Article

Vital, Sophia, and Co.—The Quest for the Legal Personhood of Robots

Ugo Pagallo

Law School, University of Turin, Lungo Dora Siena 100, 10153 Turin, Italy; ugo.pagallo@unito.it;
Tel.: +39-011-6706903

Abstract: The paper examines today's debate on the legal status of AI robots, and how often scholars and policy makers confuse the legal agenthood of these artificial agents with the status of legal personhood. By taking into account current trends in the field, the paper suggests a twofold stance. First, policy makers shall seriously mull over the possibility of establishing novel forms of accountability and liability for the activities of AI robots in contracts and business law, e.g., new forms of legal agenthood in cases of complex distributed responsibility. Second, any hypothesis of granting AI robots full legal personhood has to be discarded in the foreseeable future. However, how should we deal with Sophia, which became the first AI application to receive citizenship of any country, namely, Saudi Arabia, in October 2017? Admittedly, granting someone, or something, legal personhood is—as always has been—a highly sensitive political issue that does not simply hinge on rational choices and empirical evidence. Discretion, arbitrariness, and even bizarre decisions play a role in this context. However, the normative reasons why legal systems grant human and artificial entities, such as corporations, their status, help us taking sides in today's quest for the legal personhood of AI robots. Is citizen Sophia really conscious, or capable of suffering the slings and arrows of outrageous scholars?

Keywords: artificial intelligence; legal agent; liability; personhood; responsibility; robotics

---

1. Introduction

The legal personhood of robots has been a popular topic of today's debate on the normative challenges brought about by this technology. In 2007, for example, Carson Reynolds and Masatoshi Ishikawa explored the scenarios of Robot Thugs, namely, machines that choose to commit and, ultimately, carry out a crime: their aim was to determine whether and to what extent these machines can be held accountable [1]. Three years later, I expanded this analysis on agency and criminal responsibility, to the fields of contracts and extra-contractual liability [2]. In homage to Reynolds and Ishikawa's creature *Picciotto Roboto*, my next paper then provided a concise phenomenology on how smart AI systems may affect pillars of the law, such as matters of criminal accountability, negligence, or human intent [3]. In 2013, I summed this analysis up with my monograph on The Laws of Robots [4]. There, I suggested a threefold level of abstraction, so as to properly address today's debate on the legal personhood of robots and smart AI systems, that is:

(i) The legal personhood of robots as proper legal "persons" with their constitutional rights (for example, it is noteworthy that the European Union existed for almost two decades without enjoying its own legal personhood);

(ii) The legal accountability of robots in contracts and business law (for example, slaves were neither legal persons nor proper humans under ancient Roman law and still, accountable to a certain degree in business law);

(iii)　New types of human responsibility for others' behaviour, e.g., extra-contractual responsibility or tortuous liability for AI activities (for example, cases of liability for defective products. Although national legislation may include data and information in the notion of product, it remains far from clear whether the adaptive and dynamic nature of AI through either machine learning techniques, or updates, or revisions, may entail or create a defect in the "product").

Against this framework, the aim of the paper is to shed further light on such threefold status that AI robots may have in the legal domain, by taking into account what has happened in this domain of science, technology, and their normative challenges over the past years. Whereas most legal systems, so far, have regulated the behaviour of AI robots as simple tools of human interaction and hence, as a source of responsibility for other agents in the system [4], have advancements of technology affected this traditional framework? Do certain specimens of AI technology, such as smart humanoid robots, recommend that we should be ready to grant some of these robots full legal personhood and citizenship? Or, would such legislative action be morally unnecessary and legally troublesome, in that holding AI robots accountable outweighs the "highly precarious moral interests that AI legal personhood might protect" [5]?

To offer a hopefully comprehensive view on these issues, the paper is presented in three parts. First, focus is on current trends of AI technology and robotics, so as to stress both benefits and threats of this field. Then, attention is drawn to the confusion that prevails in most of today's debate between the legal personhood of AI robots and their legal accountability in contracts and business law. Finally, the analysis dwells on the pros and cons of granting AI robots full legal personhood, as opposed to the status of legal accountability, or as a source of responsibility for other agents in the legal system. At the risk of being lambasted for reactionary anthropocentrism, the conclusion of the paper is that such a quest for the legal personhood of AI robots should not have priority over the regulation of more urgent issues triggered by the extraordinary developments in this field.

2. Current Trends of Robotics

To shed light on today's debate on the legal personhood of AI robots, it is important to briefly put this debate into perspective. Robots materialized as a reprogrammable machine operating in a semi- or fully autonomous way, so as to perform manufacturing operations, more than fifty years ago. Inspired by the research of George Devol and Joseph Engelberger, robots were deployed in the automobile sector since 1961, i.e., the UNIMATE robot removing die-casting and performing spot welding in a General Motors factory in New Jersey. In the early 1980s, such a use of robots within the automobile sector became critical. The Japanese industry attained a strategic competitiveness through the large-scale use of this technology in their factories, reducing their costs and increasing the overall quality of their cars. This trend went on until the early 2000s, when certain individuals still had the impression that robotics was too dependent on the automobile industry. Remarkably, in the *Editorial* to the World 2005 Robotics Report of the Economic Commission for Europe and the International Federation of Robotics, Åke Madesäter raised this risk: "In the period 1997–2003, the automotive industry in Spain received 70% of all new robot installations. In France, the United Kingdom and Germany the corresponding figure amounted to 68%, 64% and 57%, respectively" [6].

