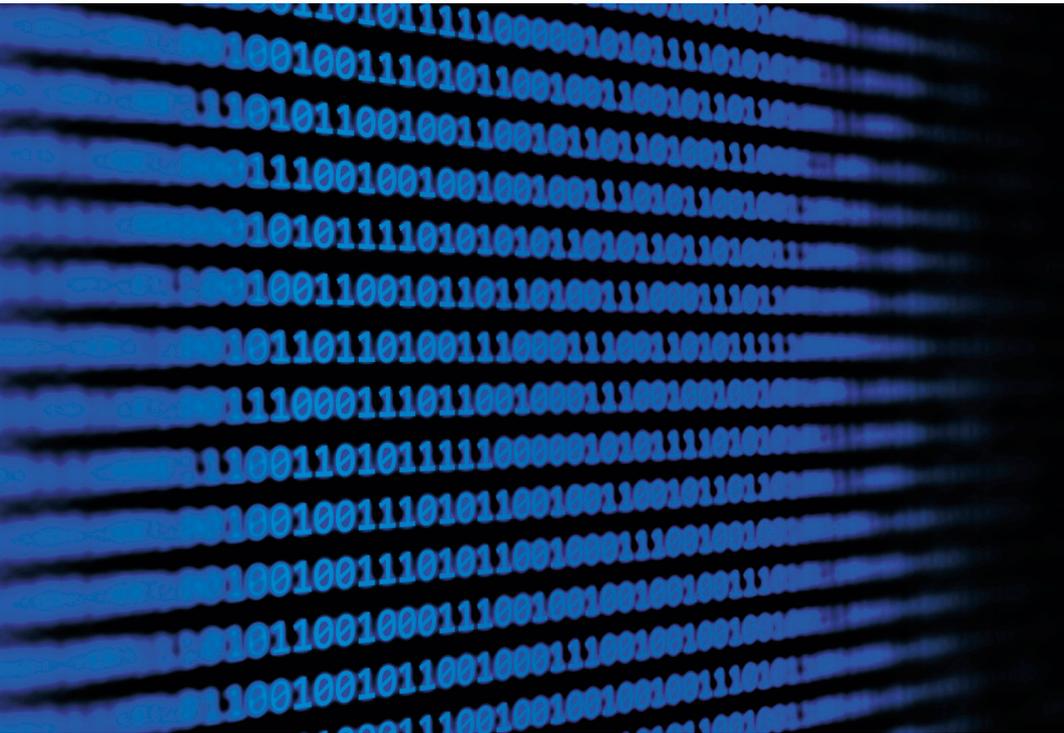


VINCENT BERTHET  
LÉO AMSELLEM

# LES NOUVEAUX ORACLES

**COMMENT LES ALGORITHMES  
PRÉDISENT LE CRIME**



CNRS EDITIONS



# Les nouveaux oracles

Comment les algorithmes  
prédissent le crime



Vincent Berthet  
Léo Amsellem

# Les nouveaux oracles

Comment les algorithmes  
prédissent le crime

**CNRS ÉDITIONS**

15, rue Malebranche 75005 Paris



*« Organiser scientifiquement l'humanité,  
tel est donc le dernier mot de la science moderne,  
telle est son audacieuse mais légitime prétention »*

Ernest Renan,  
*L'avenir de la science : pensées de 1848, 1890.*



# Préface

## L'intelligence artificielle en quête de sens

*par Paul Duan*<sup>1</sup>

L'avènement de l'intelligence artificielle (IA) a vocation à bouleverser la vie des citoyens. Il suscite à la fois de grands espoirs comme de nombreuses craintes – parfois fantasmées, parfois réelles.

Quel sens donner à l'IA ? Cette question se pose avec d'autant plus d'acuité que les algorithmes prennent une place croissante non seulement dans notre vie quotidienne, mais également dans l'organisation de nos sociétés. Beaucoup ont été écrits sur les effets secondaires des nouvelles technologies, à l'instar de l'influence des algorithmes de Facebook sur l'opinion publique, ou encore des nombreuses prises d'intérêts des entreprises privées du numérique dans la santé ou l'éducation. Plus directement encore, les États s'en approprient de plus en plus l'usage au sein de l'action publique, à commencer par les domaines régaliens de la sécurité et de la justice, pour lesquels plusieurs exemples sont donnés dans cet ouvrage.

