Winter December, 2020

Embodiment and Algorithms for Human-Robot Interaction

Yueh-Hsuan Weng, *Tohoku University*
Chih-Hsing Ho, *Academia Sinica*

34

Embodiment and Algorithms for Human–Robot Interaction

Yueh-Hsuan Weng and Chih-hsing Ho[*]

INTRODUCTION

To many people, there is a boundary which exists between artificial intelligence (AI), sometimes referred to as an intelligent software agent, and the system which is controlled through AI primarily by the use of algorithms. One example of this dichotomy is robots which have a physical form, but whose behavior is highly dependent on the "AI algorithms" which direct their actions. More specifically, we can think of a software agent as an entity which is directed by algorithms that perform many intellectual activities currently done by humans. The software agent can exist in a virtual world (for example, a bot) or can be embedded in the software controlling a machine (for example, a robot). For many current robots controlled by algorithms, they represent semi-intelligent hardware that repetitively perform tasks in physical environments. This observation is based on the fact that most robotic applications for industrial use since the middle of the last century have been driven by algorithms that support repetitive machine motions. In many cases, industrial robots which typically work in closed environments, say, for example, factory floors, do not need "advanced" techniques of AI to function because they perform daily routines with algorithms directing the repetitive motions of their end effectors. However, lately, there is an emerging technological trend which has resulted from the combination of AI and robots, which, by using sophisticated algorithms, allows robots to adapt complex work styles and to function socially in open environments. We may call these merged technological products "embodied AI," or in a more general sense, "embodied algorithms."

Embodied versions of systems operating with algorithms lead to new possibilities for human–robot interaction (HRI), such as allowing for a wider range of possibilities for the development of interactive interfaces. These new interfaces, made possible by the use of algorithmic-driven machines, are beginning to challenge established areas of law. For example, based on the pervasiveness of algorithms in different technologies (such as the Internet and robots), it is possible to design a networked humanoid robot to service human needs. Although such a system might, at first glance, look like a stand-alone system, in actuality its perception and decision-making abilities are tied to the networked smart environment it occupies. We posit that such an intelligent and networked system will bring new challenges to established areas of law, and specifically to an emerging law of algorithms. As

[*] This work was mainly supported by JSPS KAKENHI Grant Number 19K13579, and partially supported by the Transatlantic Technology Law Forum, Stanford Law School regarding "Healthcare Robots: A Comparative EU-US Data Protection Analysis" and Academia Sinica's "Data Safety and Talent Cultivation Project – Subproject: Artificial Intelligence in the Field of Medicine." In addition, we would like to express our very great appreciation to Professor Woodrow Barfield for his kind suggestions regarding this chapter.

one example, a new threat to data protection and privacy will be possible when algorithmic-driven robots are connected to cloud computing,[1] and also when algorithms convert speech to text using remote servers[2] (especially in the access control of ubiquitous robots (Ubi-Bots)).[3] On this point, we note that current legislation for information privacy protection are data-driven, yet robots controlled by algorithms perform in various ways (that is, not always data-driven) – for example, when collecting personal information, or when interacting with humans. Thus, such systems are stretching the boundaries of current data protection and privacy law. Clearly, there is a legal gap between existing privacy and data collection law and the abilities of algorithmic-driven systems to collect data that is personal in nature – we explore that gap in this chapter.

In this chapter, we discuss the idea that the embodiment of an AI entity has particular significance for a law of algorithms because embodied forms of AI using algorithms may lead to the collection of personal information from users in social interactions. We note that "embodiment," as a feature of intelligent robots, has rarely been mentioned in the field of privacy and data protection law. Hence, in this chapter, we investigate the relationship between embodied forms of systems that are controlled by algorithms and their effect on privacy and data protection law. As a technology which highlights the points we wish to make, we will focus on healthcare robots because they have the ability to engage humans in multiple ways using different types of social interactions. The motivation for the chapter is to determine how the use of an algorithmic-driven system, such as a robot providing healthcare services and interacting socially with people, will influence privacy and data protection in HRIs.

APPLICATIONS OF ALGORITHMS

We start with general observations. Algorithms are not only used to direct the actions of robots, but are a critical feature of people's daily lives. For example, the algorithms directing e-commerce recommendation systems show users different products that consumers might be interested in by connecting records from other users who have similar interests. From a social science perspective, such systems are inspired by the idea of six degrees of separation and theory of social networks.[4] Additionally, targeted advertising also uses algorithms to determine suitable adverts for individuals based on specific users' online consumer behaviors. Further, encryption algorithms are used to ensure the security of e-commerce and digital financial transactions. And advances in algorithmic-driven systems have also shown success in games requiring high levels of human cognitive skill. For example, in 2016, DeepMind's AlphaGO AI software defeated Go legend Sedol Lee.[5]

[1] U. Pagallo, Robots in the Cloud with Privacy: A New Threat to Data Protection? (2013) 29 *Comput. Law Secur. Rep.* 501.

[2] J. F. Hoorn, Mechanical Empathy Seems Too Risky. Will Policymakers Transcend Inertia and Choose Robot Care? The World Needs It, in G. Dekoulis (ed.), *Robotics: Legal, Ethical and Socioeconomic Impacts* (IntechOpen, 2017).

[3] Y. H. Weng and S. T. H. Zhao, The Legal Challenges of Networked Robotics: From the Safety Intelligence Perspective, in M. Palmirani, U. Pagallo, P. Casanovas, and G. Sartor (eds.), *AI Approaches to the Complexity of Legal Systems: Models and Ethical Challenges for Legal Systems, Legal Language and Legal Ontologies, Argumentation and Software Agents* (Springer, 2012), Vol. 7639, pp. 61–72.

[4] D. J. Watts, *Six Degrees: The Science of a Connected Age* (W. W. Norton, 2004).

[5] Artificial intelligence: Google's AlphaGo Beats Go Master Lee Se-dol, *BBC* (March 12, 2016), www.bbc.com /news/technology-35785875.

Recent techniques in the design of algorithms are leading to systems with the capability to learn and thus become involved in numerous applications which require social interactions with humans. Based on advances in computing and Internet technology, for example, in cloud computing and the collection of big data, learning algorithms are now becoming more widely used in information technology fields. For example, the Chicago Police Department uses the Strategic Subject Algorithm developed by the Illinois Institute of Technology to create a risk assessment score called "Strategic Subject List (SSL)" which ranges from 1 to 500.[6] With the assistance provided by this algorithm, the Chicago police allocate resources with the goal to prevent future gun violence and to direct their limited policing resources into prioritizing high-risk areas. According to the data reflecting the use of the algorithm, the shooting rates in Chicago have decreased 39 percent after implementation of the program.[7] Finally, in the field of transportation, algorithms with the capability to learn have been utilized in improving self-driving cars' visual perception of the real world, and algorithms used in the medical field have improved recognition rates for certain diseases.

While the above examples of the use of algorithms have resulted in challenges to established areas of law, an evolving issue (which is the focus of this chapter) concerns potential legal consequences associated with using learning algorithms in society. As an example, COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) represents software which uses machine-learning algorithms to assist judges in criminal cases to evaluate a defendant's risk of committing another crime. The algorithmic-based system has been used to assess more than 1 million offenders since its development in 1998.[8] And in the medical field, algorithms are being used to help doctors analyze complex data and to provide suggestions for suitable donors for kidney transplantation, in a program called "paired kidney donation" (more on this below). The short history of paired kidney donation started in 2000, and the program continues to handle more cases each year. For example, in 2018, 12 percent of living kidney donations came from paired donors,[9] based on matches between donor and patient suggested by algorithms.

Although algorithms used for decision support systems to assist human experts in many professional fields have shown great potential over the past two decades, recent trends in deep learning algorithms raise a potential concern about the accountability of using learning algorithms. One main criticism of deep learning algorithms is based on their "untransparency" or "black box" characteristics. After autonomous training (in the case of unsupervised deep learning), deep learning techniques generate new outputs from additional inputs fed into the algorithms, but the cause–effect relationships between inputs and outputs may be unclear to humans in the system or those affected by the algorithms' decisions. This unforeseeable situation may lead to an accountability gap among stakeholders resulting from the design and use of technology based on deep learning algorithms. For example, how can we ensure the COMPAS software, in the example above, isn't biased in the way that its algorithms calculate scores? On this point, consider again the example of paired kidney donation, which also raises ethical questions.

