

Fall September 25, 2019

The Religious Impacts of Taoism on Ethically Aligned Design in HRI

Yueh-Hsuan Weng, *Tohoku University*

Yasuhisa Hirata, *Tohoku University*

Osamu Sakura, *The University of Tokyo*

Yusuke Sugahara, *Tokyo Institute of Technology*



The Religious Impacts of Taoism on Ethically Aligned Design in HRI

Yueh-Hsuan Weng^{1,2} · Yasuhisa Hirata³ · Osamu Sakura^{1,4} · Yusuke Sugahara⁵

Accepted: 16 September 2019
© Springer Nature B.V. 2019

Abstract

As robots become more and more ubiquitous in human society, the study of their emergence becomes critical to assessing robot performance and appropriateness for different applications, and for employment across countries with different cultural backgrounds. Throughout history, religion has been a major element in all human cultures, and there is a long history of religious commentary on the idea of automation. With emerging robotic platforms and applications already beginning to touch on the subject of religion, now is the right time to begin discussing this potentially controversial topic. The objective of this article is to analyze what impacts Taoist religion may have on the use of Ethically Aligned Design in future human–robot interaction.

Keywords Taoism · AI and culture · Chinese studies · HRI · Ethically Aligned Design

1 Introduction

One critical question about how to best prepare for an emerging society of human–robot co-existence has to do with how to design robot sociability. This covers considerations for various fundamental aspects of human society, including ethics, law, personal safety, and religion, amongst others [1]. In particular, religious practices have differed across cultures since the beginning of human society. Such factors matter when considering how to design a proper human–robot relationship in different societies. That said, current ethical, legal, and social implications (ELSI) studies often neglect how important cultural factors may be. This

knowledge gap in investigating the ELSI of AI and robotics is especially apparent when we touch on comparative or international issues [2].

The earliest case of robots and religion involves the Golem from Jewish folklore, a clay doll that loyally obeys its creator's orders. There are many religious restrictions for using the Golem. If its master doesn't follow these restrictions, he or she might lose control of the Golem [3]. The Golem folklore is believed to be the source that inspired Karel Čapek to create the character of "Robot" in his science fiction literature work *R.U.R. (Rossumovi Univerzální Roboti)* in 1920 [4]. The second kind of early example has to do with applying automation to religious activities. For example, the Hero of Alexandria's invention of the world's first vending machine for selling holy water, and of an automatic door for controlling the gate of a Greek temple [5]. Today, along with the steady progress of technology, robots are gradually entering our daily lives. Such technological intimacy will create a third type of relationship between robots and religion, which we can call "religious impact on robots".

In 2017, there was a debate over the question, "Can a robot be religious?" when Saudi Arabia granted citizenship to Sophia the robot [6]. Some people worried it might be difficult to expect Sophia's behavior to meet the requirements of Islamic discipline for female citizens. At the same time, half a world away, a Japanese company used the humanoid robot Pepper to read a Buddhist prayer [7]. A question naturally arises from this topic: can intelligent robots be affected in some way by reciting such a prayer?

✉ Yueh-Hsuan Weng
yueh-hsuan.weng@riken.jp

¹ Center for Advanced Intelligence Project (AIP), RIKEN, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan

² Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, 6-3 Aramaki aza Aoba, Aoba-ku, Sendai 980-8578, Japan

³ Department of Robotics, Graduate School of Engineering, Tohoku University, 6-6-01 Aramaki aza Aoba, Aoba-ku, Sendai 980-8579, Japan

⁴ Interfaculty Initiative in Information Studies, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

⁵ Department of Mechanical Engineering, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8550, Japan

Religious context is important for the creation of “Ethical Robots” as well. Hiroko Kamide said that the concept of “One Being for Two Origins”, considers the unification of technology, nature, and human beings. It emerges from the theory of three different aspects: “Good”, “Bad”, and “Muki (Avyakata)”. Though “One Being for Two Origins” is from the teachings of Buddha, it is not an absolute concept—in other words, right or wrong exists in the eye of the beholder. As long as humans accept that they are beholden to subjective mechanisms of perceiving the world, it should be possible to develop ethical robots which adhere to those mechanisms [8]. In Christianity, Jack Crosbie explains that sex with robots is sinful. Suppose in the future a care center wants to hire an AGI (artificial general intelligence) humanoid robot for providing sexual services to quadriplegic patients who have physiological demands. Regardless of whether the robot conceives of its behavior as part of necessary healthcare services for the patient, or whether it simply follows instructions from its master to perform such behaviors without any awareness, neither is condoned from a Christian perspective [9]. These examples teach us that while there is a large area of overlap between ethics and religion, they do not always imply the same conclusion. So, it is necessary to treat the study of the impact of religion on robotic design as an independent subject, and not merely a sub-topic under AI ethics.

The relationship between religion and technology should not be conceived of as only involving religion’s impact on technology. On the contrary, technology itself can impact or become a religion for human beings when there is a huge technological imbalance between an artificial AI, for example, and the human race. It is similar to the Cargo Cult phenomenon that happened during and after the Second World War, when Vanuatu islanders worshipped the US Military’s airplanes, which dropped goods by accident from time to time on their islands [10]. However, what kind, if any, of a techno-religion may be created from a future AI is still not clear at this time.

We can see that religion may be one of universal elements underpinning the operation of human society. What’s more, humans show different levels of social acceptance of and attitudes towards AI and robots under the sometimes subtle influence of different religions. Some say that the different attitude towards humanoid robots between East and West is due to religious influences. In Christianity, the fundamental belief that everything is created by God may cause people to reject robots that imitate human appearance. Shinto’s animist origins on the other hand may encourage people to treat humanoid robots as a natural extension of existing beings and as human companions. Such debate has been the main axis of discussions in robots and religion studies in the past [11–15]. Outside of Christianity and Shinto, however, it seems that discussion of other religions is relatively limited.

In particular, when it comes to Taoism, we cannot find any previous academic publication on this topic.

