

Research Article

Artificial Intelligence, the Theory of Obligations and the Analogical Use of the Roman Servus Institution

Alonso Salinas Garcia^{1,*} , Matias Kahn Aranguiz² 

¹Law School, Pontifical Catholic University of Chile, Santiago, Chile

²Law School, Alberto Hurtado University, Santiago, Chile

Abstract

This article analyzes the legal challenges posed by artificial intelligence (AI) within the framework of general contract law, given that it lacks the essential characteristic required for the attribution of liability: the will, which is inherently human. Following the jurist Dr. Carlos Amunátegui Perelló, the solution to this difficulty is proposed to be the analogous application of the status of the Roman slave, which would allow for the easy regulation of contemporary artificial agents. Various Roman institutions are examined, such as the *actiones adiectitiae qualitatis* (*actio quod iussum*, *exercitoria*, *institoria*), *noxal actions*, the *actio de pauperie*, the *peculium*, and the *actio in rem verso*, all of which aim to establish contractual and extracontractual liability for the actions of entities lacking legal personality, with the goal of analogously applying these concepts to artificial agents—entities that can make decisions while being things and lacking the attributes of personality, and thus, ultimately, not being persons. Thus, it is concluded that an analogical use to the Roman *servus* is necessary in order to integrate AI into the theory of obligations in an organic and efficient manner, while respecting the ontological limits of human personhood, without fictitiously attributing solutions or assumptions incompatible with the artificial agent's status as a thing.

Keywords

Artificial Agents, Artificial Intelligence, Civil Liability, Contractual Liability, Non-contractual Liability, Roman Law

1. Defining the Boundaries of Artificial Intelligence

According to Carlos Amunátegui Perelló, contemporary artificial intelligence is rooted in the tradition of cybernetics and information theory. This scholar starts from the idea that many current technologies—digital platforms, AI, and neurotechnologies—are, at their core, applications of informational principles to various mechanical, biological, or social systems. From this perspective, both machines and living organisms can be conceived as systems governed by information flows that are encoded, transmitted, and processed according to formal rules,

thereby generating orderly behaviors without the need for any consciousness [1].

In this regard, the development of modern AI is closely linked to the McCulloch and Pitts model, which described the nervous system as a network of binary units that fire or do not fire when certain thresholds are exceeded, making it possible to represent neural activity as a logical calculation that can be implemented in machines [2, 3].

However, although this is a simplified model of the actual

*Correspondence: Alonso Salinas Garcia (alonso.salinas@uc.cl)

Received: 26 March 2026; **Accepted:** 20 April 2026; **Published:** 30 April 2026



brain and insufficient to explain the complexity of conscious and unconscious activity—and even limited in explaining the concept of “mind”—its positivist premises ultimately served as the foundation for artificial neural networks and, ultimately, for large language models, such as ChatGPT, which learn statistical patterns from enormous volumes of data.

In this vein, when we use the word “learn,” it is technically a misnomer, since these systems do not “understand” in the strong sense, as Amunátegui explains, but rather identify regularities and produce outputs that maximize statistical probability given an input. Under no circumstances are they an artificial embodiment of the intuition, perception, decision-making, and behavior characteristic of human beings, since the activity of the mind cannot be subsumed, limited, or described as a set of computational operations on information.

From Amunátegui’s perspective, this reinforces the idea that we are dealing with extraordinarily sophisticated digital computing machines, but not with new ontological entities endowed with an act of being of their own that allows them to possess themselves and, consequently, to have consciousness, will, or self-perception. In this context, artificial intelligence is not a personal entity, but rather a set of techniques designed to automatically solve problems that, for human beings, require intellect [1].

That said, and having understood the nature of artificial intelligence, we must make a fundamental distinction between symbolic AI—which attempts to replicate human reasoning through preprogrammed logical rules—and connectionist or subsymbolic AI, which is inspired by the functioning of biological neural networks to learn directly from data. The latter forms the basis for so-called “artificial agents,” a category Amunátegui uses to describe programs or systems that operate within the legal sphere with a certain degree of functional autonomy: they set prices, negotiate terms, execute complex orders, and interact with other systems without constant human intervention. For example, in e-commerce, there are algorithms that today not only convey the will of the parties—as was the case with email—but can ultimately shape it, optimize it, or even make decisions within the parameters pre-established by whoever uses them, whether a public or private entity [1, 3, 4].

