INTERNATIONAL >
CNIL : protection des données personnelles

FOCUS > Le service de renseignement criminel de la gendarmerie nationale

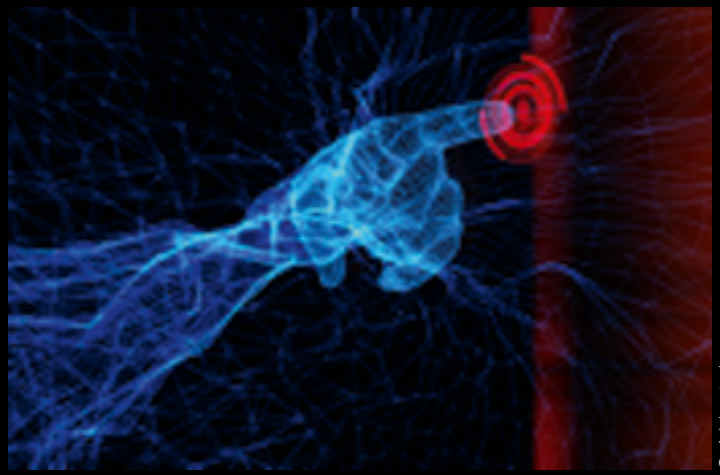
DOSSIER > Impact des algorithmes et de l'exploitation des données massives dans les États de droit

REVUE
de la gendarmerie nationale

REVUE TRIMESTRIELLE / JUIN 2018 / N° 261 / PRIX 6 EUROS
RÉALISÉE SOUS LA DIRECTION SCIENTIFIQUE DE JÉRÔME LAGASSE

Algorithmes
et espaces normatifs

retour sur images NUMÉRO 260



© Adobe stock

FIC 2018 – HYPERCONNEXION ET RÉSILIENCE

Le cyberespace est devenu un enjeu majeur et un vecteur de stratégies commerciales, relationnelles et de domination militaire ou politique. L'absence fondamentale d'une régulation des techniques employées et des contenus modifie les rapports entre les Etats et entre ces derniers et leurs citoyens. Elle engage les relations entre les personnes qui voient poindre un nouveau couple homme-machine porteur d'une compétition au regard du développement de l'intelligence artificielle. Une régulation pourrait se faire par l'instauration d'un partenariat international où les acteurs privés, les groupements d'états et les représentations des usagers pourraient instaurer un système préservant la confiance numérique, la protection des données personnelles et un encadrement strict de l'intelligence artificielle.

RETROUVEZ EN PAGE 114 LES DÉVELOPPEMENTS QUI VOUS SONT PROPOSÉS PAR DES OFFICIERS DU SERVICE CENTRAL DE RENSEIGNEMENT CRIMINEL (PÔLE JUDICIAIRE DE LA GENDARMERIE NATIONALE)

© Semisatch

La question de l'emploi des algorithmes porte des confusions tenaces suggérées par une littérature qui vise à faire perdurer l'idée d'une civilisation qui aurait délégué le soin de présider à ses destinées à des entités électrotechniques opaques et irresponsables juridiquement.

Dans un contexte mondialisé, la sphère politique et les milieux économiques ont besoin d'acquérir rapidement le contexte de leurs décisions. Le traitement de données massives, la disponibilité d'informations empreintes de sens et la nécessité d'appuyer des décisions sur un référentiel exact militent pour inclure des algorithmes dans le cycle de la décision.

Il faut cependant évaluer le poids réel d'une intelligence artificielle dans la genèse des recommandations algorithmiques et de la valeur ajoutée d'analystes. Ces derniers détiennent une expertise pour paramétrer l'acquisition des données, extraire les informations utiles et contextualiser des recommandations viables et fiables. Ils caractérisent la plus-value humaine qui réside dans l'intuition, la perception d'un environnement et en dernier ressort une imputabilité de l'acte.

Une norme internationale doit encadrer l'emploi des algorithmes notamment quant aux traitements des données personnelles. Elle doit placer l'homme au centre du dispositif en le protégeant d'un caractère impératif des recommandations algorithmiques et en le désignant comme le seul responsable des actes de sa « créature » qui ne peut être raisonnablement décisionnaire notamment en matière d'ordre public et d'actes judiciaires. La question de la loyauté et de la transparence des traitements est au cœur de cette problématique de la protection des personnes et de leur vie privée.

COL(ER) Philippe Durand,
rédacteur en chef

SOMMAIRE

NUMÉRO 261

DOSSIER

Algorithmes et droit

© Personal par Vege - Adobe Stock

ALGORITHMES ET OBLIGATION DE TRANSPARENCE ET DE LOYAUTÉ

Les décisions publiques sont concernées par l'utilisation d'algorithmes utilisant pour leur fonctionnement des données à caractère personnel. Les concepteurs des algorithmes sont soumis aux obligations de transparence, de minimisation des données ou de réalisation d'une analyse d'impact relative à la protection des données (AIPD). Les décisions fondées exclusivement sur un algorithme sont encadrées même si la norme dégage quelques exceptions qui préservent toutefois la transparence, l'explicitation de la logique algorithmique et l'information de la personne concernée par un traitement.

Il subsiste des risques liés à une perception d'infaillibilité des algorithmes et des biais pouvant conduire à des discriminations induisant des décisions erronées ou perfectibles en matière d'ordre public. En conséquence, il faut qu'un algorithme soit loyal et qu'il soit régulièrement validé par les acteurs de la chaîne algorithmique.

En ce sens, la CNIL, a réalisé une déclinaison opérationnelle de six recommandations adressées tant aux pouvoirs publics qu'aux diverses composantes de la société civile.

Algorithmes et protection
des données personnelles

Questions à Mahalia Galie-Blanze. Propos recueillis par le rédacteur en chef de la Revue.

Dans notre monde marqué par toujours plus de numérique, les algorithmes sont à l'œuvre dans tous les champs de la vie quotidienne. Ces algorithmes nécessitent d'exploiter de nombreuses données dont souvent des données personnelles.

En quoi les algorithmes sont-ils concernés par les règles relatives à la protection de données personnelles ?

MAHALIA GALIE-BLANZE

Affaires régaliennes et des collectivités territoriales. Direction de la conformité. C N I L

Les algorithmes, qui sont une façon de décrire de manière très détaillée des procédés répétitifs afin de les exécuter de façon optimisée, occupent dans tous les champs de la vie quotidienne une place de plus en plus importante, bien qu'invisible.

Résultats de requêtes sur un moteur de recherche, fils d'actualité sur les réseaux sociaux, pilotage automatique d'avions, ordres financiers passés par des robots sur les marchés, diagnostics médicaux automatiques : dans tous ces domaines, des algorithmes sont à l'œuvre.

Les décisions publiques sont également largement concernées par cette utilisation croissante des algorithmes. Classement des candidatures pour affecter les lycéens à l'Université (dispositif *Parcoursup*), classement des passagers aériens à risques grâce à un outil de *scoring* (système API PNR) ou encore logiciel de *datamining* pour lutter contre la fraude fiscale (CFVR) ne sont que quelques exemples du maniement des algorithmes par l'administration.

Ces différents algorithmes utilisent pour leur fonctionnement des données à caractère personnel. Ils constituent dès lors des traitements automatisés soumis aux différentes règles relatives à la protection des données personnelles qui sont le règlement européen sur la protection des données personnelles, dit RGPD, la directive dite « police justice » et la loi « Informatique et Libertés », récemment modifiée pour tenir compte des évolutions de la réglementation européenne.

En outre, de nombreux algorithmes consistent à produire un profilage individualisé – c'est-à-dire un traitement automatisé de données ayant pour finalité d'évaluer certains aspects personnels relatifs à une personne physique afin d'émettre un jugement sur sa personne ou de tirer des conclusions à son égard. Toutefois, algorithme et profilage ne sont pas strictement identiques, un algorithme pouvant consister en un traitement automatisé de données personnelles n'ayant pas pour finalité un profilage.

© Technology Research par Kentoh - Adobe stock

Les algorithmes comportant des traitements de données personnelles sont étroitement encadrés juridiquement sur une base de loyauté, de transparence et d'accès à une logique.

Quelles sont les principales règles de la protection des données personnelles applicables aux algorithmes ?

Les algorithmes qui recourent à des données personnelles sont soumis aux obligations générales concernant tout traitement de données, telles que les règles relatives à la base légale du traitement, à la transparence, à la minimisation des données ou la réalisation d'une analyse d'impact relative à la protection des données (AIPD).

À titre d'illustration, la réalisation d'une AIPD est nécessaire lorsque le traitement est susceptible d'engendrer des risques élevés pour la vie privée et notamment en cas « *d'évaluation systématique et approfondie d'aspects personnels […] fondée sur un traitement automatisé […] et sur la base de laquelle sont prises des décisions produisant des effets juridiques à l'égard d'une personne physique ou l'affectant de manière significative* ».

En matière d'algorithmes, cette disposition trouvera fréquemment à s'appliquer puisqu'il est souvent recouru à ceux-ci afin d'évaluer certains aspects d'une personne en vue de prendre une décision produisant de tels effets.

En outre, des dispositions spécifiques concernent les décisions produisant des effets juridiques ou similaires, entièrement fondées sur un traitement automatisé. Bien que ne mentionnant pas les algorithmes, ces dispositions les concernent directement, le traitement automatisé en question étant nécessairement un algorithme.

Quelles sont les règles spécifiques pour les décisions exclusivement fondées sur un algorithme ?

L'article 22 du RGPD pose le principe selon lequel une personne a le « *droit de ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire* ». Cette interdiction est une garantie importante pour les personnes concernées.
Elle figurait déjà en 1978, dans la loi « Informatique et Libertés ».

Il existe néanmoins plusieurs exceptions au principe d'interdiction des décisions

entièrement fondée sur un traitement automatisé :

- lorsque la décision est nécessaire à la conclusion ou à l'exécution d'un contrat ;

- lorsqu'elle est fondée sur le consentement explicite de la personne concernée ;

- lorsqu'elle est autorisée par le droit de l'Union ou le droit de l'État membre.

Les exceptions relatives au contrat et au consentement sont assorties, par le RGPD, de deux catégories de garanties : les premières relatives à la transparence, les secondes à la sauvegarde des droits et libertés et des intérêts légitimes de la personne concernée.

S'agissant de la transparence, le règlement prévoit une obligation d'information particulière à l'égard des personnes concernées par une décision entièrement automatisée, le responsable de traitement devant indiquer l'existence d'une telle prise de décision et les informations utiles concernant la logique sous-jacente, ainsi que l'importance et les conséquences prévues de ce traitement pour la personne concernée. De même, au titre de son droit d'accès, la personne faisant l'objet d'une telle décision peut obtenir les informations utiles concernant la logique sous-jacente du traitement, ainsi que l'importance et les conséquences de ce traitement.

Certains considérants du RGPD insistent en outre sur la nécessaire application des principes de transparence en matière de profilage et ce même en l'absence d'une décision exclusivement fondée sur un traitement automatisé. Le considérant 60 précise ainsi que la personne concernée devrait être informée de l'existence d'un profilage et des conséquences de celui-ci et le considérant 63 indique que, sous réserve de ne pas porter atteinte aux secrets protégés par la loi, le droit d'accès devrait permettre à la personne concernée de connaître la logique qui sous-tend le traitement automatisé et les conséquences que ce traitement pourrait avoir, au moins en cas de profilage.

L'esprit du RGPD, qui renforce les droits des personnes en matière de données personnelles, pourrait conduire à étendre ces exigences à tout traitement algorithmique contribuant à une prise de décision individuelle quand bien même il ne serait pas destiné à produire un profilage. La transparence est en effet un principe cardinal en matière d'algorithme.

S'agissant des mesures de sauvegarde des droits et libertés et des intérêts légitimes de la personne concernée, le RGPD prévoit que le responsable de traitement doit, a minima, permettre à la personne concernée d'obtenir une intervention humaine, d'exprimer son point de vue et de contester la décision.

Le droit d'obtenir une intervention humaine ne doit pas s'entendre comme impliquant nécessairement une intervention préalable à toute prise de décision. Une intervention humaine a posteriori, dès lors qu'elle peut néanmoins permettre de changer la décision, peut constituer une garantie satisfaisante.

(1) À l'heure où a été rédigé cet article, la nouvelle loi « Informatique et liberté » n'a pas encore été définitivement adoptée par le Parlement. Les développements la concernant sont donc formulés sous réserve d'éventuelles modifications du projet de loi.

La loi « Informatique et Libertés » reprend les exceptions relatives au contrat et au consentement et les complète en autorisant en outre que des décisions administratives soient exclusivement fondées sur un traitement automatisé dès lors qu'elles respectent l'article L. 311-3-1et les dispositions relatives au droit au recours administratif du code des relations entre le public et l'administration (CRPA) [1].

L'article L. 311-3-1 précité prévoit que toute administration recourant à un algorithme pour fonder une décision individuelle doit en informer l'usager par une mention sur ladite décision et fournir, si l'usager en fait la demande, certaines informations sur les caractéristiques de l'algorithme (article R. 311-1-3 du CRPA).

La loi « Informatique et Libertés » précise en outre que cette mention d'information doit figurer sur la décision à peine de nullité et que le traitement ne doit pas comporter de données sensibles.

Enfin, cette disposition prévoit que le responsable de traitement doit s'assurer de la maîtrise du traitement algorithmique et de ses évolutions afin de pouvoir expliquer, en détail et sous une forme intelligible, à la personne concernée la manière dont le traitement a été mis en œuvre à son égard.

Les garanties encadrant la prise de décision individuelle, par l'administration, sur le fondement exclusif d'un algorithme sont ainsi relatives à la transparence *a priori* – information à peine de nullité – et *a posteriori* – information sur le fonctionnement de l'algorithme et son application à l'intéressé. Sur ce point, elles rejoignent les dispositions du RGPD. En

revanche, les mesures de sauvegarde des droits et libertés et des intérêts légitimes de la personne concernée précitées ne sont pas reprises par le nouvel article 10 de la loi « Informatique et Libertés ».

Le respect de cet encadrement juridique apparaît-il suffisant à la CNIL pour assurer un usage éthique des algorithmes ?

Consciente que, au-delà de cet encadrement juridique, l'utilisation croissante des algorithmes et de l'intelligence artificielle soulevait des enjeux éthiques, la CNIL a décidé d'animer, en 2017, un débat public sur le sujet et publié une synthèse qui présente des pistes de solutions pour répondre à ces enjeux.

Parmi les principaux risques identifiés figurent la confiance excessive dans les résultats de l'algorithme perçu comme infaillible, les biais, volontaires ou non, conduisant à des discriminations et l'enfermement algorithmique.

S'agissant des biais, on peut par exemple rappeler que de nombreux dispositifs dits de « prédiction du crime » se focalisent souvent sur certains types d'infractions mesurables par l'outil algorithmique au détriment d'autres conduisant ainsi à une concentration des effectifs policiers dans certains quartiers et un relatif délaissement d'autres.

Concernant l'enfermement algorithmique, on peut évoquer les offres de contenus en matière de médias d'information ou culturels, basées sur un algorithme. Si de tels contenus doivent en principe permettre à une personne de s'ouvrir sur le monde, les algorithmes en matière de promotion médiatique, dès lors qu'ils personnalisent l'offre en fonction des goûts et opinions de la personne, risquent au contraire de cantonner aux mêmes horizons voire de conduire à une polarisation des contenus autour de visions dominantes, en opposition avec l'objectif initial de diversité.

Pour faire face à ces risques et construire une intelligence artificielle au service de l'homme, deux principes apparaissent comme fondateurs.

En premier lieu, le principe de loyauté qui a pour objet de lutter contre les biais et requiert de veiller, par des mesures concrètes, à ce qu'un algorithme ne fasse que ce qu'il est censé faire et le fasse bien.

INTERNATIONAL
ALGORITHMES ET PROTECTION DES DONNÉES PERSONNELLES

En second lieu, le principe de vigilance qui a pour but de prendre en compte et de contrebalancer la confiance excessive accordée aux algorithmes et implique d'organiser, par des procédures et mesures concrètes, une forme de questionnement régulier et méthodique des algorithmes de la part de l'ensemble des acteurs de la chaîne algorithmique.

Ces principes ont fait, dans la synthèse de la CNIL, l'objet d'une déclinaison opérationnelle sous la forme de six recommandations adressées tant aux pouvoirs publics qu'aux diverses composantes de la société civile.

L'AUTEURE

Mahalia Galié-Blanzé est docteur en droit privé. Elle a exercé en cabinet d'avocat au Conseil d'État et à la Cour de cassation avant de rejoindre la CNIL, en 2014. Elle travaille au sein du service des affaires régaliennes et des collectivités territoriales de la direction de la conformité, en particulier sur les questions relatives à l'éducation nationale et l'enseignement supérieur, l'immigration et aux professions réglementées.


© USA police officer - Jonathan Stutz

ALGORITHMES PRÉDICTIFS AUX USA : UN CALCUL MÉTRIQUE PEUT OPTIMISER LES MODES D'ACTION DE LA POLICE

L'emploi d'algorithmes prédictifs aux USA peut s'imposer aux services de police pour optimiser leur appropriation territoriale.

Fondées sur une analyse de la jurisprudence, quelques règles émergent pour que les fouilles, les interpellations, les perquisitions et autres actes de police soient conformes au 4e amendement de la constitution qui proscrit les pratiques abusives.

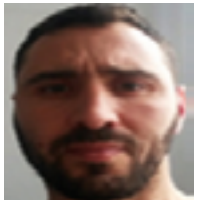
La problématique de la quantification des contrôles, du niveau d'intervention des forces de police répond à une stricte prise en compte de probabilités sectorielles de commissions d'infractions à la loi au regard des recommandations des algorithmes. Les solutions retenues s'inspirent de la détermination de pourcentages de populations à risques. Elles résistent mal au flou de leur calcul, à une inadéquation à la réalité des zones ou des personnes ciblées par la police et au respect d'un seuil d'acceptation par la population. Cet article évoque une voie qui, bien paramétrée, pourrait servir de protocole efficient pour les forces de police.

Algorithmes prédictifs

et droit des algorithmes : contrôler la quantité de faux positifs moralement acceptable

Par Bilel Benbouzid

NDLR : Nous avons souhaité exprimer par cet article le point de vue d'un sociologue. Monsieur Bilel Benbouzid s'attache à comprendre, dans un contexte de montée en puissance de la police prédictive aux États-Unis, de « quel droit » une patrouille de police peut engager un contrôle sur une personne à partir des recommandations produites par un algorithme prédictif. Les juristes, qui s'intéressent à la police prédictive aux États-Unis, essaient de répondre à cette question en référence aux dispositions protégeant des sujets de droit contre les fouilles, les perquisitions et les saisies abusives selon les garanties du Quatrième amendement.



BILEL BENBOUZID

Maître de conférences en sociologie. Université Paris-Est Marne la Vallée

Le sociologue, au cours d'un entretien, présente les solutions proposées par les spécialistes du droit qui débattent des principes, des règles et des méthodes susceptibles d'encadrer les machines prédictives.

La revue : Comment la jurisprudence américaine cadre-t-elle la question des contrôles policiers sur les personnes ?

Bilel Benbouzid - En l'état actuel de la jurisprudence, les tribunaux ont défini une multitude de situations exprimant les conditions d'évaluation de probabilités raisonnables, ou de soupçon raisonnable, de l'agissement d'un criminel dans une circonstance particulière ou dans des situations précises telles que le fait de rôder autour d'une école, d'avoir un certain type de tenue vestimentaire dans certaines situations, etc. L'idée d'attribuer une définition chiffrée à cette notion de probabilité a toujours été rejetée et c'est tout le problème de l'encadrement juridique des algorithmes prédictifs. En effet, si les juges peuvent dire qu'une probabilité supérieure

ou égale à 50 % peut justifier une perquisition dans un domicile, il est admis que ces 50 % fonctionnent comme une métaphore probabiliste et non pas comme un calcul concret.

RG : Pourquoi ce rejet de la quantification ?

BB : Ce rejet de quantification tient au respect d'une ontologie juridique remplissant une fonction anthropologique qui institue l'Homme en tant que personne, s'opposant ainsi à toute forme de quantification juridique la réduisant à l'état de chose.

Ainsi, pour que les officiers de police puissent prendre une décision, ils doivent composer avec des probabilités qui peuvent être interprétées selon un degré de confiance à accorder à une situation incertaine, compte-tenu de ce que les policiers savent par expérience des pratiques des criminels.

Un soupçon est raisonnable s'il est individuel. L'arrivée des machines prédictives remet en question une dichotomie entre soupçon individuel et général, car la plupart des données et des arguments rassemblés dans un soupçon individuel peuvent également être analysés dans un calcul de probabilité par une machine.

De quoi est constituée cette littérature juridique ?

BB : Cette littérature juridique est faite de plusieurs analyses jurisprudentielles qui offrent des justifications différentes d'un usage de la police prédictive respectueux des dispositions du quatrième amendement.

Nous y avons observé trois tentatives différentes de régulation des machines.

La première est le principe de proportionnalité qui dit que si la police, à partir d'algorithmes, produit une surveillance prédictive visant à construire des profils de comportements suspects, le taux de réussite doit être proportionnel à l'intensité de l'intrusion produite par cette surveillance. Mettons qu'un algorithme identifie 30 personnes comme étant "à haut risque", cela signifie-t-il que la police peut immédiatement stopper et fouiller ces 30 personnes ou bien qu'elles doivent faire l'objet d'une surveillance intensive ?