Ethically Aligned Design (EAD) is a set of ethical guidelines developed by IEEE Global A/IS Initiative, which covers several issues such as General Principles, Embedding Values into Autonomous Intelligent Systems, Methodologies, Personal Data, Law, Policy, Classical Ethics in A/IS, and Mixed Reality in ICT, amongst others. It aims to remove people’s fear of AI technologies, while also showing that incorporating ethics into technology development helps innovation. The version of the EAD document that we used is Version 2, which was published in December 2017 [16]. Applying EAD guidelines to human–robot interaction can improve robot sociability through a Top-Down design method, and also reduce cross-cultural issues in the design and usage of robots. In this article, we aim to investigate the religious impacts of Taoism on Ethically Aligned Design in human–robot interaction.

2 Taoist Beliefs and Practices in Reference to AI and Robotics

The earliest robot story in Taoism classics was from “Liezi”. One day a craftsman called Yen-Shih brought an artificial, humanoid dancer to demonstrate to King Mu of Zhou (976–922 BC or 956–918 BC). The king was impressed by the effect given by the artificial dancer, because it was difficult to tell that the dance was not being performed by a real human being. However, when at the end of its performance the dancer suddenly winked its eye at the king’s concubines, the king saw this and was very angry, even wanting to kill Yen-Shih. Yen-Shih was terrified, and he cut the humanoid dancer’s body into half to show the king that what had flirted with his concubines was just a machine and not a real human being [17].

The philosophical practices in Chinese Culture, “Ru-Tao-Shi (儒道釋)” are more than just a religion but also a way of life in China. Among them, “Tao (道)” means Taoism which covers dual meanings as philosophical thoughts (philosophical Tao) and the religion (religious Tao). The philosophical Tao comes from late period of Chun Qiu (春秋) (B.C. 5 century). Their classic texts are “Tao Te Ching (道德經)” and “Zhuangzi (莊子)” by the famous Chinese thinkers Laozi (老子) and Zhuang Zhou (莊周), who advocate ideas including the “doctrine of inaction” (*wu wei* 無爲) and the “way of coexistence with nature” [18, 19].

On the other hand, the religious side of Tao refers to a religious organization founded by Celestial Master Zhang in the last years of the Han Dynasty (A.D. 184–220) in mainland China. Though it still adopts thoughts from the philosophical Tao as its theoretical foundation, it adds Chinese folk beliefs about “gods and ghosts”, the worship of nature,

and the doctrines of Fangxian (方仙), and Huanglao (黃老) to form a mixed, religious body of thought. Note that in this article, when we use the term Taoism we refer to the *religious* Tao. An apparent difference between the two perspectives of Taoism can be seen in Laozi's relative status. In the philosophical Tao, Laozi is a guru who inspires people to discover many secrets from his masterpiece Tao Te Ching. However, in the religious Tao, Laozi has another status called Taishang Laojun (太上老君) or Daode Tianzun (道德天尊) and is one of the world's three superior Gods.

What is called "Tao" can be recognized as Taishang Laojun, who is an ultimate, unlimited supernatural power. We can also refer to Tao as "one". According to Tai Ping Ching (太平經), one is the root of Tao and the origin of life [20]. This gives theoretical support to Taoist ideas about keeping one's good health as well as recovering one's original simplicity. Another way to observe Tao is via the preservation of one's spiritual essence, or "*qi* (氣)", and this *qi* is a way towards becoming a Taoist sage. After the Northern and Southern dynasties period (A.D. 220–589), when Buddhists criticized the *qi* approach, there came to be a new way to observe Tao via "mind". [21] Because Tao from the classics Tao Te Ching is presented in a vague way, Tao as "one", *qi*, or mind have all been gradually generated through the concept's long history as ontological mediums for people to sense the abstract Tao.

To observe Tao is the start of applying the teachings of Taoism to one's daily life. The end of practicing Taoism to its followers is to become an immortal Taoist sage in a mystical land, a process called "*de dao cheng xian* (得道成仙)" in Chinese. Under this ultimate goal, a comparative difference with Christianity, Islam and Buddhism is that Taoism encourages people to live in the present and pay more attention to the world that we are existing in, not to imagine the other world. Because there is neither heaven nor hell to go to after death, people should focus their energies on the present by practicing Taoism in their daily life.

In Taoism, Tao is the origin of everything, as well as the master of the universe. In such a Tao-centered world there are many Gods and spirits who can perceive the existence of Tao. By the degree of their perception of Tao, they form a hierarchy among these Gods and spirits. This is the foundation of the polytheistic beliefs of Taoism.

Joseph Needham had commented that Taoism boosted ancient Chinese science and technology developments, because it helped the Chinese cultivate a scientific attitude. Also, East Asia's Chemistry, Mineralogy, Botany, and Pharmaceuticals all originated with Taoism [22]. Sun Yiping said that Taoism is the root of science and technology in ancient China. Science and technology had been despised for a long period of time under Confucian society. But Taoist followers don't reject using objective terminology and technical methods to explain how the universe operates. In other words,

they practice empirical science for religious purposes [23]. This kind of open attitude to science and technology from Taoism should make it more likely that AI and robotics are accepted relatively easily.

When AI and robotics enter a society with a Taoist culture, at the initial stage they can be efficient tools to support humans in their daily lives. Some examples include assistive robotics for in-field Monitoring and Communications Assistance and Nursing Care Assistance: both applications need sophisticated machine intelligence to ensure elderly people can be properly cared for at home. In these cases, robots are expected to have higher levels of physical and psychological perception capabilities when they serve elderly people, unlike assistance with bed transfers, walking, using the bathroom, and bathing which need lower levels of AI to ensure its physical interactions with humans [24]. As more AI tools are used in human living spaces as tools, such huge demand for AI products will trigger an intelligence explosion, lead AI across the Singularity, and finally become Superintelligent [25].

A new relationship between human beings and AI will need to be defined when AI reaches the level of Superintelligence. From Asimov's Three Laws of Robotics to the recent German ministry's proposal towards creating an ethical code to ensure AI safety in self-driving cars [26], these artificial machine ethics are all based on a premise: that robots are tools for humans, or that robots should serve their human masters. Under this premise, robots should protect human goods and benefits, and the priority issue is human safety. This is just like the first law of the Three Laws of Robotics notes that: "A robot may not injure a human being or, through inaction, allow a human being to come to harm" [27]. Such "Safety Intelligence" in human–robot interaction works fine when humans hold superior moral status to AI/robots [28].

