

# Of, For, and By the People: The Legal Lacuna of Synthetic Persons\*

Joanna J. Bryson · Mihailis E. Diamantis · Thomas  
D. Grant

August 24, 2017

**Abstract** Conferring legal personhood on purely synthetic entities is a very real legal possibility, one under consideration presently by the European Union. We show here that such legislative action would be morally unnecessary and legally troublesome. While AI legal personhood may have some emotional or economic appeal, so do many superficially desirable hazards against which the law protects us. We review the utility and history of legal fictions of personhood, discussing salient precedents where such fictions resulted in abuse or incoherence. We conclude that difficulties in holding “electronic persons” accountable when they violate the rights of others outweigh the highly precarious moral interests AI legal personhood might protect. Keywords: Legal personality : international organisations : artificial intelligence : robots : legal agency : moral subject : ethics.

## 1 Introduction

Fiction abounds with artificial human-like characters: robots, clones, and bioengineered humanoids. But fiction dwells on artists’ conceptions of the human condition, and the contexts in which that condition might or might not be altered. Human-like artefacts are no longer fiction, and humanity is now confronted by the very real legal challenge of a supranational entity considering whether to attribute legal personality to purely synthetic intelligent artefacts. The European Parliament has asked the European Commission to write legislation addressing the forthcoming challenges of artificial intelligence (AI)—a sensible and timely

---

\* Abraham Lincoln, Gettysburg Address, Nov. 13, 1863 (“[W]e here highly resolve . . . that government of the people, by the people, for the people shall not perish from the earth.”)

Joanna J. Bryson

Department of Computer Science, University of Bath, and Princeton Center for Information Technology Policy E-mail: [jjb@alum.mit.edu](mailto:jjb@alum.mit.edu)

Mihailis E. Diamantis

Department of Law, University of Iowa E-mail: [mihailis-diamantis@uiowa.edu](mailto:mihailis-diamantis@uiowa.edu)

Thomas D. Grant

Wolfson College and Lauterpacht Centre for International Law, University of Cambridge E-mail: [tdg20@cam.ac.uk](mailto:tdg20@cam.ac.uk)

suggestion. Here we address only one aspect of that proposal (part 59f, p. 19), the recommendation that the legislature should consider:

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.

The language concerning “electronic persons” indicates a clear intent to confer on some intelligent artefacts legal-person status, such as is also enjoyed by most humans.

In this article, we ask whether a purely synthetic entity could and should be made a legal person. Drawing on the legal and philosophical framework used to evaluate the legal personhood of other non-human entities like corporations, we argue that the case for AI-electronic personhood is weak. Though this article begins with philosophical premises, its orientation is ultimately pragmatic. A legal system *by* the people exists ultimately to protect the interests *of* the people. That is to say, the people currently recognized as such. In the absence of some compelling moral necessity, we should consider the likely costs and benefits of any legal change *for* the people. Welcoming AI to the class of legal persons would be a change. Our purpose here is to identify some costs that that choice would present.

We work with a historical concept of legal personhood, as set out in the excellent review of the issue by Solaiman (2016). To summarise Solaiman very briefly, legal personhood extends to the set of entities in any lawfully-regulated society that have rights and obligations under the law. The basic provisions for a legal person are:

1. that it is able to know and execute its rights as a legal agent, and
2. that it is subject to legal sanctions ordinarily applied to humans.

Historically, only a relatively small subset of humans<sup>1</sup> would have counted as legal persons. Legal personhood has been extended not only to humans, but also to corporations and (in some countries) idols and environmental objects. Creating a legal status of electronic personhood for purely synthetic intelligent entities would require that such entities could fruitfully satisfy Solaiman’s second criterion. We argue here that it is far from clear that artefacts could or should be designed so as to acquire this status.

We begin our article by demonstrating the timeliness and immediacy of our concern that robots might be made legal persons. Proposals for creating synthetic personhood are already on the table, and there are sufficient legal tools in place to implement them. We advise caution and reflection on the problems that have arisen in the past with novel legal persons. While not always a zero-sum game, sometimes extending the class of legal persons can come at the expense of the interests of those already within it. In the past, creating new legal persons has sometimes lead to asymmetries and corruptions such as entities that are accountable but unfunded, or fully-financed but unaccountable. Ultimately this means weakening the legal protections for humans vis-a-vis synthetic persons. Next we consider whether there are moral benefits to offset these risks or costs, such as achieving necessary moral objectives. We suggest that there is no moral obligation to recognize the legal personhood of AI. We recommend against the extension of legal personhood to robots, because the costs are too great and the moral gains too few.

---

<sup>1</sup> In this article, *human* refers exclusively to members of the species *Homo sapiens*. Entities holding the class of legal rights and obligations we are considering here are *legal persons*. We use *robot*, *synthetic person*, and *AI* interchangeably.

## 2 Why concern ourselves with legal personality and AI now?

Academics have written for some years about the possibility of attributing legal personality to robots,<sup>2</sup> e.g. Asaro (2007), Koops et al. (2010), and Solaiman (2016). So the idea is not new. It gained considerable currency, however, after the Committee on Legal affairs of the European Parliament on 20 January 2015 established a Working Group for legal questions related to the development of Robotics and Artificial Intelligence. On 27 January 2017, the Committee put forward a Motion for a European Parliament Resolution in respect of robotics and artificial intelligence. On 16 February 2017, this Motion was adopted as the Civil Law Rules on Robotics.

Press reports give the impression that the Motion contains “comprehensive rules for how humans will interact with artificial intelligence and robots” (Wakefield, 2017; Sulmont, 2017). It does not. There are no legally binding decisions, and the proposals it makes are not in the form of rules, much less comprehensive ones. As a recommendatory text, the Motion identifies lines for future development, for example to create a registry of “smart robots,” to use the United Nations to set regulatory standards, to allocate public money to study the “social and ethical challenges” of advanced robotics, and so on. Of particular concern here, the Motion also suggested that European law might someday attribute legal personality to robots.