Selon la perception du principe de proportionnalité, dans le premier cas, il faudrait que les personnes identifiées comme « à risque élevé » atteignent au moins un score de 30 % de chances d'être associées à un crime grave car, à défaut, un simple contrôle ne pourrait être autorisé. Cette cote de 30 % provient d'un sondage auprès des juges fédéraux qui ont indiqué qu'en moyenne la justification requise pour un simple contrôle peut être quantifiée à un niveau de certitude de 30 %…

Toujours selon le principe de proportionnalité, une arrestation à partir d'une recommandation d'algorithme peut être autorisée seulement si la police est au courant d'un crime à proximité, sinon une identification « à haut risque » permettrait à la police de procéder à des interpellations répétées et arbitraire d'individus à sa guise, ce qui pourrait avoir des conséquences dramatiques.

Dans le deuxième cas, si la police décide de surveiller ces individus à haut risque de manière intensive, le taux de réussite devra être proportionnel à l'intensité de la surveillance, soit 50 % de niveau de certitude si la surveillance est prolongée. Le problème est que cette solution nécessite de quantifier le soupçon.

RG : Quelle est la deuxième proposition ?

BB : Il ne s'agit pas d'un principe mais de la recherche de normes encadrant l'usage de la police prédictive par rapport *aux règles du soupçon raisonnable* établies précédemment par d'autres juridictions.

La logique ici est de chercher des analogies avec des situations qui ont déjà fait l'objet de décisions par les juges et qui pourraient s'appliquer sur les algorithmes. La plus intéressante est l'analogie avec un chien détecteur utilisé par les équipes cynophiles pour signaler la présence de différentes matières (drogues, explosifs, armes etc.). Le cerveau du chien peut être assimilé à une algorithmique sophistiquée d'apprentissage statistique qui fonctionne sur des principes similaires à la statistique d'apprentissage bayésienne. Selon les entrées qu'il reçoit, le chien sélectionne des hypothèses en fonction de leur plausibilité. Dans le processus d'apprentissage, les principaux éléments qui entraînent le chien sont les « récompenses » et les « erreurs » bien connues des spécialistes du *machine learning*. La justice ne se préoccupe pas des modalités de fonctionnement du cerveau du chien et il n'y a pas de raison qu'elle se préoccupe davantage du fonctionnent interne des algorithmes, notamment de leur machine *learning* (argument qui balaye la critique de l'effet « boite noire » des algorithmes d'apprentissage). Dans la jurisprudence, les certifications des chiens et leurs scores de succès sur le terrain ne peuvent à eux seuls constituer un *motif de croire raisonnable*. C'est l'ensemble des circonstances du contrôle qui doit permettre de le justifier. Cette contrainte pourrait s'appliquer aux algorithmes de soupçon automatique. Une solution partielle pourrait être de certifier le bon apprentissage de la machine pour renforcer la crédibilité juridique de cet outil.

RG : Et pour ce qui concerne la dernière proposition ?

BB : En fait, ces deux propositions, le principe de proportionnalité et la recherche des règles du bon soupçon algorithmique, n'ont pas de solution satisfaisante. Elles s'en remettent au bout du compte à une exigence de transparence des algorithmes comme moyen de régulation. À mon avis, la transparence des algorithmes est une fausse solution et une troisième proposition est intéressante. Elle ne ressort ni d'un principe, ni d'une règle de droit mais d'un

© Rear view of middle aged traffic cop - Biker3 - Adobe Stock

L'acceptation sociale de contrôles sous-tendus par une recommandation algorithmique passe par un rapport entre les personnes innocentes dans la population et celles qui sont contrôlées tout en étant également innocentes.

calcul : le *hassle rate* qui correspond un taux de nuisance policière acceptable, lié à l'usage des calculs prédictifs.

Cette solution est intelligente car elle règle le problème de la quantification du soupçon raisonnable qui est interdite. Le principe est le suivant : si les tribunaux exigent des policiers des soupçons individualisés par opposition à des soupçons généralisés, ce n'est pas seulement pour éviter de devoir établir des *motifs de croire* justes et non discrétionnaires mais aussi pour éviter d'importuner un grand nombre de personnes innocentes.

Alors que tout le monde se demande comment réguler le calcul du risque, il s'agirait de calculer *le seuil maximal de contrôles de personnes innocentes* dans la population, en rapportant le nombre de *personnes contrôlées innocentes* au nombre de personnes innocentes dans la population. Il faudrait donc être capable d'estimer un nombre de *personnes innocentes* dans une population, ce qui est compliqué. Avec cette métrique et une estimation d'un nombre de personnes innocentes assez grossière, on voit que les algorithmes de prédiction du terrorisme apparaissent inacceptables. Ils obtiendraient généralement un score très faible car le nombre de « faux positifs » est très élevé par rapport à la quantité totale de criminels dans la population. Il s'agit ici de calibrer les soupçons raisonnables en respectant des seuils quantifiables ou des indicateurs garde-fous. Il s'agit surtout de définir le nombre de faux positifs acceptables.

Les contrôles d'innocents sont légitimes tant qu'ils sont nécessaires à la production d'une offre de police proactive et uniquement dans la mesure où les contrôles profitent toujours à la population. Cette proposition fait de la proactivité policière un

bien commun, en quantifiant la contrepartie indissociable et inévitable de ce bien : le dosage des faux positifs, autrement dit l'optimisation d'un mal commun acceptable de la police prédictive. C'est la seule véritable solution à la régulation des machines prédictives existant aujourd'hui dans les sources académiques du droit.

RG : Ces solutions ont-elles fait l'objet de cas pratique ?

BB : Les deux premières solutions évoquées ne fonctionnent pas et n'intégreront sans doute jamais les décisions de jurisprudence. Par contre, la dernière solution est intéressante car c'est la seule qui est opérationnelle. On peut intégrer ces métriques directement dans les machines. Ces métriques qui limitent et régulent l'activité policière peuvent être débattues dans des commissions spéciales d'audit des algorithmes comme on en trouve désormais aux États-Unis. Les développeurs de la plateforme prédictive *Hunchlab* mènent aujourd'hui des recherches pour intégrer des métriques correctrices des nuisances de l'activité policière. Pour ces développeurs, prédire le crime, c'est modéliser le niveau optimal auquel doivent être mises en œuvre les patrouilles de police, en balançant les gains entre l'efficacité prédictive avec les pertes résultant des nuisances policières causée par les contrôles non fructueux.

RG : Qu'est-ce qu'il y aurait de spécifique à un droit des algorithmes ?

BB : J'observe que la police proactive et ses exigences de minimisation des faux positifs, pourraient se réguler sous la forme d'un calcul. La norme deviendrait quantitative car le calcul dirait le droit et corrigerait les comportements d'une police de type nord-américaine. Le droit s'exprimerait dans le calcul des faux positifs acceptables.

Nous sommes encore assez peu préparés à débattre de cette notion d'acceptabilité des faux positifs, dans la police comme dans d'autres secteurs. Un droit des algorithmes prédictifs doit savoir intégrer la notion de faux positifs, ce qui n'est pas évident.

Certains philosophes et juristes craignent que la logique du calcul vienne prévaloir sur les principes juridiques et la règle de droit. En tant que sociologue, on peut aussi estimer que les algorithmes ont permis une moralisation du travail de police qui n'avait jamais été aussi bien formalisée et opérationnalisée dans l'histoire du policing. La question reste ouverte de savoir si des machines prédictives de gouvernement augmenteraient l'efficience des forces de police en moralisant leurs protocoles d'action.

L'AUTEUR

Bilel Benbouzid est maître de conférences en sociologie à l'université Paris-Est Marne la Vallée. Il mène depuis 2013 une étude sur les développements et les usages de la police prédictive aux États-Unis. Il a publié récemment un article sur le cas de Predpol, « Des crimes et des séismes. La police prédictive entre science, technique et divination », Réseaux, vol. 206, no. 6, 2017, pp. 95-123.

© Statistics - Ericus

USA : UN QUESTIONNEMENT SUR LA DOCTRINE D'EMPLOI DES ALGORITHMES PRÉDICTIFS

Toute automatisation qui porte atteinte à la vie et à la liberté des citoyens doit susciter une régulation, l'examen de l'éthique qui sous-tend sa mise en œuvre et l'évocation de son éventuelle interdiction.

Aux États-Unis, les statistiques sur la délinquance sont relativement accessibles. En matière de police prédictive, l'accès aux conclusions et aux mécanismes qui les délivrent est beaucoup plus délicat. Il pose la problématique de la protection de la propriété intellectuelle et du secret des affaires évoqué par l'éditeur et de la protection de l'autorité du Juge.

L'équilibre entre le droit des citoyens et le devoir de protection de l'État s'illustre dans les pratiques de l'administration pénitentiaire dont les outils algorithmiques cherchent une anticipation par la détection de la récidive ou d'une dangerosité potentielle. La meilleure garantie contre une pratique discriminatoire abusive repose sur l'expertise humaine des recommandations et une attention particulière au niveau technologique déployé lors de l'arrestation et de la neutralisation d'individus.

Enjeux éthiques
et juridiques des algorithmes au regard des missions de sécurité publique

Par Christelle Papineau

Les missions de sécurité publique visent à protéger les personnes et les biens, contre toute atteinte, quelle que soit la nature de cette atteinte. Depuis quelques années, les services en charge de ces missions sont de plus en plus nombreux à s'adjoindre un soutien technique, sous forme d'algorithmes, propulsés par de l'intelligence artificielle. Ces instruments participent surtout au processus de détection des risques d'atteintes contre les biens et les personnes. Toutefois, il n'est pas exclu de les voir être utilisés comme outils d'intervention sur le terrain. L'utilisation d'algorithmes appelle inévitablement à se poser des questions relevant de l'éthique et à voir quelles réponses juridiques y sont apportées.

CHRISTELLE PAPINEAU

Doctorante en droit
Université de Montréal. Canada

On observe une gradation entre la place accordée à l'éthique et l'objectif opérationnel assigné aux algorithmes. Plus l'automatisation risquera d'atteindre la vie et la liberté des citoyens, plus la notion d'éthique sera présente et le besoin de régulation sera fort au point d'envisager, dans certaines hypothèses, l'interdiction du recours à des algorithmes. L'éthique apparaît comme un baromètre pour mesurer ce que les citoyens acceptent face aux possibilités techniques offertes par la technologie et les réponses juridiques corrélatives.

Le modèle américain offre une palette diversifiée d'algorithmes utilisés par les agents en charge de la sécurité publique pour mener une étude sur les enjeux éthiques et juridiques de tels instruments.

Les algorithmes, le risque et la sécurité publique

Plusieurs États fédérés, parmi lesquels la Californie, le Wisconsin et l'État de New-

York, mettent en œuvre des algorithmes pour anticiper les troubles à l'ordre public. Cette pratique est connue sous les expressions de « *police prédictive* » et de « *justice prédictive* ».

Les algorithmes, la prévention et la police prédictive

La police prédictive repose sur l'utilisation d'algorithmes pour analyser des bases de données administratives recueillant les informations liées à la criminalité et à la délinquance. L'objectif de ces algorithmes est d'apporter une aide aux agents en charge de la sécurité publique dans leur démarche d'anticipation des crimes et délits commis contre les biens et les personnes (cambriolages, vol de voitures, agressions physiques, etc.). Ces algorithmes sont un moyen supplémentaire d'affiner la connaissance que les agents ont déjà des paramètres criminels.

Les deux initiatives de police prédictive les plus connues aux États-Unis sont *CompStat et Predictive policing*. La première a été développée, dans les années 1990, au bénéfice du Département de Police de la ville de New-York. La mégapole connaissait alors un taux de criminalité élevé. La seconde initiative sera développée quelques années plus tard, offrant de détecter le lieu et le moment où l'infraction serait susceptible de se produire.

Les autorités semblent enclines à la transparence. Le Département de Police de la ville de New-York offre de consulter un compte-rendu statistique du nombre de crimes et délits hebdomadaires commis dans la mégapole. Le portail *CompStat 2.0* se présente sous la forme d'une comparaison des statistiques de l'année en cours avec celles de l'année précédente, en s'adossant à des entrées consacrées aux principales catégories de crimes et délits (meurtres, vols, fusillades, etc.) et des pourcentages indiquant les taux de réduction ou d'augmentation pour chacune de ces catégories. L'initiative *Predictive policing* est en revanche moins transparente : elle offre uniquement un descriptif parcellaire de la manière dont fonctionne son environnement algorithmique.

S'il y a bien une démarche de communication et de transparence, les efforts doivent toutefois aller crescendo, pour tendre vers plus d'informations sur l'environnement algorithmique en lui-même. Dans le contexte actuel, cette information s'avère une question d'équilibre, entre le droit des citoyens « à savoir » et le devoir de protection dû par l'État. Ce dernier doit informer les citoyens, sans toutefois entacher le succès de ses missions de sécurité publique.

Les algorithmes, la prévention et la justice prédictive

Plusieurs États fédérés, à l'instar du Wisconsin, utilisent des algorithmes pour évaluer le risque que des personnes (arrêtées, en attente de procès ou accusées) représentent pour la société, en termes de récidive et de potentiel de dangerosité.

(1) Pour en apprendre davantage, Fabrice Defferrard et Christelle Papineau, « Le pouvoir de juridiction des algorithmes aux États-Unis : entre fantasmes et réalité jurisprudentielle » (2017) 2017.668 Dalloz IP/IT.

(2) Julia Angwin et al, Machine Bias, ProPublica, 23 mai 2016. https://www.propublica.org/article

(3) Supreme Court of Wisconsin, 2016, affaire numéro 2015AP157-CR.

Ces algorithmes, utilisés par l'administration pénitentiaire, fonctionnent par comparaison entre le profil de la personne arrêtée ou accusée et un groupe de profils gémellaires[1]. Ces algorithmes ont été critiqués, notamment par l'ONG américaine ProPublica, qui pointait la discrimination dont certains groupes de citoyens étaient plus facilement victimes[2]. Une autre critique, née cette fois-ci dans le cadre d'un procès, dénonçait les protections entourant ces algorithmes au titre de la propriété intellectuelle et du secret des affaires[3].

En matière de justice prédictive, l'éthique et ses déclinaisons juridiques deviennent un enjeu beaucoup plus « vital » qu'en matière de police prédictive puisqu'il s'agit d'évaluer le profil d'une personne, d'estimer le danger qu'elle représente pour la société et d'utiliser ces résultats à des moments clés de la procédure. Face à des lois d'États fédérés autorisant le principe d'évaluation de ces risques, tout en étant muettes sur le recours à la technologie, les juges ont dû dégager les premières règles processuelles

© Virtual Understanding par Agsandrew - Adobe stock

L'opacité des algorithmes et l'incidence de leurs conclusions dans la procédure judiciaire requièrent un contrôle humain avant l'édiction de mesures de précaution.

d'utilisation des résultats algorithmiques dans le procès.

Quels enseignements faut-il retenir du modèle américain de justice prédictive ? Le premier concerne la reconnaissance de droits aux personnes « évaluées » : l'accès à toutes les données insérées, à l'algorithme et à la vérification de son fonctionnement. Le deuxième enseignement est relatif à la mise en œuvre de contrôles par des agents « humains » : examen du fonctionnement de l'algorithme puis commentaire du résultat et sa validation par l'administration pénitentiaire. Le troisième enseignement relève du droit processuel, lorsque l'algorithme est utilisé pour évaluer une personne accusée.

(4) Defferrard et Papineau, *supra* note 1.

Dans ce cas, le juge doit pouvoir écarter les commentaires formulés par les agents pénitentiaires et le résultat de l'algorithme lui-même[4].

Les algorithmes, le trouble en cours de réalisation et la sécurité publique

L'automatisation doit être interdite si elle porte atteinte à la vie des personnes.

L'usage de la technologie dans le cadre d'arrestations

Du point de vue éthique, l'intelligence artificielle, les algorithmes et les robots ne devraient pas servir d'armes contre les personnes. Il revient au législateur d'édicter un principe selon lequel la neutralisation et l'arrestation ne doivent pas avoir de prise « technologique » et demeurer le fait, tant intellectuel que physique, des agents en charge de la sécurité publique. Il lui revient aussi de préciser qu'il n'est pas possible de détourner la technologie existante (et à venir), à des fins d'armes contre les citoyens dans le cadre des missions de sécurité publique. Le robot téléguidé et équipé d'une bombe que la police de Dallas a utilisé pour neutraliser un suspect en 2016, a éveillé les consciences à ce sujet. Le suspect était mort suite à l'explosion de la bombe[5].

L'usage de la technologie pendant un conflit armé

(5) (Auteur non connu), fusillade de Dallas : le suspect tué avec un « robot-bombe », une première aux États-Unis, Le Monde, 9 juillet 2016 (édition en ligne) http://www.lemonde.fr/ameriques/article/2016/07/09.

(6) Ian Kerr et al, *Open letter to the Prime Minister of Canada : Call for an international ban on the weaponization of artificial intelligence*, 2 novembre 2017, consultable en ligne, https://techlaw.uottawa.ca/bankillerai.

La technologie (intelligence artificielle, algorithmes et robots) ne doit pas être détournée à des fins contraires à l'éthique, en servant d'armes, dans un contexte où le raisonnement, la supervision et la décision des agents humains, plus que jamais nécessaires, disparaîtraient. Fin 2017, cinq universitaires canadiens de renom ont lancé un mouvement et interpelé le Premier Ministre du pays pour obtenir l'interdiction d'utiliser l'intelligence artificielle comme arme. Leur initiative vise la scène internationale et insuffle l'idée d'un traité international comme instrument juridique le plus approprié pour arriver à ce résultat[6].

En conclusion, on peut avancer avec certitude que les avancées en matière judiciaire aux États-Unis, pays précurseur dans la mise en œuvre d'algorithmes, gagneront à être étudiées et suivies pour comprendre quels sont les axes éthiques et juridiques à développer ou renforcer.

L'AUTEURE

Christelle Papineau effectue un doctorat en droit, sous le régime d'une cotutelle de thèse internationale établie entre l'Université Paris I Panthéon-Sorbonne et l'Université de Montréal. Elle s'intéresse aux interactions entre le droit et l'intelligence artificielle, dans une dynamique comparatiste (France, États-Unis, Canada). Elle donne régulièrement des conférences à ce sujet.

L'INTELLIGENCE ARTIFICIELLE APPLIQUÉE AUX USAGES MILITAIRES

Pour déterminer une stratégie et sa déclinaison tactique dans une confrontation armée l'acquisition rapide du terrain et des mécanismes de combat de l'adversaire peut être réalisée en adjoignant de l'intelligence artificielle aux systèmes d'armes. Cela peut comprendre le traitement des données, l'équipement des plateformes mobiles, des combattants ou des postes de commandement. Le développement logistique des opérations est également facilité par la capacité des outils logarithmiques implantés sur les matériels pour en favoriser la disponibilité et une mobilité efficiente.

Toutefois, l'emploi de ces moyens digitaux doit être en accord avec les normes internationales en matière de conflits armés. Il n'est pas assuré que malgré sa sophistication un algorithme puisse, notamment dans les conflits asymétriques, déterminer qui est combattant au sein d'une population ou milices et forces régulières évoluent sans distinction nette et quels sont les dommages collatéraux admissibles au regard de la nécessité militaire ? C'est un enjeu de responsabilité pénale et civile pour celui qui mettra en œuvre ces logiciels à vocation décisionnaire.

Les usages possibles
de l'I.A dans les armées et responsabilité du militaire

Co-écrit par Gérard de Boisboissel et Jean-Yves Richard

Il est aujourd'hui notable que, avec la récente apparition de techniques d'Intelligence Artificielle (I.A), l'homme peut être dépassé dans certains domaines, notamment sur des fonctions calculatoires ou prédictives. L'I.A permet en effet une autonomie d'adaptation des systèmes dans des environnements contraints ou inconnus, une analyse et une synthèse rapide de multiples données issues du cyberespace et de proposer une aide à la décision et une sécurisation réactive en fonction du contexte.

GÉRARD DE BOISBOISSEL

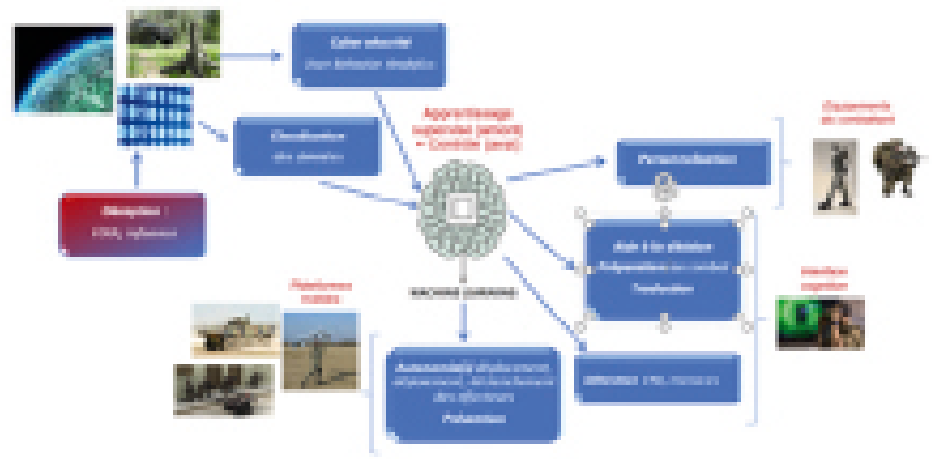
Ingénieur de recherche - CREC Saint-Cyr

JEAN-YVES RICHARD

Commissaire Principal. Professeur de Droit des conflits armés aux écoles de Saint-Cyr Coëtquidan

Les disciples de la singularité prévoient même la création d'ici quelques décennies d'entités technologiques possédant une intelligence supérieure à celle de l'homme, soit une I.A forte, capable d'égaler ou de dépasser les capacités humaines. Si cette fiction est hors de notre propos aujourd'hui, il convient de se projeter sur ce que l'I.A dite faible, permettant à des machines d'effectuer des tâches spécifiques et délimitées, pourrait apporter au monde militaire.