At this point, we are faced with one of the primary sources of difficulty in legal doctrine and statutory regulations (designed with regard to persons essentially endowed with free will in order to attribute liability, standards of conduct, specific outcomes, or express prohibitions): functional autonomy. This distinctive characteristic of these automata is highlighted by three main features: i) the ability to operate without continuous supervision, ii) the capacity to learn or update itself based on new data—that is, so-called “machine learning”—and iii) the technical complexity that renders many of its decisions unpredictable, giving rise to the well-known “black box” problem [5, 6].

This combination breaks the classic paradigm of neutral

“media” that served to regulate e-commerce in the West during the 1990s. That old regulation was based on the assumption that the technological medium did not alter either the legal structure of consent or the attribution of liability. In contrast, artificial agents are technical objects endowed with a peculiar relative autonomy: they are things—intangible assets, programs, or systems—that someone owns or uses, but which make decisions that are not entirely predetermined on a case-by-case basis by the natural or legal person using them. Added to this is their capacity for self-modification—adjustment of internal parameters, deep learning, recalibration in response to new data—which causes the agent itself to evolve over time, making it even more difficult to anchor liability to a specific human entity [5, 6].

Consequently, it is imperative to emphasize that artificial agents are not persons. Ontologically, they lack consciousness; they are “zombie intelligences” that manipulate symbols without understanding or participating in their subjective meaning, since they cannot possess the capacity to experience the world. This implies that they lack both an act of being and an intellect of their own, since, although they simulate intellectual functions, they do not and will not possess the free will or self-awareness required to be persons.

Similarly, they do not enjoy freedom or will in the legal sense, as their actions are determined by algorithms and statistical correlations; they do not choose their own ends, but rather optimize an “objective function” set by a third party. Amunátegui draws here on Searle’s critique of the Turing test through the metaphor of the “Chinese room”: a machine can generate behaviors indistinguishable from human ones by following purely syntactic rules, without any semantic understanding of the content [3].

Precisely because AI is an entity with relative functional autonomy and the capacity for self-modification, its classification within the classical categories of civil liability becomes particularly problematic and complex. This fundamental dilemma demands a rethinking of traditional frameworks of attribution without falling into the fiction of granting them full legal personality, which they cannot possess as they are objects lacking all elements characteristic of personality—such as consciousness, free will, the shared experience of reality, will or intellect as such, and, above all, the possession of their own act of being.

2. The Crisis of the Theory of Obligations in the Face of Autonomous But Will-less Agents

The emergence of artificial intelligence generates a crisis in the general theory of obligations, whose classical structure is based on human will and fault. Contemporary contractual theory stems from the famous Roman classification by Gaius and Justinian, which distinguishes four sources: obligations *ex contractu*, *quasi ex contractu*, *ex delicto*, and *quasi ex delicto*

(Gai. 3.88; IJ. 3.13.2). In this regard, in the world of Ancient Rome, a contract is defined as an agreement of wills that gives rise to obligations; a quasi-contract covers situations where a complete agreement is lacking, but the legal system imposes duties for reasons of equity; tort encompasses intentional unlawful acts that cause harm; and, finally, quasi-tort covers negligent or harmful acts that do not strictly constitute a criminal offense [1, 6, 8].

This fourfold division was transmitted to the Middle Ages through the *ius commune* across Christendom, permeating university teaching and the work of the glossators, and was systematized by modern Enlightenment doctrine—from Domat to Pothier—which decisively influenced modern civil codes. Especially through the codified reception of the Napoleonic Code and the Latin American codes, five sources of obligations were established throughout the West: contract, quasi-contract, tort, quasi-tort, and unjust enrichment or undue payment.

