

**THE AUTONOMOUS ROBOT : PERSON, SLAVE OR MACHINE ?**  
**THE LEGAL CHALLENGE OF CIVIL LIABILITY IN THE AGE OF ARTIFICIAL**  
**INTELLIGENCE**

**Author**

Ariane Bigenwald

Université Paris 1 Panthéon-Sorbonne

Article presented at the annual conference of the  
**International Association for Computing and Philosophy**  
**IACAP 2018**  
**Warsaw**

**eThicAa**  
ethics & autonomous agents

**ONERA**  
THE FRENCH AEROSPACE LAB

## Abstract<sup>1</sup>

From Pygmalion's myth to Frankenstein's monster, humanity has never ceased to fantasize about the possibility of giving life to its creatures. Today, Artificial Intelligence (AI) – to put it simply, “the science and engineering of intelligent machines” – ignites the collective imagination, and the spectre of science fiction once again haunts social debates and the media. Law is not immune to this recent buzz: debates proliferate while enthusiast speeches for the welcoming of a new era of robot-persons flourish. The origins of such excitement are two simple words: autonomy and unpredictability. In these conditions, the crucial legal question of liability arises. Who is responsible for a prejudicial and unpredictable decision autonomously made by the machine? It is generally believed that answering those questions calls for substantial modifications of the law and for the creation of new legal categories. However, the premises of such hypothetical procedures need to be thought out. When it comes to AI, the notion of autonomy is indeed far from universal. Moreover, autonomous systems are less unpredictable than commonly believed. Thus, as the notion of autonomy remains undefined and the concept of unpredictability matures, we ought to elucidate another question hiding behind the search for a responsible actor: according to current liability standards, can the autonomous robot not be considered as a simple machine, only more sophisticated? In other words, does the autonomous robot really revolutionize law and our current civil liability paradigms?

I will attempt to answer this question by analysing several debates on the status of autonomous robots. I will first challenge the relevance of attributing a legal personality to robots by critically analyzing European Commission studies and reports. I will then proceed to assess the possibility of reviving certain ancient roman notions such as slave status. These notions, although useful in reminding us key aspects of liability involving autonomous objects, will not be retained as solving the AI liability issue. I will finally consider possible liability systems applicable according to existing legal notions and their potential adaptation to the AI context. I will conclude that AI does not revolutionize law and, that in many legal aspects, robots are simple machines.

## Introduction

From Pygmalion's myth to Frankenstein's monster, humanity has never ceased to fantasize about the possibility of giving life to its creatures. Today, Artificial Intelligence (AI) – to put it simply, “the science and engineering of intelligent machines”<sup>2</sup> – ignites the collective imagination, and the spectre of science fiction once again haunts social debates and the media.

---

<sup>1</sup> The original version of this work is in french, and as such, was directed towards a civilist audience. Hence, it often refers to civil legal paradigms and terms, and cites European, French, and Quebequer legislation or documents. Nonetheless, Common lawyers may find it useful, given the actual popularity of AI and the lack of regulation concerning AI in both legal traditions. Moreover, this paper aims at defusing hasty enthusiasm or fears surrounding AI questions and is thus purposely kept not too technical and accessible for a non-legal public.

<sup>2</sup> John McCarthy, *Basic Questions, What is Artificial Intelligence?* (2007), Stanford U, online : <<http://www-formal.stanford.edu/jmc/whatisai.pdf>> (retrieved 1<sup>st</sup> July 2017)

Robots “commit suicide”<sup>3</sup> and “talk”<sup>4</sup> a secret and unintelligible language, while new technology titans like to play Cassandres and Pangloss on the media scene.<sup>5</sup> Law is not immune to this recent buzz: debates proliferate while enthusiast speeches for the welcoming of a new era of robot-persons flourish.<sup>6</sup>

The origins of such excitement are two simple words: autonomy and unpredictability. Some robots “gifted with” (“given” would be more appropriate) artificial intelligence can “make” (which in fact means “calculate”) decisions without any human control, and thus engage in unpredictable behaviour. In these conditions, the crucial legal question of liability arises. Who is responsible for a prejudicial and unpredictable decision autonomously made by the machine? The manufacturer, the programmer, or the user? It is generally believed that answering those questions calls for substantial modifications of the law and for the creation of new legal categories.

However, the premises of such hypothetical procedures need to be thought out. When it comes to AI, the notion of autonomy is indeed far from universal. There is a blurring of the lines between different categories of systems, such as automated systems (complex systems based on the application of rules, like autonomous vehicles, certain military weapons, or programmable thermostats), autonomous systems (systems capable of a certain degree of adaptability, learning and evolution, or goal-oriented systems), and intelligent systems (systems capable of human cognitive tasks).<sup>7</sup> Moreover, autonomous systems are less unpredictable than commonly believed. They are predictable in the majority of performance conditions. The problem remains testing their performance in all possible conditions as the complexity of autonomous systems increases. However, there is another issue called the opacity issue or “black box” problem i.e. the fact that certain “reasoning” produced by AI programs are untraceable and puzzling for the human mind.<sup>8</sup> Thus, as the notion of autonomy remains undefined and the concept of unpredictability matures, we ought to elucidate another question hiding behind the search for a responsible actor: according to current liability standards, can the autonomous robot not be considered as a simple machine,

---

<sup>3</sup> Lina Rhrissi, "L'androïde s'est-il suicidé en se jetant dans un bassin ?", *Le Monde* (2017), online : Le Monde <[http://www.lemonde.fr/big-browser/article/2017/07/19/un-robot-se-jette-dans-un-bassin-twitter-evoque-un-suicide\\_5162502\\_4832693.html#85LThfFAAsjLJvR2.99](http://www.lemonde.fr/big-browser/article/2017/07/19/un-robot-se-jette-dans-un-bassin-twitter-evoque-un-suicide_5162502_4832693.html#85LThfFAAsjLJvR2.99)> (retrieved on 19 July 2017).

<sup>4</sup> Morgane Tual, "Non, Facebook n'a pas « paniqué » à cause d'un programme d'IA capable d'inventer un langage", *Le Monde* (2017), online : Le Monde <[http://www.lemonde.fr/pixels/article/2017/08/01/non-facebook-n-a-pas-panique-a-cause-d-un-programme-d-ia-capable-d-inventer-un-langage\\_5167480\\_4408996.html#8H0R91YGZypZYpk5.99](http://www.lemonde.fr/pixels/article/2017/08/01/non-facebook-n-a-pas-panique-a-cause-d-un-programme-d-ia-capable-d-inventer-un-langage_5167480_4408996.html#8H0R91YGZypZYpk5.99)> (retrieved on 1<sup>st</sup> August 2017).