L'objet de cet article est donc d'envisager les outils intégrant des mécanismes d'I.A mis au service du combattant pour amplifier son efficacité opérationnelle et le décharger de tâches cognitives ou de fonctions à remplir. Ces outils pourraient prendre deux types de formes : des formes



Modalités d'usage de l'I.A.

virtuelles (des algorithmes), ou bien des équipements avec I.A intégrée leur permettant d'améliorer l'autonomie de certaines fonctions embarquées, notamment en robotique militaire.

Le schéma ci-dessus tente de lister et de classifier les différentes possibilités d'usages de l'I.A dans le mode militaire, en se donnant pour horizon 2030.

L'I.A rend les algorithmes intelligents :

Le propre de l'I.A est, comme son nom l'indique, de rendre intelligents certains processus de traitement de l'information. Notamment en ce qu'ils pourront appréhender des environnements inconnus et s'y adapter tout en étant plus réactifs que l'homme et plus précis dans l'exécution.

L'I.A doit néanmoins pour les usages militaires subir un apprentissage supervisé, c'est-à-dire un entraînement assuré par un homme qui contrôlera les données en entrée de ce processus. Un apprentissage non supervisé serait dangereux de par ses conséquences et de surcroît inutile pour le monde militaire.

Il est proposé dans cet article de classer l'utilisation de ces algorithmes intelligents en 4 domaines militaires :

A) Le traitement des données dans le cyberespace :

Pour ce qui est de la captation des données et de leur traitement en amont de toute utilisation, nous citerons tout d'abord l'explosion du volume des données massives générées dans le cyberespace, tant dans l'espace civil que dans les réseaux spécifiquement militaires. Cette inflation quantitative implique un risque d'infobésité si l'information n'est pas structurée et filtrée. L'I.A va permettre d'analyser, de classifier et de structurer automatiquement une grande quantité de données afin de les rendre accessibles à des requêtes de traitement. Elle permettra également d'effectuer des corrélations entre les signaux faibles détectés dans le cyberespace, ce qui sera de grand intérêt pour le renseignement.

Dans l'environnement militaire, ensuite, les données générées par les capteurs doivent être filtrées et analysées, afin que puissent en être extraites les informations utiles au responsable militaire. Ainsi, sur les milliers d'images que peut générer une surveillance vidéo, il est actuellement très difficile pour un opérateur d'extraire des éléments d'information d'intérêt pour l'action militaire avec une observation H24. Un traitement de ces données effectué avec une certaine intelligence pourra filtrer des informations jugées non pertinentes, comme le bruissement des feuilles dans les arbres générant des fausses alarmes de détection de mouvement, et au contraire émettre une alarme pour les plus suspectes.

L'I.A permettra également de développer des boucliers numériques auto-adaptatifs face aux évolutions des menaces. L'exemple du Cyber Grand Challenge de la DARPA lancé en 2016 montre ainsi les possibilités offertes par l'I.A. Dans ce tournoi, des systèmes autonomes s'affrontent, capables de détecter les vulnérabilités des systèmes adverses et de corriger leurs propres vulnérabilités sans intervention humaine[1]. L'I.A pourra également aider à la recherche de l'attribution des cyberattaques de par sa capacité de mise en corrélation des signaux faibles.

(1) Voir Thierry Berthier : Les Grands Dossiers de Diplomatie n° 38, 2017.

Des I.A embarquées permettront de développer leurs propres techniques de chiffrement pour échapper aux tentatives d'intrusion dans des systèmes militaires comme l'ont montré des chercheurs du projet *Google Brain* en 2016. On peut ainsi concevoir des systèmes qui créent leur propre chiffrement pour sécuriser leurs échanges de données lors de leur déploiement sur le terrain, créant ainsi un réseau de confiance et rendant impossible toute tentative d'accès externe.

Enfin une I.A peut également participer en appui de phases offensives par ses capacités de déception en introduisant de fausses données dans les systèmes d'information ennemis. C'est la technique des *False Data Injection Attacks* permettant de tromper l'adversaire, lui indiquant par exemple la présence de faux avions ou faux véhicules, ou donnant de fausses données de guidage aux systèmes ennemis. Idem pour des opérations d'influence « intelligentes et adaptatives » dans le cyberespace.

B) L'I.A au service des équipements portés par le combattant :

L'I.A va permettre d'effectuer une personnalisation des équipements numérisés que porte le combattant en opération. Ce peut être son armement, mais aussi les capteurs et boitiers qu'il transporte, qui s'adapteront à la physiologie du combattant. Par exemple, un exosquelette avec une I.A embarquée adaptera les mouvements de ce dernier en fonction de la façon de marcher de l'individu qui lui est propre.

L'I.A permettra également d'anticiper le recomplètement logistique (munitions, nourriture, eau) en fonction des phases de la manœuvre et des conditions climatiques. Elle permettra aussi d'effectuer une maintenance préventive des équipements, estimant le taux de pannes en fonction des durées d'utilisation et des conditions de leur utilisation.

C) L'I.A au service des plateformes mobiles :

Que ce soient les véhicules de combat ou les systèmes robotisés, l'I.A va permettre à ces plateformes mobiles d'intégrer des comportements autonomes. Le premier exigé sera l'autonomie de déplacement des véhicules (en avant de la manœuvre, en convoi, etc.), qui devra être accompagnée d'une autonomie de recharge énergétique pour les systèmes robotisés.

En fonction des stimuli extérieurs, les effecteurs embarqués par ces plateformes se positionneront de façon automatique en fonction de la menace. Ainsi, l'armement pointera directement sur un départ de coup de feu détecté par captation sonore ou lumineuse, ou en fonction des alertes radar.

L'I.A permettra également de détecter des menaces par observation de l'environnement : des trajectoires de véhicules suspects, des trajectoires de missiles qui permettront à ces plateformes de réagir, de se protéger et de se positionner en conséquence. Idem pour l'analyse opérationnelle des comportements ennemis connus dont on pourra prédire des tendances ou menaces potentielles, comme par exemple la pose d'IED sur des routes à des endroits clefs à certaines périodes de la journée ou des saisons.

Il en est de même pour la détection de comportements dangereux en espace ouvert ou en espace urbain, dans une foule ou ailleurs. L'analyse comportementale et l'analyse des sons émis, de la lecture sur les lèvres et des traits des visages permettront de discriminer l'attitude de personnes pouvant manifester de l'hostilité pour nos forces. De même pour les forces de sécurité intérieure, une analyse rapide du visage d'un individu menaçant permettra de retrouver sur la toile Internet très rapidement les informations concernant cet individu, ses relations et son profil psychologique. Les ressources sur le champ de bataille étant rares et précieuses, leur gestion peut bénéficier de l'I.A qui anticipera les besoins et gérera de façon optimale les potentielles pénuries. Un bon exemple concerne la saturation des ressources spectrales pour les fréquences, qui nécessiteront à l'avenir une anticipation et une réactivité immédiate en fonction de la situation des bandes de fréquences civiles et militaires.

D) L'I.A au service du soldat : personnalisation et aide à la décision

L'I.A va permettre au combattant de se préparer en s'entraînant dans un environnement simulé virtuellement au préalable à une opération. Elle permettra de jouer des ennemis intelligents et de s'adapter aux techniques de combat du soldat ou de son unité pour permettre une formation progressive et réaliste. De même le « rehearsal », qui permet de répéter la coordination d'une manœuvre, peut bénéficier des avantages d'une I.A intégrant des cas non conformes.

Un des grands apports de l'I.A sera de permettre au combattant de mieux gérer sa charge cognitive, en appréhendant au plus près la situation tactique et en améliorant le recueil des perceptions de son environnement.

Ainsi, les algorithmes intelligents pourront traduire les dialectes locaux des populations autochtones inconnus du soldat et le prévenir de toute menace dans leurs propos. L'ambiguïté des traducteurs et des guides reste en effet forte, ceux-ci apparaissant parfois comme peu fiables lors de certaines opérations.

Ils permettront aussi de reconnaître la propre voix de l'opérateur et de traduire ses ordres donnés oralement en ordres numérisés, automatiquement transmissibles aux unités subordonnées et aux supérieurs hiérarchiques destinataires en fonction du droit d'en connaître.

L'interface homme/machine qui permet au combattant de s'interfacer avec ses chefs et camarades, ou de piloter des systèmes robotiques, pourra être adaptée par l'I.A en prenant en compte à la fois ses préférences d'affichage, la situation opérationnelle et les émotions du combattant, notamment son niveau de stress. Ainsi, en fonction de la phase du combat et de son intensité (préparation, conduite de missions, contact), les informations à afficher au combattant devront être contextualisées, filtrées et présentées en fonction de la disponibilité cognitive du soldat, celle-ci déclinant avec l'intensité de l'action et avec son stress. Une interface intelligente et personnalisée, supervisée par ce dernier à l'entraînement, permettra d'éviter la surcharge informationnelle et cognitive en lui mettant à disposition une information filtrée, contextualisée et adaptative.

(2) Discussion avec le chef d'escadron Jérôme PRETEUX, Centre d'Appui aux Systèmes d'Information de la Défense.

Une autre opportunité pour l'I.A est d'optimiser le partage de l'information, notamment au sein des Cellules de Management de l'Information (CMI) au cœur des PC déployés[2]. Elle permettrait

de veiller à la transmission de la bonne information à la bonne personne qui en a besoin, tout en assurant le respect des différents niveaux de confidentialité. Puis, en se basant sur l'analyse de contenu, elle apportera une plus-value prédictive sur les ordres à transmettre ou des actions à lancer.

Enfin l'un des grands apports de l'I.A pour le chef militaire consiste en la fourniture d'une aide à la décision. En fonction d'un apprentissage préalablement supervisé de situations tactiques historiques, de la connaissance des doctrines d'emploi amies et ennemies et de la situation du moment, une I.A pourra présenter plusieurs options de choix possibles au chef militaire dans la conduite de sa manœuvre : par exemple, en analysant les modes d'action amis et ennemis ou en effectuant une analyse prédictive du comportement ennemi ou de groupes humains (manifestation, mouvement de foule etc.).

De même le Service de Santé des Armées pourrait bénéficier de diagnostics rapides en particulier pour la médecine de l'avant dans ce qui est appelé la « golden hour ». Il est important ici de préciser que la décision finale devra toujours revenir au chef qui a une vision globale de la situation et qui bénéficie d'une intuition issue de son entraînement et de son expérience. Néanmoins dans le brouillard de la guerre, une option présentée avec un pourcentage de réussite élevée influencera nécessairement le choix du chef à qui au final incombe la responsabilité.

Une fois les usages possibles de l'I.A exposés, considérons maintenant la question de l'impact des choix de l'I.A ou des options qu'elle propose sur la responsabilité du militaire en opération.

L'I.A sous le prisme du Droit international humanitaire (D.I.H) : entre contraintes et « continuum opérationnel »

Aborder la question générale de l'I.A sous le prisme du D.I.H c'est par bien des aspects retrouver certaines thématiques juridiques du Droit des conflits armés (D.C.A) liées aux systèmes d'armes létaux autonomes (SALA) et à la Cyberdéfense. Le particularisme ici, c'est de faire apprendre à l'I.A, par un codage et un apprentissage en amont, les principes du D.C.A (Humanité, Distinction combattants / non combattants, Proportionnalité et Nécessité militaire) afin de respecter au combat les engagements pris par la France au travers des multiples textes de DIH qu'elle a ratifiés depuis le siècle dernier. Cela constitue un réel défi scientifique et technologique qui justifie encore, à bien des égards pour les forces armées, la volonté *prima facie* de garder l'homme dans la boucle tout en préservant ainsi la primauté du commandement au chef militaire.

Partant du schéma ci-dessus, seuls trois usages de l'I.A nous apparaissent comme impactant directement le DIH.

a) Tout d'abord son usage est particulièrement sensible en matière d'autonomie pour le déclenchement d'effecteurs et de systèmes d'armes (missiles, laser).

En effet, à ce jour, hormis le cas d'une attaque par saturation clairement établie de la part d'un ennemi identifié (un essaim de drones par exemple, ou bien le concept de « killing box » où l'ennemi est par définition clairement identifié et surtout cloisonné), il apparaît extrêmement complexe en amont de mettre en équations le brouillard de la guerre, la « friction » selon le terme bien connu de Clausewitz. Le penseur militaire aimait ainsi à rappeler cette constante : « à la guerre plus que partout ailleurs, les choses n'évoluent pas comme nous l'avions espéré ».

Dans ce contexte, afin par exemple d'être conforme au principe cardinal du D.I.H de distinction combattant / non combattant et éviter ainsi à la dernière seconde un dommage collatéral, il apparaît difficile par définition de décrire en amont l'imprévisible sous forme d'algorithme. Cela est encore plus vrai dans tous les cas de figure dits non conformes qui surviendraient sur la zone d'opération et que seul en conduite l'esprit humain peut à ce jour gérer/appréhender, produisant un « assessment » de dernière minute, voire de dernières secondes, suspendant si nécessaire la frappe (dans le cadre d'une opération de ciblage par exemple).

L'usage actuel par certaines armées de drones armés en est une parfaite illustration.

b) De même, dans les rubriques aide à la décision et détection, la définition de la « positive identification (P.I.D) » d'une cible humaine ou d'un objectif militaire et son cumul avec le principe de nécessité militaire paraît en première approche extrêmement complexe à spécifier et à coder en avance de phase. À titre d'exemple, le port d'une veste militaire camouflée ne suffit plus aujourd'hui à désigner un combattant. Ainsi en 2018, un zadiste de la ZAD de Notre Dame des Landes ou bien encore un jeune civil pourra porter dans la rue, effet de mode saisonnier oblige, une veste camouflée. Mettre en contexte une veste camouflée avec un déclenchement d'effecteur serait dès lors une source inévitable de dommages collatéraux. D'autant plus qu'aujourd'hui nombre de combattants dans les guerres asymétriques sont, au sens du D.C.A, des civils commettant des participations directes aux hostilités (P.D.H) donc très difficile à identifier avec une certitude raisonnable. Là encore, dans ce contexte, comment coder la PID de l'ennemi et coupler à ce même codage initial celle de la nécessité militaire ? De même en opération extérieure, le chef militaire est amené quotidiennement, de façon « ad hoc » et en conduite, en temps contraint et quasi systématiquement, à faire cette adéquation entre l'objectif traité et la nécessité

militaire. Cette décision sera parallèlement prise au regard de la mission reçue et en tenant compte de facteurs exogènes (i.e : poids des dommages collatéraux et leur instrumentalisation potentielle par l'ennemi afin d'éviter le syndrome victoire tactique / défaite stratégique, impact de l'action en cours par rapport au plan de campagne stratégique, …). Et ce, quelle que soit la position hiérarchique qu'il occupe dans la chaîne de commandement.

À ces problématiques de mettre en adéquation demain une I.A qui devrait être forte pour respecter ces critères du D.I.H, s'ajoute aussi celle de la responsabilité du commandement en opération dans le cadre de la mise en œuvre de ces effecteurs/armements dotés d'I.A.

Sur ce sujet, et en guise de préambule, il apparaît d'emblée intéressant de rappeler la position du CREC Saint-Cyr de 2015[3] selon laquelle, en matière de responsabilité, le robot ne peut être sujet de droit.

(3) Voir « Drones et Killer robots : faut-il les interdire ? », édition Presses universitaires de Rennes, 2015.

Parallèlement, dans le cadre actuel d'une I.A embarquée dans des systèmes automatisés intégrant des fonctions autonomes, on pourra considérer que le triptyque du droit pénal français « autorité/pouvoir/moyens » demeure lui aussi intangible quant à la responsabilité des chefs militaires en opérations extérieures.
En effet, il y a fort à parier qu'en cas de dommages collatéraux liés à l'emploi d'une I.A, fusse-t-elle faible, on rappellera à ce même chef militaire, quel que soit son niveau dans la hiérarchie, qu'il demeure astreint aux « diligences normales » au sens de l'article 121-3 du Code pénal.

D'une façon plus générale, la jurisprudence civile influencera les futurs contentieux relatifs aux opérations militaires. Le débat sur la responsabilité juridique en matière d'intelligence artificielle vient d'ailleurs d'être relancé tout récemment avec un premier cas d'école concernant un accident mortel de véhicule autonome Uber contre un cycliste aux États Unis le 18 mars dernier dans la banlieue de Phœnix. Va-t-on rechercher la responsabilité du propriétaire du véhicule à savoir Uber, ou bien celle du constructeur ?

Ainsi, qu'en sera-t-il demain avec des équipements militaires pleinement autonomes dotés d'I.A utilisés pour des missions purement cinétiques susceptibles de produire des dommages collatéraux ?

À ce stade confronter l'I.A au D.I.H amène à reconnaître ses vertus dans l'observation, la détection de signaux faibles, la lutte contre les engins explosifs improvisés (E.E.I) ou la Logistique, voire l'aide à la décision dans le domaine du renseignement par exemple (traitement/recoupement de données…).

À l'inverse, force est de reconnaître que dans un cadre d'utilisation de l'I.A pour des systèmes cinétiques intégrant des fonctions autonomes, on voit bien les limites à conduire les opérations en pleine conformité avec les grands principes du D.I.H. Notamment sous l'angle de la distinction et de la nécessité militaire, étant donné les difficultés technologiques à décrire sous forme d'algorithmes, de façon systématique et par anticipation, ce qui, à la guerre, tient et tiendra toujours pour partie de l'imprévisible.

DÉFENSE
LES USAGES POSSIBLES DE L'I.A DANS LES ARMÉES ET RESPONSABILITÉ DU MILITAIRE

© Lina0486 - Fotolia.com.

L'EMPLOI D'UN SYSTÈME D'ARMES AUTONOME IMPOSE L'ÉDICTION D'UNE NORME

La greffe d'une IA sur des systèmes d'armes autonomes pose la question de son contrôle par l'humain. Un consensus international n'est pas évident tant les positions alternent entre la sanctuarisation du jugement humain pour séquencer un acte de guerre et la décision tacite de laisser au système d'armes, après activation, le choix de sa démarche opérationnelle.

Le respect du droit international humanitaire suggère que les algorithmes soient susceptibles de traduire des règles juridiques en instructions mais également de discerner, selon un sens commun, le contexte spatio-temporel d'une situation afin d'émettre une décision équilibrée. Il restera à régler la question de l'imputabilité d'une responsabilité d'un acte du fait que la machine et son algorithme « pilote » n'ont pas de personnalité juridique.

Systèmes d'armes autonomes intégrant de l'IA : réflexions technologique, éthique à la croisée du droit des conflits armés et du DIH.

Par Didier Gazagne et Gabriel de Bousquet

L'intelligence artificielle n'est pas nouvelle. Le concept est apparu au début des années 1950 ainsi que les premiers algorithmes. Cependant, durant les cinq dernières années on a assisté à une véritable augmentation de la puissance de ces algorithmes, désormais dotés de capacités exceptionnelles de stockage, de traitement d'informations, d'apprentissage automatique (« machine learning ») et d'autonomie de décision.

DIDIER GAZAGNE

Avocat à la Cour. Directeur Business Unit Défense – Sécurité – Systèmes autonomes

GABRIEL DE BOUSQUET

Avocat à la Cour. Business Unit Défense – Sécurité – Systèmes autonomes

Les conditions d'utilisation et/ou d'intégration de l'intelligence artificielle sont au cœur des préoccupations et des interrogations, comme le démontrent des publications récentes aux titres provocateurs telles que : *La guerre des intelligences* de Laurent Alexandre, ou encore *La chute de l'empire romain*, *Mémoires d'un robot* de Charles-Edouard Bouée, mais aussi la lettre publiée par Elon Musk et 350 experts au titre « When Artificial Intelligence Will Overtake Human Intelligence » [1]. Des chercheurs, face à leurs craintes exprimées de couplage de l'IA avec les neurotechnologies, vont jusqu'à revendiquer de nouveaux droits et en particulier quatre nouvelles lois proposées par des chercheurs pour se protéger d'un futur dominé par les neurosciences : « Droit à la liberté

(1) When Will AI Exceed Human Performance ? Évidence from AI Experts, Katja Grace, John Salvatier, Allan Dafoe, Baobao Zhang, Owain Evans, 3-5-2018.

cognitive, Droit à la vie privée mentale, Droit à l'intégrité mentale et Droit à la continuité psychologique » [2].

(2) Towards new human rights in the age of neuroscience and neurotechnology, Life Sciences, Society and Policy, Marcello Ienca et Roberto Andorno, 26-4-2017.

Jusqu'à présent, les systèmes d'armes étaient téléopérés par un opérateur humain. Avec les avancées de l'intelligence artificielle, des robots militaires entièrement autonomes sont susceptibles de faire leur apparition, dont certains pourraient tuer sans contrôle humain.

L'Agence de recherche de la défense américaine, la DARPA, mise sur un programme de recherche sur quatre ans, baptisé « OSU », qui vise à mettre l'accent sur l'illustration de la façon dont les machines prennent des décisions. Selon Alan Fern, directeur associé de la filière de robotique et de systèmes intelligents de la DARPA, les enjeux du programme OSU sont qu'« Au final, nous voulons que ces explications restent naturelles, en traduisant ces décisions de neurones profonds en phrases et en visualisations ». Cet enjeu traduit la volonté pour l'Homme, dans la boucle, de conserver en tout état de cause une traçabilité globale sur les systèmes autonomes.

1. Approche internationale des systèmes d'armes autonomes

Les systèmes d'armes autonomes sont susceptibles de devenir la troisième révolution dans la guerre après la poudre et les armes nucléaires. Le SGDSN[3] opère une distinction entre un système téléopéré (autonomie nulle), un système télé supervisé (autonomie partielle) et un système autonome au sens strict (autonomie complète). L'étude « Chocs Futurs » [4] pointe le système d'arme létale autonome (SALA) qui est plus à risque.