In this tradition, the contract is based on the consensus reflected in the maxim from the Digest 2. 14. 1. 3: “*nullum esse contractum, nullam obligationem, quae non habeat in se conventionem*,” as Amunátegui notes. The central idea is that, without a meeting of the minds, there is no true binding obligation; thus, the capacity to consent constitutes the basic prerequisite for contracting. However, before the emergence of AI as a widespread phenomenon, a discussion regarding e-commerce developed in the 1990s, whose common and decisive starting point was the premise that new media—email or web forms—did not alter the structure of consent, since these virtual mechanisms were considered mere channels for expressing a pre-existing human will. Thus, instruments such as the United Nations Convention on Electronic Communications or the European E-Commerce Directive limited themselves to equating, in principle, electronic form with traditional written form, without revising the notion of consent [1, 3, 6-9].

In contrast, the emergence of artificial agents breaks this equivalence, because we are no longer dealing with simple “electronic messengers,” but rather are exposed to systems that set essential business conditions—price, term, quantity, or even the selection of a counterparty—based on complex algorithms with a certain degree of relative autonomy.

Along these lines, Amunátegui emphasizes that, from a technical standpoint, many of these agents are now capable of acting in a manner similar to an agent, selecting, negotiating, and entering into contracts in accordance with efficiency parameters designed by their owner. However, in civil law systems, the concept of agency presupposes a person and free will in the agent: agency is a consequence of legal personality. The idea that an object—a program or a robot—could be an agent is conceptually contradictory in this tradition, which stems from Pothier and the magnum opus of the Enlightenment and positive law, the Napoleonic Code—for which the person representing another in legal transactions must be a subject of rights and obligations insofar as they are a subject endowed

with free will (unlike in Medieval Law, where there are “degrees” of will and personal statuses, centering legal activity on its relationship to the common good and the cause motivating the act) [1, 3, 6].

The result of this is an evident tension, since factually, AI acts as an agent, but legally, it remains irrevocably a thing. The question then arises: who is liable for contracts entered into by the system when these have not been individually reviewed by any human being?

One apparent alternative is the Roman quasi-contractual figures, such as the management of another’s affairs (*negotiorum gestio*), which allows for the imposition of obligations when a person, without a mandate, manages another’s affairs in a useful or necessary manner. In the contemporary technological context, this is conceivable for an automaton, since an AI system manages information, assets, or decisions for the benefit of another without a typical agreement, functionally approaching this concept. However, the management of another’s affairs necessarily and definitively presupposes the existence of a managing person who altruistically decides to intervene [1, 3, 6, 8, 9].

Consequently, applying this category to algorithms would imply projecting onto things a framework designed for conscious subjects, which would necessarily lead to conceptual chaos that could result in enormous injustices. It is also worth recalling that medieval and modern law expanded the scope of quasi-contracts—for example, with *de facto* partnerships or the collection of undue payments—but always maintained the necessary participation of persons as such at both ends of the obligatory relationship.

In the area of tort law, on the other hand, the Roman law tradition generally requires intent or negligence on the part of the agent to attribute liability for the harm, such that liability falls on those who act with harmful intent or, at the very least, with negligence, deviating from the standard of the *bonus pater familias* [8, 9].

Now, in this context, the problem is that AI lacks shared experience of reality, subjectivity, and consciousness; therefore, it neither “intends” nor “acts negligently,” but simply performs computational processes on data in accordance with its architecture and training. Thus, attempting to directly attribute a civil offense or quasi-offense to the machine, no matter how autonomous it may be, would lead to an irrational and dangerous fiction, because it would incorporate into the device categories designed to evaluate human behavior, potentially leading to reckless social behaviors harmful to community life, as well as the attribution of responsibilities that cannot be rectified, since AI is an entity lacking elements of personality [1, 3, 8, 9].

Likewise, the high unpredictability of “black box” systems sometimes makes it difficult to attribute human negligence in classical terms, since not even the designers can clearly reconstruct the chain of decisions that led to the harm—a situation exacerbated by the existence of “deep learning” and “machine learning” [10].

To address these problems, legal doctrine has explored various solutions, among which the following stand out: i) the objectification of liability, through strictly objective regimes for certain high-risk uses of AI, and ii) assimilation to liability for animals, where the owner or guardian of the animal causing harm is liable, a concept that dates back to the Roman *actio de pauperie*. In this latter model, the animal—like AI—is incapable of intent or fault, but the owner is liable for the risk it introduces into the marketplace [1, 3].