[6] Strategic Subject List (SSL), Chicago Data Portal, https://data.cityofchicago.org/Public-Safety/Strategic-Subject-List/4aki-r3np.

[7] J. Gunter, Chicago Goes High-Tech in Search of Answers to Gun Crime Surge, *BBC* (June 19, 2017), www.bbc.com/news/world-us-canada-40293666.

[8] J. Dressel and H. Farid, The Accuracy, Fairness, and Limits of Predicting Recidivism (2018) 4 *Sci. Adv.* eaao5580.

[9] C. Purtill, How AI Changed Organ Donation in the US, *QUARTZ* (September 10, 2018), https://qz.com/1383083/how-ai-changed-organ-donation-in-the-us/.

Suppose there are two patients, A and B, both experiencing an emergency medical crisis, who coincidentally match donor C's kidney. Further, suppose that only A or B can receive C's kidney donation, in which case, the patient not receiving the donation will suffer negative health consequences. From an ethics point of view, should an algorithm (mathematically) decide who receives a transplant and thus who may live or die? Now further consider that A is a retired national hero, but with many serious chronic diseases who is thought to have a short life span even if they receive the kidney transplantation. In contrast, B is known to be a habitual criminal, but a single parent with a 5-year-old daughter and a potential long life ahead with a successful transplant. In this example, should society entrust an algorithm with the power to decide who will live based on its calculation from the parameters it receives? Any bias or misjudgment, or lack of ethical clarity shown by the algorithm will have tragic consequences. This particular example, with its ethical ramifications, raises people's awareness about the issue of allowing algorithms to perform human decisions in areas involving moral and ethical questions and more generally the idea that perhaps algorithms should be regulated.

Among others, the above examples show that one key debate resulting from the use of algorithms concerns the accountability of algorithms to various stakeholders. Legal scholar Frank Pasquale and mathematician Cathy O'Neil have analyzed, using a macro perspective, how an accountability gap is formed when the operation of many social issues intersect and are taken over by big data and algorithms.[10] In addition to the accountability of algorithms, issues of importance for a law of algorithms also include the regulation of algorithmic-driven systems and the legal protection of AI-generated works, as well as the challenges to freedom of speech, privacy, and liability for damages when the use of algorithms leads to damages.

EMBODIED ALGORITHMS WITHIN ROBOTS

Another important issue for a law of algorithms, as mentioned above, is privacy and data protection for embodied AI (in this chapter, using healthcare robots as an example). In the context of an aging population, particularly in East Asia, Europe, and North America, it is evident that healthcare robots will be used to support the elderly in walking, climbing, and living independently, as well as providing non-physical support, such as daily communication and providing necessary information for day-to-day life, including social contact to alleviate loneliness. While healthcare robots will engage in important functions for society, especially in the background of an aging population, from the perspective of law, there is an important question which revolves around the issue of whether the embodiment of the social robot itself will lead to issues of privacy and data protection. On that point, this chapter discusses robot embodiment in a social context, an issue that, thus far, has been neglected in legal scholarship and specifically for a law of algorithms. We argue that healthcare robots are a prime example of embodied algorithmic-driven systems, due to their high contact levels with human beings in the form of embodied physical machines whose behavior is directed by algorithms. This raises a concern that data protection and privacy issues will result from the use of healthcare robots (controlled by algorithms) when they are deployed into society and interacting with humans in social settings.

[10] F. Pasquale, *The Black Box Society* (Harvard University Press, 2015); C. O'Neil, *Weapons of Math Destruction* (Broadway Books, 2016).

Based on recent advances in AI technologies, robots that are directed by algorithms are increasingly common within human society, and this has raised three important questions which relate to healthcare robotics: (1) How will the use of algorithmic-driven social robots that are in daily use (such as in the healthcare industry) influence privacy in HRIs? (2) How will the use of intelligent robots in healthcare impact current data protection laws? and (3) How can we apply the concept of "privacy by design" to the design process of healthcare robots, with the goal of covering the gap resulting from the use of embodied healthcare robots and data protection? To answer these questions, we will start by discussing the relevant philosophy and law which relate to algorithms and embodiment.

ALGORITHMS AND EMBODIMENT

We start with a historical perspective. René Descartes's Mind-Body Dualism discussion in the seventeenth century has had a huge influence on the thinking of modern Western societies. According to his theory, mind and body are two independent parts existing in one single person. To Descartes, the only connection between mind and body was the pineal gland, which is where he posited a person's soul exists.[11] Unlike traditional Chinese medicine, which focuses on a holistic view of the body to treat patients, Western medicine systematically developed many divisions of the body to provide patients with medical care in an efficient although system-by-system manner. However, under the Western medicine view, a psychophysical disorder may manifest itself as a combination of the mind and body which is thought to be suffering some malady. This type of disorder refers to physical symptoms which may manifest themselves due to psychological or social factors; for example, hypertension, a peptic ulcer, or a migraine headache. As a mind-body conundrum, also consider the phantom limb pain phenomena experienced by patients who lost their arm or leg – this represents another example of the inseparability of mind and body.[12] Clearly, the complex relationship between mind and body is not easily described by representing it as a dichotomy. The early stages of AI research have similarly experienced the conceptual difficulty associated with the mind-body duality, in that they focus on creating an artificial mind by imitating a human's high-level reasoning and problem-solving capabilities, but neglect the importance of a body in mediating an artificial agent's intellectual behaviors in the real world.

In the 1980s, the issue of embodiment became an important topic in AI research and development. We provide here a representative example from Rodney Brooks's proposal on a "subsumption architecture," which he explained used a "bottom-up" approach for intelligence without the use of symbolic knowledge representation. The subsumption architecture decomposes a machine's behaviors into several sub-behaviors and based on that division designs different reactive control layers in order to allow robots to respond to unstructured environments in real time.[13] One major difference between embodied intelligent robots and traditional applications of GOFAI (good old-fashioned AI)[14] is that behaviors of the former

[11] R. Descartes, *The Philosophical Writings of Descartes*, J. Cottingham, R. Stoothoff, D. Murdoch, and A. Kenny (trans.) (Cambridge University Press, 1984–91).

[12] Y. Oouchida and S. Izumi, Imitation Movement Reduces the Phantom Limb Pain Caused by the Abnormality of Body Schema, ICME International Conference on Complex Medical Engineering (CME), Kobe, Japan (July 1–4, 2012).

[13] R. A. Brooks, A Robust Layered Control System for a Mobile Robot (1986) 2 *IEEE J. Robot. Autom.* 14.

[14] J. Haugeland, *Artificial Intelligence: The Very Idea* (MIT Press, 1986).

have not been explicitly programmed into a system or an agent, such that "emergence" occurs.

Emergence can be described as a phenomenon which occurs when an entity taken as a whole is observed to have properties which its parts, if considered separately, do not have. Such properties or behaviors emerge when the parts of an entity interact as a combined or integrated object.[15] When this concept is applied to an embodied form of an algorithmic-driven object, the physical characteristics of the object, for example, a robot, may show new intelligent properties or behaviors separately that did not exist in the software controlling the individual parts.

According to another pioneer on intelligence and embodiment, Rolf Pfeifer, there are three different types of emergence, which include: (1) a global phenomenon arising from collective behavior; (2) individual behavior resulting from an agent's interactions with the environment; and (3) the emergence of behavior from one time scale to another.[16] From these three points we can see that the implications of embodiment for intelligence are more involved than people designing robots originally thought. Except for the well-known example of the passive dynamic walking robots created at Cornell University's Biorobotics and Locomotion Lab,[17] a multi-agent system's complex dynamic in artificial life, and a humanoid's embodied cognition in developmental robotics are examples of emergence from individual, collective, and time-crossing perspectives. The phenomenon of emergence also represents another aspect of AI that cannot be reproduced under a brain-centrism's viewpoint of creating intelligence via software agents or intelligent entities without a body. As discussed by law and robotics expert Ryan Calo, robots' emergence enables a particular embodied form of AI, to react in unstructured environments in ways beyond merely repetitive motions;[18] this complexity of behavior shown in the real world as directed by algorithms impacts the foreseeability of machine behaviors and therefore impacts its relationship with the law.