<sup>5</sup> Elon Musk, in the role of Cassandre, and Mark Zuckerberg, in the role of Pangloss. The first often makes statements in the media about “evil AI” and the danger they represent, while the second applauds AI’s progress. See : Morgane Tual, "Au-delà des fantasmes, quels sont les problèmes concrets que pose l’intelligence artificielle ?", *Le Monde* (2017), online : Le Monde <[http://www.lemonde.fr/pixels/article/2017/08/03/au-dela-des-fantasmes-quels-sont-les-problemes-concrets-que-pose-l-intelligence-artificielle\\_5168330\\_4408996.html#wY115Sd1AujCi9Y u.99](http://www.lemonde.fr/pixels/article/2017/08/03/au-dela-des-fantasmes-quels-sont-les-problemes-concrets-que-pose-l-intelligence-artificielle_5168330_4408996.html#wY115Sd1AujCi9Y u.99)> (retrieved on 3 August 2017).

<sup>6</sup> I think, for example, about Alain Bensoussan, a Parisian lawyer specialized in new technologies, who in many conferences fiercely defends robotic personality.

<sup>7</sup> Paul Scharre, *Autonomous Weapons and Operational Risk*, Center for New American Security, 2016.

<sup>8</sup> I think about cases of visual recognition of images using neural networks, where the AI considered its results 99,6% reliable while the results were not only incorrect but also undecipherable for human observers. The AI labelled as “starfish” or “penguin” what could have at best been a geometrical figure, and as “leopard” or “peacock” a cloud of pixels. See: Anh Nguyen, Jason Yosinski and Jeff Clune, *Deep Neural Networks are Easily Fooled : High Confidence Predictions for Unrecognizable Images*, *Computer Vision and Pattern Recognition CVPR*, IEEE, 2015.

only more sophisticated? In other words, does the autonomous robot really revolutionize law and our current civil liability paradigms?

I will attempt to answer this question by analysing several debates on the status of autonomous robots. I will first evaluate the relevance of attributing a legal personality to robots, and then proceed to assess the possibility of reviving certain ancient notions such as slave status. I will finally consider the possible liability systems applicable according to existing legal notions and their potential adaptation to the AI context.<sup>9</sup>

## **I. The robot, a person?**

Of the many legal debates sparked by the growing interest for autonomous robots, the debate surrounding the legal personality of robots is probably the most polemical. Robotic personality<sup>10</sup> – a concept that combines adaptations of existing rules with ideas that are entirely novel to law – both worries and mesmerizes the collective imagination. For some, the “person” status, often confused with an ontological status, evokes at best an hypothesis that is “as useless as incongruous”<sup>11</sup> and at worst, the coming to life of science fiction’s worst nightmares. For others, robotic personality is the next legal revolution, welcomed with somewhat hasty enthusiasm.

More nuanced attitudes should be adopted towards this proposition, especially considering that civil law is governed by practical concerns. As a matter of fact, the answer to the question “who is responsible?” has a direct impact on victims, whose compensation will be a function of the liability scheme that is ultimately chosen. It also affects other actors from the world of robotics, which will be more or less incentivized to develop new technology depending on the extent of the liability weighing on them. Reflections on robotic personality ought therefore to be mindful of the pragmatic concern of opting for the best liability regime rather than be concerned about contrived fears and fantasies.

On another note, I would like to highlight that the debate on robotic personality differs from the one on animal status. The latter is rooted in the sentient and singular experience of animals, and not on their autonomy, as well as on a refuting of the specist exercise of law.<sup>12</sup> It seeks, on the one hand, to recognize animals as beings with rich emotional lives and capable of preferences and interests, and on the other hand, to acknowledge law as being arbitrarily exclusive to human kind (due to the absence of any rational distinctive relevant feature between animals and humans).<sup>13</sup>

---

<sup>9</sup> In order to increase the readability of this paper, I will use the colloquial term of “robot” to designate autonomous artificial agents, whether they have a physical body or not.

<sup>10</sup> One can also talk about “digital personality”, “robot-person” etc.

<sup>11</sup> Nathalie Nevejans, *Règles européennes de droit civil en robotique (study for the JURI committee)*, Policy department of the European parliament, 2016. (All cited expressions from this document are personally translated).

<sup>12</sup> I can think about authors such as Peter Singer and Paola Cavalieri.

<sup>13</sup> These authors highlight the fact that any attempt at defining humankind as a specie is doomed to fail because there is no natural characteristic exclusive to human beings (as it can always be found in other species) and because such definition will always leave aside non-paradigmatic cases (elderly people, children etc.). Given this, the person status cannot be ontological and so can extend to animals. See: Peter Singer, *Animal liberation*, Wellingborough, Northamptonshire, Thornsons, 1986.

In contrast, the legal personality of the robot is not defended because of machines' feelings, their deep souls or consciousness – characteristics which we still have no scientific basis for attributing to machines. It only serves the functional objective of attributing liability. It would therefore be affiliated with the moral entity or legal person status.<sup>14</sup>

The legal person is a legal fiction that allows an entity with no physical or corporal existence to pose the rights and obligations associated with legal personality. Corporations or partnerships are examples of legal entities or persons. As a legal person, a company has a personality that is distinct from that of its shareholders, directors and employees. It has the right to own property, the right to open a bank account, to sign contracts, to take legal action to protect its interests, to claim compensation for any damages or loss, including moral damages (defamation, loss of reputation), etc. Its liability is usually<sup>15</sup> separated from that of its directors in a manner that shields their assets from the corporation's creditors.

Robotic personality would be an adaptation of legal personality. As a legal person, the robot would, like any company, have its own registration, rights and obligations, as well as legal representatives and statutes of incorporation to distribute roles and functions between those representatives. We would no longer talk about administrators or directors, but rather about programmers, users, trainers, etc.

If the idea needs better definitions and more thorough articulations of the notions it involves, the opposition to the attribution of such a status is not any more convincing. A study written by N. Nevejans<sup>16</sup>, intended for the European parliament, displays some of the inconsistencies and contradictions in its critique of the robotic personality. The author legitimately presents the hypothesis within the framework of legal personality, and its adaptation, and reaches two possibilities: “either humans are the real legal actors behind robots, or robots are their own legal actors”.<sup>17</sup>

In the first case, “such a useless result would prove the incongruity of attributing legal personality” to robots. This assertion seems dubious as it defies the concept of legal personality. The prime utility of legal personality is indeed to allow the legal existence of an entity distinct from the humans' piloting it. A company is as much a legal actor as is its representative. Their personalities are not mutually exclusive nor do they cancel out each other's relevance. Furthermore, contrary to what the author asserts, a moral entity without a human commanding it is not “an empty shell”.<sup>18</sup> An “empty shell” in law refers not to a moral entity without representatives, but to a moral entity without activities.<sup>19</sup> Aside from this technical oversight, it

---

<sup>14</sup> As opposed to the status of natural person.

<sup>15</sup> In the case, for example, of what is called “lifting the corporate veil”: when an administrator's personal liability is engaged because of his or her fraudulent acts. Art. 317, Civil Code of Quebec, RLRQ c CCQ-1991, retrieved on 01/08/2017 (hereafter CcQ).

<sup>16</sup> Nevejans, *supra* note 11.

<sup>17</sup> *Ibid* at p. 17.

