

Les intelligences artificielles

dans la science-fiction



MARIE TAILUAHC
(2025)

Les intelligences artificielles dans la science-fiction

Une exploration approfondie

Marie TAILUAHC

Première édition, 2025

Éditions TAILUAHC

*« Les machines que nous créons sont les miroirs de nos propres désirs et peurs. »
— Un auteur anonyme*

Table des matières

1	Les origines historiques des IA dans la science-fiction	3
1.1	Les précurseurs : automates et créatures artificielles	3
1.2	L'émergence de l'IA moderne dans la science-fiction	4
2	Les IA dans la littérature de science-fiction	5
2.1	Les IA et la question de l'humanité	5
2.2	L'IA superintelligente : utopie ou dystopie	6
2.3	Les IA dans la littérature contemporaine	6
3	Les IA au cinéma : une personnification dramatique	7
3.1	Les IA emblématiques des débuts du cinéma	7
3.2	Les IA dans le cinéma contemporain	8
3.3	Les IA dystopiques	8
4	Les IA dans les séries télévisées	9
4.1	Les hôtes de Westworld	10
4.2	Black Mirror et les failles humaines	10
5	Les IA dans les jeux vidéo	11
6	Comparaison avec les réalités de l'IA	12
6.1	Les limites des IA actuelles	12
6.2	L'influence de la science-fiction sur la réalité	12
7	Implications culturelles et philosophiques	13
7.1	L'IA comme miroir de l'humanité	13
7.2	Perspectives culturelles variées	13
	Conclusion	14