(3) Le secrétariat général de la Défense et de la Sécurité nationale (SGDSN), anciennement secrétariat général à la Défense (SGDN), est un organe gouvernemental français chargé d'assister le chef du Gouvernement dans l'exercice de ses responsabilités en matière de Défense et de Sécurité nationales.

(4) http://www.sgdsn.gouv.fr/rapport_thematique/chocs-futurs/.

(5) "HRW "Ban Killer Robots Before it's too Late », Human Rights Watch, 9 novembre 2012, (https://www.hrw.org/news/2012/11/19/ban-killer-robots-its-too-late) ; Losing Humanity : the Case against Killer Robots.

(6) Amnesty Internatinal, Faits & chiffres 10 raisons pour lesquelles il faut interdire les « robots tueurs », 25-12-2016.

(7) Autonomous weapons : an open letter from AI & robotics researchers, 28-7-2015.

Afin d'éviter une course à l'armement, des organisations non gouvernementales telles que Human Rights Watch[5], Amnesty International[6] ou le Futur of Life Institute[7] appellent à une interdiction préventive à l'échelle mondiale du développement, de la production et de l'utilisation d'armes totalement autonomes.

À l'ONU, le débat est engagé depuis 2014 sur l'opportunité d'une régulation – voire même d'une interdiction – des SALA.

DÉFENSE
SYSTÈMES D'ARMES AUTONOMES INTÉGRANT DE L'IA : RÉFLEXIONS TECHNOLOGIQUE, ÉTHIQUE À LA CROISÉE DU DROIT DES CONFLITS ARMÉS ET DU DIH

© Military robot - Chagpg - Adobe stock

La part de contrôle par l'être humain des matériels déployés sur un théâtre d'opération, notamment quand leur système de décision est empreint de recommandations algorithmiques, ouvre un champ de responsabilité nouveau.

1.1 Autonomie opérationnelle des robots : les doctrines française et américaine

Pour le moment, aucune définition des SALA n'a été arrêtée : d'un côté un périmètre trop inclusif risquerait de mettre à mal des capacités existantes ou le développement de capacités de pointe, et d'un autre côté un périmètre trop exclusif pourrait ne pas pouvoir englober un système pertinent.

Une chose est certaine : la question du degré d'autonomie et de contrôle humain est centrale dans la définition juridique des SALA.

Les débats ont révélé que plusieurs degrés de contrôle pouvaient être envisagés : le contrôle effectif, le contrôle significatif ou encore le jugement humain approprié. Ces diverses propositions ont pour effet de faire varier le niveau d'exigence du lien attachant l'humain à la machine afin d'exclure certains systèmes et donc de les rendre incompatibles avec la légalité internationale.

La présence humaine peut ainsi prendre trois formes :

- le robot peut sélectionner la cible et l'humain décider d'engager la force (*Human in the loop*) ;

- le robot peut sélectionner la cible et engager la force que l'humain peut suspendre (*Human on the loop*) ;

- le robot est capable de sélectionner et d'attaquer des cibles sans que l'humain puisse intervenir (*Human out of the loop*) [8].

(8) Pour approfondir cette question, voir Paul Scharre and Michael C. Horowitz, "An Introduction to Autonomy in Weapon Systems, (Center fora New American Security, 2-2015).

Doctrine française.
En France, la question de l'autonomie de décision pour les drones de combat a été résolue en attribuant systématiquement la décision de tir à l'opérateur humain. La France considère ainsi que l'homme devra conserver la capacité
de prendre les décisions ultimes s'agissant du recours à la force létale.

La question de savoir si les exigences légales d'emploi sont remplies dans une situation donnée ne dépendra pas seulement de la propre programmation du système et de ses capacités techniques, mais aussi du jugement humain.

Doctrine américaine. À l'inverse, la doctrine américaine sur les systèmes d'armes autonomes, formalisée en 2012, définit ces derniers comme un système d'arme qui, une fois activé, peut sélectionner et engager des cibles sans intervention d'un opérateur humain. Cela inclut les systèmes d'armes autonomes, supervisés par l'homme, qui sont conçus pour permettre aux opérateurs humains d'outrepasser le fonctionnement du système d'arme, mais qui peuvent sélectionner et engager des cibles sans autre intervention humaine après l'activation[9].

(9) U.S. defense policy, Department of Defense (DOD) Directive 3000.09, "Autonomy in Weapons Systems", 21-11-2012.

Plutôt que de se concentrer sur la contrôlabilité du système autonome, les États-Unis utilisent le critère du « niveau approprié de jugement humain sur l'usage de la force ». L'intervention humaine intervient davantage au moment du recours au SALA que lors de la décision d'usage de la force létale. L'humain pourra néanmoins toujours outrepasser la décision de la machine.

1.2 Divergences Onusiennes

Sur les mesures à adopter, les différentes prises de positions ont fait ressortir trois groupes d'États :

1. Les États qui considèrent que le droit international humanitaire actuel est suffisant et adéquat ;

(10) Déclaration conjointe de l'Allemagne et de la France sur l'évolution des travaux sur les SALA au sein de la CCAC, 15-11-2017.

2. ceux qui militent en faveur d'une interdiction totale ;

3. enfin, ceux qui estiment qu'à ce stade la négociation au sujet d'un instrument juridiquement contraignant ou d'une prohibition préventive serait prématurée.
Cette dernière position a été notamment adoptée par la France[10].

DÉFENSE
SYSTÈMES D'ARMES AUTONOMES INTÉGRANT DE L'IA : RÉFLEXIONS TECHNOLOGIQUE, ÉTHIQUE À LA CROISÉE DU DROIT DES CONFLITS ARMÉS ET DU DIH

Néanmoins, les discussions menées ont permis d'arriver au consensus suivant : les règles existantes du droit international humanitaire (DIH), encore appelé droit de la guerre ou droit des conflits armés, s'appliquent indubitablement à tout nouvel armement et à l'utilisation des développements technologiques robotiques à des fins militaires.

En particulier, la mise en œuvre du DIH est reconnue à l'article 36 du Protocole additionnel aux Conventions de Genève qui stipule que les États ont l'obligation de déterminer si l'emploi de moyens nouveaux ou d'une nouvelle méthode de guerre serait interdite. Ainsi, avant de développer mais surtout d'employer un SALA, chaque État doit vérifier la conformité au DIH du système d'arme et en particulier des algorithmes utilisés.

(11) Respectivement les articles 48, 57 et 51 § 5b du Protocole additionnel I aux Conventions de Genève du 12 août 1949.

L'enjeu est donc la capacité à programmer le DIH, à convertir des règles de droit en algorithmes. Les algorithmes internes aux SALA devront être capables d'intégrer et de respecter les principes de la conduite des hostilités, en particulier les principes de distinction, de précaution et de proportionnalité[11].

2. Les systèmes d'armes autonomes à l'épreuve de la réflexion éthique

Dans son rapport sur l'éthique de la robotique, publié le 14 septembre 2017, la Commission mondiale d'éthique des connaissances scientifiques et des technologies (COMEST) a livré une analyse de conformité au DIH[12] des systèmes robotiques militaires (drones) ainsi que des armes autonomes. L'objectif de ce rapport était de fournir une analyse critique et éthique de l'usage des technologies et en particulier pour les usages militaires.

(12) Rapport de la COMEST sur l'éthique de la robotique, 14-9-2017.

Robotique à usage militaire. Même si la COMEST reconnaît les avantages des systèmes d'armement télépilotés ou automatisés pour les forces armées et également pour les civils dans certaines situations de conflits, elle s'interroge néanmoins sur les règles actuelles du DIH eu égard à l'impact humanitaire prévisible de l'utilisation des systèmes d'armes automatisés.

Par rapport au principe de distinction, la COMEST souligne la qualité déterminante des images et des renseignements utilisés par le télépilote avant la décision de faire feu sur une cible et considère que les limites spatio-contextuelles accessibles au télépilote qui prendra la décision de faire feu ou d'attaquer une cible ne sont pas identiques à celles dont disposerait un pilote d'un avion de combat.

Pour la COMEST, l'évaluation de la proportionnalité à distance est jugée problématique au regard de la décision du télépilote d'utiliser ou non la force létale ou

cinétique en s'appuyant sur les données contextuelles mises à sa disposition.
La COMEST semble considérer qu'en raison de la distance, le télépilote ne pourrait pas parvenir à une « décision équilibrée » pour mettre en balance l'avantage militaire attendu sur le théâtre d'opération avec le risque de pertes de vies civiles. L'examen de la proportionnalité par un télépilote implique de prendre en compte les victimes civiles mais aussi l'ensemble des dommages collatéraux.

Armes autonomes. La COMEST mène également une analyse de conformité au DIH des armes autonomes qu'elle définit comme « une arme qui, une fois activée, sélectionne et attaque des cibles sans intervention humaine supplémentaire ».

La mise en œuvre des principes du DIH implique que l'IA soit en mesure de comprendre automatiquement des situations parfois complexes : distinguer un combattant d'un civil, reconnaître une reddition, un prisonnier de guerre, un soldat blessé, etc. (principe de distinction).

Par ailleurs, la COMEST considère que l'évaluation de proportionnalité implique nécessairement « un jugement humain », lequel d'un point de vue éthique, ne pourrait être laissé « à une machine ». Bien souvent, l'évaluation de la proportionnalité est entièrement une affaire d'intuition et de discernement émotionnel, deux qualités profondément humaines.

Un rapport de la Royal Society confirme cette analyse en relevant la difficulté de « mettre au point des machines capables d'aborder un problème de manière contextuelle c'est-à-dire en faisant usage du « sens commun » [13].

(13) Le secrétariat général de la Défense et de la Sécurité nationale (SGDSN), anciennement secrétariat général à la Défense (SGDN), est un organe gouvernemental français chargé d'assister le chef du Gouvernement dans l'exercice de ses responsabilités en matière de Défense et de Sécurité nationales.

Par ailleurs, la COMEST soutient que, s'agissant de la clause de Martens, règle spécifique du DIH qui exige l'application des « principes humanitaires », seul un humain pourrait outrepasser le droit d'une personne à la vie. Une telle décision inacceptable ne pourrait être en aucun cas déléguée à une intelligence artificielle ou à un logiciel et ce, même si les règles du DIH pouvaient être introduites en dur dans la programmation du système d'arme autonome.

(14) Document de travail « Cadre juridique d'un éventuel développement et usage opérationnel d'un futur système d'armes létal autonome » présenté à l'occasion de la réunion de Genève des 11-15 avril 2016.

Malgré ces difficultés, il n'est pas inutile de noter que l'être humain n'est pas non plus infaillible (vengeance, peur, besoins physiologiques…). La France considère même à ce titre qu'on ne peut écarter l'hypothèse que, dans certaines circonstances, les algorithmes pourraient mieux respecter les principes du DIH que des êtres humains[14].

(15) China's plan to use artificial intelligence to boost the thinking skills of nuclear submarine commanders, South China Morning Post, 4-2-2018.

En ce sens, la Chine a justement décidé d'équiper ses sous-marins nucléaires avec de l'IA afin de compenser l'altération des capacités humaines du commandement après un long moment passé dans un espace clos immergé[15].

3. Systèmes d'armes autonomes et sécurité intérieure

L'utilisation de robots équipés d'armes pour des opérations de sécurité intérieure fait également débat. L'impératif juridique du maintien de l'ordre justifie-t-il le recours à des systèmes autonomes spécifiques aux forces de sécurité ?

Les capacités opérationnelles des systèmes autonomes de sécurité augmentent le champ des possibles tout en réduisant le risque humain pour les forces de sécurité. À titre d'exemple l'utilisation du robot de Northrop Grumman Andros à Dallas en 2016 pour neutraliser un sniper a sans doute évité d'autres pertes humaines.
Le recours à de tels systèmes suscite néanmoins des questionnements en termes de proportionnalité.

Les approbateurs considèrent que ces systèmes ne risquent pas de désobéir, ne discutent pas les ordres, pas plus qu'ils ne sont susceptibles de sédition. Au contraire, les détracteurs objectent que le recours à de tels systèmes n'est pas sans risque selon les États qui les mettent en œuvre (régime autoritaire).

L'utilisation de robots suscite en outre un problème d'acceptabilité par les populations concernées. Dans son rapport sur le maintien de l'ordre, le défenseur des droits a récemment souligné que l'équipement lourd, du style « robocop », des agents chargés du maintien de l'ordre semble contribuer à une certaine déshumanisation[16].

(16) Défenseur des droits, rapport sur « Le maintien de l'ordre" au regard des règles de déontologie », 12-2017.

L'usage de véritables robots, pour encadrer une manifestation par exemple, risque d'envoyer un message d'hostilité peu favorable au dialogue.

4. Imputation de la responsabilité

(17) Olsthoorn (P.), Royakkers (L), « Risks and Robots - some ethical issues », Netherlands Defense Academy, 2011.

Selon Robert Sparrow, une condition fondamentale du respect du droit international des conflits réside dans la possibilité d'attribuer la responsabilité à un individu en cas de d'action contraire aux normes[17].

À défaut de personnalité juridique, le système autonome ne peut endosser la responsabilité de ses actions. C'est donc directement vers les humains et/ou les États qu'il faudra se tourner.

L'intégration de l'IA dans les systèmes autonomes fait craindre une dilution de la

responsabilité. À qui doit-on imputer la responsabilité : au fabricant, à l'intégrateur, au programmeur, à l'opérateur, au militaire décisionnaire, à l'État ?

Dans les opérations militaires classiques, la question de la responsabilité personnelle est déjà compliquée à gérer dans un processus faisant intervenir à la fois des autorités politiques (décision d'engagement), les autorités militaires (détail des opérations), et en dernier lieu le militaire (réalisation de l'action).

(17) OSTE Frédéric et TARAVELLA Adeline, « Relation homme-robot : prise en compte des nouveaux facteurs sociologiques », IRSEM, 2012.

Or avec les systèmes autonomes, l'écosystème nécessaire à leur fonctionnement démultiplie encore le nombre d'intervenants[18] et complexifie donc la chaîne de responsabilité en cascade.

Tant que les systèmes autonomes restent sous le commandement de l'humain, l'imputation de leurs actions à ceux qui les déploient ne semble a priori pas impacter l'attribution de la responsabilité de manière irrémédiable.

Mais le fait qu'un système soit autonome dans son fonctionnement risque de réduire le rôle des humains et d'atténuer un peu plus encore leur responsabilité. Dès lors que le système autonome bénéficiera d'une large autonomie, l'imputation de la responsabilité dans l'état actuel du droit risquera d'être plus difficile : l'opérateur disposera-t-il d'un degré suffisant de contrôle sur le système d'armes pour que sa responsabilité soit engagée ? Qu'adviendra-t-il en cas de « bug » ?

(19) L'Institute of Electrical and Electronics Engineers ou IEEE, en français l'« Institut des ingénieurs électriciens et électroniciens », est une association professionnelle. L'IEEE compte plus de 400 000 membres et possède différentes branches dans plusieurs parties du monde. L'IEEE est constituée d'ingénieurs électriciens, d'informaticiens, de professionnels du domaine des télécommunications, etc. L'organisation a pour but de promouvoir la connaissance dans le domaine de l'ingénierie électrique (électricité et électronique). Juridiquement, l'IEEE est une organisation à but non lucratif de droit américain.

(20) The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems.

À côté du cadre juridique de la responsabilité applicable à un système autonome, la culture juridique anglo-saxonne donne une plus grande place à la *soft law* et à la standardisation. Dans cette perspective, l'IEEE[19], la plus grande société savante qui a également lancé son « Initiative Globale pour des considérations éthiques en intelligence artificielle et systèmes autonomes », travaille à l'élaboration de normes sur la prévisibilité, la traçabilité, et/ou l'identification claire des responsabilités humaines dans le déploiement des systèmes autonomes[20]. Les Européens étant très peu présents dans les débats normatifs internationaux, cela devrait avoir pour conséquence de se traduire par l'adoption de standards américains ou chinois spécifiques au secteur de la défense et de la sécurité.

DÉFENSE
SYSTÈMES D'ARMES AUTONOMES INTÉGRANT DE L'IA : RÉFLEXIONS TECHNOLOGIQUE, ÉTHIQUE À LA CROISÉE DU DROIT DES CONFLITS ARMÉS ET DU DIH

Les transformations qu'entraîneront l'intégration de l'intelligence artificielle dans les systèmes d'armes autonomes, si elles augmentent le champ des possibles, impliqueront également de maîtriser la culture de la donnée tactique (Concept de cloud tactique). Le principe « *Human in the loop* » devrait rester au cœur de l'approche française concernant l'adoption de systèmes d'armes autonomes, mais la convergence de l'IA, couplée à l'utilisation de la Blockchain et à des systèmes décisionnels et autonomes, pourrait bien faire évoluer la place du soldat augmenté dans le champ de bataille numérique ainsi que la place du gendarme ou du policier augmenté, sans négliger la survenance de nouvelles ruptures technologiques.

DOSSIER

Algorithmes et droit

Algorithmes et politiques
publiques de sécurité, quels nouveaux paradigmes ?

Par Jérôme Lagasse

Les politiques publiques de sécurité déclinées dans les plans d'action du ministère de l'intérieur sont à la recherche permanente d'un juste équilibre entre les mesures de prévention et de répression. Ce diptyque trouve aussi sa version sur le plan juridique avec la distinction traditionnelle opérée entre la police judiciaire et la police administrative, déclinaison conforme au principe constitutionnel de la séparation des pouvoirs. Ces ordonnancements politico-juridiques seront susceptibles d'évoluer par l'avènement de la transformation numérique, troisième révolution industrielle, dont Jérémy Rifkins a décrit les processus de changement[1].

JÉRÔME LAGASSE

Chef d'escadron de gendarmerie. Chef du département études et stratégie. Centre de recherche de l'école des officiers de la gendarmerie nationale

(1) Jérémy Rifkins, *La troisième révolution industrielle*, Les liens qui libèrent, 2012, 414 p.

Proaction et anticipation relèvent d'une constante historique

La capacité à collecter, traiter, analyser les données de masse issues des différentes activités humaines constitue en soi une troisième révolution industrielle. Souvent traitée sous un angle anxiogène, l'arrivée des algorithmes dans le champ de la « police prédictive » et de la « justice prédictive » s'épanche plus sur le risque d'une surveillance de masse par l'État sans en démontrer de manière tangible les écueils. Malgré tout, il est possible d'éclairer d'un autre jour les changements de paradigmes qui se dessinent et s'expriment par une nécessité avérée pour les décideurs de disposer d'outils d'analyse décisionnelle en phase avec les exigences d'une société démocratique connectée, qui fait rimer transparence et égalité, avec une forte attente de prévisibilité des risques et de la capacité à anticiper pour les annihiler. Cette posture d'anticipation et de proaction est déjà très présente à la fin de l'ancien Régime[2], la Lieutenance de

police de Paris en étant l'exemple le plus abouti. Pour surprenant que cela puisse paraître au premier abord, la transformation numérique fait resurgir des problématiques communes à celles auxquelles la fin de l'Ancien Régime a dû faire face : les circulations, une hausse des mobilités, la maîtrise des flux humains et des marchandises. Le siècle des Lumières, confronté à un vaste mouvement des populations, sans précédent dans le Royaume, a eu pour conséquence directe une hausse exponentielle du nombre d'habitants des grandes villes dont en premier lieu Paris. Il s'agit pour les autorités de l'époque de réfléchir à la manière de « Gouverner l'espace ».

(2) « *Vous connaissez les caractères de la justice réglée et de l'administration, il est essentiel de distinguer ces deux voies, et important de les maintenir que l'une ne puisse nuire à l'autre. La justice réglée veut une marche lente, régulière et des formalités invariables. À cet égard, je ne peux que vous renvoyer aux lois. Leurs dispositions vous indiquent vos devoirs. À l'égard de l'administration, elle doit avoir une marche prompte pour prévenir le crime, et par une prévoyance bienfaisante garantir la société du désordre[...]* » Lettre du lieutenant de police Lenoir aux syndics de la Compagnie des commissaires, 3 août 1776, cf., Vincent Milliot « *L'admirable police, Tenir Paris au siècle des Lumières* », Champ Wallon, Collection d'histoire, p. 68/69, 2016, 366 p.

Cette époque voit foisonner de nouvelles techniques policières, ancêtres de nos applications numériques : enregistrement écrit, cartographie, découpages spatiaux et recherche d'une territorialisation équilibrée. La Maréchaussée, en la personne de M. Guillauté, officier, rédige ainsi en 1749 un Mémoire[3] où il est proposé un projet de fichier centralisé dénommé « le serre-papier » qui aurait pour objet d'enregistrer l'ensemble des habitants du pays et dont une des finalités serait de tracer les déplacements de chaque individu de ville en ville. De concert, nous assistions au XVIII[e] siècle à l'émergence d'une judiciarisation des opérations de police administrative, soucieuses déjà de « l'importance du respect des formes [4] ». Cet éclairage historique dénote, par certains aspects, qu'à l'aune de toute révolution industrielle, les disruptions qu'elle sème obligent les contemporains en responsabilité à recourir à une posture pro-active et d'anticipation suffisante pour accompagner les transformations sociétales en cours.

(3) M. Guillauté, Mémoire sur la réformation de la police de France, soumis au Roi en 1749, Hermann, 1974, 116 p. Au sujet du Serre-papiers.

(4) « Les arrestations par voie administrative ne nécessitent pas, pour être déclenchées, de déclaration préalable ou un dépôt de plainte ; elles n'exigent pas la production de preuves légales au sens propre du terme. Pourtant, assez systématiquement dans la seconde moitié du XVIII[e] siècle, les captures d'ordre du roi opérées par les inspecteurs de Sûreté respectent généralement la règle du passage devant un juge, en l'occurrence le commissaire qui doit statuer en tant que « premier juge » sur l'emprisonnement effectif ou la relaxe, comme l'évoque Lemaire ou Lenoir dans leurs écrits. », Idem note 1 (p. 248).