With regard to undue payment (*solutio indebiti*) and, more broadly, unjust enrichment, these are legal institutions designed to correct financial imbalances when someone unjustifiably enriches themselves at another's expense. In this regard, in practice, artificial intelligence systems can generate problems of this nature, such as erroneous automatic charges, improper transfers, or appropriations of informational value (particularly through training on data without sufficient legitimacy), which result in asymmetric enrichments that the law must rectify [1].

In this vein, the basic structure of unjust enrichment—an increase in wealth for one party, a corresponding impoverishment for another, and the absence of a legitimate legal cause—is fully applicable in AI scenarios, but it becomes more complex due to the lack of identification of the “author” of the payment or the enrichment. Did the system pay? Did its owner pay? Did the infrastructure provider pay? AI operates as an opaque and nebulous intermediary between assets, blurring the causal line that the classical model took for granted.

In this context, Professor Carlos Amunátegui Perelló argues that to date there is no comprehensive liability regime that allows for the satisfactory attribution of the consequences of the acts of artificial agents. Instead, various analogies have been proposed, such as applications of the regime governing animals, product liability, or even the idea of an “electronic personality.” However, no solution has achieved sufficient conceptual and systematic robustness [1].

For this reason, Amunátegui himself suggests looking to historical figures in which analogous categories existed; although these are clearly reprehensible in retrospect due to the intrinsic cruelty of their application to human beings, they are useful for forming an analytical framework applicable to AI. We refer to human beings without legal personality—in particular, Roman slaves—who created legal obligations for their dominus, functioning as “quasi-subjects” or “instrumental subjects” of the system [3].

This comparative approach promises a more organic treatment of artificial intelligence, precisely because it avoids projecting onto machines modern categories that are inextricably linked to human dignity. Thus, in the face of algorithmic opacity and the evidentiary difficulty of fitting AI into classical contractual or tort liability, the historical experience of “human automatons without legal personality” offers the most readily applicable model for rethinking, in analogical terms, the sources of liability in the era of artificial agents.

3. The Dual Status of the Roman Servus, Its Legal Nature and Its Implications for the Theory of Obligations

Amunátegui notes that the word “robot” derives from the Czech *robot*, meaning “forced labor” or “slave.” The term was popularized by Karel Čapek in his play “R.U.R. (Rossum’s Universal Robots),” where it refers to artificial creatures created to serve as a servile labor force, thus making explicit the conceptual connection between mechanical automatism and slavery. This etymology reinforces a relevant historical intuition, in which the earliest modern imaginings of intelligent automatons proceeded to conceive of them directly as substitutes for human slaves—bodies in the service of a master, devoid of moral autonomy. The connection to Roman law, where the slave is simultaneously a human person and a legal object, and to the current reality of artificial intelligence, thus emerges almost naturally [1, 3, 11].

In this regard, in classical Roman law, the slave was a *res mancipi*, a property over which the dominus exercised extensive power, but at the same time was a human being capable of acting materially in the world, conducting business, causing harm, and generating wealth. The paradox lies in the fact that, despite his humanity, he lacked legal personality; consequently, his acts did not generate obligations for him, but for his owner, to the extent and under the conditions recognized by the law. This dual status—as both an object of property and a source of legal effects—makes the slave an exceptionally useful figure for thinking about artificial intelligence. In both cases, someone acts in the marketplace without being a subject of rights and obligations, and the legal system must decide how and to what extent that action connects to another's assets.

Now, turning specifically to the sources of obligations, first, in contractual matters, the primary source of the owner's liability for the acts of his slaves was the *actio quod iussum*. This action allowed a third party who had contracted with the slave to sue the dominus directly when the latter had ordered the contract or had placed the slave in charge of a business—a presumption that case law eventually extended to cases of mere tolerance or knowledge [1, 3].

Three features of this regime are particularly relevant from the perspective of AI: i) what was decisive was not the slave's “will,” but rather his functional connection to the dominus's sphere; ii) the legal system recognized that the real economy operated through these intermediaries without legal personality, and therefore established specific remedies to protect third parties contracting with them; and iii) the dominus's liability could be modulated according to the scope of the authorization or the business entrusted to the slave [1, 3].