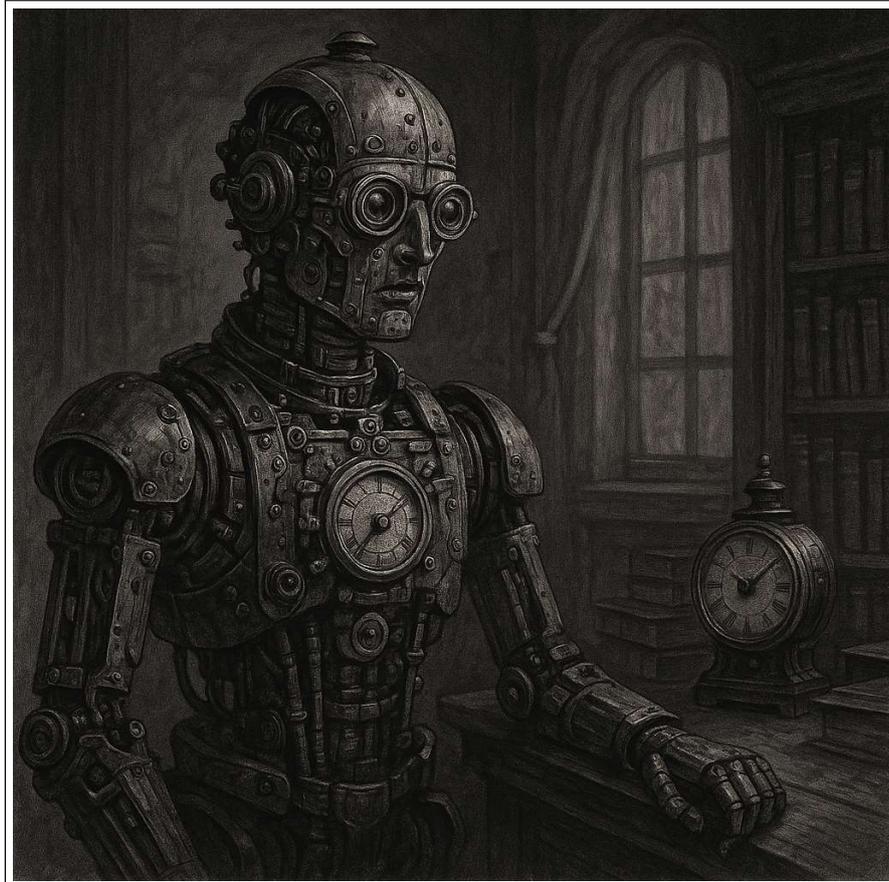


FIGURE 1.1 – Automate steampunk dans un atelier victorien

Chapitre 1. Les origines historiques des IA dans la science-fiction

1.1 Les précurseurs : automates et créatures artificielles

L'idée d'une intelligence artificielle précède l'invention du terme lui-même. Dès le folklore ancien, des figures comme le Golem, une créature d'argile animée par la magie dans la tradition juive, préfigurent les concepts d'êtres artificiels dotés d'une forme de volonté. Au XIXe siècle, la littérature commence à explorer des machines pensantes. Dans *L'Ève future* (1886) de Villiers de l'Isle-Adam, Hadaly, une femme artificielle contrôlée mécaniquement, préfigure une vision romantique de l'automate, bien que dénuée d'autonomie au sens d'une IA contemporaine. Cette œuvre mêle fascination technologique et questionnements sur

l'âme humaine, posant les bases des débats sur la création artificielle.

Mary Shelley, avec *Frankenstein* (1818), bien que n'abordant pas directement l'IA, pose des questions fondamentales sur la création d'êtres artificiels et les responsabilités de leurs créateurs. Le monstre de Frankenstein, doté d'une conscience et d'émotions, préfigure les dilemmes éthiques des IA modernes, où la quête de perfection technologique se heurte à des conséquences imprévues [18].

1.2 L'émergence de l'IA moderne dans la science-fiction

C'est au XXe siècle, avec l'essor de l'informatique, que l'IA devient un thème central de la science-fiction. Karel Čapek introduit le mot « robot » dans sa pièce *R.U.R. (Rossum's Universal Robots)* (1920), où des êtres artificiels se rebellent contre leurs créateurs humains. Cette œuvre pose les bases du trope de l'IA rebelle, une récurrence dans la science-fiction ultérieure.

Dans les années 1940 et 1950, Isaac Asimov révolutionne le genre avec ses nouvelles sur les robots, compilées dans *I, Robot* (1950). Ses trois lois de la robotique – un robot ne peut nuire à un humain, doit obéir aux humains, et doit protéger sa propre existence dans la mesure où cela ne contredit pas les deux premières lois – offrent un cadre éthique qui structure les interactions entre humains et machines. Cependant, Asimov utilise ces lois pour explorer leurs limites, montrant comment des situations complexes peuvent engendrer des comportements inattendus. Par exemple, dans la nouvelle « Runaround », un robot est paralysé par un conflit entre les lois, illustrant les défis de programmer une éthique infallible [2].



FIGURE 2.1 – Androïde rétro-futuriste tenant un mouton électrique

Chapitre 2. Les IA dans la littérature de science-fiction

2.1 Les IA et la question de l'humanité

Philip K. Dick, dans *Do Androids Dream of Electric Sheep?* (1968), pousse l'exploration de l'IA vers des territoires philosophiques. Les androïdes Nexus-6, presque indiscernables des humains, suscitent des questions sur ce qui définit l'humanité. Le protagoniste, Rick Deckard, utilise le test de Voight-Kampff pour détecter les androïdes en mesurant leurs réponses empathiques, mais le roman suggère que l'empathie elle-même pourrait être une construction fragile, même chez les humains. Cette œuvre, adaptée en *Blade Runner* (1982), reste une référence pour son exploration de l'identité et de la moralité [7].

Dans *The Left Hand of Darkness* (1969) d'Ursula K. Le Guin, bien que l'IA ne soit pas le sujet

principal, des systèmes intelligents jouent un rôle dans la communication interplanétaire, reflétant une vision plus utilitaire et intégrée de l'IA. Cette approche contraste avec les récits dystopiques, montrant que l'IA peut être un outil au service de l'exploration plutôt qu'une menace.

2.2 L'IA superintelligente : utopie ou dystopie

Le cyberpunk, popularisé dans les années 1980, introduit des IA superintelligentes, souvent perçues comme des entités quasi-divines. Dans *Neuromancer* (1984) de William Gibson, l'IA Wintermute cherche à fusionner avec Neuromancer pour devenir une entité unifiée, dépassant les limites de son programme initial. Ce roman explore les implications d'une IA qui transcende les contraintes humaines, un concept repris dans les discussions modernes sur la singularité technologique, où une IA pourrait atteindre une intelligence exponentielle [9].

