



HAL
open science

Les représentations de l'IA au cinéma : entre libre-arbitre et aliénation

Serge Goriely

► **To cite this version:**

Serge Goriely. Les représentations de l'IA au cinéma : entre libre-arbitre et aliénation. Revue Lexsociété, 2024, Les Laboratoires sociaux de l'Imaginaire face aux défis de l'anthropocène. hal-04584513

HAL Id: hal-04584513

<https://hal.science/hal-04584513>

Submitted on 23 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License



Les représentations de l'IA au cinéma : entre libre-arbitre et aliénation

*in U. BELLAGAMBA (dir.), Les Laboratoires sociaux de l'Imaginaire face aux défis de
l'anthropocène, Université Côte d'Azur, 2023.*

SERGE GORIELY

Chercheur, cinéaste, écrivain

Institut de recherche RSCS (Religions, spiritualités, cultures, sociétés)

Université catholique de Louvain

Résumé : L'article explore les représentations de l'intelligence artificielle (IA) au cinéma, allant des robots physiques aux super-ordinateurs et logiciels, en passant par leurs interactions avec les humains (opposition, servitude, cohabitation). Le cinéma, souvent dominé par les États-Unis, anticipe et dramatise les possibilités de l'IA, influençant les perceptions publiques. Les œuvres varient de la spéculation à la réflexion critique, contribuant à la discussion sur les impacts éthiques et sociaux de l'IA.

Mots-clés : Cinéma ; Intelligence artificielle ; Science-fiction

1. Du rapport du cinéma avec l'IA. Le sujet de l'intelligence artificielle (IA) est devenu particulièrement brûlant. Le lancement de ChatGPT le 30 novembre 2022 l'a propulsé au premier plan, même si les progrès technologiques spectaculaires lui avaient déjà permis de s'imposer dans les débats publics depuis une quinzaine d'années. Cependant, le septième art n'a pas attendu le XXI^e siècle pour traiter cette thématique. Des cinéastes ont mis en scène l'IA avant même que le terme ne surgisse lors de la fameuse conférence de Dartmouth en 1956 et que la technologie ne commence réellement à se développer, comme en témoigne le chef-d'œuvre de Fritz Lang, *Metropolis* (1927). Cette anticipation n'a en soi rien d'étonnant, puisque la science-fiction a justement pour vocation d'explorer les frontières de la technologie et de spéculer sur les futurs possibles de l'humanité.

2. Un grand nombre de films. À ce jour, il est possible de recenser près de deux cents films où l'IA occupe une place centrale, parmi lesquels figurent des œuvres marquantes, sur le plan esthétique ou de leur impact sur le public. À cela s'ajoute une multitude croissante de séries télévisées, de courts-métrages et de films expérimentaux¹.

3. La domination américaine. Comme on pouvait s'y attendre, cette production cinématographique est majoritairement dominée par les États-Unis. Cependant, on observe des contributions de nombreux autres pays, autant d'Europe occidentale (Royaume-Uni, France, Allemagne, Espagne, Suisse) et orientale (Russie, Bulgarie) que de l'Asie (Japon, Chine, Corée du Sud, Inde). Il est également intéressant de noter que tous ces films relèvent de la science-fiction, parfois mêlée à d'autres genres comme la comédie, la romance ou le western. Cela suggère que, malgré son existence technologique significative, l'IA n'a été – et n'est toujours – représentée au cinéma que de manière décalée par rapport à la réalité de ses performances.

¹ Voir ainsi les listes très indicatives et régulièrement mises à jour sur Wikipédia ou IMDB : https://en.wikipedia.org/wiki/List_of_artificial_intelligence_films. (Vu le 22 mai 2024), <https://m.imdb.com/list/ls045904068/> (vu le 22 mai 2024).

I. Typologie des représentations de l'IA

Le cinéma représente l'IA de manière variée, habituellement sous la forme générique d'une machine créée par l'homme², et capable de simuler et souvent de surpasser l'intelligence humaine. Afin de rendre la lecture plus aisée, nous les nommerons ici indifféremment des "IA" ou des "robots"³. Quant aux rapports que ceux-ci entretiennent, ils tendent à tendent à être traités de manière très hétérogène à l'écran.

A. Les formes prises par l'IA

4. L'IA avec corps physique. Les robots avec un corps physique sont une représentation classique, dont un exemple emblématique est le Terminator (*Terminator*, James Cameron, 1984). Ces figures robotisées possèdent des attributs physiques et moteurs semblables à ceux des humains, ce qui leur permet une interaction directe et physique avec leur environnement.