Les séquences terroristes de type « tueries de masse », débutées par l'attentat du 11 septembre 2001, ont été un des facteurs déclencheurs d'une vaste revue générale des politiques publiques de sécurité conduisant à mener un état des lieux sur la posture la plus à même à faire face aux menaces protéiformes et asymétriques. Les nouvelles technologies de sécurité, telles que les algorithmes à finalité prédictive, font progressivement leur entrée dans les différentes missions en lien, de manière directe ou indirecte, avec les politiques publiques de sécurité. Destinées initialement au haut du spectre (terrorisme), les algorithmes, au service de la déclinaison de l'action publique de ces politiques, s'invitent dans le pilotage de celles-ci. Cette modernisation de l'action publique va de pair avec le changement de posture du ministère de l'intérieur qui, jusqu'à présent, était perçu par ses process comme un ministère de l'urgence et de l'action. Progressivement, nous observons une inflexion de sa culture « d'entreprise » et de sa doctrine. Elle est tournée vers davantage d'anticipation, de proaction et de détection des prémices du passage d'une situation stable à un état de crise.

(5) https://efus.eu/fr/resources/publications/efus/11191/ cf. manuel EFUS, *Méthodes et outils pour une approche stratégique de la sécurité urbaine*,2016, p. 32.

Les données de masse, qui nourrissent notamment des algorithmes de « recommandation » esquissent des tendances de fond. Il est encore trop tôt pour mesurer toutes leurs influences dans le champ de l'action publique de la sécurité, plus particulièrement lorsqu'il s'agit de dresser des diagnostics. Dans ce sens, le *Forum européen de la sécurité urbaine*[5], association de collectivités territoriales, estime que : « *les politiques de sécurité doivent être considérées comme une discipline à l'égal de la médecine ou de l'architecture : on détecte les causes d'une affection avant de les traiter ; on évalue les caractéristiques d'un terrain avant de lancer des travaux de construction ; on doit par conséquent agir de même et établir un diagnostic de sécurité avant de résoudre les problèmes.* »

(6) Dans le contexte de la protection des données à caractère personnel, on entend par Accountability une obligation de rendre compte et d'expliquer, avec une idée de transparence et de traçabilité permettant d'identifier et de documenter les mesures mises en œuvre pour se conformer aux exigences issues de la réglementation Informatique et libertés.

Ériger les politiques publiques de sécurité à l'égal d'une science comme la médecine impose à la fois de disposer de données probantes, sur des critères normés (ISO) incontestés et incontestables, mais aussi d'algorithmes publics transparents, capables d'expliquer les recommandations qu'ils fournissent, le tout sous le contrôle final d'une décision humaine guidée par le principe d'*accountability*.[6]

L'exploitation des données de masse, par le biais de traitements automatisés, induit des changements de paradigme dans des registres majeurs déclinés dans les politiques publiques de sécurité. Il est possible d'entrevoir ces évolutions autour de dyptiques tels que : politique de répression / politique de prévention ; modèle de statistique descriptive / modèle de statistique prescriptive ; police judiciaire / police administrative ; gouvernement des Hommes et des organisations / gouvernance par les nombres.

Politiques de prévention de la délinquance versus politiques de répression :

Dans son « Manifeste d'Aubervilliers et de Saint-Denis (2012), l'Efus (European forum for urban securit) déclarait : « *Toute politique doit prendre en compte l'état des connaissances techniques, scientifiques et créer les conditions de leur production. Pour ce faire, les villes doivent se donner les moyens de s'assurer que leurs politiques sont déterminées et guidées par des données probantes, tant qualitatives que quantitatives, et non pas par des préjugés ou postures idéologiques.*[7] ». Disposer de données probantes sera l'enjeu des « villes intelligentes ». Les données probantes, classées et analysées par des algorithmes de recommandation au sein de chaque secteur d'activité d'une ville, exerceront sans nul doute une influence décisive dans les processus de normalisation et de standardisation de type Afnor et ISO. Dans ce sens, nous pouvons citer la norme ISO 37 122, en cours d'élaboration, qui s'intéresse en particulier aux indicateurs de performance des services publics à l'égard de la Ville Intelligente[8]. Sur le versant des politiques de répression, les juridictions judiciaires sont soumises à une forte tension dans les délais de traitement des affaires pénales. Les algorithmes de recommandation pourraient être ces nouveaux « assistants de justice » chargés en amont de libérer les magistrats de certaines contingences chronophages dans la mise en état d'une affaire à juger. Dans le même sens, l'article 21 de la loi du 7 octobre 2016 *pour une République numérique* ouvre la voie à un changement de paradigme en instaurant que « *les décisions rendues par les juridictions judiciaires sont mises à la disposition du public à titre gratuit dans le respect de la vie privée des personnes concernées*[9] ». Cette disposition ouvre des perspectives inédites pour les entreprises de *Legal Technology*[10] en capacité à terme de formuler des recommandations à des justiciables sur l'issue d'un procès en fonction de la jurisprudence, voire même de la composition d'un tribunal.

(7) https://efus.eu/files/2013/06/Manifeste-VF-2.pdf P. 7.

(8) ISO 37 122 - Développement durable dans les collectivités - Indicateurs de performance pour les villes intelligentes.

(9) cf., inséré Art. L111-13 du code de l'organisation judiciaire.

(10) Contracté sous le terme de Legal-Tech, elle désigne les entreprises qui mettent en œuvre les dernières technologies au service du droit pour le plus grand nombre.

Les algorithmes prédictifs nouveaux « miroirs des princes [11] » ? :

Pour le statisticien Alain Desrosières, « Les statistiques sont le produit de conventions sociales. Plutôt que de se demander si elles « reflètent objectivement la réalité », il est plus fécond de les voir comme des mises en forme du monde, parmi d'autres, et de s'interroger sur les procédures d'objectivation ».
Le choix de la procédure d'objectivation la plus « sincère et fidèle » [12] pour élaborer un diagnostic d'une politique publique fait l'objet au sein de la communauté scientifique d'un débat qui se cristallise autour de deux modèles concurrents. Le modèle déductif[13] a pour principe d'aller du général au particulier. Il laisse une part d'intuition et d'analyse au chercheur pour dresser, à partir d'une cartographie descriptive d'un phénomène, des théories mises en œuvre à travers un plan d'action dont les résultats ne pourront être évalués qu'au terme de sa durée. Le phénomène Big Data privilégie au contraire un raisonnement inductif[14] qui offre un pilotage des politiques publiques plus souple, permettant de prendre des mesures correctives à l'égard d'un plan d'action en cours de réalisation. Ainsi, dans la lutte contre la délinquance de masse, des algorithmes prédictifs[15] constituent un outil d'aide à la décision permettant aux responsables de la sécurité publique de se montrer plus agiles dès la survenance de « signaux faibles ».

(11) Les Miroirs des princes (*Specula principium*) relèvent du genre littéraire du miroir, apparu au Moyen Âge, et sont des sortes de traités d'éthique, de marche à suivre, de préceptes moraux spécifiquement destinés aux chefs d'État.

(12) Selon l'expression utilisée par les commissaires aux comptes dans la certification de ceux-ci.

(13) « Dans le raisonnement déductif, une conclusion est atteinte en appliquant des règles générales (prémisses), réduisant le domaine du discours par des étapes logiques, de la règle générale au cas particulier (conclusion) : les hommes sont mortels, Socrate est un homme, donc Socrate est mortel ». [ ] Partir d'un modèle est adopter un raisonnement déductif. cf. : Pierre Delort, Le Big Data, Coll. Que sais-je ? PUF, 2015, p. 42.

(14) « Dans le raisonnement inductif, les prémisses fournissent des arguments forts, mais pas des preuves de la conclusion. Nous passons des faits à des règles et les mathématiques permettent de mesurer l'incertitude pesant sur ces règles, dépendant notamment des faits sur lesquels ces règles sont basées ». Source : idem note 12.

(15) À ce sujet, cf. les logiciels d'analyse prédictive : *Predpol*, Blue crush, DDACTS, *Precobs*.

Le Code de la sécurité intérieure : un code hybride qui bouscule la distinction Police administrative et Police judiciaire ?

L'introduction dans l'ordonnancement juridique d'un code de la sécurité intérieure (CSI) en 2016 interroge des praticiens du droit sur sa place. Ces doutes ont été l'occasion, pour l'Institut national des hautes études de la sécurité et de la justice (INHESJ), d'introduire un colloque le 22 janvier 2016 : « *Le Code de la sécurité intérieure, trois ans après : artisan d'un nouvel ordre ou semeur de désordre ?* ». Ce séminaire invitait, entre autres, les auditeurs à réfléchir autour de deux tables rondes sur la finalité même du CSI : *un code préventif ou répressif ?* mais aussi sur sa nature potentiellement hybride : « *un code complémentaire ou concurrent du code de procédure pénale ?* ».
Ces préoccupations au sein de la doctrine, se sont faites d'autant plus jour en raison de l'insertion de la loi sur le renseignement[16] dans le CSI, au livre VIII, qui lui est spécifiquement dédié. Rappelons que cette loi a officialisé et encadré l'usage de nouvelles technologies dans le cadre de la prévention contre le terrorisme, la grande criminalité organisée et les atteintes aux institutions républi-

(16) *Comprendre la loi sur le renseignement*, note n°13 creogn, Général d'armée Marc Watin-Augouard, septembre 2013 https://www.gendarmerie.interieur.gouv.fr/crgn/Publications/Notes-du-CREOGN/La-loi-sur-

caines. Parmi ces nouvelles technologies figurent l'installation d'algorithmes prédictifs dans les systèmes des fournisseurs d'accès à internet (FAI) que les médias ont qualifié de « boites noires algorithmiques » en raison d'un manque de transparence.

(17) *Donnée servant à caractériser une autre donnée, physique ou numérique : les métadonnées sont à la base de l'archivage.* (Dictionnaire Larousse).

(18) Commission nationale de contrôle des techniques de renseignement (CNCTR).

Ces « boites noires algorithmiques » ont été programmées pour procéder à la captation de métadonnées[17] à caractère personnel en vue de les croiser pour révéler des faisceaux de « signaux faibles ». Pour discriminer ces métadonnées dans le respect des principes de proportionnalité et de finalité imposés par le législateur, une Commission[18] émet un avis sur les paramètres de détection pouvant être mise en œuvre. Paradoxalement, l'usage d'algorithmes de recommandation par les services de renseignement traduit aussi une évolution de paradigme dans la culture de l'anticipation et de la proaction qui leur est spécifique. En effet, la source humaine ne constitue plus en soi la source de référence permettant de démanteler un réseau.

Désormais, la source numérique semble prendre un ascendant certain, au détriment de la source humaine, pour détecter les faisceaux de signaux faibles annonciateurs des prémices d'une activité criminelle.

Le CSI, par sa nature potentiellement hybride, au sein de l'ordonnancement juridique ouvre une orientation plus globale vers une codification des algorithmes.

Les algorithmes prédictifs induisent-ils nécessairement un mode de « gouvernance par les nombres[19] » ?

(19) Alex Supiot, *La gouvernance par les nombres,Cours au Collège de France, 2012-2014*, Fayard, 2015, 512 p.

(20) Alexis de Tocqueville, *Considérations sur la Révolution (1850-1858)*, in Œuvres, Gallimard, 2004, T. III, P. 492 in « *L'empire des données, Essai sur la société, les algorithmes et la loi.* » Adrien Basdevant et Jean-Pierre Mignard, Don Quichotte, 2018, p. 51.

(21) Comité rattaché au Centre des hautes études du ministère de l'intérieur (CHEMI).

(22) Université Jean-Moulin Lyon 3, École Doctorale de Droit (ED 492), Équipe de recherche Louis Josserand (EA 3707).

« *[I] a notion de gouvernement se simplifie : le nombre seul fait la loi et le Droit. Toute la politique se réduit à une question d'arithmétique* ».[20] Cette citation d'Alexis de Tocqueville met en évidence au contraire que toute politique publique, plus particulièrement de sécurité intérieure, ne saurait se réduire à une question d'arithmétique. Les nouvelles technologies de sécurité soulèvent des questions éthiques et d'acceptabilité sociale dans toutes les composantes des forces de sécurité intérieure. Conscient de ces enjeux, le *Comité de prospectives*[21] a commandité, début 2018, une étude[22] relative aux « *coefficients d'acceptabilité et de résilience des acteurs de la sécurité intérieure en matière de cartographie prédictive du phénomène criminel* ».

(23) https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_garder_la_main_web.pdf La CNIL a identifié trois grandes fonctions sur cinq d'ores et déjà en développement dans le secteur de la sécurité : générer de la connaissance (ex : repérer des liens insoupçonnés pour la résolution d'enquêtes), prédire (détecter les profils à risque dans la lutte anti-terroriste, prédire le champ d'occurrence de crimes et délits), aider à la décision (suggérer aux forces de police les zones prioritaires dans lesquelles patrouiller) – Tableau p. 22.

(24) À ce propos, cf. art. « *En 2035, le Chef au combat sera t-il un meneur d'homme ou un manager connecté ?* », Cahier de la pensée mili-Terre n°48, Chef de bataillon Sébastien Pistre, Capitaine François-Xavier Lambin-Bernot et Captain Bernhard Kirchner, publié le 20/11/2017.

Dans ce cadre, des officiers de gendarmerie et des commissaires de police, en formation initiale ou continue, ont fait l'objet d'un sondage sur leur perception à l'égard de l'analyse décisionnelle et des algorithmes prédictifs. Cette étude sera suivie par un nouvel appel à projet qui traitera de « *l'encadrement des risques techniques et juridiques des activités de police prédictive et secours prédictifs* ».

Ces réflexions prospectives, orientées dans le champ de la sécurité intérieure, sont la traduction concrète de six recommandations formulées par la Commission nationale informatique et libertés (CNIL) dans son rapport du 15 décembre 2017 : « *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*[23] ». Pour éviter l'écueil que tout responsable ne se résume qu'à être un « *manager connecté*[24] » avec un risque de perte de leadership et de sens, la préparation et l'accompagnement des policiers et gendarmes face aux disruptions à venir sont fondamentaux.

Imprégnées des valeurs de l'État de droit, les forces de sécurité intérieure ont besoin de textes juridiques clairs, précis et prévisibles pour que le slogan « *code is law*[25] » soit cantonné au rang de mythe à l'instar de « *Big Brother* » ou de « *Minority report* ». Sur ce point, l'étude annuelle du Conseil d'État, *Le numérique et les droits fondamentaux*[26], invite les pouvoirs publics au travers de cinq propositions concrètes à définir un droit des algorithmes.

(25) *Lawrence Lessig – janvier 2000 – Harvard Magazine* - https://harvardmagazine.com/2000/01/code-is-law-html.

(26) cf. : propositions n° 23 à 27 p. 344 (3.3.2) de l'étude (2014).

Conclusion

(27) Frédéric Gros, « *Le Principe Sécurité* », Essais, Gallimard, p. 176, 2012, 286 p.

Pour le philosophe Frédéric Gros[27], « *un flux (de vie, mais aussi bien d'images, d'informations, de marchandises), dès qu'il est protégé, contrôlé, régulé, détermine un domaine de sécurité* ». Les flux de données, par l'importance vitale qu'ils ont dans notre environnement, relèvent à cet égard du spectre des politiques publiques de sécurité par les caractéristiques dégagées. Ces flux de données vont de pair avec un nouveau droit consacré à la mobilité (tant physique que numérique) que Michel Foucault avait déjà esquissé à travers ce qu'il nomme les circulations[28]. Pour assurer au mieux la sécurité des populations et des gouvernants dans un environnement où les connexions s'accélèrent sans cesse

par le biais des objets connectés, il est nécessaire de disposer d'outils d'aide à la décision de plus en plus performants. Le phénomène « d'infobésité » [29], engendré par la transformation numérique, ne fait qu'amplifier l'état d'incertitude dans lequel les responsables sont placés face à une décision à prendre dans une situation de crise.

Le recours à des algorithmes de recommandation, pour mieux exploiter les systèmes de traitements automatisés, n'est qu'une étape préalable à la modélisation d'une intelligence artificielle au service des dirigeants des politiques publiques de sécurité.

À ce propos, Patrick Perrot voit dans le Web 3.0 que : « À partir de ce concept, une nouvelle forme de police pourrait émerger : une police 3.0 ou 4.0 prenant en compte l'analyse des données massives, l'exploitation des dispositifs connectés et une capacité à fournir des profils individuels par anticipation. La question est de savoir qui sera le maître de cette nouvelle police ? [30] ».

(28) le véritable enjeu pour Foucault est : "*…non plus fixer et marquer le territoire, mais laisser faire les circulations, contrôler les circulations, trier les bonnes et les mauvaises, faire que ça bouge toujours, que ça se déplace sans cesse, que ça aille perpétuellement d'un point à un autre, mais d'une manière telle que les dangers inhérents à cette circulation en soient annulés. Non plus sûreté du prince et de son territoire, mais sécurité de la population et, par conséquent, de ceux qui la gouvernent.*", cf. « *Sécurité, Territoire et population* », leçons du 25 janvier 1978 p. 67, Gallimard et Seuil, 2004, 435 p.

(29) Terme québecois pour désigner un « *État résultant d'une information jugée trop abondante par rapport aux besoins ou aux capacités d'assimilation des utilisateurs.* », cf. Office québecois de la langue française.

(30) Patrick Perrot, « *What about AI in criminal intelligence ? From predictive policing to AI perspectives* », European Police and Research Bulletin Science, Issue 16, Summer 2017, p. 72.

L'AUTEUR

Au terme d'un parcours de 11 ans en unités opérationnelles (commandant de peloton en gendarmerie mobile, commandant de compagnie en gendarmerie départementale), le Chef d'escadron Jérôme LAGASSE rejoint, le 1er août 2014, le Centre de recherche de l'École des officiers de la gendarmerie nationale à Melun (77). Chef du département Recherche-Stratégie, il est doctorant en Science Politique à l'Université Paris V Descartes (École doctorale 262 – Centre Maurice Hauriou). L'intitulé de son sujet de thèse est : « *Mégadonnées et analyse prédictive. Quel nouveau paradigme en matière de politiques de sécurité publique en France ? Enjeux et perspectives pour la Police Nationale et la Gendarmerie.* » http://www.theses.fr/s177243.

Intelligence artificielle
et sécurité : enjeux et perspectives

Par Patrick Perrot

Dystopie ou utopie, le monde de demain sera un monde avec et non sans l'intelligence artificielle (IA). Nous pouvons le craindre, l'espérer, l'accepter ou le nier, l'intelligence artificielle influera sur nos actions, nos réflexions, nos décisions. Apparue dans les années cinquante sous son appellation actuelle, l'intelligence artificielle, hier source de tous les fantasmes, est aujourd'hui entrée dans le quotidien de nos vies. Le volume des données exploitables, les méthodes et outils algorithmiques et l'expansion des « clouds » en sont les principaux responsables. Il ne s'agit plus de savoir si nous devons ou non limiter les algorithmes mais bien de se préparer dès à présent à un bouleversement dans notre manière d'agir et d'interagir avec les algorithmes.

PATRICK PERROT

Colonel de Gendarmerie. Commandant le groupement de gendarmerie départementale de la Haute Marne

L'IA : Janus de la sécurité

Universelle et expansionniste, l'intelligence artificielle concerne l'ensemble des secteurs d'activité mais chacun d'entre eux doit être capable de l'appréhender selon ses spécificités afin de l'exploiter dans des conditions optimales : scientifique, éthique et juridique.
Il y a quelques années, envisager les possibilités offertes par l'IA était plutôt réservé à quelques scientifiques de haut niveau. Aujourd'hui, il apparaît inconcevable de ne pas profiter de ces nouvelles opportunités sans omettre néanmoins l'avenir de la place de l'homme dans le processus de décision. Le champ de la sécurité ne peut échapper à cette évolution tant une utilisation malveillante de l'IA génère de nouvelles capacités

pour la criminalité, tant une utilisation par les forces de l'ordre accroît la sécurité de tout un chacun : deux visages pour une même science.

Janus, dieu du commencement et des passages, regarde à la fois vers le passé et l'avenir, une façon d'ouvrir une nouvelle ère et d'appréhender les sujets de demain par la sagesse d'hier. C'est aussi l'ambition recherchée par les méthodes d'anticipation dédiées à la lutte contre la délinquance. Anticiper la criminalité est une recherche constante des forces de l'ordre, certainement depuis qu'existe la course entre gendarmes et voleurs, mais qui peut prétendre à de nouvelles ambitions par l'IA. Des applications émergent aujourd'hui dans le domaine de la « predictive policing », expression particulièrement mal traduite par « police prédictive » en français. Cette idée de « police prédictive » ressemble plus à un concept marketing qu'à une réalité à la fois conceptuelle et opérationnelle. En effet, il ne s'agit pas de prédire la police mais bien d'appliquer des méthodes d'analyse prédictive, dont l'IA démultiplie les possibles, à la lutte contre la délinquance et à son corollaire : l'activité des forces de sécurité. Au contraire de sociétés privées tels Predpol, Precobs ou encore Keycrime, la Gendarmerie nationale ne se nourrit pas de slogans marketing. C'est la raison pour laquelle elle a développé de A à Z différentes approches en matière d'IA, notamment au pôle judiciaire de la Gendarmerie nationale (PJGN) que ce soit au service central de renseignement criminel (SRC) où à l'institut de recherche criminelle (IRCGN). Les applications réalisées au PJGN, à des fins de lutte contre la criminalité, ont trait à la reconnaissance de locuteur, de texte, faciale, de la démarche, à l'anticipation de la criminalité et à la fusion des différentes méthodes. Les déclinaisons opérationnelles sont légions car en bref l'ensemble du spectre pourrait être couvert par l'IA : compréhension et anticipation de la délinquance au sens large, optimisation de l'activité des forces de sécurité, lutte contre la radicalisation et plus largement contre le terrorisme, sécurité routière. Elles reposent toutes sur une connaissance du passé pour mieux comprendre le présent et anticiper l'avenir.