In the same years as covered by the UN World report, however, the field of robotics opened up to a profound transformation, a "revolution", according to many scholars [7,8]. The first step of this diversification concerned the set of water-surface and underwater unmanned vehicles, or "UUVs", employed for remote exploration work and the repairs of pipelines, oil rigs and so on. This set of robotic applications started developing an amazing pace since the mid-1990s. Then, it was the turn of unmanned aerial vehicles ("UAVs"), or systems ("UAS"), that upset the military field in the mid-2000s [9]. A few years later, time was ripe for the advent of self-driving cars: whereas the Nevada Governor signed a bill into law in June 2011 that for the first time ever authorized the use of driverless cars on public roads, other states in the U.S. soon followed suit. In September 2017, the House of Representatives passed a bill, the Self Drive Act, which aims to provide a much-needed federal

framework for the regulation of autonomous vehicles. While the panoply of robots available out there suggests further candidates for the next robotic revolution in the field of service applications for personal and domestic use, such as robots for home security and surveillance, for handicap assistance, or just for fun and entertainment, we should not miss a crucial, twofold aspect of this trend. On the one hand, robots are progressively connected to the Internet: avoiding the shortcomings of traditional approaches, such as on-board computers for robots, the troubles with the computing power of such machines have increasingly been addressed by connecting them to a networked repository on the Internet, allowing robots to share the information required for object recognition, navigation and task completion in the real world. On the other hand, the field of robotics is more frequently intertwined with advancements of artificial intelligence ('AI'), to such an extent that even the definition of robot has evolved over the past decades. Some argue that we are dealing with machines built basically upon the mainstream "sense-think-act" paradigm of AI research [10]. Sebastian Thrun, former director of the AI Laboratory at Stanford, California, similarly reckons that robots are machines with the ability to "perceive something complex and make appropriate decisions" [8] (p. 77). Although we still have not obtained either machines that are capable of doing any work men can do, or the solution for the problem of creating proper artificial intelligence within "25 years" [11], or "the current generation" [12], we are dealing with machine-learning systems that (i) increasingly define or modify their decision-making rules autonomously; (ii) improve their knowledge and skills through the interaction with other artificial agents, smart things, or human fellows in the surrounding environment; and, (iii) respond to the stimuli of such environment, by modifying their own properties, or inner states [4]. Among the ingredients that made the convergence between computer sciences, AI and robotics possible, we should list the improvement of more sophisticated statistical and probabilistic methods, the growing availability of huge amounts of data and of massive computational power, up to the transformation of places and spaces into IT-friendly environments, e.g., smart cities and domotics.

As to the specimens of such smart machines like Blue Frog Robotics' Buddy, SoftBank's Pepper, or Asus's Zenbo, it is worth mentioning two further applications in this context. The first one is Vital, a robot developed by Aging Analytics UK, who was appointed in May 2014 as a board member of the Hong-Kong venture capital firm Deep Knowledge. The reasons for this appointment hinged on the ability of Vital to foretell good investments in the field of therapies for age-related syndromes, pinpointing market trends that otherwise would be under the human radar. Whereas AI machines will sensibly improve their adaptive decision-making rules over the next years, it seems fair to admit that trends of humans delegating complex cognitive tasks to robots and AI systems will reasonably increase as well. For example, in October 2016, a Finnish OMX-listed company, Tieto, appointed Alicia T, an AI expert system, as a member of the leadership team of its new data-driven businesses unit. According to Tieto's website, Alicia T. will not only become a full-fledged member of the management team, but also possess the capacity to cast votes: "AI will help the management team to become truly data-driven and will assist the team in seeking innovative ways to pursue the significant opportunities of the data-driven world".

Then, we have the case of Sophia, a social humanoid robot developed by Hong Kong-based company Hanson Robotics, in collaboration with Google's parent company Alphabet and SingularityNET, which provide for Sophia's voice recognition system and AI software, respectively. Activated in April 2015, Sophia made her first public appearance in Austin, Texas, in March 2016. As the Wikipedia entry claimed in early 2018, "interviewers around the world have been impressed by the sophistication of many of Sophia's responses to their questions, (although) the bulk of Sophia's meaningful statements are believed by experts to be somewhat scripted". Wikipedia entry has meanwhile been updated, insisting now on the "controversy over hype in the scientific community". Nonetheless in October 2017, as a matter of fact, Sophia became the first AI application to receive citizenship of any country, namely, Saudi Arabia; a month later, 'she' was named the first Innovation Champion of the United Nations Development Programme, the first non-human to be given any UN title.