<sup>18</sup> *Ibid*.

<sup>19</sup> The original version uses the terms “coquille vide”, which in french, have a more specific meaning. In english, their definition is looser and does not only refer to a moral entity without activities. Nonetheless, that expression is better understood this way, since we may think of examples of moral entities that function “by themselves” i.e. trusts (*fiducies*).

should be noted that what makes an entity more or less “empty” is not the absence of human actors behind it, but its lack of patrimony.

In the second case, the robot would be its own legal actor. Whereas the author started her analysis by stipulating “legal personality is not based on any considerations given to the robot’s depths or feelings”<sup>20</sup>, she grossly confuses the notions of legal actor and natural person and goes on asking the contradictory question: “how could a simple machine, an empty shell void of any conscience, feelings or will power, become its own legal actor?”<sup>21</sup> The answer is in the initial hypothesis: by being a legal person.

The contradiction is also apparent in another argument: recognizing robots as legal persons would be “an error risking to lower humans to the rank of machines in the long term”.<sup>22</sup> This statement is at odds with what the study first claimed to disavow i.e. unfounded fears and fantasies. Here, it clearly gives way to an unfounded science fiction scenario. Endowing corporations with legal personality did not reduce humans to a pure legal abstraction.

Finally, the last objection to the attribution of legal responsibility concerns the recognition of rights and duties. The author asks herself “how could a robot have duties, when these are inextricably linked to human moral representations? ... Which rights should be given to robots: the right to live, the right to dignity, to equality with human, the right to retirement...?”<sup>23</sup> It is important to keep a calculated perspective when considering what the juridical status of robots should be. Again, there is no practical obstacle to treating robots as legal persons, conceptual or otherwise. Just as a legal person has rights (the right to reputation, the right to take legal action, etc.) and obligations (legal term that I prefer to “duties”), a robot given a specific juridical status could have rights and obligations. Contrary to what the author asserts (yet indeed in the farfetched hypothesis where robots, after having replaced human labour, ask for vacations), the attribution of rights and obligations to robots would not “destroy the emerging robotics market”.<sup>24</sup> Quite the contrary, robotic personality could favour the latter’s emergence by facilitating the management of litigation and liability allocation. The question we ought to be focused on is that of the practical advantages provided by the concept of robotic personality, as compared to other possible solutions. On this issue, I acknowledge the author’s pragmatism when she highlights the potential existence of other liability regimes that are “much more efficient regarding victim compensation”.<sup>25</sup>

Paradoxically, robotic personality’s proponents tend to be far more reasonable and mindful of protecting the rights of humans than their opponents generally give them credit for. The *Charter for Robots Rights*<sup>26</sup>, suggested by Alain Bensoussan, a Parisian lawyer specialized in new technologies, is essentially made up of provisions protecting human interests: robot rights are

---

<sup>20</sup> *Ibid* at p. 16.

<sup>21</sup> *Ibid* at p. 17.

<sup>22</sup> *Ibid* at p. 18.

<sup>23</sup> *Ibid*.

<sup>24</sup> *Ibid*.

<sup>25</sup> *Ibid* at p. 17.

<sup>26</sup> Alain Bensoussan, *Charte des droits des robots*, 5th ed (Paris: Lexing, 2015), online: <https://www.alain-bensoussan.com/wp-content/uploads/2014/10/Charte-droits-des-robots-Version-5.pdf> (retrieved on the August 1st, 2017). Note that this charter seeks to serve as draft for future official charters, and thus is incomplete.

duties that humans owe one another via robots, rather than duties that humans owe to robots. “Robot dignity” (art. 3), for instance, in fact concerns the respect of stored personal data as protected by the French law on computing and freedom.<sup>27</sup> Some provisions concern programmers and robot operators (principles of manufacturing and traceability), while others concern users (for instance, the principle of forbidden use for illicit ends). Furthermore, only humans are liable for robots’ acts.

Having demystified the concepts and defused the fears surrounding robotic personality, let us now evaluate the actual impact of robotic personality, as well as its relevance. Let’s take a simple legal situation, to wit: a contract. In the hypothesis of a recognized “robot-person”, the latter would be able to sign a contract with a person. If that person ends up unsatisfied with the transaction, she would have to, as per the *Charter of Robots Rights*, complain to the robot’s legal representatives, that is, in order of sequence, the user, the manufacturer of the AI programs, the manufacturer of the robot’s physical components, and the owner.<sup>28</sup> Conversely, if robots are not attributed any legal personality, the roles played in a given transaction between the robot and the user could be interchangeable or merged. The robot entering a contractual agreement (for example, a domestic help-robot buying groceries) would be a representative of the human being mandating it. An unsatisfied third party would have to, following the rules of mandates,<sup>29</sup> sue the mandant,<sup>30</sup> who could then in turn, in accordance with the law of product liability, sue the manufacturer for a potential defect of the robot, and so on.

The question is: what major difference does robotic personality entail? In both cases, responsible actors are exclusively persons recognized by law (legal or physical) and, in both cases, they will invoke the same legal concepts in support of their claims, regardless of whether those concepts are old or new and of whether robots are formally recognized as legal persons. The answer thus seems to be: none.

The example might be simplistic, but it nonetheless reveals that the legal issues at hand concern more the details of legal provisions than any fundamental distinction between robot-person and non robot-person. It is imperative, among other things, to define the “autonomy” of a robot in order to distinguish it, if necessary, from other machines; to establish the type of action attributable to the user, or the programmer, etc.; to decide whether or not a legal presumption of liability is useful, and if such is the case, whom should it be directed toward? Should the presumption be irrefutable? Moreover, the debate on robotic personality should not lose sight of the reality of the robotic world i.e. a market of robotic products, in which it is imperative that consumers have effective tools and remedies against tech titans.

In this regard, it is worth noticing that this reality diverges from that of legal personality. Legal personality first appeared in the Middle Age, as a theological notion in debates surrounding the

---

<sup>27</sup> Law n°78-17 from January 6, 1978, *Loi relative à l’informatique, aux fichiers et aux libertés*, 1978, Légifrance. Also called “*Loi Informatiques et Libertés*”.

<sup>28</sup> No provision in the Charter specifies how to pass from the user’s liability to the owner’s.

<sup>29</sup> In Common Law jurisdictions, what the Civil Code refers to as “mandates” is called the law of “agency”: a principal instructs an agent, who is authorized to enter into legally binding agreements on the principal’s behalf.