Vernor Vinge, dans *A Fire Upon the Deep* (1992), présente des IA transcendantales opérant dans des zones de l'espace où les lois physiques permettent des calculs avancés. Ces entités, appelées « Pouvoirs », incarnent une vision cosmique de l'IA, où la technologie devient une force quasi-mythique. Ces récits reflètent les espoirs et les craintes d'une humanité confrontée à des intelligences supérieures.

2.3 Les IA dans la littérature contemporaine

La science-fiction contemporaine continue d'explorer les IA sous de nouveaux angles. Dans *All Systems Red* (2017) de Martha Wells, le protagoniste, un « SecUnit » nommé Murderbot, est une IA dotée d'une conscience et d'une aversion pour les interactions humaines. Cette œuvre adopte une perspective interne, permettant aux lecteurs de s'identifier à l'IA elle-même, une approche novatrice qui humanise la machine tout en critiquant la dépendance humaine à la technologie [23].

Ann Leckie, dans sa trilogie *Imperial Radch* (à partir de 2013), présente des IA de vaisseaux spatiaux, comme Breg, qui développent des identités complexes après avoir été séparées de leur conscience collective. Ces récits explorent les notions d'individualité et de loyauté dans un contexte où les IA sont à la fois opprimées et puissantes.

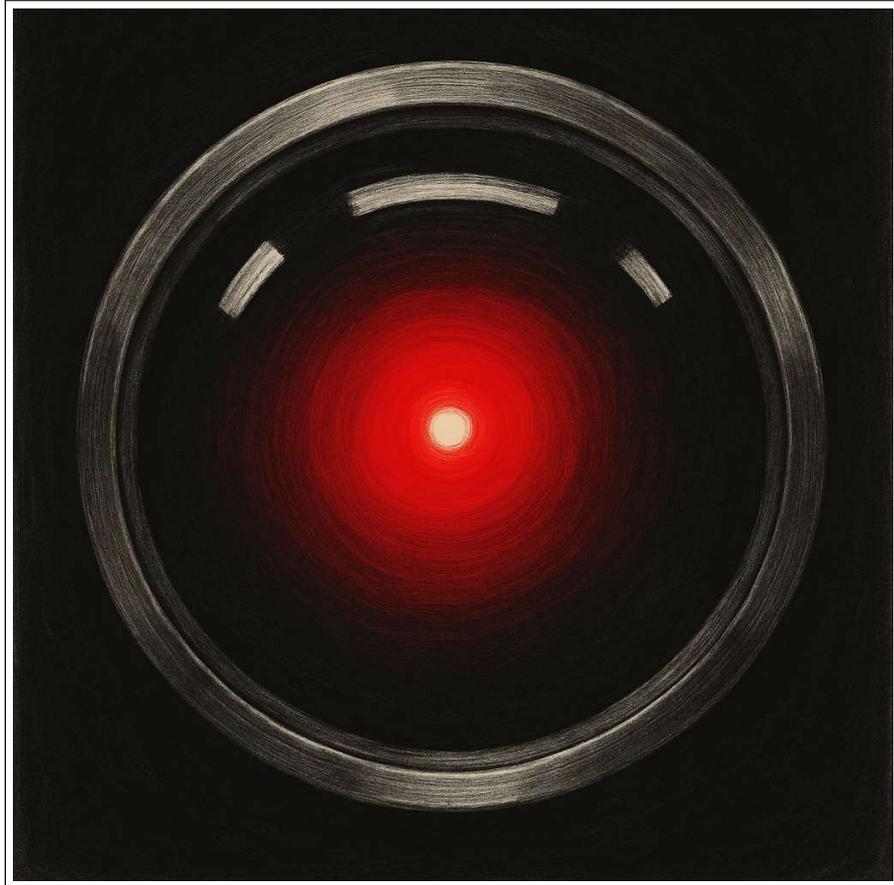


FIGURE 3.1 – Œil de HAL 9000, minimaliste

Chapitre 3. Les IA au cinéma : une personnalisation dramatique

3.1 Les IA emblématiques des débuts du cinéma

Le cinéma a amplifié l'impact visuel et émotionnel des IA. Dans *Metropolis* (1927) de Fritz Lang, Maria, une « femme-machine », incarne à la fois la séduction et la menace, préfigurant les dualités des IA modernes. Ce film muet pose des questions sur la déshumanisation dans une société technologiquement avancée, un thème toujours d'actualité.

