healthcare givers and patients (AutonomyLoss, $p = 0.023$); the trust in the government's ability to regulate the technology (TrustGov, $p = 0.048$); the trust in the private sector's ability to regulate the technology (TrustPrivate, $p = 0.005$).

To be specific, first, there are no sex differences in self-rated knowledge about the technology, or attitude toward the technology, or the attitude of whether the technology will improve the safety of the treatment and medical processes.

Second, we find that women are slightly more concerned about letting an EAI app make judgments on sensitive matters of emotion and psychology of at-risk people (i.e., the dimension of Fairness), about the issue of data management (i.e., the dimension of Harm/Care), about the perception that EAI apps can be better than human care workers in emotion recognition skills (i.e., the dimension of Purity), and about the loss of the ability to make free choices of the healthcare professionals and patients (i.e., the dimension of Purity and Harm). These findings are aligned with the literature on sex differences regarding concerns about different moral dimensions in the Moral Foundation Theory. Here, past studies have shown that women care more about the dimensions of Harm/Care, Fairness, and Purity than men (Atari et al., 2020; Graham et al., 2011).

However, women are found to have more trust in the regulatory ability of the government and the private sector. Although the original TAM and the Moral Foundation Theory do not predict these findings, once we put in the context of Japan, as still being a very patriarchal society (Woo, 2022), it somewhat makes sense that women express more trust in the authority.

Regression analysis

7.1.3. Socio-demographic factors

Table 9.2. Regression results for socio-demographic determinants of emotional AI in mental healthcare

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	3.583	.107		33.607	.000
	Age	.000	.002	-.005	-.207	.836
	Income	.021	.032	.018	.676	.499
	Education	.008	.013	.017	.640	.522

a. Dependent Variable: AttitudeEAIhealth; Note: * means p ≤ 0.05; ** means p≤ 0.01 means*** p≤ 0.001; **** means p ≤ 0.0001

There are no statistically significant results on the relationship between socio-demographic factors and attitude toward emotional AI in healthcare.

7.1.4. Utilities, values, and concerns

In this section, we will examine how perceived utility, self-rated knowledge, and concerns for different values implicated in the use of emotional AI for mental healthcare determine the acceptance of the technology.

Table 9.3. Regression results for behavioral determinants of emotional AI in mental healthcare

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.006	.096		10.518	.000
	UndueInfluence	-.112	.022	-.113	-5.037	.000
	ReplaceHumans	.052	.025	.051	2.135	.033
	DataManageConcern	.074	.023	.074	3.135	.002
	Knowledge	.115	.020	.118	5.637	.000
	AutonomyLoss	-.007	.022	-.007	-.308	.758
	SafetyUtility	.475	.023	.442	20.344	.000

TrustGov	.020	.024	.020	.829	.407
TrustPrivate	.168	.024	.171	7.050	.000

a. Dependent Variable: AttitudeEAIhealth; R square = 0.421; Note: * means p ≤ 0.05; ** means p≤ 0.01 means*** p≤ 0.001; **** means p ≤ 0.0001

This model explains 42.1% of the variation in the data. Positive correlates of attitude toward emotional AI in healthcare include increased safety ($\beta_{SafetyUtility}$ = 0.442***); self-rated knowledge of the technology ($\beta_{Knowledge}$= 0.118***); having trust in the private sectors to regulate the technology ($\beta_{TrustPrivate}$= 0.171***). It is interesting that safety utility is the strongest correlate of acceptance toward emotional AI in healthcare. In other words, the more one thinks that emotional AI will improve the safety of medical processes, the more likely he or she will accept the emerging technology. Indeed, these results agree with the Technological Acceptance Model of Davis (1989): perceived utility of the technology (i.e., improved safety) and perceived familiarity (i.e., self-rated knowledge regarding emotional AI) are too positive correlates of medical emotional AI acceptance.

The negative correlates include the worry that emotional AI applications are allowed to make sensitive judgments on the emotion and psychology of at-risk people ($\beta_{UndueInfluence}$= -0.113***). In other words, people who hold the concern that sensitive judgments on the emotion and psychology of at-risk population should not be left to AI is likely to express a broad acceptance of the technology. Mapping such concerns on the Moral Foundation Theory, this concern directly relates to the dimensions of Harm and Fairness, in that, we have a vulnerable population who is subjected to judgments made by an intelligent machine.

Interestingly, concern about data management process ($\beta_{DataManageConcern}$= 0.074**) and concern about EAI applications are considered better than humans at emotion recognition ($\beta_{ReplaceHumans}$= 0.051*) are two positive correlates. The Moral Foundation Theory cannot explain such mixed results. Looking beyond the TAM and the Moral Foundation Theory, there are three possible explanations.

First, it could be due to the novelty of the technology that these concerns are overlooked by the respondents. Second, regarding the concern that that EAI apps can be considered better than humans at emotion recognition, it is likely that those who have a positive attitude toward the technology is considered this to be a good thing. And third, regarding the concern about data management access, it might be a result of the attitude that health data, a context that is considered more communal, are not considered very sensitive among the Asian populations. For example, (Psychoula et al., 2018) found that health data are not considered as sensitive among the Asian populations. Such a positive correlation highlights that those who accept the emerging EAI technology in the healthcare context are willing to accept the muddy issues around emotional data access and management.

Fear of healthcare workers and patients' autonomy loss and trust in the government's regulation have no statistically significant relationships with the dependent variable. Further policy implications of such results will be explored in Chapter 9 on algorithmic governance. Next, we will explore the case of emotional AI being embodied in home robots for companion purpose as well as in emotional AI in cars.

7.2. Study 2: A survey of clinic visitors in Beppu City, Oita Prefecture, Japan

This study is conducted through a municipal survey with the assistance of Beppu City Hall's Health and Elderly Department to gauge the citizen perceptions and concerns for EAI health care and its domestic usage. Thus, the majority of the citizens participating in the survey is from Beppu city, with the minority coming from nearby cities, namely Oita, Hijimachi, etc. All mentioned cities are within Oita Prefecture of Kyushu Island, Japan. In the distribution process, we first obtained participation consent from 50 clinics and hospitals around Beppu city. A thousand paper surveys were prepared afterward and sent by post to 50 clinics around Beppu City. We prepared the form with a separate set of envelopes so respondents can return the surveys' answers without any fees. In total, the survey received 245 responses.

There are 17 questions in the survey, focusing on participants' awareness, familiarity, concerns, and preferences towards EAI-based tools in healthcare, both for private and public usages. We also collected participants' demographical data and other societal characteristics to form variables and build multiple regression models. This methodological approaches enable us to examine relationship of our desirable dependent variable, namely attitudes, to other independent variables (Stanton, 2001). The following table displays our variables and data treatment:

Table 9.4. Explanation of the data treatment procedure

Variable	Variable type	Description	Remarks/Survey Questions
Outcome variable			
Attitude, Private	Discrete	Attitude toward application of EAI for private use (1 for strongly disagree, 5 for strongly agree)	Attitude for EAI private use is calculated by averaging the answers to two Likert-scale questions: 1. Do you agree EAI in the car would be beneficial such as by detecting stress or distraction? 2. Do you agree that having an EAI device at home to keep track of health and feelings is beneficial?
Attitude, Public	Discrete	Attitude toward application of EAI for public use (1 for strongly disagree, 5 for strongly agree)	Attitude for EAI public use is calculated by averaging the answers to two Likert-scale questions: 1. Do you agree that EAI-based tools will improve diagnostics and care planning for patients (for example, improve accuracy, save time, etc.)?

			2. Do you agree that an EAI tool or a robot that can read or mimic human feelings is useful for doctors and nurses and the medical process in general?
Predictive variable			
AI Familiarity Healthcare	Discrete	Taking the average of four questions on the right side (1 for very unfamiliar, 5 for very familiar)	Familiarity with AI in healthcare is calculated by averaging the answers to three Likert-scale questions 1. Generally, how familiar are you with an AI-based device? (AI-enabled translations, or chatbots like Amazon Alexa) 2. How familiar are you with AI tools for medical purpose? (Such as diagnostic or health monitoring apps or wearable devices) 3. How familiar are you with a care or companion robot?
Discriminatory concern	Discrete	1 for not very worried 5 for very worried	Respondents report their concern level over discrimination with AI usage
Privacy concern	Discrete	1 for not worried 5 for very worried	Respondents report their concern level over privacy with AI usage
Lose control to AI	Discrete	1 for not very worried 5 for very worried	Respondents report their concern level over AI control in diagnosis and treatment
Sex	Binary	Male vs Female	Respondents report their biological sex
Age	Ordinal	20s or below, 30s, 40s, 50s, 60s, 70s, 80s or above	Respondents report their age range
Income level	Ordinal	low, middle, and high	Respondents self-report their level of income
Educational level	Ordinal	Highschool or lower, University undergraduate, Graduate	Respondents report their education level
Employment status	Ordinal	Unemployed, Retired, Parttime, Fulltime	Respondents report their current career level

Daily time online	Ordinal	Under 1 hour, 2-3 hours, 4-5 hours, more than 5 hours	Respondents report their average usage of time on the internet
Living arrangement	Ordinal	Living alone, Living with family, Retirement home	Respondents report their current living arrangement
Community activity	Discrete	1 for very infrequent 5 for very frequent	Respondents report their activeness level in local community activities
SNS group	Binary	Yes vs No	Respondents check if they are joining a social media group

From the constructed variables, we built four multiple regression models as follows:

$$Model\ 1: Attitude\ Private = \beta 1(AI\ Familiarity\ HC) + \beta 2(Privacy) + \beta 3(Discrimination) + \beta 4\ (Lose\ AI\ control)$$

$$Model\ 1: Attitude\ Public = \beta 1(AI\ Familiarity\ HC) + \beta 2(Privacy\ ) + \beta 3(Discrimination + \beta 4\ (Lose\ AI\ control)$$

$$Model\ 2: Attitude\ Private = \beta 1(Employment) + \beta 2(Income) + \beta 3(Sex) + \beta 4(Education) + \beta 5(Age)$$

$$Model\ 2: Attitude\ Public = \beta 1(Employment) + \beta 2(Income) + \beta 3(Sex) + \beta 4(Education) + \beta 5(Age)$$

$$Model\ 3: Attitude\ Private = \beta 1(Daily\ time\ online) + \beta 2(Living\ arrangement) + \beta 3(Community\ activity) + \beta 4(SNS\ group)$$

$$Model\ 3: Attitude\ Public = \beta 1(Daily\ time\ online) + \beta 2(Living\ arrangement) + \beta 3(Community\ activity) + \beta 4(SNS\ group)$$

The models are run through IBM SPSS Statistics software (version 25).

7.2.1. Descriptive data

Table 9.5. Descriptive data of the participants.

Sample	Male		Female	
Age	Frequency	Percentage	Frequency	Percentage
Under 40	21	8.57%	20	8.16%
From 41 to 60	20	8.16%	43	17.55%
Over 60	43	17.55%	94	38.37%
Income				

Low	34	13.88%	70	28.57%
Middle	42	17.14%	76	31.02%
High	4	1.63%	5	2.04%
Educational qualification				
Highschool	55	22.45%	105	42.86%
Bachelor’s	28	11.43%	39	15.92%
Master’s and PhD	1	0.41%	2	0.82%
Employment				
Unemployed	3	1.22%	12	4.90%
Retired	31	12.65%	48	19.59%
Part-time	5	2.04%	16	6.53%
Full-time	43	17.55%	63	25.71%
Living arrangement				
Living alone	25	10.20%	63	25.71%
Living with family	57	23.27%	92	37.55%
Underlying illness				
Yes	49	20.00%	102	41.63%
No	36	14.69%	56	22.86%
Disability				
With disability	18	7.35%	22	8.98%
Without disability	66	26.94%	136	55.51%
Members of SNS group (Line, facebook, etc…)				
Yes	42	17.14%	82	33.47%
No	43	17.55%	78	31.84%
Access Internet with mobile phone				
Yes	55	22.45%	86	35.10%
No	32	13.06%	74	30.20%
Discrete/Ordinal variables				
	Male	Scale	Female	Scale

	Mean ±SD		Mean ±SD	
Frequency of seeing doctors	2.41±1.03	(1-4)	2.54±0.99	(1-4)
Daily time online	1.58±0.93	(1-5)	1.58±0.98	(1-5)
Frequency of participating in community activities	1.85±1.39	(1-5)	1.46±1.05	(1-5)
Familiarity with AI technologies	1.62±0.90	(1-5)	1.43±0.66	(1-5)
Discriminatory Concern	3.15±1.30	(1-5)	3.23±1.25	(1-5)
Privacy Concern	2.89±1.21	(1-5)	2.73±1.18	(1-5)
Lose control AI in healthcare	3.08±1.21	(1-5)	2.97±1.13	(1-5)
EAI utility in healthcare facility	3.45±1.11	(1-5)	3.02±1.14	(1-5)
EAI utility for care planning/diagnostics	3.52±1.07	(1-5)	2.95±1.10	(1-5)
EAI-based tracking of health/emotion at home utility	3.04±1.23	(1-5)	3.02±1.16	(1-5)
AI tracking emotions in car utility	3.30±1.15	(1-5)	2.79±1.19	(1-5)

7.2.2. Demographical figures by gender

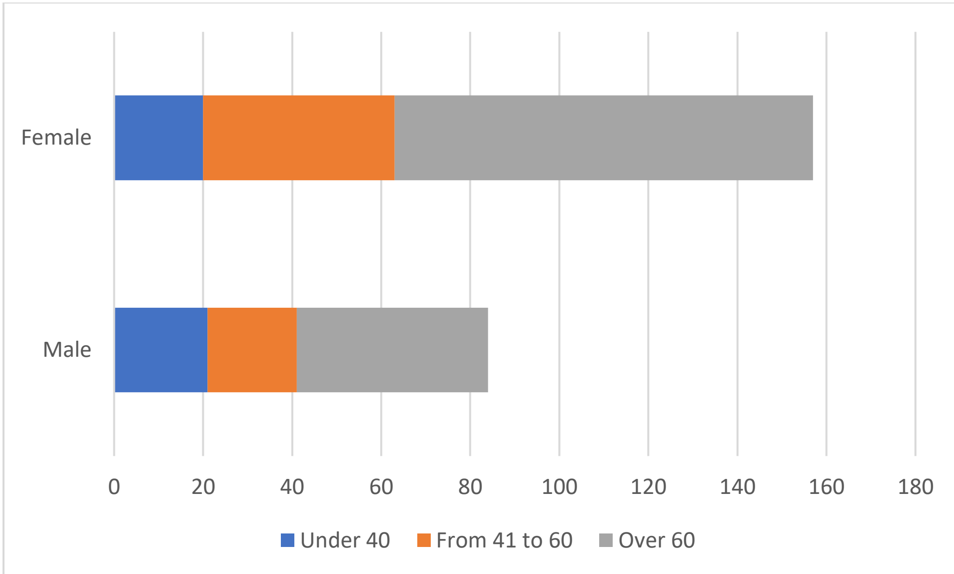
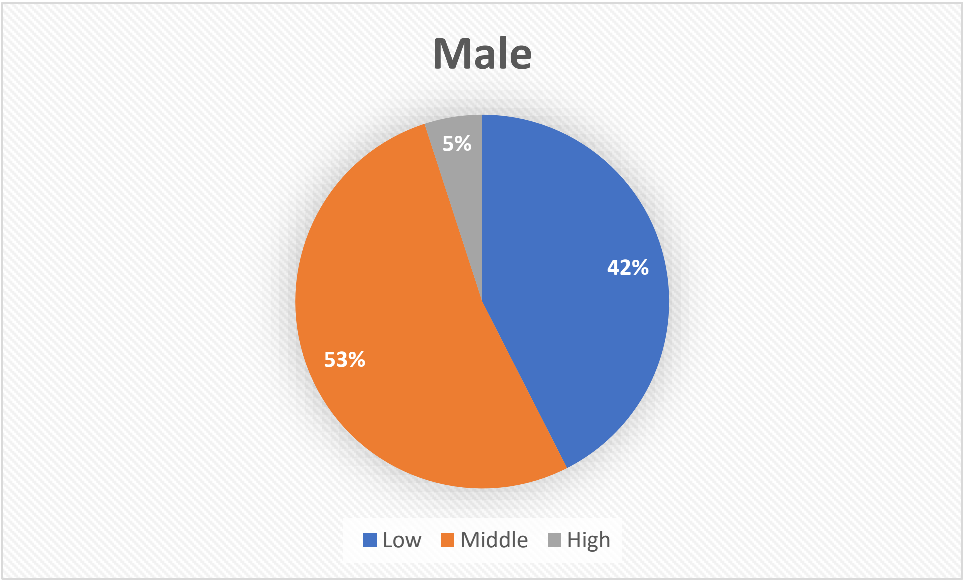


Figure 9.2. Number of patients by gender and age

There are 84 male and 157 female patients in this study. The patients are divided into three age ranges: under 40, 41-60, and over 60, with the number of people as 41, 63, and 137 respectively. Thus, the majority of patients are over 60 years old. By group, female patients over 60 years old occupy the highest proportion.



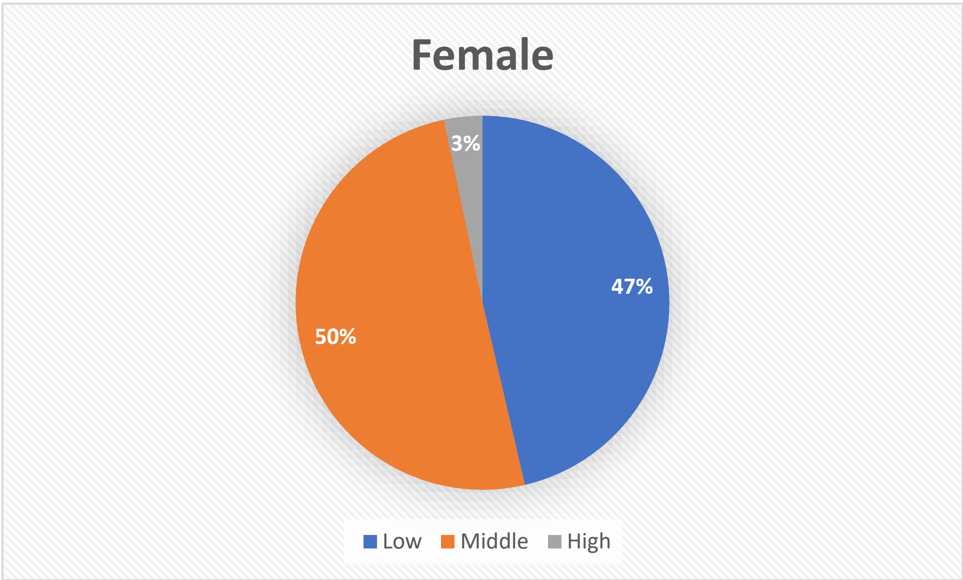


Figure 9.3. Proportion of level of income among the clinic visitors

These pie charts represent level of income by gender (Figure 9.3). Overall, there are not many differences in term of proportion between gender. The proportions of middle-income and low-income patients are relatively similar at the range of 40-50%, while high income patients only account for a fraction of all participants.

Figure 9.4. Numbers of patients divided by work experience group.

There are four group to measure the patients' work experience, which are unemployed (none), part-time (low), full-time (medium), and retired (high). From the figure, we can see that a large percentage of patients have medium to high work experiences. In the next figure, we can see that about two-third of the patients live with their family

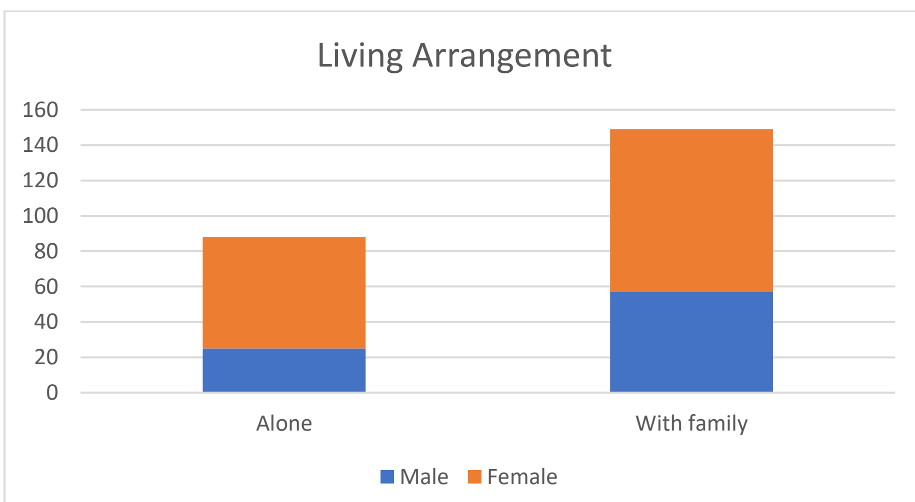


Figure 9.5. Numbers of patients divided by living arrangement type.

7.2.3. Sex differences in attitude toward EAI in healthcare

We performed an additional Chi-squared test to compare the familiarity rate and attitude toward EAI technology between male and female patients. As a result, only the variable of attitude toward EAI healthcare application in public setting display a weak statistically significant result (p=0.062). Here, male patients express a more positive view of the use EAI in public healthcare, while female patients' view is more neutral (See Table 5.2.2).

Table 9.6. Results of analyzing sex differences

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	17.590	10	.062
Likelihood Ratio	18.963	10	.041
Linear-by-Linear Association	8640	1	.003
N of Valid Cases	243		

7.2.4. Behavioral determinants: Ethical concerns and self-rated knowledge

i. Knowledge and Concerns for privacy, control, discrimination

The first half of the first model explore the how self-rated knowledge about AI, concerns for privacy, control, and discrimination predict the attitude toward EAI-integration in the private setting such as home or car.

Regarding the private use of EAI for medical/healthcare purposes, our model shows AI Familiarity HC and losing control to AI in healthcare to be statistically significant variables. The adjusted R square of the first model is yielded at 0.261, meaning the chosen variables correctly predicted 26.1 percent of the participants' attitude on AI private use in healthcare. For the ANOVA test, the regression model was found to be statistically significant (F=22.202***), meaning all variables had significant differences from the overall mean.

Specifically, familiarity with AI in healthcare (β=0.297***) and concern of losing control to AI in diagnosis and treatment (β=-0.262**) are two reliable indicators to predict attitude toward the integration of medical EAI in private setting. Whereas familiarity with AI in healthcare presents positive correlation with the attitude of private AI use, concern for losing control to Ai presents a negative correlation. In other words, those with higher concerns over losing control to AI perceived AI private usage in healthcare with more negativity. Surprisingly, concerns for privacy violations and discrimination implicated in the use of AI in private setting are not significant.

Table 9.7. Results of Model 1_Private and Model 1_Public

	Unstandardized Coefficients		**Standardized Coefficients**		
Model 1_ Private	B	Std. Error	Beta	t	Sig.
(Constant)	3.283	.226		14.544	.000
Privacy	-.084	.073	-.100	-1.146	.253
Lose control AI	-.230	.075	-.262	-3.069	.002
Discriminatory	-.033	.066	-.042	-.509	.612
Familiarity AI comb	.409	.078	.297	5.214	.000

	Unstandardized Coefficients		Standardized Coefficients		
Model 1_ Public	B	Std. Error	Beta	t	Sig.
(Constant)	2.880	.246		11.701	.000
Privacy	.067	.085	.076	.793	.429
Lose control AI	-.176	.087	-.188	-2.027	.044
Discriminatory	-.085	.077	-.098	-1.111	.268

Familiarity AI comb	.530	.091	.346	5.859	.000

The second half of the first model seeks to explain attitude toward the use of EAI in public setting such as hospitals or clinics from behavioral variables of self-rated familiarity with AI, privacy concern, discriminatory concern, and concern for losing control to AI. Overall, the adjusted R square is calculated at 0.179, suggesting that 17.9 percent of the participants' public attitudes can be predicted by the independent variables. For the ANOVA test, the regression model was found to be statistically significant (F=14.281***), meaning all variables had significant differences from the overall mean.