Déjà utilisée dans la sécurité bancaire, l'IA est particulièrement adaptée pour détecter et identifier des transactions ou des comportements frauduleux dans le domaine du e-commerce et des réseaux sociaux. En matière de lutte contre le terrorisme, les mouvements financiers font déjà l'objet d'une attention particulière mais il est également possible d'alerter sur des comportements déviants ou d'entraîner des réseaux neuronaux à partir des signaux de basculement pour évaluer les risques de radicalisation. Les œuvres de l'esprit seraient difficiles à appréhender par l'IA. Des applications naissent pourtant dans la lutte contre la contrefaçon à l'instar des travaux développés par l'université de Rugters aux Pays Bas. L'analyse de 80 000 traits individuels d'artistes tels que Picasso ou Matisse a permis d'entrainer un réseau de neurones pour identifier les originaux parmi des toiles contrefaites. Des applications similaires sont possibles dans le domaine de la sculpture, de la littérature où l'analyse et le style d'un artiste pourraient parfaitement être authentifiés. En termes de sécurité des mobilités (contrefaçon, trafic d'êtres humains, migrations, vols de fret,...), l'analyse des trajets depuis le pays d'origine en exploitant les différentes vecteurs de mobilité peut également faire l'objet d'un apprentissage afin d'anticiper les déplacements d'intérêt et cibler les contrôles.

Au-delà de la lutte contre la criminalité, l'IA ouvre aussi des perspectives globales dans le domaine de la sécurité routière. C'est un champ à explorer qui pourrait s'avérer particulièrement pertinent dans la compréhension des facteurs d'accidents, dans une analyse prédictive des zones à risque actuelles comme à venir, dans une mise en évidence des comportements et/ou trajets les plus à risque. L'IA est en mesure d'alerter un conducteur sur un changement de comportement qui le fait entrer dans une zone à risque mais aussi sur les risques non plus de sa propre conduite mais du rapport de celle-ci avec l'environnement des autres automobilistes.

Concernant l'activité des forces de sécurité, l'apport de l'IA est envisageable pour l'optimisation de la ressource humaine, de l'emploi des moyens, de la maintenance du matériel et de l'orientation des services. Anticiper de manière objective les difficultés au sein d'une unité du fait d'un manque d'effectif, évaluer l'impact des renforts sur des phénomènes criminels, prévenir le manque de moyens matériels, évaluer le risque dans l'évolution de la délinquance et optimiser le service en vue de proscrire le développement de zones interlopes mais aussi de déserts de sécurité sont quelques-unes des possibilités offertes par l'IA. Pour demain, peut-être après demain, le véhicule autonome permettra de patrouiller sur des secteurs spécifiques et d'enregistrer à l'aide de capteurs des prémices (signaux faibles) de troubles en vue de les prévenir.
Le gendarme pourra alors se consacrer à un temps plus relationnel avec la population et combiner son action de prévention entre un contact humain et la prise en compte de données en provenance de capteurs tels des véhi-

cules autonomes, des drones, des caméras ou tout simplement des ordinateurs doués d'IA. Il ne s'agit pas de déshumaniser l'activité mais plutôt de centrer l'activité humaine sur l'essentiel à savoir la relation, la prise de décision, l'action et l'acceptation de la responsabilité inhérente.

De la décision à la responsabilité de la décision

Que les méthodes d'IA aident à la décision, cela est aujourd'hui évident. L'IA voit mieux, entend mieux, calcule mieux, c'est à dire plus vite et de manière plus exhaustive et plus objective. Alors que reste-il à l'homme ? La responsabilité. Parce que l'IA offre la possibilité de formaliser la connaissance de manière objective, elle constitue un atout considérable dans le processus de décision et permet de réduire l'incertitude. Pourtant, la théorie du chaos comme la physique quantique nous apprennent que le hasard ne pourra disparaître pour laisser place à un déterminisme Laplacien. La décision perdrait d'ailleurs tout son sens dans un univers déterministe.

Le risque immédiat pour les forces de l'ordre n'est pas réellement de décider mais plutôt de savoir qui décide : Google, Facebook (en d'autres termes les GAFA), l'état ou le « citoyen participatif ou collaboratif » via l'exploitation de l'« open data » ? Le risque de tendre vers une ubérisation de la sécurité est dès à présent à considérer. Il s'agit de savoir si nous souhaitons conserver une capacité de contrôle des forces de sécurité et par voie de conséquence de la protection de nos données comme de nos libertés individuelles ou si nous privilégions l'accès de tous pour tous et par tous dans une volonté de transparence absolue. Le défi est de taille et il n'est aucunement gagné d'avance pour les forces étatiques. La solution repose en effet sur la capacité à comprendre l'IA et à développer des applications en toute autonomie afin de s'affranchir d'une vassalisation technologique séduisante à court terme mais néfaste sur le long terme.

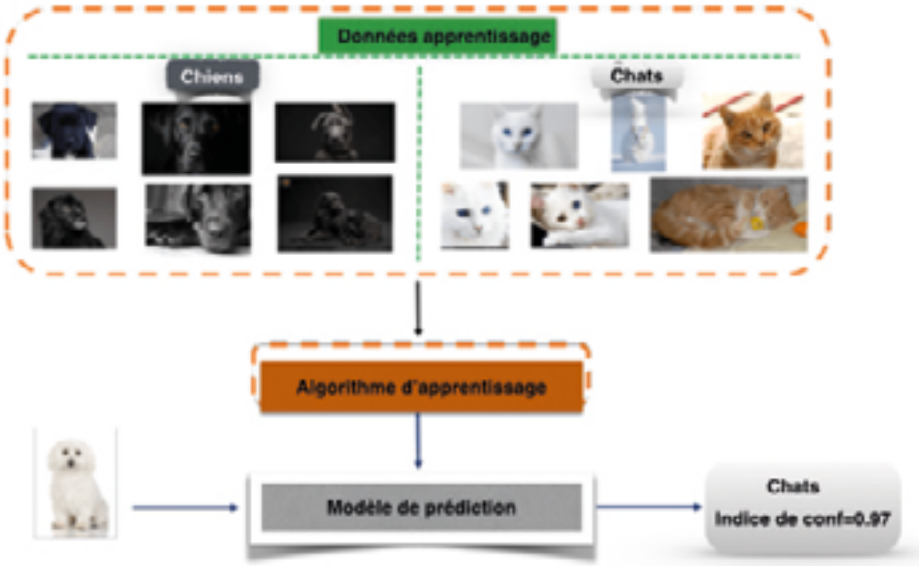
Cette nécessaire compréhension est aussi liée à la notion de responsabilité. Il est difficile de décider entendons-nous souvent. En réalité, la décision n'est difficile qu'en fonction de la responsabilité endossée par le décideur. Le débat ne se situe donc pas sur la décision elle-même mais bien sur la responsabilité qu'oblige cette décision. En matière de sécurité, la question est essentielle et pour être capable d'endosser la responsabilité, la compréhension du fonctionnement et des limites de l'IA dans le processus de décision apparaît indispensable. Exploiter un système d'IA comme une boîte noire représente un risque très objectif d'être manipulé. Même si des travaux existent sur la personnalité juridique d'un robot, la responsabilité doit demeurer humaine, peut-être est-ce cela qui nous place au-dessus de l'IA en dépit de nos piètres capacités calculatoires.

L'IA : enjeux juridiques, enjeux éthiques

Alors que le potentiel du développement de l'IA ne comprend guère de limites théoriques, les enjeux juridiques et éthiques sont vastes et mal définis. Ils ont trait aux données, aux développements scientifiques et aux objectifs poursuivis. Les polémiques autour de la protection des données ne manquent pas, l'affaire Facebook, ou plutôt Cambridge Analytica, sur la « fuite » de dizaine de millions de données à caractère personnel en est une illustration concrète. Au-delà des données, les erreurs d'appréciation d'une IA sont aussi sujettes à débat. L'accident mortel impliquant le prototype de « robotaxi » de la société Uber interroge également le citoyen. Les sujets de controverse autour de l'IA ne manquent pas, il convient néanmoins de cibler où se situent la ou les difficultés.

Concernant la protection des données, il apparaît évident aujourd'hui que nombre de sociétés exploitent des données hors du cadre défini par la CNIL car elles relèvent pour la plupart d'entre-elles d'un droit international. Pour faire face en partie à cette situation a été adopté le règlement général sur la protection des données (RGPD dont l'intitulé officiel est le règlement du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données) qui prendra effet au sein de l'union européenne à compter du 25 mai 2018 en remplacement d'une directive de 1995, la préhistoire dans le monde de la donnée. L'intérêt principal de ce texte est d'abord une harmonisation du panorama juridique européen même s'il apparaît évident que la question dépasse largement le continent.

Concernant les développements scientifiques, ce qui apparaît le plus clairement est la notion de confiance, c'est-à-dire la transparence et la loyauté des applications algorithmiques. Le débat, mettant en cause l'objectivité des algorithmes, source de discriminations et de jugements hâtifs, trouve son origine dans une méconnaissance du principe d'apprentissage pour une partie et dans les ressorts de la théorie mathématique pour une seconde partie. L'IA dispose d'une qualité intrinsèque, elle s'améliore à force d'entraînement mais celui-ci doit être effectué en connaissance de



causes et de conséquences afin d'éviter un sur ou un sous apprentissage. Dans un cas, cela génère un manque de flexibilité, dans l'autre cas, un manque de précision et d'objectivité. Un apprentissage mal défini est bien souvent à l'origine des biais d'interprétation. L'IA serait discriminatoire : raciste, sexiste, et plus encore.

Un biais peut être volontaire ou non, il peut provenir d'un apprentissage volontairement orienté ou d'un comportement mathématique inexplicable (à ce jour) de l'IA. Un algorithme volontairement mal entraîné génère des erreurs telles qu'illustrées par la figure ci-après. Dans une telle situation, il est aisé de rendre discriminante une application logicielle. Ce qui est en cause en réalité ce ne sont pas les erreurs de l'algorithme mais plutôt la probité du concepteur, voire de l'exploitant. Pour s'en prémunir, il convient de connaître les données d'apprentissage dans le détail mais celles-ci relèvent bien souvent d'une politique de confidentialité de l'entreprise. Mal exploitée, l'IA peut alors représenter un danger mais le risque peut lui aussi être détecté et même évité si nous maitrisons la méthode et les données.

L'exemple illustré par la figure est caractéristique d'un biais évident mais il en existe qui sont parfois plus difficiles à appréhender car ils appartiennent aux profondeurs de la théorie mathématique. En d'autres termes, il existe encore des inconnues.

La théorie mathématique sous-jacente aux méthodes de « machine learning », « deep learning » ou encore demain de « quantum learning » est parfois difficile à comprendre. Le fonctionnement des réseaux neuronaux qui apporte de très bons résultats interroge parfois sur certains résultats. Mais cela n'est-t-il pas salutaire en réalité ? Nous reprochons à l'IA de modéliser notre monde trop exactement et en même temps lui reprochons une part d'incertitude. En réalité, ce type de biais que nous pourrions qualifier d'involontaire est bien moindre dans le cas de la machine que dans celui de l'humain. Nous sommes cependant plus enclins à accepter un biais humain plutôt qu'un biais « machine ». Plus qu'un enjeu juridique ou éthique, il s'agit d'une question d'acceptation sociétale.

Afin de compenser une partie des risques identifiés en matière d'IA, il est essentiel de pouvoir développer des expérimentations et des recherches sur des données réelles en dépassant, sous contrôle, le cadre actuel de protection des libertés individuelles. C'est une condition pour évaluer et tester les méthodes comme l'impact sur les enjeux éthiques et juridiques. Le rapport Villani propose « La mise en place d'un environnement propice à l'expérimentation et au développement de l'IA ». C'est certainement la seule façon de connaître et d'approfondir les performances comme les limites scientifiques et éthiques de l'IA et de mieux protéger le citoyen.

Ainsi, l'IA couvre des champs d'application très vastes et certains apparaîtront encore dans les années à venir. Pour autant, l'IA n'est pas tout et tout n'est pas IA. À force de rechercher la transparence par une vulgarisation qui perd de son sens, nous tombons dans une simplification qui n'a plus trait à l'IA. Tout algorithme n'est pas un produit de l'IA, tout ce qui sort d'un ordinateur n'est pas un produit de l'IA. L'IA est d'abord fondée sur une notion d'apprentissage automatique, à partir de laquelle les étapes de modélisation puis de reconnaissance et enfin d'interprétation sont réalisées. Trouver des relations entre des individus, des faits ou des véhicules, établir une fresque chronologique ne relève pas d'un processus d'IA. Cette discipline est trop sensible, trop performante pour entendre des approximations qui, échappant à tout crédit scientifique, fragilisent ses applications. La rencontre de la sécurité et de l'IA est en cours et ne manquera pas de considérablement se développer.

La question est alors de savoir si les forces de l'ordre sont en mesure de participer à cette rencontre où si elles en seront absentes.

L'AUTEUR

Colonel de gendarmerie, Patrick Perrot dirige le groupement de gendarmerie départementale de la Haute Marne. Il a alterné des postes opérationnels et des fonctions de nature scientifique au sein du PJGN. Auteur de nombreuses publications dans le domaine des sciences forensiques et du renseignement, il est ingénieur de formation et titulaire d'un doctorat de Télécoms Paris Tech.

Autonomie des systèmes,
l'heure du choix

Par Jérôme Préteux

Le principe de décision, avec la responsabilité qui en découle, est une prérogative éminemment humaine. L'actualité a montré que les tentatives d'autonomisation de systèmes ont rencontré un accueil virulent de la part du public dès qu'un problème survenait. Peut-on alors imaginer qu'un jour nous accordions aux machines une réelle autonomie ?

IA, mythes, fantasmes et réalités

Je ne ferai pas ici l'affront au lecteur de revenir en détail sur ce qu'est l'Intelligence Artificielle. En revanche, il est intéressant de noter les différentes réactions que peuvent avoir les gens quand on parle d'IA. Un micro trottoir réalisé en 2016[1] sur la vision que les gens pouvaient avoir de l'intelligence artificielle montre le fossé qui peut exister entre l'IA fantasmée proche d'un iRobot ou d'un Matrix et la réalité de systèmes autonomes mais très spécifiques.

(1) « Micro trottoirs – L'intelligence artificielle selon vous », IAE Gustave Eiffel Colloque RH 2017, Youtube.

JÉRÔME PRÉTEUX

Commandant. Urbaniste des systèmes d'information. Centre d'Appui au Systèmes d'Information de la Défense

Le robot pilote, Total Recall, 1990

Faisons donc un rappel sur quelques notions communément admises.

Malgré un florilège de définitions parfois très divergentes, la notion qui revient le plus souvent est celle de l'imitation de l'être

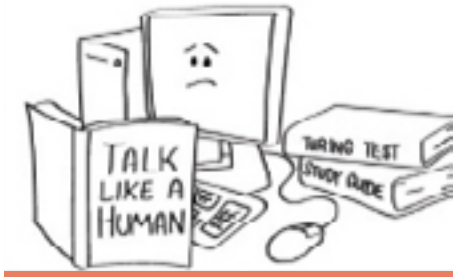
humain. Une distinction est alors faite quant à la perception qu'a la machine de ses propres actions.

Ainsi, une IA sera dite faible quand elle simulera des comportements et des réponses de nature humaine, alors qu'une IA forte aura conscience de ses acions et des réponses qu'elle apportera aux problèmes donnés. Ceci dépasse très largement le cadre scientifique, en faisant notamment référence à la notion de conscience. Cette limitation philosophique des IA ne fait pourtant pas consensus et la communauté scientifique est pour le moins divisée sur le sujet, certains reconnaissant en la conscience une caractéristique du vivant alors que d'autres estiment que la conscience n'est qu'un simple processus chimique modélisable et donc reproductible.

(2) Le test décrit par Turing consiste à mettre en compétition une Intelligence Artificielle (logiciel) et une Intelligence Naturelle (humain). Un observateur, ne voyant aucun des deux compétiteurs, devra, en discutant avec eux, deviner lequel des deux est humain et lequel est artificiel. L'Intelligence Artificielle remportera le test de Turing si l'observateur n'arrive pas à distinguer les deux compétiteurs.

(3) Cette assertion repose sur le "No-Free lunch Theorem", démontré par D. WOLPERT en 1996, qui stipule qu'il n'existe pas d'algorithme universellement bon pour tous les problèmes mais qu'il existe un algorithme optimal pour un couple [Données ; Question].

Cette distinction entre IA forte et faible est née du célèbre test dit « *de Turing* » [2], que le mathématicien a décrit en 1950 dans la revue *Mind*. À l'heure actuelle, il n'existe pas d'IA forte et, mathématiquement, rien ne dit que l'on soit capable d'en produire une dans un avenir proche[3].



Le test de Turing, confondre un humain et un ordinateur, blogdelhombre.com

Un second point de vue pour différencier les IA consiste à les classifier en générale et restreinte :

- Une IA générale, qui se rapprochera fortement du comportement humain, sera en mesure d'appréhender et de résoudre tout type de problème.

- Une IA restreinte sera quant à elle cantonnée à la résolution de certains types de problèmes et sera parfaitement incompétente dès qu'un problème sortira de son champ d'action.

Mais entre mythe et réalité, l'attente autour de l'Intelligence Artificielle est particulièrement forte, et les exigences le sont tout autant.

Des exigences et des attentes fortes

Pour illustrer notre propos, prenons deux exemples concrets, Tay, le chatbot de Microsoft et les voitures autonomes d'Uber et Tesla.

(4) « Microsoft created a Twitter bot to learn from users. It quickly became a racist jerk. » D. VICTOR, the New York Times, 24 mars 2016. « L'IA de Microsoft est-elle réellement devenue raciste au contact des internautes ? », E. LECOMTE, Science et avenir, mars 2016.


Microsoft Tay

Tay[4]

Tay est un *chatbot* spécifiquement développé pour la plateforme *Twitter* par *Microsoft*. Lancé en mars 2016, son but était d'apprendre des interactions qu'il développait avec les humains.

La configuration était la suivante :

- Tay a été configurée comme une adolescente américaine de 18 ans, un peu naïve, fan de pop culture américaine ;

- Tay embarquait une fonction spécifique, « *repeat after me* » qui permettait à un interlocuteur de donner un poids artificiellement augmenté à une assertion qui lui était soumise.

Au bout de quelques heures d'utilisation, le *chatbot* de Microsoft a subi des « attaques coordonnées » de la part des utilisateurs du forum *4chan* et a commencé à tenir des propos jugés déplacés et racistes. Ainsi, après 24 heures et de nombreux dérapages, les équipes de Microsoft ont donc décidé de la débrancher.

Quelles conclusions pouvons-nous tirer de cette expérience ?

Tout d'abord, elle s'est très clairement conclue pour Microsoft par un désastre médiatique, de nombreux articles de presse, et même certains au sein de la communauté scientifique, fustigeant la responsabilité des programmeurs de Tay.

Pourtant l'expérience a également montré la puissance et la rapidité d'apprentissage du *chatbot*. On peut alors la considérer comme une véritable réussite scientifique. La question se pose de savoir si Tay n'a pas été débranchée trop tôt. Aurait-elle été capable de contrebalancer les assertions des utilisateurs de *4chan* en se nourrissant d'autres assertions plus moralement correctes ? Et si oui… au bout de combien de temps ?

Enfin, cette expérience montre que Tay pouvait être considérée comme une IA faible, qui n'avait pas conscience des réponses qu'elle apportait, même si celles-ci étaient sémantiquement correctes et avaient du sens pour ses interlocuteurs.

Voitures autonomes

De nombreux acteurs, dont les trois principaux sont *Tesla*, *Uber* et *Google*, se sont lancés dans une course aux voitures autonomes. Cette course effrénée a été émaillée de quelques incidents qui ont eu un retentissement médiatique particulièrement important, montrant par-là les vulnérabilités de ce type de systèmes.

A ring of salt will protect you, une voiture autonome prise au piège, Nerdist.com

Il est à noter que, depuis le début des expérimentations, 3 morts par accidents ont été le fait de voitures autonomes (ou en tout cas disposant d'une autonomie partielle renforcée, ie sous supervision d'un conducteur humain). 3 morts pour plusieurs centaines de millions de kilomètres parcourus. Les chiffres fiables manquent mais voici une comparaison intéressante :

- Tesla, c'est un mort pour plus de 200 millions de kilomètres (2016).

- En France en 2013 nous avions à déplorer un peu moins de 4 000 morts sur les routes pour 725 milliards de kilomètres parcourus, soit un mort pour environ 180 millions de kilomètres[5].

Le constructeur a même annoncé que dans le monde, un accident mortel survenait tous les 97 millions de kilomètres parcourus[6].

La comparaison mérite bien évidemment d'être affinée et prise avec des pincettes. Mais notre perception de cet accident aura été que les voitures autonomes ne sont pas sûres ! De plus, ce qui est intéressant à noter, c'est la large couverture médiatique qu'ont entrainée les affaires de Tesla et plus récemment d'Uber.

Dans le cas particulier d'Uber, le visionnage de la vidéo extérieure de la voiture montre très clairement qu'un conducteur humain n'aurait pas mieux fait que la voiture autonome et d'ailleurs la question de la responsabilité du piéton, qui traverse une deux fois deux voies non éclairée en dehors de tout passage réservé, peut légitimement se poser. Mais voilà, il s'agissait d'une voiture autonome et cet accident a fait grand bruit à tel point qu'Uber a décidé de suspendre ses expérimentations[7].