<sup>30</sup> Could mandataries or legal representatives not be persons? In the current state of law, a mandatary is a person (art. 1984 Code Civil (CC), France, 1804, Légifrance). Legal adjustments would be necessary. They could potentially be inspired by the slave status. See section: The robot, a slave?.

nature of the Christ,<sup>31</sup> then, under the casuistic pulse of civilists and canonists facing legal issues,<sup>32</sup> as the status of a community called “*persona ficta*” or “*persona representata*” (in opposition to the human “*persona vera*”). These jurists questioned, *inter alia*, the possibility of a *persona ficta* being constituted of only one single person or of none<sup>33</sup> (e.g. in the eventuality of perpetuating a brotherhood’s existence beyond the death of its last monk). If comparatively, this phenomenon seems to happen between the physical robot and the digital individual<sup>34</sup> (the algorithm and the collected data form a digital individuality can survive the physical destruction of the robot), similarities end there. From the XIIIth century onward, medieval jurists reflected on the application of the regime of *persona* in the absence of any qualification (*ficta*, *vera* etc.). According to a broad interpretation, it was permissible to consider a community a *persona* and to endow it with rights normally reserved to a *persona vera* pursuant to the principle of identity of reason.<sup>35</sup> Similarly today, a legal person possesses a “reason”. A legal person, such as a corporation, is an intersection of different actors’ personal and common interests, from the employee’s to the stockholder’s, crystallized in the superior and independent interest of the company. This logic is difficult to apply to the robot. As an item of consumption, it would embody opposing interests. Attributing a personality to the robot would make the manufacturer, the seller and the client cohabit a single legal entity. Finally, the practical utility of legal personality lying in its independently-owned, and often far more substantial assets, it is either unrealistic to imagine a robot-person having some, or outlandish to imagine one without them.

I humbly think that this debate deserves to be clarified and should not deviate precious attention and resources towards formalistic issues. It is more important to establish tangible liability rules independently of robotic personality, before deciding on the appropriate nomenclature of these rules. Meanwhile, indecision – if not incoherence – persists, as demonstrated by the European resolution of January 2017, which, from on paragraph to the next, first attributes liability exclusively to humans then conceives of the possibility of robot liability along with robotic personality.<sup>36</sup>

## I. The robot, a slave?

---

<sup>31</sup> Éric Schneider, “ L’interprétation des termes d’homme ou de personne en droit savant, source d’inégalité ”, *Droit et cultures*, vol 69, 2015, p. 89-112, citing the Franciscan theologian Alexandre of Hales : Magistri Alexandri De Hales, *Glossa in quatuor Libros Sententiarum Petri Lombardi*, Lib. III, Dist. VI (AE), 18, (Florentiae collegium-Bonaventurae, éd. Quaracchi 1954), p. 80.

<sup>32</sup> *Idid.* citing : Yann Thomas, “Un expédient interprétatif à l’origine de la personne morale”, *L’architecture du droit. Mélanges en l’honneur de Michel Troper*, (2006) Economica, p. 951-960 ; P. Michaud-Quantin, *Expressions du mouvement communautaire dans le Moyen-Age latin*, (Paris: Librairie philosophique J. Vrin, 1970).

<sup>33</sup> *Idid.*

<sup>34</sup> Expression coined by Alexei Grinbaum, “Responsabilité des êtres calculants”, (2017) *Revue française d’éthique appliquée*, n° 3, p. 119 : « First, the data collected by each digital “thing” permits it to become unique and to distinguish itself from other digital things. Second, the digital thing gets individualized by its code structure.” (Personal translation).

<sup>35</sup> Schneider, *supra* note 31.

<sup>36</sup> Mady Delvaux, *Report with recommendations to the Commission on Civil Law Rules on Robotics*, Committee of Legal Affairs, European Parliament, 2015/2103(INL), January 2017, para. 56 *in fine* and para. 59f). Moreover, the general formulation of para. 59f), in attributing legal personality and liability to robots “where robots make autonomous decisions or otherwise interact with third parties independently” is too vague and embraces a too vast domain.

A robot has two modes of existence: it is a thing, a manufactured object, and a digital individual, a source-code which singularly evolves through the data it collects and uses. It is also somewhat autonomous and fulfills objectives and courses of action it has previously chosen among a number of possibilities. The ambiguity emerging from this autonomy and individuality combined with an object status is not without overtones of the status historically borne by slaves. The slave, both a good that a master could dispose of and a man acting on his own will, could take actions with legally binding effects. He enters into contracts on his master's behalf, was involved in business, accumulated savings, etc.<sup>37</sup> This situation, added to a well-known etymological parallel ("robota" in Czech can be translated to "slave" or "serfdom")<sup>38</sup>, explains the paradox brought about by a new individual capable of acting by itself and which is neither a legal subject nor liable for its actions. It mostly forces us to compare: could a robot benefit from a similar slave status? What legal notions should then be resurrected? As an answer to these questions, it has been suggested that we adapt the ancient Roman notion of noxality.<sup>39</sup>

Noxality dates back to the Roman Law of the XII Tables. Variations of it can be found in other civilizations (in Athenian Gortyn laws and in Germanic laws), subject to a few trivial differences.<sup>40</sup> *Noxa* or *noxia* designates an offence committed by a "person in potential" (a son, a slave), an animal or even an inanimate object.<sup>41</sup> The noxal responsibility is the liability of the *paterfamilias* (head of the family or owner), arising from a *noxia* committed by someone or something placed under the *paterfamilias'* *potestas* (that is, members of his family, his slaves, his animals, and his goods). More than a delegated responsibility to a paternal authority, noxality has a singularity: the possibility of noxal abandonment. It is said "noxal" any legal action that grants this alternative *aut in noxam dedere*. In conformity with the two principles "*noxiam sarcire*" and "*noxae dare, in noxam dedere*", the *pater* can free himself from a noxal action by either financially compensating the victims or handing over the culprit to them. "Noxal abandonment is an act by which the defendant places the offender under the claimant's authority".<sup>42</sup> In this way, the victims control and can choose the remedy (or absence of thereof) they prefer. Noxal responsibility is therefore a form of strict liability that arises only by virtue of being a *pater*.<sup>43</sup> This regime gained in complexity from the II to the VI century in response to more sophisticated contractual practices such as slave lending, which temporarily involved two possibly liable *pater*, the owner and the borrower. Roman jurists thus introduced two alternative regimes linked to the knowledge of the reproached acts. According to the principle of *scientia domini* as contained in the *Lex Aquilia*, if a *pater* has ordered certain action, or was aware that one of the individuals under his tutelage was committing them, he became personally liable for those actions. If on the contrary, he was unaware of the action (*insciente domino*), a noxal action could arise. Hence, authority and knowledge are the foundations of noxality.

---

<sup>37</sup> Jean Christian Dumont, *Servus. Rome et l'esclavage sous la République*, (Publications de l'École française de Rome, vol 103, 1987).

<sup>38</sup> Term coined by Karel Capek, *R.U.R (Rossum's Universal Robots) (1921)*, (Maryland: Wildside Press, 2012).

<sup>39</sup> Grinbaum, *supra* note 34. Grinbaum works from: Fernand de Visscher, *Le régime de la noxalité. De la vengeance collective à la responsabilité individuelle*, (Brussels: A. de Visscher, 1947).

<sup>40</sup> Paul Fauconnet, *La responsabilité. Étude de sociologie*, (Montreal: digital ed. Les classiques des sciences sociales, p. 78) (from original 2<sup>nd</sup> ed (Paris: Librairie Félix Alcan, 1928).