Obviously, one may wonder why on earth Saudi Arabia has enrolled Sophia as a citizen of her own, much as the UN celebrating Sophia as an innovation champion. Yet, some months earlier, in February 2017, the European Parliament adopted a proposal, in which the EU institution invites the European Commission "to explore, analyze and consider the implications of all possible legal solutions, (including) … creating a specific legal status for robots in the long run, so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause, and possibly applying electronic personality to cases where robots make autonomous decisions or otherwise interact with third parties independently" (§59f of the document).

Admittedly, current trends of AI and robotics have suggested some pessimistic views. In 2015, for instance, the Future of Life Institute released an open letter addressing the challenges and threats posed by this technology: "Its members—and advocates, among which Bill Gates, Elon Musk, and Stephen Hawking—are concerned that as increasingly sophisticated achievements in AI accumulate—especially where they intersect with advances in autonomous robotics technology—not enough attention is being paid to safety". A year later, the White House Office of Science and Technology Policy (OSTP) conducted a series of public workshops on questions of AI and policy, releasing a final report on how to tackle such issues, as fairness, accountability, or social justice, through means of transparency [13]. While the European Parliament's Committee on Legal Affairs and the UK House of Commons presented similar reports in 2017, an Industry Connections Program within the IEEE Standards Association issued its own document, namely, *The Global Initiative on Ethical of Autonomous and Intelligent Systems* from December 2017, in which the normative challenges and ethical threats of this kind of technology are similarly taken into account. In light of the manifold AI robotics applications and of the multiple, and even opposite, normative views of legislators, experts, and opinion makers, on whether or not legal systems should grant AI robots their "electronic personhood", is there any chance to orient ourselves?

The next step of the analysis has to set the proper level of abstraction, in order to take sides in today's debate. From a methodological viewpoint, the aim is to determine the interface that makes an investigation of some crucial aspects of the legal system possible, so as to comprise a set of features representing the observables and variables of the analysis, the result of which provides a model for the field [4]. From a substantial perspective, we should distinguish the analysis of the technology that is subject to legal regulation, and the set of legal notions that are at stake with matters of accountability, liability, and responsibility. Next sections explore this twofold aspect of the problem separately.

3. Levels of Abstraction

A main source of misunderstandings in today's debate on the legal personhood of AI robots has to do with the ways in which the different kinds of issues, interests, or goods that are at stake with their behaviour, are confused in a Hegelian night where all kinds of legal status look grey. Going back to the aforementioned European Parliament's proposal from February 2017, for example, it is unclear whether "the status of electronic persons" refers to the full legal personhood of robots as proper legal "persons", or regards their legal accountability in contracts and business law, or both. This confusion reappears with certain scholars. Some claim, "that for a computer agent to qualify as a legal agent it would need legal personhood. Both meanings of 'agency' raise questions as to the desirability of legal personhood of bots" and other artificial agents such as robots [14]. Others argue that granting robots legal personhood would prevent "the debates over slavery" that "remind us of uncomfortable parallels with the past" and "reflect ongoing tension over humanity's role in an increasingly technologized world" [15] (p. 186). More recently, this confusion between legal personhood and accountability of AI robots reappears in [5]. Here, the reason legal systems should not confer legal personhood on "purely synthetic entities" has to do with moral reasons and the abuse of legal person status by robots and those that make them, i.e., either robots as liability shields, or robots as themselves unaccountable rights violators.

A proper level of analysis has thus to be set, in order to stop comparing apples and oranges, namely, the apples of legal accountability and the oranges of legal personhood. (In this context, a third hypothesis on AI robots as a source of responsibility for other agents in the system is set aside, e.g., current work for the amendment of the EU directive no. 374 from 1985 on liability for defective products.) Correspondingly, attention should be drawn to three different kinds of reasoning, which can be summed up as follows: (i) if apples, i.e., robots as accountable agents, then we have oranges; (ii) if apples, then we should have oranges; and, (iii) if we do not have apples, then neither oranges as a result.

As to the first kind of argument, according to which the legal agency of AI robots would require their legal personhood [5,14,15], it is not necessary to resort to the example of the legal status of slaves under ancient Roman law to show that forms of dependent or restricted legal status, such as agents in contract law, are not essentially intertwined with forms of independent legal personhood. For example, the European Union existed for almost two decades without enjoying its own legal personhood. Therefore, scholars may discuss about different types of apple, namely, registering AI robots like corporations, or bestowing them with capital, or making the financial position of such smart machines transparent, without resorting to any kind of AI personhood. From the different types of apples under scrutiny in today's research, in other words, it does not follow that AI robots necessarily turn out to be an orange, thereby enjoying some form of full legal personhood.