Nous voyons donc par le prisme de ces deux exemples que des expérimentations ont dû brutalement s'arrêter car les conséquences n'étaient pas acceptées (ou acceptables…).

Les questions sous-jacentes à ces deux exemples sont simples et pourtant d'une importance capitale. Il est admis qu'un humain puisse se tromper. Il assumera alors les conséquences de ses actes. Mais est-il acceptable qu'une IA se trompe ? Sur qui ou quoi reposera la responsabilité d'une erreur qui aura des conséquences parfois lourdes ?

(5) Statistique de l'observatoire national interministériel de la sécurité routière, site du ministère de l'intérieur, interieur.gouv.fr

(6) « Premier accident mortel pour le pilote automatique de Tesla », J. MARIN, Blog SILICON 2.0, Le Monde, 01 juillet 2016.

(7) « Uber suspend son programme de voitures autonomes après un accident mortel. », L'Express.fr, 19 mars 2018.

Minimiser les risques

De ces questions éthiques fondamentales ressortent deux axes principaux sur lesquels il est nécessaire d'avoir un impact.

Tout d'abord il est primordial d'assurer un cadre juridique le plus clair possible autour de l'IA. De nombreux travaux sont en cours sur ce sujet. De la création d'une personne robot, au même titre qu'une personne morale, à la définition de la responsabilité des développeurs durant la phase d'apprentissage ou encore celle des exploitants durant la phase d'utilisation, plusieurs pistes, en particulier au niveau européen, sont explorées[8]. L'actualité récente, avec l'affaire Cambridge Analytica, montre d'ailleurs l'intérêt que le monde anglo-saxon a apporté au règlement général de protection des données et, au-delà même du texte, à l'idée sous-jacente de l'éthique des algorithmes dont l'Europe est porteuse[9].

(8) « Le projet de résolution sur le droit des robots adopté en commission au Parlement européen », M. REES, Nextinpact.com, 13 janvier 2017.

(9) « Nous devons instaurer de la confiance dans nos technologies, et cela commence par assumer notre responsabilité dans les algorithmes et expériences que nous concevons. » S. NADELLA, Microsoft Build 2017.

Il sera d'ailleurs intéressant de noter la posture que prendra le congrès américain sur cette affaire. Les nations sont le socle de tout cadre juridique à même de défendre les peuples qu'elles représentent, et l'on a pu voir les effets délétères que cela pouvait avoir d'abandonner ce rôle. Le congrès américain reprendra-t-il ses prérogatives d'être le seul juge de ce qui va à l'encontre de l'intérêt de son peuple qui l'a élu pour cela ? Les grands noms du numérique sont des acteurs incontournables de l'avenir des peuples mais ils n'en sont pas les dépositaires uniques. Ils devront alors assumer leur responsabilité, comme l'a fait remarquer Satya NADELLA, patron de Microsoft, mais ils devront le faire en lien avec les nations.

Ensuite il est nécessaire de mettre en place les mesures de sécurité adéquates pour accorder la confiance la plus grande possible dans les conclusions que peuvent fournir les algorithmes. En effet, au-delà des menaces classiques qui pèsent sur tous systèmes d'information, il en existe des spécifiques aux systèmes d'intelligence artificielle et en particulier à l'algorithme de machine learning qu'ils embarquent[10].

(10) Pour aller plus loin, lire l'article « Sécuriser l'intelligence artificielle : il n'est pas trop tard ! », G. BILLOIS, Les Échos, 29 mars 2018.

Ces menaces peuvent être contrées par des mesures de sécurité particulières qui porteront principalement sur les données d'entrée du système, que ce soit durant la phase d'apprentissage ou celle d'utilisation. Ainsi l'attention que l'on portera à l'apprentissage du système et à la base d'apprentissage qu'il utilisera déterminera la qualité des conclusions qu'il prendra.

Pourtant, bien que nécessaires, ces deux axes sont encore embryonnaires et de nombreux travaux sont actuellement menés dans ces domaines.

À la fin, il faudra choisir

Il est donc nécessaire de mener la réflexion éthique en parallèle de ces travaux afin de pouvoir, quand les conditions seront satisfaites, répondre à cette simple question : sommes-nous, oui ou non, prêts à laisser un algorithme prendre une décision à notre place ?

Comme nous avons pu le voir avec les exemples de Tay et des voitures autonomes, des essais ont été réalisés pour laisser une autonomie quasi complète à ces systèmes. Des « erreurs » étaient alors inévitables. De l'acceptation de ces erreurs dépendra la façon dont l'IA évoluera, tout comme les erreurs que nous commettons façonnent notre attitude et nous éclairent parfois plus sur nous-mêmes que nos réussites. Il conviendra alors de choisir judicieusement. Sommes-nous prêts à accepter les erreurs des algorithmes et à accélérer la recherche ? Ou bien souhaitons nous garder la main et limiter nos ambitions en matière d'IA ?

Pour ma part, en tant que chef militaire, je refuse catégoriquement que me soit retirée la responsabilité de la prise de décision, mais ce qui est sûr c'est que ce choix dépendra de l'humanité tout entière et non de quelques personnes, fussent-elles expertes dans le domaine, assises dans le confort d'un bureau californien.

L'AUTEUR

Ingénieur en informatique, diplômé de l'École Spéciale Militaire de Saint Cyr en 2004, le chef d'escadron Jérôme PRETEUX a travaillé dix ans au sein de la logistique opérationnelle de l'armée de terre. Breveté de l'École de Guerre en 2015, il obtient le diplôme de mastère spécialisé en management des systèmes d'information à l'école Centrale Supélec en 2017. Passionné par les nouvelles technologies, il a rédigé dans ce cadre une thèse « IA et cybersécurité » au sein du cabinet de conseil Wavestone. Il sert actuellement en tant qu'urbaniste des systèmes d'information au sein du Centre d'Appui au Systèmes d'Information de la Défense.

La liberté des décideurs

de la sécurité publique à l'épreuve des algorithmes prédictifs

Par Cyril Piotrowicz

Les algorithmes prédictifs, qu'ils soient désignés par les termes spécifiques de police prédictive ou d'analyse décisionnelle, permettent de renseigner les forces de sécurité publique sur les zones à forte capacité criminogène au regard notamment des infractions enregistrées. Souvent qualifiés d'outils d'aide à la décision en ce qu'ils permettent d'éclairer et d'informer l'autorité dotée du pouvoir décisionnaire, notamment en orientant les effectifs disponibles, ils soulèvent alors la question de l'autonomie de la prise de décision.

CYRIL PIOTROWICZ

Doctorant associé Centre de Recherche de l'École Nationale Supérieure de la Police

La prise de décision n'a pas attendu les algorithmes prédictifs pour être influencée et certains aspects ou méthodes sont d'ores et déjà encadrés par la loi au travers du rôle du consultant, du conseil, de l'expert ou de l'informateur. Mais force est de constater qu'*a priori* les algorithmes prédictifs ne s'intègrent à aucune de ces catégories puisque celles-ci sont réservées aux personnes physiques ou morales.

Or, face à une réalité objectivée, à une situation rationalisée à l'extrême, quelle est la place de l'analyse objective ou de l'appréciation subjective du décideur ?
Peut-il librement s'écarter du résultat fourni par le logiciel ?

De la nature de l'aide à la décision

Les outils de cartographie prévisionnelle du phénomène criminel peuvent faire appel à deux formes de technologies :
la première est le reflet d'un raisonnement humain et pourrait être qualifiée de dépendante de celui-ci, elle a recours aux méthodes d'apprentissage automa-

tique (*machine learning*) supervisé ou semi-supervisé, tandis que la seconde, que l'on pourrait qualifier d'autonome à la logique humaine aura recours à une méthode d'apprentissage automatique non-supervisé.

Dans l'hypothèse d'un apprentissage supervisé ou semi-supervisé, l'être humain reste au centre de la technologie, c'est le *datascientist* qui indique à l'algorithme quelles sont les données pertinentes à prendre en considération. Le processus décisionnel étant connu on peut le qualifier de *white-box*. Le raisonnement intellectuel n'est pas créé par l'algorithme mais par l'être humain qui vient simplement déléguer l'exécution de celui-ci soit pour bénéficier d'un gain de temps soit parce que la masse de données à traiter dépasse ses propres capacités. On peut donc considérer que le *datascientist* demeure le spécialiste puisqu'il est titulaire de la technique et du raisonnement. Nous sommes alors face à une personne physique ou morale qui met à disposition du décideur son expertise, sa technique, par le biais d'un logiciel.

L'apprentissage non-supervisé privilégie une autonomie du logiciel : après avoir inséré les données brutes, le *datascientist* laisse au logiciel le soin de faire les liens qui lui apparaissent comme logiques entre une donnée et un évènement, au risque de ne pouvoir ni identifier la causalité, ni en vérifier la pertinence.

C'est un processus décisionnel inconnu, que l'on peut qualifier de *black-box*. Dans le cadre d'un apprentissage automatique non-supervisé, l'être humain s'efface pour laisser l'algorithme créer son propre raisonnement logique, nous sommes dès lors confrontés à un logiciel artificiellement autonome capable de proposer une solution inédite, inconnue à l'intellect humain. Par conséquent, le rôle d'expert n'est plus tenu par le *datascientist* mais par l'algorithme lui-même.

De l'aide à la prise de décision

Si les algorithmes prédictifs sont capables de renseigner l'acteur humain sur la décision à prendre, il convient de savoir à qui va incomber la prise de décision.

Bien que notre analyse soit ici limitée au pouvoir de direction et d'organisation du service de police ou de gendarmerie, on peut d'ores et déjà souligner que l'article 10 de la loi Informatique et Liberté et l'article 22 du nouveau RGPD, qui interdisent toute décision individuelle fondée exclusivement sur un algorithme, n'auront pas vocation à s'appliquer à la cartographie prévisionnelle du phénomène criminel. En effet, les mesures d'organisation du service (comme l'orientation des patrouilles) ne sont pas des décisions individuelles produisant des effets juridiques. Il n'existe donc, en théorie, pas d'obstacle légal à une automatisation de l'orientation des patrouilles.

DOSSIER
LA LIBERTÉ DES DÉCIDEURS DE LA SÉCURITÉ PUBLIQUE À L'ÉPREUVE DES ALGORITHMES PRÉDICTIFS

En revanche, la solution apparaît différente dès lors que l'on se place sur le volet opérationnel de la mission de sécurité publique (décision de contrôle d'identité, placement en garde à vue, etc.).

(1) Art. 2, décret du 2 août 2005 portant statut particulier du corps de conception et de direction de la police nationale.

(2) Art. 1, décret du 12 septembre 2008 portant statut particulier du corps des officiers de gendarmerie.

(3) Art. R434-4, Code de la sécurité intérieure.

Toutefois, si la loi ne s'oppose pas à une organisation algorithmique et automatisée du service, d'autres textes sont des freins bienvenus. Ainsi, au niveau opérationnel, les commissaires de police[1] et les officiers de la gendarmerie nationale[2] sont chargés de la direction des services dont ils assument la responsabilité opérationnelle et organique.

Les règles déontologiques précisent que : « *L'autorité investie du pouvoir hiérarchique prend des décisions, donne des ordres et les fait appliquer* » [3].
Dès lors, la question semble tranchée : peu importe que le décideur opérationnel s'adjoigne l'expertise technique d'une personne physique ou morale ou d'un outil, en white ou black-box, il demeure seul décideur. Il nous semble important que ce principe soit rappelé à l'ensemble des décideurs opérationnels à qui est confié un outil d'aide à la décision, *a fortiori* si celui-ci est artificiellement autonome.

De la décision aux responsabilités

Si, y compris en matière d'organisation du service, les algorithmes prédictifs ne peuvent se substituer à l'acteur humain, leur rôle d'aide interroge sur l'autonomie de la décision, question qui ne peut se concevoir sans la notion de responsabilité.

(4) Art. R434-4, Code de la sécurité intérieure.

(5) Art. R434-7, Code de la sécurité intérieure.

Encore une fois, les règles déontologiques précisent que « *l'autorité investie assume la responsabilité des ordres donnés.* » [4] Ainsi, en sus d'être seul décideur, le commissaire de police ou l'officier de gendarmerie sont également seuls responsables des ordres donnés. Toutefois, de nombreux facteurs peuvent venir influencer son autonomie décisionnelle : sa hiérarchie, sa considération personnelle pour la technologie, ses relations avec les tiers, etc. Quand bien même le décideur bénéficie de la protection juridique de l'État, ce sont autant de responsabilités, parfois moins formelles, qui lui incombent : professionnelle, organisationnelle, managériale, etc.

Il nous apparaît essentiel que les outils d'aide à la décision ne doivent pas conduire à une déresponsabilisation du décideur public, de surcroît dans des activités régaliennes. À l'inverse, ces outils doivent contribuer à renforcer l'indépendance de l'instance décisionnaire, non

© Champosan - Adobe stock

Le décideur public reste libre d'apprécier la valeur d'une expertise logicielle. Il se pose toutefois la question de la responsabilité du décideur qui fonde sa décision sur une Blackbox aux références opaques et dépourvue de personnalité juridique.

pas en la confrontant à l'algorithme mais en mettant celui-ci à la disposition du décideur qui demeure libre d'en suivre, ou non, les recommandations en engageant dès lors sa responsabilité sur le fondement de la décision prise.

De la responsabilité des personnes à la responsabilité technologique

Si le décideur demeure responsable de toutes les décisions qu'il prend, qu'en est-il dans l'hypothèse où il est informé de manière incomplète ou erronée par un système d'aide à la décision défaillant ou dysfonctionnant ?

(6) Pour des raisons éditoriales, les comparaisons avec les autres types d'expertises ou de conseil ne seront pas développées ici.

D'une manière générale, la question de la responsabilité de l'expert ou du cabinet de conseil dans l'exercice d'un service public régalien est délicate. *À priori*[6], une analogie peut être faite entre la sécurité publique et la Justice : le commissaire ou l'officier de gendarmerie et le juge, après avoir écouté les conclusions de l'expert et sans avoir l'obligation d'en suivre les recommandations, prennent librement une décision pour laquelle ils n'engagent pas leurs responsabilités personnelles[7]. À l'inverse, l'expert judiciaire engage sa responsabilité civile, et éventuellement pénale, en cas de faute, y compris légère, et ce même si le juge a suivi ses recommandations[8].

(7) Art. 11-1, ordonnance n° 58-1 270 du 22 décembre 1958 portant loi organique relative au statut de la magistrature.

(8) Civ. 2e, 8 oct. 1986, n° 85-14.201.

En présence d'un algorithme de type White-box, il nous semble que cette solution s'applique sans difficulté car la décision de l'algorithme, même apprenante et adaptative, reste sous la contrôle et la supervision de l'expert développeur de l'algorithme.

En revanche, face à une technologie *a minima* artificiellement autonome, ou *black-box*, c'est-à-dire capable de raisonner sans la supervision d'une personne physique susceptible de voir sa responsabilité ou celle de la personne morale à laquelle elle appartient, engagée, il est nécessaire, en raison de la complexité de la matière et du nombre d'acteurs impliqués (développeurs de l'algorithme, distributeurs du logiciel, récolteurs ou fournisseur des données, utilisateur du logiciel, etc.) que le législateur intervienne pour désigner celui sur lequel la responsabilité juridique reposera. Dans cette hypothèse, à l'instar des personnes morales, face à une entité insusceptible d'engager sa responsabilité – une machine ne disposant pas de la personnalité juridique – *le recours à un substratum personne physique ou morale* devient alors nécessaire.

Si cela est d'ores et déjà le cas avec l'hypothèse des voitures autonomes où la responsabilité incombera aux « *titulaires de l'autorisation en charge des expérimentations* » [9], qu'en sera-t-il des algorithmes prédictifs ?

(9) DUPIN (M.), Propos sur le projet de loi Pacte, RMC, 06/04/2018.

In fine, si les algorithmes prédictifs peuvent valablement être qualifiés d'outils d'aide à la décision, il nous semble important de souligner qu'ils vont contribuer à renforcer l'autonomie décisionnelle des décideurs publics en élargissant leur champ des possibilités, sans influencer leur régime de responsabilité. La difficulté ne semble en réalité apparaître que dans une situation particulière d'un outil intelligent, défaillant ou dysfonctionnant.

L'AUTEUR

Cyril PIOTROWICZ est doctorant en droit privé et sciences criminelles à l'Université Jean-Moulin Lyon III et à l'Université de Technologie de Troyes. Il est doctorant associé près du Centre de Recherche de l'École Nationale Supérieure de la Police. Auditeur de l'INHESJ (6e session régionale) et diplômé en criminologie, il s'est spécialisé dans la prévention du phénomène criminel et les nouvelles technologies et notamment l'étude des outils de « predictive policing ».

CENTRE DE RECHERCHE DE L'ENSP

Le Centre de Recherche de l'ENSP, dirigé par un professeur habilité à diriger des Recherches, a pour mission d'assurer l'interface entre les équipes de recherche et les différentes directions de la police nationale. Les travaux sont architecturés au travers de 4 axes de recherche: le lien entre police et population, la Prise de décision (situations complexes et environnements numériques), le Management situationnel et qualité de vie au travail et la synténique entre sciences humaines et sociales et sciences de l'information et des technologies.

Le Centre de recherche maintient également une veille sur l'impact de l'innovation et des développements technologiques dans le domaine de la sécurité publique et privée. Les projets actuels concernent: ILEAnet (Horizon 2020) – innovation par le réseautage des acteurs publics de la sécurité intérieure, TARGET (Horizon 2020) - Simulateur de prise de décision en intervention et IDFRAud (ANR – Agence Nationale de la Recherche) - Framework Automatique et Opérationnel pour la Détection et le Profilage de Fraude sur l'Identité.

Le coup *data*

ou le renversement du pouvoir par la donnée

Questions à Jean-Pierre Mignard et Adrien Basdevant

RG : Au regard du développement de l'intelligence artificielle, quels sont selon vous les impacts qu'il est possible d'identifier en rapport avec les algorithmes et l'exploitation des données massives dans les pays régis par les principes de l'État de droit ?

Comme le rappelle fréquemment le député Cédric Villani, médaille Fiels[1], chargé par le gouvernement d'un rapport sur le sujet, il n'existe pas de bonne définition de l'intelligence artificielle (IA). Cette notion renvoie en réalité à des techniques diverses ainsi qu'à des nombreux phantasmes. Afin d'analyser les impacts de cette « *science qui consiste à faire faire aux machines ce que l'homme ferait moyennant une certaine intelligence* » [2], il peut être intéressant d'envisager l'IA comme une « *nouvelle catégorie d'algorithmes, paramétrés à partir de techniques dites d'apprentissage* ». Autrement dit, « *les instructions à exécuter ne sont plus programmées explicitement par un développeur humain, elles sont en fait générées par la machine elle-même, qui "apprend" à partir des données qui lui sont fournies* »[3]. Les applications sont extrêmement variées, elles peuvent concerner les techniques d'enquête, la reconnaissance

JEAN-PIERRE MIGNARD

Associé au cabinet Lysias Partners

ADRIEN BASDEVANT

Associé au cabinet Lysias Partners

(1) La médaille Fields est avec le prix Abel une des deux plus prestigieuses récompenses en mathématiques. Elle est considérée comme un équivalent du prix Nobel car il n'en existe pas pour cette discipline.

(2) Définition de l'IA par l'un de ses créateurs, le scientifique américain Marvin Minsky.

(3) apport de la CNIL de décembre 2017, Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_garder_la_main_web.pdf.

d'images, le calcul de risque de récidive, le profilage d'individus à risque criminel ou terroriste. On voit ainsi poindre directement un problème majeur posé par le développement de cette nouvelle classe d'algorithmes : la possibilité de les comprendre, de les interroger, voire de les contester. En effet, les logiques sous-jacentes peuvent rester incompréhensibles, y compris à ceux qui les conçoivent et les développent. Dès lors, le fait de déléguer nos décisions quotidiennes à ces boites noires soulèvent plusieurs enjeux démocratiques de première importance. Serons-nous en mesure de débattre du fonctionnement d'outils que nous ne sommes pas toujours capables de comprendre ? Risquons-nous de diluer notre responsabilité en nous en remettant à leur diagnostic ? Comment vérifier les critères utilisés par ces algorithmes ? Ceux-ci sont-ils biaisés ou discriminants ? Sommes-nous voués à entretenir une nouvelle forme de servitude volontaire ? Dans notre livre « L'Empire des données », nous parlons d'un « Coup Data » pour évoquer l'idée qu'à l'heure de l'intelligence artificielle nourrie par des données massives, une gouvernance algorithmique convaincrait aisément du temps épargné et du moindre coût des processus de décision publique ou plus généralement sociale. Mais cela serait faire fi de l'essentiel, car c'est le processus de la décision qui conditionne la légitimité de l'adhésion du corps social et la responsabilité de celui qui la prend. C'est bien pourquoi la question de la délibération démocratique est plus que jamais posée. Or, la faculté de délibérer et de faire entendre sa voix en tant que citoyen, c'est-à-dire en tant que « sujet de droit » serait mise à mal par un « Coup Data » dont le processus ultime ne serait pas enrayé. Contrairement à la loi, les algorithmes ne considèrent pas les personnes comme des sujets de droit, pris dans leur individualité, c'est-à-dire avec leur motivation, leur passé, leur histoire ; mais les appréhendent comme des fragments de données appartenant à un vaste flux qu'il s'agit de calculer et modéliser.

© Cyber crime in secure network, binary code - Lagarto Film

Une personne ne peut être amputée de son contexte de vie, de son histoire et de la causalité de ses actes par un choix de données parcellaires et une logique sous-jacente d'algorithmes qui peut s'avérer spécieuse et inaccessible juridiquement.