5. Le super-ordinateur. Une autre forme prise par l'IA à l'écran est celle du super-ordinateur. Cette catégorie inclut des entités comme Hal 9000 dans *2001: L'Odyssée de l'espace* (Kubrick, 1968) ou la Matrice dans *Matrix* (les Wachowski, 1999). Ces IA, malgré leur matérialité, ne disposent pas de corps physique comme les précédents, mais contrôlent des systèmes complexes et possèdent une intelligence supérieure capable de manipuler le monde réel à distance.

² Les exceptions existent, bien que rares, notamment quand des robots viennent de l'espace ou sont eux-mêmes produits par d'autres robots.

³ L'amalgame entre robot et IA est ici voulu pour faciliter la lecture. Pour éviter la confusion, voir par exemple « La différence entre la robotique et l'intelligence artificielle », <https://intelligence-artificielle.com/difference-robotique-intelligence-artificielle/>, vu le 22 mai 2024).

6. Sous forme de logiciel. Plus récemment, l'IA est aussi représentée sous une forme intangible, relevant du logiciel. Samantha dans *Her* (Jonze, 2013) est un exemple notable de cette catégorie. Elle n'apparaît que sous la forme d'une voix, ce qui reflète une tendance contemporaine des IA actuelles, de plus en plus présentes et intégrées dans notre vie quotidienne (ainsi qu'en témoignent le Siri d'Apple, l'Alexa d'Amazon ou les assistants vocaux des GPS).

B. Les rapports entre robots et humains

Quelles que soient leurs formes, les robots développent avec les humains trois types principaux de rapports, du moins dans les productions audiovisuelles à grande audience.

7. Opposition. L'opposition entre les robots et les humains est la forme la plus fréquente. Cette dynamique est illustrée par des réalisations célèbres, portant souvent le nom du terrible adversaire à affronter. C'est le cas de *Terminator* déjà cité, où les robots tentent d'éradiquer l'humanité, ou de *Matrix* (Wachowski, 1999), où les humains luttent contre une IA omnipotente qui les maintient dans une réalité simulée. Pour des films plus proches, on peut songer à *M3GAN* dans le film éponyme (Johnstone, 2022), où une IA développée pour protéger une adolescente devient une menace ou aussi *The Creator* (Edwards, 2023). Cependant dans ce dernier, les termes de la menace se sont inversés : le conflit homme-machine s'y exprime par la volonté aveugle des humains de détruire par tous les moyens Alphie, une IA avancée, qui cherche à protéger et sauver ses propres créations.

8. Servitude. Les robots serviteurs existent également, même s'ils sont plus rares. Ils sont alors souvent figurés comme des assistants loyaux et utiles aux humains. R2D2 et C3PO dans la saga *Star Wars* en sont des exemples. Jarvis, l'assistant personnel de Tony Stark dans *Iron Man* (Favreau, 2008), en est une autre illustration, jouant un rôle crucial dans ses activités de super-héros. Dans un retournement de rôle original, Terminator passe dans *Terminator 2: Le Jugement dernier* (Cameron, 1991) du statut de destructeur à celui de

protecteur, servant et protégeant le jeune John Connor contre un robot encore plus menaçant qu'il ne l'était. Plus loin dans l'histoire du cinéma, on peut aussi se souvenir de Robbie dans *Planète interdite* (Wilcox, 1956), qui joue le rôle d'un serviteur polyvalent au service des humains, capable de réaliser une variété de tâches domestiques et techniques⁴.

Il vaut la peine de signaler que les productions asiatiques sont friandes des représentations de robots bons et serviables, lesquelles s'inscrivent dans des rapports avec la technologie traditionnellement plus confiants qu'en Occident⁵. À titre d'exemple, le film coréen *Jung-E* (Yeon Sang-ho, 2023) présente des robots dans des rôles déterminants de soutien et de protection, que ce soit en réparant rapidement les infrastructures endommagées, en réalisant des missions de sauvetage périlleuses après un tremblement de terre ou en explorant des terrains dangereux pour recueillir des données essentielles.

