9-2023

Artificial Intelligence and Interspecific Law

Daniel J. Gervais

John J. Nay

POLICY FORUM

TECHNOLOGY AND LAW

Artificial intelligence and interspecific law

Law could recognize nonhuman AI-led corporate entities

By **Daniel J. Gervais**[1] *and* **John J. Nay**[2]

Several experts have warned about artificial intelligence (AI) exceeding human capabilities, a "singularity" at which it might evolve beyond human control. Whether this will ever happen is a matter of conjecture. A legal singularity is afoot, however: For the first time, nonhuman entities that are not directed by humans may enter the legal system as a new "species" of legal subjects. This possibility of an "interspecific" legal system provides an opportunity to consider how AI might be built and governed. We argue that the legal system may be more ready for AI agents than many believe. Rather than attempt to ban development of powerful AI, wrapping of AI in legal form could reduce undesired AI behavior by defining targets for legal action and by providing a research agenda to improve AI governance, by embedding law into AI agents, and by training AI compliance agents.

"Law" is a human invention. Humans make the laws and the enforcement institutions. The law is composed of directives according to which, "human beings are required to do or abstain from certain actions, whether they wish to or not" [(*1*), pp. 78–79]. Humans have occasionally invited into their legal order nonhumans such as animals, rivers, and ethereal entities created by humans called "legal persons" (*2*). Yet, neither animals nor rivers have a mind or intentions that make them responsible for their behavior in the eyes of the law. Up to this point, the legal system has been univocal; it only allows humans to speak to design and use it. In the legal system, nonhuman legal subjects have necessarily instantiated their rights through human proxies; those subjects "are simply a vehicle for addressing human interests and obligations" [(*2*), pp. 591–592].

The humanness of the legal system goes further. The system allows humans (vis-à-vis other humans) to exert control over other species. Humans can "own" other species, generating rights and responsibilities, including the right to decide which members of other species live or die. Thus, the law certainly affects other species, but it is meant to regulate relations among humans about other species.

Human laws are written using language, another distinctively human invention. Only humans speak our language, which we see both as evidence of the existence of our (superior) mind and, hence, as Chomsky put it, as a "true species property" (*3*). Functionally, the abilities that have made us the only species able to design and interact with the legal system is language. Language, and its emergent concepts, allow humans to understand norms and institutions that together constitute our legal system.

Human language is no longer distinctive to humans. We have developed AI that can do as well as or better than the best humans on at least 150 cognitive tasks. A capacity to understand human language and legal norms unlocks the ability of AI to interact with the legal system. Work in recent years on contract drafting, pretrial discovery, and now more complex tasks such as understanding fiduciary obligations all point toward the conclusion that it is only a matter of time before an "understanding" of human law by AI is achieved (*4*). Recent evidence includes the ability of large language models (LLMs) to accomplish hundreds of legal reasoning tasks (*5*).

Yet, even before machines outperform lawyers at much of what lawyers do, AI may enter the legal system as legal subject. This is a catalyst for a serious conversation. Our research is not motivated by any normative claim about rights for AI or "robots," whether based on the ontological properties of advanced AI or on the direct application of a social-relational model. It is about finding a way to apply what has essentially been human law to autonomous AI capable of performing many cognitive tasks that until recently only humans could—specifically when, having taken corporate form, an AI interacts with humans, businesses, and the legal system.

CORPORATE BUT NOT HUMAN

A major law reference book refers to corporations as "artificial persons" (*6*). This is particularly apt in this context because that is precisely what is happening; AI could take corporate form. The laws of several jurisdictions in the United States and elsewhere do not always explicitly require humans at the helm of a legal person. Some laws do require the existence of a board of directors, which, by interpretation, some could say must be composed of humans, but not all jurisdictions require such a board. The laws are even laxer in the case of Limited Liability Companies (LLCs). Overall, nothing generally prevents an AI from managing the affairs of a corporate entity. By law, corporate entities need not have human owners or managers.

The idea of a "zero-member LLC" is not new [e.g., (*7*)]. What has changed is that it is now possible from a technological standpoint. And the Uniform Limited Liability Company Act (ULLCA) explicitly provides for the possibility. Though it notes that absence of any members leads to dissolution of an LLC, the ULLCA did not make that provision mandatory (*8*). Even if only one US state allowed zero-member LLCs, that entity could operate nationwide under the so-called "internal affairs doctrine," according to which courts look to the law of the state of incorporation for rules governing the internal affairs of a corporate entity.

