DLSU
BUSINESS
NOTES AND BRIEFINGS
ISSN (Print): 2345-8216 | ISSN (Online): 2350-6814
Volume 12 | Number 4 JUNE 2024
Ramon V. del Rosario College of Business
CBRD
Center for Business Research & Development

THE ROBO CORP: A.I. ROBOTS IN PHILIPPINE CORPORATE GOVERNANCE

Written by:

Atty. James Keith Heffron
Department of Commercial Law
College of Business
DLSU-Manila

Abstract

The era of Artificial Intelligence (AI) has undeniably arrived, permeating nearly every facet of human existence. As AI develops, it can take on several forms, one of which is robots, defined essentially as physical manifestations of AI that can directly interact with their environment. Following Ray Kurzweil's Law on Accelerating Returns, it is only a matter of time before AI finds its way into the corporate boardroom, participating in decisions that will not only shape the organization it serves but also impact the lives of the people that it affects. The article seeks to examine the legal challenges of having artificial intelligence (i.e. robots) in corporate governance roles, focusing particularly on the 2019 Revised Corporation Code, the premier legislation on corporate governance in the Philippines. Using the doctrinal approach, the paper recommends suitable amendments to the law, creating a general legal framework for artificial intelligence in corporate governance based on the seven guiding principles of the 2024 ASEAN Guide on AI Governance and Ethics.

I. Introduction

"It seems probable that once the machine thinking method had started, it would not take long to outstrip our feeble powers… They would be able to converse with each other to sharpen their wits. At some stage, therefore, we should have to expect the machines to take control." - Alan Turing

Artificial Intelligence and Robotics

There is no one set definition of Artificial Intelligence or AI. Scientists, engineers, and technology opinion makers though agree that artificial intelligence is a broad concept - ranging from technology that can solve simple mathematical problems (i.e. the lowly calculator) to whole sets of algorithms that can allow a machine to predict outcomes and thereby solve problems before they even exist.

In the mid-20th century, scientist Alan Turing essentially developed a system or an "algorithm" that would pave the way for the brain behind the modern-day computer. In his seminal paper *Computing Machinery and Intelligence*, where he introduced the Turing Test, Turing posited that artificial intelligence is where "intelligent" machines would not be taught how to think like adults but rather like children where they could be taught to eventually "think" for themselves (Turing, 1950). This is essentially the beginnings of machine learning. Today, Turing's philosophy and legacy have served as the foundation of modern-day artificial intelligence.

In his article *The AI Revolution: The Road to Superintelligence*, Tim Urban suggested that artificial intelligence is best defined by describing it through categories or what he calls "calibers": Artificial Narrow Intelligence (ANI), Artificial General Intelligence (AGI), and Artificial Super Intelligence (ASI) (Urban, 2015). ANI or "weak AI" is one that is only focused on one area alone and is not programmed to do anything else. Like for instance those chess AI programs. AGI or "strong AI" or "human level AI" is one that can think like a human. Meaning those that can perceive and solve problems like any normal human can. ASI on the other hand is one that has the intelligence that is above that of a human. ASI can learn and tackle problems in a more intelligent way than that of the smartest humans. Urban observes that artificial intelligence is currently at ANI. Humans have already achieved this level from the garden variety phone apps that assist users in everyday living to programs that are used in the military, financial markets, and the medical fields – all having specially designed algorithms tasked at addressing very specific problems and objectives. Urban further observes that this level of AI will not last long. As more developments in hardware happen, from data storage to processing speed, it is not improbable that in a matter of years artificial intelligence will reach AGI and eventually ASI levels. History has proven this as shown by the exponential jump in technology in just one human lifetime or as what futurist Ray Kurzweil calls – the Law on Accelerating Returns (Kurzweil, 2005).

Enter now robots.

It is not surprising therefore that from the time that the first computers were developed in the late 1940s and early 1950s, humans have already found a way to somehow create an evolution from mere formless data computing to robots in a matter of just a few decades. Today, most of these robots have an ANI level, trained to do specific tasks that humans may or may not necessarily do effectively. Ideally, robots are the physical manifestation of artificial intelligence in the real world. Urban calls them a "container" for AI, which may or may not have a human form, with the AI as the "brain" and the robot as the "body". In his article *Robotics and the Lessons of Cyberlaw,* Ryan Calo observes that robots are basically "artificial objects or systems that sense, process, and act upon the world to at least some degree" (Calo, 2015). However, Calo posits that tough this may be a technical definition, robots are more interestingly defined by the following qualities that make them a disruptive form of technology: embodiment, emergence, and social valence. Calo explains that a robot is designed to physically interact with the real world – sensing and acting on and within it – a sort of embodiment, that raw artificial intelligence cannot necessarily do. Further, a robot can be programmed to perform original actions permuted and processed from the data it gathers – a sort of machine learning or autonomy, or which Calo calls emergence. Finally, robots, especially those created with at least familiar human forms, can dictate or even trigger behavioral responses from humans in the same manner as how humans treat each other, which Calo refers to as social valence. These qualities are what perhaps will define how robots will be socially, morally, and legally

treated or "classified" in the future.