Le débat est complexe car, d'une part, l'IA reste une technologie *nouvelle*, où les bonnes pratiques en matière

d'algorithmes publics sont découvertes au fur et à mesure des réussites – et des ratés – des implémentations d'États ou d'institutions précurseurs. D'autre part, si ces expérimentations partent généralement d'une volonté de mettre l'innovation au service de problèmes sociétaux, elles sont fréquemment soumises à des contraintes budgétaires ou temporelles qui en biaisent la mise en œuvre, notamment lorsque la réduction des dépenses est, plus qu'une contrainte, l'objectif premier de l'usage de l'IA.

Ces facteurs limitent la capacité de la société à s'emparer du sujet afin de prendre le recul nécessaire pour repenser l'organisation des services publics. Le risque est de ne voir en l'IA qu'une simple façon d'automatiser des processus déjà existants, comme de remplacer un médecin ou un professeur d'école, c'est-à-dire de vouloir simplement substituer des algorithmes aux humains.

\*\*\*

Pour répondre à la question du sens de l'IA, il faut sortir du discours purement technique, et s'interroger plutôt sur ses aspects sociétaux : comment créer un système de justice plus juste *et* plus efficace ? Où mettre le curseur, lorsque le second se fait au détriment du premier, ou *vice-versa* ? Comment concilier enjeux sécuritaires et protection de la vie privée ? Cette analyse demande une approche nuancée et pragmatique, car elle est politique. La réserver au domaine technique, comme cela est trop fréquemment le cas, c'est encourir le risque de manquer le débat important – celui des valeurs –, et de réduire à une joute d'experts ce qui devrait être un véritable questionnement démocratique.

Pour ce faire, commençons par démystifier ce qui est entendu par « intelligence artificielle ». Le terme évoque de nombreux fantasmes, notamment sur le caractère intelligent des machines. Cette notion, généralement regroupée sous le terme « intelligence artificielle forte » ou « intelligence artificielle générale », est intéressante du point de vue philosophique – et a tendance à capturer une grande partie de l’imagination publique autour de ces enjeux. Mais la réalité d’aujourd’hui est que l’IA désigne surtout des algorithmes d’optimisation capables uniquement de répondre à des questions simples et formulées de manière très spécifique.

Il ne faut pas pour autant minimiser l’impact structurel de ces technologies : le cycle de Gartner explique justement que les effets de court terme d’une nouvelle technologie ont tendance à être surestimés, alors que ceux de long terme sont sous-estimés. Aujourd’hui déjà, les algorithmes développés par d’importantes entreprises de la Silicon Valley leur permettent, pour certaines applications, d’avoir un impact qui semble parfois dépasser celui des États et de la puissance publique. Car l’effet de levier pour toucher leurs clients par millions est réel et concret<sup>2</sup>. Si ces algorithmes ne sont pas « intelligents » au sens courant du terme, et ne peuvent répondre qu’à des questions précises, ils restent néanmoins extrêmement efficaces car ils permettent de le faire avec une rapidité et une échelle inédites. C’est là la source des débats qui entourent les algorithmes de Facebook, rendant possible l’optimisation des contenus affichés pour leurs milliards d’utilisateurs, et leur donnant ainsi un pouvoir d’influence démesuré sur l’opinion publique. Qu’il s’agisse de biais involontaires ou d’effets pervers naturellement liés au modèle économique de Facebook – plus un contenu engrange de vues, plus il rapporte de revenus publicitaires –,

ces algorithmes ont déjà aujourd'hui un impact fort sur nos sociétés.

Dans le même temps, nos États ne semblent pas encore dotés du corpus idéologique permettant de penser cette nouvelle donne numérique dans le service public. Pour le construire, il est nécessaire de sortir du manichéisme : en matière d'IA, l'opinion tombe généralement dans un clivage binaire où les algorithmes sont soit la solution miracle à tous nos maux, soit une technologie survendue et mal maîtrisée aux effets principalement néfastes.