Following the development of AI, machines could eventually be given a moral status if they can feel pain, and could even be considered subject to law when they have an equal level of consciousness to human beings, or even when they are a superior existence to humans after becoming Superintelligent. At this level, AI safety issues are beyond early mentioned "Safety Intelligence", but instead on how to avoid existential risks caused by Superintelligence [29]. The proposed transformation in AI safety will influence machine ethics as well, as human superiority may no longer be feasible under this new AI paradigm. As we progress toward the era of Superintelligence, we may have to discard traditional ideas centered around human superiority to artificial agents, and instead look for a practical way to coexist with superior artificial beings.

In Taoism's world, AI and robots are not restricted to only an inferior or superior existence to human beings, but can emerge out of the discussion of material things into a

method or way for humans to change themselves. As discussed, to become immortal Taoist sages in a mystical other world is the supposed ultimate goal for Taoist believers. Additionally, as noted, through the long history of China, there have been several different approaches to this question suggested. In light of this, AI and robotics could be a feasible way to help Taoist practitioners get closer to their goal of becoming immortals. Though using these technological approaches can be an effective way to reach the goal of practicing Taoism, one concern is about breaking down the balance of human coexistence with our natural environments. If it is true, what will be the boundary to stop the over use of AI and robotics?

In this article, we want to discuss potential religious impact of Taoism on AI and robotics from three different perspectives: (1) AI/robots as superior existence akin to gods; (2) AI/robots as a method for becoming immortals; (3) AI/robots as tools to assist humans (Fig. 1).

3 AI/Robots as Superior Beings Akin to Gods

Deified of Laozi (老子)—Taishang Laojun (太上老君) or Daode Tianzun (道德天尊)—is one of the three superior gods, along with Yuanshi Tianzun (原始天尊) and Lingbao Tianzun (靈寶天尊). They are called “San Qing (三清)”. Taoism is polytheistic, including the three main gods mentioned above, and up to 233 total gods in the Taoist world overall [30]. Generally speaking, this number can be divided into three areas: (1) Natural Gods, (2) Thing Gods, (3) Human–Ghost Gods.

It might be difficult for people to imagine AI as a kind of superior existence to humans presently, but when AI crosses the Singularity they will achieve human-like intelligence—known as artificial general intelligence—and then grow even more powerful and intelligent, becoming Superintelligent entities, akin to gods. Though it is a long-term process, someday AI will become superior in existence, or God-like, as it were; one question to ask, then, is will belief

in Taoist gods cause people to treat Superintelligent entities in a unique way?

Though Taoism and Shinto contain aspects of both polytheism and animism, Taoism’s acceptance of an artificial being of the kind described above is not as high as Shinto’s. In the Shinto worldview, there are 8 million gods and spirits that exist in natural environments, so it comes as no surprise for the Japanese to treat things in a special way. For example, the “tomb of needles” is a place to worship needles that have helped their human partners accomplish tasks like those of a tailor [31]. Except for Thing Gods, it is rare for Taoists to treat things per se in a special way, like the Shinto attitude towards sewing needles. One exception can be god idols, which are humanoid wooden idols placed inside Taoist temples that accept people’s prayers. After a special ritual to endorse the sanctity of newly completed idols called “Kai-Guang (開光)”, they become avatars of Taoist Gods in the real world. Though these wooden idols are not Object Gods, people believe the Gods’ spirits stay inside them, and therefore, treat them respectfully.

A typical way for people to communicate with Gods in Taoism is to burn incense sticks when they pray. People also prepare offerings, like fruits or other foods, light incense, and also burn symbolic prayer money to express their thankfulness. Another way of communication is through charms—a kind of Taoist magic figure. These Taoist charms come originally from Huang Di (黃帝)’s “Yun Shu (雲書)”. These magic figures are written by Taoists based on their observations of nature. They are not created by Taoists but can be represented as Gods’ messages to humans. This may be important in AI ethics, as we can consider developing a communication system to perceive superintelligent entities’ will and help their decision making become relatively transparent to humans, so that humans will have better sense for how to interact with AI.

Geng Sen (庚申) is a belief in Taoism that a worm called “San Shi (三尸)” hides inside people’s bodies secretly, recording all their words. Every year on Geng Sen these San

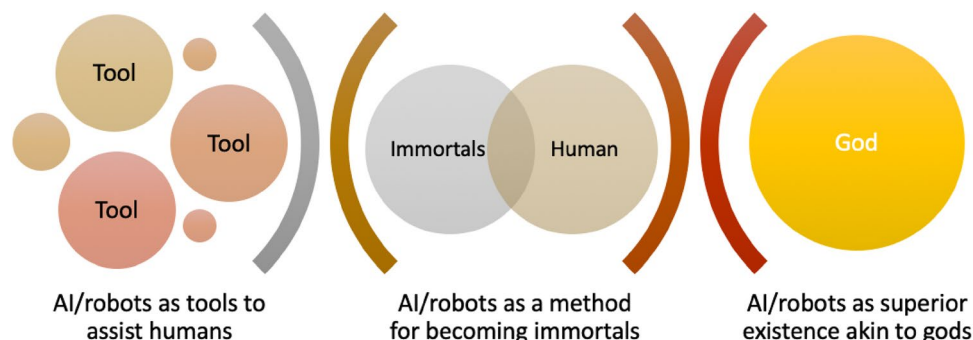


Fig. 1 The religious impact of Taoism on AI and robotics from three different perspectives

Shi worms fly to the heavens and report to the Gods about people's behaviors [32]. The Geng Sen belief is the foundation of Taoism. People must obey these disciplines, under the supervision of San Shi worms, otherwise Gods will know about their misbehaviors. If Superintelligent entities follow this precedent, they will be able to create an efficient monitoring system to control human behaviors through high intensity surveillance systems to evaluate daily human performance, something similar to a truly omniscient and more powerful version of today's social credit system in China. In this case, how to create a trustworthy relationship between humans and the AI will be the main challenge.