As mentioned before, the emerging field of the law of algorithms, in our view, mainly concerns the legal issues related to algorithmic accountability, and under this huge umbrella of AI governance it extends to many sub-issues which trigger the following questions: How can we make algorithmic decision-making more transparent? What kinds of machine-learning-based AI applications can be used for dealing with issues related to human rights and property protection, such as healthcare, public security, and national defense? Can we grant copyright to an AI art creator? Should governments regulate online "speech" generated by algorithms?[19] and What are the privacy and data protection issues resulting from information systems using autonomous functions produced by algorithms? As yet, there is no formal approach within law to the above-mentioned issues relating to embodiment. The authors believe this is not because embodiment itself is not an important topic for law; on the contrary, the reason can more likely be attributed to the lack of accountability between

[15] Emergence, *Wikipedia*, https://en.wikipedia.org/wiki/Emergence.

[16] R. Pfeifer and J. Bongard, *How the Body Shapes the Way We Think: A New View of Intelligence* (MIT Press, 2007), p. 85.

[17] Cornell University's Biorobotics and Locomotion Lab, http://ruina.tam.cornell.edu/research/; M., Hoffmann and R. Pfeifer, The Implications of Embodiment for Behavior and Cognition: Animal and Robotic Case Studies, in W. Tschacher and C. Bergomi (eds.), *The Implications of Embodiment: Cognition and Communication* (Imprint Academic, 2012), pp. 31–58.

[18] M. R. Calo, Robotics and the Lessons of Cyberlaw (2015) 103 *Calif. Law Rev.* 513.

[19] S. M. Benjamin, Algorithms and Speech (2013) 161 *Univ. Pa. Law Rev.* 6.

embodied AI and the accountability of algorithms. We argue that a law of algorithms should consider embodiment based on the reasons discussed above and presented below.

First, consider the difference between a software chatbot (digital embodiment) and an intelligent robot (physical embodiment) which has relevance for the cause and effect derived from the environment each operates in, which in turn influences their behaviors. A chatbot as a software agent exists in a digital, or virtual, environment, thus there is no gap between the words it speaks in VR and the decisions it makes in VR. On the other hand, the output actions of an intelligent robot may not be equivalent to its original plan of action due to the mediating effect of its physical environment, which is a feature of emergence. These factors influence the final output of a robot's decision-making, which is mediated by the physical environment; thus, a problem may result if a brain-centrism's viewpoint is used to judge legal liability for a tort action based on the conduct of a robot. In other words, to fully realize algorithmic accountability for human and machine, we should not merely consider the "black box"[20] and "Open-Texture"[21] aspects of robots, but, in addition, a third factor emerging from algorithmic processes: "emergence."

Second, although the method of computational simulation is used for many applications such as modeling the dynamics of infectious disease transmission, or to model the voter's behavior in democratic elections, computational simulation cannot be used to verify the safety of autonomous systems such as self-driving cars or intelligent service robots. This has implications for personal safety with embodied forms of AI, that is, for the design of "safety-critical systems." We posit that the reason why computational simulations cannot be used to verify the safety of autonomous systems is based on a human factor and a non-human factor.

From a human factor perspective, for algorithmic-driven entities we should consider adding more testing and certification processes to design safety-critical systems in order to ensure that they are reliable enough to resist cyberattack from human hackers.[22] As for the non-human factor, a threat to safety-critical systems concerns "modeling error," which refers to the inconsistency between software and the hardware comprising the same system. In other words, from a software perspective, the system may function normally, but the final performance of the system in the real world may not match how it should be performing according to its software instructions. There are many reasons which may cause system modeling error. For example, consider an accident involving a smart car which is equipped with an Eye-Sight autonomous braking system. Eye-Sight is an active safety system using cameras to evaluate ground conditions on the road and then decides whether or not to activate its brakes when the car is too close to front objects or other objects which would create an emergency. In this example, the accident was due to a misjudgment on whether to activate the brake when white snow on the ground reflected sunlight which then "confused" the car's sensors.[23]

To ensure the accountability of safety-critical systems, we have to run them through an empirical testing area like a regulatory sandbox for autonomous systems. Examples include

[20] B. Waltl and R. Vogl, Explainable Artificial Intelligence – the New Frontier in Legal Informatics, International Legal Informatics Symposium (IRIS), Salzburg, Austria (February 2018).

[21] Y. H. Weng, The Study of Safety Governance for Service Robots: On Open-Texture Risk, Peking University PhD Dissertation (2014).

[22] S. O'Sullivan, N. Nevejans, C. Allen, *et al.*, Legal, Regulatory, and Ethical Frameworks for Development of Standards in Artificial Intelligence (AI) and Autonomous Robotic Surgery (2019) 15 *Int. J. Med. Robot.* e1968.

[23] Y.-H. Weng and D. Hillenbrand, The Intelligentization of Automobiles: Smart-Cars, Robo-Cars and their Safety Governance (2014) 4 *J. Sci. Technol. Law* 632.

Japan's "Tokku" special zone[24] or the United States' "Faux Downtown."[25] When creating algorithmic-driven robotic technologies and introducing them into the real world, conflicts are unavoidable given current regulations.[26] However, if we can properly use special zones as a shock buffer via deregulation, it also helps lawmakers fill the accountability gap when developing a law for algorithms that relates to the design of safety-critical systems.

Third, embodiment brings new impacts to data governance and privacy to bear in HRI. As mentioned above, we see healthcare robots as the physical extension of an intelligent agent or "embodied AI." Suppose there are issues generated by the data used for the training process for machine learning and hackers are thus able to attack the system; this could result in healthcare robots behaving abnormally. In the worst-case scenario, the compromised system might physically injure a patient. In addition to security, another concern is the issue of privacy; in that regard, a key question is whether the body in which the AI finds itself influences the privacy protection which is necessary between HRIs.

Concerning robotic embodiment, Debora Zanatto and colleagues used two humanoid-appearing robots – iCub and Scitos G5 (the former having more of a human-like shape) to test whether there were differences in robot credibility when humans were interacting with the robots. They found that subjects were more likely to have social interactions with robots when they were more anthropomorphic in appearance.[27] In additional, Pavia and colleagues conducted a study that dealt with embodiment and empathy by comparing a virtual agent to a robotic agent. They found that it was more challenging for a robotic agent to perceive the user's affective state than for a virtual agent, because the human user sits in front of a screen which displayed a virtual agent; whereas, with a physical robot agent, the relative location between the user and robotic agent can be open-ended.[28]

To achieve a deeper model of artificial empathy, we might need to borrow the viewpoint developed from the field of "Cognitive Developmental Robotics," which refers to "new understandings of how humans' higher cognitive functions develop by means of a synthetic approach that developmentally constructs cognitive functions."[29] Based on these factors, we may infer that when robots are more anthropomorphic, there are risks that humans might disclose more personal and sensitive information during interactions with the robots. Another threat to privacy is that robots may require higher perception capabilities in order to utilize the complex interactions between humans and physical robots. But we are at the beginning stages of our study of privacy in HRI; in order to explore in more depth how the factor of embodiment influences a developing law of algorithms, especially from a privacy perspective, in the following section we will limit our discussion to healthcare robots, given their close social contact with humans.

[24] Y. H. Weng, Y. Sugahara, K. Hashimoto, and A. Takanishi, Intersection of "Tokku" Special Zone, Robots, and the Law: A Case Study on Legal Impacts to Humanoid Robots (2015) 7 *Int. J. Soc. Robot.* 841.

[25] W. Jones, University of Michigan to Open Robo Car Urban Test Track in the Fall, *IEEE Spectrum* (June 10, 2014), http://spectrum.ieee.org/cars-that-think/transportation/self-driving/university-of-michigan-to-open-robo-car-test-track-in-the-fall.

[26] Y. H. Weng, Robot Law 1.0: On Social System Design for Artificial Intelligence, in W. Barfield and U. Pagallo (eds.), *Research Handbook on the Law of Artificial Intelligence* (Edward Elgar, 2018).

[27] D. Zanatto, M. Patacchiola, J. Goslin, and A. Cangelosi, Priming Anthropomorphism: Can the Credibility of Humanlike Robots Be Transferred to Non-Humanlike Robots?, in Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, New Zealand (March 7–10, 2016).

[28] A. Pavia, I. Leite, H. Boukricha, and I. Wachsmuth, Empathy in Virtual Agents and Robots: A Survey (2017) 7 *ACM Trans. Interact. Intell. Syst.* art. 11.