Il faut reconnaître aux nouvelles technologies la capacité d'avoir un impact positif en rendant plus efficace des services d'intérêt général existants, ou encore en rendant possible ce qui ne l'était pas auparavant. Bien utilisée, l'IA peut permettre de prédire et d'éviter les feux de forêts en analysant les images satellites. Elle peut aider à personnaliser les parcours de santé en analysant les dossiers médicaux tout comme les parcours d'apprentissage des élèves. Elle peut également permettre la démocratisation de domaines auparavant réservés aux experts, mettant ainsi le pouvoir entre les mains d'usagers à même aujourd'hui de détecter les signes de cancer sur leur peau depuis leur smartphone, par exemple.

Il faut cependant noter les limites de ces technologies. Parfois, leur puissance réelle est plus faible qu'on ne le pense, comme dans l'exemple de la police prédictive analysé dans le présent ouvrage ; parfois, elles sont dangereuses justement de par leur trop grande puissance, notamment en matière de reconnaissance faciale. De même, elles peuvent être sujettes à des risques de biais introduits par erreur, comme lorsque

les IA utilisées en justice considèrent systématiquement les minorités raciales comme plus dangereuses. Elles peuvent aussi amplifier les biais – voire les desseins – de leurs créateurs, posant ainsi la question de la concentration des pouvoirs aux mains des concepteurs d'intelligences artificielles.

Il est donc crucial de ne pas seulement poser la question du « comment », mais celle bien plus importante du « pourquoi ». En tant que *data scientist* ayant travaillé dans la Silicon Valley et entrepreneur social, j'ai eu l'occasion de voir à de maintes reprises comment ce questionnement permanent pouvait faire la différence entre les algorithmes ayant des effets négatifs, et ceux permettant d'imaginer des solutions innovantes aux enjeux de société. Par exemple, les algorithmes peuvent faire preuve de biais racial dans le domaine de la justice, mais ces biais existent également chez l'humain. Comment alors imaginer des solutions et des designs alliant humain et machine, de manière à compenser les faiblesses de l'un et de l'autre, et d'en amplifier les forces ?

Ma conviction est que cette redistribution des cartes est une formidable opportunité de repenser le rôle de l'action publique, et d'utiliser les nouvelles technologies pour donner plus de pouvoir et de meilleurs services aux citoyens. C'est aussi l'occasion pour ces derniers de créer à partir de ces technologies, des innovations sociales qui contribuent au bien public. Pour ce faire, le rôle de l'État est double. Il doit en premier lieu servir de garant pour réguler les usages privés ou publics de l'IA pour s'assurer que ceux-ci sont conformes à l'intérêt général, et en deuxième lieu, contrôler les dérives lorsqu'elles ont des effets secondaires néfastes – par erreur ou par design. Mais aussi et surtout, il doit

susciter des incitations positives pour favoriser l'émergence de nouveaux services utilisant les algorithmes pour le bien commun. Ceux-ci n'ont pas besoin d'être uniquement créés à l'initiative des États ; ils peuvent et doivent également être développés par la société civile, comme autant de services publics citoyens. Cela nécessite de concevoir une réelle doctrine autour de l'utilisation positive de l'IA.

Le présent livre de Vincent Berthet et Léo Amsellem pose de manière très juste la question des technologies prédictives, en discutant de façon nuancée de leurs bienfaits comme de leurs risques au travers de nombreux exemples concrets. Parce qu'il ne tombe pas dans la caricature ni dans la béatitude, parce qu'il résume simplement des enjeux complexes, cet ouvrage pose les bases d'une réflexion commune sur la place de l'IA dans le service public. En ce sens, il constitue un outil important pour nourrir le débat public. Car la question du rôle des algorithmes dans la société ne relève pas que du domaine de la technique : il s'agit surtout d'une question politique. À nous, collectivement, de nous en saisir.

# Introduction

« – Pourquoi l'avez-vous rattrapée ?