As artificial intelligence and robotics technology further evolve, and with Kurzweil's inevitable Law of Accelerating Returns, robotics have now found its way in most aspects of human life. It therefore is not surprising that soon, robotics will nevertheless be a critical facet of human living for which most human activity will entirely be dependent on.

Traditional Corporate Governance

Humans have engaged in commerce ever since the dawn of civilization. This is one of humanity's unique abilities. It is not just a feature to address humanity's need for self-preservation but is a trait attributable to it being a social animal. In the evolution of commerce, humans eventually have figured out that the better way to promote commercial activity (and therefore profit from it) is to do it with the help of others. Over millennia, humans have designed numerous commercial vehicles to help realize this goal. In modern times, humans bonded themselves into groups, having a common aim, adopting certain rules, with the end-goal of promoting a trade or business. As early as the mid-12th to early 13th century, humans have used the term "company" to describe these groups. The word comes from the old French term *compagnie* which means "companionship, consort of persons one with another, intimate association" (Harper, 2001). By the 14th and 15th century, the word evolved to mean "a number of persons united to perform or carry out anything jointly," which was the earliest reference to a trade or business association. (Harper, 2001). By the latter part of the millennia, companies have referred to business associations created by law that may have a distinct legal personality separate from its human members, enjoying certain rights and privileges, and having the unique ability to perpetually exist despite the death of any of its members. In the modern era, such a definition has been referred to as a "corporation", which has been defined as "a specific legal form of organization of persons and material resources, chartered by the state, for the purpose of conducting business" (Britannica money, 2024).

In the Philippines, a corporation is defined as "an artificial being created by operation of law, having the right of succession and the powers, attributes, and properties expressly authorized by law or incidental to its existence" (Republic Act 11232, Section 2, 2019). Corporations are usually created to serve a specific set of objectives. Under the law, these are termed as primary and secondary purposes that must clearly be defined under its Articles of Incorporation ((Republic Act 11232, Section 13, 2019). Corporations, being artificial beings, cannot act on its own. It is not a corporeal body. Thus, the law provides that corporations may generally act through its corporators, - its stockholders (for stock or profit corporations) or members (for non-profit or non-stock corporations (Republic Act 11232, Section 5, 2019).

Corporate governance is defined and regulated under the law. In the Philippines, the primary law in relation to this is Republic Act 11232 or the Revised Corporation Code, a 2019 law that was a revision of the old corporation code that was originally enacted in 1980. Under this law, corporate governance is essentially exercised through a corporation's Board of Directors (Republic Act 11232, Section 22, 2019). Though the law allows artificial beings like corporations to be corporators - even adding in its 2019 revision their capacity to be incorporators - directors however should be natural persons. The powers of directors in relation to corporate governance are likewise defined and limited under the law. Under the Revised Corporation Code, directors have the power to decide on matters relating to three major areas: exercise of corporate powers, conduct of corporate business, and control of corporate properties.

These corporate powers are exclusively defined under the law (Republic Act 11232, Title IV, Sections 35 to 43, 2019). These include the power to sue and be sued in its corporate name; have perpetual existence; adopt and use a corporate seal, amend its articles of incorporation and adopt bylaws; in case of stock corporations, to issue or sell stocks; and to admit members to the corporation if it be a nonstock corporation; to exercise real rights including ownership over real and personal property, to enter into a partnership, joint venture, merger, consolidation, or any other commercial agreement with natural and juridical persons; to make reasonable donations; establish pension, retirement, and other plans for the benefit of its directors, trustees, officers, and employees; and to exercise such

other powers as may be essential or necessary to carry out its purpose or purposes as stated in the articles of incorporation; to extend or shorten the corporate term; increase or decrease capital stock or incur, create or increase bonded indebtedness; deny preemptive rights to its stockholders; to sell or dispose corporate assets; acquire its own corporate shares; to invest corporate funds in another corporation or business or for any other purpose; declare dividends; and enter into management contracts.

As for the conduct of corporate businesses and control of corporate properties, the powers of directors in relation to these are provided for by law and augmented by the corporation's by-laws. In other words, directors cannot act beyond those powers provided for by law and the corporation's articles and by-laws. Any act that violates this policy will be deemed as ultra-vires and therefore shall be considered as void (Republic Act 11232, Section 44, 2019). It will have no legal consequence.