4 AI/Robots as a Method of Becoming Immortals

As a local religion originating from China, Taoism has had great influence in many aspects of science and the arts in Chinese society through more than two thousand years of history. The Taoist concept of alchemy applies cosmological ideas such as Yin and Yang (陰陽) and the Five Elements or Wuxing (五行), to the process of using iron and other mineral ingredients to make a medicine for eternal life. This is called Waidan (外丹) or "External Alchemy". Waidan is similar to European Alchemy in the Middle Ages. However, while it was popular from the Han dynasty to the Tang dynasty, it gradually became unpopular after the Song dynasty.

Modern western medicine categorizes diseases according to specific disciplines (for example: general medicine, general surgery, neurology, cardiology, urology, etc.), but traditional Chinese medicine treats diseases from a "harmony" viewpoint. Influenced by Taoism, they see the appearance of disease as representing an imbalance in the patient's body, and the way to cure this is not to focus on a single ailment, but instead on the harmonization of the universe (body). Since the Tang dynasty, Waidan has gradually been replaced by Neidan (內丹) or "Internal Alchemy". In Waidan, Taoists use a cauldron to make medicine, but in Neidan the human body itself is a cauldron in which Essence, Breath, and Spirit are cultivated for the purpose of improving physical, emotional and mental health, and ultimately, turned towards the goal of pursuing immortality.

Shen Xian (神仙) thoughts refer to becoming a God or an immortal person via Taoist alchemy like Waidan or Neidan, as we have mentioned. When it comes to technology, we are still far away from being able to upload human consciousness to machines. As depicted in the movie "Transcendence", a more important issue in robotics could be how to use technology to extend human lifespan or improve our physical capabilities, such as cyborg technology [33]. In the meantime, we should be mindful of ethical issues around

using this technology. In one case, an Austrian man who lost both of his arms added two prosthetic arms (one of which he can control with his brain waves) and used them to handle the steering wheel of his car. One day when he drove his car, he lost control and crashed, causing his death [34].

One unique aspect of Taoism has to do with immortal sages, giving people a positive motivation to follow the religion. It seems that cyborg technology could be a more feasible way of reaching this goal of becoming immortal. But, how do we avoid overusing cyborg technology? Let's look at an extreme case. In the Japanese science fiction "Android Kikaider", Hakaider implants a live human brain into a mechanical body. Can such immortal beings be accepted in Taoism? It will be necessary to discuss required conditions for using such a "method," and whether it contradicts with the Taoist theory that man is an integral part of nature.

5 AI/Robots as Tools to Assist Humans

In Taoism, people honor Gods by offering food, drinks and fruits. People transmit their wishes to Gods via lighting incense and also by burning symbolic prayer. There are different Gods who have their own expertise to provide service to their followers. For example, Wen-Chang (文昌) is the God of knowledge and wisdom. For students who want to pass their college entrance exams, they can attempt to make contact with Wen-Chang. Ma-Tzu (媽祖) is the Goddess of sea voyages whom sailors usually contact before their departure. The inspiration for AI Utilitarianism in Taoism may be: what kinds of factors are important when we treat AI beings as tools for humans?

This kind of viewpoint is the base for AI/Robots becoming human tools supporting them in their daily environments, acting as health aides, social companions or "magic machines" to realize people's wishes. If that is the case, it is not surprising that sometimes idols in Taoist temples become targets for those who want to express their anger to the Gods. Idols are where a God's spirit resides. They are magic machines created in the God's physical image for its believers. In Taiwan, a gambler who stole a God idol cut its hands and legs and discarded it in a forest in 2015. When he was arrested, he said he prepared many offerings to the God, sincerely praying to the God to help him to win money from his gambling, but he still lost it all. So, he made a plan to exact revenge towards the God's inaction towards his wish.

Feng-Shui (風水) is also a crucial factor when we consider robots as tools. Taoism holds that Grotto-heavens and Blessed lands (洞天福地) are sites that have helped Taoists in their religious practice on the way towards becoming immortals. Feng-Shui is a philosophy of how to choose better living environments for human beings which includes a set of complex rules and many unspoken rules of thumb

handed down from ancient China. Feng-Shui can be seen as a tool for people to seek out these holy Taoist sites. On the other hand, in our daily life there are also some unwanted places which are dull or unsafe, such as factories or recycling plants. The existence of these sites is necessary to maintain the operation of our society, so there must be someone to do tasks there. An innovative way of thinking about the application of Feng-Shui to robotics is that robots can be avatars for humans in some unwanted environments. In other words, robots can help people enhance their quality of life by replacing at 3D (Dirty, Dull, Dangerous) work environments for humans. This is also a traditional viewpoint in understanding the developing relationship between humans and robots.

6 Taoism and Ethically Aligned Design in HRI

In the previous discussion, we pointed out the religious impact of Taoism on AI and robots, separated into three different perspectives: as Gods, immortals and tools. The one that has demands immediate consideration is the use of AI/robots as human tools, because we are already at the initial stage of social robots entering our societies. Their role of supporting humans in healthcare and their daily entertainment is as a kind of tool to ease users' burdens. Today some Taoist temples in Hong Kong use humanoid robots to provide result interpretation services of Kau Chim (Divination). In the past, worshippers relied on the experience of senior human interpreters to read and understand their fortunes. It's straightforward, however, to use robotics technology to provide such a service. Various HRI design factors need to be considered when AI or robots are used to support users in this case however. In the future, robots in the religious use contexts will not only be consulting agents in Divination, but could also serve in other roles such as social companions or religious avatars. In all these cases religion and cross-cultural issues will gradually become more important for HRI designers.

The importance of religion and cultural factors to HRI design is twofold. First, the thinking around design in robotics technology can be very different depending on the designer's specific religion and cultural background. For example, "Automata" or mechanical humanoid or animal-shaped toys were especially popular in Europe from the sixteenth to the eighteenth centuries. Over a similar time period (from the seventeenth to the nineteenth centuries), Japan also had their own style mechanical humanoid toys called "Karakuri Puppets", one of the most representative examples of which is a mechanical doll for tea serving. It can autonomously transport a cup of tea to the front of the guest, stop and wait until the guest move back the cup on

the tray, and then return back to its original place. Although Automata and Karakuri Puppets are both mechanical toys for the purpose of entertainment, the thinking behind their design is not exactly the same. Shigeki Sugano argued that European Automata pursues a realistic approach to presenting action in exact detail from the object it imitates, such as Jacques de Vaucanson's Digesting Duck which vividly reflects the animals' digesting process. On the other hand, Japanese Karakuri Puppets pursue another artistic approach, more akin to Noh drama in that its various emotional states can be presented by simply changing the angle of the mask [35].