– Parce qu'elle allait tomber.

– Vous en êtes certain ?

– Oui.

– Mais elle n'est pas tombée, vous l'avez rattrapée. Le fait que vous ayez empêché que cela se produise n'enlève rien au fait que cela allait se produire. »

*Minority Report* de Steven Spielberg, 2002.

La film *Minority Report* de Steven Spielberg sorti en 2002, tiré de la nouvelle éponyme de Philip K. Dick, décrit un monde dystopique en 2054 dans lequel Washington D.C. est parvenue à éradiquer la criminalité. La ville utilise un système où trois humains dotés du don de précognition (les « précogs ») alertent utilement une unité de police spécialisée (« Précrime ») en cas de vision de crimes futurs, qui peut alors intervenir avant leur survenance. Au-delà de la fiction, *Minority Report* met en jeu un principe clé : *agir sur le fondement d'une prédiction avant qu'un événement se produise*. Pour illustrer ce principe, John Anderton, le chef de l'unité, envoie une boule rouler sur une planche à un agent

du ministère de la Justice rendant visite à l'unité Précrime. L'agent l'attrape alors au vol. Cette *action* est la conséquence de son *anticipation* du fait qu'elle allait tomber.

Anticiper l'avenir permet de prendre ses dispositions dans le présent de façon à obtenir la situation la plus souhaitable. Anticiper la demande aide une entreprise à déterminer le niveau de production optimal. Anticiper la croissance permet à un Gouvernement de fixer la politique budgétaire adaptée. Prédire qu'un détenu condamné pour une agression violente a de fortes chances de récidiver oriente la décision du juge.

Si les commissariats de police ne sont pas encore dotés d'une section Précrime et les précogs ne font pas encore partie de ce monde, le principe d'actions préventives est déjà à l'œuvre en matière de sécurité. Il sous-tend par exemple le *Children Act* adopté au Royaume-Uni en 1989, une loi qui vise à protéger les enfants d'éventuels actes de maltraitance dans leur foyer en les en retirant lorsque les services sociaux identifient une « probabilité de faire du mal » trop élevée. C'est cette même logique qui a dicté la loi française du 25 février 2008 sur la rétention de sûreté prévoyant des mesures de privation de liberté après l'exécution de sa peine pour un détenu présentant « une particulière dangerosité caractérisée par une probabilité très élevée de récidive ». C'est par ailleurs une des missions premières des services de renseignement, leaders de la prédiction par algorithme, que d'estimer la probabilité qu'ont certains individus ou de certains groupes de commettre des actes criminels, en fonction des informations qu'ils peuvent recueillir.

Si la logique de la prédiction tend à se généraliser, une question essentielle s'impose : d'où viennent les prédictions et dans quelle mesure sont-elles fiables ? Lorsqu'elles proviennent des précogs dans *Minority Report*, les prédictions correspondent à de pures visions de l'avenir – à l'instar des oracles dans la mythologie grecque ou des palantirs, les pierres de clairvoyance, dans l'univers de J. R. R. Tolkien. Le système de *Minority Report* repose sur la fiabilité de ces prédictions : agir par anticipation est pertinent seulement dans la mesure où le futur est correctement anticipé (dans la fiction, les prémonitions des précogs ne sont pas infailibles : il arrive que l'un d'eux voie une scène future différemment des deux autres).

Dans la mythologie, Apollon est le dieu de la lumière, des arts, des chants, de la beauté masculine, et également le dieu des oracles. La plupart des sites oraculaires de grande envergure lui sont dédiés, au premier rang desquels celui de Delphes, le plus renommé. La Pythie, prêtresse delphique enivrée par de mystérieuses exhalaisons, y délivre les prophéties divinatoires que lui insuffle le dieu Apollon. C'est le poète grec Pindare qui établit un lien entre ces oracles comme visions de l'*avenir* et le *passé*. Dans son recueil d'odes intitulé *Pythiques*, Pindare prête ces paroles au plus célèbre des centaures Chiron, adressées au dieu Apollon : « Toi qui sais le terme fatal de toutes choses et tous les chemins qu'elles prennent ; toi qui peux compter les feuilles que la terre fait pousser au printemps et les grains de sable que, dans la mer ou les fleuves, roulent les vagues et les souffles des vents ; toi qui vois clairement l'avenir et son origine ! » La connaissance du passé, de « l'origine », serait donc indissociable de toute intuition du futur.