The Civil Law Rules are cautious and non-committal on the question of whether robots should be legal persons. Nevertheless, calling on the European Commission to consider the place of robotics in the European legal order gives the question unprecedented stature. Paragraph AB, in the introductory recitals of the Motion, says as follows:

“[T]he more autonomous robots are, the less they can be considered simple tools in the hands of other actors (such as the manufacturer, the owner, the user, etc.); . . . this, in turn, questions whether the ordinary rules on liability are insufficient or whether it calls for new principles and rules to provide clarity on the legal liability of various actors concerning responsibility for the acts and omissions of robots. . . .”<sup>3</sup>

The paragraph goes on to say:

“[U]ltimately, the autonomy of robots raises the question of their nature in the light of the existing legal categories or whether a new category should be created, with its own specific features and implications.”<sup>4</sup>

<sup>2</sup> Most authors focus on robots rather than general purpose artificial intelligence (AI), presumably because robots are easier to identify with or feature more sympathetically in fiction. Here we take *intelligence* to be a process for doing the right thing at the right time, where what is ‘right’ depends on context; and a *robot* to be an artefact that perceives and acts in the analogue, physical world, in contrast to software that operates only in the context of a digital artefact.

<sup>3</sup> The original language in the Final Motion, Paragraph S, read as follows: “[T]he more autonomous robots are, the less they can be considered simple tools in the hands of other actors (such as the manufacturer, the owner, the user, etc.); . . . this, in turn, makes the ordinary rules on liability insufficient and calls for new rules which focus on how a machine can be held—partly or entirely—responsible for its acts or omissions; . . . as a consequence, it becomes more and more urgent to address *the fundamental question of whether robots should possess a legal status*” (emphasis added).”

<sup>4</sup> The original language in the Final Motion, Paragraph T, read as follows: “[U]ltimately, robots’ autonomy raises the question of their nature in the light of the existing legal categories—of whether they should be regarded as natural persons, legal persons, animals or objects—or whether a new category should be created, with its own specific features and implications as regards the attribution of rights and duties, including liability for damage.”

As adopted, these recitals do not prescribe a particular future status for “autonomous robots.” They are without prejudice to whether European law should attribute legal personality to them. Nevertheless, to identify as “fundamental” the question “whether robots should possess a legal status” strongly implies that the door is open to that innovation.

As part of a list of “general principles concerning the development of robotics and artificial intelligence for civil use,” the Motion draws further attention to the possibility of attributing legal personality to robots. The Motion in particular calls on the European Commission “when carrying out an impact assessment of its future legislative instrument [on robots], to explore, analyse and consider the implications of all possible legal solutions.” In Paragraph 59, it includes giving robots “the status of electronic persons” among “possible legal solutions.”

Again, the Motion is not a statement of law-in-force. Nor does the Motion espouse a particular solution. It does, however, call on the European Commission to “explore” the attribution of legal personality to robots as a possible solution. Invoking the expressions “electronic persons” and “electronic personality,” it gives the idea a higher profile than ever before. The idea of legal personality and AI accordingly merits particular scrutiny at this time.

### 3 Legal persons: fictive, divisible, and not necessarily accountable

Before we can talk sensibly about legal personality for robots, we need to know what the expression “legal personality” means in general. *Legal personality* is a term of art in legal scholarship and practice. Jurists in multiple countries have set out definitions of it. This one, from the *Yale Law Journal* in 1928, is serviceable: “To be a legal person is to be the subject of rights and duties. To confer legal rights or to impose legal duties, therefore, is to confer legal personality. . . .” (Smith, 1928, 283). This definition is congruent with Solaiman’s characterization of legal personhood, discussed above.

Three observations about legal personality, so defined, are pertinent to the question of a possible electronic legal personality. First, legal personality is an artifice. When we say that an actor has legal personality, we mean that a legal system addresses its rules to the actor, both to give the actor rights and to subject it to obligations. Legal personality is not necessarily correlated with a metaphysical or ethical notion of personhood. While we should want our legal system to bear the metaphysical and ethical concepts in mind, at different times legal systems have conferred legal personhood on much less and much more than the set of metaphysical or ethical persons. Legal personality results from a legal system’s decision to recognize that a particular entity has it. We may thus think of legal personality as a kind of fictional status, which the law may confer when doing so suits its ends.

Second, legal personality is an aggregate of legal rights and obligations, and thus it is divisible. Legal people need not possess all the same rights and obligations, even within the same system. A legal system might treat a given actor as a legal person in respect of some rights and some obligations but not in respect of others. It may even be helpful to think of legal personhood as a scalar concept, so that an entity can be more or less of a legal person as it possesses more or fewer rights and obligations.

Third, the legal personality of an actor, even if it entails that the actor has extensive rights and obligations, does not necessarily entail the actor’s effective engagement with the legal system. Though the actor may be the beneficiary of certain rules that give it rights, or the addressee of others that impose obligations on it, this does not in itself tell us what opportunities the legal system provides to that actor to take advantage of the rules or to other

actors to hold it to account for breaches. That is to say, the rights and obligations that a legal person may have as a matter of law may not match those it has as a matter of fact.

We now consider in detail how each of these observations about legal personality bears on the possibility of extending legal personhood to robots.

### 3.1 Legal personality is a fiction of a given legal system

An entity's inherent characteristics do not determine whether it is a legal person. It is true that legal systems are less likely to confer legal personality on inanimate objects, and more likely to confer it on entities that are people in the ethical and metaphysical sense. This may be because most legal systems wish to recognize and give effect to the rights and obligations that true people possess. But this rough generalization can be misleading. To determine whether an entity is a legal person, one must look to the approach a given legal system takes toward it.

Because of the rough generalization that legal people are in fact people, that the legal rights and obligations correspond to real rights and obligations, it is natural to think of legal personality as a fiction pretending to be something real. When a legal system confers legal rights and obligations on an entity, it has determined to treat that entity as though it were a person in fact. It is a kind of pretense in which legal systems can decide to engage, regardless of whether an entity really is a person (See examples in Solaiman, 2016, 3–4). Calling legal personality “a fiction” does not mean that it lacks real effects. To the contrary, the purpose of conferring legal personality on an actor is to enable that actor to have certain effects in, and to be effected in certain ways by, the legal system.