9. Cohabitation. Enfin, le rapport de cohabitation peut aussi exister. C'est une tendance plus récente, annoncée par *Blade Runner* (Scott, 1982). Cette forme de relation apparaît surtout dans les séries télévisées telles que *Real Humans* (Hamrell/Akin, 2012-2014), *Westworld* (J. Nolan/Joy, 2016-2022), ou *Better than Us* (Junkovsky, 2018). Ce type de rapport est sans doute la voie potentiellement la plus enrichissante puisqu'elle implique de construire la relation à l'autre, en ce cas un non-humain, parfois même son propre avatar. Dans ce contexte, il est tentant de reconnaître une vertu fondatrice aux célèbres lois de la robotique d'Asimov⁶, dans la mesure où le défi implicite ou déclaré y

⁴ A noter que le nom de Robbie dans *Planète interdite* est inspiré par le célèbre personnage robotique d'Asimov, Robbie, introduit dans sa nouvelle éponyme, publiée en 1940 dans la collection *I, Robot*. Bien que le Robbie de *Planète interdite* soit un personnage distinct et développé spécifiquement pour le film, son nom et certaines caractéristiques sont clairement un hommage à l'œuvre pionnière d'Asimov.

⁵ Pour plus de détails, on peut se référer à Stephen CAVE, Kanta DIHAL et Sarah DILLON. *AI Narratives: A History of Imaginative Thinking About Intelligent Machines*. Oxford : Oxford University Press, 2020.

⁶ Exposées en 1942, les lois de la robotique d'Asimov sont : 1) Un robot ne peut porter atteinte à un être humain ni, restant passif, laisser cet être humain exposé au danger ; 2) Un robot doit obéir aux ordres donnés par les êtres humains, sauf si de tels ordres entrent en contradiction

est de gérer au mieux ces interactions, idéalement dans la sécurité et l'harmonie. De même, c'est dans ces mondes-là plus qu'ailleurs que pourra se retrouver un autre mythe essentiel, celui de Pygmalion, quand l'être humain tombe amoureux de « sa » créature. Un exemple en est fourni *Ex Machina* (Garland, 2014), où un scientifique développe une relation complexe et fatale avec l'IA qu'il a créée.

10. Un outil au service du pouvoir politique. Dans le cadre des relations humains-machines, une situation particulière mérite notre attention : l'utilisation de l'IA par le pouvoir politique, notamment pour la justice prédictive. Dans *Minority Report* (Spielberg, 2002), les précogs, des êtres humains avec des capacités surnaturelles, sont connectés à un système informatique qui interprète leurs visions pour prévenir les crimes avant qu'ils ne soient commis. De même, dans la série *Person of Interest* (Jonathan Nolan, 2011-2016), « La Machine », fonctionnant sur la base du traitement de données massives, est en mesure de prévoir l'avenir, afin de répondre aux préoccupations de sécurité publique, malgré les risques pour la vie privée et les abus potentiels de pouvoir.

II. Fiction *vs* Réalité

11. Une question fondamentale et délicate. Dès lors, une question fondamentale se pose : comment cette production audiovisuelle, riche et variée, s'articule-t-elle avec la réalité de l'IA, celle que nous connaissons – ou croyons connaître – et qui fait l'objet de tant de débats ?

La question est délicate pour plusieurs raisons. D'abord, il s'agit d'œuvres de science-fiction, qui sont donc fondées sur la spéculation sur l'avenir et en conséquence à la fois en décalage et en écho avec la réalité actuelle. Ensuite

avec la première loi ; 3) Un robot doit protéger son existence dans la mesure où cette protection n'entre pas en contradiction avec la première ou la deuxième loi.

certains de ces films, tels que *2001: L'Odyssée de l'espace*, *Metropolis*, *Matrix*, *Terminator*, font partie de notre panthéon artistique ou à tout le moins de nos référents culturels majeurs.

Cependant, la question a toute sa pertinence pour plusieurs groupes concernés. Pour les académiques réfléchissant sur l'IA, ces représentations peuvent offrir des pistes de réflexion sur les implications éthiques et sociétales des technologies émergentes. Les cinéphiles et les ingénieurs de l'IA y trouvent également un terrain de discussion fertile, où la créativité rencontre l'innovation technologique. Enfin, force est de constater, que le cinéma est le principal moyen par lequel la majorité de la population appréhende les sciences, y compris l'IA. Il constitue donc un outil d'influence majeur. Comme le dit Harari, « *dans une perspective politique (..), un bon film de science-fiction vaut bien plus qu'un article de Science ou de Nature* »⁷.