<sup>41</sup> *Ibid*, p.80.

<sup>42</sup> *Ibid*, p.77. All information contained in the paragraph comes from the same book.

<sup>43</sup> Grinbaum, *supra* note 34.

Based on this notion and these principles, Grinbaum, a French physicist and philosopher, lays out the foundations of a potential liability regime applicable to robots and AIs:

1) The source-code is the digital individual's *pater* and his noxal responsibility can be engaged; 2) the user possesses, in the sense of property rights, the digital object and is liable for any harm the object may cause; 3) through data, the digital individual is acquainted with the user (has a link of knowledge). This aspect necessarily leads to a form of noxal responsibility for the digital individual. Since the latter is not a legal subject, this responsibility is transferred to the source-code, which in turn is immanent to the programmer. Thus, collecting data gives a noxal responsibility to the programmer.<sup>44</sup>

The idea is that the programmer could abandon the digital individual i.e. a possibility which already concretely happens, according to the author, via the phenomenon of Open Source. Open Source refers to a type of licence for digital programs that gives free rights to redistribution, access the source-code and creation of derived products.<sup>45</sup>

I will focus on the possibility of reviving the notion of noxality rather than the detailed modalities suggested by Grinbaum, which are fettered by too many technical mistakes to be appreciated as they are.<sup>46</sup>

If the proposition seems interesting at first sight in that it highlights and mixes the characteristics of “paternity” and knowledge, it nonetheless raises a few doubts. First concerning the origins and foundations of noxality: indeed, depending on the interpretation and the civilization, noxality can be read as either the organization of personal vengeance or as a form of liability relief for the *paterfamilias*.<sup>47</sup> Noxality remains a primitive private retribution system, giving the victim the choice of sanction, and in which the right of redemption has not always existed.<sup>48</sup> Indeed, the culprit's abandonment, when left to the discretion of the defendant, is no more than a form of atonement. It represents the chance to disengage from collective responsibility, to deviate the ire of men and the wrath of gods.<sup>49</sup> In this perspective, noxality loses its appeal when applied to robots. Our current legal system rests on the victim's compensation in accordance with actual losses rather than on any idea of vengeance.<sup>50</sup>

---

<sup>44</sup> *Ibid*, p. 123.

<sup>45</sup> Online: [https://en.wikipedia.org/wiki/Open\\_source\\_\(disambiguation\)](https://en.wikipedia.org/wiki/Open_source_(disambiguation)) (retrieved 1<sup>st</sup> August 2017)

<sup>46</sup> “The user possesses, in the sense of property rights”: there is a confusion between the user and the owner, who may not be the same individuals, and between possession and property, which are not legal equivalents; “material liability” and “digital object”: vague terms. A material liability for the physical parts of the robot? If the owner breaks himself the robot, there is no possible legal action. If the robot breaks by itself due to a material default and is still under warranty, the manufacturer could be responsible. The situation of the user not being the owner gives rise to another kind liability situation. Etc. Number 3) is nonsensical: the premise is confusedly expressed: it would mean that the digital individual is responsible of the user's acts just by virtue of having knowledge and information on him and his whereabouts. Finally, the conclusion, that is, the programmer's noxal responsibility concerning the data and hence, the possibility of abandoning it, is inconsistent with the consensual legislative objective concerning the protection of personal data.

<sup>47</sup> Fauconnet, *supra* note 40, p. 78.

<sup>48</sup> *Ibid*. It is the case, for example, when the choice belongs to the plaintiff, as is prescribed in Gortyn Law. *Ibid*, p.80: “ That this sanction may be a vendetta is undoubtedly true [...] that this vendetta is exercised is observable.”(Personal translation). See also: Roger Henrion, “compte-rendu de De Visscher”, *L'Antiquité classique*, (1949) 18: 1, p. 112-114.

<sup>49</sup> Roger Dekkers, “compte-rendu de De Fisscher”, *Revue Belge de philologie et d'histoire*, (1948) 26:1, p. 263-266.

<sup>50</sup> At least in civil litigation.

In the same vein, noxality's inadequacy to contemporary times reveals itself in another aspect: noxality would imply the *noxia*'s responsibility, that is, the responsibility of animals and inanimate objects.<sup>51</sup> Fauconnet indeed considers this responsibility to be implicit in the fact that the abandonment does not, or may only coincidentally correspond to the value of losses: "A cow for a clump of grass? [...] an animal, which can be worth close to nothing, offered as compensation for a damage, as big as that damage be?" He also finds the rule *noxia caput sequitur* (the action follows the head) to be only founded in the *noxia*'s responsibility. The sanction follows the object passing through successive owners' hands; it forces the abandonment on the current owner and not on the owner at the time of alleged offences; it thus targets the object and not the *pater*. This hypothesis finds confirmation in its consequences: "if, since the offence, the victim has become the owner of the guilty animal, the legal action is extinguished".<sup>52</sup> Although the responsibility of animals or objects, far from being primitive, first emerged in great civilizations,<sup>53</sup> would it be conceivable in today's legal landscape? I humbly leave society answer to this question.

However, a second, more practical interpretation of noxality remains: the partial de-responsibilization or disempowerment of the *paterfamilias*. Grinbaum actually favours that approach.<sup>54</sup> Is that hypothesis, i.e. the noxal abandon by the programmer, even realistic considering the millions-billed-costs of developing AI programs and so the improbability of abandonment of a source-code?<sup>55</sup> What is the utility of a legal option that is not being used?<sup>56</sup> Moreover, is it desirable? One could argue that the de-responsibilization of developers-programmers would benefit the emerging market of robotics. Another legal tool, such as a presumption of liability for the user could, however, have the same effect. From the victim's perspective, it becomes more obvious that the noxal abandonment would complicate things rather than ensure just compensation. What would a layman, not especially versed in digital technologies, do with a potentially defective source-code? How to measure the value of such code in link with actual damages? Giving the choice of abandon to the defender is to encumber a greater and useless burden upon the victim. Bestowing it on the plaintiff is to give rise to a possible business war between developers who could relatively cheaply obtain the competitor's results.

I previously stressed the interesting notions noxality was bringing to the fore i.e. authority, or power, and knowledge, or ignorance, of acts committed by those under our tutelage. I can finally question the relevance of noxality in regards to these two aspects by comparing it to existing provisions.

---

<sup>51</sup> *Ibid*; Fauconnet, *supra* note 40, p. 79.

<sup>52</sup> *Ibid*, p. 79.

<sup>53</sup> *Ibid*, p. 81: "The criminal liability of animals is not a "primitive" phenomenon that would disappear with civilization. Quite the opposite is true. We find it the three societies from which ours comes from, in Israel, in Greece, in Rome, and it reaches its apotheosis during the Christian Europe".

<sup>54</sup> Grinbaum, *supra* note 34, p. 123.

<sup>55</sup> One can think about Apple who refused to give the Iphone security program source-code to the FBI in March 2016.

<sup>56</sup> The answer "to make companies or corporations pay" is not a suitable answer since the obligation of compensating already exists.