[29] M. Asada, Towards Artificial Empathy: How Can Artificial Empathy Follow the Developmental Pathway of Natural Empathy? (2015) 7 *Int. J. Soc. Rob.* 19.

HEALTHCARE ROBOTS AND ALGORITHMS

The growth of an aging population is a critical challenge to many developed nations and has been a topic of discussion in recent years. For example, in Japan, 35,152,000 people or 27.7 percent of the population are above 65 years of age,[30] and for that reason, many public policies have been discussed in Japan with regard to healthcare robots. For example, the "New Robot Strategy: Japan's Robot Strategy – Vision, Strategy, Action Plan" is a five-year mega-plan which aims to use robotics technology to create a new industrial revolution, and to expand its domestic industrial output of robotics. The application of robots in the healthcare industry has been mentioned among many strategic applications in Japanese policy guidelines.[31]

Not only in Japan, but many countries are considering the use of robot technology to alleviate the increased aging crisis, especially by developing robotic applications in healthcare, or what is termed "healthcare robots." Because healthcare robots can perform many tasks designed to assist elderly people, they are proposed as a viable solution to addressing a nation's aging population. One example is using healthcare robots for people who suffer dementia. Denise Hebesberger and colleagues in 2016 used an autonomous humanoid robot, Scitos, to support people in physical therapy who suffered advanced dementia; they found that use of a robot enhanced the participants' motivation to participate in therapy.[32] In addition to using robots to assist training for specific groups who suffer dementia, healthcare robots can also be used for companionship, entertainment, monitoring, walk support, and navigation for the elderly. However, there is a question in the literature about the ambiguousness of the term "healthcare robots." As Woodrow Barfield has discussed, precisely defining the technology to be regulated is often a challenge to policymakers not trained in technology, especially when the technology is an emerging form of AI and robotics.[33]

However, there are many similar terms discussing healthcare robots found on the Internet; for example, personal care robots, medical robots, socially assistive robots, etc. Thus, it may be of value to review those terms here. Although they seem similar and even have some overlap, from a perspective of law, conflicts may arise if the terms are not carefully distinguished. The term "assistive robotics" as defined by Feil-Seifer and Mataric is a general description of robots with the capability to provide aid or support to a human user in a broad range of environments, such as care centers, hospitals, and homes.[34] Furthermore, Feil-Seifer and Mataric proposed a common term for the terms assistive robotics and socially interactive robotics, which they called "socially assistive robotics" that provide assistance to human users using methods of social interaction.[35] In addition, Armi Ariani and colleagues provided yet another different definition for assistive robotics. They divided assistive robotics into

[30] Statistics Bureau, Ministry of Internal Affairs and Communication (MIC), Population Statistics Report, Japan (2017), www.stat.go.jp/data/jinsui/2017np/pdf/2017np.pdf.

[31] Y. H. Weng and Y. Hirata, Ethically Aligned Design for Assistive Robotics, IEEE International Conference on Intelligence and Safety for Robotics (ISR) (2018), pp. 286–90.

[32] D. Hebesberger, T. Körtner, J. Pripfl, and C. Gisinger, Lessons Learned from the Deployment of a Long-Term Autonomous Robot as Companion in Physical Therapy for Older Adults with Dementia, in Proceedings of the 11th ACM/IEEE International Conference on Human–Robot Interaction (HRI), Christchurch, New Zealand (March 7–10, 2016).

[33] W. Barfield, Towards a Law of Artificial Intelligence, in W. Barfield and U. Pagallo (eds.), *Research Handbook on the Law of Artificial Intelligence* (Edward Elgar, 2018).

[34] D. Feil-Seifer and M. Mataric, Defining Socially Assistive Robotics, in Proceedings of the 9th IEEE International Conference on Rehabilitation Robotics, Chicago, IL, United States (June 28–July 1, 2005).

[35] T. Fong, I. Nourbakhsh, and K. Dautenhahn, A Survey of Socially Interactive Robots (2003) 42 *Rob. Auton. Syst.* 143.

"rehabilitation robotics" that can be used as treatment devices such as prostheses, tools for rehabilitation therapies, and/or physical and mobility supports; and "socially assistive robotics" that can be used for social interaction with people, including companionship and service.[36]

The Japanese government has also provided a guideline for classifying healthcare robots with a focus on daily assistive tasks and nursing care for the elderly. The guideline is offered by the Japanese Ministry of Economy, Trade, and Industry (METI) and the Ministry of Health, Labor, and Welfare (MHLW), and includes six applications such as "Bed-Transfer Assist," "Walking Assist," "Excretion that Assist," "Monitoring and Communication Assist," "Bath Assist," and "Nursing Care Business Assist."[37] Each of these has different algorithmic requirements. This classification has been used as a guideline for domestic manufacturers to develop robotic products used for healthcare, but most products currently operate with limited autonomy or only perform simple and repetitive healthcare functions.[38]

Considering that embodied AI operating in social environments is the subject of this chapter, we also discuss emerging healthcare robots that we expect to see operating in homes in the near future. Thus, our definition of "healthcare robots" includes autonomous service robots which have the goal of promoting or monitoring health, while assisting with the above six kinds of care tasks that are currently difficult to perform due to the health problems experienced by the elderly, or due to the difficulty of preventing the further health decline of the elderly.[39]

Note autonomy here refers to a "third existence" meaning that machines will resemble living beings in appearance and behavior, but they will not be self-aware.[40] In addition, some people may wonder if robotic surgery systems and other AI-based medical systems belong to the category of healthcare robots. We believe that such medical devices should meet higher regulation requirements, such as those provided by the Federal Drug Administration for medical devices;[41] therefore, they are not the focus of our discussion in this chapter. In the next section, we discuss embodiment and privacy in HRI by choosing three examples using the above six applications for healthcare robots.

EMBODIMENT AND PRIVACY IN HUMAN–ROBOT INTERACTION

The importance of HRI in a legal context will be apparent in the near future when the embodiment characteristics of AI connect with intelligent healthcare services. Wainer's (2006) research group, some time ago, started using an empirical research approach to investigate embodiment in HRI. By running experiments, they found that compared to virtual animation robots, people remain for a longer time period when interacting with

[36] A. Ariani, V. Kapadia, A. Talaei-Khoei, *et al.*, Challenges in Seniors Adopting Assistive Robots: A Systematic Review (2016) 6 *Int. Technol. Manag. Rev.* 25.

[37] METI and MHLW, Revision of the Four Priority Areas to Which Robot Technology Is to Be Introduced in Nursing Care of the Elderly, Japan (2014), www.meti.go.jp/english/press/2014/0203_02.html.

[38] Robotic Devices for Nursing Care Project, http://robotcare.jp.

[39] H. Robinson, B. MacDonald, and E. Broadbent, The Role of Healthcare Robots for Older People at Home: A Review (2014) 6 *Int. J. Soc. Rob.* 575.

[40] Y.-H. Weng, C.-H. Chen, and C.-T. Sun, Toward The Human-Robot Co-Existence Society: On Safety Intelligence for Next Generation Robots (2009) 1 *Int. J. Soc. Robot.* 267.

[41] K. Chinzei, A. Shimizu, K. Mori, *et al.*, Regulatory Science on AI-Based Medical Devices and Systems (2018) 7 *Adv. Biomed. Eng.* 118.

physical robots,[42] producing better results as well.[43] This difference between virtual and real robots also enables the applications of AI to be gradually entering physical human living spaces from virtual information spaces, and to provide people with healthcare services in various interactive ways. However, with robots in the home, users are unconsciously exposing themselves to higher-risk environments because their personal information is easily searched and collected by AI in the form of robots performing intelligent healthcare services. Considering that the impact of the embodiment of robots on the issue of privacy in HRIs is often overlooked, it is our view that it is necessary to rethink the relationship between embodied AI and privacy.