Though corporate governance is primarily exercised through a corporation's set of directors, the law provides that some of these powers may be delegated to and exercised by certain corporate officers as defined under the law such as a President, a Treasurer, Corporate Secretary, and such other officers as provided for in its by-laws (Republic Act 11232, Section 24, 2019).

II. The Case for Robots in Corporate Governance

In 2014, a Hong Kong venture capitalist firm named Deep Knowledge Ventures was on the brink of bankruptcy because their capital investments suffered considerable failures. The company then decided to hire a team of big data analysts who created an artificial intelligence system called VITAL or Validating Investment Tool for Advancing Life Sciences. VITAL assisted management and the board of directors to make logical and sound business decisions specifically on investment risk management. The experiment proved to be a success that VITAL was made part of the management team and appointed as a director on "observer status" with positive board investment decisions being made with its inputs (Burridge, 2017). Today, the trend seems to point to the usage of artificial intelligence in boardroom decisions, especially when those decisions are in relation to capital investments or risks. Many companies, particularly those in the finance and capital markets sectors, are now incorporating artificial intelligence into most of their decision-making processes. Nowadays, decisions are rarely made without the use of big data analytics, with machine learning being one of the key techniques for processing this data.

With artificial intelligence becoming mainstream, is it now feasible to 'contain' AI within a robot that can be appointed as a director or manager on a company's management team?

Advantages

Purely Scientific Decision Making

The most obvious advantage of having a robot as part of the management team is the benefit of precise scientific decision-making based on cold data. As a computer, a robot's mind can seamlessly and accurately make decisions that are in line with the objectives of the company. For instance, if the objective is most profit with the least risk, a robot can efficiently make decisions for a business investment based on these variables with the least margin of error compared to a human. The power of computing of a robot's mind and its ability to efficiently use that power is significantly better than that of a human. Its memory is perfect. It will have no problems or issues in remembering or applying formulas or variable equations. Its use of its algorithms is efficient, and it can be programmed to learn and adapt to novel situations. Therefore, there is that notion that a robot's decision will always be the "smartest" decision since this was reached through pure logical and mathematical calculations.

Decisions are Bereft of Emotions, Bias, and Prejudices

Since robots do not have the weight of emotion or prejudice their actions are perceived to be impartial. Robots make decisions purely on the objectives of their programming. Humans may input all necessary variables and a robot will decide based on the parameters of those variables, nothing more and nothing less. Those variables may be controlled, some may even be left out. Since there is this element of control, a robot will not decide based on a hidden or lacking variable such as those available for humans like emotion or bias.

It also is supposedly ignorant of prejudice. Again, unlike humans, robots have no notions of race, religion, politics, or class unless these are intentionally made variables in their programming. To do so however will of course not be attributable to the robot but to the human, or the deployer, who made such program.

Decisions are Bereft of Fallacious Logic

Since machine learning is usually based on pattern analysis, logical comprehension, and mathematical computations, robot decisions are purely based on acceptable logical inferences. There is simply no room for fallacious conclusions. Humans may fall into the trap of fallacious arguments when making decisions. Their judgment may be clouded or worse influenced by points or arguments that simply do not compute. For instance, a robot cannot be appealed to emotionally. Many injurious decisions have been made by managers based on fallacious appeals to emotions (*argumentum ad misericordiam*) sometimes without the benefit of logical judgment.

Cost-Benefit efficiency

Since robots make decisions with minimal margin of errors, the cost-benefit ratio of its existence, maintenance, and operation is higher than that of human managers. Human managers can make erroneous decisions that may entail huge losses or an inefficient use of resources. Though one can argue that human labor is supposedly cheaper than capital assets, this will even out in the long run when robots make decisions with lesser margin of errors.

Moreover, from a labor perspective, employing robots as directors or managers in corporations will not only enhance decision-making efficiency and increase productivity but will also foster the development of an entire economy centered around robotic technology. For instance, it is theorized that the robotics industry will create more unique jobs than those that the robots intend to replace. Though there may be wide job displacements, this can be solved by a refocus and retraining of personnel (Mathiason, et.al., 2016). In fact, there are several reasons why robotics in the workplace should be seen with hope and not with doom: advances in technology may displace certain types of work but historically they have been a net creator of jobs; we will adapt to these changes by inventing entirely new types of work, and by taking advantage of uniquely human capabilities; technology will free us from day-to-day drudgery, and allow us to define our relationship with "work" in a more positive and socially beneficial way; and ultimately, we as a society control our own destiny through the choices we make (Smith, et.al., 2016).

Overall, these collective advantages have ensured credibility, trust, and integrity in management decisions made by robots. Something that may not necessarily be present in decisions made by its human counterparts. These advantages therefore increase the belief of company stakeholders in the vision and direction of the company. This can have an exponential effect – ranging from an increase in productivity and profits to greater investor trust. These tempting benefits are simply too hard to ignore.