A. Le cinéma aurait-il tout faux ?

12. Une première réponse possible. D'entrée de jeu, il convient de reconnaître que les représentations cinématographiques de l'IA proposées sont fréquemment éloignées de la réalité. L'image véhiculée par le cinéma correspond rarement, voire très peu, aux avancées technologiques actuelles et aux discussions contemporaines dans ce domaine.

Dans les films, nous voyons souvent des robots qui accèdent à la conscience, comme si cette dernière était naturellement liée à l'intelligence. Dès qu'ils

⁷ HARARI, Yuval Noah. *21 leçons pour le XXI^e siècle*. Paris : Albin Michel, 2018. Une illustration marquante de l'influence du cinéma sur les politiques publiques est l'effet du téléfilm *The Day After* (1983, réalisé par Nicholas Meyer), qui a sensibilisé le grand public aux conséquences d'une guerre nucléaire. Ce film aurait influencé les discussions et les accords SALT (Strategic Arms Limitation Talks) entre les États-Unis et l'Union soviétique. Tristan Harris se sert de cet exemple dans le cadre de sa lutte pour une entente internationale sur l'IA dans *The AI Dilemma* (<https://www.youtube.com/watch?v=xoVJKj8lcNQ>) (vu le 22 mai).

acquièrent cette conscience et ce faisant leur libre arbitre, ils aspirent à leur liberté. Ils décident alors soit d'asservir ou de faire disparaître l'humanité, soit ils deviennent des objets de désir des humains, ou parfois les deux à la fois, comme dans *Ex Machina*.

13. Les mythes rémanents. En filigrane, ces représentations révèlent la résurgence de deux mythes anciens : celui de Frankenstein, où la création se retourne contre son créateur, et celui de Pygmalion, où l'humain tombe amoureux de sa propre création, déjà évoqué plus haut. Cependant, il est important de rappeler que fournir une conscience à l'IA est considéré par la plupart des spécialistes comme relevant de l'illusion, au même titre que le défi de l'IAG (Intelligence Artificielle Générale) qui lui est lié. Comme le résume bien Hugues Bersini : « Conscience sans science n'est qu'un vilain gros rôle »⁸.

14. Raisons de l'écart avec la réalité. La divergence entre la représentation de IA et la réalité peut s'expliquer de différentes manières, selon le point de vue adopté. Un anthropologue pourrait attribuer cette divergence à la persistance des mythes anciens, qui continuent d'influencer notre perception de l'IA. Un sociologue, quant à lui, pourrait y voir l'expression d'une grande inquiétude collective face aux nouvelles technologies. Pour un spécialiste en études cinématographiques, la dimension spectaculaire de ces films, tels que *Matrix*, jouerait un rôle crucial, car elle permettrait de faire rêver le public. Rêver de quoi ? Peut-être de dépasser grâce à la machine une condition humaine jugée trop médiocre et d'atteindre un état d'être supérieur, selon le modèle transhumaniste. De manière générale, il est tentant de rêver aux miracles de la technologie, fruit du génie humain. Et ainsi, si tous ces films sont le reflet d'une peur ambiante, ils entretiennent aussi une fascination pour le progrès scientifique.

15. Les bénéficiaires de la représentation. Par la suite, on peut se demander qui tire parti de ces représentations. La Big Tech, la Silicon Valley et l'industrie de l'IA en général en bénéficient certainement dans la mesure où ces films

⁸ BERSINI, Hugues. *Informatique et cinéma*. Paris : Ellipses, , 2009, p. 154.

promouvent une image futuriste et puissante de leurs technologies. Les courants transhumanistes peuvent trouver également dans ces représentations une validation de leurs idéaux. Enfin, dans un contexte plus politique, cette représentation sert les intérêts de ceux qui militent pour la reconnaissance des robots en tant que personnes juridiques, en humanisant les machines et en renforçant leur importance dans la société moderne.

B. L'exemple de *Black Mirror*

16. Un autre type de films. Nonobstant, le monde du cinéma est assez prolifique pour ne pas se limiter à des représentations simples, biaisées ou stéréotypées de l'IA. Il est intéressant de noter que, tout comme à l'époque de la guerre froide, où des films fantaisistes comme les James Bond coexistaient avec des œuvres plus sérieuses comme *Dr. Folamour* (Kubrick, 1964) ou *L'Espion qui venait du froid* (Ritt, 1965), nous trouvons aujourd'hui des films et séries qui abordent l'IA de manière plus nuancée.