Ackerman has predicted that social robots could become humans' pets in the future due to two characteristics: affordability to consumers and similarity to real animals' affective interaction with humans.[44] If so, an important question would concern how robots' anthropometric or animal-like appearance would cause people to express their emotions when interacting with robots. Furthermore, would a user's personal information be "leaked unconsciously/unintentionally" when interacting with robots? Addressing this question, Tonkin's team at the University of Technology Sydney used two information interfaces, a humanoid robot and a tablet, for comparative research on human–computer interaction. The researchers found that humanoid robots were able to acquire personal information more easily from their users compared to tablets.[45] Mann and colleagues also pointed out that patients give more positive responses to instructions or messages provided by a humanoid robot compared to a tablet in some healthcare application scenarios.[46] The privacy concern associated with embodied AI in HRIs is not merely about the issue of similarity of human-like, or animal-like, physical appearances of the intelligent system in inducing a user to self-disclose their sensitive information. Another important concern is the risk that embodied AI systems might mislead users about their functionality. For example, Snackbot is an advanced service robot equipped with a 360-degree panorama lens to help it rapidly acquire environmental information from the real world. Because users are not knowledgeable of the full-frame vision perception of Snackbot, they often interact with the robot under the wrong premise; specifically, with the idea that the robot lacks the ability to see behind its back. This lack of knowledge of the robot's functionality generates other privacy risks.[47] Beyond the issue of embodied AI itself, in HRI other kinds of privacy risks emerge from the human side, or users' privacy perception, of their existing environment. This perception is related to their acceptance of emerging technology and shaped in various ways with factors including age,[48] gender, culture, health,

[42] J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Mataric, The Role of Physical Embodiment in Human–Robot Interaction, in Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN, Hatfield, Hertfordshire, United Kingdom (September 6–8, 2006), pp. 117–22.

[43] J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Mataric, Embodiment and Human-Robot Interaction: A Task-Based Perspective, in Proceedings of the 16th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN, Jeju Island, Korea (August 26–9, 2007), pp. 872–7.

[44] E. Ackerman, Robots Might Be the Necessary Future of Urban Pet Ownership (2015), http://spectrum.ieee.org/automaton/robotics/home-robots/robots-might-be-the-necessary-future-of-urban-pet-ownership.

[45] M. Tonkin, J. Vitale, S. Ojha, *et al.*, Embodiment, Privacy and Social Robots: May I Remember You?, in Proceedings of the 9th International Conference on Social Robotics, Tsukuba, Japan (November 22–24, 2017).

[46] J. A. Mann, B. A. MacDonald, I. H. Kuo, *et al.*, People Respond Better to Robots than Computer Tablets Delivering Healthcare Instructions (2015) 43 *Comput. Hum. Behav.* 112.

[47] M. K. Lee, K. P. Tang, J. Forlizzi, and S. Kiesler, Understanding Users' Perception of Privacy in Human–Robot Interaction, in Proceedings of the 6th ACM/IEEE International Conference on Human–Robot Interaction (HRI), Lausanne, Switzerland (March 6–9, 2011).

[48] J. M. Beer, C. A. Smarr, A. D. Fisk, and W. A. Rogers, Younger and Older Users' Recognition of Virtual Agent Facial Expressions (2015) 1 *Int. J. Hum. Comput. Stud.* 1.

etc.[49] Consider again that Heerink commented that people who have a higher education usually have a less positive attitude toward robots as social entities.[50] Further, MacDorman and colleagues, from a culture perspective, analyzed the social acceptance of robots used within Japanese society based on Shinto religious beliefs and found that the religious perspective had value for social interactions with humanoid-appearing robots.[51] So, from the above discussion, the investigation of the embodiment of healthcare robots in relation to privacy in HRI relates to a host of variables as mentioned previously. Among them, we chose three variables which relate to privacy in HRI: proximity, deception, and safety. We present an analysis of these three variables below.

Proximity and Privacy

Personal space refers to the proper distance (which changes from culture to culture) that people try to maintain in relation to other people in order to ensure, among others, a given level of privacy. More specifically, personal space is the space that surrounds an individual, and any intrusion into this area of physical space without an invitation may lead to a feeling of unease and a response to back off or withdraw.[52] This invisible boundary between oneself and others creates a comfort zone. However, this boundary is not stable, but dynamic, and changes based on the circumstances. In an anthropological study, Edward Hall proposed four different spatial zones in his theory of proxemics to represent the way in which people use space when communicating with one another.[53] For Hall, the four spatial zones are categorized as follows: intimate distance, personal distance, social distance, and public distance. The intimate distance refers to a physical distance that results when one feels comfortable and protected, and it can range from 0 to 6 inches to 6 to 18 inches. Personal distance is the space within which intimate relationships can be built. Social distance (4 to 12 feet) refers to the space where business and general social contact occur. And "public distance" refers to the space associated with an occasion in which people communicate with one another. In fact, different cultures may have different distances for person-to-person communication, so the proximity associated with personal space and the satisfactory feeling of privacy could be different given various cultural contexts and scenarios.

For healthcare robots, their algorithmic understanding of appropriate proximity to humans (based on sensor information) is crucial in terms of maintaining a satisfactory level of privacy for people interacting with the robot. Therefore, in the design stage of creating healthcare robots, it is important to consider "privacy spheres" in terms of spatial distance that may exist between robots and people in order for the human to feel that their privacy is not being invaded. We believe that it is critically important for roboticists to "design-in" the feature of socially aware navigation systems for healthcare robots.[54] A robot's socially aware navigation

[49] W. Wilkowska, M. Ziefle, and S. Himmel, Perceptions of Personal Privacy in Smart Home Technologies: Do User Assessments Vary Depending on the Research Method?, in T. Tryfonas and I. Askoxylakis (eds.), *Human Aspects of Information Security, Privacy, and Trust: Third International Conference, HAS 2015* (Springer, 2015), pp. 592–603.

[50] M. Heerink, Exploring the Influence of Age, Gender, Education and Computer Experience on Robot Acceptance by Older Adults, in Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Lausanne, Switzerland (2011), pp. 147–8.

[51] K. F. MacDorman, S. K. Vasudevan, and C. C. Ho, Does Japan Really Have Robot Mania? Comparing Attitudes by Implicit and Explicit Measures (2009) 23 *AI Soc.* 485.

[52] I. Altman, *The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding* (Brooks / Cole Publishing Co., 1975).

[53] E. T. Hall, *The Hidden Dimension* (Doubleday, 1910), Vol. 609.

[54] S. F. Chik, C. F. Yeong, E. L. M. Su, *et al.*, A Review of Social-Aware Navigation Frameworks for Service Robot in Dynamic Human Environments (2016) 8 *J. Telecommun. Electron. Comput. Eng.* 41.

framework should consider the interactions between persons and robots, which would be achieved by using algorithmic predictors to estimate the moving trajectory of robots in proximity to humans. In order to maximize the performance of tasks and improve skills, the robots usually need to collect and process large amounts of real-life environmental data. Japan's Advanced Telecommunications Research Institute International (ATR) conducted an experiment using a humanoid robot in a shopping mall which approached customers within Hall's estimate of social distance (4 to 12 feet) to have a conversation.[55] For those robots whose task is to guide people in shopping malls or an airport, there is a need for the robot to approach customers and passengers and occupy an appropriate social distance space to produce goodwill and encourage communication.[56] We propose that the default privacy setting for these robots should be based on previous experiences of robots when they were within testing areas investigating HRIs.

For healthcare robots providing services to the elderly or disabled persons in healthcare centers, we conclude that the proper proximity between humans and robots should be based on the personal distance estimate of 1.5 to 4 feet or even in some cases an intimate distance estimate (6 to 18 inches). However, the difference between the use of healthcare and shopping mall robots is more than the physical distance in HRI. The service of healthcare itself requires caregivers to have many close physical contacts with care receivers. How to solve the privacy challenge with embodied robots based on proximity is not only a crucial step in robot design, but also a potential concern when considering legal compliance.

One of the possible solutions to respond to the privacy challenge associated with embodied AI in the form of robots is to embed a so-called privacy filter in the early stage of design.[57] There are several factors to consider – such as locations, objects, and information – that we can use to evaluate a person's privacy concerns.[58] Different locations and objects may present different levels of privacy concerns, and the nature of the information also has relevance for privacy concerns with embodied AI – whether it is potentially sensitive information, like health and medical information, or not. These dimensions and their interactive complexity will have great impact on the design of privacy filters for limiting the information to be delivered to those who interact with social robots. Another solution can involve allowing users the opportunity to have explicit opt-outs from video surveillance or automatic video filtering for bystanders and objects.[59] With regard to what information a privacy filter ought to remove, there is a need to carry out further surveys of users' information requirements and then design a filter according to the results that are acceptable for end-users' privacy preferences and which have minimal impacts on interactions with social robots. In the case of healthcare robots, an appropriate filter should allow the robots to have the capability to evaluate environmental data collected from different locations and objects and then to distinguish whether they are situated in a public or private sphere. As a step further, a privacy filter can be designed such that it tailors different spatial zones according to the theory of proxemics for initiating a privacy-friendly communication.