Disadvantages

However, the use of robots in management decision-making is not without its follies. Often, the very advantages that make robots efficient decision makers can be its greatest weaknesses.

Potential for Error

Though robots may be seen as almost infallible when it comes to algorithmic decision making, sometimes a heavy reliance on a robot's perceived algorithmic infallibility may cause humans to overlook its potential capacity for error. At the end of the day, robots perform through its code or programming; a code that was introduced by a human. Humans are of course fallible and are not immuned from introducing codes or programs that are imperfect. Sometimes, such errors in programming may be discovered only later, and may even be masked. This heavy dependence on artificial intelligence on management decisions may therefore cause blind adherence which in turn might cause more harm than good. A good example of this is robot "racism." In 2017, a United States court based its risk assessment analysis on prisoners on a program that eventually turned out to be biased against African Americans (Buranyi, 2023). One may think, how can that happen? A robot does not and cannot perceive biases unless such variable was placed there by its bias programmers. The controversial story created a serious

concern on the credibility of artificial intelligence to make fair and just decisions.

In fact, there seems now to be a glaring basic design issue that may be perceived as a flaw when it comes to the physical appearances of robots. Programmers and designers may wittingly or unwittingly design and program robots that infuse racial biases. Some experts believe that the need of infusion of racial identity in robot design is a legitimate concern for robots to replicate important aspects of human interactions in relation to race based attitudes, beliefs, and behaviors. This in turn may have chilling effects on how robots will act or decide and on how they will be perceived (Bartneck, et.al., 2018).

Robots are Bereft of Moral, Emotional, and Humane Decisions

Emotions are not always a bad thing. Sometimes, even in businesses, there are certain problems that can be resolved using the heart rather than the mind. Take for instance a company that owns land that is being leased for free to an orphanage and another that is being leased for valuable consideration to a manufacturing factory. Both leases expired and the company must decide on what lease to terminate so it can use and invest on the land for its own venture. When the data is fed to the artificial intelligence program, the robot found that terminating the orphanage lease, though it will ensure the closure of the orphanage and the displacement of hundreds of orphans, will nevertheless be more profitable for the company in the short and long run. Terminating the lease with the manufacturing company on the other hand will not just be less profitable but will also close it, having an adverse effect on the hundreds of jobs that it provides. In other words, following the company's directive for maximum profit with the least risk, the termination of the lease of the manufacturing company will be the smartest choice. True, but in situations like these, is the smartest choice really the right choice? This may not necessarily be a direct moral dilemma since there is no exact morally "right" or "wrong" decisions in this example. Either decision is valid, socially, morally, and legally. However, in instances such as these, how would human decision-makers decide? We know what the robot will do. As to the humans? That is something not exactly definite. When the relevant human emotion is exercised, the company's human directors may decide differently from the robots. They may choose to not at all terminate but merely extend both leases for the sake of compassion and mercy – human emotions that a robot will not consider as a variable in its algorithmic processes.

Or take the infamous Trolley Test, a moral thought experiment usually applied to artificial intelligence like self-driving vehicles, whereby a runaway train approaching a fork in its tracks will be heading into either of two paths - one where a person is tied up on one track and another where five are tied up on the other. To save the five, one must control the track to direct the train to where the single person is tied up (Mascini, 2020). Or for instance, what will a robot do when it must make a decision that involves the demise of ten convicted rapists versus one innocent person? Feeding it mere data and following the usual human self-preservation principle or "programming" of greater good for the greater many, it is not surprising that a robot will decide in favor of the five people on the tracks or the ten rapists because this will save more human lives. However, is this moral or even right considering the circumstances?

This is where the unique power of humanity may be overlooked. Certain situations and certain problems are solved not necessarily through the cold and impartial process of mathematical computations. Instead, some decisions do not require logic at all nor do they require math. What they need is a human's ability to instinctively decide based on a certain "code" that has been "programmed" to them by millions of years of physical and social evolution - through something called ***conscience***. A robot does not have it. And there lies the rub.

Data Breach and Security

At the end of the day, a robot is still a machine. A machine created by humans and therefore not immune from malicious tamper. Data breach and hacking is a real threat to the artificial intelligence and robotics industry. It goes without saying that a robot's programming may be hacked, and its code be surreptitiously altered to make decisions based on the hacker's own interests and objectives. Sometimes, these intrusions may again be discovered only after the damage has already been done.

As more people have access to artificial intelligence and cyber technology, there are more chances that this technology will be prone to attacks and security breaches. Further, data that are used for analysis may not likewise necessarily be immuned from data privacy breaches.