The second kind of argument is normative, for it claims that, once AI robots are conceived of as agents in contracts and business law, then they should be treated as legal persons. The normative ground of this reasoning rests on the reasons why legal systems grant human persons full legal personhood. As we will see later in the next section, such reasons have to do with the moral status of humans, their intrinsic worth and capability to suffer, their consciousness, and so forth. In the tradition of human rights declarations, the reference value is given by the idea of dignity, e.g., Article 1 of the Universal Declaration of Human Rights (UDHR), and Protocol 13 of the European Convention on Human Rights (ECHR). The problem with this second kind of argument is that different types of AI robotic agenthood can be examined, regardless of whether or not such artificial agents are conscious, capable to suffer, or experience emotions, desires, pleasures, or pain. What is at stake with the legal agenthood of AI robots and their accountability has to do with matters of efficiency in transactions and economic affairs, rather than any kind of AI robotic dignity. Advocates of the argument, according to which the legal agency of AI robots should require their legal personhood e.g., [15], have thus to preliminarily demonstrate that such artificial agents possess some of the requisites that legal systems usually take into account, in order to grant humans their full legal personhood.

Yet, there is a variant of the second argument, which likens the status of robots to the legal personhood of corporations. In both cases, so goes the reasoning, once we admit that AI robots and corporations are agents in contracts and business law, then—for reasons that hinge on the efficiency of economic affairs and transactions—they should be considered as full legal persons. Admittedly, the idea of registering AI robots just like corporations is popular among scholars [16–18]. As claimed by others, moreover, "existing laws might provide a potentially unexpected regulatory framework for autonomous systems" [19]. According to this view, we should not amend the current law, to admit that AI robots may "inhabit" a company and "thereby gain some of the incidents of legal personality" [19].

However, going back to the variant of our second argument, there are three problems. First, the corporate solution for the legal agenthood of AI robots is only one among several technical options that scholars have suggested over the past years, in order to tackle problems of accountability for AI robotic behaviour. Scholars have proposed registries for artificial agents, insurance policies, or modern forms of the ancient Roman legal mechanism of *peculium*, namely, the sum of money or property granted by the head of the household to a slave or son-in-power [4]. What all these cases illustrate is that legal systems can properly address the challenges of the agenthood of AI robots in contracts and business law, without embracing any form of corporation and hence, any kind of legal personhood of AI robots. In light of the examples in the previous section, we can thus say that the status of Vital or

Alicia T. may even make sense, without endorsing any kind of citizenship status for Sophia. Second, the extent of the legal personhood of corporations dramatically varies among legal systems. Contrary to the US tradition, for example, most EU companies do not enjoy their own privacy rights, or their own political rights, such as freedom of speech [20]; corporations cannot be held criminally responsible [21], and so forth. This latter scenario is at odds with that which advocates of the legal personhood of AI robots usually claim: at least in Europe, the corporate solution for the legal personhood of AI robots would be a Pyrrhic victory. Third, we have to take into account the opinion of those who oppose granting robots the status of legal persons just like corporations. According to this view, "there are two kinds of abuse that might arise at the expense of human legal rights—humans using robots to insulate themselves from liability and robots themselves unaccountably violating human legal rights" [5] (p. 285).

　　　This latter argument on the "two abuses" can be understood both as a critique of the just like corporation-view and as an illustration of the third kind of confusion in today's debate between the legal agenthood and the legal personhood of AI robots. As to the first aspect of this stance, its advocates claim, the personhood of artificial agents could be a means to shield humans from the consequences of their conduct. In light of the *International Tin Council case* before the House of Lords in October 1989, "the risk (is) that electronic personality would shield humans actors from accountability for violating rights of other legal persons, particularly human or corporate" [5] (p. 287). Although this possibility is for real, we should also pay attention to the other way around, namely, cases in which the intricacy of the interaction between humans and computers can make it extremely difficult to ascertain what is, or should be, the information content of the natural or artificial entity, as foundational to determining the responsibility of individuals. Such cases of distributed responsibility that hinge on multiple accumulated actions of humans and computers may lead to cases of impunity that already have recommended some legal systems to adopt new forms of criminal accountability. Think of the collective knowledge doctrine, the culpable corporate culture, or the reactive corporate fault, as ways to determine the blameworthiness of corporations and their autonomous criminal liability [22].

　　　Still, in addition to the risk of AI robots as liability shields, advocates of the "two abuses"-doctrine raise the further threat of robots as themselves unaccountable right violators. In a nutshell, the problem revolves around who could represent the artificial agent in a legal dispute and moreover, how we should deal with issues of robot insolvency. Although as we have seen above in this section, legal systems could establish mechanisms for AI robots to own property or hold accounts, much as requiring the creators of robots to place initial funds in such accounts, "money can flow out of accounts just as easily as it can flow in; once the account is depleted, the robot would effectively be unanswerable for violating human legal rights" [5] (p. 288). Traditional punitive sanctions of the law, such as jail time for criminal insolvency, would be unavailable, unsatisfying, or ineffective. As a result, we may envisage the malfunctioning of AI robots or their manipulation that cause or fuel human wrongdoing, if not properly detected and recovered, thus making people vulnerable to systematic recourse to such artificial systems. In addition, we should expect a novel wave of AI crimes and wrongdoing, after the 1990s generation of computer crimes set up by national legislators, such as new forms of AI Ponzi schemes [22].