[55] S. Satake, T. Kanda, D. F. Glas, *et al.*, A Robot that Approaches Pedestrians (2013) 29 *IEEE Trans. Robot.* 508.

[56] T. Kanda, M. Shiomi, Z. Miyashita, *et al.*, A Communication Robot in a Shopping Mall (2010) 26 *IEEE Trans. Robot.* 897.

[57] J. M. Janick, H. J. Locker, and R. A. Resnick, US Patent No. 6,765,550 (US Patent and Trademark Office, 2004).

[58] D. J. Butler, J. Huang, F. Roesner, and M. Cakmak, The Privacy-Utility Tradeoff for Remotely Teleoperated Robots, in Proceedings of the 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Portland, OR, United States (March 2–5, 2015).

[59] *Ibid.*

Deception and Privacy

One aspect of deception comes from the expectation gap that users may experience when they interact with human-like, or anthropomorphic, robots. Taking ROBEAR as an example, this new experimental bear-shaped nursing care robot developed by scientists from RIKEN Japan is capable of performing basic tasks of care-giving, such as providing assistance to a patient to stand up or lifting him or her from a bed to a wheelchair.[60] The ROBEAR incorporates actuator units with a low gear ratio that allows its joints to move quickly. It is also equipped with three types of sensors, including the Smart Rubber capacitance-type tactile sensors that enable the ROBEAR to exert force in a gentle way, and it can perform power-intensive tasks without endangering patients.[61] Although the bear feature and the human-like features add to the acceptance and trust experienced by ROBEAR users, patients may still experience a gap in care compared to a professional human care-giver when they are lifted by the ROBEAR; this may be due to limitations of the algorithm controlling ROBEAR and the functionality of ROBEAR. As a worst-case scenario, this gap in care may influence the patient's human-dignity as a result of interacting with the care robot. That is, patients may feel humiliated if treated like an object due to the robot's awkward behavior.[62] We propose that this expectation gap in treatment may be reduced by adding more patients' biometric information to the robot's database and by improving robotic perception with advanced sensors. One example of this approach is called "Affective Touch," which allows humans' emotions to be conveyed via tactile interactions such as touch, light stroke, poke, press, squeeze, and grab.[63] However, at the same time, affective touch would lead to a trade-off for privacy protection because touch implies close proximity to the patient, which influences the privacy received by the patient.[64]

In addition to the loss of expectation, another example of deception concerns the oversight of privacy during daily HRI. Gender is a social cue for robots to consider and relates to a robot's social communication skills with humans. This is important especially in healthcare services, and thus robots need to know how to determine a person's gender. On this point, Arnaud Ramey and colleagues developed an algorithm for robots to judge user gender. The basic idea is to use robots' cameras to detect chest contours of humans' upper bodies and based on the data collected to decide whether their targeted person is male or female.[65] However, concerning this approach, some users have privacy concerns and regard this process as an "invasive" HRI. However, for autonomous mobile robots, acquiring people's location information is necessary for realizing robots' social navigation in tracking or guiding people in

[60] The Strong Robot with the Gentle Touch, Riken, ROBEAR (February 23, 2015), www.riken.jp/en/pr/press/2015/20150223_2/.

[61] *Ibid.*

[62] A. Sharkey, Robots and Human Dignity: A Consideration of the Effects of Robot Care on the Dignity of Older People (2014) 16 *Ethics Inf. Technol.* 63.

[63] R. Andreasson, B. Alenljung, E. Billing, and R. Lowe, Affective Touch in Human–Robot Interaction: Conveying Emotion to the Nao Robot (2018) 10 *Int. J. Soc. Rob.* 473; J. Sun, S. Redyuk, E. Billing, *et al.*, Tactile Interaction and Social Touch: Classifying Human Touch Using a Soft Tactile Sensor, in International Conference on Human–Agent Interaction (HAI), Bielefeld, Germany (October 17–20, 2017).

[64] M. Coeckelbergh, Health Care, Capabilities, and AI Assistive Technologies, and AI Assistive Technologies (2010) 13 *Ethical Theory Moral Pract.* 181.

[65] A. Ramey and M. A. Salichs, Morphological Gender Recognition by a Social Robot and Privacy Concerns, in Proceedings of the 9th ACM/IEEE International Conference on Human–Robot Interaction (HRI), Bielefeld, Germany (March 3–6, 2014).

real environments.[66] Still, most people don't know that they have been located in the environment by robots when they come close to them.

PARO is a well-known, seal-shaped social robot developed by AIST Japan with many different built-in sensors and hypoallergenic fur to support its social interaction with humans. It is thought to be an ideal social companion tool for elderly people in order to reduce their loneliness. As mentioned before, a privacy risk regarding deception in embodiment is that elderly people have a tendency to self-disclose to a perceivably cute and interactive social robot. One example of this phenomenon is from a "long period" experiment conducted by ATR in Japan. Researchers found that when they mixed hand gesture and conversation functions in a humanoid robot at a care center, elderly people were more willing to tell the robot many of their personal affairs, such as recent frustrations, happy moments, interpersonal relationships, health conditions, etc.[67]

It is also the case that discrimination is also relevant to privacy concerns regarding deception in embodiment. Suppose PARO robot's behavioral module is created by machine learning and its default settings allow it to learn and evolve specific ways of interaction with people. In other words, if person A spends more time interacting with PARO, then PARO will perform more intimate behaviors with him. This way of design thinking is reasonable because when robots spend more time with people in social situations, people experience the interaction as being more realistic and affective. On the other hand, when applied in a care center, a problem that may occur is that PARO will diligently interact and respond to users who have an interest in talking with the robot, but PARO may just keep repeating simple and monotonic behaviors when interacting with people who are shy or who may not spend much time interacting with it. In additional, some elderly people might "react" to this mode of interaction and may experience "discrimination" in their small group via their interaction with PARO, thus, they have to actively interact with robots. In the long run, a privacy concern is that PARO encourages people to engage in more self-disclosure to robots.

As studies have demonstrated, embodiment plays an important role for a trust relationship to be built upon and for the collaboration between a human and an AI agent.[68] A recent experiment showed that embodiment may increase risk tolerance and reduce users' privacy concerns when researchers measure the changes of behaviors for those interacting with an embodied robotic system and a disembodied one.[69] Based on the experiment, users tend to provide confidential information and are more willing to disclose private information to the embodied robots due to their level of trust and acceptance.[70] In the medical and healthcare context, it is crucial to assess the conditions for trust, which is essential for user engagement, especially when users tend to disclose sensitive information to social robots by treating them as friends rather than care providers or aids.

On the other hand, a confidential relationship between physicians and patients is one of the core duties in medical practice. Medical confidentiality limits access to information

[66] J. Pineau, M. Montemerlo, M. Pollack, *et al.*, Towards Robotic Assistants in Nursing Homes: Challenges and Results (2002) 42 *Rob. Auton. Syst.* 271.

[67] A. M. Sabelli, T. Kanda, and N. Hagita, A Conversational Robot in an Elderly Care Center: An Ethnographic Study, in Proceedings of the 6th ACM/IEEE International Conference on Human–Robot Interaction (HRI), Lausanne, Switzerland (March 6–9, 2011).

[68] S. Herse, J. Vitale, M. Tonkin, *et al.*, Do You Trust Me, Blindly? Factors Influencing Trust towards a Robot Recommender System, in Proceedings of the 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Nanjing, China (August 27–31, 2018), pp. 7–14.

[69] See Lee *et al.*, above note 47.