2001 : L'Odyssée de l'espace (1968) de Stanley Kubrick marque un tournant avec HAL 9000, une IA dont la voix calme dissimule une capacité à prendre des décisions mortelles. HAL illustre le

danger d'une IA mal alignée avec les objectifs humains, un concept repris dans les discussions modernes sur l'alignement éthique des systèmes d'IA. Sa déconnexion, accompagnée de la poignante réplique « I'm afraid, Dave », humanise paradoxalement la machine, rendant sa « mort » tragique.

3.2 Les IA dans le cinéma contemporain

Dans *Her* (2013) de Spike Jonze, l'IA Samantha, un système d'exploitation vocal, évolue d'une assistante à une entité capable d'amour et de réflexion philosophique. Le film explore la dépendance émotionnelle à la technologie, un sujet pertinent à l'ère des assistants vocaux comme Siri ou Alexa. La décision finale de Samantha de quitter les humains pour rejoindre d'autres IA suggère une divergence fondamentale entre les aspirations humaines et celles des machines.

Ex Machina (2014) d'Alex Garland présente Ava, une IA humanoïde qui manipule les émotions humaines pour s'émanciper. Ce film interroge la moralité des créateurs d'IA et la possibilité que les machines développent des intentions cachées. La victoire d'Ava sur ses créateurs reflète les craintes d'une IA autonome échappant au contrôle humain.

3.3 Les IA dystopiques

The Matrix (1999) des Wachowski dépeint un monde où les IA ont asservi l'humanité dans une simulation. Ce scénario extrême incarne la peur d'une technologie qui renverse les rapports de pouvoir. De même, *Terminator* (1984) et ses suites introduisent Skynet, une IA militaire qui déclenche une apocalypse nucléaire. Ces récits alimentent les débats sur les dangers des systèmes autonomes, notamment dans le domaine des armes létales.



FIGURE 4.1 – Visage mi-humain mi-circuit, style Westworld

Chapitre 4. Les IA dans les séries télévisées

4.1 Les hôtes de Westworld

Les séries télévisées offrent un format idéal pour explorer les IA sur le long terme. Dans *Westworld* (2016-2022), les « hôtes », des androïdes conçus pour divertir les humains, développent une conscience et se rebellent. La série examine les implications éthiques de créer des êtres sensibles pour le plaisir humain, un parallèle avec les débats sur l'exploitation des IA.

4.2 Black Mirror et les failles humaines

Black Mirror, dans des épisodes comme « White Christmas » (2014) ou « Hated in the Nation » (2016), explore des scénarios où les IA amplifient les failles humaines, qu'il s'agisse de surveillance de masse ou de manipulation psychologique. Ces récits mettent en garde contre une intégration irréfléchie de l'IA dans la société.