Every legal system must decide to which entities it will confer legal personhood. Legal systems should make this decision, like any other, with its ultimate objectives in mind. The most basic question for a legal system with respect to legal personhood is whether conferring legal personhood on a given entity advances or hinders those objectives. Those objectives may (and, in many cases *should*) be to give legal recognition to the rights and obligations of entities that really are people. In many cases, though, the objectives will not track these metaphysical and ethical truths. Sometimes legal personhood may be denied to real people in order to serve odious ends, like perpetuating privileges for some smaller group of people. Other times, a legal system may grant legal personhood to entities that are not really people because conferring rights upon the entity will protect it or because subjecting the entity to obligations will protect those around it.

In this regard, the discourse and practice of recognizing legal personhood fits the kind of structure that philosophers call *fictionalism*. A domain of discourse is fictionalist if it seeks to represent something other than the literal truth (Eklund, 2011). Participants in a fictionalist discourse engage in a sort of pretense (whether wittingly or not) by assuming a stance according to which things said in the discourse, though literally false, refer to real entities and describe real properties of them. Discourse about fictional narratives is one easy example. When someone asks whether Daenerys Targaryen has two or three dragons, they are not asking after some fact in the world. Rather, they mean to ask whether the statement is true within the fiction *Game of Thrones*. Many modern philosophers think fictionalism offers the best account of some familiar domains of discourse, from math (e.g. Field, 1989), to morality (e.g. Joyce, 2001), to truth (e.g. Burgess and Burgess, 2011). When they argue that these domains of discourse are fictionalist, philosophers take on the burden also of saying why we would go to the effort of earnestly saying things that are literally false. Usually, this involves giving an account of why the discourse is useful—e.g., talk of fictional narrative is

fun, talk of numbers allows us to build airplanes, and talk of morality allows us to organize socially.

In the legal context, there is a long history of conferring legal personhood on corporations, and recognizing that the discourse surrounding corporate legal personhood is fictional. The United States has perhaps the most thoroughly developed legal discourse on the matter. Under U.S. federal law, the term *person* is defined to include corporations.<sup>5</sup> Participants in the legal system recognize that the discourse surrounding corporate personhood is fictional. As the U.S. Supreme Court wrote, “[T]he corporate personality is a fiction, although a fiction intended to be acted upon as though it were a fact. . . .”<sup>6</sup> Scholars for the most part take such statements at face value (Dewey, 1926, 655-73; Laufer, 1994, 647, 650). Creating a fictional discourse according to which corporations are people was a useful shorthand for conferring on them the legal rights and obligations possessed by human people within the legal system. These include, for example, the corporate right to bind others through contract and the corporate obligation to satisfy its commitments under contract. Without an extensive suite of rights and obligations characteristic of legal personhood, corporations could not be the engines of economic progress they have become.

Sometimes legal systems will even confer legal personality on an *ad hoc* basis to individual entities. This happened, for example, with the Bank for International Settlements. In a case involving claims against the Bank, an arbitral tribunal noted that the international instruments that created and empowered the Bank—part of a Convention concluded in 1930 by Germany, Belgium, Great Britain, Italy, Japan and Switzerland—confirmed that the Bank was to be an international law entity. The arrangement was novel, a company limited by shares and, apparently, generally recognized as a person under international law. Some of the participants doubted that this was legally tenable, and so they set up a rather tangled structure to give the Bank a Swiss law status—even as Swiss law was expressly not the Bank’s governing law for its most important purposes.<sup>7</sup> The Bank was intended to be an international legal person, and the states participating in the Bank communicated their intention by adopting a treaty.<sup>8</sup> The Bank’s personality was confirmed (the tribunal went on to observe) by explicit statements in other international agreements.<sup>9</sup>

We are concerned here about possible future cases concerning the legal personality of robots. Some academic writings about robot legal personality address questions of personhood in other than a legal sense, e.g., what does it take to constitute a person in a social, biological or even theological sense (Foerst, 2009). Legal personality, however, results from a decision in the legal system to confer legal personality on a given entity. This decision may, but need not, be informed the status of robots as persons *vis-a-vis* these non legal senses. Legal personality is a highly elastic concept. The range of actors on which a system might confer legal personality is large, a point understood since at least the 1930s (see Nékám, 1938, p. 34). The European Parliament Motion of 27 January 2017 to consider the possibility of conferring legal status on robots, accordingly, is not trivial. Nothing in the character of

<sup>5</sup> Dictionary Act, 1 U.S.C. 1.

<sup>6</sup> See *Int’l Shoe Co. v. Wash., Office of Unemployment Compensation and Placement*, 326 U.S. 310, 316 (1945).

<sup>7</sup> *Reineccius, First Eagle SoGen Funds Inc., et al v. Bank for International Settlements*, Tribunal established Pursuant to Article XV of the Agreement Signed at the Hague on 20 January 1930 (Reisman, Frowein, Krafft, Lagarde, and van den Berg), Partial Award, 22 November 2002, (2004) 23 Reports of International Arbitral Awards 184, 212-213 (paras. 105–106).

<sup>8</sup> *Ibid.*, p. 214 (para. 112).

<sup>9</sup> *Ibid.*, p. 215 (paras. 115-16) (“[R]ecognition [of the Bank as a legal person] clearly flows from the provisions of the Agreements.”).

legal systems as such forecloses the possibility, and there is significant precedent to enable it.

### 3.2 Legal personality is divisible

Legal personhood is not an all-or-nothing proposition. Since it is made up of legal rights and obligations, entities can have more, fewer, overlapping, or even disjointed sets of these. This is as true of the legal personhood of human beings as it is for non-human legal persons. Every legal system has had, and continues to have, some human legal persons with fewer legal rights and different obligations than others. The world-wide struggle for equal rights for women, ethnic and religious minorities, and other disadvantaged groups in many nations bears continuing witness to this fact. The disparity is not always invidious; sensible policy can ground different rights and obligations (in some ways more, in others less) for non-citizens, felons, and children. (Asaro, 2007, p. 3).

