

# « ROBOTS TUEURS »

Que seront les soldats  
de demain ?



Brice Erbland

# « ROBOTS TUEURS »

Que seront les soldats  
de demain ?

**ARMAND COLIN**

Conseiller éditorial : Jean-Vincent Holeindre

Illustration de couverture : © Jonas Birkebæk Poulsen

Photo de l'auteur : © Véronique Ricco

Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.

Le Code de la propriété intellectuelle du 1<sup>er</sup> juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autori-

sation des ayants droit. Or, cette pratique s'est généralisée dans les établissements

d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour

les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée.

Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du

Centre français d'exploitation du droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).



© Armand Colin, 2018

Armand Colin est une marque de

Dunod Éditeur, 11 rue Paul Bert, 92240 Malakoff

ISBN 978-2-200-62007-3

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

# Sommaire

<b>Introduction</b>	<b>9</b>
<b>1 Conséquences de l'emploi des robots soldats</b>	<b>21</b>
<b>Rupture sociologique de la guerre</b>	<b>22</b>
Emploi	22
Morphologie	27
Interaction SALA-humain	29
Terrorisme robotique	35
<b>Biais stratégique et conséquences tactiques</b>	<b>36</b>
<b>2 Émotions vs algorithmes</b>	<b>43</b>
<b>Faiblesses humaines au combat</b>	<b>45</b>
Vengeance	46
Addiction à la destruction	48
Effet Lucifer	51
Distanciation	54
Esclavage technologique	57
Peur	60
Stress post-traumatique	61
Besoins physiologiques	62
<b>Vertus humaines au combat</b>	<b>66</b>
Courage	67

Instinct ou intuition ?	68
Créativité	70
Cohésion	72
Discernement émotionnel	73
<b>Objectifs d'une éthique artificielle</b>	<b>78</b>
<b>3 Programmation d'une éthique artificielle</b>	<b>87</b>
<b>Le processus décisionnel humain</b>	<b>88</b>
Mission et règlements	89
Procédures et expérience	91
Intuition et créativité	93
Éthique	94
Le phénomène check-list	100
Le schéma décisionnel humain	101
<b>Proposition d'architecture d'un module d'éthique artificielle</b>	<b>103</b>
Approche déontologique	103
Logique déontique	106
Logiques non monotones	111
Réseaux de neurones	116
Une combinaison des moyens pour approcher au mieux la décision humaine	125
Une « éducation » nécessaire	130
<b>Quel test pour valider cette éthique ?</b>	<b>134</b>
<b>4 Que nous apprend la littérature ?</b>	<b>139</b>
<b>Conclusion</b>	<b>151</b>

<b>Synthèse des déductions de l'étude</b>	<b>157</b>
Apparence	157
Emploi	157
Intelligence artificielle	158
<b>Remerciements</b>	<b>163</b>
<b>Bibliographie</b>	<b>165</b>





## Introduction

Dans la mythologie grecque, le sculpteur Pygmalion se voue au célibat parce que les femmes de l'île de Chypre le déçoivent. Aucune d'entre elles ne présente assez de qualités à ses yeux, et leur conduite ne satisfait pas sa morale exigeante. Il se réfugie donc dans son atelier, où il sculpte son idéal féminin. La précision et la beauté de son travail sont telles qu'il tombe éperdument amoureux de sa sculpture. Mais sa femme parfaite est composée de pierre de taille et de cire, elle n'a aucune âme, et sa peau et ses lèvres demeurent désespérément froides. Pygmalion prie alors la déesse Aphrodite de donner vie à l'objet de son amour. La déesse, touchée, accède à sa demande et la statue s'éveille : Galatée prend vie et devient la femme de Pygmalion.