　　　There are, however, two further problems with this narrative on AI robots representing a threat as unaccountable right violators. The first issue is empirical, and has to do with scholarly work and legislative measures as to how to hold such artificial agents accountable. Whereas, as previously stressed, scholars have suggested different kinds of strategies, as registries, insurance policies, modern forms of *peculium*, and the like, some institutions, as the Japanese government, have worked out a way to address these issues through the creation of special zones for robotics empirical testing and development, namely, a form of living lab, or *Tokku*. Significantly, the special zone of Tsukuba was set up in 2011, in order to understand how AI safety governance and tax regulation could be disciplined [23]. Thus, we can dismiss this part of the "two abuses"-doctrine as an empirical issue

concerning how legal systems could properly keep the legal agenthood of AI robots under control, and make them accountable.

However, how about the further claim of the "two abuses"-doctrine? Should we buy the idea that once AI robots are a mere liability shield, or rather potential unaccountable rights violators in contracts and business law, no legal personhood shall follow as a result?

The fallacy of the argument concerns once again the confusion between apples and oranges, that is, between the legal agenthood of AI robots in contracts and business law, and their legal personhood. There are several instances of how legal systems might grant rights of personhood, independently of matters that regard accountability in economic affairs. As to the rights of human persons, think about minors and people with severe psychological illnesses, who cannot be deprived of their legal personhood as espoused in certain rights despite such illnesses, or emotional and intellectual immaturity, as occurs with e.g., the 1989 UN Convention on the Rights of the Child. As to the set of legal persons, consider the legal personhood that is enjoyed by such non-human entities, as the Whanganui river and Te Urewera national park in New Zealand, the Ganges and the Yamuna rivers in India, up to the entire ecosystem in Ecuador. Therefore, the question is not about whether the legal agency of AI robots would—or should—require their legal personhood, or whether the legal personhood of AI robots should vice versa be subject to their accountability in contracts and business law. Rather, the problem has to do with the reasons why legal systems usually grant humans their legal personhood, and whether AI robots meet such requirements.

4. AI as Legal Persons

The previous section has stressed some of the reasons why legal systems usually grant humans their legal personhood. According to the philosophical stance, or ideological options scholars adopt, such motives are often referred to either the moral status of humans and the protection of their dignity, or their capability to suffer, along with further elements, such as human consciousness, intentionality, desires, and interests. Some of these requisites, e.g., capability to suffer and (a certain degree of) consciousness, have been evoked to extent the sphere of legal personhood to other natural agents, such as animals [24]. Others have debated whether such extension could comprise some artificial agents. In his seminal 1992 article on the *Legal Personhood for Artificial Intelligences*, for example, Lawrence Solum examines three possible objections to the idea of recognizing rights to those artificial agents, or intelligences (AIs), namely, the thesis that "AIs Are Not Human" [25] (pp. 1258–1262); "The Missing-Something Argument" [25] (pp. 1262–1276); and, "AIs Ought to Be Property" [25] (pp. 1276–1279). Remarkably, according to Solum, there are no legal reasons or conceptual motives for denying the personhood of AI robots: the law should be entitled to grant personality on the grounds of rational choices and empirical evidence, rather than superstition and privileges. "I just do not know how to give an answer that relies only on a priori or conceptual arguments" [25] (p. 1264).

We may accept Solum's argument that scholars cannot, on conceptual grounds, rule out in advance the possibility that AI robots should be given the status of legal personhood; still, we have to face three different kinds of problem with this new legal status. First, attention should be drawn to "the missing-something problem". Although certain scholars claim that AI robots would already have the capability of fulfilling the awareness requirements in criminal law, together with "the mental element requirements of both intent offenses and recklessness offenses" [26] (p. 99), current AI robots lack most requisites that usually are associated with granting someone, or something, legal personhood: such artificial agents are not self-conscious, they do not possess human-like intentions, or properly suffer. This does not amount to say that the levels of autonomy, self-consciousness, and intentionality—which arguably are insufficient to grant AI robots their full legal personhood today—are inadequate to produce relevant effects in other fields of the law, e.g., the legal agenthood of artificial agents in the field of contracts and business law, as previously explored above in Section 3. Otherwise, we would incur in the same kind of confusion that has been stressed apropos of, say,

the "two abuses"-doctrine, by simply reversing the terms of such argumentation, that is, if AI robots do not meet the requisites of legal personhood, then they cannot be legal agents either.

The second kind of problem concerns the consequences of granting AI robots legal personhood. Once we admit there being artificial agents capable of autonomous decisions similar in all relevant aspects to the ones humans make, the next step would be to acknowledge that the legal meaning of "person" and, for that matter, of crimes of intent, of constitutional rights, of dignity, etc., will radically change. Even Solum admits that, "given this change in form of life, our concept of a person may change in a way that creates a cleavage between human and person" [25] (p. 1268). Likewise, others warn that "the empirical finding that novel types of entities develop some kind of self-consciousness and become capable of intentional actions seems reasonable, as long as we keep in mind that the emergence of such entities will probably require us to rethink notions of consciousness, self-consciousness and moral agency" [27] (pp. 558–559). At the end of the day, nobody knows to where this scenario may lead. For instance, would a strong AI robotic lawyer accept the argument that "evil is not part of the components of criminal liability" [26] (p. 93)? What if the AI robot, rather than an advocate of current exclusive legal positivism, is a follower of the natural law tradition?