In this regard, Professor Amunátegui observes here a clear analogy with the regime proposed by the Uniform Computer Information Transactions Act (UCITA) in the United States, which considers transactions carried out by an “electronic

agent” to be binding on the user of that agent, even if no individual specifically reviewed each transaction. The logic is analogous to that of the *actio quod iussum*: whoever uses an autonomous instrument to enter into a contract assumes the consequences of its actions [1, 3, 12].

Meanwhile, secondly, in the non-contractual sphere, Roman law recognized noxal actions, which allowed the dominus to be held liable for damages caused by his slaves or by his children under his parental authority. The owner could choose between responding with his own assets or handing over the slave through *noxae deditio*, transferring him to the injured party as a form of partial compensation. The logic of tort actions stems from the idea that the community cannot be left unprotected against harm caused by those who, even though lacking legal personality, act materially within it, that is, the risk created by keeping a slave—such as the risk of keeping a dangerous animal or an autonomous machine today—is attributed to the owner, who is the one who decides to introduce that risk factor into the public sphere [1, 3, 8, 9].

In addition, there was also the *actio de pauperie*, concerning damage caused by animals, which contemporary legal scholarship has proposed as an analogous model for liability for acts of artificial intelligence, precisely because it disregards the subjectivity of the direct causer and focuses on custody and the assumption of risk.

Another key institution is the *peculium*, whereby the dominus could entrust the slave with a set of assets to manage, constituting a separate estate against which certain creditors could bring the *actio de peculio*. This concept functioned as a proto-corporate or limited liability mechanism, insofar as it allowed the economic transactions assigned to the slave to have their own asset base, without automatically compromising the entirety of the owner’s estate. Complementarily, the *actio in rem verso* allowed the dominus to be required to return any enrichment obtained through the slave’s acts, even when these had not been expressly ordered [1, 3].

Thus, by combining *actio quod iussum*, *actio exercitoria* and *actio institoria* (when the slave was put in charge of a specific business, such as a ship or a shop), *actio de peculio*, *actio in rem verso*, and tort actions, the law of Ancient Rome constructed a flexible framework to channel both contractual and non-contractual liability arising from subjects without legal personality [1, 3].

That systemic flexibility—capable of calibrating the dominus’s liability based on the order, the entrusted business, the *peculio*, and actual enrichment—is precisely what seems to be lacking today in the regulation of artificial intelligence, where we still lack a set of organically articulated actions to deal with these new “automata” without legal personality that operate in economic and social transactions.

4. Conclusions

The analysis conducted in the preceding sections allows for several conclusions that, taken together, point toward the need

for a comprehensive rethinking of the theory of obligations in the face of artificial intelligence.

In the first place, artificial intelligence is not an ontological entity endowed with consciousness, free will, or an act of being of its own. It is, rather, an extraordinarily sophisticated system for processing information, capable of simulating the results of intentional activity without being grounded in any form of subjectivity.

In that sense, necessarily it follows that artificial agents cannot be persons in any juridically meaningful sense, and that attributing to them rights or obligations of their own, would constitute a conceptual fiction with dangerous practical consequences: the potential erosion of the conditions under which human beings can be held responsible for the systems they design, deploy, and profit from.

Secondly, this ontological clarity reveals the inadequacy of the classical categories of civil liability when applied without modification to the acts of artificial agents. The Roman tradition, transmitted through the *ius commune* and ultimately codified in the Napoleonic model, built the entire structure of obligations around the will of a person: a will capable of intending harm, of consenting to a contract, of managing the affairs of other persons, etc. Artificial agents clearly possess none of these features.

They don’t intend, consent or choose: they only compute. Consequently, forcing their acts into categories designed for conscious subjects produces, at best, strained analogies and, at worst, outcomes that are both conceptually incoherent and practically unjust.

In the third place, however, the absence of an adequate modern framework does not require us to begin from zero. Roman law itself confronted a structurally analogous problem and responded with a flexible and articulated toolkit: the *actiones adiecticiae qualitatis*, the noxal actions, the *peculium*, and the *actio in rem verso*. Although slavery was and is absolutely and unequivocally condemned, the legal architecture built around it in Roman law is susceptible of a moral inversion: by placing only genuine *res* (AI systems that truly are things) in the structural position once occupied by misclassified human beings, the law avoids both the error of the Romans and the inverse error of modernity.