As discussed above, legal systems can confer legal personhood on non-human entities. In almost every case, these will have both fewer rights and fewer obligations. Consider the legal personhood that environmental features now have in several countries—the Whanganui river and Te Urewera national park in New Zealand (Rousseau, 2017), the Ganges and the Yamuna rivers in India, (Safi, 2017) the entire ecosystem in Ecuador.<sup>10</sup> Of necessity, the legal rights and obligations accorded to these environmental features differ from those given by their respective nations to human beings. In the case of the Whanganui River, for example, the primary concern was to ensure the rights of the river not to be owned (Calderwood, 2017). Corporations in the United States may be the legal persons with the suite of legal rights and obligations most closely approximating those given to human beings. A detailed constitutional jurisprudence has grown around the issue. While the U.S. Supreme Court seems on track to affirm that corporations have nearly every constitutional right and obligation, it has balked in some rare instances, such as the right against self-incrimination at criminal trial.<sup>11</sup>

In some cases, courts have had to address the divisibility of legal personhood head on. The General Assembly, in 1948, asked the International Court of Justice whether the UN had the capacity to bring an international claim against a State. The Court advised in the affirmative. In so advising, the Court drew attention to the varied character of persons in a legal system:

“The subjects of law in any legal system are not necessarily identical in their nature or in the extent of their rights, and their nature depends upon the needs of the community. . . [T]he [UN] is an international person. That is not the same thing as saying that it is a State, which it certainly is not, or that its legal personality and rights and duties are the same as those of a State. . . Whereas a State possesses the totality of international rights and duties recognized by international law, the rights and duties of an entity such as the [UN] must depend upon its purposes and functions as specified or implied in its constituent documents and developed in practice.” (Liang, 1949)

The Court understood that legal personality is a divisible concept. It is not necessary in any legal system for there to be one uniform and unified status of legal person.

<sup>10</sup> Ecuador Const., title 10 (“Nature shall be the subject of those rights that the Constitution recognizes for it.”), available at <http://pdba.georgetown.edu/Constitutions/Ecuador/english08.html>.

<sup>11</sup> *Hale v. Henkel*, 291 U.S. 43 (1906).

The divisibility of legal personhood raises the question of which rights and duties a legal system should confer on a legal person, once it has decided to recognize the legal person as such. We should resolve the issue of the legal personhood of robots at this level, rather than treating legal personhood as an all-or-nothing black box (Koops et al., 2010, p. 556). Edsger Dijkstra has noted, “A convincing demonstration of correctness being impossible as long as the mechanism is regarded as a black box, our only hope lies in not regarding the mechanism as a black box” (Dijkstra, 1970). A legal system, if it chose to confer legal personality on robots, would need to say specifically which legal rights and obligations went with the designation. If it does not, then the legal system will struggle, as happened with the Bank for International Settlements, to make sense of what it has done. To try to confer “legal personality,” without being more specific, is to regard legal personality as a black box. In line with the fictionalist paradigm, and as the ICJ opined with respect to the UN, the legal system should determine the legal rights and obligations of a new legal person by reference to how the legal person relates to the legal system’s purposes.

### 3.3 The gap between *de jure* and *de facto* legal personality

Even once a legal system has determined which rights and obligations to confer on a legal person, practical realities may nullify them. Legal rights with no way to enforce them are mere illusion. Standing—the right to appear before particular organs for purposes of presenting a case under a particular rule—is crucial to a legal person seeking to protect its rights in the legal system. Standing does not necessarily follow from the existence of an actor’s legal personality. An entity, even when its legal personality is not in doubt, must exercise its standing before it can avail itself of relevant procedures (van Vollenhoven et al., 1926). When an entity tries to invoke newly conferred rights, challenges to its standing are all the more likely (Shah, 2013).

Consider the legal right of “integral respect” that Ecuador gave to its ecosystem. While the ecosystem may have the right as a matter of law, it clearly lacks the non-legal capacities it would need to protect the right from encroachment. To effectuate the right, the Ecuadorian constitution gave standing to everyone in Ecuador to bring suits on behalf of the ecosystem. Thus, in 2011, private Ecuadorians successfully sued the Provincial Government of Loja to halt expansion of a roadway that was damaging an important watershed (Greene, 2017). The outcome would have been very different if Ecuador had provided no mechanism for protecting nature’s legal right of integral respect. Nature cannot protect itself in a court of law.

Just as legal rights mean nothing if the legal system elides the standing to protect them, legal obligations mean nothing in the absence of procedure to enforce them. The advisory opinion of the ICJ establishing that the UN has legal personality was in 1948, but this resolved only whether the UN could bring a claim. It said nothing about an obvious correlate: the legal capacity of the UN to bear responsibility and answer for its own breaches. Affirmation that the UN indeed can be responsible for its breaches did come—but over half a century later (Wickremasinghe and Evans, 2000, para. 66). Despite the efforts of international lawyers, there is still no reliable procedure for suing an international organization.<sup>12</sup>

We could never anticipate *ex ante* all the ways robot legal people would interact with other legal persons and with the institutions of the legal system (courts, administrative agen-

<sup>12</sup> Ibid. See also *Behrami v. France*, No. 71412/01 and *Saramati v. France, Germany and Norway*, No. 78166/01, ECtHR, 2 May 2007, Decision (admissibility), 45 EHRR SE 10, para. 149 and comment by Knoll (2008, p. 444).



cies, legislatures, police, etc.). In its first encounters with the legal system, every rule invoked on a robot's behalf or against it would require novel and controversial developments in law. Courts and other organs would struggle to decide how, if at all, the rules—heretofore addressed to other legal persons—address the robot. Both the robot's standing against other actors and other actors' standing against the robot would be sharply contested. If the topic of electronic personality is to be addressed, as directed in the European Parliament's 27 January 2017 Motion, standing—both of robots to sue and of others to sue them—is a further matter that would need to be considered.

### 3.4 Summary

The intricacies described in this section are not just inevitable 'bugs' to be eventually worked out. They are crucial questions that we must answer before introducing novel legal personhood. Concerns about legal accountability, and the way electronic persons might affect accountability, are our main motivation in writing this paper. We now turn to consider the impacts of offering some form of personhood status to robots.