Aujourd'hui, la science a donné raison à Pindare. Les prédictions les plus précises proviennent désormais d'algorithmes qui analysent de larges quantités de données passées afin de tirer parti des régularités statistiques qu'elles contiennent. L'intuition de l'avenir, mystique et infondée, est délaissée au profit de la prédiction algorithmique, bien plus efficace. Cette nouvelle donne est la conséquence d'une révolution technologique majeure : l'avènement de l'intelligence artificielle (IA), appuyée sur les données massives (big data). Le traitement par algorithmes d'une quantité importante de données permet de repérer des répétitions, des tendances dans les événements. Dès lors qu'une tendance stable existe, le passé prédit – en partie – l'avenir. Les *algorithmes prédictifs* sont donc une formidable caisse de résonance de la logique prédictive. Celle-ci s'applique désormais pour anticiper aussi bien une maladie (médecine prédictive), une panne (maintenance prédictive), le départ d'un employé (recrutement prédictif), ou encore la blessure d'un joueur (analyse prédictive dans le sport).

La gestion par prédiction algorithmique, résolument issue du secteur privé ces dernières décennies, impacte également l'État même dans ses fonctions les plus régaliennes, au premier rang desquelles la sécurité.

La prédiction algorithmique est tout d'abord utilisée dans la justice pénale, aux États-Unis notamment, pour estimer le risque de récidive d'un détenu. En comparant le cas d'espèce à des millions de cas antérieurs suivant un ensemble de critères (l'âge, les antécédents judiciaires, le nombre d'arrestations antérieures, etc.), des logiciels calculent son risque de récidive. Une telle information éclaire utilement la décision de libération conditionnelle prise par le juge.

La prédiction algorithmique est aussi utilisée par la police dans son action contre la délinquance afin de connaître à l'avance les probabilités d'occurrence d'un crime ou d'un délit dans un certain lieu ou par un certain individu. Ces outils de « police prédictive », originaires de la Silicon Valley, permettent une meilleure gestion et affectation des unités de police. Les algorithmes prédictifs sont enfin utilisés par les services de renseignement dans la lutte contre le terrorisme afin de détecter les « signaux faibles » contenus dans les données, révélateurs de menaces potentielles. À l'heure actuelle, ces moyens mis en œuvre en matière de sécurité intérieure et extérieure proviennent essentiellement d'entreprises privées, en particulier *Palantir Technologies* qui développe et vend aux entreprises et aux États des solutions d'aide à la décision et à la prédiction fondées sur l'exploitation de données. Cette société, créée par Peter Thiel, a ainsi proposé ses services à la police de Los Angeles et de la Nouvelle Orléans, à la U.S. Army en Irak et en Afghanistan, ou encore à plusieurs services de renseignement américains ou étrangers.

Quand la police et la justice appuient leur action sur des prédictions issues d'algorithmes sur le comportement criminel futur, c'est le premier pas vers *Minority Report 2.0*. Que penser de l'usage des technologies prédictives par l'État ? Par leur capacité à être déployés à grande échelle, les algorithmes offrent un véritable effet de levier, pour le meilleur et pour le pire.

D'un côté, la prédiction algorithmique charrie la promesse d'une efficacité accrue, d'une sécurité renforcée. Par exemple, mieux prédire la récidive des détenus permet de réduire le taux d'incarcération (et la dépense publique correspondante) tout en améliorant la sécurité publique. Sous

cet angle, les technologies prédictives favorisent un retour à l'essence de l'État, car assurer la sécurité des citoyens – en contrepartie du renoncement à une part de liberté individuelle – est bien l'objectif qui fonde le contrat social. Observant l'éventualité de l'émergence d'un État omniscient par la quantité d'informations qu'il serait capable de traiter, le père dominicain Dubarle avait anticipé dès 1948 « le surgissement d'un prodigieux Léviathan politique<sup>1</sup> ». Anticipant l'avènement d'une gouvernance algorithmique, syndrome de la « quantophrénie » de pouvoirs publics de plus en plus technocratiques, Dubarle avait compris comment la technologie allait affecter l'action publique.

