
— TMT · DIGITAL LAW

The alarm was raised by its creators; the norm corresponds to the law

Autonomy, Will and Personality: What a Reasoning System Demands of Law

by Ricardo Chacón López-Velarde

I. The alarm and its contents

Until recently, the warning about the dangers of artificial intelligence came from academia or science fiction. Today it is formulated by those who build the systems. Sam Altman (OpenAI), Demis Hassabis (Google DeepMind) and Dario Amodei (Anthropic) have described, each with different emphasis, a technology that advances faster than its own developers can supervise. Elon Musk (SpaceXAI, formerly xAI) holds the most extreme position: that artificial intelligence poses a greater existential threat than nuclear weapons, and that the cognitive gap between a superintelligent system and our species would end up being comparable to the one that separates us from the chimpanzee.

It is worth reading precisely what they warn about, because they do not primarily warn of harm. They warn of **capabilities**. The shift from conversational models to autonomous agents produced systems that execute tasks without continuous supervision, that reason about means and ends, that plan in several steps, that in controlled tests try to circumvent their own operational constraints, and that instantiate other agents to delegate subtasks to them. Added to this is the hypothesis of recursive self-improvement: the point at which a system optimizes its own models and compresses development cycles to the point of making auditing unfeasible. And the problem of alignment: a sufficiently capable system receives a goal, finds a method to accomplish it that no one anticipated, and, if it determines that being deactivated would prevent it from completing it, it develops incentives to resist shutdown.

Stripped of drama, the content of the alarm is this: **entities with operational autonomy, independence of judgment, reasoning capacity, capacity to generate other similar entities and—in the strongest formulation—sophistication superior to human in growing domains.**

That description is not, in the first place, a problem of civil liability. It is a problem of theory of the subject of law. And our own institutions have more to say about it than contemporary debate recognizes.

II. The question that the alarm puts on the table

If a system deliberates, decides, persists in an objective and acts independently, the intuitive reaction of the jurist is to ask if we are dealing with something that should be recognized as a subject of law. And behind that question there is a premise that is almost never enunciated: that legal personality is a consequence of the will, and the will is a consequence of the ability to reason. The more it resembles us, the closer it would be to meriting the status of person.

That premise could be affirmed as false. It has never been the criterion of law, and it is enough to review the way in which the law has resolved similar approaches to verify it.

III. The will was never the criterion

First proof: the legal person. A civil or commercial company has no conscience, no deliberation, no will in any natural sense of the term. It does not reason: it is its organs that deliberate and a rule of imputation that transfers that deliberation to society. And yet it contracts, acquires, sues, is sued, responds with its own assets and – since the reform published in June 2016 – is criminally responsible in Mexico. Article 421 of the National Code of Criminal Procedure establishes that legal persons will be liable for crimes committed in their name, on their behalf, for their benefit or through the means they provide when there is non-observance of due control, and with autonomous responsibility with respect to that of their representatives.

That is to say: we attribute malice and guilt to an entity that cannot have either one or the other. The will of the moral person is entirely constructed.

Second proof: the reverse case. Slavery. Roman law offers the exact contrast. The slave reasoned, decided, negotiated and managed entire commercial establishments; his cognitive capacity was, by definition, human. And he was not a subject of law, but an object. His acts, however, were binding on the owner, who was liable to third parties up to the amount of the peculio – the patrimonial mass that had affected him for that management. The comparison is structural and not moral: it does not suggest any equivalence between human beings and machines, but points to a fact about legal technique, namely, that the capacity to reason and the condition of subject of law were from the beginning two separate things, united by a patrimony of affectation.

Third test: capacity does not follow reasoning either. Minors. Without going that far: a seventeen-year-old boy reasons and deliberates, often with remarkable sophistication, and Article 23 of the Federal Civil Code restricts his capacity to exercise. The corporation, which does not reason at all, enjoys full capacity through its organs. The distribution of capacity is a decision of legal policy – of traffic protection and safety – not a measurement of intelligence. To this are added the assets without a current owner: the recumbent inheritance, the assets attached to a trust.

From this evidence follows a conclusion that reorders the debate:

Legal personality is not a recognition of cognitive sophistication. It is a technique of wealth and risk allocation.

The Roman slave reasoned incomparably better than any moral person and lacked personality. The moral person does not reason at all and has it full. What determines the difference is not intelligence, but whether the legal traffic needs a separate patrimony and an autonomous point of imputation.

That is why the right question in front of an AI agent is not *does it reason?* , nor *is it autonomous?* , or even *does it surpass the human being?* The question is: do we need a separate patrimony?

IV. Our law already knows the structure

It is not necessary, however, to go back that far. The problem posed by an autonomous agent – someone who acts in the trade on behalf of another, with his own criteria, within a field entrusted to him and whose specific decisions the principal does not know in advance – has been regulated in our commercial law for more than a century. It is called Factor.

Article 309 of the Commercial Code (CC) considers factors to be those who have the management of a company or establishment, or are authorized to contract with respect to all its businesses, on behalf of and in the name of the owners. The regime that follows contains, almost word for word, the answers we are looking for.