## 4 Human purposes and synthetic personhood

According to the fictionalist paradigm, the advisability of conferring legal personhood on robots is ultimately a pragmatic question—Does endowing robots with this legal right or that legal obligation further the purposes of the legal system? It is an exercise with which judges and legal scholars are familiar from their extensive experience with corporate legal personhood. As Garrett (2014) and Blair and Pollman (2015) have separately argued, U.S. courts approach individual legal rights and obligations claimed by and against corporations using such a consequentialist framework. We should do the same for each of the divisible legal rights and obligations at issue for robot legal personhood.

A full treatment of the advisability of conferring legal personhood on robots would step methodically from one legal right or obligation to the next. Our primary concern in this paper is to raise a cautionary flag in the face of what seems to be international enthusiasm for extending legal personhood to robots. Elon Musk has recently renewed his apocalyptic predictions about the "existential risk" AI poses to human beings (Domonoske, 2017). Our concern is somewhat different, and arises internally to legal systems and how robotic legal persons would interact with human legal persons. Robotic legal personhood raises concerns about a sort of abuse within the legal system: While robot legal persons would enjoy a host of rights against human legal persons, it is unclear how corresponding legal obligations could be enforced against them.

A crucial step in the analysis will be to specify what are the purposes of the legal system in relation to which robot legal personhood should be assessed. Legal systems can be presumed to serve many purposes, and any claim as to what those are is sure to be deeply controversial. Cast at a general enough level, though, much of the controversy about the purposes of legal systems should dissipate. To that end, we claim that the basic purposes of human legal systems are:

1. to further the material interests of the legal persons it recognizes, and
2. to enforce as legal rights and obligations any sufficiently weighty moral rights and obligations, with the caveat that

3. should equally weighty moral rights of two types of entity conflict, legal systems should give preference to the moral rights held by human beings.

We think this statement of purpose reflects the basic material and moral goals of any human legal system, with what we hope will be an uncontroversially light thumb on the scale in favour of human interests. Yes, this is speciesism. But a kind that allows for deference to the weighty interests of other entities, via the mechanism of human investment in those entities (cf. the Solaiman, 2016, discussion of idols). If there is even the faintest shadow of truth to Musk's prediction, a much stronger version of speciesism would be justified vis-a-vis AI. However, the weaker statement here suffices to make our arguments below.

#### 4.1 Robot legal personhood as a moral imperative

If robots have, or were on course to acquire, moral rights, then granting them legal personhood by conferring some legal rights would further the purposes of legal systems. But there is great room for skepticism about first the possibility of ever designing robots that would hold moral rights, and second whether that possibility—were it to exist—should be realised.

The very grounds of moral rights is highly uncertain for any kind of entity. Some academics suggest that consciousness could be the litmus test for possessing moral rights. Consciousness itself is a tricky notion, and scholars frequently conflate numerous disjoint concepts that happen to be currently associated with the term *conscious* (Dennett, 2001, 2009). In the worst case, this definition is circuitous and therefore vacuous, with the definition of the term itself entailing ethical obligation (Bryson, 2012). If we could settle on a universal metric for moral patiency, that metric could inform whether and when we should give robots legal personhood. At present, any plausible metric should tell against synthetic legal personhood—there is no widespread acceptance that current robots can consistently satisfy these metrics.

Nevertheless, many consider that AI will progress to the point that it can pass any behaviourally observable metric for human-like consciousness. Note that the commonly-suggested Turing Test—requiring that a person interacting with an entity over a communications device mistakes them for human—is already routinely passed at least for limited periods by AI. If AI became capable of mimicking human intelligence, the tides may shift as academics and laypeople alike will come to identify empathetically with robots. For some—the *transhumanists*, who see technology as a mechanism to become themselves superhuman, even immortal (Goertzel, 2010; Geraci, 2010)—the identification will be even more immediate. Some even self-identify as robots already.

But there is no guarantee or necessity that AI will be developed in this way. It is far from clear that such an AI system would be desirable, and some scholars have suggested that designing such AI would be immoral (Bryson, 2009). There is no inevitable point at which AI systems must replicate their makers in becoming functionally similar to human adults. We can therefore ask the question whether such an effort should be attempted. Two options are that it could, like human cloning, be banned altogether; or that human-like AI development should be limited to small-scale, individual, artisanal work, and in particular not be tenable as legal products or business entities that would require fundamental changes to the law (Bryson, 2017).

Even if robots were to be constructed on the mass scale and to acquire moral rights, this would not fully settle the question of whether the law should recognize them as legal persons. Legal systems are flexible as to what actors they confer legal personality upon, and

they need no evidence of supposed inherent qualities of an actor in order to do so.<sup>13</sup> Similarly, the inherent qualities of non-human entities do not dictate the final word on whether they should be recognized as legal persons. Beforehand, we must also check for potential conflicts between possible legal rights of the non-human entity and those already held to be legal persons.

#### 4.2 Abuse of legal person status by robots and those that make them

As Solaiman (2016) emphasizes, it is important that legal persons have legal obligations as well as legal rights. If robots were recognized as legal persons capable of entering into complex legal relationships with with legal persons, there would inevitably arise situations where the acts of robots would interfere with the rights of humans and other legal persons. Without an obligation to respect the rights of other legal persons, those rights would, at least vis-a-vis robotic actors, be rendered a nullity. The solution may seem clear—Impose legal obligations on robots. But legal obligations are meaningless if there is no way to hold robots accountable for them. It is not clear that there is.

In seeming recognition of this, the United States Department of Defence has proactively declared in their Law of War Manual<sup>14</sup> that robotic weapons are never responsible legal agents.

*Law of War Obligations of Distinction and Proportionality Apply to Persons Rather Than the Weapons Themselves.* The law of war rules on conducting attacks (such as the rules relating to discrimination and proportionality) impose obligations on persons. These rules do not impose obligations on the weapons themselves. . . The law of war does not require weapons to make legal determinations, even if the weapon (e.g., through computers, software, and sensors) may be characterized as capable of making factual determinations, such as whether to fire the weapon or to select and engage a target. . .