Questions on Robot Liability

One of the most controversial issues on robots is the concern for legal liability. When robots make decisions and these decisions turn out to cause losses, or worse, injuries or damages to stakeholders, who then should be liable? Gabriel Hallevy, in his seminal paper entitled *"I, Robot – I, Criminal"—When Science Fiction Becomes Reality: Legal Liability of AI Robots committing Criminal Offenses"* exhaustively tackled and presented theories on the issue of robot liability (Hallevy, 2010). Hallevy presented three models on robot criminal liability. First is the Perpetration by Another model where the robot is referred to as an "innocent agent." As an innocent agent, the robot is treated like an inanimate tool and any act committed by it is in reality committed by its principals - its software/hardware designers, programmers, and/or its user, who in turn should have intended to commit the crime *(mens rea)*. Second, the Natural Probable Consequence liability model, where the liability on the injurious acts of the robots will still be attributed to the programmer, designer and/or user who may not have any knowledge or *mens rea* but nevertheless failed to exercise the required degree of diligence in its programming and/or use. Last is the Direct Liability model where the robot will be responsible on its own when it commits criminal acts, akin to how natural or even artificial (i.e. corporations) persons are liable for criminal offenses. Hallevy states that these models may in fact be used or applied inclusively with each other depending on the circumstances of each case - like for instance when the software or hardware designer or even the user is a robot itself.

Hallevy further points out that robot punishment for criminal offenses or even the legal consequences of tortious acts may not be as clear-cut as one may think. For instance, unlike humans, robots are essentially immortal. To "jail" a robot and deprive it of liberty will be ridiculous since it is not obviously fully self-aware on the significance of its freedom. Even if it were aware, its incapacity to age is an advantage and incarceration or even deactivation will not necessarily affect it. Under the Positivist Theory in criminal law, .criminal punishment tends to be preventive and corrective – it is designed to correct the behavior of the perpetrator so they can eventually be a good member of society when they are released. This may not be as effective as to robots.

Therefore, unlike human directors and officers of a corporation, there is difficulty in attributing liabilities to robots. For human directors the law is clear. When a director commits certain criminal acts or violations of the Corporation Code there are legal mechanisms in place to determine liability and their legal consequences. For instance, under the Classical Theory in criminal law, a director who commits a crime will be liable if he commits it through *rea or culpa*. It is not the act but the malicious intent or the gross lack of care that are punished. These are fundamental principles. As human beings are self-aware, it is easy to attribute liability when humans intentionally commit a crime or fail to exercise the standard degree of diligence. A robot however is not self-aware and therefore does not exactly know or understand the meaning of its actions. It just acts without reflecting if what it is doing is right or wrong. To apply Hallevy's first and second models where liability may be attributable to a robot's programmers or users may not be as effective. Directors are integral members of the corporation and thus inherently have a vested interest in its success. It goes without saying that they will do everything they can to protect that interest, and this logically serves as a valid deterrent for crime or a reason for their exercise of due diligence. Programmers, designers, or users are not necessarily invested in the corporation's interest unlike its stockholders or directors. This then may affect how they program or use these robots. To the corporate stakeholders, a director's inherent interest in the corporation that they serve is an attribute that is necessary to foster stakeholder trust and confidence. No less than the law itself provides that a director should be at least a stockholder of the corporation, a glaring change from the old corporation code when directors may be outsiders (Republic Act 11232, Section 22, 2019).

III. Theoretical Framework: Ethics of Artificial Intelligence

With the revolutionary impact of artificial intelligence in the corporate world, it is crucial that decisions made by AI should align with universally accepted ethical and social norms. In its 4th Digital Ministers Meeting on February 2, 2024, the Association of Southeast Asian Nations (ASEAN) established the ASEAN Guide on AI Governance and Ethics, a living document that provides a set of guidelines for governments and businesses in the use of artificial intelligence. These guidelines serve as a powerful ethical oversight in the use of AI supposedly without compromising its social and economic benefits.

The guidelines are based on seven key principles that are rooted on universal ethical norms: transparency and explainability, fairness and equity, security and safety, robustness and reliability, human centricity, privacy and data governance, accountability and integrity (ASEAN, 2024).