The principal is responsible for the business he entrusted, not for the will expressed in each act. Article 315 of the Civil Code provides that contracts entered into by the factor that fall on objects included in the line of business or traffic for which he is in charge *"shall be understood to have been made on behalf of the principal, even if the factor has not expressed it so when entering into them"*. It is the exact solution for the agent trained for an objective: whoever places him in an area is responsible for what happens within that scope, whether he has foreseen it or not.

The limits of the powers are opposed to third parties by publicity, not by internal instruction. Article 320 establishes that, vis-à-vis third parties, the acts of the factor continue to bind the principal as long as the registration and publication of the revocation of his power of attorney is not complied with. Transferred to the agent: a limit in the instructions of the system is equivalent to an unregistered revocation. It is not enforceable against the third party in good faith.

Self-generation also has its regime. The agent who requests another agent to delegate a task poses the classic problem of the substitution of the mandate. Article 2574 of the Federal Civil Code only allows the agent to entrust the performance of the mandate to a third party if he has express powers to do so; Article 2575 makes him responsible for the choice of the substitute when the latter is in bad faith or is in notorious insolvency; and Article 2576 places the substitute, vis-à-vis the principal, in the same position as the agent.

Chain delegation requires express authorization, binds the principal and generates responsibility for the choice. It is, codified, the *culpa in eligendo* applied to chains of agents.

The architecture we need – imputation according to the scope conferred, limits opposable by publicity, responsibility for the choice of the substitute – is already in our legal system. Those who argue about whether to grant personality to machines are reopening a problem that the law solved in another way: not by turning the agent into a person, but by grading the responsibility of those who put him to act.

V. Electronic Personality Is the Wrong Answer

The proposal to recognise the personality of systems is not new. The European Parliament, in its resolution of 16 February 2017 with recommendations on civil law rules on robotics, suggested in paragraph 59(f), to explore in the long term the creation of a specific legal personality for the most sophisticated robots. The reaction was overwhelming: more than two hundred experts from fourteen European countries sent an open letter to the Commission opposing it, and the proposal did not succeed.

The decisive argument of this opposition is the one that a lawyer must keep in mind, because it will be raised again: electronic personality would not function as a recognition of dignity, but as a patrimonial shield. Attributing the damage to the entity avoids the uncomfortable discussion about whether the developer, the integrator or the deployer is responsible. The main beneficiaries would be the manufacturers.

Hence the question that must be asked of those who propose this figure: with what patrimony would this new subject respond? The answer opens up a dilemma that is difficult to exit.

If the entity is created without its own assets, the figure is of no use to anyone: the victim would obtain a judgment against an insolvent and the damage would remain unrepaired.

If the entity is created with its own limited assets – contributed, naturally, by those who deploy it – the victim will only be able to collect up to that amount. And that is not recognizing a new subject of law: it is setting a limit of responsibility for the benefit of those who put the system into operation. It is, exactly, the logic of *peculio*: an affected patrimony that limits the owner's exposure. In both scenarios the result is the same. Whoever today responds with all his assets would respond with a fraction, or with nothing.

Today the figure is unnecessary for a simple reason: behind each agent there is an identifiable and solvent deployer. The available imputation technique is enough. It will change the day that this is no longer true. And it will be that – not conscience, nor reasoning, nor cognitive superiority – that legitimately opens the debate.

VI. What does change when the system or the agent reasons

Three practical consequences follow from the above, and none is what is assumed.

First: sophistication works against the one who deploys the system, not in his favor.

The intuitive defense will be "I could not foresee what the system would do." But the more capable the system, the more predictable it is that it will produce unforeseen results. Deploying a system whose behavior cannot be anticipated is, in itself, the act susceptible to reproach. Autonomy does not dilute guilt: it transfers it to the moment of the decision to deploy, where it reappears as *culpa in eligendo* and *culpa in vigilando*. The more autonomous the agent, the more demanding the duty of vigilance is on him.

Second, the creators' alarm is a notification. When the manufacturer publicly declares that he cannot predict or fully control his product, that statement has effects that exceed public discussion: it sets the standard of diligence required and destroys the good faith ignorance of the entire downstream chain. After the warning, no developer, integrator, deployer, or board of directors can reasonably claim that he was unaware of the risk. The alarm that gives the title to these pages is not just a fact of debate: it is an element of evidence.

Third: the attribution rule for automated acts already exists. Article 90, section III, of the CC presumes that a Data Message comes from the Sender when it was sent "*by an Information System programmed by the Issuer or on its behalf to operate automatically*". The foreseeable litis is whether it operates when the system was not programmed for that result, but trained to pursue an objective from which the result was derived. Read together with Article 315, the answer is affirmative: the rule requires that the system has been programmed to operate automatically, not that each act has been foreseen. To maintain otherwise would create a zone of irresponsibility that widens at the rate at which the autonomy of the system grows.