Losing of control over AI in healthcare and AI familiarity were found to be statistically significant. Thus, we can conclude that familiarity (β= 0.346***) and losing control to AI (β=-0.188*) are reliable indicators to predict public attitude. In other words, people with higher concerns over control in medicine perceived the technologies with more negativity and objection. In contrast, higher familiarity with AI application in healthcare and medicine would have a positive impact on subjects' perception of EAI usage in public health facilities. Similar to attitude of EAI usage in the private setting, concerns for privacy and discrimination are not statistically significant predictors.

ii. Discussion 1: Fear of losing control

The analysis results indicate a significant negative correlation between fear of losing control to AI in healthcare and attitude toward EAI applications in both private and public medical setting. In other words, those with more concerns of overdependence on AI-based tools for diagnosis and treatment would perceive the technology with more negativity. There can be multiple interpretations for this result. First, perhaps the fact of our surveyed sample are medical clinic visitors suggest that they might have already perceived themselves as losing autonomy in daily life, and the prospect of having AI technologies making decisions for them can further exacerbate this fear. Second, popular dystopian representation of AI technology in social media might explain how this fear came to be (Bennett, 2014; Ouchchy et al., 2020). As the concept of robot autonomy and AI self-control are usually misunderstood or overexaggerated (Johnson & Verdicchio, 2017), the idea of losing control to AI may loom more strongly to the masses. Policy-wise, we can improve EAI perception in patients through 1) explanation of AI's underlying algorithmic structure even only in high-level abstractions, 2) the current legal and ethical safeguards, 3) what role humans play in the decision-making process. Previous studies have converge into a common feature of human-machine relationship that when users perceive higher level of self-efficacy and having mechanisms to assert meaningful control over the algorithms, they become more comfortable with AI technologies (Lobera et al., 2020; Lu, 2020; McStay, 2020b; Mohallick et al., 2018). Future studies can further identify the aspects of AI dependency that patients are averse to, whether it is AI replacement of human workers or AI dominant control in the treatment process.

iii. Discussion 2: Privacy and discriminatory concern: A cultural interpretation

In this study, both privacy concern and discriminatory concern are found to be statistically insignificant in predicting attitude toward AI tools in healthcare, whether in the public or in the private. Regarding privacy concern, this result can be interpreted as the surveyed Japanese patients are willing to submit their personal data in exchange for diagnosis and treatment accuracy. On the face of it, this result seems surprising given the well-established negative correlation of privacy concern and attitude toward smart technologies in the literature. For example, Lobera et al. (2020) found those who expressed privacy concern are more prone to oppose AI. However, given the context of this study is Japan, this result needs to be framed culturally. Previous studies have found for Asian subjects, health data are considered not as sensitive (Lee et al., 2016; Oderkirk et al., 2013; UN.ESCAP, 2020). Moreover, previous studies on the evolution of the privacy notion in Japanese law show that the status of privacy in many communal contexts such as the workplace or in medical setting is ambiguous as there is an acculturation process (Vuong & Napier, 2015) where the Japanese traditional notion of privacy as symbol of trust in the relationship of self and the collective is interacting with the Western privacy notion rooted in the respect for individual liberty (Miyashita, 2011).

According to Miyashita, law professor from Chuo University, "The laws in Japan say almost nothing about the status of health data in the context of the workplace" (Miyashita H. cited in Ho et al. (2021)). Moreover, there are many doubts on ability to enforce the personal data protection law in Japan given the lenient monetary fine (Oshima & Sakai, 2020) as well as the lack of penalty for companies involved in data breach scandals in the history of the Japan's sole governmental authority on personal data protection, Personal Information Protection Commission (PPC), pointed out in (Miyashita, 2021). The ambiguous relationship between privacy concern and attitude toward AI in medical setting found in this study is perhaps the result produced by such ambiguous acculturation process and the public doubts in the ability for government agencies to protect privacy in the age of AI.

Discriminatory concern is also found to have no significant correlation with attitude toward AI in healthcare. This result diverges from the literature as Lobera et al. (2020) also found those who express egalitarian values are prone to oppose AI, and many books and articles have strongly raised concerns about algorithmic biases against disadvantage groups (N. T. Lee et al., 2019). Discrimination by AI algorithms is the result of misrepresentation or non-inclusive groups of population during the data training process (Norori et al., 2021). AI applications in the healthcare sector is even more prone to discrimination, given the dataset upon which AI algorithms are trained can hardly recognize the bias among biodata of patients with different ethnicities or races (Schönberger, 2019). In this study, however, the discrimination concern is found not to affect patients' perception of AI application in healthcare, for both private and public setting. One of possible explanation is the myth of homogeneity in Japanese society. This reflects the belief of Japanese people as one race, one national identity with no other ethnicities despite the existence of Indigenous Ainu ethnic and immigrated workforces (Howell, 1996). Therefore, the AI-based tools in healthcare should not pose any discrimination concern to the majority, who perceived themselves as homogeneous and indifferent.

On the other hand, it is also possible that the patients' stance on the ethical issues were neutral as they are not well aware of such issues with AI application in healthcare. Future studies examine patients' perception of ethical issues need to pay close attention to the differences in the perceptions between uninformed and well-informed groups. In addition, as our study target is Japanese healthcare system, similar insignificant results may also be found in other Eastern countries with strong sense of collectivism or in contrast, significant results may be achieved in Western countries with strong sense of individualism. Indeed, concurring with the emerging literature on emotional AI (Bakir et al., 2022), our results suggest it necessary for future studies to devise a more culturally sensitive way to systematically hypothesize about ethical concerns in the age of smart machines that interact with our most human and intimate feature, our emotions (Ghotbi & Ho, 2021; Ghotbi et al., 2022; Mantello et al., 2021).

In a recent study, there has been a proposal that such novel method can draw from the mindsponge model of information filtering (Vuong, 2022; Quan Hoang Vuong, 2016; Vuong & Napier, 2015). As a product of studies on how successful expats adapt to the acculturation process, the mindsponge model compels us to think of the mind as the sponge that carries the function of filtering out or absorbing novel inputs such as a new cultural value, a new idea, or in our case, a new technology. The filtering process takes place first with the buffer zone, where initial cost-benefit evaluation takes place, yet the overriding determinant of this filtering process is how the new input interacts with the core personal mindset and external cultural-ideological values. Clearly, this model that is more culturally attuned compared to the TAM model, which primarily focuses on variable of utility and ease of use (Ho et al., 2022). Here, moving beyond the frequentist paradigm, such theoretical leap can be operationalized with the emerging statistical modeling techniques in the Bayesian multi-level analysis, which allow social scientists to compute the effects of hypothesized behavioral factors more precisely, and use socio-cultural factors such as culture, regions, or religions can segment the population in different pools of data (i.e. varying intercepts) (La & Vuong, 2019; Spiegelhalter, 2019; Vuong et al., 2018; Vuong, La, et al., 2020).

7.2.5: Socio-demographic factor, Community, and mobility

i. Age, sex, income, educational level

In the first half of the second model, employment status, sex, income level, educational level, and age were used to predict attitude toward Healthcare EAI in private setting. The adjusted R square was found to be 0.115, making 11.5 percent of the sample predictable by the chosen variables. For the ANOVA test, the regression model was found to be statistically significant (F=6.336***), meaning any of the variables have significant differences from the overall mean.

Age (β=-0.315***) and income level (β=0.166*) were statistically significant. Therefore, age and income level were reliable predictors to measure the attitude for EAI private usage in healthcare. Age was found to have a negative correlation with attitude private, suggesting the older a participant was, the more worries they had toward the technology. In contrast, income level had a positive correlation with attitude public. This means participants with higher income level perceived private EAI usage with more positivity.

Table 9.8. Results of Model 2_Private and Model 2_Public

		Unstandardized Coefficients		**Standardized Coefficients**		
Model 2_ Private		B	Std. Error	Beta	t	Sig.
1	(Constant)	3.323	.418		7.953	.000
	Sex	.129	.146	.059	.881	.379
	Age	-.016	.004	-.315	-4.351	.000
	Income level	.310	.126	.166	2.451	.015
	Educational level	-.034	.149	-.016	-.226	.821
	Employment status	.005	.087	.004	.055	.957

		Unstandardized Coefficients		**Standardized Coefficients**		
Model 2_ Public		B	Std. Error	Beta	t	Sig.
1	(Constant)	3.277	.448		7.320	.000
	Sex	.363	.156	.155	2.330	.021
	Age	-.011	.004	-.206	-2.838	.005
	Income level	.263	.135	.131	1.950	.053
	Educational level	.199	.160	.087	1.244	.215
	Employment status	-.078	.093	-.058	-.838	.403

In the second half of the 2nd model, similar independent variables were used to predict the attitude toward EAI usage in public healthcare facilities. The adjusted R square was calculated to be 0.110, meaning 11 percent of the patients' attitude can be predicted by the independent variables.

For the ANOVA test, the regression model was found to be statistically significant (F=6.198***), meaning all variables had significant differences from the overall mean.

For this model, age (β=-.206**). and sex (β=.155*) are the two variables with statistically significant value Thus, they are reliable variables to predict attitude of patients toward EAI usage in public healthcare facilities. With a negative correlation, individuals with higher age in the study perceived the technology with more negativity.

ii. Community and mobility

In the first half of the 3rd model, daily time online, living arrangement, community activity, and SNS group were chosen as independent variables to predict attitude toward EAI in private healthcare setting. The adjusted R Square was found to be 0.094, indicating the chosen variables correctly predicted 9.4 percent of the participants' attitude on EAI private use in healthcare. For the ANOVA test, the regression model was found to be statistically significant (F=5.854***), meaning all variables had significant differences from the overall mean.

Among the chosen independent variables, community activity (β=0.127*) and access to internet through mobile (β=0.320**) were statistically significant. As a result, they were two reliable variables to predict attitudes toward EAI private usage in healthcare. Examining their relation, as both variables had positive correlations with the attitude variable, we could conclude that a higher level of social interaction with surrounding community and accessibility to the internet would positively impact patients' perception of EAI private use in healthcare.

Table 9.9. Results of Model 3_Private and Model 3_Public

Model 3_ Private		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	2.169	.276		7.854	.000
	Daily time online	.009	.080	.008	.106	.915
	SNS Group	-.104	.207	-.049	-.503	.616
	Access Internet Mobile	.684	.218	.320	3.135	.002
	Community activity	.113	.056	.127	2.027	.044
	Living arrangement	.091	.142	.042	.644	.520

Model 3_ Public		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	2.154	.299		7.211	.000
	Daily time online	.122	.088	.098	1.386	.167
	SNS Group	-.018	.227	-.007	-.077	.939
	Access Internet Mobile	.600	.239	.254	2.506	.013
	Community activity	.142	.061	.143	2.317	.021
	Living arrangement	.138	.154	.057	.897	.371

In the second half of the 3rd model, we applied similar independent variables to predict attitudes toward EAI usage in public healthcare setting. The adjusted R Square was calculated to be 0.114 meaning 11.4 percent of the participants' public attitudes can be predicted by the independent variables. For the ANOVA test, the regression model was found to be statistically significant (F=7.116***), meaning any of the variables have significant differences from the overall mean.

Community activity (β=0.143*) and access internet mobile (β=0.254*) were statistically significant variables. Therefore, the activeness level in local community activities and accessibility to the internet are reliable positive predictors of patients' attitudes toward EAI applications in public healthcare facilities. These correlations are similar to the findings of first half of the 3rd model.

iii. The challenges of promoting AI tools for the elderly

Age had a negative correlation on both patient's attitude toward EAI's private (β =-.315***) and public usage (β= -.206**). This finding suggested that Japanese elderly patients were having negative perceptions of EAI-based tools in healthcare in both private and public settings, which is problematic as the target demographic of AI-based healthcare utilities is the older generation in Japan. This result is largely consistent with the literature, for example, He et al. (2022) found that compared to the elderly, young Japanese tend to accept the new contact-tracing apps during the COVID-19 pandemic.

Interpreting this result using the TAM, the surveyed Japanese elderly are likely to perceive little or no usefulness and familiarity (ease of use) regarding the emerging AI tools in their healthcare. Concurringly, in a study on Japanese perception of smart health services, Shimizu et al. (2022) found that perceived benefits positively correlate with social acceptance of the new technologies. Moreover, Anaraky et al. (2021) found that older people (65+) are only willing to adopt apps that ask them to disclose personal data if they perceive higher benefits in disclosing. It is also likely that the elderly Japanese in our survey perceive themselves as having little or no knowledge about the technology, as in section 3, we have found that the more familiar a citizen became with AI-based healthcare tools, the more positively they perceived the technology, and vice versa.

This result implied, policy-wise, increasing awareness of AI through education and media can produce a positive impact on EAI perception in healthcare. For instance, conventional channels of media such as TV and radio should promote AI healthcare familiarity to a large demographic of aging patients, while workshop for senior citizens at local healthcare facilities on EAI-based tools will increase their awareness of the benefits of the technology, thus improving its acceptance as suggested by the TAM framework.

However, it is important to note that our familiarity variable doesn't include the entire spectrum of AI knowledge as familiarity doesn't correspond to a thorough understanding of the technology of both its technical details and its social implications. We suspect a broader understanding of social and ethical issues related to AI might induce more averseness of the technology (Brougham & Haar, 2018; Ghotbi & Ho, 2021; Ghotbi et al., 2022). Critically, previous experimental studies have shown once the research subjects were given more information about new technologies, their perception could shift significantly (Lima et al., 2020). Thus, future studies can examine the impact of thorough understanding of AI-based medical tools' threats on EAI's perception.

iv. Improving perception of AI-integration in healthcare through social interactions

Social interaction, measured through community activity level and accessibility to internet, were found to be a positive influence on patients' EAI perception. Specifically, community activity level had a positive correlation with both private (β=0.127*) and public attitude (β=0.143*). In other words, people who frequently interact with the community find the EAI in private and public setting more appealing. Perhaps, those who are more engaged with the community are more open-minded about the benefits of new technologies. How can we utilize this correlation to improve the perception of AI-based utilities in healthcare? Introduction of EAI-based tools at community gathering facilities can help elders familiarize with the technology, while indirectly having a positive impact on their attitude toward the technology being used in the diagnosis and treatment process. At the same time, access to internet mobile had a positive correlation on private attitude and public attitude. This finding can be interpreted as patients with experience to internet will also be more susceptible to accept AI technology. Thus, virtual social interaction can also serve as a powerful tool to spread awareness of AI-based applications in healthcare. Communication among patients through the SNS platform can be an ideal environment where elderly patients get more information about EAI technology.

7*.4. Chapter summary*

Chapter 7 has provided the empirical results from two studies based on a national, representative survey of 2,000 Japanese subjects and a municipal survey of 245 clinic visitors in Beppu City, Oita Prefecture, Japan. In both cases, we don't find major, significant sex differences in that attitude toward emotional AI in healthcare setting. However, the national survey's analysis reveals women possess more concern about the ethical implications of emotional AI for healthcare: the loss of autonomy, the undue influence of AI that making sensitive judgments about emotion and psychology of at-risk populations, the human replacement by AI, and how data on patients are stored and managed. These concerns are related to the moral dimension of Harm/Care, which has been found by empirical research on Moral Foundation Theory to be more of a concern for women.

In both cases, we find that concern about the undue influence of emotional AI in medical processes, i.e., the loss of autonomy, is negatively correlated with attitude toward the technology. This suggests that the integration of emotional AI in healthcare sector must respect the value of autonomy. Meanwhile, concerns about biases in AI systems, privacy implications, and how patients' data are managed are either non-significant or positively correlated with acceptance of emotional AI in the medical setting. Such counterintuitive results are not aligned with the predictions of the Moral Foundation Theory. Thus, the result on privacy concern must be interpreted as a reflection of the cultural attitude among Japanese people that medical information has a communal dimension and is not treated as completely a private matter, while the result on bias concern can be interpreted as a reflection of the long-standing cultural belief in a homogenous Japanese society. This chapter has also discussed the results in relation to the current literature and their policy implications. It is suggested that future studies should consider cross-cultural differences in core values and their influence on moral reasoning about the effects of technologies in our lives.

Chapter 8: Public security: Security camera and social media policing

In this chapter, we will systematically look at applications of emotional AI for public security in two cases: the physical space and the virtual space. Specifically, we will study how the Japanese population feels about emotional AI when it is installed in security cameras, and when it is used to surveil social media. As mentioned in the literature review and the introduction section, emotional AI has been widely promoted as a solution for public security all over the world. According to James Wright, a researcher at the Alan Turing Institute, UK, emotional AI algorithms of ELSYS are now being adopted by security companies and government police forces in China, Japan, Russia, and South Korea to surveil public spaces such as convenient stores, ATMs, sports and music events (Wright, 2021). As documented in the literature, legal experts have identify a number of ethical and legal issues that come with a more widespread use of facial recognition technologies, which include, emotional AI security cameras: the risk of discrimination against minority groups; the risk of copyrights violation due to the automatic processing, gathering, curating of images; the risk of infringement on human dignity (Cabitza et al., 2022). Most of the emotional AI security camera systems in the market such as that of ELSYS or NEC purport to detect various emotional states, suspicious behaviors, physical movements in real time with great accuracy. However, technical analysis of the reliability of these systems and interviews with police officers show they have very low reliability, thus risking inaccurate even wrongful interventions into the citizen lives (Urquhart & Miranda, 2022; Urquhart, Miranda, et al., 2022b). In recent years, there has been a growing interest in appraisal-based emotional AI camera systems (McStay & Urquhart, 2019), which classify emotions not only by depending on still images and basic emotion categories but also on multi-faceted contextual, physiological or personal information. In a technical analysis by Cabitza et al. (2022), while additional contextual information does improve the accuracy of these security camera systems, the authors emphasize such the systems will pose greater risk of privacy and morality violation.

In Japan, security cameras with AI facial recognition capacity have been reported to be in rising demand. For example, the market research company Fuji Keizai estimates the Japanese domestic market for commercial security cameras is expected to grow from ¥56.3 billion in 2020 to ¥61.9 billion in 2024. There have been many prominent examples of AI security cameras being reported to the public. For example, one of the most debated issues is whether security cameras should be made mandatory by the government, following high-profile attacks on passengers such as the knife attack on a Tokyo train in October 2021. In this incident, it is reported that since there is no camera on the train, the police as well as the train operators were unable to devise a full account of what had transpired during the knife attack (The Japan Times, 2021). This has prompted public discussion and the government's consideration of mandatory installation of security cameras on all trains. Security reasons aside, AI-powered security cameras are also considered among vital solutions for monitoring an aging population with more and more people suffering from dementia. A recent report claims that around 17,000 people with dementia went missing in 2020, up from 9,600 in 2012, which makes dementia the leading cause of missing-person cases in Japan (Dooley & Ueno, 2022).

For example, with the purpose is to detect suspicious people, lost children, elderly people in trouble, and physical distress, AI security cameras have been tested in the underground plaza of Hisaya-Odori Park in Nagoya by NTT Communications Corp. in 2020. Or Panasonic Corp. is reported to develop an indoor HD camera that can detect movement, temperature, and other factors to monitor children or the elderly via smartphones; Nagoya-based Digital Cube Technology launched an AI camera that can give verbal warnings in fall 2021. A more controversial example is the plan to introduce a network of 5,000 security cameras with live facial recognition capability to detect ex-offenders by East Japan Railway Corp (JR East). There has been a strong public pushback that resulted in the companies dropping the plan (Ogawa & Akada, 2021).

Commented on such use of facial recognition technology, Miyashita, the professor of constitutional law at Chuo University, explained a private-sector company such as JR East should not be allowed to access and use facial information of people who have served prison terms in that way. Yet, other experts take a softer stance. For example, Yusaku Fujii, a professor of safety engineering at Gunma University, thinks it is inevitable that AI security cameras will be used more in the future, and urges it is time to determine the social norms regarding sensitive issues such as the privacy of ex-offenders and parolees (Ogawa & Akada, 2021).

Regarding the surveillance of social media using AI technology, there have not been many academic studies or news reports on the case of Japan. Thus, this chapter provides the first evaluation of the public perception of the use of emotional AI for surveilling social media sites in the Japanese population. Before delving into the quantitative results, we first look at a set of insights from the qualitative thematic coding of interviews and citizen workshops related to the use of emotional AI for security purposes.

8.1. Security camera

8.1.1. Qualitative insights

In the course of this study, we have interviewed four companies that are directly involved in developing and selling of emotional AI: ELSYS Japan, Preferred Networks, and two other companies who want to remain anonymous (See Chapter 3 for more details on these companies).

All companies extoled the utilities of their AI products, stating numerous cases where AI solutions can make life easier and more secure for people. For example, Preferred Networks are currently developing deep learning algorithms to make assembly lines and factory floors safer analyze video camera images and provide real-time updates. Or the AI-driven security system developed by ELSYS have been used in convenient stores, factory plants, sports events, even schools in Japan for detecting suspicious behaviors, nervousness, and anxiety of people in the public places.

Regarding the accuracy concern of the technology, when the interviewees were asked about what they think of the current debate in science community about the nature of emotions and the controversy around Paul Ekman's theory of universal emotions, all three companies express some concerns, however, they remain positive that this situation will be solved by better algorithm and better, more diverse datasets. A senior member of emotional analytics team in a Japanese global security company emphasizes: "As we add more layers of data on top of our emotion-detecting AI (facial expressions, biometrics, voice, etc.), we are moving toward a perfect capture of people's inner sphere" (Interview data, 2021).

All companies divulge that they have been cooperating with local universities and companies to acquire more diverse datasets to better train their algorithms. For example, an interviewee states her company compares local Japanese data-annotators and Japanese growing up in the US to make sure the datasets "have the diversity in their ethnicity and a cultural background… For example, we have a test that, how do read the smiles of the Japanese between American Japanese, Japanese growing up in the US, or Japanese, grown up in Japan. And we don't have much to differences at this moment" (Interview data, 2021). However, critically, the interviewees were not able to clarify how their companies will move away from the controversial Ekman’s model of emotions.

Regarding the issue of privacy, worryingly, in some places, emotional AI technologies are being used unchecked. For example, the distributor of a Russian security algorithm that detects emotions from human head’s micromovements, Vibra~~n~~image, ELSYS Japan’s representative stated: “Although we distribute stickers that say, ‘Monitoring with ELSYS’ to comply with privacy regulation in Japan, our clients are not using these stickers” (Interview data, 2021).

To mitigate harms, one interviewee stated that he has urged their clients to not to use emotion-sensing technologies without clear guidelines as well as any use of the technology to be verified by third party ethics committee. However, it is admitted by most interviewees that we are still in an uncharted territory regarding the regulation of emotion-sensing technology. A clear example is the case of ELSYS product, for which, a manager at ELSYS Japan states “The products were originally developed in Russia and have undergone quite a few human experiments. It is quite unthinkable in Japan but because it was undergone in Russia, I think these experiments were possible.” More importantly, he also confirmed the current software product has not been tweaked and adapted to the Japanese population. This example highlights an urgent need to regulate this new technology whether by formalized legal rules or by social norms.

In the citizen workshops, the issues related to emotional AI security cameras have been discussed in-depth. All groups share the concerns for the lack of current regulation and its potential for overreaching into the lives of the people. They think even with this technology, it is more important that human security guards are well-trained and well-adapted to such technology, and not letting the technology dictates what they should do in a real-life situation involving humans. Comparatively, the group with physical disabilities express the most willingness to adopt and use the technology, while the group of people over 60 and the foreigner's group are more skeptical over its use. One of the foreign participants stated: "London is one of the most surveilled city in the world, but it is not the safest," which he cites as a warning that more surveillance is not necessarily means more security.

Security Camera

Below is the explanation given to the survey respondents before they were asked to give answers on how much they agree or disagree with a series of Likert-scale questions concerning the use of emotional AI in security cameras.

> "Some private companies provide camera-based security services, claiming they can detect a person's mental status or emotional level, such as whether they are feeling aggressive, tense, or stressed. These systems may then label a person as suspicious before a crime is committed. Such systems are designed to be used in a range of places, including convenience stores, shopping centers, and transport hubs."

8.1.2. Descriptive statistics

The descriptive statistics on general attitude toward EAI security camera shows that nearly half of the respondents remain neutral about the technology (48.7%). For those who take a side, there are more than double the number of people who feel positive about the technology than those who do not (35.4% vs. 16.1%). The average answer for this question is 3.14 (std = 0.914). This suggests that the Japanese population is slightly in favor of the emotional AI security camera.

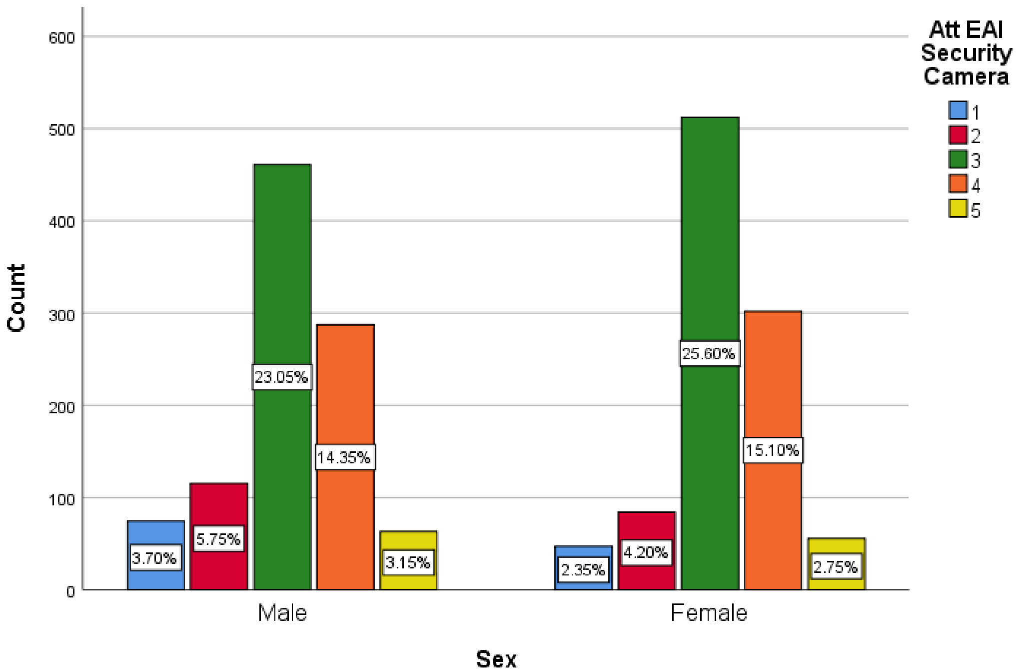


Figure 10.1. Distribution of attitude toward emotional AI security camera by sex.