Rather, it is persons who must comply with the law of war. . . [I]n the situation in which a person is using a weapon that selects and engages targets autonomously, that person must refrain from using that weapon where it is expected to result in incidental harm that is excessive in relation to the concrete and direct military advantage expected to be gained. . . [T]he obligation. . . may be more significant when the person uses weapon systems with more sophisticated autonomous functions. . .

The concern here is not a necessarily conceptual one. Through very careful planning, we may discover mechanisms by which robots could be held accountable for legal obligations imposed on them. But the planning would have to be careful indeed. Without it, there are

<sup>13</sup> This is another point visible in the *Bank for International Settlements* case. One party observed that the functions of the Bank were, in part at least, commercial in character and, from that argued that the Bank was not an international legal person. The tribunal rejected the party's argument: the acceptance of the Bank as an international legal person by the (decentralized) mechanisms of decision in that legal system established that it was a legal person: *Reineccius, First Eagle SoGen Funds Inc., et al v. Bank for International Settlements*, Tribunal established Pursuant to Article XV of the Agreement Signed at the Hague on 20 January 1930 (Reisman, Frowein, Krafft, Lagarde, and van den Berg), Partial Award, 22 November 2002, (2004) 23 Reports of International Arbitral Awards 184, 216 (paras. 116-117). See also Smith's skepticism toward "legal philosophers and students of jurisprudence" who "have not been content with so simple an explanation" and who "have sought for the 'internal nature' of legal personality. . .": (Smith, 1928, 284).

<sup>14</sup> U.S. Department of Defense, LAW OF WAR MANUAL, ¶6.5.9.3 (2015), <http://www.defense.gov/Portals/1/Documents/pubs/Law-of-War-Manual-June-2015.pdf>

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.

#### 4.2.1 Robots as liability shields

It is to be assumed that if decision-makers in the system say that they are ready to consider the possibility of “electronic personality,” then human actors will seek to exploit that possibility for selfish ends. There is nothing objectionable in itself about actors pursuing selfish ends through law. A well-balanced legal system, however, considers the impact of changes to the rules on the system as a whole, particularly so far as the legal rights of legal persons are concerned. We take the main case of the abuse of legal personality to be this: natural persons using an artificial person to shield themselves from the consequences of their conduct. Recognition of robot legal personhood could present unscrupulous actors with such “liability management” opportunities.

The law has a way to address this kind of difficulty: It can look behind the artificial person and reach a real one. Veil-piercing—i.e., going behind the legal form and helping or (more usually) sanctioning the real people behind the form—is well-known in various legal systems (Huang, 2012). A U.S.-Great Britain arbitral tribunal in the 1920s put the matter like this:

“When a situation legally so anomalous is presented, recourse must be had to generally recognized principles of justice and fair dealing in order to determine the rights of the individual involved. The same considerations of equity that have repeatedly been invoked by the courts where strict regard to the legal personality of a corporation would lead to inequitable results or to results contrary to legal policy, may be invoked here. In such cases courts have not hesitated to look behind the legal person and consider the human individuals who were the real beneficiaries.”<sup>15</sup>

The situation had been “anomalous” because the Cayuga tribe had legal personality as a corporate entity in New York State but not under international law. That is, the law that the tribunal had power to apply did not recognize the tribe as an entity to which that law could be applied. “[R]ecognized principles of justice and fair dealing” came to the rescue: The tribunal addressed the individuals comprising the tribe to get around its inability to address the tribe.

Solutions like this are not available in every case. Lawmakers contemplating legal personhood must consider the matter and provide for it. The arbitrators in the *Cayuga* case had an express invitation to apply equitable principles, the jurisdictional instrument (a treaty) having stipulated equity to be part of the applicable law.<sup>16</sup> Where equity or a similar principle is not part of the applicable law, a judge or arbitrator well might not be able to “look behind the legal person.” In a situation like that, the “human individuals” who were meant to answer for injury done remain out of the picture.

<sup>15</sup> *Great Britain (for the Cayuga Indians in Canada) v. USA*, Tribunal under Special Agreement of 18 August 1910 (Nerinx, President; Pound & Fitzpatrick, Arbitrators), Award, 22 January 1926, (1955) 6 *Reports of International Arbitral Awards* 173, 179. To similar effect a little later, see *Shufeldt Claim (USA/Guatemala)*, (Sisnett, Arbitrator), Decision, 24 July 1930, (1949) 2 *Reports of International Arbitral Awards* 1083, 1098. (“International law will not be bound by municipal law or by anything but natural justice, and will look behind the legal person to the real interests involved.”)

<sup>16</sup> *Ibid.*

The *Tin Council* case provides an illustrative warning. The case involved the International Tin Council, a public international organization constituted by a group of states (broadly an entity like the International Bank for Settlements). The states, using the Council, aimed to corner the world market for tin. When the prospects for success looked solid, the Council contracted debts. But the price of tin collapsed, and the Council went insolvent. When the creditors sought to sue and collect what they could on the debts, they found an empty shell and no procedural recourse. The Tin Council could not be sued in English court, and it would have been useless to sue anyway. The Council's creditors sought compensation from the member states, but this was to no avail either: The creditors' contractual relationship was with the Council, not with those who had called it into being. Apart from the possibility of a diplomatic solution—i.e., the states agreeing *ex gratia* to replenish the Council or pay the creditors—the creditors had no recourse.<sup>17</sup>

A difficulty in the *Tin Council* case was that the legal relations involved were novel, and so the court's precedents offered no guide for effectuating the creditor's rights:

“None of the authorities cited by the appellants [the creditors] were of any assistance in construing the effect of the grant by Parliament of the legal capacities of a body corporate to an international organization pursuant to a treaty obligation to confer legal personality on that organization.”<sup>18</sup>

Nor did the creditors adduce “any alleged general principle” in the English law sources that would have allowed the court to pierce the veil and attach liability to the states that had constituted the Council.<sup>19</sup> As for international law, “[n]o plausible evidence was produced of the existence of such a rule of international law” (i.e., a rule holding the constituents of the Council responsible for the Council's debts).<sup>20</sup> In short, unlike the tribunal in the *Cayuga* claims, the House of Lords found no way to avert “inequitable results.” The unusual and novel character of the entity led the court to a dead end.