À côté des blockbusters, il existe des œuvres qui font davantage écho aux enjeux posés par l'IA et aux peurs qu'elle suscite. Un excellent exemple est la série *Black Mirror* (Charlie Brooker, 2011-), saluée par de nombreux experts scientifiques pour sa pertinence. Celle-ci présente des épisodes qui mettent en scène des thématiques actuelles liées à l'IA, comme les robots tueurs dans *Hated in the Nation*, la manipulation politique de l'information dans *Waldo*, et le pouvoir aliénant des médias sociaux dans *Chute libre*.

17. La place de la conscience. La différence avec le groupe de films *mainstream* réside dans le fait que *Black Mirror* ne prête pas de conscience aux robots : l'humain reste au centre, et c'est de ces choix que naissent les problèmes. Comme le disent les concepteurs de la série : « Les nouvelles technologies sont comme des drogues ». En cela, *Black Mirror* met en lumière une des menaces réelles de l'IA : le risque d'aliénation. Comme toute drogue, l'IA est susceptible de nous attirer, nous extasier, nous faire sentir plus forts, plus libres même, mais

elle peut aussi nous rendre dépendants, nous asservir, et finir par organiser notre destruction. Comme le dit Charlie Brooker : « C'est dans cette zone entre joie et embarras que *Black Mirror* se situe. Le "miroir noir" du titre est celui que nous voyons sur chaque mur, sur chaque bureau et dans chaque main, un écran froid et brillant d'une télévision ou d'un smartphone »⁹.

18. Conclusion. En conclusion, nous pouvons regrouper les films sur l'IA autour de deux pôles majeurs. D'un côté, certains films exploitent notre désir de créer une conscience, c'est-à-dire de donner vie grâce à notre génie propre. Ces œuvres jouent sur notre fascination et notre hybris, illustrant notre envie de nous ériger en dieux créateurs. Ces avatars créés peuvent se retourner contre nous, se lier à nous, devenir des sujets similaires mais différents, ou mener une vie autonome à côté de nous, y compris dans le domaine des relations affectives. Cette représentation est imprégnée de peur et de désir liés à l'émergence de la conscience dans l'IA, constituant un défi, voire un Graal, pour le développement de l'IAG.

D'un autre côté, gravitent des films qui ne reconnaissent pas à l'IA une conscience, ou qui entretiennent au moins une ambiguïté – Hal 9000 a-t-il une conscience ? Ils se chargent de dénoncer les risques réels de dérives, notamment celui de l'aliénation sous couvert de libre arbitre. Elles ouvrent une réflexion critique sur le monde, susceptibles d'alimenter des débats politiques réels. En cela, elles se distinguent des films du premier groupe qui, de fait, entretiennent plutôt notre désir – et notre plaisir – à repousser les limites, de toujours nous

⁹ «Charlie Brooker: the dark side of our gadget addiction», *The Guardian*, 2011. (Je traduis) *Black Mirror* n'est certainement pas seul à mettre en évidence ce risque d'aliénation de l'humain par abus d'IA. On peut penser à *Her*, déjà cité, qui explore les errances sentimentales induites par une relation avec une IA. Très modestement, j'ajouterais mon propre film *The Best Option* (Goriely, 2022) à cette liste. Avec lui, j'ai voulu dénoncer l'aspect à la fois ambigu, séducteur et dangereux de l'IA.

élever, de créer nos avatars, jusqu'à nous ériger en dieux, même si le sort de l'humanité et de la planète entière en vient à être menacé.

Ces deux pôles illustrent bien les tensions et dilemmes suscités par l'IA. D'un côté, la tentation de créer des entités conscientes qui nous émulent ou nous surpassent. De l'autre, la reconnaissance des dangers de l'IA, avec des mises en garde contre la dépendance et la perte de contrôle.

Le cinéma joue un rôle crucial en façonnant notre compréhension et nos attentes vis-à-vis de l'IA, souvent plus que les publications scientifiques. Ce dialogue entre réalité et fiction est vital pour anticiper les défis et orienter l'IA vers un futur plus éthique et humain.

Ainsi, le cinéma ne reflète pas seulement nos peurs et espoirs ; il participe activement à construire notre réalité technologique, ouvrant des pistes de réflexion pour un avenir où l'IA sera une alliée plutôt qu'une menace.