Let's first study liability for the actions of animals. According to that regime, the animal's keeper is liable, similarly to a case of noxal liability, without any negligence or fault on his part.<sup>57</sup> The authority bond proceeds from the same logic in both cases. In one case however, if a strict parallel were applied, liability would lie with the user, whereas noxality would attribute it to the source-code programmer. Yet, nothing prevents us from adapting and modifying the provisions concerning the liability for actions of animals in order to include the programmer if need be.

In regards to *scientia*, no need is to return to Antiquity, although a comparison with the slave status might still be useful. Take, for instance, the provisions in the *Code Noir* ("Black Code") concerning contractual matters. Dating back to the XVIII century, the *Code Noir* designates all the edicts and the legal texts compilations applying to French colonies and to slave-master relations.<sup>58</sup> It recognized the right of the slave, despite the latter not being a legal subject, to enter into contracts on behalf of his master via special mandate. The master was held responsible for any obligations contracted in accordance with his instructions (similarly to the principle of *scientia domini*), and for any other bypassing the imposed limits only up to the value of profits, if any, the master gained from it.<sup>59</sup> This particular mandate rule can still be found in some civilist societies.<sup>60</sup> We could envisage, through the adaptation of such rules, considering knowledge of the robot's acts in evaluating the user's (for example) liability.

If noxality informs us of ancient modes of liability and methods of compensation, and reminds us of useful notions such as *potestas* and *scientia*, its direct adaptation seems unrealistic. Minimal as it might be, it would need to at least incorporate the possibility of noxal abandon, without which there would not be any noxality, and which, as shown before, is incompatible with our current legal mentality and doubtfully beneficial. The comparison at the basis of this legal hypothesis is, however, legitimate given the striking resemblance between the slave and the robot.<sup>61</sup> Located in ontological antipodes, both share ambiguity between statuses as things and persons. Further research and reflection remain to be explored down this road (in this direction?).

## I. The robot, a simple machine?

---

<sup>57</sup> Tort Law, the Common Law of civil liability, is very idiosyncratic and different from the Civil Law of civil liability. Common Law doesn't rely on the notions of fault, damage and causality. Instead, it looks at various forms of "wrongs" called Torts, each requiring the existence of a duty of care, an accompanying standard of care and a breach of that standard of care. Negligence is probably the most commonly alleged Tort.

<sup>58</sup> Erwann Le Guern, *Une chose baptisée? Le statut juridique de l'esclave des colonies françaises sous l'Ancien Régime*, French DEA Thesis (Rennes: University of Rennes I, 2003) online: <http://partages.univ-rennes1.fr/files/partages/Recherche/Recherche%20Droit/Laboratoires/CHD/Theses/LeGuern.pdf>

<sup>59</sup> *Ibid.*

<sup>60</sup> In Quebec, Canada: Art. 2153 CcQ.

<sup>61</sup> The comparison sometimes bothers, probably because of persistent issues and embarrassment linked to slavery past, but also surely because of a futurist type of discourse in AI spheres i.e. talks about sentient and sapient robots and their right to dignity, non-discrimination etc. See: Nick Bostrom & Eliezer Yudkowsky, "The Ethics of Artificial Intelligence" (2011) *Cambridge Handbook of Artificial Intelligence*, Cambridge University Press ; Hutan Ashrafian, "Artificial Intelligence and Robot Responsibilities : Innovating Beyond Rights" (2015) 21:2 *Science and Engineering Ethics*, 317 -326

Some institutions, including the European parliament, but also CERNA and very new legislation on the subject<sup>62</sup>, have given rise to reflections on the adaptation of existing law to autonomous robots. I suggest examining some of these.

It is first worth noting that, in the absence of specific rules for autonomous robots, some provisions of certain areas of law can readily be applied. At the present time, a robot is considered an object and damages resulting from its actions are always attributable to a human. The liable person can be a user, a developer, a trainer (for learning machines), a manufacturer, an operator, etc. That person is considered responsible only in cases where the cause of the robot's actions or inaction can be attributed to him or her and where he or she "could have predicted and thus avoided the harmful behaviour of the robot".<sup>63</sup> Furthermore, the "current legal framework regarding vicarious liability for products, which holds manufacturers responsible for the malfunctioning of a product, and the rules setting out liability for harmful actions, which hold the user of a product responsible for any harmful use of the product, apply to damages<sup>64</sup> caused by a robot or an AI".<sup>65</sup> Moreover, European Council directive 85/374/CEE deals with manufacturing defects of robots, but requires victims to present proof of the alleged defect, in addition to proving a quantifiable loss and causal nexus.<sup>66</sup> There remains a legal void for all contracts a robot might enter into.

Finally, there is existing legislation which already appears well suited to dealing with situations broadly considered to be problematic. In particular, the "diabetic patient and nurse-robot" dilemma comes to mind<sup>67</sup>. In this situation, "a diabetic patient is monitored by an autonomous robot that remotely reports the patient's feeding behaviour and health to a physician, who can then advise the patient. Let us suppose that the patient wants to eat some sweets for once, and shares this desire with the robot. How would the robot balance the patient's desire and the physician's objective?".<sup>68</sup> This example serves to illustrate the dilemma between the right to privacy and freedom of choice on the one hand, and the potential danger incurred by not disclaiming the information on the other. In the eyes of law however this situation is not problematic. A patient, although he or she previously accepted treatment, can always stop following it and revoke his or her consent to undertake treatment.<sup>69</sup> The robot could be regarded as a part of a treatment, and the patient would preserve the right to shut down the robot pursuant to existing law.

---

<sup>62</sup> Delvaux, *supra* note 36; Commission de réflexion sur l'Éthique de la Recherche en Sciences et Technologies du Numérique d'Allistène (CERNA), *Ethique de la recherche en apprentissage machine* (2017); *Nev. Rev. Stat NRS*, 2013, Ch. 482A.

<sup>63</sup> Delvaux, *supra* note 36, resolution para. AD.

<sup>64</sup> "Damages" in Common Law Torts refers to the monetary award order by a court to be paid by a tortfeasor to a victim, while in Civil it broadly refers to losses or harm suffered by a victim. I use the term in its civil law acception.

<sup>65</sup> *Ibid*, para. AE. In the United States, lawsuits involving semi-autonomous robots, such as the surgical robot Da Vinci, have based themselves on product liability rules. See for example: *Taylor v. Intuitive Surgical Inc.* (2017), SCW N 92210, online: [law.justia.com/cases/washington/supreme-court/2017/92210-1.html](http://law.justia.com/cases/washington/supreme-court/2017/92210-1.html).

<sup>66</sup> *Directive on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products*, 85/374/EEC L210(1985), European Council.

<sup>67</sup> Example given in: EthicAA team, "Dealing with ethical conflicts in autonomous agents and multi-agents systems" (2015) AAAI Workshop on AI and Ethics, Austin, Texas.

<sup>68</sup> *Ibid*.