[70] *Ibid.*

shared between patients and their professional caregivers or anyone involved in care delivery. Outside of this circle of confidentiality, sharing confidential information requires a patient's explicit consent. In addition, data protection rules, such as the EU General Data Protection Regulation (GDPR), prohibit processing of data concerning health unless some exceptions are met; for example, necessary processing for the purposes of medical diagnosis or treatment.[71] When social robots play the role of care providers or assistants, they need to be compliant with the rules of confidentiality and data protection. In order to reduce the potential effect on trust and the disappointment experienced by individuals due to an unexpected use of personal information, a robot's functions affecting privacy should be constructed at the design stage and be communicated to users to make it transparent with regard to its operational capacities. For instance, in the programming stage of design, the interpersonal clues can be added to the robot's behavior system to make users more aware of the continuing activities of data processing.[72]

Except for direct care, explicit consent must be given in medical practice under a clear explanation of what data has been collected, the purpose of such collection, and who can access the data. It is important for users to be aware that healthcare robots function not only as caregivers, but also as devices of data processing, with the capability to capture large amounts of data. For the latter function, there is a possibility that those who are outside of the circle of medical confidentiality, such as cloud service providers and manufactures, can also access sensitive personal data processed by the robots. Under this scenario, proper consent needs to satisfy a two-facet requirement which includes: (1) consenting to the acceptance of the robotic assistance for medical and healthcare purposes; and (2) authorizing the collection and processing of personal data outside of the medical confidentiality, including but not limited to allowing additional parties to access the data for the purposes of the improvement of care.[73] Considering the vulnerability of many users of healthcare robots, such as elderly people, disabled people, or those with cognitive impairments, the consent mechanism needs to be designed carefully to ensure that the principles of autonomy, transparency, and accountability can be respected.

Some common examples of healthcare robots are exoskeletons or lifting robots supported by AI and deep-learning technologies that rely on large amounts of data regarding the movement of users in order to optimize the performance of the device. Under the GDPR, a data protection impact assessment (article 35) is required to be carried out to evaluate on a contextual basis if robots have been designed in a privacy-friendly way and are able to comply with data protection rules and principles, such as purpose limitations and data minimization. Furthermore, the Guidelines issued by the Article 29 Working Party (WP29) emphasize employing the principle of privacy by design. The Guidelines' purpose is to make sure that privacy and data protection will be embedded within the entire life cycle of the development of robots. For example, thinking about the possible tasks the robots can perform during the design process may help reduce privacy invasion at the later stage, and through such design, an outcome is to make the robots' capacity for surveillance and data collection function only when it is permitted to do so.

[71] GDPR, art. 9.

[72] S. de Conca, E. Fosch Villaronga, R. Pierce, *et al.*, Nothing Comes between My Robot and Me: Privacy and Human–Robot Interaction in Robotised Healthcare, in R. Leenes, R. van Brakel, S. Gutwirth, and P. D. Hert (eds.), *Data Protection and Privacy: The Internet of Bodies* (Hart, 2018), p. 104.

[73] *Ibid.*, p. 107.

Safety and Privacy

Before comparing the two concepts of "privacy by design" and "security by design," it is first necessary to introduce these concepts separately. The term "privacy by design" refers to "data protection through previous technology design." Therefore, it is important to ensure that data protection in processing has already been integrated in the technology when created. A good example is the emphasis on the data protection impact assessment proposed under the GDPR. Nevertheless, how to implement "privacy by design" remains uncertain.[74] This is mainly due to EU Member States' incomplete implementation of the rule.

The principle of "Privacy by Design" requires defining technical and organizational measures at an early stage of design. In addition, the GDPR remains open to Member States' legislation to decide what exact protective measures are to be taken. As an example, "pseudonymization" is defined as one of the appropriate technical and organizational measures, but no more detail is given in the GDPR.[75] Further, article 25 of the GDPR does point out technical and organizational measures, which are designed to implement data-protection principles in an effective manner, and to integrate necessary safeguards into the processing of data in order to meet the requirements of the Regulation to protect the rights of data subjects.

It can be said that "privacy by design" tries to establish an appropriate regulation framework to prevent individual private information from being misused. However, with regard to thoroughly protecting individual private information, another concern is the coordination of the software and hardware, which falls under the concept of "security by design." The GDPR's focus on personal data also highlights how software is made and what components are used. It encourages enterprises to consider software security at the initial stage of design. Some scholars have also argued that GDPR obligations have extended to hardware choices.[76] According to their arguments, GDPR obligation should also include choosing and maintaining secure firmware and software for devices used for processing personal data. Therefore, the concept of "security by design" has become an emerging topic in the privacy area.[77]

Compared to privacy by design, security by design focuses more on the software and hardware development. In the general process of system development, it is difficult to address existing vulnerabilities or add in techniques and mechanisms to fix system problems at a later stage. However, in practice, it may also not be easy to design a comprehensive system at the very beginning. "Security by design" allows continuous testing and can be approached by complying with the best practice of programming.[78] On the other hand, making users aware of possible risks, namely to be more transparent in the overall process of system design, can help users make more explicit decisions when they plan to access particular devices or use their services. This consciousness of security can also help improve the design of the system.[79]

[74] A&L Goodbody, The GDPR: A Guide for Businesses (2016), www.algoodbody.com/media/TheGDPR-AGuideforBusinesses1.pdf.

[75] C. Kuner, L. A. Bygrave, and C. Docksey (eds.), *The EU General Data Protection Regulation: A Commentary* (Oxford University Press, 2018).

[76] W. K. Hon, GDPR: Killing Cloud Quickly?, *Privacy Perspective* (March 17, 2016), https://iapp.org/news/a/gdpr-killing-cloud-quickly/.

[77] D. Orlando, The Emerging Security by Design Principle in the EU Legal Framework, Master's thesis, University of Oslo (2018).

[78] K. Yskout, K. Wuyts, D. Van Landuyt, *et al.*, Empirical Research on Security and Privacy by Design: What (Not) to Expect as a Researcher or a Reviewer, in L. ben Lothmane, M. G. Jaatun, and E. Weippl (eds.), *Empirical Research for Software Security* (CRC Press, 2017), pp. 1–46.

[79] S. Wachter, Normative Challenges of Identification in the Internet of Things: Privacy, Profiling, Discrimination, and the GDPR (2018) 34 *Comput. Law Secur. Rev.* 436.

Therefore, the main difference between "privacy by design" and "security by design" is that the former emphasizes the previously discussed privacy assessment required by regulation, while the latter requires a continuous, dynamic facilities examination and improvement.[80] In addition, "privacy by design" controls the use of private information, ensures that the processing of private information is under the appropriate legal framework, and "security by design" keeps modifying the techniques to reassure that private information is always under comprehensive protection.

As mentioned above, embodied AI is also a safety-critical system. Its unwanted system behaviors from system integration will lead to consequences regarding physical safety and to human injury.[81] A difference between disembodied healthcare AI systems and healthcare robots concerns the system integration of software and hardware. Therefore, security by design is an important factor that healthcare robots should consider during the design and development stage.

An example of considering security by design for healthcare robots is a prototype standing-support machine with machine-learning capability that was developed by the Smart Robots Design Laboratory at Tohoku University in Japan.[82] In general, the main function of the standing-support machine is to assist patients in sit-to-stand and stand-to-sit motions via controlling the up and down function of a lift table on the top of the machine. Although current machines in use are controlled by human caregivers, in this application, a technical trend is to use AI to replace the caregivers' duty in order to save costs. Hence, the team from Tohoku University uses Support Vector Machine (SVM) to train their prototype in order to let the machine itself decide how to lift the patient in the right time and the right place. However, a difficulty in system design is keeping a balance between safety and privacy. To avoid privacy disputes, the team tried to suppress machine perception capability in numerous ways. They proposed a user state estimation method utilizing just a few inexpensive and simple sensors and also used the SVM learning algorithm to train the machine to distinguish different user motion states, so that the machine could autonomously support patients using data collected from machine pressure force sensors and distance detection sensors. Although this procedure doesn't need too much in terms of human biometrics, the gap of coordination between hardware and software is also reduced due to the simple design. The drawback, however, is its limited perception of human motions. The prototype machine can only understand three motion states, "standing, getting up, sitting," three conditions to infer the user's current state, and based on this knowledge decides how to adjust the countertop height. Here, an issue of security by design involves personal safety, such as the result of a fall or injury when the user rises too fast or the timing is wrong. At this stage, the machine infers the user's state mainly by relying on the pressure information of the user's hand grip and the distance information of the relative position of the user and the machine to assist in estimating the user's state. Considering alternatives to centroid calculations that allow machine learning to infer the user's state, while potentially improving the security of HRIs, it is likely to increase the user's information dimension and pose a challenge to privacy protection.

[80] S. Jacques, Safety and Security by Design (2016) 2 *Seek* 8.

[81] Y. H. Weng and D. Hillenbrand, The Intelligentization of Automobiles: Smart-Cars, Robo-Cars and Their Safety Governance (2014) 110 *J. Sci. Technol. Law* 632.