The third kind of problem has to do with how we should mediate today's state-of-the-art and Leibniz's warning about our own ignorance: "every mind has a horizon in respect to its present intellectual capacity but not in respect to its future intellectual capacity" [28] (p. 115). On the one hand, some popular claims of today's debate can be deemed as simply non-sense, such as the awareness of AI robots, thus subject to retribution and deterrence, rehabilitation and incapacitation, down to capital penalty e.g., [26]. Yet, on the other hand, the breath-taking advancements of technology in this field recommend being prepared as to how we shall rethink notions of consciousness, self-consciousness and moral agency. As previously stressed above in Section 3, some legislators and policy makers have adopted forms of legal experimentation, e.g., the Japanese government's special zones set up over the past 15 years, as a way to address the normative challenges of AI robotics in a pragmatic way, that is, through empirical testing and development [29]. Admittedly, some sort of Western *Tokku* can increase our comprehension of risks and threats triggered by robots and smart AI systems, in order to prevent their undesirable actions and keep them in check. In addition, we can grasp how such systems may react in different contexts and whether robots and other AI agents ultimately meet human needs. From a legal viewpoint, the set up of special zones for robotics empirical testing and developing appears particularly relevant, because we can properly address on this basis the set of potential issues brought about by the advancement of AI and robotics, e.g., a fleet of self-driving cars for "public car sharing" [30]. In the traditional world of human drivers, many legal systems had to introduce—in addition to compulsory insurance policies—public funds for the victims of road accidents, e.g., the Italian legislative decree no. 209 from 2005. In the foreseeable world of autonomous vehicles, hypotheses of accountable AI car systems may make sense because a sort of digital peculium, embedded in the design of the system, can represent the smart AI counterpart to current public funds for the victims of road accidents. Along these lines, it is worth mentioning that the former Italian data protection authority has suggested that robots may soon become "data processors" pursuant to Article 28 of the EU regulation on data protection [31].

Whether or not this kind of legal experimentation and pragmatic approach will end up granting AI robots full legal personhood, thereby transforming pillars of the law and such basic concepts, as the idea of legal person, is of course a question that goes beyond our current Leibniz's "horizon". In the meanwhile, we already mentioned the case of our first robotic citizen, Sophia, who is the first non-human ever to be awarded by the UN too. *Pace* the empirical remarks and conceptual exercises that have been summed up throughout this paper—taking into account more than 25 years of scholarly debate in the field—what would the reasons for these legal steps be? By taking UDHR's Article 1 and ECHR's Protocol 13 seriously, where would the dignity of Sophia lie? Is she really conscious, or capable to suffer the slings and arrows of outrageous scholars?

The time is ripe for the conclusions of this paper.

5. Conclusions

Scholars properly stress, time and again, that the notion of person is a fiction in the legal domain. While Ancient Roman lawyers recovered the term 'person' from the theater's mask that actors used to wear on stage, none of their definitions of legal person resembles today's meaning of personhood. Contrary to Roman ideas on the role of the parties in legal acts, or in a process, on the status of free people, or enslaved persons, and so on, we currently associate the notion of person with the legal subject that has rights (and duties) of its own. This is the definition that we find in chapter 16 of Thomas Hobbes's *Leviathan*, which has an origin in the ideas discussed by scholars of Canon Law in the 1200s and 1300s. In his *Commentary on Digestum Novum* [32], for instance, Bartolus de Saxoferrato (1313–1357) affirms that an artificial person, such as a monastery, is not a real person but rather, a fiction that nonetheless "stands in the name of the truth" (*pro vero*) [32]. This equalization between natural persons and artificial persons, between humans and legal entities, such as a mission, or a corporation, triumphed with the tenets of legal positivism and formalism in the 1800s. The great Roman scholar, Friedrich August von Savigny, makes this point clear in his *System of Modern Roman Law* (1840–1849) ed. (1979). Here, Savigny admits that only humans have rights and duties of their own and still, the Law has the power to grant such right of personhood to anything, be it monasteries or corporations, governments or ships in maritime law, rivers in New Zealand or India, down to the entire ecosystem in Ecuador.

Legal fictions have real effects, though. Since Ancient Roman times, they concern for example the procedural mechanisms that allow individuals to enforce their own rights, e.g., the *actio in personam* which gives an individual the role of a party in a process or legal act. In addition, fictions may regard the family status of an individual, e.g., adoptions, or when such individual should be considered deceased, e.g., Article 4 of the Italian civil code on cases in which some legal effect depends on whether someone outlived someone else and yet, it is impossible to determine such circumstance, so that both persons are considered by the law as deceased at the same time.