For where Rome reduced persons to things, the contemporary temptation runs in the opposite direction: to elevate things to persons, thereby dissolving the qualitative distinction on which the dignity of every human being, and the coherence of the entire legal order, ultimately rest. The Roman legal experience, for all its moral contradictions, at least perceived that tension and attempted, however imperfectly, to manage it through calibrated instruments. Whether contemporary law will be capable of lucidity when confronting its own new challenges, depends on how agency, will, consciousness and, consequently, responsibility, are understood when regulating emergent complex phenomena like AI and its many potential repercussions on legal relations.

Abbreviations

AI	Artificial Intelligence
UCITA	Uniform Computer Information Transactions Act

Acknowledgments

We would like to thank Carlos Amunátegui Perelló, Doctor of Law, for his academic work as a researcher and lecturer; without him, this essay would not have been possible.

Author Contributions

Alonso Salinas Garcia: Conceptualization, Investigation, Writing – original draft, Writing – review & editing

Matias Kahn Aranguiz: Conceptualization, Investigation, Writing – original draft, Writing – review & editing

Funding

This research did not receive any funding from any public or private organization or institution.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Amunátegui Perelló C. *Arcana Technicae. El Derecho y la Inteligencia Artificial*. Valencia, España: Editorial Tirant Lo Blanch; 2021.
- [2] McCulloch, W., Pitts, W. A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*. 1943, 5, 115-133. <https://doi.org/10.1007/BF02478259>
- [3] Amunátegui Perelló C. *Arcana Futura. From Information Theory to Neural worlds*. Milton Park, South Oxfordshire, England: Routledge; 2026.
- [4] Ilkou, E., Koutraki, M. Symbolic vs Sub-symbolic AI Methods: Friends or Enemies? In *Proceedings of the CIKM 2020 Workshops*, Galway, Ireland, 2020. <https://ceur-ws.org/Vol-2699/paper06.pdf>
- [5] Sheikh, H., Prins, C., Schrijvers, E. *Artificial Intelligence: Definition and Background*. In *Mission AI: The New System Technology*, 1st ed. Cham, Switzerland: Springer; 2023, pp. 15–41. https://doi.org/10.1007/978-3-031-21448-6_2
- [6] Rodríguez Ennes, L. La "obligatio" y sus fuentes. *RIDROM. Revista Internacional de Derecho Romano*. 2009, 1(2), 90–126. <https://reunido.uniovi.es/index.php/ridrom/article/view/17922>
- [7] Smith, S. E. The United Nations Convention on the Use of Electronic Communication in International Contracts (CUECIC): Why It Should Be Adopted and How It Will Affect International E-Contracting. *SMU Science and Technology Law Review*. 2008, 11(2), 133–162. <https://scholar.smu.edu/scitech/vol11/iss2/4>
- [8] Mousourakis, G. Grounds of Delictual Liability in Classical Roman Juridical Literature. *Hiroshima Law Journal*. 2022, 46(1), 59–96. <https://doi.org/10.15027/52821>
- [9] Marcín Balsa, F. Notas para el estudio del Derecho de los contratos en el Derecho Común europeo del medievo, con especial atención a la tradición romano-castellana. *Revista Mexicana de Historia del Derecho*. 2011, 23, 191-208. <https://doi.org/10.22201/ijj.24487880e.2011.23.10115>
- [10] Razzaq, K., Shah, M. Machine Learning and Deep Learning Paradigms: From Techniques to Practical Applications and Research Frontiers. *Computers*. 2025, 14(3), Article 93. <https://doi.org/10.3390/computers14030093>
- [11] Čapek, K. *R.U.R. (Rossum's Universal Robots)*. Translated by Paul Selver and Nigel Playfair. New York, United States of America: Samuel French Edition; 1923.
- [12] Wiant, S. K. Uniform Computer Information Transactions Act (UCITA). In *Encyclopedia of library and information sciences*, Vol. 7. 2010, pp. 5328-5336.