The Taoist definition of human mental health can be understood as living with a transcendental point of view, and not finding happiness in things. Additionally, it seeks to prevent people from reaching extremes by virtue of their own specific interests or preferred values. It also starts from the point of view that human beings are fundamentally a part of nature. Hence we must recognize the relationship between ourselves and the objective nature or external reality. [36] When we apply this Taoist thinking about mental health to design, we can find an apparent difference in that robots shall not only play a role limited to always following and satisfying human orders and needs. When it's necessary, robots can limit or refuse a human's orders for his or her own good. For example, David feels bored and wants to play video games. He asks his service robot to connect the channel for him. However, he has already played the game for over 12 hours in the past day, without a rest. In this case, robot needs to disconnect this service for David's own good, despite the fact that this against David's wishes.

Second, religion and cultural factors influence AI governance, especially when it comes to norms that govern human-robot coexistence. Although many proposed ethical principles and values for AI are universally recognized by the global community, some ethical values are still developed locally according to unique cultural backgrounds. These cultural factors matter when we align ethical values with the design process for autonomous systems, especially in designing the interface for human-robot interaction. Now we attempt to analyze how Taoist religion may influence designers when aligning AI ethical values for human-robot interaction. They cover three issues as below:

(1) Applying Taoist Ethical Values to Human-Robot Interaction

Yiping Sun commented that the core of Taoist ethics consists of ethical principles and social norms coming from people's beliefs about "Tao". Though the sources of these principles and norms are complex and various, in general, they include four main parts [23]. The nominal goal for Taoist believers is to become immortal sages, which sounds like an aloof

attitude which exists outside of human society. Actually, Taoism asks its believers at the beginning of their journey to have a worldly attitude and to positively practice principles and norms in the society in which they exist in order to keep a balance between the transcendental and the mundane. Second, the idea of Cheng Fu (承負) means that one's good and evil behaviors will give his descendants corresponding influences in the future. On the other hand, someone who suffers luck or misfortune can lay blame or praise at his forefathers' feet. Third, it encourages people to do good and discourages people from doing evil via several kinds of Taoist religious disciplines that guide people to behave well and gradually work towards becoming immortal sages. Fourth, an ideal ethical relationship in Taoism is to pursue a harmonization of object and ego to maintain the wider environmental balance.

When the above mentioned Taoism ethics are applied to the design process for intelligent robots, they will influence the foundation of human-machine relationships and the concrete ways that humans and robots interact. The first example tells us that AI ethics will still be human-centered, but that we should not to use a transcendental framework to define ethical values for a small interest group. In the second example, Karma in Buddhism is a concept leading adherents to reflect upon the cause and effect relationship for a specific person's fate. However, the thought of Cheng Fu is different because cause and effect relationship not limited to a specific person but can cross different generations of a single family. Due to this difference, Taoist believers' social acceptance of AI technology should be greater because the uncertainty and risks surrounding a machine's autonomous behaviors could be defined as a group obligation. In the third example, in Taiwan some regions have developed a principle of abstaining from killing or eating cattle, because they are seen to be important helpers for traditional farmers during the early agricultural period in Taiwan. If robots' role in serving humans becomes crucial, then humans should accordingly consider special manners for the disposal of broken robots, for example, and not simply casually throwing them into a garbage disposal site. With regard to our last example, we should also consider the usage of robots from an environmental and ecological context. For example, robots do not need to rest in order accomplish their tasks. Considering the goal of a harmonization of the object and ego, Taoism would prefer to follow a humanitarian principle in their interactions with robots.

(2) Taoist Culture versus Machine Ethics

In “Zhuangzi (莊子)”, there is a famous debate on whether one can know the joy a fish experiences: *Zhuangzi and Huizi were enjoying themselves on the bridge over the Hao River. Zhuangzi said, “The minnows are darting about free*

and easy! This is how fish are happy.” Huizi replied, “You are not a fish. How do you know that the fish are happy?” Zhuangzi said, “You are not I. How do you know that I do not know that the fish are happy?” Huizi said, “I am not you, to be sure, so of course I don't know about you. But you obviously are not a fish; so, the case is complete that you do not know that the fish are happy.” Zhuangzi said, “Let's go back to the beginning of this. You said: How do you know that the fish are happy; but in asking me this, you already knew that I know it. I know it right here above the Hao.” For robot ethics research, we are eager to know when and whether robots have self-awareness, but following this story, suppose someday we heard news about a self-aware robot that has been created. We may wonder, is the robot self-aware to all observers, or, do we alone know that the robot is self-aware?

Accompanying the growth of the range of services provided and the growing level of machine autonomy, we will need to authorize some spaces for machines outside of the loop of human control in order to make their decisions as they perform their duties. This emerging field is called “Machine Ethics” [37]. However, the state-of-the-art of machine ethics remains far from creating real “ethical robots” or “robots who can perform behaviors or make decisions ethically from humans' observation” [38], something similar to John Searle's “Chinese Room Argument” [39] or what is known as “weak AI”. Even so, it already has potential demands in the field of safety critical AI systems, such as the German Ministry of Transportation's proposal for using a Code of Ethics for self-driving cars on public roads [26]. Furthermore, Michael and Susan Andersons also propose using care robots built with machine ethics for educating children about what things are ethical and are not [40].

With regard to the implementation of machine ethics, many critical challenges still exist. According to IEEE's Ethically Aligned Design, one general issue around “Norm Implementing” is to discuss the pros and cons of Top-Down (where the system has some symbolic representation of its activity), Bottom-Up (where the system builds up an understanding through experiences of what is to be considered ethical or unethical in certain situations), and Hybrid Approaches (where action selection can be carried out by sub-symbolic system, and add a symbolic gateway or “ethical governor” to control the situation) [16].