D'un autre côté, les technologies prédictives présentent des risques à la mesure de leur potentiel. En particulier, Si un algorithme présente un biais à l'encontre d'un groupe social particulier, il se produira alors de façon systématique à grande échelle. De tels algorithmes imparfaits agiraient alors comme des « armes de destruction massive » selon l'expression de Cathy O'Neil<sup>2</sup>. Par ailleurs, la prédiction algorithmique fait peser un risque considérable sur les libertés fondamentales et, au-delà, sur les principes et les valeurs mêmes qui fondent notre société. Prenons l'exemple de la prédiction algorithmique du risque de récidive des détenus. Peut-on juger un détenu sur la base d'un calcul statistique de son risque de récidive ? Est-il juste de considérer un cas particulier sur la base de cas similaires passés ? Cette pratique traduit un changement de paradigme, le glissement de la notion de « culpabilité à celle de dangerosité<sup>3</sup> », syndrome des aspirations sécuritaires de nos sociétés.

Trois critères sont mobilisés pour évaluer les algorithmes d'IA : leur efficacité, leur équité (leur capacité à traiter les

# Table des matières

<b>Préface. L'intelligence artificielle en quête de sens, <i>par Paul Duan</i> .....</b>	<b>9</b>
<b>Introduction .....</b>	<b>15</b>
<b>Chapitre 1. L'avenir dans une formule, <i>Vincent Berthet</i> .....</b>	<b>23</b>
Prédiction humaine vs. Prédiction statistique....	24
Comment l'IA a décuplé la prédiction statistique....	36
De l'usage des algorithmes prédictifs .....	43
Algorithmes prédictifs : mythes et réalités .....	47
<b>Chapitre 2. Prédire et juger, <i>Vincent Berthet</i>.....</b>	<b>55</b>
Évaluer le risque de récidive d'un détenu.....	57
L'émergence des outils d'évaluation du risque...	63
Réduire l'erreur humaine et améliorer l'efficacité de la justice .....	67
Comment fonctionnent les outils d'évaluation du risque.....	71
Des algorithmes imparfaits .....	76
Les algorithmes prédictifs sont-ils solubles dans le droit ? .....	85
Droit et algorithmes : le spectre de la justice prédictive .....	90

<b>Chapitre 3. Anticipés coupables, <i>Léo Amsellem</i>.....</b>	101
Les devins des soldats du feu.....	104
Quand la réalité rattrape la science-fiction .....	107
L'avènement de la police prédictive.....	109
La police prédictive : une efficacité limitée .....	118
Les discrets pionniers de la prédiction.....	126
<b>Chapitre 4. La reconnaissance faciale : surveillance libérée, liberté surveillée, <i>Léo Amsellem</i> .....</b>	143
Comment s'utilisent reconnaissance et comparaison faciales ? .....	148
Reconnaissance faciale prédictive : les apparences sont-elles trompeuses ?.....	152
Ces systèmes sont-ils pleinement opérationnels aujourd'hui ? .....	156
Le spectre de Big Brother plane .....	161
Crédit social chinois ou panoptique benthamien ?.....	167
<b>Chapitre 5. Prédiction, sécurité et liberté, <i>Vincent Berthet et Léo Amsellem</i> .....</b>	179
Le compromis efficacité-liberté.....	180
Principes <i>vs.</i> pragmatisme .....	182
Choisir la sécurité pour les bonnes raisons.....	187
Sortir du compromis efficacité-liberté.....	190
Garder la main sur les algorithmes prédictifs ....	192
<b>Conclusion : Vers une société de la prédiction ?.....</b>	199
<b>Bibliographie .....</b>	209
<b>Remerciements .....</b>	217
<b>Notes .....</b>	219