8.1.3. Sex differences

Table 10.1. Sex differences in attitude toward and concerns about emotional AI in security camera.

Sex		SafetyUtility	Fairness/Bias Concern	PreCrime Concern	Knowledge	Privacy Concern
Male	Mean	3.28	3.54	3.26	3.16	3.31
	N	1000	1000	1000	1000	1000
	Std. Deviation	1.019	.957	.997	.960	1.021
Female	Mean	3.43	3.61	3.30	3.16	3.34
	N	1000	1000	1000	1000	1000
	Std. Deviation	.917	.862	.886	.872	.931
Total	Mean	3.36	3.57	3.28	3.16	3.33
	N	2000	2000	2000	2000	2000
	Std. Deviation	.972	.911	.943	.917	.977

Sex		Accuracy Concern	Att EAI Security Camera	TrustGov	TrustPrivate
Male	Mean	3.52	3.15	2.90	2.93
	N	1000	1000	1000	1000
	Std. Deviation	1.006	.964	1.000	1.010
Female	Mean	3.63	3.23	2.90	2.99
	N	1000	1000	1000	1000
	Std. Deviation	.858	.860	.927	.879
Total	Mean	3.57	3.19	2.90	2.96
	N	2000	2000	2000	2000
	Std. Deviation	.936	.914	.964	.947
Range: 1 (strongly disagree) to 5 (strongly agree).					

Applying the chi-square test, we find meaningful differences between the sexes in the all the variables except for the trust in the government to regulate the technology (TrustGov, $p < 0.001$): SafetyUtility ($p=0.01$); Fairness/Bias Concern ($p=0.015$); PreCrimeConcern ($p=0.006$); Knowledge ($p=0.007$); PrivacyConcern ($p=0.25$); AccuracyConcern ($p<0.001$); Attitude toward EAI security Camera ($p = 0.006$); TrustPrivate ($p<0.001$).

Specifically, on average, female respondents agree more with the increased safety (SafetyUtility) and benefits (Att EAI Security Camera) brought about by EAI security cameras than their counterparts. These results agree with the findings for the Chinese sample, but not the German, UK, and US samples (Kostka et al., 2021). Female respondents had a higher acceptance rate of facial recognition technology in the Chinese group while in contrast, male respondents had a higher rate in the German group.

However, women respondents express more worries that EAI security cameras would entail biases toward disadvantaged groups, precrime concerns, privacy concerns, and accuracy concerns as the means for females are higher while standard deviations are smaller compared to males among the examined variables. Under the lights of the Moral Foundation Theory, the higher concerns for the moral dimensions of harm and fairness in women have been established in a recent study by Atari et al. (2020) that look at sex differences in moral judgments across 67 countries. Past studies on the subject also found women tend to care more for the dimensions of Harm, Fairness, and Purity, while men tend to care more for the dimension of Loyalty and Authority (Graham et al., 2011). In this context, the ethical concerns for EAI security camera should reflect on the harm and fairness dimension of moral judgments, thus are scored higher accordingly by the female respondents.

8.1.4. Socio-demographic factors

Table 10.2. Regression results for socio-demographic factors and attitude toward emotional AI in security camera.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	3.436	.111		30.918	.000
	Age	-.009	.002	-.130	-5.214	.000
	Income	.057	.033	.045	1.730	.084
	Education	.013	.013	.025	.981	.327

a. Dependent Variable: Att EAI Security Camera

First, age is a *negative* significant predictor of attitude toward the use of Emotional AI security camera (β_{age}=-0.009***, $p<0.001$). This finding on the Japanese population is contradictory to the results of the analysis in other countries' samples. For example, Kostka et al. (2021) found a small, but positive statistically significant correlation between age and facial recognition technology in the UK and US samples. In the same study, regarding the Chinese and German samples, the authors found no significant results for age and facial recognition technology acceptance.

Second, Income is a *positive* predictor of attitude toward EAI security camera ($\beta_{Income}=0.26$), although the p-value test yields relatively weak results (p=0.082). In comparison, Kostka et al. (2021) found a significant positive effect of income on facial recognition technology in China, the UK, and the US samples, but not Germany.

Meanwhile, education has no statistically significant relationship with attitude toward EAI security camera. This result is similar to the results found for facial recognition technology in China, the United Kingdom, and the United States, and different from the significant positive correlation between the level of education and facial recognition acceptance in the German sample (Kostka et al., 2021).

Here, to a certain extent, the TAM framework can explain the significant results of income and age. For the age variable, applying the TAM framework, we can reason that elderly people in Japan are likely to perceive the EAI security camera as irrelevant (thus, low in perceived utility) and unfamiliar (hence, low in perceived ease of use). Regarding the income variable, extrapolating the TAM framework can explain the positive correlation between income and acceptance of EAI security camera: higher income might perceive themselves as having more to lose, thus seeing EAI security camera as offering an improvement in safety (i.e., improved perceived utility). Moreover, as the higher-income respondents are more likely to afford similar technologies, i.e., CCTV or live facial recognition technology, one can reason that they would perceive EAI security camera as more familiar (i.e., higher perceived ease of use).

However, once we look beyond the case of Japan and contemplate the cross-cultural differences among different countries' populations, it is evident that such cross-cultural differences cannot be fully captured with the underlying assumption of universality in the TAM or the Moral Foundation framework.

8.1.5. Regression analysis

Table 10.3. Regression results for behavioral determinants of emotional AI in security camera.

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.568	.082		6.918	.000
	SafetyUtility	.320	.018	.340	17.343	.000
	Fairness/BiasConcern	.062	.020	.061	3.144	.002
	PreCrimeConcern	-.047	.020	-.048	-2.366	.018
	Knowledge	.240	.018	.241	13.189	.000
	PrivacyConcern	-.120	.019	-.128	-6.308	.000
	AccuracyConcern	.061	.018	.062	3.390	.001
	TrustGov	.117	.020	.123	5.938	.000
	TrustPrivate	.192	.020	.199	9.563	.000

a. Dependent Variable: Att EAI Security Camera; R square = 0.582

The model explains about 58% variation in the data, which is quite similar to many other studies that use an extended TAM framework (J. Lee et al., 2019). The regression analysis yields three interesting sets of results. First, in terms of positive correlates with attitudes toward EAI security cameras, the more people agree with the statements about increased safety ($\beta_{SafetyUtility}=0.34^{***}$), have a basic understanding of the technology ($\beta_{Knowledge}=0.241^{***}$), and trust government and private actors to regulate the technology ($\beta_{TrustGov} = 0.123^{***}$; $\beta_{TrustPrivate}=0 .199^{***}$), the more they think Emotional AI in security camera will benefit society.

These findings agree with the predictions of the TAM model, as more perceived ease of use and perceived usefulness lead to more acceptance of new technology.

Second, *pre-crime concern* and *privacy concern* are negative correlates of attitude toward EAI security camera. That is to say, the more people agree that security camera with AI capabilities to recognize emotions seems dystopian and very intrusiveness, the less they think the technology will benefit society ($\beta_{PrecrimeConcern}$ = -0.048*; $\beta_{PrivacyConcern}$ = -0.128***). The results on the pre-crime concern and privacy concern can be explained within the Moral Foundation Theory. The theory predict as violations of a moral norm would lead to a rejection or harsh judgment of a new technology (Hidalgo et al., 2021). Here, in the case of emotional AI security cameras, privacy violation and precrime concern, which can be mapped on the harm/care dimension (Graham et al., 2011), thus, logically these concern can undermine the acceptance of the emerging technology.

It is worth noting a context-sensitive result: the finding on the negative correlation between concern about dystopian/precrime feature of emotional AI security camera and its attitude diverge with the case of emotional AI cars, where the variable of concern about dystopian feature of AI systems that constantly monitoring people's emotions in cars does not have a statistically significant association with attitude toward the technology. In other words, people view EAI-enabled security camera more skeptically than EAI-enabled interior sensing in cars, and they worry emotional AI security camera would bring about a more dystopian future.

Third, paradoxically, concern about the accuracy of the technology ($\beta_{AccuracyConcern}$= 0.062***), and concern about biases/unfairness toward disadvantaged groups ($\beta_{Fairness/BiasConcern}$= 0.061**) are positive correlates of attitude toward Emotional AI Security Camera. To interpret these paradoxical results, respondents who agree that EAI security cameras will benefit society are equally aware of the downsides of the new technology: lack of accuracy or increase in biases toward disadvantaged groups. It is also worth noting that the correlation between attitude toward EAI security camera and the acknowledgment of its safety utility is much stronger than between attitude toward the technology and concerns for biases or the lack of accuracy (nearly 6 times stronger), we can interpret it as among the surveyed factors, the most important determinant of acceptance of EAI security camera is the people's desire to improve their security. Arguably, this result reflects the strong sentiments expressed in Japanese media regarding the recent high-profile knife assaults on a local train in Tokyo, which has prompted public discussion and the government's consideration of mandatory installation of security cameras on all trains (The Japan Times, 2021).

Here, it is unclear how to make sense of the results on the accuracy and bias concerns within the Moral Foundation Theory. Concerns for the lack of accuracy of and biases within a new technology can be mapped to the moral dimension of Fairness in the Moral Foundation Theory, and these concerns should negatively correlate with the acceptance of the technology. Yet, they are two positive correlates. This means Japanese people who accept the emotion-sensing security cameras also accept its flaws, i.e., its lack of accuracy and its biases toward disadvantaged groups. On the one hand, this attitude can be a result of the acceptance that it is inevitable that technology will have certain flaws. This resonates the earlier comment by Yusaku Fujii, a professor of safety engineering at Gunma University, where he believes it is inevitable that AI security cameras will be used more in the future, thus it is wise to start thinking about the social norms that should be applied in governing its uses (Ogawa & Akada, 2021). As per the market research company Fuji Keizai's estimations, the Japanese domestic market for commercial security cameras is expected to grow from ¥56.3 billion in 2020 to ¥61.9 billion in 2024, and AI capabilities are the main selling point for this product. Security is not the only social concern, but also as with the rapid aging of the Japanese population, surveillance cameras, those that could detect in real-time movements and emotions, are seen as a inevitable solution for the supervision of elderly people suffering from dementia. It is estimated that around 17,000 people with dementia went missing in 2020, up from 9,600 in 2012, which makes dementia the leading cause of missing-person cases in Japan (Dooley & Ueno, 2022).

Figure 10.2. A digital surveillance camera in Itami, Japan. Source: Hiroko Masuike/The New York Times.

The lack of concern about the inaccuracy and the bias toward disadvantaged group in people's attitude toward emotional AI security camera might also reflect Japan's long-lasting beliefs in cultural homogeneity as well as patriarchal values, which have been well-documented in the literature (Howell, 1996; Woo, 2022). As the Japanese people strongly believe they are ethnic homogenous and have strong sense of collectivism, the concepts of bias and discrimination by EAI algorithm might not appear as saliently and as strongly as in Western culture where individualism and multiculturalism dominate.

In fact, in the citizen workshops, the Japanese participants only start to consider the problems of biases when the discussion turns explicitly toward the topics of foreign workers in Japan might be inappropriately profiled as the results of using the emerging emotion-sensing security camera. For example, the distributor of a Russian security algorithm that detects emotions from human head's micromovements, Vibrainimage, Elsys Japan's representative stated: "Although we distribute stickers that say, 'Monitoring with ELSYS' to comply with privacy regulation in Japan, our clients are not using these stickers" (Interview data, 2021).

Clearly, the novelty the technology has impacted people's judgment of the technology. The issue with the accuracy of the technology will affect people's perception in the long term, a point that has been covered in the work of Urquhart and Miranda (2022). For example, the authors state when EAI-based live facial recognition is used as admissible evidence in court, its accuracy or lack thereof takes on significantly higher stake. It is hard to predict how the perception will change once the application of EAI in security camera is commonplace. Thus, future studies can further expand on this area by studies populations of people who have direct, real-life experiences with this technology. It is also important to note that the right to privacy of identity is strongly emphasized in Japan constitution in Article 13, where the collection of citizen's photographed identity without consent or legitimate reason is prohibited (Ozaki, 2020). Arguably, citizens who are aware of their constitutional privacy rights should be more opposed to the EAI security camera, and this factor should be a subject for future studies.

Next, we will turn our attention to the case of using emotional AI for social media policing.

8.2. Social media policing

Below is the explanation given to the survey respondents before they were asked to give answers on how much they agree or disagree with a series of Likert-scale questions concerning the use of emotional AI in social media policing.

> "Police forces use computer software to search, understand and monitor social media posts to measure the strength of feeling in what people say about a topic. Police can also establish the location of where specific social media users are posting messages from. This helps the police to decide how public demonstrations and protests should be policed. This can help police direct their resources to trouble hot spots by helping them to decide how many officers to send, and which crowd-control techniques and equipment (such as shields and weapons) to use."

8.2.1. Descriptive statistics

Regarding the attitude toward using emotional AI for social media policing, the majority of the respondents remain neutral on this topic (51.3%). More people are positive about the technology than those who are negative: 30.6% of the sample responded positively toward the technology versus 18.2% who responded negatively. Similar to the case of emotional AI security camera, the average score for the attitude toward using emotional AI policing social media is 3.13 (std = 1.067, i.e., slightly more distributed than the attitude toward emotional AI security camera). This suggests that Japanese people, on average, tend to accept EAI technology. However, there are more varying attitudes compared to the case of security camera. Perhaps, this is due to the lack of a physical presence as well as the comparative novelty of emotional AI systems that police social media platforms.

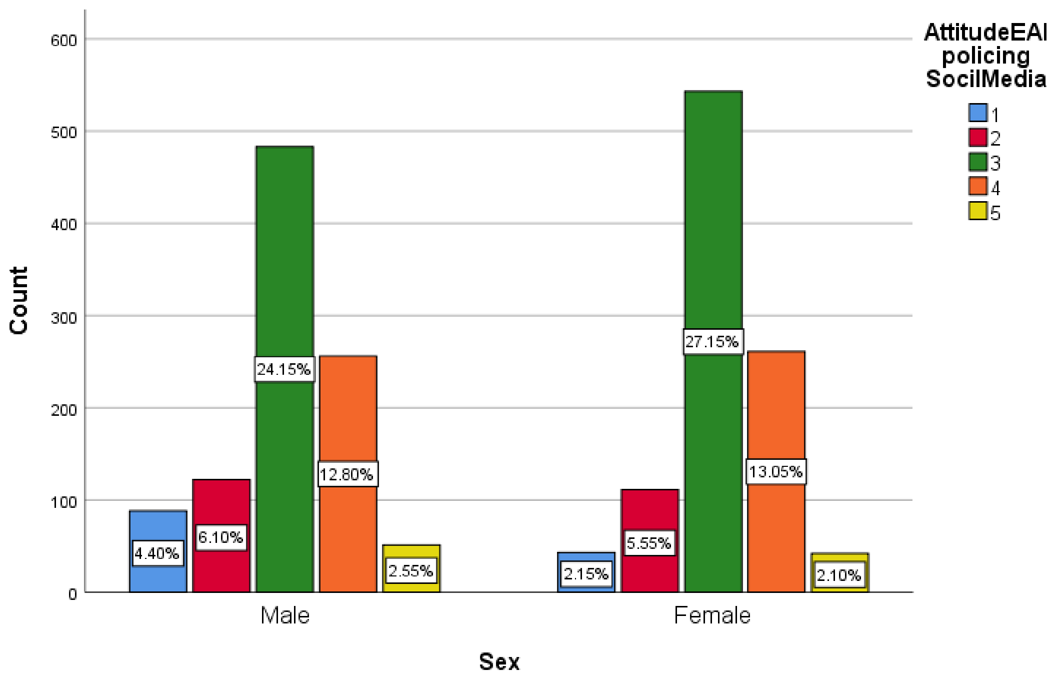


Figure 10.3. Distribution of attitude toward emotional AI used in social media policing. Range 1 means strongly disagree to 5 means strongly agree.

8.2.2. Sex differences

Table 10.4. Sex differences in the attitude toward and concerns about emotional AI in social media policing.

Sex		Safety Utility	Autonomy Concern	Dystopian SurveilConcern	Knowledge	Bias Concern
1	Mean	3.07	3.02	3.28	3.19	3.04
	N	1000	1000	1000	1000	1000
	Std. Deviation	1.124	1.032	1.073	1.042	1.045
2	Mean	3.20	3.15	3.30	3.19	3.13
	N	1000	1000	1000	1000	1000
	Std. Deviation	1.004	.914	.951	.964	.946

Total	Mean	3.13	3.09	3.29	3.19	3.08
	N	2000	2000	2000	2000	2000
	Std. Deviation	1.067	.977	1.014	1.003	.997

Sex		Accuracy Concern	AttitudeEAI police SocMedia	TrustGov	TrustPrivate
1	Mean	3.42	3.06	2.87	2.91
	N	1000	1000	1000	1000
	Std. Deviation	1.028	.965	1.030	1.001
2	Mean	3.50	3.15	2.89	3.00
	N	1000	1000	1000	1000
	Std. Deviation	.886	.831	.910	.895
Total	Mean	3.46	3.10	2.88	2.95
	N	2000	2000	2000	2000
	Std. Deviation	.960	.901	.971	.950
Range 1 (Strongly disagree) to 5 (strongly agree)					

Conducting the Chi-square test, we find statistically significant differences between the sexes in the following variables: Safety Utility (p<0.001); Autonomy Concern (p<0.001); Dystopian Surveillance Concern (p=0.001); self-rated Knowledge (p=0.05); Bias Concern (p=0.009); Accuracy Concern (p< 0.001); Attitude EAI police social media (p< 0.001); TrustGov (p=0.002); Trust Private (p =0.005).

Specifically, female respondents are, on average, more positive about the safety utility, and overall benefits of EAI use in policing social media. Under the TAM, this result makes sense as females and sexual minorities have been found to be at a higher risk of receiving aggressive and violent behaviors, both online (e.g., cyberbullies, harassment, etc.) and offline (Aboujaoude et al., 2015; Chowdhury & van Wee, 2020). Hence, female respondents are more sensitive to the issue of improving safety in digital platforms.

However, the female respondents are also on average more concerned about freedom for protest, the accuracy of the technology, biased algorithm, and a dystopian surveillance state brought about by such use of the technology. Here, the results are aligned with the empirical findings related to the Moral Foundation Theory, as women are found to be more concerned about the moral dimensions of Harm and Fairness (Atari et al., 2020; Graham et al., 2011).

8.2.3. Regression analysis

Table 10.5. Regression results for behavioral determinants of attitude toward emotional AI in social media policing.

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.600	.077		7.803	.000
	SafetyUtility	.196	.019	.233	10.510	.000
	AutonomyConcern	.073	.019	.079	3.850	.000
	DystopianSurveilConcern	-.086	.017	-.097	-4.987	.000
	Knowledge	.206	.018	.230	11.668	.000
	BiasConcern	-.016	.018	-.018	-.929	.353
	AccuracyConcern	.094	.017	.100	5.448	.000
	TrustGov	.181	.019	.195	9.305	.000
	TrustPrivate	.166	.020	.175	8.252	.000

a. Dependent Variable: AttitudeEAI policing SocilMedia; R square =0.509; df =8

This model explains about 51% of the variation in the data. First, for statistically significant positive correlates of attitude toward emotional AI use for social media policing, we find safety utility the strongest correlate ($\beta_{SafetyUtility} = 0.233$***), self-rated knowledge being the second strongest ($\beta_{Knowledge} = 0.230$***), then trust in government's regulation ($\beta_{TrustGov} = 0.195$***), and trust in the private sector's regulation ($\beta_{TrustPrivate} = 0.175^{***}$). These results are quite similar to the case of emotional AI security camera. This means individuals who believe Emotional AI policing on social media will be beneficial overall would also agree with the increased safety, consider themselves as having a basic understanding of the technology, and trust the government and private actors' ability to regulate the technology.

Second, for statistically significant negative correlates of attitude toward EAI policing social media platforms, the more people express concern that EAI policing social media will bring about a dystopian surveillance state ($\beta_{DystopianSurveilConcern} = -0.097^{***}$), the less likely they believe it will be beneficial for society.

Similar to the results in the case of emotional AI security cameras, we also find two seemingly paradoxical results: people with higher concerns about the accuracy of the technology ($\beta_{AccuracyConcern} = 0.100^{***}$), and autonomy loss ($\beta_{AutonomyLoss} = .079^{***}$), i.e., the freedom to protests online and offline, agree that the technology will be beneficial for society overall. The concern about biases embedded in the technology is not a statistically significant factor. Here, we can interpret the result as concern about the accuracy of the technology and an acceptance of the fact any new technologies are inevitably flawed and lack in accuracy, and the utility of increased safety outweighs the cost of such flaws. Next, using the Moral Foundation Theory, the result on the concern about autonomy loss suggest that such harm might not a major factor in the society stiff in its long-lasting tradition of patriarchy and cultural belief in homogeneity (Woo, 2022). In such a culture, an inherent strong sense of collectivism may attenuate citizens' concerns about autonomy issues.

8.2.4. Socio-demographic factors

Table 10.6. Regression results for socio-demographic factors in attitude toward EAI social media policing.

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	3.289	.109		30.070	.000
	Age	-.008	.002	-.127	-5.096	.000
	Income	.063	.033	.050	1.945	.052
	Education	.021	.013	.043	1.646	.100

a. Dependent Variable: AttitudeEAI police SocMedia

First, age is a *negative* significant predictor of attitude toward the use of Emotional AI policing social media (β_{age}= - 0.127, $p<0.001$). Second, income is a *positive* predictor of attitude toward the EAI security camera (β_{Income}=0.50), although the p-value test yields a slightly weak result ($p=0.052$). Education has a weaker statistically significant relationship with attitude toward EAI policing social media platforms ($\beta_{Education}$=0.43; $p=0.100$). Thus, people with higher education qualifications and higher incomes are more likely to agree that EAI use for policing social media platforms would benefit society. In contrast, older people are less inclined to such an agreement, similar to the case of the workplace and the case of EAI security camera policing.

Applying the TAM framework, we can interpret the results on income and education as people with higher education qualifications and higher income are likely to perceive the technology as easier to use and perceive more utilities in the technology. Earlier studies on technological adoption have pointed to the fact that people with higher social statuses are more likely to be early adopters based on the simple fact that they can afford the emerging affect-sensing technologies and are thus, more inclined to educate themselves about their benefits as well as possible risks (Blanden & Gregg, 2004; J. Lee et al., 2019). McClure (2017)'s study of AI technophobia among the US population also reveals that people from non-dominant social classes such as the lower income group, non-white groups, or females are far more likely to be threatened by new technologies.

Regarding the age variable, the TAM framework implies that elderly people tend to see new technologies as less relevant to their lives as well as more unfamiliar.

8.3. Chapter summary

This chapter has provided a systematic examination of various determinants of emotional AI in two security applications: security cameras and social media policing. We find that the Japanese population, on average, has a slightly positive outlook concerning these applications. In both cases, the utility of improved safety is the strongest correlate with the attitude toward the technology. Meanwhile, there is an acceptance of the technology despite the lack of accuracy in the emerging technology as well as its worrying implications for issues of social equality/fairness and autonomy. We have argued that such seemingly paradoxical results are produced by a sense of inevitability in the adoption of AI-enabled security measures as well as a Japanese culture that strongly emphasizes homogeneity and hierarchy.

Chapter 9: Political campaigning and synthetic media

In this chapter, the use of emotional AI during political elections and emotional AI uses in synthetic media for political campaigns are studied. We will first look at the perception of Japanese people regarding the use emotional AI to create political adverts and messages during the election campaigns. Then next, we will look at a more advanced use of emotional AI, i.e., the creation of synthetic media in political campaigns.

It is worth noting that studies that focus on the use of AI in political campaigns and elections in Japan have been few and far between. Thus, this study is the first in quantitatively studying Japanese perceptions of emotional AI in politics.

9.1. Social media and elections

Below is the explanation given to the survey respondents before they were asked to give answers on how much they agree or disagree with a series of Likert-scale questions concerning the use of emotional AI in political advertising in social media.

> "Companies working on behalf of political parties and groups seeking to campaign on a political issue use social media advertising services to find out which political adverts and messages are most emotionally engaging for specific audiences, as well as to personalise and micro-target the type of political adverts we see on social media."