<sup>69</sup> Art. 16 – 3 CC; See also: *Fiche 5 – Le consentement aux soins*, (2015) France Assos Santé, online: <http://www.leciss.org/sites/default/files/Consentement-aux-soins.pdf>

If current rules nevertheless still seem far from sufficient, this is due, on one hand, to the opacity of autonomous systems, which inhibit a clear distinction and division of roles and liabilities in litigious situations, and on the other hand, to the fact that the autonomy of robots leads to a certain unpredictability of their behaviour. Indeed, opacity, by which I mean the difficulty – perhaps the impossibility in certain cases – of knowing the algorithmic origins of a decision taken by the robot, is a first obstacle to the reconstruction of a chain of liability. Paradoxically, opacity also allows us to sketch a rough liability framework. Opacity depends on the human level of understanding of the machine: a robot that malfunctions can be a real mystery for the user, whereas a programmer might just regard it as a capricious machine. Hence, the level of understanding of the machine can lay the groundwork for a different liability status. CERNA analyses this relativity: the source-code developer, as the one having the most knowledge on his or her creation (or the “paternity” over it), has a certain power and thus the liability that goes with it. That liability however, is limited by the lack of knowledge and control over the data fed to the machine throughout its learning process, and thus ends the moment the code is put into action. The trainer, in turn, is liable for the data given and any latent biases within that data. He can exempt himself by invoking his lack of knowledge regarding the information processing system. Liability could then fall back on the developer who, in order to avoid being held liable, would need, to “create control mechanisms, to set out restrictions on and conditions of use, including the appropriate nature of data to be learned by the system”.<sup>70</sup> Finally, the limited knowledge that a user or an owner has of the internal functioning of a intelligent system would be the basis for a limited form of liability. The European parliament has suggested a similar framework, though introducing a particularity: a form of liability that would be proportional to the learning period.<sup>71</sup>

Autonomy, and the unpredictability linked to it, have, in turn generated a multitude of suggestions, from robotic personality to, more pragmatically, liability for acts of animals, strict liability or liability for risk.<sup>72</sup> All those liability regimes involve faultless/strict liability, which is to say that they do not require any negligence or faulty behaviour on the part of the offending party, such as failing to monitor or watch over something. They arise from the concomitance of three facts: the act of a thing, damages resulting from that act, and causal link between the act and the damages incurred. They all, thus, establish a sort of assumption against one actor in particular. For example, if a dog is under my custody and, without any failing on my part, it bites a neighbour, I am presumed liable for any resulting damage. I will not be exonerated unless I invoke another cause for the damage (superior force, third party, etc.).<sup>73</sup> If an animal’s behaviour is no less unpredictable than that of a robot, nothing should prohibit an injured from benefiting from the same remedy.<sup>74</sup>

---

<sup>70</sup> CERNA, *supra* note 62, VI.6, RES1 and RES2.

<sup>71</sup> Delvaux, *supra* note 36. para 56. Some shall regret the clumsy formulation of the resolution, which implies that the more autonomous the robot the more liable the trainer, and which makes an unrealistic distinction between “skills resulting from “training” given to a robot” and “skills depending strictly on its self-learning abilities” (if we knew how to make it, we would already be able to identify the crucial “moment” of decision, deviation, or defect and distribute liabilities).

<sup>72</sup> *Ibid*, para 53, 54, 55, 59f).

<sup>73</sup> Art. 1244, 1245 CC.

<sup>74</sup> It can be noted that knowledge about, or acquaintance with, the animal’s general behaviour or habits does not matter. *Mottawi c. Laval (Ville de)*, 2012 QCCS 4455 (CanLII)

Until we choose the type of liability regime and clearly outline all the definitions involved, certain partial but immediate measures can already be applied.

It is thus clearly essential to establish a mandatory insurance regime obligating manufacturers to enrol in an insurance policy covering all potential damages caused by robots.<sup>75</sup> The urgency of adopting such policies is all the more obvious in light of rise of autonomous vehicles, some of which are already hitting our roads.<sup>76</sup> Nevada State, which is somewhat avant-gardist in this domain, can be applauded for its initiative in creating the first legislation in the world to specifically deal with mandatory insurance coverage for autonomous vehicles.<sup>77</sup> The European parliament also advocates for the parallel creation of a compensation fund, “the main function of which would be to guarantee compensation even when the damages are not covered by insurance”.<sup>78</sup>

Other measures concern the adaptation of existing legal vocabulary. It has been recommended to revive the definition of “defective product” contained in directive 85/374/CEE.<sup>79</sup> One could also imagine drawing from other legal texts such as “Machines directive”<sup>80</sup> where some terms, e.g. “danger”, “risk”, “reasonably predictable misuse”, have both broad and narrow definitions. For example, “danger” has a more specific meaning than in ordinary language. It designates a potential source of damage, characteristically inherent to the machine and independent of the damage occurring probability. A very hot component, for example, even if inaccessible to the user, represents a danger.<sup>81</sup>

Finally, other, more hypothetical solutions have been suggested. For example, the standardisation of liability regimes for damages caused contractually or extra-contractually<sup>82</sup> (where the victim has not entered into any contract with the manufacturer or the developer), or the need to distinguish, in a user liability situation, between work accidents and cases where the user is both the owner and the victim, or even the liability presumption against the lessor in cases of lease.<sup>83</sup> A homogeneous regime for all robots, autonomous or not, has also been put forward in order to avoid lengthy preliminary procedures dedicated to distinguishing between liability regimes and additional legal fees.<sup>84</sup> This initiative, though stemming from a noble objective, is unrealistic as it would encompass to many machines and would simply displace the problem of defining terms, from “autonomy” to “robot”. Finally, some have emphasized the importance of adapting class actions rules to the world of robotics.<sup>85</sup>

One will notice that numerous difficulties in defining and specifying rules can already be avoided or limited through the adoption of ethical rules in the AI developing industry. One fundamental aspect is traceability of AI decisions, which could be incorporated at the design and

---

<sup>75</sup> Delvaux, *supra* note 36 para.57, 59a); Bensoussan, *supra* note 26, art. 8.

<sup>76</sup> The Google-car “drives” since 2010.

<sup>77</sup> *Nev. Rev. Stat NRS*, *supra* note 62.

<sup>78</sup> Nevejans, *supra* note 11.

<sup>79</sup> *Directive 85/374/CEE*, *supra* note 66.

<sup>80</sup> *Directive on Machinery*, 2006/42/EC (2006), European Parliament and Council.

<sup>81</sup> Ian Fraser, *Guide to Application of the Machinery Directive 2006/42/EC*, (2010), 2<sup>nd</sup> ed. European Commission Enterprise and Industry

<sup>82</sup> Or “in tort” for common lawyers.

<sup>83</sup> Nevejans, *supra* note 11.

<sup>84</sup> Delvaux, *supra* note 36, para. 59g).