Even when the law does explicitly provide for veil-piercing, judges and arbitrators have tended to apply it cautiously and as an exception. Easterbrook and Fischel (though defending the economic rationale for veil-piercing) memorably described veil-piercing as happening “freakishly”; they likened it to “lightning... rare, severe, and unprincipled” (Easterbrook and Fischel, 1985).

The *Tin Council* case foreshadows the risk that electronic personality would shield some human actors from accountability for violating the rights of other legal persons, particularly human or corporate. Without some way around that shield, we would surely see robots designed to carry out activities that carry high legal risk for human or corporate legal persons. Though this might benefit the humans behind the robots, it would come at the expense of human legal interests more generally.

#### 4.2.2 Robots as themselves unaccountable rights violators

Even if the legal system sensibly provided mechanisms for veil piercing in the case of robot legal persons, that solution could only go so far. By design, collective legal persons like

<sup>17</sup> *International Tin Council Case, JH Rayner (Mincing Lane) Ltd. v. Department of Trade and Industry*, 26 October 1989, House of Lords, Lord Griffiths, para. 178; reprinted (1990) 81 *International Law Reports* 670.

<sup>18</sup> *Ibid.*, Templeman, LJ, para. 166.

<sup>19</sup> *Ibid.*, para. 167.

<sup>20</sup> *Ibid.*, para. 168.

corporations and international organisations have legal persons behind them, who might stand to answer for violations of the rights of human legal persons. Advanced robots would not necessarily have further legal persons to instruct or control them. That is to say, there may be no human actor directing the robot after inception. The principal-agent model that veil piercing rests upon would then be hard to apply.

Autonomous or semi-autonomous robots interacting with humans will inevitably infringe the legal rights of humans. Giving robots legal rights without counter-balancing legal obligations would only make matters worse. In the conflict between robot and human legal rights, only the former would be answerable to the latter; humans would have no legal recourse. This would not necessarily be a problem, if

1. the other problems of legal personality—like standing and availability of dispute settlement procedures—were solved; and
2. the electronic legal person were solvent or otherwise answerable for rights violations.

But it is unclear how to operationalize either of these two steps.

In the case of corporate legal persons, humans composing the corporation can manage dispute settlement on behalf of the corporation in which they have an interest. But what we are imagining here is a robot legal person, untethered from an interested human principal. Who will represent the robot in the dispute? With the right AI, the robot might be able to represent itself. But we are sure to encounter this problem well before AI capable of effective court advocacy develop. Conceivably, the robot could hire its own legal counsel, but this brings us to the second step: robot solvency.

It is unclear what it would mean for a robot to hold assets, or how it would acquire them. It is possible that the law could contemplate mechanisms for robots to own property or hold accounts, as it does for corporate legal people. The law could also require the creators of robots to place initial funds in these accounts. But 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. When insolvent human legal persons violate others' legal rights, other tools are available to hold them to account—anything from apology to jail time. In the case of robots, these options are unavailable, unsatisfying, and/or ineffective.

Good faith efforts, like designing robots in order to avoid infringement of human legal rights, would not solve all the problems either. A machine made to endeavour to avoid breaches of legal obligation still would present risks. Any actor in society will encounter frictions and mischances resulting in legal incident. This is an unavoidable feature of the complex legal and social space that proponents of robot legal personhood would have robots enter.

## 5 Conclusion

We have shown that it is completely possible to declare a machine a legal person. The impulse to do so exists both at the individual level with academic proponents, and at the level of international governance with the European Parliament recommending consideration. We have also argued here that conferring legal personality on robots is morally unnecessary and legally troublesome. While it may, either now or in the future, have emotional and economic appeal, so do many superficially desirable hazards against which the law protects us. The basic concern is for protecting human and corporate legal rights against abuse by—or more accurately, by exploiting—robots. Trying to hold an electronic person to account, claimants

would experience all the problems that have arisen in the past with novel legal persons. There almost inevitably would arise asymmetries in particular legal systems, situations like that of the investor under investment treaties who can hold a respondent party to account but under the same treaties is not itself accountable. Future claimants, if they were to sue an electronic person, likely would confront the accountable but empty, like the International Tin Council; the fully-financed but unaccountable, like the United Nations; and *sui generis* arrangements like the Bank for International Settlements that novel legal persons tend to instigate.

Perhaps a robot could be likened to a force of nature—a storm or avalanche. But this would not be satisfactory either: Natural forces are not legal persons. They affect our legal relations, but we do not speak of them as *having* legal relations. The electronic person by contrast, would engage in some or all of the legal relations available under the legal system, and yet, for those with whom it transacts or third parties whom it encounters, it would be difficult to hold to account. We have insurance schemes to address floods and fires. You can sue its owner if a dog bites you. The constituent states of the Tin Council, if the court had been willing to pierce the veil, would have stood exposed to the debts it had accrued. An electronic person by contrast might prove to be a legal black hole, an entity that absorbs a human actor's legal responsibilities and from which no glint of accountability is seen. Unfortunately, there is no question that such a readily-manufacturable legal lacuna would be exploited as a mechanism for avoiding and displacing legal liabilities and obligations.

It could be in theory that the benefits justify the costs of introducing robot persons to a legal system. Both need to be considered with proper care before moving further toward such an innovation. But in summary of our own investigation, we find the idea could easily lead to abuse at the expense of the legal rights of extant legal persons. We currently have a legal system that is, first and foremost, of, for, and by the (human) people. Maintaining the law's coherence and capacity to defend natural persons entails ensuring that purely synthetic intelligent entities never become persons, either in law or fact.

### Acknowledgements

All authors contributed equally to this project; their names are in alphabetical order. We thank the anonymous reviewers for useful comments, as well as the feedback from the participants of the AI Personhood workshop at the Princeton Center of Information Technology Policy in June 2017, particularly Bendert Zevenbergen. Thanks also to Rob Wortham and Andreas Theodorou for useful comments.