Transparency refers to the duty of organization to inform all stakeholders of the use of AI, which includes its processes, in its decision making. Explainability refers to communicating the reasoning, in a clear manner, behind the decision or recommendations made by AI. These factors ensure the AI's trustworthiness and likewise enables relevant stakeholders to make an informed choice when dealing with organizations that utilize AI. Fairness and equity mean that safeguards must be put into place to ensure that decisions made by AI systems are not biased or discriminatory. These systems must be regularly checked and if at all adjusted as needed. Security and safety refer to placing mechanisms that ensure against malicious cyberattacks and/or unsafe decisions made by AI systems. These mechanisms should likewise include those for the benefit of the human deployers, developers, and users of AI. Robustness and reliability mean that AI systems must be able to cope with any kind of environment especially when there are unexpected errors. They must be reliable and consistent in their outputs and decisions regardless of the data they are made to process. Human centricity means that AI systems should not violate human values and should be geared towards human benefit. AI processes and decisions should never cause any kind of human harm and should instead always promote their safety and well-being. Privacy and data governance ensures that all data processed by AI systems should be protected and that their integrity are secured. This means that AI systems should at all times conform to data privacy regulation. Finally, accountability and integrity mean that all actors involved in the development, deployment, and use of AI should be made accountable for the consequences of AI processes and decisions that may violate any of these guidelines. This includes ensuring that all actors comply with all existing legal regulations in all stages of the AI lifecycle (ASEAN, 2024).

The guidelines emphasize that the management of the use of AI systems in organizations should have a holistic approach. That is taking into consideration all aspects of the AI experience, which includes its development, deployment, and use by human actors, the Internet of Things (IoT), robotics, technological developments, and all others related to it (ASEAN, 2024).

IV. The Legal Hurdles of Robot Governance

Juridical Capacity and Personality

Clearly, the Philippine corporate legal framework does not allow robot directors, managers or officers. There is a basic reason for this. Philippine law does not recognize robots as having a separate and distinct legal personality. Under the Philippine Civil Code, juridical capacity may only be granted to natural persons and to those artificial beings that are exclusively enumerated under the law, namely the State and its political subdivisions, public and private corporations, partnerships and associations having a separate and distinct juridical personality (Republic Act 386, Articles 37 and 44, 1949).

Understandably, since the Civil Code was enacted in 1949, the concept of modern-day artificial intelligence and robotics have yet to exist. Thus, this is the first step. The law should be amended to consider giving limited separate juridical personalities to robots that possesses Calo's three attributes of embodiment, emergence, and social valence. The conferring of legal personalities to robots for purposes of autonomous and independent governance is a valid point. For one, it will legally separate itself from its programmers, designers, and users. Likewise, it will be legally distinct from the corporation itself. It will not be treated as a mere tool or capital asset

but as a separate artificial person with limited rights and powers. This is necessary to address so many legal issues, primary of which is robot liability.

Rights, Powers, and Functions

The rights, duties, and powers of directors and officers of a corporation are clearly enumerated under the Revised Corporation Code. These powers are broad and do not necessarily directly relate to capital investments or profit/loss concerns. Robot directors may be given limited powers or functions that only directly relate to capital investments or business concerns. These functions may also be further limited to those that necessarily entail: (1) the analysis of big data; and (2) within the ANI expertise of the robot director/officer. In other words, there need not be a blanket authority given to a robot director/officer that is the same as that of its human counterpart.

Robot Liability and the Business Judgment Rule

Under the Revised Corporation Code, directors are said to be men of "recognized ability in financial affairs and it is believed that they have long foreseen all possible future contingencies and that under such men the interests of the stockholders are duly protected" (Philippine Supreme Court, 1927). As such, the law has given due weight to their decisions and vowed not to interfere as long as these decisions are made in good faith. Thus,

> A corporation is but an association of individuals, allowed to transact under an assumed corporate name, and with a distinct legal personality. In organizing itself as a collective body, it waives no constitutional immunities and perquisites appropriate to such a body. As to its corporate and management decisions, therefore, the state will generally not interfere with the same. Questions of policy and of management are left to the honest decision of the officers and directors of a corporation, and the courts are without authority to substitute their judgment for the judgment of the board of directors. The board is the business manager of the corporation, and so long as it acts in good faith, its orders are not reviewable by the courts (Philippine Supreme Court, 1997).

This legal principle is the *business judgment rule*. The rule emphasizes on the human attribute of good faith. Meaning as long as a director exercised due diligence and was in good faith when they came out with a decision, then the decision cannot be reviewed or overturned by the courts, nor the director be liable even if such a decision proved to be damaging to the corporation.

It is difficult to straightly apply the *business judgment rule* to robot directors under its present context in Philippine law. For one, how can the concept of good faith, a human attribute, be ascertained and ascribed on a robot? Following Hallevy's first and second models, it seems that the concept of good faith can only be applied to the humans behind the robots – the programmers, designers, and the users. However, since these people are (1) not necessarily part or stakeholders of the corporation having direct interests thereon; and (2) not the decision makers but merely enablers of the robot director to make the decision, it is therefore difficult to apply this rule to them. As to the third model where the robot is solely liable for its actions, then again one goes back to the basic question on how a robot's "good faith" can be ascertained let alone established.