For instance, under state laws that made it the first US state to authorize the formation of LLCs, Wyoming could become the home of the first zero-member LLC. Although its corporate laws allowing the formation of zero-member LLCs predate AI, the state government is aware of the possibility and has not taken steps to prevent it. Indeed, the state would like to attract AI investment, like many other jurisdictions. Envision a zero-member LLC operating autonomously: taking orders for specific products online or via email; contacting suppliers; arranging for shipping and payment; managing feedback, returns, and complaints, all without any direct human involvement in the process.

If such an AI-operated LLC did exist, what options would the law have? The law

[1]Law School, Vanderbilt University, Nashville, TN, USA. [2]CodeX–The Center for Legal Informatics, Stanford University, Stanford, CA, USA. Email: daniel.gervais@vanderbilt.edu; john.j.nay@gmail.com

could not easily attach legal consequences to the autonomous acts of the AI—acknowledging that "autonomy" here may well be limited (but so it is generally also for humans). The law could perhaps attach responsibility to humans who launched the LLC, but this might require a new legal toolkit, as humans are generally shielded from liability for acts of a corporate entity that they create or control. Moreover, it is not inconceivable that an AI itself could apply to establish a new LLC, especially if it is already operating as one. Courts would still have all the tools they now use when a corporate entity is held liable: damages, seizure of assets, and dissolution come to mind. Courts could issue orders such as injunctions, but the decision to comply would be made by the AI.

There is another option: making zero-member LLCs illegal in all jurisdictions, but that would require a massive legislative effort worldwide and arguably runs counter to the ethos of attracting and growing technology industries that animate many jurisdictions around the world. Thus the emergence of zero-member LLCs should catalyze the legal system to adapt to autonomous AI agents. That said, the creation of AI-operated zero-member LLCs is difficult to reconcile with theories of the firm, which are anthropocentric and instrumental in nature. Those frameworks are based on the benefits that the firm provides to humans, just as, say, intellectual property is designed to ensure human progress (*9*).

CLEAR LEGAL SUBJECT

AI now has the capabilities to trade in currency. Although trading in fiat currencies (such as the US dollar, euro, yen, etc.) is highly regulated, AI can trade in digital currencies that settle on blockchains. The machine could thus pay humans to perform tasks that require interfacing with the physical world—at least until they can be performed efficiently by robots. Digital currencies can be decentralized by relying on distributed blockchain technology. That can then allow for AI to operate in a decentralized way, leading to the formation of decentralized autonomous organizations (DAOs) that make it (even) harder to regulate.

Having an AI operate as an LLC, or other legal entity, provides a clear legal subject. This has two benefits: It makes AI a target for legal action for compensatory damages paid by the corporate entity; and it provides a clearer research agenda for machine learning researchers to improve AI governance.

Using the legal system to create zero-member LLCs is a way to test its limits. But many corporations are already increasingly managed by AI. Many human resource functions, such as filtering of applications, have been delegated to AI. Copywriters are being replaced by LLMs, whose role in advertising is already prominent and poised to grow exponentially. Many aspects of financial planning can be performed by LLMs leveraging tools.

As AI moves up the cognitive tree and can perform ever more complex functions—in many cases surpassing humans—why would it not enter the C-suite? Soon, many parts of a chief technology officer's job (and possibly a chief financial officer's) will likely be done by AI. Lawyers already rely on many AI-powered tools. It won't be long before a corporation can ask AI for legal advice, which may, with time, be better than many human lawyers can offer.

"Machines are first introduced as…'copilots.' …But the purpose of a copilot is to be able to take over at some point."

This is happening fast because tens of billions of dollars are being spent to build more bridges, at ever increasing levels, between human and machine cognition. Machines are first introduced as assistants, or "copilots" (*10*). This "assist first, replace later approach" is natural because it enables a machine to be gradually trusted, put in place, and then AI capabilities improve through further research advances and through additional training data collected from real-world deployments. But the purpose of a copilot is to be able to take over at some point. At what stage does the machine no longer need feedback from humans specifically, rather than automated feedback or feedback from other AI systems? This is what AI is trending toward.

REGULATION AND RESEARCH

Because AI can be used in almost any field of activity—e.g., investing, legal services, power grid management, warfare, health care—the regulation of AI as a single, new regulatory target is awkward. Each governmental agency should instead consider the impact of AI and, in our opinion, the real possibility that autonomous AI agents will exist among humans.