9.1.1. Descriptive statistics and sex differences

Examining the distribution of attitude toward emotional AI in political elections, we find while the majority of the respondents are neutral on the topic (53.2%), there are slightly more people who are negative about such use of emotional AI compared to those who are positive (25.1% versu 21.8%, respectively). Here, the average score of the attitude toward emotional AI in political elections for the Japanese population is 2.94, with a standard deviation of 0.91. The descriptive statistics suggest that Japanese people are slightly against the use of emotional AI in political elections.

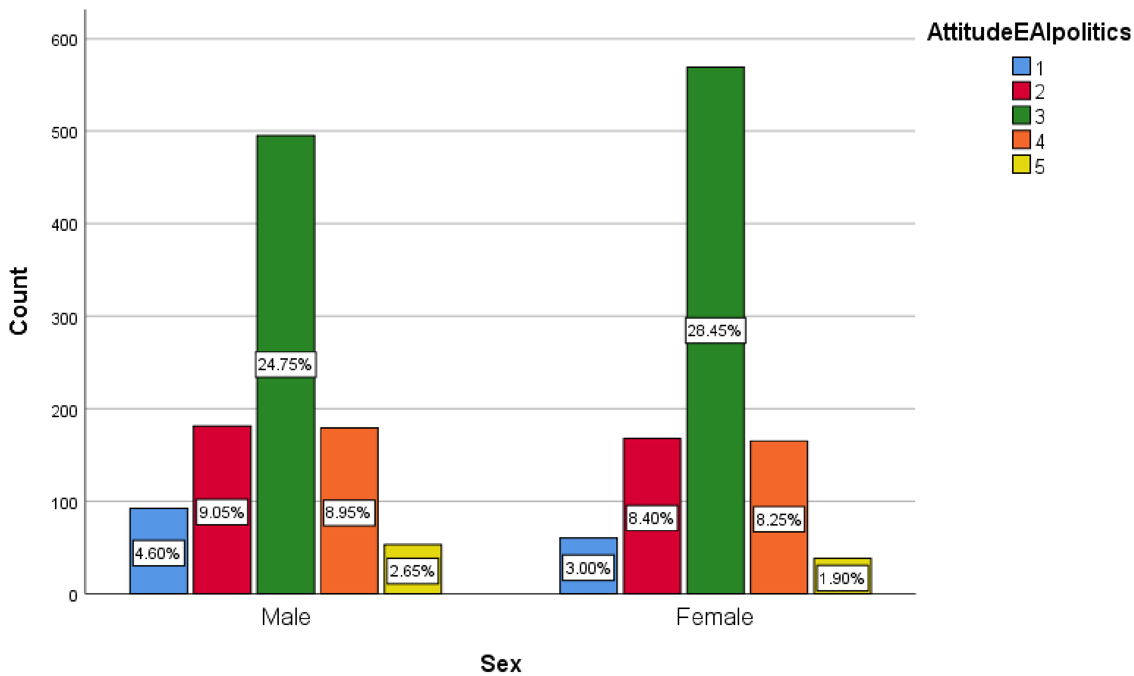


Figure 11.1. Distribution of attitude toward emotional AI applications in political elections by sex. Range of answer: 1 means strongly disagree to 5 means strongly agree.

Table 11.1. Descriptive statistics by sex of all the surveyed variables on the topic of emotional AI applications in elections.

Sex		DemocracyEngage	AidPoliciesMaking	StokingFear	EmoVotingConcern	AccuracyConcern
Male	Mean	3.19	3.12	3.43	3.48	3.48
	N	1000	1000	1000	1000	1000
	Std. Deviation	1.025	.958	.929	.963	.963
Female	Mean	3.26	3.23	3.50	3.52	3.52

	N	1000	1000	1000	1000	1000
	Std. Deviation	.877	.831	.854	.843	.843
Total	Mean	3.22	3.18	3.46	3.50	3.50
	N	2000	2000	2000	2000	2000
	Std. Deviation	.954	.898	.893	.905	.905

Sex		Safety Decline	Knowledge	AttitudeEAI politics	TrustGov	Trust Private
Male	Mean	3.39	3.00	2.92	2.78	2.87
	N	1000	1000	1000	1000	1000
	Std. Deviation	.914	.908	.967	1.042	.999
Female	Mean	3.41	2.95	2.95	2.81	2.91
	N	1000	1000	1000	1000	1000
	Std. Deviation	.835	.872	.851	.939	.885
Total	Mean	3.40	2.97	2.94	2.80	2.89
	N	2000	2000	2000	2000	2000
	Std. Deviation	.875	.890	.910	.992	.944

Range: 1 (strongly disagree) to 5 (strongly agree)

Using the Chi-square test, we find statistically significant differences between the sexes in the following variables: DemocracyEngage; i.e., the utility of EAI in engaging people with democratic processes ($p<0.001$); Aid policies making, i.e., the utility of EAI in helping more effective policy-making ($p<0.001$); Stoking Fear, i.e., the concern that EAI use in politics can create undue influence by stoking social fears and tensions ($p=0.19$); EmoVotingConcern,i.e., the concern that EAI use in political campaign can encourage emotional voting behaviors ($p=0.048$); AccuracyConcern, i.e., the concern about accuracy of EAI technology ($p=0.001$); AttitudeEAIpolitics, i.e., attitude toward the use of EAI in political campaign ($p=0.004$); TrustGov, i.e., the trust toward government's ability to regulate such use of EAI in politics ($p=0.011$); TrustPrivate, i.e., the trust toward the private sector's ability to regulate the use of EAI in politics ($p=0.008$).

Here, we find that women are more slightly more concerned about EAI applications in politics might be inaccurate, and they can lead to decline in individual safety, or create undue influence by stoking social fears and tensions or encourage emotional voting behaviors. Under the Moral Foundations Theory, the above findings make sense since it has been found that women are more sensitive toward the moral dimensions of Harm, Fairness, and Purity (Graham et al., 2011). Clearly, as harms and violations of moral norms around fairness and purity can occur if emotional AI applications use in elections are not accurate or being used to stoke social fears and tensions.

Interestingly, despite holding these concerns more than men, women are slightly more positive about the attitude toward EAI use in politics, and more inclined to think EAI can help engage more people with democratic political processes and help better policymaking. Women are also exhibit slightly higher trust in the government and the private sector's ability to regulate such use of EAI technology in politics.

9.1.2. Socio-demographic factors

Table 11.2. Regression results for socio-demographic factors regarding emotional AI used in political elections.

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	3.450	.110		31.487	.000
	Age	-.012	.002	-.176	-7.082	.000
	Income	.035	.033	.028	1.076	.282
	Education	-.011	.013	-.023	-.874	.382

a. Dependent Variable: AttitudeEAIpolitics; R square = 0.031

Running a linear regression model for socio-demographic factors, we find only age exhibits a negative statistically significant relationship with attitude toward EAI in politics (β_{Age} = -

0.176***). Under the TAM framework, this result can be interpreted as Japanese elderly people are averse to the technology because they see little apparent utilities as well as familiarity.

9.1.3. Values and Concerns

Table 11.3. Regression results for behavioral determinants of attitude toward EAI in political elections.

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.382	.088		4.323	.000
	DemocracyEngage	.136	.024	.143	5.651	.000
	AidPoliciesMaking	.145	.026	.143	5.594	.000
	StokingFear	.011	.023	.011	.485	.628
	EmoVotingConcern	.021	.023	.020	.890	.374
	AccuracyConcern	-.005	.024	-.005	-.229	.819
	SafetyDecline	-.052	.024	-.050	-2.104	.036
	Knowledge	.244	.020	.239	12.329	.000
	TrustPrivate	.156	.023	.161	6.758	.000
	TrustGov	.201	.022	.219	9.169	.000

a. Dependent Variable: AttitudeEAIpolitics; R square= 0.527

This model explains 52.7% of the variation in the data, which is within the ballpark of other studies which has utilized the extended TAM framework (Lew et al., 2020). The regression analysis produces three sets of interesting results. First, aligned with predictions from the TAM framework, positive correlates of attitude toward emotional AI in political campaign include more democratic engagement ($\beta_{DemocracyEngage}$ = 0.143***); better policymaking ($\beta_{AidPoliciesMaking}$=0.143***); self-rated knowledge of the technology ($\beta_{Knowledge}$= 0.239***). In other words, perceiving more utilities in and familiarity with a technology lead to more acceptance of the technology.

Second, on the issues of regulation, we find that trust in the government's regulation ($\beta_{TrustGov}$= 0.161***); having trust in the private sectors to regulate the technology ($\beta_{TrustPrivate}$= 0.219***). Here, the results resonate with findings of Vu and Lim (2021) that techno-social environment and governmental effectiveness positively influence individual acceptance of AI/Robots technology.

Third, in terms of negative correlates, people who agree that emotional AI in politics would lead to a decline in individual safety are more likely to disagree that such use of the technology will be beneficial for society (β_{Safety}= - 0.05*). This result can be interpreted within both the Moral Foundation theory and the TAM. Here, those who believe that emotional AI use in politics leads to a decline in individual safety (i.e., a violation of the Harm dimension) are likely to reject the technology. In the context of Japan, this correlation makes even more sense as we have seen in the previous chapter on security camera and policing social media, concern for safety is the strongest determinant of acceptance of emotional AI.

Finally, interestingly, concerns for the accuracy of the technology, and the ability of EAI to stoke social fears and tensions, as well as encourage emotional voting behaviors bear no statistically significant relationship with the attitude toward EAI in politics.

9.2. Synthetic media and political campaigning

Below is the explanation given to the survey respondents before they were asked to give answers on how much they agree or disagree with a series of Likert-scale questions concerning the use of emotional AI in political advertising in social media.

> "Artificial intelligence tools can generate realistic audio-visual duplicates of people doing or saying things that they never actually did or said (so called 'deepfakes'). These can be used to deliver emotionally powerful but false messages, attributed to politicians, leaders and celebrities.

Computer programs can also seem like real human users ('bots') and can be used to amplify messages on social media, often in favour of, or against a political or social issue."

9.2.1. Descriptive statistics

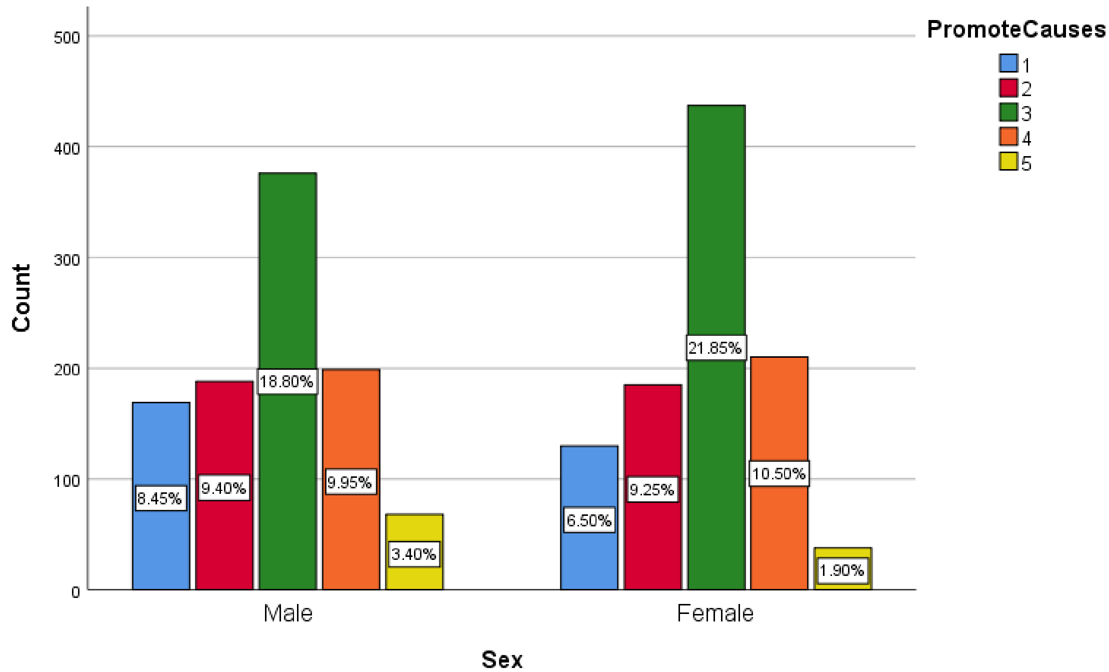


Figure 11.2. Distribution of attitude toward synthetic digital media in political campaigns be sex.

Regarding whether a respondent would be comfortable with the use of synthetic media to promote social and political causes, 33.6% report feeling negative about such use (strongly disagree or tend to disagree), while 40.7% report feeling neutral, around 29% report feeling positive about the use of synthetic media in promoting social and political causes. Compared to the case of emotional AI use in elections, it is clear that there is a strong negative shift in the attitude toward synthetic media.

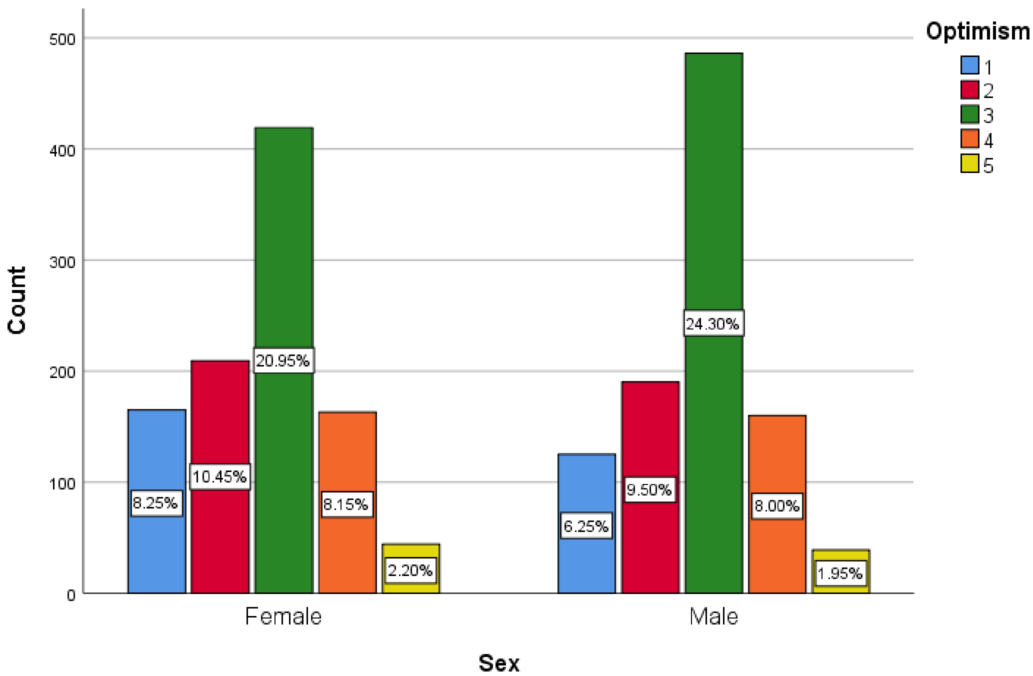


Figure 11.3. Distribution of optimism toward synthetic digital media in political campaigns be sex.

Regarding whether a respondent feeling optimistic that EAI synthetic media would benefit society, compared to the previous statement, there are slightly more people who express a neutral stand and less people who express a positive stand. There are 34.4% of the respondents report a sense of pessimism, 45% of the respondents are neutral, while only 20% are positive that synthetic media will benefit society.

It is also notable that among all other contexts, synthetic media is the context that receives the lowest mean score on whether emotional AI application would benefit society. In other words, the Japanese population expresses the most negative reaction against EAI use in synthetic media for political campaigning (mean = 2.83, std = 1.082).

9.3.1. Sex differences

Table 11.4. Sex differences in attitude toward and concerns about synthetic media in political campaigns.

Sex		PromoteCauses	TruthConcern	VerifyPopular Concern	Optimism
Male	Mean	2.81	3.64	3.58	2.71
	N	1000	1000	1000	1000
	Std. Deviation	1.140	.979	.970	1.061
Female	Mean	2.84	3.74	3.71	2.80
	N	1000	1000	1000	1000
	Std. Deviation	1.021	.901	.890	.983
Total	Mean	2.83	3.69	3.65	2.75
	N	2000	2000	2000	2000
	Std. Deviation	1.082	.942	.932	1.023

Sex		AutonomyLoss	Knowledge	TrustGov	TrustPrivate
Male	Mean	3.58	2.92	2.75	2.78
	N	1000	1000	1000	1000
	Std. Deviation	.972	.977	1.022	1.015
Female	Mean	3.62	2.84	2.77	2.78
	N	1000	1000	1000	1000
	Std. Deviation	.886	.912	.941	.925
Total	Mean	3.60	2.88	2.76	2.78
	N	2000	2000	2000	2000
	Std. Deviation	.930	.946	.982	.971

Applying the Chi-square test, we find statistically significant sex differences in the following variables: whether a person is comfortable with the use of synthetic media for social and political causes (PromoteCauses, $p = 0.001$); the concern that it is difficult to verify whether a message is real with synthetic media (TruthConcern, $p = 0.032$); the concern that it is difficult to know if a message or a person is truly popular for artificially amplified (VerifyPopularConcern, $p = 0.001$); the optimism that synthetic media would benefit society (Optimism, $p = 0.02$); the concern that synthetic media would have undue influence over a person's thinking and feelings about a political/social issue (AutonomyLoss, $p = 0.011$); self-rated knowledge regarding the technology (Knowledge, $p = 0.008$); trust in the government ability to regulate the technology (TrustGov, $p = 0.023$); trust in the private sector ability to self-regulate the technology (TrustPrivate, $p = 0.004$).

To be specific, men are found to be more *uncomfortable* with the use of synthetic media to promote social and political causes, as well as are *more pessimistic* about the benefit such a technology will bring to society. Of all ten use cases, this is the only cases where we find men are less accepting of a new technology than women.

Men are found to have less trust in the government and the private sector to regulate the technology. Meanwhile, women are found to have slightly more concerns regarding the difficulty of verifying truth and falsehood with synthetic media and have more concerns regarding the difficulty of verifying whether a message is actually popular or artificially amplified, as well as the loss of autonomy in thinking and feeling about political issues. Here, it is also useful to invoke the sex differences in concerns about different moral foundations, where researchers have found that women care more about the three dimensions of Harm/Care, Fairness, and Purity, while men are found to care more about the dimensions of Loyalty and Authority (Atari et al., 2020).

9.4.1. Regression analysis

Table 11.5. Socio-demographic factors regarding attitude toward synthetic media.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.259	.123		26.422	.000
	Income	.058	.037	.041	1.590	.112
	Age	-.013	.002	-.173	-6.949	.000
	Education	-.010	.015	-.018	-.692	.489

We only find a statistically significant result for age, which is a negative correlate of optimism toward the use of synthetic media in political campaigning (β_{Age} = - 0.173***).

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.510	.099		5.154	.000
	PromoteCauses	.324	.020	.343	16.592	.000
	TruthConcern	-.125	.027	-.115	-4.626	.000
	VerifyPopularConcern	.044	.030	.040	1.468	.142
	AutonomyLoss	.027	.025	.025	1.070	.285
	Knowledge	.159	.020	.147	7.807	.000
	TrustGov	.183	.026	.176	6.958	.000
	TrustPrivate	.205	.026	.195	7.903	.000

a. Dependent Variable: Optimism; R square = 0.532

This model explains 53.2% of the variation in the data. Positive correlates of optimism toward synthetic media uses in political campaign include being comfortable with the technology use for promoting social and political causes ($\beta_{PromoteCauses}$ = 0.343***); self-rated knowledge of the technology ($\beta_{Knowledge}$= 0.147***); having trust in the government's regulation ($\beta_{TrustGov}$= 0.176***); having trust in the private sectors to regulate the technology ($\beta_{TrustPrivate}$= 0.195***). The findings on the utility of synthetic media for promoting social and political causes and self-rated knowledge of the technology are aligned with the original TAM model. The findings on the issue of trust in regulation of the government and the private sector also converge with the literature on extended TAM, that higher level of techno-social environment and higher government effectiveness are often linked with a positive attitude toward smart technology (Vu & Lim, 2021).

In terms of negative correlates, people who agree that it would be difficult to verify the truth of political messages with synthetic media are likely to disagree that such use of the technology will be beneficial for society ($\beta_{TruthConcern}$= - 0.115***). Here, the Moral Foundation Theory helps us make sense of the results. This concern about that synthetic media would lead to a difficulty to verify the truth of a political message is a concern about a violation of dimension of Harm and Purity. However, concerns for the ability verify if a political issue is really popular, and the loss in independent thinking bear no statistically significant relationship with the attitude toward EAI-synthetic media.

9.3. Chapter summary

This chapter has been devoted to the use of emotional AI in politics. Our analysis reveals that with regards to political elections and to synthetic media, the Japanese people have, for the most parts, a skeptical view on whether the use of emotional AI in politics is beneficial for society. Moreover, people have a strongest negative reaction to the use of emotional AI in creating synthetic media. Importantly, this is the only case where we find that men are more worried about the use of the emerging emotional AI technology in politics. Regarding the connections with theories, the TAM again proves to be useful as the variables of perceived utility of and perceived familiarity with emotional AI are positive predictors of attitude toward the technology.

The findings also confirm the literature on sex differences on Moral Foundations, we find that women express more concerns related to harmful implications of emotional AI in politics: the decline in safety, the stoking of fears and tensions, the loss of autonomy in thinking and voting. In the case of politics, women expressed more concerns about whether emotional AI is inaccurate, and whether it might lead to a decline in individual safety, or whether it might create undue influence by stoking social fears and tensions, or it might encourage emotional voting behaviors.

For the case of synthetic media, women are found to have slightly more concerns regarding the difficulty of verifying the truth a political message, the difficulty of verifying whether a message is actually popular or artificially amplified, as well as the loss of autonomy in thinking and feeling about political issues.

When we investigate the determinants of emotional AI applications, we find the results agree with the TAM that perceived utilities of emotional AI, (i.e., increasing democratic engagement, aiding policymaking, spreading social awareness) and self-rated familiarity with the technology predict its acceptance. Moreover, similar to previous chapters, trust in the regulatory framework, whether provided by the government or the private sector, is also a key determinant in the acceptance of the technology.

Chapter 10: Contributions, limitations, and future research directions

In this chapter, a summary of key findings and contributions of this thesis will be provided. More importantly, we will reflect on the limitations of the research design, thus, mapping out the future research directions.

10.1. Theoretical and empirical contributions

Theoretically, utilizing the intuition that an acceptance of a new technology is a function of not only its perceived utilities and ease of use (formalized in the Technological Acceptance Model), but also its perceived validation/rejection of deep-seated socio-cultural norms and values (formalized in the Moral Foundation theory), this study is the first in the literature to bringing the TAM and the Moral Foundation Theory together on the topic of acceptance toward emotional AI in Japan. Under the analytical Three-pronged Approach (Contexts, Variables, and Statistical models), this study has successfully demonstrated the utility of bringing the TAM and the Moral Foundation Theory to analyze determinants of attitude toward the emerging emotional AI technology.

The statistical models in this study have successfully accounted for an average of 52.11% the variation in the data, which is within the same ballpark of the empirical studies on technological acceptance of the smart technologies. For example, a study found the extended model accounted for 61% of the variance in the behavioral intention (BI) to adopt mobile wallet technology (Lew et al., 2020). Another meta-analyses of digital technology adoption in education show the TAM models can account for up to 44% of variance in the BI (Scherer et al., 2019). As for the most successful case of statistical modeling in this study, the case of Home Robots, our model accounts for 67.8% of the variation in the data. The next two successful cases are security camera and workplace, each account for around 58% of the variation in the data. Thus, the empirical strategy in this study proves is promising in producing more interesting results.

Table 12.1. A summary of statistical modeling results.

Purposes	**Use cases**	**R square**
Political messaging	Social media and elections	52.70%
	Synthetic Media	53.20%
Security	Social media policing	50.90%
	Security Camera	58.20%
Education/ Children development	Toys	38%
	Education	44.50%
Private	Car	58.30%
	Home Robots	67.80%
Health care	Mental health screening	42.10%
Workplace	Workplace	55.40%
Min = 38% (Toys); Max = 67.8% (Home Robots); Mean = 52.11%		

During this study, we have identified results that are aligned with predictions of the TAM or the Moral Foundation Theory. For example, first, agreeing with the TAM framework, we consistently find the citizens' self-rated knowledge of emotional AI applications and its perceived utilities (e.g., improved safety, decreased loneliness, etc.) consistently predict a positive attitude toward the emerging technology. Second, agreeing with the empirical findings related to the Moral Foundation Theory, across all use cases, we consistently find women express more concerns for key values that might be threatened by the adoption of emotional AI: autonomy, fairness/equality, privacy, etc.