### References

- Asaro, P. (2007). Robots and responsibility from a legal perspective. In *The IEEE Conference on Robotics and Automation, Workshop on Roboethics*, Rome. unpublished proceedings, paper available <http://www.peterasaro.org/writing/ASARO%20Legal%20Perspective.pdf>.
- Blair, M. M. and Pollman, E. (2015). The Derivative Nature of Corporate Constitutional Rights. *William & Mary Law Review* 1673, 56.
- Bryson, J. J. (2009). Building persons is a choice. *Erwägen Wissen Ethik*, 20(2):195–197. commentary on Anne Foerst, “Robots and Theology”.

- Bryson, J. J. (2012). A role for consciousness in action selection. *International Journal of Machine Consciousness*, 04(02):471–482.
- Bryson, J. J. (2017). The meaning of the EPSRC Principles of Robotics. *Connection Science*, 29(2):130–136.
- Burgess, A. and Burgess, J. P. (2011). *Truth (Princeton foundations of contemporary philosophy)*. Princeton University Press, Princeton, NJ.
- Calderwood, K. (2016, last accessed 18/08/2017). Why New Zealand is granting a river the same rights as a citizen. *Australian Broadcasting Corporation*. <http://www.abc.net.au/radionational/programs/sundayextra/new-zealand-granting-rivers-and-forests-same-rights-as-citizens/7816456>.
- Dennett, D. C. (2001). Are we explaining consciousness yet? *Cognition*, 79:221–237.
- Dennett, D. C. (2009). The part of cognitive science that is philosophy. *Topics in Cognitive Science*, 1(2):231–236.
- Dewey, J. (1926). The historic background of corporate legal personality. *The Yale law journal*, 35(6):655–673.
- Dijkstra, E. W. (1970). Notes on structured programming. circulated privately.
- Domonoske, C. (2017, last accessed 18/08/2017). Elon Musk Warns Governors: Artificial Intelligence Poses ‘Existential Risk’. *National Public Radio*. <http://www.npr.org/2017/07/17/537686649/elon-musk-warns-governors-artificial-intelligence-poses-existential-risk>.
- Easterbrook, F. H. and Fischel, D. R. (1985). Limited liability and the corporation. *The University of Chicago Law Review*, 52(1):89–117.
- Eklund, M. (2011). Fictionalism. *The Stanford Encyclopedia of Philosophy*.
- Field, H. (1989). *Realism, mathematics & modality*. Blackwell.
- Foerst, A. (2009). Robots and theology. *Erwägen Wissen Ethik*, 20(2).
- Garrett, B. L. (2014). The Constitutional Standing of Corporations. *University of Pennsylvania Law Review*, 163.
- Geraci, R. M. (2010). The popular appeal of apocalyptic AI. *Zygon*, 45(4):1003–1020.
- Goertzel, B. (2010). AI against ageing: AIs, superflies, and the path to immortality. In *Singularity Summit*, pages 14–15, San Francisco, CA, USA.
- Greene, N. (2011, last accessed 18/08/2017). The first successful case of the Rights of Nature implementation in Ecuador. <http://therightsofnature.org/first-ron-case-ecuador/>.
- Huang, H. (2012). Piercing the Corporate Veil in China: Where Is It Now and Where Is It Heading? *The American Journal of Comparative Law*, 60(3):743–774.
- Joyce, R. (2001). *The myth of morality*. Cambridge University Press.
- Knoll, B. (2008). Rights without remedies: The european court’s failure to close the human rights gap in kosovo. *Zeitschrift für ausländisches öffentliches Recht und Völkerrecht*, 68:431–451.
- Koops, B.-J., Hildebrandt, M., and Jaquet-Chiffelle, D.-O. (2010). Bridging the accountability gap: Rights for new entities in the information society. *Minnesota Journal of Law, Science & Technology*, 11:497–561.
- Laufer, W. S. (1994). Corporate bodies and guilty minds. *Emory LJ*, 43:647.
- Liang, Y.-L. (1949). Reparation for Injuries Suffered in the Service of the United Nations. *The American Journal of International Law*, 43(3):460–478.
- Nékám, A. (1938). *The personality conception of the legal entity*. Harvard University Press.
- Rousseau, B. (2016, last accessed 18/08/2017). In New Zealand, Lands and Rivers can be People too (Legally Speaking). <https://www>.



- [nytimes.com/2016/07/14/world/what-in-the-world/in-new-zealand-lands-and-rivers-can-be-people-legally-speaking.html](https://www.nytimes.com/2016/07/14/world/what-in-the-world/in-new-zealand-lands-and-rivers-can-be-people-legally-speaking.html).
- Safi, M. (2016, last accessed 18/08/2017). Ganges and Yamuna rivers granted same legal status as human beings. <https://www.theguardian.com/world/2017/mar/21/ganges-and-yamuna-rivers-granted-same-legal-rights-as-human-beings>.
- Shah, S. (2013). Questions Relating to the Obligation to Prosecute or Extradite (Belgium v Senegal). *Human Rights Law Review*, 13(2):351–366.
- Smith, B. (1928). Legal personality. *The Yale Law Journal*, 37(3):283–299.
- Solaiman, S. M. (2016). Legal personality of robots, corporations, idols and chimpanzees: a quest for legitimacy. *Artificial Intelligence and Law*, pages 1–25.
- Sulmont, R. (2017, last accessed 18/08/2017). Les robots auront bientôt une responsabilité civile; le parlement européen examiner le mois prochain un rapport qui fixe le premier statut juridique des robots. <http://www.rtl.fr/culture/futur/les-robots-auront-bientot-une-responsabilite-civile-7786755275>.
- van Vollenhoven, C., MacGregor, G. F., and Parker, E. B. (1926). North American Dredging Company of Texas v. United Mexican States. *The American Journal of International Law*, 20(4):800–810.
- Wakefield, J. (2017, last accessed 18/08/2017). MEPs vote on robots' legal status - and if a kill switch is required. <http://www.bbc.co.uk/news/technology-38583360>.
- Wickremasinghe, C. and Evans, M. D. (2000). Difference Relating to Immunity from Legal Process of a Special Rapporteur of the Commission on Human Rights. *The International and Comparative Law Quarterly*, 49(3):724–730.