V. Conclusion and Recommendation

Artificial intelligence is not infallible. Therefore, there is still a need to have a system in place that could serve as a check and balance for processes and decisions made by robot directors. This protocol naturally should be supported by law. Following the guiding principles of the ASEAN Guide on AI Governance and Ethics, a suitable legal framework should be adopted and incorporated under Philippine law with at least the following attributes:

Transparency and Explainability

If a corporation chooses to have a robot director, then it should be mandatory that stakeholders such as directors and stockholders/corporators be informed on the capacity, processes, and limitations of that robot. In other words, transparency should be an inherent right of every stockholder/corporator in relation to all

information that will affect the robot's decision-making. Failure of the company to comply should be met with stiff penalties (ICGN, 2020). Such information should also be likewise promptly reported and included in the corporation's reportorial requirements before government regulatory bodies such as the Securities and Exchange Commission, especially if the corporation is publicly listed.

Further, there should be a mandatory requirement that at least one human director should understand and be sufficiently well-versed on the algorithms and processes of the robot director, and in those moments of decision making, the corporation can make it mandatory for that director to further discuss and explain the processes and analyses that were used to come up with the decision. It can be mandated as an obligatory inclusion in the official minutes of the meetings of stockholders/corporators and the board of directors. (Republic Act 11232, Sections 35 to 43, 2019).

Fairness And Equity

One of the key components of the ASEAN Guide on AI Governance and Ethics centers on the adoption of internal policies and measures that would ensure the responsible use of AI in governance. As recommended by the ASEAN Guide, corporations using AI in governance should have an internal AI Ethics Advisory Committee as part of top management and this committee would ensure that AI systems do not render decisions that are bias or discriminatory. This includes the data sets that these systems process to produce these decisions. The AI algorithms used in both the processes and decisions should be in line with and not violate anti-discriminatory legislation. A corporation's adoption of these measures should be made mandatory under the law.

Security and Safety

The law should mandate that a corporation should set measures that would ensure the security and safety of all its stakeholders. Examples of these measures are protocols against malicious cyberattacks, deepfakes, and rules on the use of generative AI. Human oversight protocols should be set in place to legally override AI decisions. This can be done by retaining the usual legislated concept of stockholder/corporator oversight through votation. To be specific, for board decisions participated in by a robot director, the law can make stockholder/corporator concurrence mandatory and instead of the mandated two-thirds overturning vote, the threshold can be lowered (i.e. majority rule). This system will ensure that stockholders/corporators will have an easier way to junk robot decisions especially if these are decisions against general sentiment.

Robustness and Reliability

Just like any highly technical industry where systems are subject to stringent regulation, corporations that use AI in governance should likewise be subject to strict technical protocols. Measures should be mandated to guarantee the reliability and safety of the AI system for the purpose of preventing harm to all stakeholders. Periodic maintenance checks and strict reportorial requirements on AI system safety and reliability should be mandated. For instance, under the law, corporations are required to submit annual reports of their general and financial status to the SEC. If a corporation is listed, additional reportorial requirements are mandated to be submitted before the Philippine Stock Exchange. Periodic reports on the integrity of AI systems can be incorporated in these reports to keep the public and all other stakeholders updated of their status. This requirement serves a two-prong purpose - it ensures that corporations will see to it that the reliability of their AI systems are solid and up to date and it enables all stakeholders to make informed choices when dealing with them.

Human centricity

The processes and decisions of robot directors should be aligned with human values and not, at all times, contravene law, morals, good customs, public order and safety. To achieve this, corporations that use AI in governance should be subject to licensure the way certain corporations or businesses are subject to government permission before they can be allowed to operate. This way, government agencies such as the Department of Information and Communication Technology together with the SEC and other like agencies can be mandated to ensure that the algorithm, programs, and processes of these AI systems are in line with law. These corporations, together with its AI deployers can be

mandatorily subject to stringent government certifications to ensure the public's trust in their decision making.

Privacy and data governance

With the passage of the Data Privacy Act of 2012, it is the State's responsibility to ensure that *personal information in information and communications systems in the government and in the private sector are secured and protected (*Republic Act 10173, 2012). Thus, AI systems in corporate governance, including robot directors, should be mandated to strictly adhere with the provisions of the Data Privacy Act and all other related data privacy regulation. Failure to do so should subject not just the corporation itself, but the deployers and users of these AI systems, to strict penalty in accordance with the law.

Accountability and integrity

A separate set of policies should be in place to ascertain liability of robots that will be part of corporate governance. The *business judgment rule* may be tweaked to consider establishing guidelines to ascertain what exactly shall constitute good faith and due diligence in relation to robot decisions. For instance, Hallevy's first and second models may be adopted as fundamental policy to determine liability. Further, the other sitting human directors who agreed with the robot director's decision may be held solidary liable if it can be established that they failed to exercise the required diligence or committed bad faith in their acquiescence.