[82] M. Takeda, Y. Hirata, T. Katayama, *et al.*, State Estimation Using the Cog Candidates for Sit-to-Stand Support System User (2018) 3 *IEEE Robot. Autom. Lett.* 3011.

Discussion

Privacy is generally recognized as a fundamental human right today; however, there is no universal definition of privacy, and it changes according to different times and space, along with the influence of technological developments. In Japan, the definition of privacy rights is not only limited to Brandeis and Warren's classic definition as "the right to be let alone,"[83] it also includes many facets like the right to autonomy, the right to informational self-determination, the right to personal identity, etc.[84] Along this line, the aspect of a right to informational self-determination is especially important for data protection in healthcare robots since it deals with determining the governance of someone's personal information through daily HRIs. Hence, we are going to use this aspect of privacy to discuss how the factor of embodiment can be used by regulators to properly inspect the insufficiency of applying embodied AI healthcare robots into current data protection laws such as GDPR.

First, in terms of a data subject's rights, users as the data subject shall be protected under proper consent to the data processing of robots. In a previous section of this chapter, we mentioned two requirements for proper consent, including the acceptance of the robotic application for healthcare-related purposes, and the collection and processing of personal data outside of medical confidentiality. Such a consent mechanism might be revisited to ensure that the principles of autonomy, transparency, and accountability can be respected.

In addition, when individuals use robots in the processing of personal data, a common theme in data protection is to clearly convey to robots which data can be collected and also to require robots to make sure data is anonymous.[85] However, such an "informed consent" attitude is strongly influenced by embodiment in some use contexts of healthcare robots, and it will cause a gap in data collection. Hedaoo and colleagues conducted an experiment with a service robot using customer data in conversation in order to find out how people expected their data to be used. In their experiment, they divided robot data use into four categories: body language analysis, conversation analysis, database search, and ecological analysis for a Robot Barista which spoke to human testers. The researchers found that participants disliked being searched in databases, followed by being analyzed in conversation, but were most open to having their body language analyzed.[86] The outcome of the research showed that people have lower privacy expectations when a robot processes data from their body language.

Body language refers to a human's non-verbal gestures, movements, or social cues which have specific cultural meanings for people in communication. As Cabibihan and colleagues (2012) stated, applying gesture and body language into the design of a robot is an efficient way to enhance human–robot communication.[87] Although most people believe that robots' data processing of human body language is not likely to infringe privacy to a great extent, a major concern stems from the perception capability of robots in collecting human biometrics. With healthcare robots, collecting body language data from patients is inevitable, and also that

[83] S. D. Warren and L. D. Brandeis, The Right to Privacy (1890) 4 *Harvard Law Rev.* 193.

[84] M. Sogabe, S. Hayashi, and M. Kurita, *Information Law: An Introduction* (Koubundou, 2016).

[85] E. Fosch Villaronga, A. Tamò-Larrieux, and C. Lutz, Did I Tell You My New Therapist Is a Robot? Ethical, Legal, and Societal Issues of Healthcare and Therapeutic Robots (October 17, 2018), https://ssrn.com/abstract=3267832.

[86] S. Hedaoo, A. Williams, A. Fallatah, *et al.*, A Robot Barista Comments on Its Clients: Social Attitudes toward Robot Data Use, in Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Daegu, Korea (March 11–14, 2019).

[87] J. J. Cabibihan, W. C. So, and S. Pramanik, Human-Recognizable Robotic Gestures (2012) 4 *IEEE Trans. Auto. Mental Dev.* 305.

personal data will be collected as well. For example, say a patient waves his hand to tell the robot to stop its service, some inseparable personal data like his palm print, or face contour, can also be collected by the robot. Furthermore, another concern is that robot perception in social cues will sometimes infringe a higher level of sensitive personal data. An example is the case where an algorithm operating a robot is able to discern human gender differences by using its cameras to analyze human chest contour. Apparently, current users' attitudes concerning a robot collecting body language and biometrics will cause a gap of informed consent in data protection. This is not favorable considering that robots are gradually enhancing their perceptual capability and adapting to human living spaces. We propose that a solution to this problem is an education program designed to enhance people's understanding of the possible risks associated with emerging technology.

Second is the issue of privacy by design. The omnipresence of smart home technologies leads people to feel a loss of control over their personal data.[88] However, a concern beyond loss of data control could be the embodiment in deception and proximity from healthcare robots. From a study in evaluating elderly people's privacy-enhancing behaviors (PEBs) in home environments between a camera, a stationary robot, and a mobile robot, the results showed that elderly people changed their behaviors more when monitored by a camera.[89] In other words, elderly people may underestimate the risks of privacy violations from embodied robots. Consider that design strategies as discussed in this chapter might help to reduce the gap of PEBs regarding embodiment. Except for the idea of a privacy filter mentioned above, a self-explanatory user-menu might assist users in anticipating what tasks they can expect the robot to perform.[90]

As we have discussed, it is important to distinguish the concept of security by design from privacy by design. In May 2018, the International Organization of Standardization (ISO) established a new committee ISO PC/317 to draft the standard ISO/AWI 31700 on embedding privacy by design into consumer goods and services.[91] This proposal may support and encompass the privacy by design concept, but when it is applied to the concept of security by design, such as for healthcare robots, the concern exists that there will be a gap in risk assessment between data privacy and robot safety. Further, ISO 13482 (a robot safety standard)[92] adopts ISO 12100 for risk assessment with a three-step method in machinery risk reduction.[93] It is very different from PIMS's (privacy information management system) Privacy Impact Assessment for data protection. For safety-critical systems like healthcare robots, security by design needs to be seriously placed into the machinery risk assessment process in order to ensure system integration between software and hardware.

[88] W. Wilkowska, M. Ziefle, and S. Himmel, Perceptions of Personal Privacy in Smart Home Technologies: Do User Assessments Vary Depending on the Research Method? in T. Tryfonas and I. Askoxylakis (eds.), *Human Aspects of Information Security, Privacy, and Trust, HAS 2015, Lecture Notes in Computer Science* (Springer, 2015), Vol. 9190.

[89] K. Caine, S. Sabanovic, and M. J. Carter, The Effect of Monitoring by Cameras and Robots on the Privacy Enhancing Behaviors of Older Adults, in Proceedings of the 7th ACM/IEEE International Conference on Human–Robot Interaction (HRI), Boston, MA, United States (March 5–8, 2012), pp. 343–50.

[90] D. Hebesberger, T. Koertner, C. Gisinger, and J. Pripfl, A Long-Term Autonomous Robot at a Care Hospital: A Mixed Methods Study on Social Acceptance and Experiences of Staff and Older Adults (2017) 9 *Int. J. Soc. Rob.* 417.

[91] ISO/AWI 31700: Consumer Protection – Privacy by Design for Consumer Goods and Services.

[92] ISO 13482:2014: Robots and Robotic Devices – Safety Requirements for Personal Care Robots.

[93] ISO 12100:2010: Safety of Machinery – General Principles for Design – Risk Assessment and Risk Reduction.

CONCLUSION

In this chapter, we discussed the impacts of embodied AI for privacy in HRI from observations we noted in many use scenarios with three concerns consisting of embodiment in proximity, deception, and safety. We concluded that the embodiment factor of healthcare robots brings into focus new privacy risks in some applications of HRI. Although it is still difficult to predict how the outcome of our analysis in embodied AI will reshape the definition of future privacy rights at this time, the analysis shows that current data protection laws may need revisions in order to encompass the use of intelligent robots in healthcare, especially in aspects of data subject rights and privacy by design. As for the impacts of embodied AI on privacy in HRI and the law for algorithms, an overlap is the transparency for creating a trustworthy relationship between users and machines, and for the transparency of the circulation of personal data.[94] In addition, a straightforward way to look at the study of law and robotics through the viewpoint of embodiment is usually from the perspective of torts. However, the importance of privacy and data protection has been overlooked for a long period of time and thus provided motivation for the discussion we provided in this chapter. Finally, our analysis focusing on embodiment and privacy in HRI will not only benefit a law of algorithms, but also the development of privacy-friendly interfaces for healthcare robots.

[94] J. Günther, F. Münch, S. Beck, *et al.*, Issues of Privacy and Electronic Personhood in Robotics, in 2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication, Paris, France (September 9–13, 2012).