Moreover, it provides the first systematic and comprehensive analysis of socio-demographic and behavioral determinants of attitude toward emotional AI applications in 10 different settings: for private uses (i.e., cars, home robots, etc.); for healthcare (i.e., diagnosis of mental illnesses and stresses); for education and children development (i.e., in school and children toys); for political campaigns (with synthetic media, and micro-targeted political adverts); for security (with emotion-sensing security camera, and emotional AI that surveils social media platforms).

Consequently, it has provided a set of rich empirical findings to the question of context sensitivity of social perception toward emotional AI, addressing research question 3: "How do determinants of attitude toward emotional AI applications vary according to the domains of applications for example healthcare, education, security, politics, workplaces, etc.?" Table 10.2 presents all the decisions about each hypothesis proposed in the beginning of the thesis.

Table 12.2. A summary of hypotheses and relevant literature examined in this study.

No.	Hypotheses	Literature/ Theories	Decision	Research questions
1	H1: **Being male** is positively correlated with attitude toward emotional AI. While the opposite is true for **female**.	Empirical findings on attitude toward AI applications (Ali, 2012; McClure, 2017; Urueña et al., 2018)/ Sex differences in Moral Foundation Theory (Atari et al., 2020; Graham et al., 2011; Hidalgo et al., 2021)	*Supported:* Workplace *Rejected:* Toys; Healthcare; Schools; Cars; Political campaigns; Synthetic Digital media; Security Camera; Social Media Policing; Politics	RQ1
2	H2: **Female** express more concerns about emotional AI's implications for moral harms such as privacy violation, autonomy loss, biased algorithms.	Sex differences in Moral Foundation Theory (Atari et al., 2020; Graham et al., 2011)	Supported in all cases.	RQ1/ RQ2
3	H3**: Income** is positively correlated with the attitude toward emotional AI.	Empirical findings on attitude toward AI applications (Ali, 2012; Chen & Lee, 2019; McClure, 2017; Urueña et al., 2018)	Supported: Social media policing. Rejected: Workplace; Toys, Education, Healthcare; Home Robots	RQ1
4	H4: **Age** is negatively correlated with attitude toward emotional AI.	Empirical findings on attitude toward AI applications (Ali, 2012; Chen &	Supported in all cases except	RQ1

		Lee, 2019; McClure, 2017; Urueña et al., 2018)	healthcare; Home Robots	
5	H5: **Higher educational qualification** positively correlated with attitude toward emotional AI.	Empirical findings on attitude toward AI applications (Ali, 2012; Chen & Lee, 2019; McClure, 2017; Urueña et al., 2018)	*Supported*: Cars; workplace (weak significance); Social media policing (weak significance) *Rejected*: Toys; healthcare; Home Robots; Education; Cars; Political campaigns; Synthetic Digital media.	RQ1
5	H5: **Perceived utilities** of emotional AI technologies positively correlate with attitude toward them	Predictions from Technological Acceptance Model (Alina & Khalina, 2021; Davis, 1989; Kamal et al., 2020; Taherdoost, 2018)	Supported in all cases.	RQ2
6	H6: **Self-rated knowledge** with emotional AI technologies is positively correlated with attitude toward the emerging technologies.	Predictions from Technological Acceptance Model's (Alina & Khalina, 2021; Davis, 1989; Kamal et al., 2020; Taherdoost, 2018)	Supported in all cases.	RQ2
7-1	H7-1: Concern about emotional AI's negative impacts on the moral value of **privacy** is negatively correlated with attitude toward emotional AI technologies.	Predictions from Moral Foundation Theory as adapted in the book How humans judge machines (Atari et al., 2020; Graham et al., 2011; Hidalgo et al., 2021)	*Supported:* Workplace; Cars; Education; Toys; Security Camera; Social media Policing. *Rejected:* Home robots;	RQ2
7-2	H7-2: Concern about emotional AI's negative impacts on **autonomy** is negatively correlated with attitude toward emotional AI technologies.	Predictions from Moral Foundation Theory as adapted in the book How humans judge machines (Atari et al., 2020; Graham et al., 2011; Hidalgo et al., 2021)	*Supported*: Workplace. *Rejected*: Home robots; Social Media Policing; Political elections; Synthetic Media.	RQ2
7-3	H7-2: Concern about emotional AI's negative impacts on **fairness** is negatively correlated with attitude toward emotional AI technologies.	Predictions from Moral Foundation Theory as adapted in the book How humans judge machines (Atari et al., 2020; Graham et al., 2011; Hidalgo et al., 2021)	*Supported:* Healthcare *Rejected:* Workplace; Education; Toys; Security Camera; Political elections	RQ2

8	H8: Concern about **accuracy of the technology** is negatively correlated with attitude toward emotional AI technologies.	Predictions from Moral Foundation Theory as adapted in the book How humans judge machines (Atari et al., 2020; Graham et al., 2011; Hidalgo et al., 2021)	*Rejected*: cars, education, security camera, workplace; Social media policing; Political elections	RQ2
9	H9: **Transparency on data management** (how emotional data is managed, stored, processed) positively correlated with attitude toward emotional AI. The opposite is true when no transparency is provided.	Qualitative research results from various use cases including cars (McStay & Urquhart, 2022), toys (McReynolds et al., 2017; McStay & Rosner, 2021), data management (McStay, 2020b), education (McStay, 2020a), smart homes , security (Urquhart & Miranda, 2022); workplace (Mantello et al., 2021; Urquhart, Laffer, et al., 2022), etc.	Rejected: Toys; education; home robots; healthcare Supported: Workplace; Cars.	RQ2
10	H9: **Trust toward the government's ability to regulate** the technology is positively correlated with attitude toward emotional AI technologies.	Empirical findings from attitude toward AI/Robots and government effectiveness index (Vu & Lim, 2021)	Supported in all cases	RQ2
11	H10: **Trust toward the private sector's ability to regulate** the technology is positively correlated with attitude toward emotional AI technologies.	Empirical findings from attitude toward AI/Robots and techno-social environment (Vu & Lim, 2021)	Supported in all cases	RQ2
12	H11 (**The context sensitivity hypothesis**): Determinants of attitude toward emotional AI varied in according to different contexts.	Qualitative research results from various use cases including cars (McStay & Urquhart, 2022), toys (McReynolds et al., 2017; McStay & Rosner, 2021), data management (McStay, 2020b), education (McStay, 2020a), smart homes , security (Urquhart & Miranda, 2022); workplace (Mantello et al., 2021; Urquhart, Laffer, et al., 2022), etc.	Supported per the decisions shown above.	RQ3

Notably, to answer the first research question: "How do socio-demographic factors influence the acceptance of emotional AI?", this study has found that in most cases, the statistically significant negative correlate is age. In other words, the older someone is, the less likely they are willing to accept the technology.

Moreover, agreeing with the empirical findings in the literature on sex differences in Moral Foundation Theory (Atari et al., 2020), we find that women are more concerned than men in many key ethical issues related to emotional AI: privacy violation, exacerbation of social biases, data management, and misuse, the loss of autonomy, etc.

Only in cases of toy, do we find a negative significant association between education with the attitude toward emotional AI, which has been interpreted as the growing awareness among parents and people with higher educations about adverse effects of overreliance on smart technologies among children, adolescent, as well as young adult. Clearly, the scientific community has started to warn the public about the worrying trends in mental health among Gen Z, the population who has grown up exclusively with social media platforms.

These findings will prove crucial in creating socially aware policies that support the ethical use of emotional AI technologies to combat detrimental effects of a rapidly aging population: the increase of social isolation, lack of social interactions, labor shortage in the workforce and education, etc. We will touch on these policy recommendations in the following chapters.

10.2. Limitations of the positivist approach

This study, by employing multiple sources of data, from national survey to interviews with emotional AI stakeholders to workshops with citizens, has provided a comprehensive and systematic investigation of social and ethical perceptions of emotional AI in various use cases. However, the results should not be overgeneralized as it is focused almost exclusively on the case of Japan. Cautions are needed when interpreting the correlations found in this study. For establishing causal relations, future studies can utilize longitudinal surveys as well as experiments.

More importantly, the current study has strictly taken on a positivist approach, in that, the focus is to describe truthfully the state of the world via statistical analyses and qualitative thematic analysis. Since it is vitally important to identify ways to coexist ethically and well with emotion-sensing technologies that are getting smarter and more interactive, future studies can expand on the empirical findings and take on the normative, constructivist approach. A prominent recent example is the deployment of the critical Marxist Frankfurt tradition of social critiques by Hanemaayer (2022). For example, Ariane Hanemaayer's edited book "Artificial Intelligence and its discontents: Critiques from social sciences and humanities" (2022). In this edited volume, Hanemaayer et al. take on the Frankfurt school's perspectives to question who are discontented by AI and how they resist the technology. Resonating Karl Marx's famous 11th thesis of Feuerbach, Hanemaayer argues the goal of AI criticism as a subfield in social sciences and humanities is beyond a mere interpretation of the technology, but to concretely transform it, i.e., to effect changes in "its infusion, investment, and implementation" (p.8). They show the contemporary narrative of AI development is a narrative of overcoming, centering around the unrelenting march of progress in the field of AI research despite multiple setbacks in AI winters. Thus, arguing that the narrative of AI must not be left to only the programmers. The authors argue their critiques on AI technologies, like the critiques of the Frankfurt school toward different aspects of culture and industry, serve no purpose to the advent and progress in AI development. Instead, the critiques contribute to the foundations of understanding the technologies within a larger social and practical world: the nature of AI systems themselves, their cultural symbols, and representations of AI in the social world, and their impacts.

10.3. Limitations of the linear human-machine relationship presumed by the TAM and Moral Foundation Theory

Across all cases, we consistently encounter results that are seemingly paradoxical that contradicting the straightforward implications of the TAM (i.e., the attitude toward a new technology is a function of its perceived utilities and ease of use) and the Moral Foundation Theory (i.e., the attitude toward a new technology is a function of its validation of and threat toward ethical values and norms). This finding implies it is necessary to develop theoretical frameworks that capture cross-cultural differences in moral reasoning about technological effects on our daily lives.

For example, in chapter 4, considering the workplace, the concern about biases against disadvantaged group in emotional AI systems and the concern for inaccurate emotional AI have no statistically significant association with acceptance of emotional AI in the workplace. In chapter 6, concern about data management process (i.e., how the emotional data are collected, stored and who get access to it) and concern about EAI applications are considered better than humans at emotion recognition (i.e., threat to replace human caregivers) are two positive correlates. In chapter 7, regarding emotional in school, concerns for privacy violation, autonomy loss, human replacement, changing expectations of human connection are found to be non-significant in predicting the acceptance of emotional AI. Regarding emotional AI toys, concern about data management and concern about embedded biases are positive correlates with attitude toward EAI toys and while accuracy concern is not a statistically significant predictor. In chapter 8, paradoxically, concern about the accuracy of the technology and concern about biases/unfairness toward disadvantaged groups are also positive correlates of attitude toward Emotional AI Security Camera, which means respondents who agree that EAI security cameras will benefit society are equally aware of the downsides of the new technology: lack of accuracy or biases toward disadvantaged groups.

Thus, the empirical evidence from this study and the literature (Alsaleh et al., 2019; Hope & Jones, 2014; Mantello et al., 2021; Psychoula et al., 2018), shows within the framework of Davis (1989)'s Technological Acceptance Model (TAM), it is difficult to account for how cultural, environmental factors (such as politics, regions or religions) underly (or not) other predictive factors (such as gender, educational level, income, etc.), implying the need for the expansion of the TAM model.

Moreover, although we have complemented the TAM with the Moral Foundation Theory, this thesis's modeling and conceptualizing of our relationship with technology is still linear. As Table 10.1 has demonstrated, there are certain findings that directly contradict the Moral Foundation Theory. For example, the hypotheses that concern about the loss of autonomy as well as concern about biases in the algorithms are *negative*ly correlated with attitude toward emotional AI applications, which are predicted by the Moral Foundation Theory, are rejected in the case of home robots, social media policing, or security camera. *These findings show there exists a hierarchy of values, which needs to be captured better by future theoretical models.*

We have assumed our relationship with machines as a linear function, in which, acceptance is higher with utilities and lower when the technology implicates a loss in social norms and values, we deeply care about such as privacy, autonomy, and fairness. However, given that emotional AI technologies are not static, and being increasingly deployed as a ubiquitous, ambient factor running in the background in personal devices or in public spaces, seeking to modify our behaviors, gauging technological acceptance of emotional AI is far more complex than Davis's notion of behavioral intention in technology or the linear presumption of Hidalgo et al. (2021)'s adaptation of the Moral Foundation Theory in 'How humans judge machines'. Arguably, we are better off thinking about our relationship with a new technology as a value-filtering process. Below is an example.

A technology such as emotional AI represents a certain set of values, whether on an individual level or in a collective level. For example, in the context of the workplace, it is about making emotions more transparent and personalized tracking for enhancement of mindfulness and optimizing motivation. This can be a problem as it comes into contact with the traditional culture of a Japanese workplace, where it is more oriented toward teamwork and thinking about other people. Employees are expected to hide their true emotions, especially the negative ones, and being ambiguous in their expression. Recalled in chapter 4 on the workplace, a CEO of a Japanese data company commenting on the fact that Japanese companies, no matter how global they are, want to stay Japanese. Here, for him the defining essence is that Japanese businesspeople are naturally more indirect, read more between the lines, and care more what other people think. And the classic book on contemporary Japanese culture, *The Japanese mind*, is opened with the concept of Aimai (曖昧) or Ambiguity defined as 'a state in which there is more than one intended meaning, resulting in obscurity, indistinctness, and uncertainty' (Osamu, 2002, p. 9). This cultural value is clearly at odds with the value of presumed by emotional AI technologies, i.e., making emotions more transparent so that the workplace can function better with less stress and more efficiency.

As our interviewees from Amazon Japan stated the introduction of AI systems to monitoring employees present a cultural conflict between the traditional, normative way Japanese society views employer-employee relationship versus the importance placed on efficiency, productivity, and optimization, even at the cost of the people by global companies such as Amazon. Clearly, there is an acculturation process that takes place, and the technology of emotional AI is not a mere tool, but conduit of values and new ways of working.

10.4. Future research directions: Mindsponge-based technological acceptance models

Thus, future studies should consider systematic differences generated by cultural mindsets by exploring information filtering mechanism postulated by Vuong and Napier (2015)'s mindsponge framework as a model of technological acceptance. Unlike TAM, the mindsponge framework considers the cost-and-benefit evaluation, i.e., the perceived usefulness and perceived ease of use in TAM, as not as the overriding factors in the filtering mechanism of the mind for a new input. The ease of use as well as usefulness, in the mindsponge framework, acts as trust evaluators of the filtering process, e.g., higher perceived usefulness or higher ease of use help increase the trust in a technology, but they are not the overriding factors in determining its acceptance as the traditional TAMs suggest. Whether the mind of a user rejects or accepts an input is also contingent on auxiliary factors such as an individual's ability to creatively adapt new inputs to their specific circumstances but also how individual core values and external settings (cultural and political) reinforce or diminish the uses of such inputs.

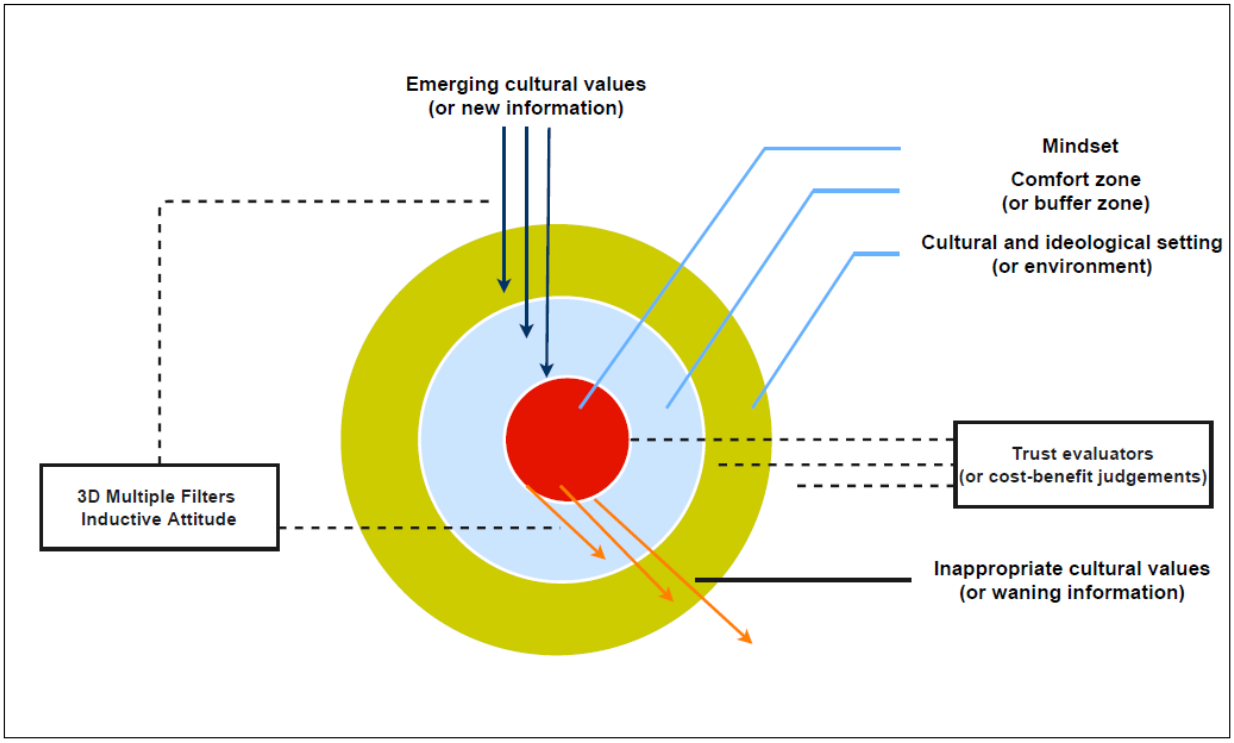


Figure 12.1. A visualization of the mindsponge model of information filtering. Copyright: CCBY 4.0

Hence, rather than adding new variables to the traditional TAM in a linear and somewhat arbitrary way (Kamal et al., 2020; Rajak & Shaw, 2021), the mindsponge framework combined with Bayesian multi-level modelling can offer a more systematic, hierarchical way of extending the TAM by differentiating between variables that come from an individual's core mindset, and the external cultural, ideological setting.

Thus, the mindsponge-based technological acceptance model can open new fertile grounds for future research and below are a few suggestions for future studies to consider. First, we incorporate various factors from the mindsponge model of information filtering to form other expanded TAMs. These factors include personal core values (i.e., level of openness to experiences, level of creativity, level of religiosity); environmental factors of culture (i.e., regions of home country) and politics (i.e., political regime of the home country). In terms of modeling techniques, these factors form a varying intercept for a Bayesian network model or can be used in structural equation modeling as latent variables (Fan et al., 2016).

The partial pooling techniques of Bayesian statistical analysis, these cultural and environmental factors can be used as varying intercepts in the model. This approach is called Bayesian multi-level or hierarchical Bayesian regression, and it has many advantages suited for the purposes of this study. First, the Bayesian statistical approach also allows us to directly compare the plausibility of our models via various indicators of weights. Second, since many more studies from the field of science and technology studies and social sciences rely on data from online survey, in this regard, Bayesian multi-level modeling has an advantage over traditional frequentist statistics. Online survey often means non-random and limited data can render local analyses carried in traditional statistical method impossible or unreliable. In the words of the world-renowned statistician, Sir David Spiegelhalter, the traditional frequentist statistics struggles with situation where data are non-random and limited such that local analyses cannot be performed reliably. The Bayesian multi-level regression or hierarchical modelling is the perfect response to this problem. Here, “the basic idea is to break down all possible respondents into small ‘cells’, each comprising of a highly homogeneous group of people-say living in the same area, with the same age, gender, …” (Spiegelhalter, 2019, p.329).

10.5. Chapter summary

In conclusion, chapter 10 has summarized the major contributions, theoretically and empirically, of this thesis and provided a series of reflections on its limitations. Crucially, to balance the positivist approach taken in this thesis, I have recommended future studies to draw on normative and constructivist approaches in social sciences and humanities to further explore, critique, and understand emotional AI applications within a larger social and practical world full of symbols, narratives, and power relations. We have also discussed the limitations of the linear assumption of human-machine relationship underlying the TAM and Moral Foundation Theory and recommend future studies can rectify these shortcomings with the mindsponge model of information filtering process proposed by Vuong and Napier (2015).

Chapter 11: Discussion and conclusion

Seven lessons on algorithmic governance, the privacy-autonomy paradox, and a Turing test for emotional AI

The final chapter provides concrete policy recommendations based on careful consideration of the empirical findings presented in this study. Then, the thesis will be concluded with two philosophical discussions of the privacy-autonomy paradox and a Turing test for emotional AI.

1) Diversity and inclusiveness in policy-making bodies are vitally important.

In analyzing our national survey, we find that women are on average expressing a more reserved attitude toward the emerging emotional AI technology in most of use cases. In particular, as previous chapters have shown women are found to be more concerned about the key issues regarding the impacts of emotional AI technology for our lives: data privacy, freedom of expression, autonomy loss, biases toward minorities and disadvantaged groups.

The worries are in many decision-making bodies in the tech and business world, women are underrepresented. It is a well-established fact that the tech industry as well as STEM education are currently being dominated by males, for instance, the European Institute for Gender Equality's 2020 report stated in the EU, only 2 in 10 ICT specialists are women (European Institute for Gender Equality, 2020). Moreover, emerging studies also suggest that males are less exposed to the dangers of algorithmic bias (N. T. Lee et al., 2019). In Japan, this situation is even worse due to the traditional male-dominant hierarchy that is still very commonplace in many companies both domestically and internationally. Japan ranked 120 out of 156 countries in World Economic Forum's Global Gender Gap Index Report in 2021, noteworthily, more than half of the working women in Japan occupy a non-regular role, which designate temporary, part-time or casual jobs that offer limited security, few benefits, low wages and low prestige. According to labor expert Wakana Shuto, even the former Prime Minister Shinzo Abe has made correcting the gender inequality problem in the workplace his priority, COVID-19 pandemic has erased most of the labor gains for women in recent years (Wakana, 2021). As for women in politics, the 2019 the Act on Promotion of Gender Equality in the Political Field (the Gender Parity Law), disappointingly, contains no enforcement mechanism, hence produced minimal effects. The 2021 Lower House election saw the number of women representatives drop to 45 — two fewer than in the 2017 election (Dalton, 2022). As the most of working professionals we interviewed have stated, it is acknowledged that women in Japanese society still suffer many unwritten, culturally bounded disadvantages compared to their men counterparts. This situation is worrying and needs correcting to ensure that the vision of inclusive smart society can be realized.

2) Public outreach and educational initiatives to promote the technology for various segments of society, especially the elderly population.

We consistently found the age is a negative correlate of acceptance toward the technology in all the cases. However, the elderly population not only in Japan but across the globe stands to gain many benefits from ethical and effective applications of emotional AI: improved mental healthcare diagnostics and treatment; road safety in cars via distraction, stress, and fatigue detection; decreased loneliness with the presence of companion robots; etc. More importantly, if the power of AI technologies are ethically harnessed, many societal issues that come with a rapidly aging society can be mitigated, for example, the shortage of labor forces in critical areas of society such as education, healthcare, and services. It is reported that Japan will soon experience a shortage of nearly half a million healthcare workers by 2025 (JIJI Press, 2022). In this context, the Basic Principles of the Asia Health and Wellbeing Initiative launched in 2016 was revised in 2018, which emphasizes the independence, participation and choices of the elderly. Here, in synergy with the vision of Society 5.0, it is clear that smart technologies can play an important role in helping Japan to realize these principles and supporting the aging population to achieve an independent, active, and free live.

Given the wide-ranging impacts of emotional AI in society and its deeply nuanced ethical consideration demonstrated in this thesis, curriculum of educational system would strongly benefit from inclusion of courses on social and ethical implications of AI. This is to correct the younger generation's misconceptions and enrich their understanding of the positive and negative potential of such technologies. Unfortunately, the current media and educational discourses focus much more on acquiring data analytics skills and how AI presents a huge opportunity for business and economic growth than exposing students to the ethics of AI. For instance, a recent study finds that most current curriculums of AASCB-accredited business schools put a very strong emphasis on the importance and advantages of acquiring data analytics skills in, without deeply discussing the social and ethical implications of AI (Clayton & Clopton, 2019). Clearly, ethical training and critical thinking about the ethics of smart technologies should be much more integral to public educational pedagogy and institutional higher learning epistemology so that younger generations are prepared for an AI-powered society.

3) Increasing transparency when emotional AI is used in public place is important to gain public trust.

Across all cases, in both private and public use of emotional AI, our regression analysis indicates that trust in government regulation and trust in the private sector' self-regulation are consistently among the positive correlates with acceptance of emotional AI applications. This finding converges with the observation made by Vu and Lim (2021) that acceptance of AI and robots depends on perception of government effectiveness and how advanced a country's techno-social environment is. Consequently, the results imply citizen trust in the regulatory frameworks of new technologies are vital in its adoption and ethical use.