VI. Future Research and Policy Direction

Indeed, there is no denying that the age of the robot is here. Keyser Soze, in the seminal 1995 film The Usual Suspects - taking inspiration from French poet Charles Baudelaire - declared, "the greatest trick the devil ever pulled was convincing the world he didn't exist." The last thing we want is for AI and robots to trick us in thinking the same way.

Every aspect of human life from personal and intimate activities to activities that can shape the destiny of humankind are assisted or even directed in one way or the other by artificial intelligence. It is not therefore hard to imagine that the day will come that corporations would be run top-level by robots. The time therefore to create legislation is now. As discussed earlier, specific legal recommendations may range from laws defining robot liability, ensuring the reliability and safety of AI systems, to human oversight legal protocols. Future legislation can turn to the guiding principles of the ASEAN Guide on AI Governance and Ethics as a welcome start.

References:

ASEAN, (2024), ASEAN Guide on AI Governance and Ethics, ASEAN Main Portal. Retrieved June 2, 2024, from https://asean.org/book/asean-guide-on-ai-governance-and-ethics/.

Bartneck, C., Yogeeswaran, K., Qi. S, Woodward, G., Sparrow, R., Wang, S., Eyssel, F. (2018), Robots and Racism, In Proceedings of 2018 ACM/IEEE International Conference on Human Robot Interaction (HRI '18), ACM, New York, NY, USA. Retrieved June 2, 2024, from https://doi.org/10.1145/3171221.3171260.

Britannica Money. (2024). Retrieved on June 2, 2024, from https://www.britannica.com/topic/corporation.

Buranyi, S. (2023), Rise of the racist robots – how AI is learning all our worst impulses, The Guardian. Retrieved June 2, 2024, from https://www.theguardian.com/inequality/2017/aug/08/rise-of-the-racist-robots-how-ai-is-learning-all-our-worst-impulses.

Burridge, N. (2017), Artificial intelligence gets a seat in the boardroom, Nikkei Asia. Retrieved June 2, 2024, from https://asia.nikkei.com/Business/Artificial-intelligence-gets-a-seat-in-the-boardroom.

Calo, R. (2015), Robotics and the Lessons of Cyberlaw, 103 Calif. L. Rev. pages 513-63.

Hallevy, G. (2010), I, Robot – I, Criminal—When Science Fiction Becomes Reality: Legal Liability of AI Robots committing Criminal Offenses, 22 Syracuse Science & Tech. Law Reporter 1 (Spring 2010), pages 7-37.

Harper, D. (2001), Company, Etymology. Retrieved on June 2, 2024, from https://www.etymonline.com/word/company.

ICGN (2020), Artificial Intelligence and Board Effectiveness. Retrieved June 2, 2024, from https://www.icgn.org/artificial-intelligence-and-board-effectiveness.

Kurzweil, R. (2005), The Singularity Is Near: When Humans Transcend Biology. Penguin Books, New York.

Mascini, L. (2020), It's utopian to think that people can make infallible AI, IO. Retrieved June 2, 2024, from https://innovationorigins.com/en/its-utopian-to-think-that-people-can-make-infallible-ai/.

Mathiason, G., Pierce, N., Cerilli, J., Gordon, P., Kennedy, P., Lee, T., Lotito, M., Notestine, K., Ryu, E., Schuman, I., Weiner, P., Weissman, W., Wolff, R., Brown, G., Hwang, J., Mossavar, M., Ros, S. (2016) Littler Reports: The Transformation of the Workplace Through Robotics, Artificial Intelligence, and Automation: Employment and Labor Law Issues, Solutions, and the Legislative and Regulatory Response. Retrieved June 2, 2024, from https://www.littler.com/files/2016_wp_transformation_of_the_workplace_through_robotics_ai_and_automation_2.pdf/.

Philippine Supreme Court (1927), The Government of the Philippine Islands v. El Hogar Filipino, G.R. No. L-26649.

Philippine Supreme Court (1997), Philippine Stock Exchange, Inc., v. The Honorable Court of Appeals, Securities and Exchange Commission and Puerto Azul Land, Inc., G.R. No. 125469.

Republic Act 10173, Data Privacy Act, 2012.

Republic Act 11232, Revised Corporation Code, (2019).

Republic Act 386, Civil Code of the Philippines, 1949.

Smith, A. and Anderson, J. (2014), AI, Robotics, and the Future of Jobs. Retrieved June 2, 2024, from https://www.pewresearch.org/internet/2014/08/06/future-of-jobs/

Turing, A. (1950), Computing Machinery and Intelligence, Mind, Volume LIX, Issue 236, pages 433–460.

Urban, T. (2023), The AI Revolution: the road to superintelligence. Wait but why. Retrieved June 2, 2024, from https://waitbutwhy.com/2015/01/artificial-intelligence-revolution-1.html.