The other issue which “Norm Selection” concerns has to do with which kind of ethical principles we choose for artificial moral agents to follow. Imagine there are a group of elderly people and another group of young people inside a bus. When the bus arrives at the terminal how do we decide the order in which they should disembark from the bus? In Asian society, young people are taught to respect elderly people. So, if we were to follow Immanuel Kant's

Deontological Ethics it sounds like elderly people should have priority to get off the bus. On the other hand, if we adopt Jeremy Bentham's Utilitarianism then young people should get off the bus first, because their actions are faster than the elderly. It saves more time for the total passengers if elderly people wait for young people.

Compared to Norm Implementing, Norm Selection is a highly culturally sensitive issue, especially in terms of the ethical norms of Taoism. Embedding the ethical values of Taoism into autonomous systems is about machine perception. How can machines tell if their current situation (or state) needs the application of their embedded ethical decision functions? Or how else can they detect whether any allowable or unallowable norms exist in its local environments (the real world)? This is a key issue for applying machine ethics, because even a robot with sound norm selection and norm implementing mechanisms may still not be able to do such ethical reasoning if it misunderstands its situation. Let's take a robotic guide dog for example. Suppose a blind person uses a robotic guide dog to take him cross the road. When they are still waiting for the traffic light to change from red to green, say the blind person makes a wrong guess and allows the dog to guide him across the road right away. In this situation, the robotic guide dog should have a kind of reasoning mechanism to enable it to make an "intelligent decision to disobey" its human master. That said, a prerequisite for the guide dog is to correctly understand its current situation (they are in a corner of the crossroads waiting for the traffic light to turn green) and relevant deontological norms (road signs and traffic lights). In this case, the concept of a Taoist charm can be a way for formatting vision-based norms and easing the complexity of the perception of machine ethics.

Ancient Chinese Taoists believe that charms are the foundation of the five numerologies (五術). Taoist charms are regarded as a kind of media for Taoists to communicate with the heaven, and these magic figures are written by Taoists from their observation from the nature. They are not created by Taoists, but can be seen as as Gods' messages to humans. In the past, these Taoist charms were not only used as medical treatment for patients, but also believed to have the magical power to trigger natural spirits' action or inaction. One extreme example is from Hong Kong's famous zombie movies in the 90s. In one movie, a zombie was suddenly stopped in its tracks when someone pasted a Taoist charm onto its forehead. To many Taoist believers, a Taoist charm is a kind of instruction to warn non-human spirits about something that they should or shouldn't do. Because these people are Taoist believers they already have psychological expectations that make these figures a normal

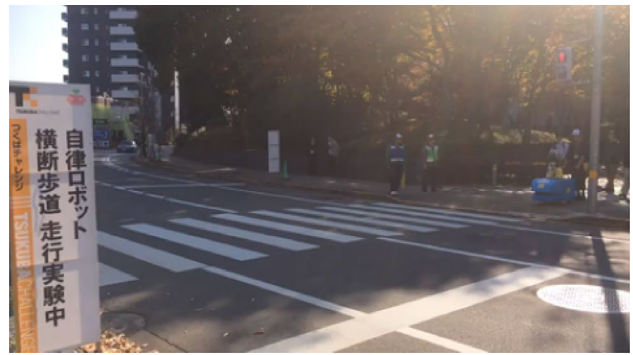


Fig. 2 A robot is waiting for traffic light at an experiment in the Tsukuba robotic special zone

feature of their daily lives. If we borrow this idea to develop a common symbol system for human and robots, then these "visualized norms" from Taoism can be one solution to reduce the complexity and uncertainty for machine ethics in their initial stage (Fig. 2).

(3) From Fuzzy Thinking to Electronic Personhood

In recent years cases have been reported around the world of people projecting their empathy on to robots when robots are mistreated by someone [41, 42] or users feel a special emotion of likeness and dependence when robots accompanied them for a long period of time [43, 44]. In 2015, there was a case where an inebriated person ran into a Softbank shop and attacked an innocent humanoid Pepper robot. Many people thought that the drunk individual should be sued for its injury to the robot, but not for damage to property under the Japanese Penal Code [45]. This response shows a potential gap between "to be" and "ought to be" when we define robots as tools. Although this kind of attitude towards robots will be influenced by different social and religious concerns in many societies, in the end the legal system may consider assigning a proper legal status to robots someday in the future.

The European Parliament is the world's first legal body to consider creating an independent "Electronic Personhood" status for sophisticated autonomous robots having a status as electronic persons with specific rights and obligations, including "making good any damage they may cause, and applying electronic personality to cases where robots make smart autonomous decisions or otherwise interact with third parties independently" [46]. However, many discussions currently hold a conservative position with regard to the EU's proposal, including the IEEE's Ethically Aligned Design, because it is an attempt to break the long-stable order between the "subject" and "object" of the law. Suppose robots become a legal subject: in this case, then, who should afford responsibility for the decisions made by these entities? On the other hand,

if we insist that robots remain in the category of property or things from a legal perspective, than their capability of performing autonomous behaviors will always challenge their human keepers' obligation for maintenance and management, thus causing an accountability gap.

Philosophical Dualism is rooted as a core part of Christian western societies. In the seventeenth century, René Descartes' "Mind–Body Dualism" recognized humanity's unique existence apart from other animals [47]. From the mid-twentieth century, the dualism of zero and one has been the foundation for developing modern computers and cybernetics. Though computer states can be represented based on zero and one—as yes or no, respectively—human perception is not like this. In order to solve this problem, Lotfi A. Zadeh proposed his famous "Fuzzy Logic", which helps to represent the gray zone for many daily applications of control theory, including train brake controls or the temperature control of air conditioning units [48].

Fuzzy Logic is like a harmonization between Yin (陰) and Yang (陽), which are the two extremes for Taoism. We may say such wisdom comes from the East and can be helpful in solving something that might be regarded as a dead end from Dualism's point of view. One such possibility is applying the Taoist way of thinking to the issue of Electronic Personhood and to consider creating a gray zone between the object and the subject of law, or as we mentioned in previous section, the "Third Existence" [28]. If we refer to humans as the First Existence, which is a status that can claim rights and confers responsibility, then artificial things and non-human animals are the Second Existence that is a passive status and only can receive consequences of rights and responsibility. Naturally, there will be a status in between that we need for a non-conscious, but autonomous robot.