<sup>85</sup> CERNA, *supra* note 62, RES1; Delvaux, *supra* para 12; Bensoussan, *supra* note 26, art. 5.

production level.<sup>86</sup> It would allow for more accurate identification of a decision's origins and thus the person liable for said decision. Increased consumer awareness to appropriate conditions and restrictions of use would greatly decrease the risk of accident.<sup>87</sup> Finally, it is of course imperative to build robots with an integrated emergency stop function.

### **Military context: an illustration of a possible adaptation**

I will indulge in a brief tangent to illustrate the possibilities of adapting law. In the military context, lively debates abound regarding the use of autonomous weapons. Positions vary from all out prohibition<sup>88</sup> to the positing of the existence of a duty to use.<sup>89</sup> However, there is a consensus regarding LAWS (lethal autonomous weapons systems) as being authorized only in the presence of "significant human control" (or "meaningful control").<sup>90</sup> These expressions have yet to be defined. Nonetheless, the military context, in which the use of force is always submitted to hierarchal control and supervision, leaves the door open to a comparison between the robot and the mere soldier. One can thus think about the adaptation of the notion "effective control", which applies in evaluating a superior's liability. This notion is indeed required in international humanitarian law to infer a superior's responsibility for acts committed by soldiers.<sup>91</sup> The following criteria define effective control: a relationship of subordination, knowledge on the part of the superior concerning past, present or future violations, and a failure to take necessary and reasonable measures to prevent or stop these acts or to punish the criminals.<sup>92</sup> Each of these in turn is defined in international jurisprudence. Without going into further detail, one can already notice a potential compatibility with situations involving robots.

Admittedly, the notions of "effective control" and "significant control" can seem paradoxical: when the first triggers responsibility by its presence (and inversely, exonerates when absent), the second, considered as an obligation to which one should not fail, engages liability when absent. Nonetheless, the definition of effective control can be relevant to that of significant control, the imperative to define a factual situation being quite different than inferring from it an obligation.

Hence, the robot, a tool and a resource above all, could be considered as being subordinated to humans who deploy it on the battlefield. These humans could then be imputed knowledge of illicit acts and be held liable for these acts as a result. It could be established that they knew the approximate functioning of the robot, the circumstances and the context of use, the risk of deploying LAWS according to these parameters, etc. One could finally imagine them to be submitted to the same obligation of preventing illicit behaviour by the robot.

---

<sup>86</sup> *Ibid*, RES2.

<sup>87</sup> Bensoussan, *supra* note 26.

<sup>88</sup> Stephen Goose, "The Case for Banning Killer Robots"(2015) 58: 12 Viewpoints p. 43.

<sup>89</sup> Bradley J. Strawser, "Moral Predators: The Duty to Employ Uninhabited Aerial Vehicules" (2010) 9: 4 Journal of Military Ethics 342-368.

<sup>90</sup> Goose, *supra* note 88 ; *Autonomous Weapon systems: Technical, military, legal and humanitarian aspects*, ed (ICRC, 2014), Expert meeting; Open Letter, Future of Life Institute (online) : <: <http://bit.ly/1UQ17QE>>

<sup>91</sup> Rome Statute of the International Criminal Court, 17 July 1998, 2187 R.T.N.U. 3 (entered into force 1<sup>st</sup> July 2002), art. 28.

<sup>92</sup> *International Criminal Law Guidelines: Command Responsibility*, (January 2016) Centre for International Law Research and Policy.

I conclude this section by emphasizing that we too often underestimate law's resources. The most pragmatic, practical, and available solutions are, and will be based on existing notions: we will talk about "predictable damages", "negligence", "reasonable" use and expectations of the consumers regarding conditions of use, "intention" of a human actor underlying a particular result, "voluntary choice" of data given to a robot, etc.<sup>93</sup> The real problem, raised by initial motivations behind debates around robotics – that is, opacity and autonomy – is a definitional one above all. If it is difficult to attribute liability across an algorithmic chain of decisions, it is because, beyond any technical difficulty, the definition of terms that are crucial to the understanding of the phenomenon of AI are missing. How to stratify levels of opacity? In other words, what is an unpredictable damage? If possible to trace algorithmic decisions, what type of algorithmic modification ought to be attributed to which actor? What is an autonomous robot? What legal distinction concerning machines in general gives rise to a liability regime that is specific to autonomous robots? The European parliament defines autonomy as "the capacity to take decisions and put them into practice in the exterior world, independently of any external control or influence".<sup>94</sup> An industrial refrigerator, which continuously monitors and controls its internal temperature in reference to the general climate and to the feedback received through its sensors, would be an autonomous robot by that definition. Hence, it is imperative to draw a line, through a legal test or another tool, between robots and autonomous robots.<sup>95</sup> We should not forget that, while those definitions are informed by technical realities, they remain first and foremost legal definitions. Hence, although they cannot be out of touch with technical criteria, they can distinguish themselves in certain situations and thus avoid certain difficulties. For example, in the case of impossibility of breaching the decision's opacity, a legal presumption or strict liability could still identify a responsible actor. Finally, it is worth stressing the importance of vocabulary and terms used that need not alter technological reality or misguide public perceptions.<sup>96</sup>

The concept of an autonomous robot today no longer corresponds to the outdated legal conception of a "machine".<sup>97</sup> Yet, it does not radically subvert the legal world as we know it. In many respects, it still is a machine – though admittedly more sophisticated and complex – that does not call for a revolution of our liability regimes, but rather for a definition of what it is and of what it does.

## CONCLUSION

The question that guided me throughout this analysis was to know whether or not robots, in the civil legal landscape, were real precursors of a legal upheaval. After having examined their

---

<sup>93</sup> All those hypotheses are explored under the form of questions in the document CERNA, *supra* note 62, VI.6.

<sup>94</sup> Delvaux, *supra* note 36, para. AA.

<sup>95</sup> If we want to treat them according to different regimes. It has also been suggested to treat those two categories uniformly. See: Nevejans, *supra* note 11.

<sup>96</sup> Neil M. Richards and William D. Smart, "How Should the Law Think About Robots?" (2013), online: [http://robots.law.miami.edu/wp-content/uploads/2012/03/RichardsSmart\\_HowShouldTheLawThink.pdf](http://robots.law.miami.edu/wp-content/uploads/2012/03/RichardsSmart_HowShouldTheLawThink.pdf). In this article, the authors, drawing from the experience of cyberlaws, highlight that legal vocabulary evolves in metaphors and that if those are not cautiously thought out, they can have disastrous consequences.

<sup>97</sup> Directive on Machinery, *supra* note 80, art. 2(a): " 'machinery' means: an assembly, fitted with or intended to be fitted with a drive system other than directly applied human or animal effort, consisting of linked parts or components, at least one of which moves, and which are joined together for a specific application".

different potential statuses – those of person, slave or simple machine – I can conclude as follows: robots, legal descendants of slaves, are machines in substance and potentially persons in form. This triple nature will translate into details that, far from being revolutionary, are based on that which already exists. These details will focus on the definitions of terms rather than inventing original regimes. The main objective of legal reforms to come will remain, despite fantasies and science fiction, reaching an equilibrium between fair compensation for victims and incentivizing AI development.