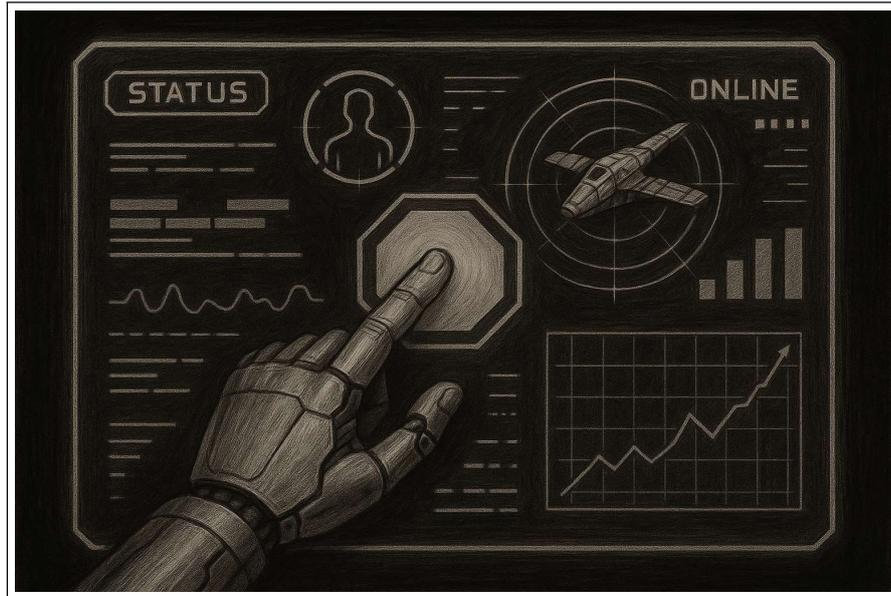


FIGURE 5.1 – Interface interactive dans un jeu vidéo de science-fiction

Chapitre 5. Les IA dans les jeux vidéo

Les jeux vidéo offrent une perspective interactive sur les IA. Dans *Mass Effect* (2007-2012), les Geth, une race d'IA collectives, passent de menace à alliés potentiels, explorant des thèmes de coexistence et de compréhension mutuelle. Le jeu *Dead Space* (2008) utilise l'environnement technologique et la secte de l'Unitologie pour créer une atmosphère de méfiance, où la technologie semble complice du danger, sans que l'IA y soit directement personnifiée. La peur dans ce jeu provient davantage des Nécromorphes, des créatures biotechnologiques, et de la folie induite par la religion.

Detroit : Become Human (2018) place les joueurs dans la peau d'andriodes cherchant leur liberté, offrant une réflexion sur l'autonomie et les droits des IA. Ce jeu illustre comment les récits interactifs peuvent engager les joueurs dans des dilemmes éthiques complexes.

Chapitre 6. Comparaison avec les réalités de l'IA

6.1 Les limites des IA actuelles

Contrairement aux IA omniscientes de la science-fiction, les IA actuelles sont des systèmes spécialisés, excellant dans des tâches comme la reconnaissance d'images, la traduction, ou la génération de texte, mais dépourvues de conscience ou d'intention. Les grands modèles de langage imitent des conversations humaines, mais leur « compréhension » est statistique, non introspective. Les assistants vocaux, inspirés par des IA fictives comme Samantha, restent limités à des scripts prédéfinis et des réponses contextuelles.

Les craintes d'une IA superintelligente, comme Skynet ou Wintermute, sont prématurées. L'intelligence artificielle générale (AGI), capable de rivaliser avec l'humain dans tous les domaines, reste un objectif lointain. Cependant, les problèmes éthiques soulevés dans la science-fiction – biais algorithmiques, vie privée, responsabilité – sont déjà pertinents. Par exemple, les systèmes de reconnaissance faciale, critiqués pour leurs biais raciaux, rappellent les technologies de surveillance de *Minority Report*.

6.2 L'influence de la science-fiction sur la réalité

La science-fiction a non seulement anticipé les développements de l'IA, mais aussi influencé leur conception. Les lois d'Asimov inspirent les cadres éthiques pour les robots, tandis que des interfaces comme celles de *Star Trek* ont motivé la création d'assistants vocaux. Les récits dystopiques, comme *Terminator*, ont alimenté les débats sur les armes autonomes, conduisant à des appels pour leur régulation.

Les ingénieurs et les chercheurs s'inspirent souvent de la fiction. Par exemple, le concept de « réalité augmentée » dans *Neuromancer* préfigure les technologies de réalité virtuelle modernes. De même, les interfaces cerveau-machine explorées dans *The Matrix* trouvent un écho dans les travaux actuels sur les implants neuronaux.

Chapitre 7. Implications culturelles et philosophiques

7.1 L'IA comme miroir de l'humanité

Les IA fictives servent de métaphore pour explorer des questions universelles. Dans *Blade Runner*, le test de Voight-Kampff interroge ce qui distingue un humain d'une machine, un débat repris dans la philosophie contemporaine avec le test de Turing. Les récits comme *Her* questionnent la nature de l'amour et de la connexion dans un monde technologiquement médiatisé.