This crucial real-life impact of legal fictions has thus recommended scholars to carefully mull over whether and to what extent AI robots may create loopholes in the legal system and hence, whether current provisions of the law should be modified, or amended. By distinguishing between the legal agenthood of AI robots in contracts and business law, and the legal personhood of AI robots with their constitutional rights, the paper has insisted on the reasons why we should not confuse such legal statuses, so that two different outcomes follow as a result.

On the one hand, as to the legal agenthood of AI robots, it makes sense to consider new forms of accountability and of liability in the field of contracts and business law, such as registries, or modern forms of *peculium*. The aim is to prevent both risks of robotic liability shield and of AI robots as unaccountable rights violators, while tackling cases of distributed responsibility that hinge on multiple accumulated actions of humans and computers that may lead to cases of impunity. On the other hand, as to the legal personhood of AI robots, current state-of-the-art has suggested that none of today's AI robots meet the requisites that usually are associated with granting someone, or something, such legal status. Although we should be prepared for these scenarios through manifold methods of legal experimentation, e.g., setting up special zones, or living labs, for AI robotic empirical testing and development, it seems fair to concede that we currently have other types of priority, e.g., the regulation of the use of AI robots on the battlefield [4].

Therefore, going back to the February 2017 proposal of the European Parliament, which was mentioned above in Section 2, the final recommendation of this paper would be threefold: (i) in the mid term, skip any hypothesis of granting AI robots full legal personhood; (ii) take seriously into account the possibility of new forms of accountability and liability for the activities of AI robots in contracts and business law, e.g., new forms of legal agenthood in cases of complex distributed responsibility; and, (iii) test such new forms of accountability and liability through methods of legal experimentation.

However, going back to the current debate on the legal personhood of AI robots, we should recognize that granting someone, or something, legal personhood is—and always has been—a highly

sensitive political issue. In addition to rivers in New Zealand and India, or the entire ecosystem in Ecuador, consider the legal jungle of the status, or condition, of individuals as legal members of a state, e.g., people's citizenship. As shown by the legal condition of Turks in Germany, or of some Brazilian football players in Italy, or of young immigrants in US, this is the realm of political discretionary power that sometimes turns into simple chaos, or mere sovereign arbitrariness. The recent case of Saudi Arabia enrolling Sophia as a citizen of her own is hence unsurprising. It reminds us of Suetonius' *Lives of the Twelve Caesars* (121 AD), in which we find Caligula planning to make his horse, Incitatus, a senator, and "the horse would invite dignitaries to dine with him in a house outfitted with servants there to entertain such events".

From Incitatus to Sophia, the paper has stressed the normative reasons, according to which we can evaluate whether granting legal personhood makes sense, or turns out to be a simple matter of sheer chance and political unpredictability. In the case of legal persons, such as corporations, political decisions have to do with matters of efficiency, financial transparency, accountability, and the like. In the case of human fellows, the reference is to their dignity, consciousness, intrinsic worth, and so forth. Certainly, we cannot prevent on this basis the odd decisions of legislators making robots citizens, or horses senators. Yet, from Caligula's horse to current Sophia, basic legal principles make clear when political decisions on "persons" are incongruous, so that courts may one day overturn them for having no rational basis.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Reynolds, C.; Ishikawa, M. Robotic Thugs. In *Ethicomp*; Bynum, T.W., Rogerson, S., Murata, K., Eds.; Global e-SCM Research Center & Meiji University: Tokyo, Japan, 2007; pp. 487–492.
2. Pagallo, U. The Human Master with a Modern Slave? Some Remarks on Robotics, Ethics, and the Law. In *The "Backwards, Forwards and Sideways" Changes of ICT*; Arias-Oliva, M., Bynum, T.W., Rogerson, S., Torres-Corona, T., Eds.; Universitat Rovira I Virgili: Tarragona, Spain, 2010; pp. 397–404.
3. Pagallo, U. The Adventures of Picciotto Roboto: AI and Ethics in Criminal Law. In *The Social Impact of Social Computing*; Bissett, A., Light, A., Lauener, A., Rogerson, S., Ward Bynum, T., Eds.; Sheffield Hallam University: Sheffield, UK, 2011; pp. 349–355.
4. Pagallo, U. *The Laws of Robots: Crimes, Contracts, and Torts*; Springer: Dordrecht, The Netherlands, 2013.
5. Bryson, J.J.; Diamantis, M.E.; Grant, T.D. Of, for, and by the People: The Legal Lacuna of Synthetic Persons. *Artif. Intell. Law* **2017**, *23*, 273–291. [CrossRef]
6. UN World Robotics. *Statistics, Market Analysis, Forecasts, Case Studies and Profitability of Robot Investment*; Edited by UN Economic Commission for Europe and Co-Authored by the International Federation of Robotics; UN Publication: Geneva, Switzerland, 2005.
7. Gogarty, B.; Hagger, M. The Laws of Man over Vehicle Unmanned: The Legal Response to Robotic Revolution on Sea, Land and Air. *J. Law Inf. Sci.* **2008**, *19*, 73–145.
8. Singer, P. *Wired for War: The Robotics Revolution and Conflict in the 21st Century*; Penguin: London, UK, 2009.
9. Pagallo, U. Robots of Just War: A Legal Perspective. *Philos. Technol.* **2011**, *24*, 307–323. [CrossRef]
10. Bekey, G.A. *Autonomous Robots: From Biological Inspiration to Implementation and Control*; The MIT Press: Cambridge, MA, USA; London, UK, 2005.
11. Simon, H. *The Shape of Automation for Men and Management*; Harper & Row: New York, NY, USA, 1965.
12. Minski, M. *Computation: Finite and Infinite Machines*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1967.
13. OSTP; National Science and Technology Council Networking and Information Technology; Research and Development Subcommittee; National Science and Technology Council Networking and Information Technology. *The National Artificial Intelligence Research and Development Strategic Plan*; OSTP: Washington, DC, USA, 2016.
14. Hildebrandt, M. From Galatea 2.2 to Watson and back? In *Human Law and Computer Law: Comparative Perspectives*; Hildebrandt, M., Gaakeer, J., Eds.; Springer: Dordrecht, The Netherlands, 2011.