Les ingénieurs en robotique militaire sont en quelque sorte des Pygmalions modernes. Ils ont pendant longtemps équipé les soldats humains, mais leurs outils technologiques sont devenus si rapides et si précis que les capacités cognitives des hommes ne suffisent plus... En revanche, la sculpture de la statue idéale reste possible : ils cherchent donc à créer un robot soldat qui serait le guerrier parfait, sans défaut sur le champ de bataille. Hélas, le problème reste le même que dans l'Antiquité, car si les connaissances scientifiques actuelles permettent peut-être de sculpter la statue rêvée, cette dernière n'a

toujours pas d'âme. Pourtant, si la déesse Aphrodite ne peut rien pour les ingénieurs, les promesses du progrès informatique sont de bon augure. À tel point que l'agence de recherche et développement militaire américaine, la DARPA<sup>1</sup>, estime qu'une intelligence artificielle assez accomplie pour pouvoir raisonner de manière autonome, comme un être humain sur un champ de bataille, pourrait voir le jour d'ici 2030<sup>2</sup>. L'autonomie intellectuelle de la technologie serait donc le Graal de l'éveil de conscience du robot soldat, l'intervention divine qui donnerait vie à la statue révérée par les sculpteurs du combat moderne.

Mais qu'est-ce qu'une technologie autonome au combat ? Est-ce un simple logiciel « intelligent » gouvernant différents systèmes automatisés, à l'image de l'ordinateur HAL 9000 de 2001, *L'Odyssée de l'espace*, ou est-ce un robot tueur humanoïde à l'image de *Terminator* ? L'imaginaire collectif associe aisément technologie autonome et robot, et l'on donne rapidement ce qualificatif à toute machine qui se déplace ou effectue des tâches sans pilote. Peter W. Singer propose d'ailleurs, en abordant les technologies autonomes dans son ouvrage *Wired for War*, une définition du robot qui repose sur quatre critères :

- le robot est fabriqué par l'homme ;
- il possède des senseurs pour appréhender son environnement ;

---

1. *Defense Advanced Research Projects Agency*.

2. KRISHNAN Armin, *Killer Robots: Legality and Ethicality of Autonomous Weapons*, Burlington, Ashgate Publishing Company, 2009, 204 p.

- il contient des programmes pour définir une réponse à une situation donnée ;
- il a les moyens de mettre en œuvre cette réponse<sup>1</sup>.

Si cette définition donne une première approche intéressante, elle n'est malheureusement pas assez précise, et donc pas assez restrictive pour cerner la nature d'une technologie autonome. Car un missile antiaérien, doté d'un système auto-directeur, serait selon cette définition un robot : fabriqué par l'homme, possédant un senseur infrarouge, contenant un programme pour déterminer sa trajectoire afin de suivre sa cible, possédant des ailettes amovibles pour appliquer les changements de trajectoires établis. Ce n'est pourtant pas le missile qui décide d'ouvrir le feu, et encore moins de la cible à traiter. Il faut donc préciser cette définition.

L'organisation non gouvernementale *Human Rights Watch*, quant à elle, définit le robot<sup>2</sup>, qu'elle associe à une arme dans sa lutte contre les robots tueurs, comme une machine pouvant opérer de manière autonome dans la sélection des cibles. Mais elle classe le niveau d'autonomie, et donc d'intervention humaine dans le processus d'ouverture du feu, selon trois catégories. Les armes « *human-in-the-loop* » sont celles où le robot sélectionne les cibles mais pour lesquelles l'homme commande

---

1. SINGER Peter W., *Wired for War: the Robotics Revolution and Conflict in the 21st Century*, New York, Penguin Books, 2009, 499 p.

2. HUMAN RIGHTS WATCH, *Losing Humanity: the Case against Killer Robots*, HRW, 2012, 49 p.

l'ouverture du feu. Les armes « *human-on-the-loop* » sont celles où le robot sélectionne les cibles et ouvre le feu sur elles, mais pour lesquelles l'homme peut intervenir dans le court laps de temps avant l'ouverture du feu et annuler la décision. Enfin, les armes « *human-out-of-the-loop* » sont celles où le robot sélectionne et traite les cibles sans qu'aucun humain ne puisse intervenir. Le professeur américain Armin Krishnan reprend en quelque sorte ces trois catégories<sup>1</sup>, en les nommant respectivement autonomes préprogrammée, supervisée et complète, mais ajoute une catégorie première qui est l'autonomie télé-opérée. Il embrasse ainsi tous les systèmes d'armes existants et peut les classer selon leur degré d'autonomie. Mais c'est là un biais de jugement, car le terme autonomie est alors écarté de son sens premier. La définition du Larousse spécifie en effet la « capacité de quelqu'un à être autonome, à ne pas être dépendant d'autrui ; caractère de quelque chose qui fonctionne ou évolue indépendamment d'autre chose ». Un système entièrement et uniquement télé-opéré ne peut donc pas, par définition, être autonome.