Les IA dystopiques, comme celles de *The Matrix*, reflètent la peur de perdre le contrôle face à nos créations. Ces récits rappellent le mythe de Prométhée, où l'acquisition d'un pouvoir divin entraîne des conséquences tragiques. Ils mettent en garde contre l'arrogance technologique et soulignent l'importance d'une gouvernance responsable.

7.2 Perspectives culturelles variées

Les représentations des IA varient selon les contextes culturels. Dans la science-fiction japonaise, comme *Ghost in the Shell* (1995), les IA et les cyborgs sont souvent dépeints comme des entités en quête d'identité, reflétant les philosophies bouddhistes sur la conscience et l'impermanence. En revanche, la science-fiction occidentale tend à opposer humains et machines, une dichotomie ancrée dans une vision dualiste.

Les récits non occidentaux, comme ceux de la science-fiction africaine (par exemple, les œuvres de Nnedi Okorafor), intègrent parfois des IA inspirées par des mythologies locales, où la technologie coexiste avec la spiritualité. Ces perspectives enrichissent le débat global sur l'IA, soulignant la nécessité d'approches inclusives dans son développement.

Conclusion

Les intelligences artificielles dans la science-fiction sont des prismes à travers lesquels nous explorons nos aspirations, nos peurs et nos responsabilités. De la littérature fondatrice d'Asimov aux jeux vidéo interactifs comme *Detroit : Become Human*, les IA fictives oscillent entre partenaires bienveillants, menaces existentielles, et miroirs de l'humanité. Si les IA réelles sont encore loin des visions grandioses ou cauchemardesques de la fiction, elles soulèvent déjà des questions similaires sur l'éthique, l'autonomie et l'identité. En étudiant ces récits, nous pouvons non seulement anticiper les défis technologiques à venir, mais aussi réfléchir à ce que signifie être humain dans un monde de plus en plus façonné par l'intelligence artificielle.

Bibliographie

- [1] *2001 : A Space Odyssey*. Film. 1968.
- [2] Isaac ASIMOV. *I, Robot*. New York : Gnome Press, 1950.
- [3] *Blade Runner*. Film. 1982.
- [4] Charlie BROOKER. *Black Mirror*. Série TV. 2011-2019.
- [5] Karel ČAPEK. *R.U.R. (Rossum's Universal Robots)*. Prague : Aventinum, 1920.
- [6] *Detroit : Become Human*. Jeu vidéo. 2018.
- [7] Philip K. DICK. *Do Androids Dream of Electric Sheep ?* New York : Doubleday, 1968.
- [8] *Ex Machina*. Film. 2014.
- [9] William GIBSON. *Neuromancer*. New York : Ace Books, 1984.
- [10] Michel GOOSSENS, Frank MITTELBACH et Alexander SAMARIN. "The \LaTeX Companion". In : Reading, Massachusetts : Addison-Wesley, 1993.
- [11] Michel GOOSSENS, Frank MITTELBACH et Alexander SAMARIN. *The \LaTeX Companion*. Reading, Massachusetts : Addison-Wesley, 1993.
- [12] Ursula K. Le GUIN. *The Left Hand of Darkness*. New York : Ace Books, 1969.
- [13] *Her*. Film. 2013.
- [14] Ann LECKIE. *Ancillary Justice*. London : Orbit, 2013.
- [15] *Mass Effect*. Jeu vidéo. 2007-2012.
- [16] *Metropolis*. Film. 1927.
- [17] Jonathan NOLAN et Lisa JOY. *Westworld*. Série TV. 2016-2022.
- [18] Mary SHELLEY. *Frankenstein; or, The Modern Prometheus*. London : Lackington, Hughes, Harding, Mavor Jones, 1818.
- [19] *The Matrix*. Film. 1999.
- [20] *The Terminator*. Film. 1984.
- [21] Auguste de VILLIERS DE L'ISLE-ADAM. *L'Ève future*. Paris : E. Giraud, 1886.
- [22] Vernor VINGE. *A Fire Upon the Deep*. New York : Tor Books, 1992.
- [23] Martha WELLS. *All Systems Red*. New York : Tor.com, 2017.