15. Chopra, S.; White, L.F. *A Legal Theory for Autonomous Artificial Agents*; The University of Michigan Press: Ann Arbor, MI, USA, 2011.

16. Karnow, C.E.A. Liability for Distributed Artificial Intelligence. *Berkeley Technol. Law J.* **1996**, *11*, 147–183.

17. Lerouge, J.-F. The Use of Electronic Agents Questioned under Contractual Law: Suggested Solutions on a European and American level. *John Marshall J. Comput. Inf. Law* **2000**, *18*, 403.

18. Weitzenboeck, E.M. Electronic Agents and the Formation of Contracts. *Int. J. Law Inf. Technol.* **2001**, *9*, 204–234. [CrossRef]

19. Bayern, S.; Burri, T.; Grant, T.D.; Häusermann, D.M.; Möslein, F.; Williams, R. Company Law and Autonomous Systems: A Blueprint for Lawyers, Entrepreneurs, and Regulators. *Hastings Sci. Technol. Law J.* **2017**, *9*, 135–162. [CrossRef]

20. Pagallo, U. The Group, the Private, and the Individual: A New Level of Data Protection? In *Group Privacy: New Challenges of Data Technologies*; Taylor, L., Floridi, L., van der Sloot, B., Eds.; Springer: Dordrecht, The Netherlands, 2017; pp. 159–173.

21. Pagallo, U.; Quattrocolo, S. The Impact of AI on Criminal Law, and Its Twofold Procedures. In *The Research Handbook of the Law of Artificial Intelligence*; Barfield, W., Pagallo, U., Eds.; Elgar: Cheltenham, UK, 2018.

22. Pagallo, U. AI and Bad Robots: The Criminology of Automation. In *The Routledge Handbook of Technology, Crime and Justice*; McGuire, M.R., Holt, T.J., Eds.; Routledge: London, UK; New York, NY, USA, 2017; pp. 643–653.

23. Pagallo, U. LegalAIze: Tackling the Normative Challenges of Artificial Intelligence and Robotics through the Secondary Rules of Law. In *New Technology, Big Data and the Law. Perspectives in Law, Business and Innovation*; Corrales, M., Fenwick, M., Forgó, N., Eds.; Springer: Singapore, 2017; pp. 281–300.

24. Singer, P. *Practical Ethics*; Cambridge University Press: Cambridge, UK, 2011.

25. Solum, L.B. Legal personhood for artificial intelligence. *N. C. Law Rev.* **1992**, *70*, 1231–1287.

26. Hallevy, G. *Liability for Crimes Involving Artificial Intelligence Systems*; Springer: Dordrecht, The Netherlands, 2015.

27. Hildebrandt, M.; Koops, B.-J.; Jaquet-Chiffelle, D.-O. Bridging the Accountability Gap: Rights for New Entities in the Information Society? *Minn. J. Law Sci. Technol.* **2010**, *11*, 497–561.

28. Coudert, A.P. *Leibniz and the Kabbalah*; Kluwer Academic: Boston, MA, USA; London, UK, 1995.

29. Pagallo, U. From Automation to Autonomous Systems: A Legal Phenomenology with Problems of Accountability. In Proceedings of the International Joint Conferences on Artificial Intelligence Organization (IJCAI-17), Melbourne, Australia, 17–23 August 2017.

30. Pagallo, U. Three Lessons Learned for Intelligent Transport Systems that Abide by the Law. JusLetterIT. Available online: http://jusletter-it.weblaw.ch/issues/2016/24-November-2016/three-lessons-learne_9251e5d324.html (accessed on 24 November 2016).

31. Pizzetti, F. *Intelligenza Artificiale, Protezione dei Dati e Regolazione*; Giappichelli: Torino, Italy, 2018. (In Italian)

32. De Saxoferrat, B. Commentary on Digestum Novum. In *Commentaria*; Il Cigno: Rome, Italy, 1996.