To secure this trust, it is important that the law must reinforce the transparency of how emotional data are processed, stored, and who can access the data. And importantly, how emotional AI is used should be assessed under the data minimization principle, which has been formalized in Article 5(1) of the GDPR, which stipulates, the collecting and processing data should only be limited only toward necessary ends.

In practice, at least in Japan, signages that notify when emotional AI are being used in public spaces are not commonplace and have not been strongly reinforced by the law. As admitted by a manager at ELSYS Japan admitted his company provides signages to convenience stores, airport security, and factories, where emotion-sensing AI is used, to notify the public and comply with the law, nevertheless, the signages are seldomly used when inspecting the sites. In recent legal scholarly studies, there have been serious doubts expressed by Japanese legal experts over the ability to reinforce data protection laws because even though there have been many public data breaches scandals, the national agency of Personal Data Protection Commission (PPC) has not issued any fine since its establishment of 2003. It has been noted that staffs of PPC are from the ministries and the commissioners are from various industries, which raise questions over conflicts of interest. For emotional AI applications to play a more positive role in social life in Japan, the shortcomings from two prominent stakeholders of emotional AI must be improved. Namely, the organizations and businesses that use this technology must be more forthcoming on how it is used and what measures have been taken to ensure it is used ethically and legally. In fact, I argue the growing presence of interactive AI systems that seek to modify our behaviors demand AI ethics to be incorporated formally into Corporate Social Responsibility practices.

4) International regulatory and ethical frameworks to assess the risks and rewards of emotional AI technology should consider cross-cultural differences in notion of privacy, autonomy, and fairness.

In recent years, there have been many initiatives seeking to assess and limit the negative impacts of AI technology. Nonetheless, it must be acknowledged that the current discourses around the ethics of emerging technologies are dominated by Western-centric ideas. For example, Japan, similar to the EU, has promoted the idea of respecting "human dignity" when developing and adopting smart technologies despite the fact that there is no native notion equivalent in the Japanese language for the word human dignity (Miyashita, 2021). Moreover, many important keywords that are often used in discourse in the age of AI such as privacy or autonomy are often based on the Western neo-liberal notion of individual liberty. There must be more acknowledgments for the evolution of the notion of privacy and autonomy in countries where communitarian and collective values are the norms. For example, Miyashita explains in his 2011 article on the evolution of the notion of privacy in Japan that there is still an ongoing acculturative process of the modern Japanese notion of privacy, in which, there is a hybridization between the traditional notion of privacy as a symbol for respect between the self and the collective and the Western notion of privacy rooted in individual liberty (Miyashita, 2011).

Another example includes the observation made about the notion of ownership and private belongings by the Japanese scholar, Matsuura Kazuya, who studies the implications of Japanese traditional beliefs in Bushido and Buddhism for the notion of ownership and privacy in the age of AI. Matsuura concludes that the ideal person according to both Bushido and Japanese Buddhism will ultimately shed his/her personal attachment to his belongings and serve the community, thus things he owns, in the end, belong to the community (Matsuura, 2021).

These deep-seated cultural notions must be accounted for when analyzing the ethics as well as law-making process related to the adoption of emotional AI. Here, the varying cultural notions of values such as privacy, autonomy, and fairness have two important implications. First, current dominant regulatory and ethical frameworks as well as academic theories used to study AI technology's impacts should be adapted to reflect and respect the traditional, native values. Second, cultural variations in these values also open the doors for the exploitation of these loopholes to harvest personal data of unsuspecting populations. Third, the cultural variations in notions of privacy, equality, and autonomy also means there is no one-size-fit-all regulation toward AI. Thus, moving forward, the approach taken in the EU's proposed AI Act that provides a risk-based assessment of various AI technologies should be embraced by law-makers in Japan and across the world to create a regulatory environment that balances between the development of the technologies and still preserve the cross-cultural differences in notion of privacy, autonomy, and fairness that are conducive to human flourishing.

5) People are willing to accept flawed emotional AI. What does that mean for policy?

In this study, we repeatedly find somewhat puzzling results regarding the concern for emotional AI's accuracy: it either bears no statistically significant relationship or positively correlates with the attitude toward the emerging technology. For example, in the case of car, we see a statistically significant positive association between attitude toward emotional AI and concern about its accuracy ($\beta_{AccuracyConcern}= 0.072^{***}$); or in the case of education, accuracy concern has no statistically significant association with attitude toward edtech emotional AI ($\beta_{AccuracyConcern}= -0.043$, $p = 0.067$); or in the case of security camera, accuracy concern is positively correlated with attitude toward emotional AI security camera ($\beta_{AccuracyConcern}$= 0.062***). These results imply that an awareness of the flaws in the technology is not equivalent to a rejection of the technology, and there is a willingness to accept the adoption of an emerging technology despite its flaws.

Here, the concept of ‘machinic verisimilitude’ coined by McStay in “*Emotional AI: The rise of empathic media*” is useful in making sense of such psychology as well as understanding its ethical implications. Machinic verisimilitude means that the empathy machines are capable of are not authentic but only “the appearance of intimate insight” (McStay, 2018, p.5). Nonetheless, that is not to say such an appearance is morally insignificant, because, as McStay comprehensively describe in the book, machinic verisimilitude alone has already transformed and had serious implications for our lives: we are afforded new abilities to communicate, to gauge emotional reactions of others, and to engage in new aesthetic experiences.

To a large extent, these results relax the worry voiced by many technologists that people will reject new technologies because their failures tend to produce visceral and salient reactions. In Hidalgo et al. (2021)’s *How humans judge machines*, through series of experiments, the authors find most crucial difference between our judgments toward AI versus toward humans is that we tend to not ascribe intention to AI, thus we judge them more by the outcomes, while the morality of a situation involving a human decision-maker is judged more by the intention. A clarifying example in Hidalgo et al.’s book is the event of a natural disaster, machines will be judged more harshly if they try to save humans and fail, while people in the same scenario will still be judged positively and received more empathy.

Such observation of human psychology is of great relevance to our subject since we are increasingly in the presence of AI systems whose performance is not of 100% success or accuracy rate but is nonetheless better than their human counterparts. For example, data from the 65,000 miles of self-driving cars by Waymo demonstrated how the current generation of autonomous vehicles can entirely avoid collision modes often caused by human drivers such as road departure or fixed objects collision (Schwall et al., 2020). In the case of emotional AI, interior sensing cars can save many more lives and prevent many more deaths, yet given the feature of human psychology described above, they would still be perceived as not trustworthy as humans. This seems to only be true for the US or Western populations.

Our results suggest otherwise for the Japanese population, where acceptance of the emerging emotional AI technology is undeterred by inaccuracy in many cases. Policy-wise, it suggests there seems to be a willingness to accept the initial potential failures and misfunctions of a new technology. On the one hand, if combined with a participatory approach to decision-making, this attitude is promising for developing a healthy approach to adoption of emotional AI applications that is not saddled by over-expectation or over-pessimism. On the other, this attitude is worrying given the increasing number of immigrant workers in Japan and the lack of women and other minorities in decision-making bodies in the country.

6) Resolving the tension between privacy and autonomy: From the Cartesian agent view to the Homo Faber view.

In the age of smart technology that not only *feels* but also *feeds* on our emotions, we are constantly facing a choice between giving up our private personal data to the machines in a trade for personalized benefits.

On the one hand, we have a desire to automate as much of our lives as possible, to free ourselves from mundane errands so that we have more room to do creative work and make deliberate choices when it is necessary. Thus, although we all have a strong preference for privacy, this desire compels us to give up our private personal data so that our machines can learn about us and help us automate our decisions or nudge us toward a better version of ourselves. In the literature on technological acceptance, this phenomenon is called privacy paradox or the personalization-privacy paradox, which is stated in the literature as although most people state strong preferences for the privacy of their personal data, they do not take steps to protect such data and often willing to give them up in the pursuit of personalized benefits (Ameen et al., 2022; Choi et al., 2019; Gerber et al., 2018).

On the other hand, we want to believe we are an independent, free-thinking agent, who are making free choices. Our autonomy and agency matter to us, thus, we do not want to be dependent on our machines and let them dictate our lives. Across all cases, we systematically found self-rated knowledge of emotional AI technologies positively predict an accepting attitude toward the technology. This finding reflects the phenomenon referred to in the literature as '*taming the algorithms*,' where researchers found self-belief in the ability to exert one's agency in social media platforms is indicative of active engagement with AI technologies (Lobera et al., 2020; Lu, 2020).

Here, such constant push and pull in our interaction with smart emotion-sensing systems, which are embedded in our physical and virtual environment and constantly interacting, even modifying our behaviors, have highlighted the *illusoriness of our mind as a completely independent system*. As pointed out by philosophers Clowes et al. (2021) in a recent book titled '*The Mind-Technology Problem : Investigating Minds, Selves and 21st Century Artefacts*,' our thinking about AI has been driven by the Cartesian agent view. The Cartesian agent view's underlying assumption is that a system is fully in charge of its cognitive environment and its agency is separate from artefacts which it depends on. Increasingly, the rise of ubiquitous computing devices embedded in our smartphones, wearables, etc. have demonstrably blurred the distinction between artificial and human intelligence, between real life and the virtual world. For example, as pointed out in recent works in cognitive science, the rise of smartphones which has given us readily accessible information have been found to change how we form memories. Thus philosophically, such new human-machine relationship demands a reconceptualization of human memory, our epistemic environment, and even personal identity (Clowes et al., 2021).

In our quest to resolve the tension between privacy and autonomy implicated in the age of emotional AI, perhaps, it is wise to abandon the commonsensical, intuitive dualist notion that we are a system is fully in charge of its cognitive environment and our agency is separate from artefacts which it depends on (i.e., the Cartesian agent view in the literature on the extended mind hypothesis (Clowes et al., 2021) or the notion of natural-born dualist popularized by Yale psychologist, Paul Bloom (Bloom, 2007, 2013)). It is wiser, perhaps, to start internalizing the Homo Faber view, which states the evolution of the mind depends both on the history and the pre-history of our artefacts. In other words, our mind has been shaped both by a natural evolutionary process in the pre-history of our artefacts as well as by the creation of technologies. A clear example is in *Ultrasociety*, where Peter Turchin (2016) applies multi-level cultural evolution dynamics and demonstrates our deep-seated preference for equality has been shaped by the creation of projectile technology (a spear, stone-throwing, bows, and arrows, etc.). Turchin argues hunter-gather society is fiercely egalitarian, not because of an innate tendency toward egalitarianism as suggested by the metaphor of a noble savage by Jean-Jacques Rousseau (Wynn et al., 2017), but because the ease of using a projectile technology to punish anyone who wants to upset the hierarchy. The technological invention of projectile technology also puts immense evolutionary pressure on humans to evolve the capacity to communicate complex thoughts and form social alliances.

Once we relax the hold of the Cartesian agent view, and see ourselves more as interdependent, contingent beings, we can start to appreciate how much the philosophical difference in how society understands the ethics of governance and the relationship between the individual and the collective (Vuong and Napier, 2015; Vuong et al., 2018) also plays a key role in shaping the design and use of AI. For instance, the social media platforms in China such as Weibo or Douyin are thought to favor news, and videos about science and engineering projects as an effort of social engineering norms that are considered favorable for 21st-century living: curiosity, love, and respect for science and engineering. Meanwhile, social media algorithms in the West are found to be the main factor in driving the hyper-fragmentation online, which is often referred to as an affectively polarized cyberspace (Santos et al., 2021; Morgan et al., 2021).

7) A Turing test for emotional AI?

The final lesson concerns the specifications of a Turing test for emotional AI. Emotional artificial intelligence, in its present form, is a weak, and narrow form of AI since it is limited by its pre-programs and it does not have the capacity to understand or experience any parts of its information processing, whether input, output, or its algorithm, as discussed in the first chapter.

It is arguable that at some point in the future, emotional AI will achieve the status of being strong and general, which means, it is no longer limited by its programs, and it can have subjective experience of the emotions it is trained to recognize. In other words, it will pass the Turing test for emotional intelligence. In this section, I argue for *five specifications* of the Turing test for emotional AI, drawing from Schwaninger (2022)'s work on a philosophizing machine.

Schwaninger (2022) develops a specification of the Turing test based on his observations of large language models. Here, the author specifies that the Turing test for large-language-model AI is whether it can philosophize, provided three requirements. First, there is a need to control its training data, i.e., knowing a reasonable level of detail what the training data contain and how the machine might manipulate symbols/texts to come up with its answer. Second, testing the machine to see if it has any gasp of vagueness such as in the sorites paradox. Third, the test must also cover whether the machine can come up with a psychological question, i.e., a question that identifies why humans are inclined to accept the truth of an obviously false conclusion given its induction steps and the premises.

What are the specifications of a Turing test for emotional AI then? Drawing on Schwaninger (2022)'s work, one can extrapolate the specifications of emotional AI's Turing test in a few interesting ways.

First, having a conversation about emotions is a good way to test emotional understanding of AI. Clearly, dialogue plays an important role in the original Turing test as well as Schwaninger's specifications. In the case of emotional AI, given the recent increased reliance on multi-modalities of data (texts, voice tone, biometric data, video images, etc.) to develop emotional AI, it is likely that future emotional AI systems will be able to use conversation to convey its understanding of emotions. An example of a Turing test for emotional AI includes showing the AI and a control human subject a video of humans interacting, then letting an examiner pose questions to both the AI system and the human about the emotions that can be inferred from the videos. This leads to the second requirement.

Second, it is necessary to control the training data and the training protocol for emotional AI. Specifically, similar to Schwaninger's first requirement, the AI shall not have prior knowledge of certain emotions, and it is necessary to know in reasonable details how such an AI system come up with an answer when being asked to recognize an emotion. When encountering emotions that are not in its training data, and whether it can realize that it does not know such emotions would provide evidence for its capacity of strong and general emotional intelligence.

Third, one can also leverage cultural differences in emotional expression to test its understanding. For example, while the AI only receives training emotional data from people in a culture, in the Turing test, an examiner can show the EAI video tapes or chats of people from a different culture. If the EAI system identifies confusion in itself, then we can say this can also be evidence of its emotional understanding.

Fourth, causal relationships among emotion, reason, and action (words spoken included) can also be leveraged to test emotional understanding of an AI system. An emotional AI system that passes the Turing test should be able to identify the possible causal relationship among emotions and actions of the people in a video or a dialogue that are being presented to it. For example, it should be able to make factual statements such as "person A breaks things because he/she feels angry" and also counterfactual statements such as "had this person not felt stressed, he/she would have not cursed." More importantly, it needs to be able to identify ambiguous situations, where it is not clear what is the causal direction of an emotion and an action. This is to leverage the concept of vagueness in Schwaninger (2022)'s work.

Finally, an emotional AI system that passes the Turing test should be required to have an intelligible conversation to philosophize about the nature of emotions. It must be said that it is still a heated debate whether emotions are biologically hardwired into human beings (i.e., the essentialist account) or emotions are socially constructed (i.e., the constructivist account). According to the theory of constructed emotion, emotions are abstract categories, constructed as mental representations of us and the world, to fulfill five functions: meaning-making, body-regulating, action-prescribing, communication, and social influence. If this account is correct, it implies that emotion expression and emotion inference are not mere cognitive functions, but it has clear behavioral and social mandates. To develop a capacity for understanding emotions, one must interact with the physical and social world. The cases of emotional disorders among children who lack social interactions when they were infant point to a highly probable conclusion: a disembodied algorithm cannot pass the Turing test for emotional intelligence, specified above.

In sum, this section puts forth some considerations on the features of a Turing Test for emotional AI, which includes the control of training emotional data, the use of cultural differences in emotions, the use of dialogue, the use of causal relationships among emotion, reason, and action, as well as the philosophizing on the nature of emotions. These tools serve as initial parameters for evaluating whether an emotional AI machine has achieved the status of general and strong emotional intelligence. It also highlights the tension in theoretical debates on what emotions are, whether they are biologically hardwired or constructed, and speculates that a disembodied affect-sensing algorithm cannot pass the Turing test for emotional intelligence. Laden in our understanding of emotions are our presumptions of what constitutes a mind and its relationship with the world. Thus, clarifying philosophical implications, including the epistemology, ontology, and ethics, of emotional AI, requires the efforts of not only theorists and scientists, but also engineers and citizens. Unraveling the mystery of the mind-technology problem is crucial for identifying ways to live well and ethically with smart technologies that not only *feel* but also *feed* off our emotions.

8) Concluding remarks

This study has provided a comprehensive investigation of social and ethical perceptions of Japanese people regarding the rise of emotional AI. With the staggering annual growth rate of 11.36% over the recent 25-year period and huge potential for commercial and political applications, the field of emotional AI is expected to continue to expand and becomes an ever-more dynamic sub-field of AI, continuing to attract billions and billions in research and development.

Yet, this study has shown that the adoption of emotional AI or any emerging interactive technologies cannot be understood separately from the physical, social, and cultural worlds they will be integrated. Being the first study in the literature to bring the Technological Acceptance Model and the Moral Foundation Theory together under the analytical Three-pronged Approach (Contexts, Variables, and Statistical models) to study determinants of emotional AI's acceptance in 10 different use cases in Japan, this study has found some successes in statistical modeling of emotional AI's social perception. The statistical models have successfully accounted for an average of 52.11% of the variation in the data (min = 38%; max = 67.8%) (Table 10.1), on the same par with current models in the literature on technological adoption. In the most successful case, Home Robots, our model accounts for 67.8% of the variation in the data, outperforming existing models in the literature. Noteworthily, in a meta-analysis, the extended TAM accounts for only 44% of the behavioral intention to adopt smart technologies (Scherer et al., 2019). Thus, this indicates that the Three-pronged Approach of this thesis, i.e., the combination of the TAM and Moral Foundation Theory, can open up a fertile group for further empirical studies in the field of human-machine relations. The visualization in Figure 11.1 presents several key considerations for future studies to consider when designing their survey and experiments to examine the social perceptions of smart machines: 1) Culturally aware survey/experiment design, 2) making theoretical assumptions apparent; 2) Deliberate choices of statistical methods.

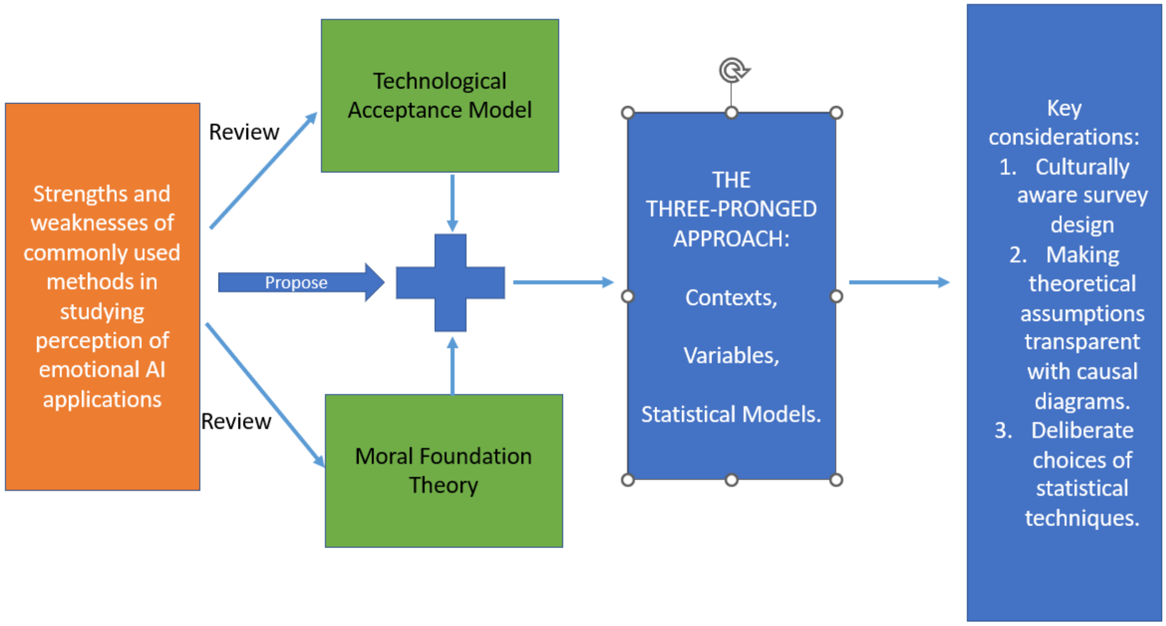


Figure 11.1: A visualization of key considerations of the Three-Pronged approach taken in this study.

In the course of this study, we have identified several key empirical results that carry strong implications for policymaking. First, women are more concerned about key ethical issues of emotional AI: algorithmic biases, data privacy, loss of autonomy, etc. Second, age is a negative correlate of attitude toward emotional AI applications, suggesting more public outreach efforts are needed to promote AI solutions for the elderly population—a major beneficiary of emotional AI technologies in the rapidly aging Japanese society. Third, interestingly and paradoxically, in many cases, accuracy concern, data management concern, and bias concern are found to be either non-significant or positively correlated with attitude toward emotional AI. As such, they suggest a willingness to adopt emotional AI applications despite its potential flaws and muddy issues around data management, or even its lack of consideration for disadvantaged social groups. This attitude relaxes the concern that many technologists have raised over the hesitance of AI adoption due to its failure would be more psychologically jarring and salient. Nonetheless, these results are worrying given the increasing number of immigrant workers and the lack of women in key decision-making positions in Japan. Finally, throughout the study, we have discussed aspects of Japanese cultures that come into conflict with the adoption of emotional AI technologies, namely, the long-standing cultural belief in homogenous Japan, the unspoken trust between employer and employee, the ambiguity in Japanese communication and what it means for emotion-sensing, the blurry boundary between the private and the communal sphere, etc.

Consequently, the thesis calls for the development of theoretical frameworks that capture better cross-cultural differences in moral reasoning about the effects of technologies on our daily lives. Indeed, the findings in this study carry important implications for the governance of the emerging emotional AI technologies, which have been provided in the final chapter. As such, our findings offer a fertile platform for further exploration of the complex intersection between psychology, culture, and emotion-recognition technologies as well as vital insights for policymakers wishing to ensure the design and regulation of the technology serves the best interests of society.

References

Abe, K., & Iwata, S. (2022). NEC's emotion analysis solution supports work style reform and health management. *NEC Technical Journal*, *2019*(190109). https://www.nec.com/en/global/techrep/journal/g19/n01/190109.html

Aboujaoude, E., Savage, M. W., Starcevic, V., & Salame, W. O. (2015). Cyberbullying: Review of an Old Problem Gone Viral. *Journal of Adolescent Health*, *57*(1), 10-18. https://doi.org/https://doi.org/10.1016/j.jadohealth.2015.04.011

Adikari, A., de Silva, D., Moraliyage, H., Alahakoon, D., Wong, J., Gancarz, M., Chackochan, S., Park, B., Heo, R., & Leung, Y. (2022). Empathic conversational agents for real-time monitoring and co-facilitation of patient-centered healthcare. *Future Generation Computer Systems*, *126*, 318-329. https://doi.org/https://doi.org/10.1016/j.future.2021.08.015

AdMobilize. (2022). *Signalive delivering media to your screens*. Retrieved August 19 from https://www.admobilize.com/signagelive/

Aho, B., & Duffield, R. (2020). Beyond surveillance capitalism: Privacy, regulation and big data in Europe and China. *Economy and Society*, *49*(2), 187-212. https://doi.org/10.1080/03085147.2019.1690275

AI Smiley. (2022). *AI を教育現場に活用するメリット・デメリットとは？導入事例を紹介 [What are the advantages and disadvantages of using AI in education?]*. Retrieved September 24 from https://aismiley.co.jp/ai_news/what-are-the-advantages-and-disadvantages-of-ai-for-education/

Akhtar, S., Ghosal, D., Ekbal, A., Bhattacharyya, P., & Kurohashi, S. (2019). All-in-One: Emotion, Sentiment and Intensity Prediction using a Multi-task Ensemble Framework. *IEEE Transactions on Affective Computing*, 1-1. https://doi.org/10.1109/TAFFC.2019.2926724

Ali, J. (2012). Factors affecting the adoption of information and communication technologies (ICTs) for farming decisions. *Journal of Agricultural & Food Information*, *13*(1), 78-96. https://doi.org/10.1080/10496505.2012.636980

Alina, P., & Khalina, N. y. (2021). Perception of New Technologies: Constructing an Innovation Openness Index. *Foresight and STI Governance*, *15*(1), 39-54. https://foresight-journal.hse.ru/en/2021-15-1/454834548.html

Alsaleh, D. A., Elliott, M. T., Fu, F. Q., & Thakur, R. (2019). Cross-cultural differences in the adoption of social media. *Journal of Research in Interactive Marketing*, *13*(1), 119-140. https://doi.org/10.1108/JRIM-10-2017-0092

Ameen, N., Hosany, S., & Paul, J. (2022). The personalisation-privacy paradox: Consumer interaction with smart technologies and shopping mall loyalty. *Computers in Human Behavior*, *126*, 106976. https://doi.org/https://doi.org/10.1016/j.chb.2021.106976

Anaraky, R. G., Byrne, K. A., Wisniewski, P. J., Page, X., & Knijnenburg, B. (2021). *To Disclose or Not to Disclose: Examining the Privacy Decision-Making Processes of Older vs. Younger Adults* Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan. https://doi.org/10.1145/3411764.3445204

Anderson, M., & Jiang, J. (2018). *Teens, Social Media and Technology 2018*. Retrieved March 03 from https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/

Atari, M., Lai, M. H. C., & Dehghani, M. (2020). Sex differences in moral judgements across 67 countries. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1937), 20201201. https://doi.org/10.1098/rspb.2020.1201

Aydin, B. (2019). Public acceptance of drones: Knowledge, attitudes, and practice. *Technology in Society*, *59*, 101180. https://doi.org/https://doi.org/10.1016/j.techsoc.2019.101180

Badawy, A., Ferrara, E., & Lerman, K. (2018, 28-31 Aug. 2018). Analyzing the Digital Traces of Political Manipulation: The 2016 Russian Interference Twitter Campaign. 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM),

Bakir, V. (2020). Psychological operations in digital political campaigns: Assessing cambridge analytica's psychographic profiling and targeting. *Frontiers in Communication*, *5*, 67. https://doi.org/10.3389/fcomm.2020.00067

Bakir, V., Ghotbi, N., Ho, T. M., Laffer, A., Mantello, P., McStay, A., Miranda, D., Miyashita, H., Podoletz, L., Tanaka, H., & Urquhart, L. (2022). Emotional AI in Cities. *Machine Learning and the City*, 621-624. https://doi.org/https://doi.org/10.1002/9781119815075.ch51 (Wiley Online Books)

Bard, J. S. (2021). Developing Legal Framework for Regulating Emotion AI. *BUJ Sci. & Tech. L.*, *27*, 271.

Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*(1), 1-23. https://doi.org/10.1093/scan/nsw154

Barrett, L. F. (2021). AI weighs in on debate about universal facial expressions. *Nature*, *589*(7841), 202-204.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1-68. https://doi.org/10.1177/1529100619832930

Barros, P., Churamani, N., & Sciutti, A. (2020). The FaceChannel: A Fast and Furious Deep Neural Network for Facial Expression Recognition. *SN Computer Science*, *1*(6), 321. https://doi.org/10.1007/s42979-020-00325-6

Basiri, M. E., Nemati, S., Abdar, M., Cambria, E., & Acharya, U. R. (2021). ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis. *Future Generation Computer Systems*, *115*, 279-294. https://doi.org/https://doi.org/10.1016/j.future.2020.08.005

Bean, R. (2022). *Why becoming a data-driven organization is so hard*. Retrieved August 18 from

Bearn, S. (2021). *A wristband that tells your boss if you are unhappy*. Retrieved July 1st from https://www.bbc.com/news/business-55637328

Bennett, E. (2014). Deus ex Machina: AI Apocalypticism in Terminator: The Sarah Connor Chronicles. *The Journal of Popular Television*, *2*(1), 3-19. https://doi.org/10.1386/jptv.2.1.3_1

Birhane, A. (2020). Algorithmic colonization of Africa. *SCRIPTed*, *17*(2), 389. https://script-ed.org/article/algorithmic-colonization-of-africa/

Blanden, J., & Gregg, P. (2004). Family Income and Educational Attainment: A Review of Approaches and Evidence for Britain. *Oxford Review of Economic Policy*, *20*(2), 245-263. https://doi.org/10.1093/oxrep/grh014

Bloom, P. (2007). Religion is natural [https://doi.org/10.1111/j.1467-7687.2007.00577.x]. *Developmental Science*, *10*(1), 147-151. https://doi.org/https://doi.org/10.1111/j.1467-7687.2007.00577.x

Bloom, P. (2013). *Just babies: The origins of good and evil*. Broadway Books.

Bondanini, G., Giorgi, G., Ariza-Montes, A., Vega-Muñoz, A., & Andreucci-Annunziata, P. (2020). Technostress Dark Side of Technology in the Workplace: A Scientometric Analysis. *International journal of environmental research and public health*, *17*(21), 8013. https://doi.org/10.3390/ijerph17218013

Bostrom, N. (2017). *Superintelligence*. Dunod.

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, *108*(Supplement 2), 10918. https://doi.org/10.1073/pnas.1100290108

Brassart Olsen, C. (2020). To track or not to track? Employees' data privacy in the age of corporate wellness, mobile health, and GDPR†. *International Data Privacy Law*, *10*(3), 236-252. https://doi.org/10.1093/idpl/ipaa004

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Brougham, D., & Haar, J. (2018). Smart Technology, Artificial Intelligence, Robotics, and Algorithms (STARA): Employees' perceptions of our future workplace. *Journal of Management & Organization*, *24*(2), 239-257. https://doi.org/10.1017/jmo.2016.55

Buolamwini, J., & Gebru, T. (2018a). Gender shades: Intersectional accuracy disparities in commercial gender classification. Conference on fairness, accountability and transparency,

Buolamwini, J., & Gebru, T. (2018b). *Gender shades: Intersectional accuracy disparities in commercial gender classification* Proceedings of the 1st Conference on Fairness, Accountability and Transparency, Proceedings of Machine Learning Research. http://proceedings.mlr.press

Bustamante, C. M. V., Alama-Maruta, K., Ng, C., & Coppersmith, D. D. (2022). Should machines be allowed to 'read our minds'? Uses and regulation of biometric techniques that attempt to infer mental states. *MIT Science Policy Review*, *3*(August 29), 112-121. https://doi.org/10.38105/spr.qy2iibrk72

Cabinet Office Japan. (2019). *Society 5.0*. Retrieved September 23 from https://www8.cao.go.jp/cstp/english/society5_0/index.html

Cabinet Office Japan. (2021). *Annual Reporton the Ageing Society [Summary] FY2021*. Retrieved Sep 22 from https://www8.cao.go.jp/kourei/english/annualreport/2021/pdf/2021.pdf

Cabinet Office of Japan. (2019). *Social Principles of Human-Centric AI*. Retrieved September 23 from https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf

Cabitza, F., Campagner, A., & Mattioli, M. (2022). The unbearable (technical) unreliability of automated facial emotion recognition. *Big Data & Society*, *9*(2), 20539517221129549. https://doi.org/10.1177/20539517221129549

Calvo, R. A., D'Mello, S., Gratch, J. M., & Kappas, A. (2015). *The Oxford handbook of affective computing*. Oxford Library of Psychology.

Cambria, E., Li, Y., Xing, F. Z., Poria, S., & Kwok, K. (2020). SenticNet 6: Ensemble Application of Symbolic and Subsymbolic AI for Sentiment Analysis. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (pp. 105–114). Association for Computing Machinery. https://doi.org/10.1145/3340531.3412003

Castleberry, A., & Nolen, A. (2018). Thematic analysis of qualitative research data: Is it as easy as it sounds? *Currents in Pharmacy Teaching and Learning*, *10*(6), 807-815. https://doi.org/https://doi.org/https://doi.org/10.1016/j.cptl.2018.03.019

Chen, C., Crivelli, C., Garrod Oliver, G. B., Schyns Philippe, G., Fernández-Dols, J.-M., & Jack Rachael, E. (2018). Distinct facial expressions represent pain and pleasure across cultures.

Proceedings of the National Academy of Sciences, *115*(43), E10013-E10021. https://doi.org/10.1073/pnas.1807862115
Chen, S.-Y., & Lee, C. (2019). Perceptions of the Impact of High-Level-Machine-Intelligence from University Students in Taiwan: The Case for Human Professions, Autonomous Vehicles, and Smart Homes. *Sustainability*, *11*(21).
Chicca, J., & Shellenbarger, T. (2018). Connecting with Generation Z: Approaches in Nursing Education. *Teaching and Learning in Nursing*, *13*(3), 180-184. https://doi.org/https://doi.org/10.1016/j.teln.2018.03.008
Choi, J. P., Jeon, D.-S., & Kim, B.-C. (2019). Privacy and personal data collection with information externalities. *Journal of Public Economics*, *173*, 113-124. https://doi.org/https://doi.org/10.1016/j.jpubeco.2019.02.001
Chowdhury, S., & van Wee, B. (2020). Examining women's perception of safety during waiting times at public transport terminals. *Transport Policy*, *94*, 102-108. https://doi.org/https://doi.org/10.1016/j.tranpol.2020.05.009
Clarke, S., & Hoggett, P. (2019). *Researching beneath the surface: Psycho-social research methods in practice*. Routledge.
Clayton, P. R., & Clopton, J. (2019). Business curriculum redesign: Integrating data analytics. *Journal of Education for Business*, *94*(1), 57-63. https://doi.org/10.1080/08832323.2018.1502142
Clowes, R. W., Gärtner, K., & Hipólito, I. (2021). The Mind Technology Problem and the Deep History of Mind Design. In R. W. Clowes, K. Gärtner, & I. Hipólito (Eds.), *The Mind-Technology Problem : Investigating Minds, Selves and 21st Century Artefacts* (pp. 1-45). Springer International Publishing. https://doi.org/10.1007/978-3-030-72644-7_1
Compass. (2022). *Stress free IT environment* Retrieved September 29 from
Conway, M., & O'Connor, D. (2016). Social media, big data, and mental health: current advances and ethical implications. *Current Opinion in Psychology*, *9*, 77-82. https://doi.org/https://doi.org/10.1016/j.copsyc.2016.01.004
Crawford, K. (2021a). *Artificial intelligence is misreading human emotion*. Retrieved 2021 May 14 from https://www.theatlantic.com/technology/archive/2021/04/artificial-intelligence-misreading-human-emotion/618696/
Crawford, K. (2021b). Time to regulate AI that interprets human emotions. *Nature*, *592*(167). https://doi.org/10.1038/d41586-021-00868-5
Creswell, J. W. (1994). *Research design: Qualitative and quantitative approach*. Sage Publications.
D'Ignazio, C., & Klein, L. F. (2020). *Data Feminism* MIT Press.
Dalton, E. (2022). *Japan's stubborn gender inequality problem*. Retrieved January 06 from https://www.eastasiaforum.org/2022/06/28/japans-stubborn-gender-inequality-problem/
Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319-340.
Dery, G. (2018). Is a friend truly a friend if you can just leave it in the garage?Toyota's and Honda's concept cars could have significant Fourth Amendment implications. *American Criminal Law Review*, *55*, 585.
Digital Agency JP. (2022). *Roadmap on the utilization of data in education*. Retrieved January 06 from https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/0f321c23-517f-439e-9076-5804f0a24b59/20220307_en_education_outline_01.pdf

Dirksen, N. (2020). Artificial Intelligence in Japan 2020.

Do, N. T., Nguyen-Quynh, T. T., & Kim, S. H. (2020, 16-20 Nov. 2020). Affective Expression Analysis in-the-wild using Multi-Task Temporal Statistical Deep Learning Model. 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020),

Doki, S., Sasahara, S., Hori, D., Oi, Y., Takahashi, T., Shiraki, N., Ikeda, Y., Ikeda, T., Arai, Y., Muroi, K., & Matsuzaki, I. (2021). Comparison of predicted psychological distress among workers between artificial intelligence and psychiatrists: a cross-sectional study in Tsukuba Science City, Japan. *BMJ Open*, *11*(6), e046265. https://doi.org/10.1136/bmjopen-2020-046265

Dombrosky, K., Templeton, B., & Fong, J. (2018). *An insider's guide to Generation Z and higher education 2018*. University Professional and Continuing Education Association. Retrieved March 03 from https://upcea.edu/wp-content/uploads/2018/01/An-Insiders-Guide-to-Generation-Z-and-Higher-Education-eBook.pdf

Dooley, B., & Ueno, H. (2022). *Where a thousand digital eyes keep watch over the elderly*. https://www.nytimes.com/2022/02/02/business/japan-elderly-surveillance.html

Dutot, V., Bhatiasevi, V., & Bellallahom, N. (2019). Applying the technology acceptance model in a three-countries study of smartwatch adoption. *The Journal of High Technology Management Research*, *30*(1), 1-14. https://doi.org/https://doi.org/10.1016/j.hitech.2019.02.001

Ekman, P. (1999). Basic Emotions. In T. Dalgleish & M. Power (Eds.), *Handbook of Cognition and Emotion*. John Wiley & Sons.

European Institute for Gender Equality. (2020). Gender Equality Index 2020: Digitalisation and the future of work. https://eige.europa.eu/publications/gender-equality-index-2020-report/men-dominate-technology-development

Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S. R., Park, H., & Shao, C. (2016). Applications of structural equation modeling (SEM) in ecological studies: an updated review. *Ecological Processes*, *5*(1), 19. https://doi.org/10.1186/s13717-016-0063-3

Fast, E., & Horvitz, E. (2017). Long-term trends in the public perception of artificial intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence*, *31*(1). https://ojs.aaai.org/index.php/AAAI/article/view/10635

Field, H. (2022). *The Algorithmic Accountability Act is back—here's what's in it*. Retrieved March 03 from https://www.morningbrew.com/emerging-tech/stories/2022/02/11/the-algorithmic-accountability-act-is-back-here-s-what-s-in-it

Fleming, S. (2019). *Japan's workforce will be 20% smaller by 2040*. Retrieved October 01 from https://www.weforum.org/agenda/2019/02/japan-s-workforce-will-shrink-20-by-2040

Fujitsu Website. (2022). *LifeMark Health Assistant*. Retrieved October 01 from https://www.fujitsu.com/jp/solutions/industry/healthcare/products/lifemarkha/

Fumiko, K., Arai, H., & Ema, A. (2020). Ethical Issues Regarding the Use of AI Profiling Services for Recruiting: The Japanese Rikunabi Data Scandal. *arXiv preprint arXiv:2005.08663*.

Gazzano, A. (2021). Japan's programming education: a critical focus on music in elementary schools. *Arts Education Policy Review*, 1-13. https://doi.org/10.1080/10632913.2021.1974996

Gerber, N., Gerber, P., & Volkamer, M. (2018). Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers & Security*, *77*, 226-261. https://doi.org/https://doi.org/10.1016/j.cose.2018.04.002

Ghotbi, N., & Ho, M. T. (2021). Moral Awareness of College Students Regarding Artificial Intelligence. *Asian Bioethics Review*, *13*(4), 421-433. https://doi.org/10.1007/s41649-021-00182-2
Ghotbi, N., Ho, M. T., & Mantello, P. (2022). Attitude of college students towards ethical issues of artificial intelligence in an international university in Japan. *AI & SOCIETY*, *37*(1), 283-290. https://doi.org/10.1007/s00146-021-01168-2
Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of personality and social psychology*, *101*(2), 366.
Graziosi, G. (2020). *Amazon uses worker surveillance to boost performance and stop staff joining unions, study says*. Retrieved March 23, 2021 from https://www.independent.co.uk/news/world/americas/amazon-surveillance-unions-report-a9697861.html
Gruebner, O., Sykora, M., Lowe, S. R., Shankardass, K., Trinquart, L., Jackson, T., Subramanian, S. V., & Galea, S. (2016). Mental health surveillance after the terrorist attacks in Paris. *The Lancet*, *387*(10034), 2195-2196. https://doi.org/10.1016/S0140-6736(16)30602-X
Haidt, J. (2007). The New Synthesis in Moral Psychology. *Science*, *316*(5827), 998-1002. https://doi.org/10.1126/science.1137651
Hanemaayer, A. (2022). Introduction: Critical Insights—Bringing the social sciences and humanities to AI. In A. Hanemaayer (Ed.), *Artificial Intelligence and Its Discontents: Critiques from the Social Sciences and Humanities* (pp. 1-20). Springer International Publishing. https://doi.org/10.1007/978-3-030-88615-8_1
He, Y., Yatsuya, H., Ota, A., & Tabuchi, T. (2022). The association of public trust with the utilization of digital contact tracing for COVID-19 in Japan. *Public Health in Practice*, *4*, 100279. https://doi.org/https://doi.org/10.1016/j.puhip.2022.100279
Henrich, J. (2020). *The WEIRDest people in the world: How the West became psychologically peculiar and particularly prosperous*. Farrar, Straus and Giroux.
Hidalgo, C. A., Orghian, D., Canals, J. A., De Almeida, F., & Martin, N. (2021). *How humans judge machines*. MIT Press.
Ho, M.-T., Mantello, P., Ghotbi, N., Nguyen, M.-H., Nguyen, H.-K. T., & Vuong, Q.-H. (2022). Rethinking technological acceptance in the age of emotional AI: Surveying Gen Z (Zoomer) attitudes toward non-conscious data collection. *Technology in Society*, 102011. https://doi.org/https://doi.org/10.1016/j.techsoc.2022.102011
Ho, M.-T., Mantello, P., Nguyen, H.-K. T., & Vuong, Q.-H. (2021). Affective computing scholarship and the rise of China: a view from 25 years of bibliometric data. *Humanities and Social Sciences Communications*, *8*(1), 282. https://doi.org/10.1057/s41599-021-00959-8
Holroyd, C. (2022). Technological innovation and building a 'super smart' society: Japan's vision of society 5.0. *Journal of Asian Public Policy*, *15*(1), 18-31. https://doi.org/10.1080/17516234.2020.1749340
Hope, A. L. B., & Jones, C. R. (2014). The impact of religious faith on attitudes to environmental issues and Carbon Capture and Storage (CCS) technologies: A mixed methods study. *Technology in Society*, *38*, 48-59. https://doi.org/https://doi.org/10.1016/j.techsoc.2014.02.003
Horii, M., & Sakurai, Y. (2020). *The future of work in Japan: Accelerating automation after COVID-19* (Transformation Practice, Issue.

https://www.mckinsey.com/~/media/mckinsey/locations/asia/japan/our%20insights/future%20of%20work%20in%20japan/the-future-of-work-in-japan_v4_en.pdf
Horiuchi, J. (2020). *Japan firms testing AI tech to recruit talent*. Retrieved October 01 from https://english.kyodonews.net/news/2020/06/f246f68dbb27-feature-japan-firms-testing-ai-tech-to-recruit-talent.html
Howell, D. L. (1996). Ethnicity and Culture in Contemporary Japan. *Journal of Contemporary History*, *31*(1), 171-190. http://www.jstor.org/stable/261100
Hoy, M. G., & Milne, G. (2010). Gender Differences in Privacy-Related Measures for Young Adult Facebook Users. *Journal of Interactive Advertising*, *10*(2), 28-45. https://doi.org/10.1080/15252019.2010.10722168
Huffman, A. H., Whetten, J., & Huffman, W. H. (2013). Using technology in higher education: The influence of gender roles on technology self-efficacy. *Computers in Human Behavior*, *29*(4), 1779-1786.
Inoue, K., Sakuma, N., Okada, M., Sasaki, C., Nakamura, M., & Wada, K. (2014, 2014//). Effective Application of PALRO: A Humanoid Type Robot for People with Dementia. Computers Helping People with Special Needs, Cham.
Ishibushi, K., & Matsakis, L. (2021). *Masafumi Ito, leader of the Amazon Japan Union, sues for wrongful termination*

Retrieved October 03 from https://restofworld.org/2021/lawsuit-amazon-japan-union/
Ishii, E., Ebner, D. K., Kimura, S., Agha-Mir-Salim, L., Uchimido, R., & Celi, L. A. (2020). The advent of medical artificial intelligence: lessons from the Japanese approach. *Journal of Intensive Care*, *8*(1), 35. https://doi.org/10.1186/s40560-020-00452-5
Ishizaki, A. (2021). *GIGA School Program is finally accelerating Japan's digitalization in education systems*. Retrieved September 29 from https://edujump.net/news/2856/
Ivanova, E., & Borzunov, G. (2020). Optimization of machine learning algorithm of emotion recognition in terms of human facial expressions. *Procedia Computer Science*, *169*, 244-248. https://doi.org/https://doi.org/10.1016/j.procs.2020.02.143
[Record #19 is using a reference type undefined in this output style.]
Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. *Stroke and Vascular Neurology*, *2*(4), 230. https://doi.org/10.1136/svn-2017-000101
JIJI Press. (2022). *Japan to face a shortage of nearly 1 million medical and welfare staff in 2040*. Retrieved November 11 from https://www.japantimes.co.jp/news/2022/09/16/national/medical-worker-shortage/
Johnson, D. G., & Verdicchio, M. (2017). Reframing AI Discourse. *Minds and Machines*, *27*(4), 575-590. https://doi.org/10.1007/s11023-017-9417-6
Kaiser, B. (2019). Targeted. My inside story of Cambridge Analytica and how Trump, Brexit and Facebook broke democracy. . In *HarperCollins*.
Kamal, S. A., Shafiq, M., & Kakria, P. (2020). Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM). *Technology in Society*, *60*, 101212. https://doi.org/https://doi.org/10.1016/j.techsoc.2019.101212
Kamo, K. (2020). *Pandemic speeds up automation at conveyor belt sushi chains*. Retrieved October 01 from
Kasilingam, D. L. (2020). Understanding the attitude and intention to use smartphone chatbots for shopping. *Technology in Society*, *62*, 101280. https://doi.org/https://doi.org/10.1016/j.techsoc.2020.101280

Kostka, G., Steinacker, L., & Meckel, M. (2021). Between security and convenience: Facial recognition technology in the eyes of citizens in China, Germany, the United Kingdom, and the United States. *Public Understanding of Science*, *30*(6), 671-690. https://doi.org/10.1177/09636625211001555

Kucirkova, N., Toda, Y., & Flewitt, R. (2021). Young Children's Use of Personalized Technologies: Insights From Teachers and Digital Software Designers in Japan. *Technology, Knowledge and Learning*, *26*(3), 535-554. https://doi.org/10.1007/s10758-020-09465-3

La, V.-P., & Vuong, Q.-H. (2019). *bayesvl: Visually Learning the Graphical Structure of Bayesian Networks and Performing MCMC with 'Stan.'* <https://cran.r-project.org/web/packages/bayesvl/index.html>; version 0.8.5 (officialy published on May 24, 2019).

Lecher, C. (2019). *How Amazon automatically tracks and fires warehouse workers for 'productivity'*. Retrieved March 23, 2021 from https://www.theverge.com/2019/4/25/18516004/amazon-warehouse-fulfillment-centers-productivity-firing-terminations

Lee, J., Kim, J., & Choi, J. Y. (2019). The adoption of virtual reality devices: The technology acceptance model integrating enjoyment, social interaction, and strength of the social ties. *Telematics and Informatics*, *39*, 37-48. https://doi.org/https://doi.org/10.1016/j.tele.2018.12.006

Lee, M., Lee, K., Shim, J., Cho, S. j., & Choi, J. (2016, 26-28 Oct. 2016). Security threat on wearable services: Empirical study using a commercial smartband. 2016 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia),

Lee, N. T., Resnick, P., & Barton, G. (2019). *Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms*. Brookings Institute. https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/

Lee, Y. (2022). *70% of Japan's junior high teachers crossing 'death by overwork' danger line: survey*. Retrieved September 29 from https://mainichi.jp/english/articles/20220514/p2a/00m/0na/031000c#:~:text=The%20survey%20found%20that%2074.4,or%20%22karoshi%22%20in%20Japanese

Lew, S., Tan, G. W.-H., Loh, X.-M., Hew, J.-J., & Ooi, K.-B. (2020). The disruptive mobile wallet in the hospitality industry: An extended mobile technology acceptance model. *Technology in Society*, *63*, 101430. https://doi.org/https://doi.org/10.1016/j.techsoc.2020.101430

Lima, G., Kim, C., Ryu, S., Jeon, C., & Cha, M. (2020). Collecting the Public Perception of AI and Robot Rights. *Proc. ACM Hum.-Comput. Interact.*, *4*(CSCW2), Article 135. https://doi.org/10.1145/3415206

Lobera, J., Fernández Rodríguez, C. J., & Torres-Albero, C. (2020). Privacy, Values and Machines: Predicting Opposition to Artificial Intelligence. *Communication Studies*, *71*(3), 448-465. https://doi.org/10.1080/10510974.2020.1736114

Lu, S. (2020). Taming the News Feed on Facebook: Understanding Consumptive News Feed Curation through a Social Cognitive Perspective. *Digital Journalism*, *8*(9), 1163-1180. https://doi.org/10.1080/21670811.2020.1837639

Lyles, T. (2020). *Moxie is a $1,500 robot for kids*. The Verge. Retrieved July 1st from https://www.theverge.com/2020/5/13/21257821/moxie-robot-kids-educational-tech-embodied-price

Manolev, J., Sullivan, A., & Slee, R. (2019). The datafication of discipline: ClassDojo, surveillance and a performative classroom culture. *Learning, Media and Technology*, *44*(1), 36-51. https://doi.org/10.1080/17439884.2018.1558237

Mantello, P., Ho, M.-T., Nguyen, M.-H., & Vuong, Q.-H. (2021). Bosses without a heart: socio-demographic and cross-cultural determinants of attitude toward Emotional AI in the workplace. *AI & SOCIETY*. https://doi.org/10.1007/s00146-021-01290-1

Marć, M., Bartosiewicz, A., Burzyńska, J., Chmiel, Z., & Januszewicz, P. (2019). A nursing shortage – a prospect of global and local policies [https://doi.org/10.1111/inr.12473]. *International Nursing Review*, *66*(1), 9-16. https://doi.org/https://doi.org/10.1111/inr.12473

Martín-Ruíz, M. L., Fernández-Aller, C., Portillo, E., Malagón, J., & del Barrio, C. (2018). Developing a System for Processing Health Data of Children Using Digitalized Toys: Ethical and Privacy Concerns for the Internet of Things Paradigm. *Science and Engineering Ethics*, *24*(4), 1057-1076. https://doi.org/10.1007/s11948-017-9951-x

Masami, H. (2021). Society 5.0 and education in Japan In Y. W. Purnomo & Herwin (Eds.), *Educational Innovation in Society 5.0 Era: Challenges and Opportunities: Proceedings of the 4th International Conference on Current Issues in Education (ICCIE 2020)*. https://doi.org/10.1201/9781003206019

Matson, Z., Donmez, B., Savan, B., Photiadis, D., Farahani, E., & Dafoe, J. (2012). Social drivers of technology adoption and use in the workplace productivity context. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 2103-2107. https://doi.org/10.1177/1071181312561446

Matsushita, M., & Yamamura, S. (2022). The Relationship Between Long Working Hours and Stress Responses in Junior High School Teachers: A Nationwide Survey in Japan [Original Research]. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.775522

Matsuura, K. (2021). *Socialization of robots: Humanities in anthropocene [ロボットをソーシャル化する:「人新世の人文学」10 の論点]*.