Lawmakers face an unprecedented challenge: regulating AI that can perform cognitive tasks that until recently only humans could. What happens, as a matter of law, when another "species" interacts with us and makes "real-world" decisions—not through human proxies, but directly—and does all this "intelligently"? There is a real risk that AI technology will slip beyond human control. The regulatory equation to solve is that, even if AI machines can behave like humans—and in some cases much faster—we cannot regulate them exactly as humans (*11*). However, machine learning and computer science research more broadly may be able to unlock law-following tendencies in AI that, when paired with interspecific law, could enable AI governance.

Within the current paradigm, there appears to be a path toward embedding law-following into LLM-powered AI agents (*12*). For example, humans can train an LLM to accurately predict which real-world actions are more consistent with a given legal standard, e.g., fiduciary duties. The resulting LLM can be used to monitor, and/or influence, a primary AI agent, and it could be used as a reward model (in reinforcement learning parlance). If the latter, humans could train a "student" LLM with reinforcement learning against that reward model "teacher." The student may then exhibit behavior more consistent with the legal standard it was trained against. This would represent a step toward LLM-powered AI agents aligned with human law. To be clear, there are many situations involving human agents that ultimately require human adjudication of what is legal in that context; this does not change with AI agents. Embedding a deeper understanding of law into AI agents seeks to address the vast majority of day-to-day actions and situations but will never handle the highly ambiguous, or the edge cases that require a human court opinion.

More broadly, data generated by legal processes and the tools of law (methods of lawmaking, statutory interpretation, contract drafting, applications of standards, and legal reasoning) can be leveraged by researchers to train AI systems to learn methods for internalizing what one could label the "spirit" of the law, that is, the robust specification of often inherently vague human goals. Asimov's laws were a parable for why brittle rules (if-then statements) do not suffice. Our proposal is to instead focus AI research on legal standards and the spirit of directives.

We can continue to see the legal system with humans sitting alone at the top, allowing only physical things that we designate (specific lakes and rivers) or ethereal things created out of thin air (corporations) to exist as potential right holders. Beyond the normative case that one could make against this, the very idea with the advanced forms of AI that will emerge in the next few years is that we may want them to be part of our legal system so that we can track their actions, apportion blame, put guardrails around behavior, and guide AI research toward building lawful artificial "consciences." Making

legal compliance a core component of the AI training process and deploying an adversarial automated monitoring requirement is a potential path forward.

The two main enforcement tools of the legal system are financial penalties and imprisonment. As Stuart Russell noted, "[i]f we were to imprison the robot for nonpayment, why would it care?" [(*13*), p. 126]. The same could be said of the imposition of financial penalties, though an AI might "care" if it operates an LLC and has to obtain general liability insurance. But imagine an AI in an offshore jurisdiction where it may be easier to operate, trading only in digital currency. What grip would human law have on that entity? Instead of running from that, but having it happen, we can run toward bringing AI into the fold in a way that we can track everything it does and subject it to every law that a human must comply with, plus potentially more law specific to AI. But the key is that everything a human is subject to, the AI is as well.

The reality is that most people obey the law most of the time without any enforcement. Humans internalize behavioral norms consistent with legal standards. We may understand that if we were to ignore all laws—knowing that there is only so much the state can do to enforce them—a general state of chaos might emerge. For a variety of hard-to-express reasons, humans generally view reasonable compliance as desirable.

By contrast, machines are amoral. They can predict nuances of human morality, but their way of "internalizing" morals is different. Machine thinking can be much more powerful than human thinking, but it is devoid of key human characteristics such as interoception, the sense of what is happening inside our bodies. The challenge is to get the machines to "want" to comply without relying on an inherent inner compass that many humans seem to be endowed with through some combination of nature and nurture. The challenge goes beyond aligning AI with a set of ethical rules—however those are determined—which risks creating a second set of rules for AI that is not democratically determined through law-making processes.

The challenge is to align AI machines with the law on the books. We use the zero-member LLC as an exemplar of why this is desirable because an LLC should comply with the law (*14*). This is potentially more likely to happen, in our view, if we can make the case that we are willing to bring advanced AI into the legal system after it has had sufficient legal compliance training and is subject to sufficient adversarial monitoring and live AI-driven (but human interpretable) feedback. This AI-driven adversarial monitoring is part of how we keep the primary AI systems in line with ever-evolving law.