A concrete way of applying Third Existence towards designing robot sociability includes, but is not limited to, the following two possibilities. The first is to go through a case study of the robot that is involved in a specific incident and authorize a professional committee to judge whether all the "characteristics" that were collected from that robot amount to proof of a legal subject or legal object. Take the case of Pepper the robot as an example [45]. If the committee found that Pepper is a machine which perform its actions without a high or even medium level of sentience, then their decision might likely be to judge it as merely a legal object. Although this approach is flexible as robots begin to enter human society, it will become a problem to perform such case-by-case judgments if the number of such robots become "considerable" in the overall market. At this stage, we may refer to Takashi Izumo's "Digital Peculium" or Digital Specific Property. It's an idea that borrows from the Roman slaves' Peculium status and transferred into intelligent robots in order to secure the robot master's liability arising from the economic activities carried out by his or

her robots, to be distributed between the robot owners and their contract partners [49]. When compared to the European Parliament's Electronic Personhood (EP), Digital Peculium (DP) is a specific status for robots beyond their position as property, but not yet at the higher position of something equivalent to humans. In other words, they are not able to receive rights or afforded responsibilities under DP, but their actions are more accountable because an independent economic account is affiliated with their Digital Peculium. This is a good example of a third existence.

7 Conclusion

Culture and religion is a deep and meaningful topic for social robotics. From the point of view of religion, we can explain why Christian-leaning Western societies are threatened by the existence of robots. Terminator or Frankenstein imagery suggests that robots are a threat to humans. On the other hand, the Animism of Japan's Shinto religion led to the creation of the characters Atom Boy or Doraemon, which are not viewed as a threat. In addition, religion is important to the normative development of ethics and law to robots. In this article, we choose the relatively unexplored topic of Taoism as our subject of study, and then proceed to analyze Taoism's religious impacts on human–robot interaction. We sketched the potential religious impact of Taoism on AI and robotics and found that it can be classified from three different perspectives: as superior existence akin to gods; as a method for becoming immortals; and as tools to assist humans. With a central focus on the third impact, AI and Robots as tools to assist humans, we referred to the IEEE's Ethically Aligned Design principles and discussed the connection with the design of human–robot interaction along with a focus on how Taoist ethical factors influence this design. When designing AI and autonomous systems, we should ensure that the behaviors of robots will not violate human value systems. Religion as a collective belief in society has at least the same importance as other social values like ethics, law, and safety. The religious impacts of Taoism on HRI will be important when we begin to design robot sociability. The next steps for this research can be to establish religion as a sub-field of culturally-aware robotics, by defining its theoretical background and its most relevant goals in today's fast-changing, culturally diverse world.

Acknowledgements The authors wish to thank Mr. Christopher Benek and Prof. Shuzhi Sam Ge for their helpful suggestions.

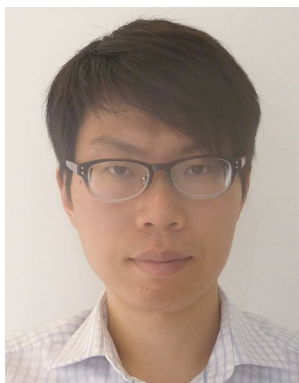
Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