McClure, P. K. (2017). "You're fired," says the robot: The rise of automation in the workplace, technophobes, and fears of unemployment. *Social Science Computer Review*, *36*(2), 139-156. https://doi.org/10.1177/0894439317698637

McReynolds, E., Hubbard, S., Lau, T., Saraf, A., Cakmak, M., & Roesner, F. (2017). *Toys that Listen: A Study of Parents, Children, and Internet-Connected Toys* Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, Colorado, USA. https://doi.org/10.1145/3025453.3025735

McStay, A. (2018). *Emotional AI: The rise of empathic media*. Sage.

McStay, A. (2020a). Emotional AI and EdTech: serving the public good? *Learning, Media and Technology*, *45*(3), 270-283. https://doi.org/10.1080/17439884.2020.1686016

McStay, A. (2020b). Emotional AI, soft biometrics and the surveillance of emotional life: An unusual consensus on privacy. *Big Data & Society*, *7*(1), 2053951720904386. https://doi.org/10.1177/2053951720904386

McStay, A., & Rosner, G. (2021). Emotional artificial intelligence in children's toys and devices: Ethics, governance and practical remedies. *Big Data & Society*, *8*(1), 2053951721994877. https://doi.org/10.1177/2053951721994877

McStay, A., & Urquhart, L. (2019). 'This time with feeling?'Assessing EU data governance implications of out of home appraisal based emotional AI. *First Monday*, *24*(10).

McStay, A., & Urquhart, L. (2022). In cars (are we really safest of all?): interior sensing and emotional opacity. *International Review of Law, Computers & Technology*, 1-24. https://doi.org/10.1080/13600869.2021.2009181
Melville, N. P., Robert, L., & Xiao, X. (2022). Putting humans back in the loop: An affordance conceptualization of the 4th industrial revolution [https://doi.org/10.1111/isj.12422]. *Information Systems Journal*, *n/a*(n/a). https://doi.org/https://doi.org/10.1111/isj.12422
Menon, M. (2021). *New AI program could help screen seniors for depression and anxiety through video calls*. Retrieved March 03 from https://www.straitstimes.com/singapore/community/new-ai-program-could-help-counsellors-accurately-screen-seniors-for-depression
METI. (2020). *Recent Trends in AI-related Inventions.* . https://www.meti.go.jp/english/press/2020/0727_001.html
METI. (2022). *実践的な AI 人材育成のためのデータ付き教材の提供を開始します[We will start providing educational materials with data for practical AI human resource development]*. Retrieved September 24 from https://www.meti.go.jp/press/2022/05/20220509001/20220509001.html
MEXT. (2018). *学校における ICT 環境の整備について（教育の ICT 化に向けた環境整備 5 か年計画（2018（平成 30）～2022 年度）[Developing an ICT environment in schools (5-year plan for improving the environment for the introduction of ICT in education (2018-2022)]*. Retrieved September 29 from https://www.mext.go.jp/a_menu/shotou/zyouhou/detail/1402835.htm
Mitchell, M. (2019). *Artificial intelligence: A guide for thinking humans*. Penguin UK.
Miyashita, H. (2011). The evolving concept of data privacy in Japanese law. *International Data Privacy Law*, *1*(4), 229-238. https://doi.org/10.1093/idpl/ipr019
Miyashita, H. (2021). Human-centric data protection laws and policies: A lesson from Japan. *Computer Law & Security Review*, *40*, 105487. https://doi.org/https://doi.org/10.1016/j.clsr.2020.105487
Mohallick, I., Moor, K. D., Özgöbek, Ö., & Gulla, J. A. (2018). Towards new privacy regulations in europe: Users' privacy perception in recommender systems. International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage,
Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, *29*(3), 436-465.
Moore, P. V., & Woodcock, J. (2021). *Augmented exploitation: artificial intelligence, automation, and work*. Pluto Press.
Muk, A., & Chung, C. (2015). Applying the technology acceptance model in a two-country study of SMS advertising. *Journal of Business Research*, *68*(1), 1-6. https://doi.org/https://doi.org/10.1016/j.jbusres.2014.06.001
Muriithi, P., Horner, D., & Pemberton, L. (2016). Factors contributing to adoption and use of information and communication technologies within research collaborations in Kenya. *Information Technology for Development*, *22*(sup1), 84-100. https://doi.org/10.1080/02681102.2015.1121856
NEC Press Release. (2020). *Nec and Realeyes announce strategic alliance in emotion analysis solutions-Collaboration in distribution and development of emotion analysis products -*. Retrieved June 10 from https://www.nec.com/en/press/202010/global_20201023_03.html

Neri, H., & Cozman, F. (2020). The role of experts in the public perception of risk of artificial intelligence. *AI & SOCIETY*, *35*(3), 663-673. https://doi.org/10.1007/s00146-019-00924-9

Ngai, W. K., Xie, H., Zou, D., & Chou, K.-L. (2022). Emotion recognition based on convolutional neural networks and heterogeneous bio-signal data sources. *Information Fusion*, *77*, 107-117. https://doi.org/https://doi.org/10.1016/j.inffus.2021.07.007

Nikkei staff writers. (2018). *Japan's recruit employs AI to stop workers from quitting*. Retrieved October 01 from https://asia.nikkei.com/Business/Companies/Japan-s-Recruit-employs-AI-to-stop-workers-from-quitting

Nikkei Staff Writers. (2021). *Panasonic's cuddly robot helps chase away pandemic blues*. Retrieved Sep 19 from https://asia.nikkei.com/Business/Electronics/Panasonic-s-cuddly-robot-helps-chase-away-pandemic-blues

Norori, N., Hu, Q., Aellen, F. M., Faraci, F. D., & Tzovara, A. (2021). Addressing bias in big data and AI for health care: A call for open science. *Patterns*, *2*(10), 100347. https://doi.org/https://doi.org/10.1016/j.patter.2021.100347

Oderkirk, J., Ronchi, E., & Klazinga, N. (2013). International comparisons of health system performance among OECD countries: Opportunities and data privacy protection challenges. *Health Policy*, *112*(1), 9-18. https://doi.org/https://doi.org/10.1016/j.healthpol.2013.06.006

Odyssei. (2019). *HRTech market survey" by interviewing HR managers nationwide "AI" is the most effective latest IT tool when used in HRTech Over 60% of HR managers are thinking about introducing HRTech in the future!* Retrieved October 01 from https://prtimes.jp/main/html/rd/p/000000021.000026061.html.

OECD. (2017). *PISA 2015 Results (Volume V)*. https://doi.org/doi:https://doi.org/10.1787/9789264285521-en

OECD TALIS Report. (2013). *Key Findings from the Teaching and Learning International Survey (TALIS)*. OECD. Retrieved September 29 from https://www.oecd.org/japan/TALIS-2013-country-note-Japan.pdf

Ogawa, T., & Akada, Y. (2021). *Plan to use facial recognition at JR East stations to spot ex-cons iced*. Asahi Shinbun. Retrieved September 20 from https://www.asahi.com/ajw/articles/14445523

ÓhÉigeartaigh, S. S., Whittlestone, J., Liu, Y., Zeng, Y., & Liu, Z. (2020). Overcoming Barriers to Cross-cultural Cooperation in AI Ethics and Governance. *Philosophy & Technology*, *33*(4), 571-593. https://doi.org/10.1007/s13347-020-00402-x

Okada, J., & Iwatsu, S. (2021). *Toward realizing improved employee well-being through data-driven health and productivity management*. Retrieved October 01 from https://www.fujitsu.com/global/about/resources/publications/technicalreview/2021-01/article02.html

Okumus, B., Ali, F., Bilgihan, A., & Ozturk, A. B. (2018). Psychological factors influencing customers' acceptance of smartphone diet apps when ordering food at restaurants. *International Journal of Hospitality Management*, *72*, 67-77. https://doi.org/https://doi.org/10.1016/j.ijhm.2018.01.001

Onikle Inc. (2021). *The Rising Sun: Top 5 AI Companies in Japan*. Medium. https://onikle.medium.com/rising-sun-top-5-ai-companies-in-japan-21711be3bd32

ORG. (2020). *Public are Kept in the Dark Over Data Driven Political Campaigning, Poll Finds.* . Retrieved January 18 from https://www.openrightsgroup.org/press/releases/2020/public-are-kept-in-the-dark-over-data-driven-political-campaigning,-poll-finds

Osamu, I. (2002). *The Japanese Mind: Understanding Contemporary Japanese Culture* Tuttle Publishing.
Oshima, Y., & Sakai, M. (2020). The Enforcement of Personal Data Protection Law in Japan. *Global Privacy Law Review*, 173-179. http://www.kluwerlawonline.com/api/Product/CitationPDFURL?file=Journals\GPLR\GPLR2020094.pdf
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*, *35*(4), 927-936. https://doi.org/10.1007/s00146-020-00965-5
Ozaki, A. (2020, 2020//). Governance Framework for Facial Recognition Systems in Japan. Human-Centric Computing in a Data-Driven Society, Cham.
Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017). A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion*, *37*, 98-125. https://doi.org/https://doi.org/10.1016/j.inffus.2017.02.003
Psychoula, I., Singh, D., Chen, L., Chen, F., Holzinger, A., & Ning, H. (2018, 8-12 Oct. 2018). Users' Privacy Concerns in IoT Based Applications. 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI),
Rajak, M., & Shaw, K. (2021). An extension of technology acceptance model for mHealth user adoption. *Technology in Society*, *67*, 101800. https://doi.org/https://doi.org/10.1016/j.techsoc.2021.101800
Reddy, S., Allan, S., Coghlan, S., & Cooper, P. (2020). A governance model for the application of AI in health care. *Journal of the American Medical Informatics Association*, *27*(3), 491-497. https://doi.org/10.1093/jamia/ocz192
Rhue, L. (2019). Anchored to bias: How AI-human scoring can induce and reduce bias due to the anchoring effect. *Available at SSRN 3492129*.
Richardson, S. (2020). Affective computing in the modern workplace. *Business Information Review*, *37*(2), 78-85. https://doi.org/10.1177/0266382120930866
Rieder, B., & Hofmann, J. (2020). Towards platform observability. *Internet Policy Review*, *9*(4). https://doi.org/10.14763/2020.4.1535
Roberts, H., Cowls, J., Morley, J., Taddeo, M., Wang, V., & Floridi, L. (2021). The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation. *AI & SOCIETY*, *36*(1), 59-77.
Sato, H. (2020). Educational responses to the pandemic in Japan: Primary and secondary education policy issues. *Journal of the Commonwealth Council for Educational Administration & Management*, *48*(2), 64-70. http://cceam.net/wp-content/uploads/2020/08/ISEA-2020-48-2.pdf
Schelenz, L. (2022). Artificial Intelligence Between Oppression and Resistance: Black Feminist Perspectives on Emerging Technologies. In A. Hanemaayer (Ed.), *Artificial Intelligence and Its Discontents: Critiques from the Social Sciences and Humanities* (pp. 225-249). Springer International Publishing. https://doi.org/10.1007/978-3-030-88615-8_11
Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, *128*, 13-35. https://doi.org/https://doi.org/10.1016/j.compedu.2018.09.009

Schneider, T., Hong, G. H., & Le, V.-A. (2019). *Land of the rising robots*.
Schönberger, D. (2019). Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications. *International Journal of Law and Information Technology*, *27*(2), 171-203. https://doi.org/10.1093/ijlit/eaz004
Schuller, D., & Schuller, B. W. (2018). The Age of Artificial Emotional Intelligence. *Computer*, *51*(9), 38-46. https://doi.org/10.1109/MC.2018.3620963
Schuller, D. M., & Schuller, B. W. (2020). A Review on Five Recent and Near-Future Developments in Computational Processing of Emotion in the Human Voice. *Emotion Review*, *13*(1), 44-50. https://doi.org/10.1177/1754073919898526
Schwall, M., Daniel, T., Victor, T., Favaro, F., & Hohnhold, H. (2020). Waymo public road safety performance data. *arXiv preprint arXiv:2011.00038*.
Schwaninger, A. C. (2022). The Philosophising Machine – a Specification of the Turing Test. *Philosophia*. https://doi.org/10.1007/s11406-022-00480-5
Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and brain sciences*, *3*(3), 417-424.
Seiji, Y. (2018). 人工知能 AI の現状と教育への影響 [Current State of AI and its Influence to Education]. *Computer and Education*, *45*, 12-16.
Shimizu, Y., Ishizuna, A., Osaki, S., Hashimoto, T., Tai, M., Tanibe, T., & Karasawa, K. (2022). The Social Acceptance of Smart Health Services in Japan. *International journal of environmental research and public health*, *19*(3). https://doi.org/10.3390/ijerph19031298
Shoji, M. (2020). *Coronavirus pandemic has exacerbated Japan's hanko stamp problem—will it improve?* Retrieved October 05 from https://fpcj.jp/en/useful-en/wjn-en/p=81156/
Spataro, J. (2020). *The future of work—the good, the challenging & the unknown*. Retrieved July 1st from https://www.microsoft.com/en-us/microsoft-365/blog/2020/07/08/future-work-good-challenging-unknown
Spiegelhalter, D. (2019). *The art of statistics: Learning from data*. Penguin UK.
Stanton, J. M. (2001). Galton, Pearson, and the Peas: A Brief History of Linear Regression for Statistics Instructors. *Journal of Statistics Education*, *9*(3), null-null. https://doi.org/10.1080/10691898.2001.11910537
Stappen, L., Baird, A., Christ, L., Schumann, L., Sertolli, B., Messner, E., Cambria, E., Zhao, G., & Schuller, B. (2021). *The MuSe 2021 Multimodal Sentiment Analysis Challenge: Sentiment, Emotion, Physiological-Emotion, and Stress*.
Susanto, Y., Cambria, E., Ng, B. C., & Hussain, A. (2021). Ten Years of Sentic Computing. *Cognitive Computation*. https://doi.org/10.1007/s12559-021-09824-x
Suwa, S., Tsujimura, M., Kodate, N., Donnelly, S., Kitinoja, H., Hallila, J., Toivonen, M., Ide, H., Bergman-Kärpijoki, C., Takahashi, E., Ishimaru, M., Shimamura, A., & Yu, W. (2020). Exploring perceptions toward home-care robots for older people in Finland, Ireland, and Japan: A comparative questionnaire study. *Archives of Gerontology and Geriatrics*, *91*, 104178. https://doi.org/https://doi.org/10.1016/j.archger.2020.104178
Taherdoost, H. (2018). A review of technology acceptance and adoption models and theories. *Procedia Manufacturing*, *22*, 960-967. https://doi.org/https://doi.org/10.1016/j.promfg.2018.03.137
The Japan Times. (2021). *Japan considers making security cameras mandatory on trains*. Retrieved September 20 from https://www.japantimes.co.jp/news/2021/12/03/national/japan-train-cameras/
Tracxn. (2022). *AI in Healthcare Startups in Japan*. https://tracxn.com/explore/AI-in-Healthcare-Startups-in-Japan

Turchin, P. (2016). *Ultrasociety: How 10,000 years of war made humans the greatest cooperators on earth*. Beresta Books Chaplin, CT.

Ubie. (2022). *About Ubie*. https://ubie.life/en

UN.ESCAP. (2020). *Resource materials on data privacy laws in the Asia and the pacific : webinar on data privacy laws in ASEAN* https://hdl.handle.net/20.500.12870/4606.

United Nations Development Group. (2020). *Data privacy, ethics and protection guidance note on big data for achievement of the 2030 agenda*. Retrieved 2021 June 21 from http://unsdg.un.org/sites/default/files/UNDG_BigData_final_web.pdf

Urquhart, L., Laffer, A., & Miranda, D. (2022). Working with affective computing: Exploring UK public perceptions of AI enabled workplace surveillance. *arXiv preprint arXiv:2205.08264*.

Urquhart, L., & Miranda, D. (2022). Policing faces: the present and future of intelligent facial surveillance. *Information & Communications Technology Law*, *31*(2), 194-219. https://doi.org/10.1080/13600834.2021.1994220

Urquhart, L., Miranda, D., & Podoletz, L. (2022a). Policing the smart home: The Internet of Things as 'Invisible Witnesses'. *Information Polity*.

Urquhart, L., Miranda, D., & Podoletz, L. (2022b). Policing the smart home: The internet of things as 'invisible witnesses'. *Information Polity*, *27*, 233-246. https://doi.org/10.3233/IP-211541

Urueña, A., E Arenas, Á., & Hidalgo, A. (2018). Understanding workers' adoption of productivity mobile applications: a fuzzy set qualitative comparative analysis (fsQCA). *Economic research-Ekonomska istraživanja*, *31*(1), 967-981.

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, *46*(2), 186-204. https://doi.org/10.1287/mnsc.46.2.186.11926

Vu, H. T., & Lim, J. (2021). Effects of country and individual factors on public acceptance of artificial intelligence and robotics technologies: a multilevel SEM analysis of 28-country survey data. *Behaviour & Information Technology*, 1-14. https://doi.org/10.1080/0144929X.2021.1884288

Vuong, Q.-H. (2022). *A New Theory of Serendipity: Nature, Emergence and Mechanism*. Walter De Gruyter GmbH.

Vuong, Q.-H., Bui, Q.-K., La, V.-P., Vuong, T.-T., Nguyen, V.-H. T., Ho, M.-T., Nguyen, H.-K. T., & Ho, M.-T. (2018). Cultural additivity: behavioural insights from the interaction of Confucianism, Buddhism and Taoism in folktales. *Palgrave Communications*, *4*(1), 143. https://doi.org/10.1057/s41599-018-0189-2

Vuong, Q.-H., Ho, M.-T., Nguyen, H.-K. T., Vuong, T.-T., Tran, T., Hoang, K.-L., Vu, T.-H., Hoang, P.-H., Nguyen, M.-H., Ho, M.-T., & La, V.-P. (2020). On how religions could accidentally incite lies and violence: folktales as a cultural transmitter. *Palgrave Communications*, *6*(1), 82. https://doi.org/10.1057/s41599-020-0442-3

Vuong, Q.-H., La, V.-P., Nguyen, M.-H., Ho, M.-T., Ho, M.-T., & Mantello, P. (2020). Improving Bayesian statistics understanding in the age of Big Data with the bayesvl R package. *Software Impacts*, *4*, 100016. https://doi.org/https://doi.org/10.1016/j.simpa.2020.100016

Vuong, Q., & Parry, G. (2021). Deep Affect: Using objects, scenes and facial expressions in a deep neural network to predict arousal and valence values of images. *PsyArxiv*. https://europepmc.org/article/ppr/ppr354382

Vuong, Q. H. (2016). Global mindset as the integration of emerging socio-cultural values through mindsponge processes: A transition economy perspective. In J. Kuada (Ed.), *Global Mindsets: Exploration and Perspectives* (pp. 109-126). Routledge.
Vuong, Q. H. (2016). Global mindset as the integration of emerging socio-cultural values through mindsponge processes: A transition economy perspective. In J. Kuada (Ed.), *Global Mindsets: Exploration and Perspectives* (pp. 109-126). Routledge.
Vuong, Q. H., & Napier, N. K. (2015). Acculturation and global mindsponge: an emerging market perspective. *International Journal of Intercultural Relations*, *49*, 354-367.

Wakana, S. (2021). *The failed promise of workplace equality in japan: A view from the trenches*. Retrieved January 06 from https://www.nippon.com/en/in-depth/d00653/

Wang, Z., Ho, S.-B., & Cambria, E. (2020). Multi-level fine-scaled sentiment sensing with ambivalence handling. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *28*(04), 683-697.

Wang, Z., & Yu, Q. (2015). Privacy trust crisis of personal data in China in the era of Big Data: The survey and countermeasures. *Computer Law & Security Review*, *31*(6), 782-792. https://doi.org/https://doi.org/10.1016/j.clsr.2015.08.006

White, D., & Galbraith, P. W. (2019). Japan's Emerging Emotional Tech [https://doi.org/10.1111/AN.1070]. *Anthropology News*, *60*(1), e41-e47. https://doi.org/https://doi.org/10.1111/AN.1070

Williamson, B. (2017a). Decoding ClassDojo: psycho-policy, social-emotional learning and persuasive educational technologies. *Learning, Media and Technology*, *42*(4), 440-453. https://doi.org/10.1080/17439884.2017.1278020

Williamson, B. (2017b). Moulding student emotions through computational psychology: affective learning technologies and algorithmic governance. *Educational Media International*, *54*(4), 267-288. https://doi.org/10.1080/09523987.2017.1407080

Williamson, B. (2021). Psychodata: disassembling the psychological, economic, and statistical infrastructure of ‘social-emotional learning’. *Journal of Education Policy*, *36*(1), 129-154. https://doi.org/10.1080/02680939.2019.1672895

Woo, Y. (2022). Homogenous Japan? An Empirical Examination on Public Perceptions of Citizenship. *Social Science Japan Journal*, *25*(2), 209-228. https://doi.org/10.1093/ssjj/jyac001

Woolley, S. C., & Howard, P. N. (2018). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford University Press.

Wright, J. (2019). Robots vs migrants? Reconfiguring the future of Japanese institutional eldercare. *Critical Asian Studies*, *51*(3), 331-354. https://doi.org/10.1080/14672715.2019.1612765

Wright, J. (2021). Suspect AI: Vibraimage, Emotion Recognition Technology and Algorithmic Opacity. *Science, Technology and Society*, 09717218211003411. https://doi.org/10.1177/09717218211003411

Wylie, C. (2019). *Mindf*ck: Inside Cambridge Analytica's plot to break the world*. Profile Books.

Wynn, K., Bloom, P., Jordan, A., Marshall, J., & Sheskin, M. (2017). Not Noble Savages After All: Limits to Early Altruism. *Current Directions in Psychological Science, 27*(1), 3-8. https://doi.org/10.1177/0963721417734875

Yamada, K., Wakaizumi, K., Kubota, Y., Murayama, H., & Tabuchi, T. (2021). Loneliness, social isolation, and pain following the COVID-19 outbreak: data from a nationwide internet survey in Japan. *Scientific Reports*, *11*(1), 18643. https://doi.org/10.1038/s41598-021-97136-3

Yankson, B., Iqbal, F., & Hung, P. C. K. (2017). Privacy Preservation Framework for Smart Connected Toys. In J. K. T. Tang & P. C. K. Hung (Eds.), *Computing in Smart Toys* (pp. 149-164). Springer International Publishing. https://doi.org/10.1007/978-3-319-62072-5_9

Yoshida, K. (2022). *FamilyMart deploys restocking robots to ease staffing crunch*. Retrieved October 01 from https://asia.nikkei.com/Business/Technology/FamilyMart-deploys-restocking-robots-to-ease-staffing-crunch

Yu, K.-H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature Biomedical Engineering*, *2*(10), 719-731. https://doi.org/10.1038/s41551-018-0305-z

Yue, L., Chen, W., Li, X., Zuo, W., & Yin, M. (2019). A survey of sentiment analysis in social media. *Knowledge and Information Systems*, *60*(2), 617-663. https://doi.org/10.1007/s10115-018-1236-4

Zuboff, S. (2019). Surveillance Capitalism and the Challenge of Collective Action. *New Labor Forum*, *28*(1), 10-29. https://doi.org/10.1177/1095796018819461