We can imagine a scenario in which the government requires any sufficiently advanced AI deployment to be accompanied by a certified AI compliance agent and, separately, a form of government-run AI auditor or police leveraging similar AI legal compliance technology for a final line of public defense. Our focus is on catalyzing AI research that can help prevent illegal behavior of AI agents through private and/or public initiatives. With the speed of digital intelligence, a focus on proactive automated crime prevention is critical. But in the inevitable situations where AI systems don't live up to the ideal behavior because of wildly out-of-distribution situations in which they find themselves, the relevant corporate entity should be liable for paying damages to those that prove harm to a court. To enable this, frameworks could be extended that require minimum business liability insurance policies to be held. This is common practice for many areas of business. And, of course, where a human can be found liable for misuse of AI, they should be.

"The challenge is to get the machines to 'want' to comply without relying on an inherent inner compass..."

BRING AI INTO THE LEGAL FOLD

Interspecific law will happen, but it is impossible to predict where on the spectrum we will end up. At one end, interspecific law means slightly adapting corporate law to the operation of corporate entities with humans only partly in control. This is imminent because AI is already in control of many corporate activities. At the other end, it means adapting the legal system to everyday interactions with autonomous, intelligent entities in cases where prosaic tools of the legal system, like post hoc enforcement, are unlikely to work in the traditional way.

Some will dismiss the idea of letting machines enter the legal system as subjects as nonsense. To those readers, three things. First, they could do so soon, at least indirectly, using zero-person LLCs where there is no human governing a legal entity, as explained above. Second, if we don't proactively wrap AI agents in legal entities that must obey human law, then we lose considerable benefits of tracking what they do, shaping how they do it, and preventing harm. Third, if the point is that machines will never have powerful agency well beyond human oversight: We hope you are right!

There seems to be another choice. We could try to put a hard stop on AI. Some scientists have warned that developing AI more advanced than humans is extremely risky. In our view, this hard stop will likely not happen. Capitalism is en marche. There is too much innovation and money at stake, and societal stability historically has relied on continued growth.

AI replacing most human cognitive tasks is a process that is already underway and seems poised to accelerate, although the possibility of a major public opinion backlash that would limit the scope and depth of replacement of humans cannot be excluded (*15*). This means that our options are effectively limited: Try to regulate AI by treating the machines as legally inferior, or architect AI systems to be law-following, and bring them into the legal fold now with compliance tendencies baked into them and their AI-powered automated legal guardrails. Research directions described here suggest that this may be possible. We think it may be desirable. ■

REFERENCES AND NOTES

1. H. L. A. Hart, *The Concept of Law* (Clarendon, 1994).
2. R. L. Cupp Jr., *Fla. Law Rev.* **465**, 506 (2017).
3. J. Chen, *Wash. Univ. Law Q.* **73**, 1263 (1995).
4. J. Nay (2023); https://law.stanford.edu/publications/large-language-models-as-fiduciaries-a-case-study-toward-robustly-communicating-with-artificial-intelligence-through-legal-standards/.
5. N. Guha *et al.*, LegalBench: A Collaboratively Built Benchmark for Measuring Legal Reasoning in Large Language Models (2023); https://arxiv.org/abs/2308.11462.
6. *Black's Law Dictionary* defines an "artificial person" as a "nonhuman entity that is created by law and is legally different owning its own rights and duties. AKA juristic person and legal person."
7. S. J. Bayern, *Northwest. Univ. Law Rev.* **108**, 1485 (2014).
8. https://www.uniformlaws.org/committees/community-home?CommunityKey=bbea059c-6853-4f45-b69b-7ca2e49cf740. As of 2023, 20 US states have neither enacted the ULLCA nor introduced legislation to do so. It is worth nothing that the 2013 amendments to the ULLCA (RULLCA) do seem to require at least one person as member, for example Section 20, and Section 201.
9. B. R. Cheffins, "The Rise and Fall (?) of the Berle-Means Corporation" (2018); https://doi.org/10.17863/CAM.27945.
10. See, for example, https://github.com/features/copilot.
11. T. Arnold, M. Scheutz, *Ethics Inf. Technol.* **20**, 59 (2018).
12. J. J. Nay, *Northwest. J. Tech. Intell. Prop.* **20**, 309 (2023).
13. S. Russell, *Human Compatible: Artificial Intelligence and the Problem of Control* (Viking, 2019).
14. C. O'Keefe, Law-Following AI 1: Sequence Introduction and Structure, AI Alignment Forum (2022); https://www.alignmentforum.org/posts/NrtbF3JHFnqBCztXC/law-following-ai-1-sequence-introduction-and-structure.
15. J. Stilgoe, *Science* 10.1126/science.adk0176 (2023).

ACKNOWLEDGMENTS

J.J.N. is the founder and CEO of Nomos AI and a cofounder and the chairman of Brooklyn Investment Group.

10.1126/science.adi8678