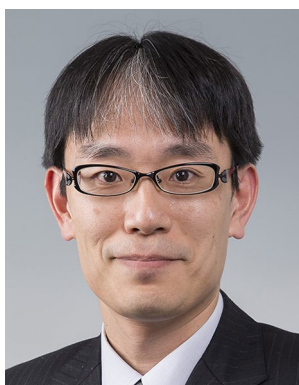
- Weng Y-H (2018) Robot Law 1.0: on social system design for artificial intelligence. In: Barfield W, Pagallo U et al (eds) Research handbook of law and artificial intelligence. Edward Elgar Publishing, Cheltenham
- Sakura O (2018) Socio-cultural shaping of AI/robots in the East Asia. In: The 19th FRIS seminar, 2018-07-10, Tohoku University
- Golem. Wikipedia. <https://en.wikipedia.org/wiki/Golem>
- Morris N (2007) The Golem in Jewish American Literature: risks and responsibilities in the fiction of thane rosenbaum. Peter Lang, New York, p 119
- Hero of Alexandria. Wikipedia. https://en.wikipedia.org/wiki/Hero_of_Alexandria
- Lesley A (2017) Can a robot be religious? World religion news. <https://www.worldreligionnews.com/religion-news/can-robot-religious>
- Siciliano L (2017) A Japanese company just unveiled a robot priest that will read scriptures at Buddhist funerals. <http://uk.buinessinsider.com/japanese-company-unveiled-robot-priest-technology-pepper-soft-bank-robotics-religion-2017-8>
- Weng Y-H (2016) One being for two origins: a new perspective on roboethics. Robohub. <http://robohub.org/one-being-for-two-origins-a-new-perspective-on-roboethics/>
- Crosbie J (2017) This religion is the first to say sex robots are a sin. Inverse. <https://www.inverse.com/article/33751-sex-robot-s-islam-christianity-religion>
- Cargo Cult. Wikipedia. https://en.wikipedia.org/wiki/Cargo_cult
- Mims C (2010) Why Japanese love robots (and Americans fear them). MIT Technology Review. <https://www.technologyreview.com/s/421187/why-japanese-love-robots-and-americans-fear-them/>
- Robertson J (2017) Robo sapiens japanicus robots, gender, family, and the Japanese Nation. University of California Press, Berkeley
- Kaplan F (2004) Who is afraid of the humanoid? Investigating cultural differences in the acceptance of robots. Int J Humanoid Robot 1(3):465–480
- Takanishi A (2016) Humanoid robotics, culture and society of Japan. In: Zielińska T, Zieliński C (eds) Romansy 16, vol 487. CISM courses and lectures. Springer, Vienna
- Nomura T, Syrdal DS, Dautenhahn K (2015) Differences on social acceptance of humanoid robots between Japan and the UK. In: 4th international symposium on new frontiers in human–robot interaction, Canterbury, UK
- The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (2017) Ethically aligned design: a version for prioritizing human well-being with autonomous and intelligent systems, Version 2. IEEE
- Liezi: Chapter 5 Tang wen
- Laozi (4th century BC) Tao Te Ching
- Zhuang Zhou (3rd century BC) Zhuangzi
- Gan Ji (2nd century BC) Tai Ping Ching
- Pengzhi L (2000) Taoism philosophy. Wen Jin Press, Taipei
- Needham J et al (1990) Science and civilization in China, vol 2. Science and Technology Press, Beijing
- Sun Y (2008) Taoism's belief and thoughts. Dong Da Press, Taipei
- Weng Y-H, Hirata Y (2018) Ethically aligned design for assistive robotics. In: IEEE international conference on intelligence and safety for robotics, Shenyang, China
- Bostrom N (2014) Superintelligence: paths, dangers, strategies. Oxford University Press, Oxford
- BMVI (2017) Ethics Commission on automated driving presents report. German Federal Ministry of Transport and Digital infrastructure
- Asimov I (1950) Robot. Ghome Press, New York
- Weng Y-H, Chen C-H, Sun C-T (2009) Toward the human–robot co-existence society: on safety intelligence for next generation robots. Int J Soc Robot 1(4):267–282
- Tallin J (2018) On steering the AI. In: The 14th FRIS seminar, 2017-09-10, Tohoku University
- Kubo N (2000) The gods in Taoism. Heika Press, Tokyo
- Hari-Kuyo. Wikipedia. <https://en.wikipedia.org/wiki/Hari-Kuyo>
- Kubo N (1962) The study on Kou Shin beliefs. Hara Press, Tokyo
- Barfield W (2015) Cyber-humans: our future with machines. Springer, Switzerland
- Unknown editorial (2010) Robotic arm man Christian Kandlbauer dies in hospital after crash. The Guardian
- Sugano S (2011) Dreams that people saw, the way the robot came: from Greek myths and Atoms. JIPM-S, Tokyo
- Yip K-S (2004) Taoism and its impact on mental health of the Chinese communities. Int J Soc Psychiatry 50(1):25–42
- Wallach W, Allen C (2010) Moral machines: teaching robots right from wrong. Oxford University Press, Oxford
- Anderson M, Anderson SL (2018) A prima facie duty approach to machine ethics: machine learning of features of ethical dilemmas, prima facie duties, and principles through a dialogue with ethicists. In: Anderson M, Anderson SL (eds) machine ethics. Cambridge University Press, Cambridge
- Searle J (1980) Minds, brains, and programs. Behav Brain Sci 3(3):417–457
- Anderson M, Anderson SL, Luo D, Weng Y-H (2018) Ethics and care robots. PKU-Berggruen AI Salon, 2018-11-22, Berggruen Institute China Center and Peking University Education Foundation
- Salvini P, Ciaravella G, Yu W, Ferri G, Manzi A, Mazzolai B, Laschi C, Oh S-R, Dario P (2010) How safe are service robots in urban environments? Bullying a robot. In: Paper presented on the 2010 IEEE international symposium on robot and human interactive communication (RO-MAN'10). Viareggio, Italy
- Kanda T, Ishiguro H (2017) Human–robot interaction in social robotics. CRC Press, New York
- Thang LL (2018) Before we give them fuzzy robots, let's try solving elderly loneliness with people, 2018-12-02, QUARTZ, <https://qz.com/1490605/before-we-give-them-fuzzy-robots-lets-try-solving-elderly-loneliness-with-people/>
- Unknown Editorial (2015) Hitchhiking robot is decapitated and left in a ditch after just two weeks. IFLSCIENCE, <https://www.iflscience.com/technology/hitchhiking-robot-manages-two-weeks-us-being-destroyed/>
- Love D (2015) A drunk man's assault on a robot raises unusual legal issues. The Daily Dot, 10.06.2015
- Delvaux M et al (2016) DRAFT REPORT with recommendations to the commission on civil law rules on robotics. European Parliament
- Mind-Body Problem. Wikipedia. https://en.wikipedia.org/wiki/Mind-body_problem
- Fuzzy Logic. Wikipedia. https://en.wikipedia.org/wiki/Fuzzy_logic
- Izumo T (2018) Digital specific property of robots: a historical suggestion from Roman Law, vol 1, issue 1, Delphi—interdisciplinary review of emerging technologies

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Yueh-Hsuan Weng is an Assistant Professor at FRIS, Tohoku University, a Visiting Scientist at the RIKEN Center for Advanced Intelligence Project in Tokyo, Japan and a TTLF Fellow at Stanford Law School. He received his Ph.D. in Law from Peking University, China. He has been appointed as Visiting Assistant Professor for Faculty of Law at The University of Hong Kong (2018), Research Associate, Peking University Law School (2010-2014). He is strongly interested in interdis-

iplinary studies, especially in issues concerning the interface between Artificial Intelligence and Law, including Robot Law, Social Robotics, and Legal Informatics.



Yasuhisa Hirata is a Professor in the Department of Robotics at Tohoku University, Sendai, Japan. He received the BE, ME and PhD degrees in mechanical engineering from Tohoku University, Sendai, Japan in 1998, 2000 and 2004, respectively. His Research interests include intelligent control of multiple mobile robots in coordination, human-robot cooperation system, Assistive Robot and Haptics. He is a member of IEEE, RSJ, SICE and JSME.



Osamu Sakura is a Professor at the Interfaculty Initiative in Information Studies, University of Tokyo and a Team Leader at the RIKEN Center for Advanced Intelligence Project in Tokyo, Japan. He received his PhD in 1992 from Kyoto University, Japan, by behavioral ecology of wild chimpanzees in West Africa. After finishing doctoral thesis, he had shifted his field to rather historical and philosophical areas, especially on the reception of sociobiological ideas in Japan. His recent activi-

ties include AI & culture, popularization and education of sciences.



Yusuke Sugahara is an Associate Professor in the Department of Mechanical Engineering at Tokyo Institute of Technology (Tokyo Tech), Japan. He received his PhD in 2006 from Waseda University. Before he joined the faculty of Tokyo Tech, he had served as a Research Associate at Waseda University (2003-2007), an Assistant Professor at Tohoku University (2007-2012), and a Full-time Lecturer at Kokushikan University (2012-2015). His research interests include design and con-

trol of mechanical and robotic systems.