Jean-Baptiste Jeangène Vilmer, en rappelant la distinction fondamentale entre autonomie et automaticité<sup>2</sup>, confirme qu'une technologie autonome doit être différenciée des systèmes d'armes automatisés déjà existants. L'automaticité est en effet la caractéristique des systèmes

---

1. KRISHNAN Armin, *op. cit.*

2. JEANGÈNE VILMER Jean-Baptiste, *Diplomatie des armes autonomes : les débats de Genève*, in *Politique étrangère* Automne, n° 3, 2013, pp. 119-130.

qui sont programmés pour réagir d'une certaine façon à une situation donnée et perçue par des capteurs. Le système réagira donc toujours de la même façon face aux mêmes *stimuli*, quel que soit le contexte. L'autonomie, quant à elle, implique que le système soit capable de mener un raisonnement à chaque fois unique, quelle que soit la situation perçue, pour décider de la meilleure action à mener. Il s'agit donc pour ce qui nous concerne, en respectant la distinction de notion, de parler de *systèmes d'armes autonomes*. Et étant donné qu'il s'agit de technologie destinée au combat, c'est-à-dire capable de délivrer un feu mortel, nous parlerons de *système d'armes létal autonome* (SALA). Jean-Baptiste Jeangène Vilmer en propose une définition bien plus précise et réaliste, en reprenant les notions déjà abordées : « un système d'armes qui, une fois activé, est capable de décider seul, c'est-à-dire sans intervention ni supervision humaine, du ciblage et du déclenchement de la frappe, en fonction d'un environnement changeant auquel il s'adapte<sup>1</sup> ». Ce qui sous-entend premièrement, par les termes « une fois activé », que le SALA n'appartient pas exclusivement à l'une ou l'autre des catégories d'autonomie proposées par *Human Rights Watch*, mais qu'il les embrasse toutes par un système hybride où l'humain peut être, selon les cas, *in/on/out-of-the-loop* ; deuxièmement, par les termes « auquel il s'adapte », que l'autonomie exige

---

1. JEANGÈNE VILMER Jean-Baptiste, « Terminator Ethics : faut-il interdire les “robots tueurs” ? », in *Politique étrangère* Hiver, n° 4, 2014, pp. 151-167.

une intelligence artificielle capable d'apprendre par expérience et de s'adapter à une situation nouvelle, ce qui est un des défis majeurs des recherches en intelligence artificielle.

Car le champ de bataille n'est pas, contrairement à ce que pourrait laisser croire l'analogie habituelle, aussi simple et normé qu'un plateau d'échecs. Si l'intelligence artificielle est capable depuis peu de surpasser l'homme au plus classique des jeux de réflexion, elle n'est pas encore prête à affronter la complexité mouvante d'une situation de guerre. Le champ de bataille est désormais multidimensionnel, puisqu'aux espaces de conflit habituels, terrestre, maritime et aérien, s'ajoutent l'espace exo-atmosphérique et le cyberspace. Un SALA pourra donc être conçu pour combattre dans l'une ou l'autre de ces dimensions, voire être capable de le faire dans plusieurs d'entre elles, successivement ou simultanément, notamment pour le cyberspace. Mais comment définir le champ de bataille ? La doctrine militaire ne le définit pas en tant que tel, mais l'inclut dans la définition de l'espace de manœuvre d'une unité opérationnelle. Littéralement, il s'agit du lieu où se livre une bataille. Or, selon le Larousse, la bataille est un « combat livré entre deux armées ou deux flottes, dont l'issue influe sur la conduite de la guerre ». Mais si l'on considère la portion d'espace géographique, dans la ou les dimensions envisagées, couverte par la portée des différentes armes, il est rapidement possible d'atteindre la planète tout entière. Il faut donc restreindre la définition pour que la question du remplacement du soldat humain par une technologie autonome ait un sens. Le