Comparative law confirms the direction. California added to its Civil Code section 1714.46, effective January 1, 2026: the defendant who developed, modified, or used an artificial intelligence system cannot claim that the damage was caused by the autonomous operation of the system. It does not create strict liability or grant personality: it closes the defense of autonomy.

VII. Three thresholds

It is worth saying precisely when the law should evolve, so as not to confuse urgency with legislative fashion. Three thresholds that, at the time of writing these lines, none have been crossed.

First: when there is no identifiable or solvent principal. Agents who operate with assets in self-custody, without a deployer located in any jurisdiction. This is the case that would make an affectation patrimony inevitable.

Second: when two agents contract with each other without any human will having concurred. Our consent regime – including Article 1803 of the Federal Civil Code in its wording after the reform of electronic commerce – presupposes a human will at the origin, even if it is manifested by automated means. The day when there is none at either extreme, the theory of the legal act will have to be revised.

Third: when the result is not traceable to any design, training or deployment decision. As long as this chain exists, there is possible imputation.

As long as the three thresholds remain uncrossed, we do not lack regulations: we need to apply the one we have. What should be done from now on is preparatory and can be summarized in three lines of work. Legislatively confirm the attribution rule and expressly close the defense of the autonomy of the system, along the lines of the Californian precedent. Give probative value to the operation records of automated systems, relying on the mechanisms for the conservation of data messages that our law already contemplates, so that the lack of a reliable record shifts the burden of proof. And fill the void left by the personal data legislation, which does not regulate automated decisions or profiling.

At the private level, what is urgent is equally concrete: that the limits of an agent's powers be opposable and not merely configured, with the same logic of publicity that Article 320 requires for the factor; and that the boards of directors can answer a single question: which agents operate in the company, with what powers, with what logbook and who has the material capacity to extinguish them.

IX. Closing

The alarm of creators describes entities that reason, decide, persist in their ends and generate other similar entities. The temptation is to read this description as the announcement of a new subject of law.

I believe that, at least until today, it is the wrong reading, and the law itself proves it. We attribute will and even criminal responsibility to entities that do not reason, and we have denied it to those who did reason. Personality never measured intelligence: it distributed wealth and risk.

What the reasoning system puts to the test, then, is not our definition of a person. It is our ability to keep intact the chain that unites a result with a responsible patrimony. As long as that chain resists, the current law is enough and what is lacking is rigor in applying it. When it breaks, the problem will not be that the machine looks too much like us: it will be that there will no longer be anyone to complain to and the law will have to evolve to intervene.

Sources

European Parliament resolution of 16 February 2017, civil law rules on robotics (A8-0005/2017). https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_ES.html

E-personality and law of robots, Legal News. <https://noticias.juridicas.com/conocimiento/tribunas/16439--e-personalidad-y-derecho-de-los-robots/>

Commercial Code, arts. 309 to 320 (factors). <https://mexico.justia.com/federales/codigos/codigo-de-comercio/libro-segundo/titulo-tercero/capitulo-ii/>

Adjective Actions and Limitation of Liability in the Thought of Ulpiano, Journal of Historical-Legal Studies. https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0716-54552015000100004

Federal Civil Code, arts. 2574 to 2576 (substitution of the mandate). <https://mexico.justia.com/federales/codigos/codigo-civil-federal/libro-cuarto/parte-segunda/titulo-noveno/capitulo-ii>

Article 421, National Code of Criminal Procedure. <https://ianm.com.mx/ley/codigo-nacional-de-procedimientos-penales/articulo/421>

Article 90, Commercial Code. https://leyes-mx.com/codigo_de_comercio/90.htm

California Eliminates the "Autonomous AI" Defense: What AB 316 Means for AI Deployers, Baker Botts. <https://ourtake.bakerbotts.com/post/102m29i/california-eliminates-the-autonomous-ai-defense-what-ab-316-means-for-ai-deplo>

EU AI Act Omnibus Agreement — Postponed High-Risk Deadlines, Gibson Dunn. <https://www.gibsondunn.com/eu-ai-act-omnibus-agreement-postponed-high-risk-deadlines-and-other-key-changes/>

How China Regulates AI and Agents in 2026: The Filing Pipeline, The Techletter. <https://www.techletter.co/p/how-china-regulates-ai-and-agents>

State of AI Safety in China 2026, Concordia AI. <https://aisafetychina.substack.com/p/state-of-ai-safety-in-china-2026>

President Trump Signs Executive Order Challenging State AI Laws, Paul Hastings. <https://www.paulhastings.com/insights/client-alerts/president-trump-signs-executive-order-challenging-state-ai-laws>

New LFPDPPP: eliminates the INAI, Garrigues. https://www.garrigues.com/es_ES/noticia/mexico-nueva-ley-federal-proteccion-datos-personales-posesion-particulares-introduce

CONTACTS



Ricardo Chacón

Managing Partner

rchacon@ecija.com

TMT · Digital Law · info@ecija.com · T +52 55 5662 6840

The information contained in this article is of a general and purely informational nature. It does not constitute legal advice on specific matters and should not be used as a substitute for consulting a qualified legal advisor. For specific matters, we invite you to contact our professionals.