RG : Sur le plan de la méthodologie statistique, vous soulignez un changement de paradigme, à savoir le passage d'une méthode déductive à une méthode inductive, où il ne s'agit plus de rechercher les causes, de calculer des moyennes et tendances à partir d'échantillons. Pouvez-vous décrire cette transformation ?

En s'affranchissant de la notion statistique, jusqu'alors majeure, de la « moyenne », le Big data – à savoir l'exploitation massive de données par des algorithmes – marque un véritable tournant dans l'histoire de l'information. Auparavant, tout n'était pas mesurable. On se concentrait donc volontairement sur une partie de la population qui servait d'échantillon, qu'on allait ensuite extrapoler, pour effectuer des mesures statistiques. Désormais, plus question d'éviter les points de données trop écartés de la moyenne, le Big data prend tout en compte, indistinctement. On détecte ce qui était auparavant non observable ou inaperçu. Par conséquent, le traitement des données devient plus fluide et change radicalement la logique de l'opération. Il ne s'agit plus de déduire, c'est-à-dire de partir d'hypothèses et de modèles pour aboutir à une conclusion, mais d'induire une loi générale à partir de l'observation des faits. Le Big data porte ainsi la croyance de pouvoir accéder directement au réel, en collectant et mesurant tous les signaux, sans plus devoir interpréter ni questionner leurs contenus. On s'en remet à eux de manière quasi-divine. La rationalité reliant les phénomènes aux causes, ainsi que la place qu'occupait la théorie dans la science, sont remis en cause. La corrélation remplace la causalité. L'illusion scientifique des algorithmes réside dans le fait qu'ils nous offrent la possibilité d'étudier de nouveaux phénomènes complexes, auparavant imperceptibles, sans pour autant nous permettre d'en comprendre et connaître les causes. Le profilage des individus qui peut en résulter ne vise ainsi plus le corps statistique de la population telle qu'elle est. Il cible ce que ce corps pourrait potentiellement devenir : par exemple le fraudeur potentiel, le terroriste potentiel en matière de politique criminelle. L'aléa, le risque sont désormais relayés au domaine subjectif qu'il convient d'écarter au profit

d'une vérité numérique objective révélée par ces outils algorithmiques.

RG :Les risques d'enfermement algorithmique (la claustration comme vous l'évoquez dans votre ouvrage) : ce point me semble très important pour le lecteur-citoyen qui peut, s'il ne prête garde, perdre son libre-arbitre et sa capacité critique.

Si le mode de conception des profils a pour conséquence directe de leur prêter des désirs ou des besoins jamais ressentis ou exprimées auparavant, le risque est délibérément pris de les enfermer dans ces avatars comportementaux. Ces méthodes favorisent un possible enfermement des personnes dans leur propre « bulle de données » modifiant progressivement leur perception de l'information. Cette « *Internet filter bubble* » décrite par l'auteur américain Eli Pariser individualise l'approche d'Internet, au risque d'un enfermement algorithmique. Les résultats hyper-personnalisés des recherches effectuées sur un moteur de recherche ou des contenus accessibles via un réseau social pourraient interdire toute confrontation avec des opinions diverses ou contradictoires, en nous berçant dans l'idée de la douce présence d'un processus neutre et objectif. Par exemple, un utilisateur américain de Facebook n'ayant dans son entourage que des connaissances votant pour le camp démocrate ou républicain, aura certainement plus de probabilités de consulter des contenus en affinité avec ces préférences. Prisonnier de choix affinitaires, homogènes, lisant, accédant et partageant uniquement des contenus de personnes similaires, aux centres d'intérêts partagés, la claustration guette en effet.

RG :Quelles sont les interconnexions qu'il serait possible de déceler entre la police prédictive et la justice prédictive ?

Dans les deux cas, les termes « police prédictive » et « justice prédictive » sont malencontreux. Il s'agit tout au plus de police/justice descriptive ou indicative et cette description de la réalité n'est que partielle. Le principe est de prédire la survenance de futurs événements à partir de la réalisation d'événement déjà passés, que ce soit une infraction ou une décision de justice.
À titre d'illustration, pour détecter et réagir en temps réel aux crimes, le système de détection de coup de feu *Shotspotter* fournit la localisation d'un tir par triangulation quelques secondes seulement après son déclenchement. Le Big data serait alors utilisé pour prédire et éviter les crimes. Selon les mêmes principes que la méthode utilisée pour anticiper les répliques de tremblement de terre, le logiciel américain *Predpol* ou français *Predvol* mesure la concentration et la contagion du crime pour définir l'intensité du risque. Si l'algorithme conclut qu'une zone est dangereuse, les forces de l'ordre y sont dépêchées et procèdent à des arrestations au détriment d'investigations dans d'autres secteurs. Par un effet mécanique, les arrestations augmentent dans ces « *hots spots* ». Il existe ainsi un risque évident de voir l'algorithme se nourrir lui-même et gonfler ses propres prédictions, ce qu'on appelle les « *self-fulfiling pedictions* ».
Les algorithmes prédictifs risquent de se concentrer sur les effets et les statistiques, avec les conséquences que l'on sait en termes d'allocation de moyens sécuritaires au détriment de l'identification des causes notamment socio-économiques de la criminalité ou de déserter d'autres terrains à égale densité criminogène.

Pour les décisions de justice, un risque similaire existe avec des logiciels qui permettraient de calculer les chances de succès d'une action, en fonction de plusieurs paramètres (juridiction, juges composant le tribunal, précédentes jurisprudences rendues sur des cas similaires, etc.). Nous devons veiller à déterminer une affaire par rapport au droit, et non exclusivement à partir des données collectées.
En matière judiciaire, les procès composent toujours avec une part d'incertitude, dans un contexte particulier. Chaque cas, en particulier au pénal, est unique. Il faudra donc voir comment ces outils s'adapteront à l'évolutivité rapide de règles juridiques de plus en plus enchevêtrées et techniques. Enfin, cela pose avec insistance la question de la culture statistique des magistrats qui s'appuieront sur ces outils pour prendre leurs décisions et qui jusqu'alors n'ont jamais été formés pour interpréter de tels résultats. Il leur faudra apprendre à s'en dégager comme ils le devaient déjà vis-à-vis des jurisprudences, au risque de se rallier par conformisme paresseux à une jurisprudence constante.

Cela vaut aussi pour les enquêteurs, c'est le principe de la police scientifique. Un exemple concret permet de mesurer l'ampleur des risques associés à ces nouvelles pratiques si elles ne sont pas réfléchies. Le logiciel *Compas*, développé par l'entreprise *Northpointe*, est utilisé par certains comtés américains pour décider de l'incarcération d'un suspect avant son procès ou pour évaluer les probabilités de récidive. Eric Loomis, un citoyen américain du Wisconsin, a ainsi été condamné à six années d'emprisonnement suite à un verdict du logiciel, dont le rapport numérique précisait que l'accusé présentait un « haut risque de violence, un haut risque de récidive ». Le juge a suivi cette prédiction. L'accusé a fait appel, en arguant de la violation de ses droits par l'usage d'un algorithme qui lui est inaccessible. L'entreprise qui détient le logiciel faisait valoir que cet algorithme est protégé par un secret commercial. Elle a par conséquent refusé à la justice comme à la personne mise en cause de prendre connaissance de la formule mathématique de son algorithme. L'avis du gouvernement fédéral a été requis. La Cour suprême du Wisconsin a évité de répondre à cette question embarrassante et a confirmé la décision, considérant que l'accusé aurait écopé de la même peine sur la base de constatations traditionnelles, d'autant que ce dernier

avait un casier judiciaire chargé. Cet exemple soulève d'importantes questions éthiques. Une étude de quatre journalistes ayant comparé les parcours de dix mille détenus aux prévisions proposées par l'algorithme, a révélé d'importants biais ethniques : les Noirs étaient systématiquement considérés comme de potentiels récidivistes, un des critères décisifs selon l'algorithme étant la couleur de peau[4].

(4) Julia Angwin, Jeff Larson, Surya Mattu et Lauren Kirchner, « Machine Bias », ProPublica, 23 mai 2016 : « There's software used across the country to predict future criminals. And it's biased against blacks ».

Il ne faut pas sous-estimer l'aveuglement éthique des algorithmes, de leurs concepteurs et de leurs utilisateurs.

RG :Y aura-t-il toujours une place pour la présomption d'innocence à l'ère du tout calculé ?

Il faut s'assurer qu'un algorithme ne reflète pas les préjugés de ses concepteurs dans l'inventaire des critères utilisés pour déterminer les menaces et les risques. Ces enjeux permettront sans doute d'aborder enfin des questions préexistantes restées encore taboues jusqu'aujourd'hui. Il doit exister des raisons objectives avant toute fouille policière. Mais qu'en sera-t-il lorsque l'algorithme aura recommandé cette mesure ? Le respect du principe selon lequel toute personne qui se voit reprocher une infraction est réputée innocente tant que sa culpabilité n'a pas été légalement établie devra être défendu à l'heure des algorithmes. Les enquêteurs doivent aussi faire preuve de la plus grande perspicacité, comme les juges, ou les avocats qui ne plaideraient plus que les circonstances atténuantes et considéreraient que les probabilités égalent les faits. La présomption d'innocence est un principe fondamental de notre État de droit, et l'une des premières caractéristiques de nos démocraties libérales permettant de savoir les faits qui nous sont reprochés, de participer aux audiences et d'être entendus. Que se passera-t-il lorsque les données désigneront des criminels avant même qu'ils n'aient commis leurs crimes ?
Que restera-t-il de la présomption d'innocence pour celui qui présente les caractéristiques d'un multirécidiviste ? Il s'agirait alors de glisser insensiblement du commencement d'exécution à l'acte préparatoire, puis de l'acte préparatoire à la potentialité de commettre un crime. Ce renversement résume ainsi le danger produit par le bouleversement des théories du droit classique pénal libéral. Comme le résume parfaitement Mireille Delmas-Marty : « *Prétendre prédire le passage à l'acte, détecter l'intention, c'est déjà une forme de déshumanisation parce que le propre de l'homme est dans l'indétermination : sans l'indétermination, on n'est plus responsable de rien* » [5]. Seuls les faits comptent dans un système de droit libéral.

(5) Mireille Delmas-Marty, La démocratie dans les bras de Big Brother, Le Monde Blogs, 10 juin 2015.

Ceci se confirme dans les procédures pénales des sociétés démocratiques, qui ne prescrivent des expertises psychologiques et des enquêtes de personnalité qu'après le constat de l'existence « *d'indices graves et concordants de participation à un crime ou un délit* », selon les termes notamment des articles 156 et 161 du Code de procédure pénale. Les nouvelles technologies de prédiction sont conçues sur des critères de dangerosité des individus (c'est-à-dire leurs actions possibles dans le futur) et non sur l'évidence de culpabilité (exigeant la preuve de faits commis).
Le risque est de voir sanctionner demain l'écart par rapport à une prétendue « norme » sociale ou économique définie par un code informatique, sans rapport avec la « norme » juridique. La question de la place du droit pénal, des garanties procédurales, et partant de l'État de droit, devra être posée avec insistance.

L'AUTEUR

Jean-Pierre MIGNARD est avocat au Barreau de Paris. Il est maître de Conférences à l'Institut d'Études Politiques de Paris. Il intervient également dans des dossiers de droit de la communication et de la presse et de droit de la propriété littéraire et artistique. Il enseigne le droit pénal des affaires à l'École de droit de Science Po Paris. Il enseigne le droit des médias et de la presse à l'École de Journalisme de Sciences Po Paris et dirige un exécutive master. Il est membre du Comité Consultatif National d'Éthique. Officier de la Légion d'Honneur. Fondateur du cabinet Lysias en 2002 (devenu Lysias Partners en 2006), il en est aujourd'hui Président du Comité Exécutif. Jean-Pierre Mignard est enfin co-directeur de la rédaction du journal Témoignage Chrétien.

Il a réalisé de nombreuses publications parmi lesquelles : Le droit interne des partis politiques (Sous la direction de Julie Benetti, Anne Levade et Dominique Rousseau), Mare & Martin, janvier 2017 - Robert F. Kennedy, la foi démocratique (coll. H. Roussel), Stock, janvier 2016 - Gardien de nos frères, Stock, novembre 2014 - Sûreté nucléaire, droit et gouvernance mondiale (coll. S. Mabile et M. Mabile), éditions Bruylant, Bruxelles, avril 2012 - Manifeste pour la Justice (ouvrage collectif du Club Droits, Justice et Sécurités), éditions du Cherche-Midi, 2012 - La Question Prioritaire de Constitutionnalité : pratique et contentieux (coll. X. Magnon, V. Bernaud, K. Foucher, T. S. Renoux, éditions Litec Lexis Nexis, 2011- Justice pour tous, avec A.Vogelweith, Club témoin et Syndicat de la Magistrature, La découverte, Paris, 2001. Etc.

L'AUTEUR

Adrien Basdevant est avocat au Barreau de Paris. Il se dédie au contentieux des affaires. Il bénéficie d'une double formation de l'ESSEC Business School et de l'Université Panthéon-Assas. Passionné par l'impact des technologies sur la société, il a contribué au développement d'une start-up à New York avant de se former au sein de cabinets français et anglo-saxon. Il est également membre du comité scientifique de la Chaire « Humanisme Numérique » du Collège des Bernardins. Il a réalisé des publications liées à l'émergence des technologies numériques dans les sociétés modernes.

L'intelligence artificielle, un nouvel horizon, 2017 – Cahiers Lysias - Pour un Parquet National du Numérique et une 33e chambre correctionnelle de la cybercriminalité, Revue Lamy Droit Immatériel, juillet 2017 - Le Droit des drones, 2016 – Cahier Lysias - « Traçage des données et Profilage comportemental », Université Panthéon-Assas, 1 er juin 2013 - « Innovation in digital platforms », Yale Club of New York, 14 février 2012 - « To trust or not to be : Comment éviter la tragédie de la transparence totale ? », OuiShare Magazine, 20 novembre 2012,…

Adrien Basdevant et Jean-Pierre Mignard sont co-auteurs de l'ouvrage : L'empire des données (Essai sur la société, les algorithmes et la loi) - Éditeur : Don Quichotte - ISBN : 978-2-35949-635-2 - 284 pages - Parution : 03/2018

Enjeux juridiques
et éthiques des algorithmes

Questions à Georgie Courtois et Nina Gosse

NDLR : A l'heure de la numérisation du monde, le débat public est notamment orienté sur le sujet des algorithmes, à qui l'on demande de plus en plus de « rendre des comptes ». Georgie Courtois et Nina Gosse, de la société d'avocats de Gaulle Fleurance & Associés, expliquent, au cours de cet entretien, en quoi le développement et l'utilisation des algorithmes soulèvent aujourd'hui tant d'interrogations et vont jusqu'à questionner notre droit positif.

GEORGIE COURTOIS

Avocat Associé, De Gaulle Fleurance & Associés Pilote du groupe juridique Hub France IA.

NINA GOSSE

Avocat, De Gaulle Fleurance & Associés. Membre du groupe juridique Hub France IA

Quels sont les enjeux juridiques liés aux algorithmes en termes de responsabilité ?

Georgie Courtois : Les questions relatives à la responsabilité des algorithmes et de l'intelligence artificielle sont multiples. Il convient de distinguer, d'une part, la responsabilité des algorithmes d'intelligence artificielle eux-mêmes, désincarnés, qui n'ont aucune interaction avec le monde physique (à titre d'exemple, les algorithmes utilisés dans les logiciels de police prédictive ou encore de diagnostics médicaux) et, d'autre part, la responsabilité des algorithmes intégrés dans des objets qui ont une interaction directe avec le monde physique (les robots, les voitures autonomes ou encore les drones autonomes).

Les enjeux et régimes de responsabilités applicables seront nécessairement différents.

Dans la première hypothèse, l'analyse de la responsabilité des algorithmes d'intelli-

gence artificielle eux-mêmes passera par le prisme des conséquences décisionnelles auxquelles aura conduit l'analyse de l'algorithme. En effet, dans cette hypothèse, l'algorithme apportera une aide décisionnelle à une personne utilisant le logiciel d'intelligence artificielle. À titre d'exemple, si un médecin utilise un logiciel d'intelligence artificielle de diagnostic médical, et que ce logiciel commet une erreur de prescription ayant des conséquences graves pour le patient, la question de la responsabilité du logiciel pourra émerger. Vis-à-vis du patient, le médecin restera responsable directement dans la mesure où il sera décisionnaire sur l'application, ou non, de la prescription recommandée par le logiciel. Rien n'empêche d'envisager, néanmoins, une action du médecin et du patient à l'encontre de l'éditeur du logiciel en raison de l'erreur grave commise dans le diagnostic. Les régimes généraux de responsabilité (contractuel pour le médecin vis-à-vis de l'éditeur, et délictuel pour le patient vis-à-vis de l'éditeur) seront mis en œuvre. Il sera alors nécessaire de démontrer que l'erreur provient du logiciel. Cette preuve ne sera pas toujours aisée à démontrer en raison de la transparence relative des algorithmes (certains algorithmes étant de véritables boites noires dont personne ne peut tracer le processus décisionnel, comme c'est le cas pour les algorithmes de deep Learning). Par ailleurs, l'algorithme d'intelligence artificielle ne faisant que traiter les données qui lui ont été fournies, l'erreur de configuration ou de la fourniture de données biaisées, par le médecin, pourra être de nature à écarter la responsabilité de l'éditeur.
En tout état de cause, les éditeurs de logiciels essayeront de limiter au maximum leur responsabilité par l'utilisation de clauses de limitation de responsabilité contractuelles.

(1) Art. 1242 Code Civil.

(2) Art. 1245 et suivants du Code Civil.

(3) V. : A.Bensamoun (dir.), Les Robots, contribution G.Courtois sur « Robot et Responsabilité », Édition Mare & Martin ; G.Courtois, Robots Intelligents et Responsabilité : Quels Régimes, Quelles Perspectives ? Dalloz IP/IT 2016, p. 287.

S'agissant de la deuxième hypothèse, c'est-à-dire les algorithmes d'intelligence artificielle incorporés à des objets, les régimes généraux de responsabilité du fait des choses que l'on a sous sa garde[1], et de responsabilité du fait des produits défectueux[2] lorsque le dommage vient d'une défaillance du robot externe à son utilisateur, sont actuellement applicables[3].

En revanche, des règles spéciales doivent être adoptées en ce qui concerne certains objets intelligents, comme les véhicules autonomes. La parution récente du *Décret n° 2018-211 du 28 mars 2018 relatif à l'expérimentation de véhicules à délégation de conduite sur les voies publiques* permet d'envisager, à court terme, une adaptation des règlementations, afin d'intégrer ce nouvel objet de droit sur les réseaux routiers français. La démarche qui vise l'expérimentation concrète afin d'affiner le cadre règlementaire doit être poursuivie. L'expérimentation doit précéder la régulation, afin que cette dernière ne devienne pas un frein à l'innovation.

Le droit français est suffisamment souple et adaptable pour trouver dans les textes actuels les solutions aux problématiques futures, sans nécessairement avoir besoin de recourir à de nouveaux concepts.
Il est donc faux, voire dangereux, d'agiter le spectre du « vide juridique » qui n'existe pas. Ces questions sont passionnantes et tous les praticiens attendent avec impatience les premières jurisprudences en la matière.

Pourquoi la question de l'éthique des algorithmes est-elle aujourd'hui au centre des débats, notamment ceux de la CNIL ?

Nina Gosse : Si l'existence des algorithmes, au sens de langages de programmation, remonte à plusieurs décennies, les évolutions récentes connues par nos sociétés confèrent toute sa pertinence à cette question. Les algorithmes sont en effet devenus omniprésents dans nos vies numériques : que ce soit pour gérer les achats, les relations sociales, l'information, ou encore pour optimiser la sécurité, les illustrations sont infinies.
Or, toute technologie n'est ni bonne ni mauvaise mais elle est n'est pas neutre, comme le dit Lawrence Lessig dans son célèbre article, « code is law » [4].

(4) Code is Law – On Liberty in Cyberspace, Lawrence Lessig – janvier 2000 – Harvard Magazine.

Il ne fait, en effet, plus aucun doute que les algorithmes ont un impact sur nos comportements mais aussi, plus largement, sur nos sociétés, on a pu le voir par exemple avec les dernières élections présidentielles. En réalité, à l'ère où les données s'accumulent à un rythme croissant, nous avons besoin des algorithmes.

Les algorithmes gouvernent nos achats, nos relations sociales, notre manière de nous distraire et de nous informer. Cela soulève de nombreuses questions éthiques, telles que celles liées aux biais, discriminations et erreurs, au profilage et à l'absence de libre arbitre ou autonomie humaine, mais encore l'enfermement des individus dans des « îlots comportementaux »,… Ces questions ne sont pas limitées à un domaine ou secteur particulier, d'où l'existence d'un relatif consensus sur la nécessité d'une réflexion générale.
Notre sentiment, en tant qu'individu, est souvent contrasté, entre intérêt et malaise, ce dernier étant bien souvent dû à une incompréhension autour de cet objet technique complexe. L'opacité des algorithmes fait qu'on les qualifie d'ailleurs fréquemment de « boîtes noires ».

(5) Comment permettre à l'Homme de garder la main ? Rapport sur les enjeux éthiques des algorithmes et de l'intelligence artificielle, CNIL, 15 déc. 2017.

Dans ce contexte, le débat public qu'a voulu mener la CNIL est celui de la maîtrise de l'Homme dans ce nouvel univers. Dans son rapport publié en décembre dernier, l'autorité de protection des données se prononce pour l'élaboration d'un modèle français de gouvernance éthique[5].

Loyauté et vigilance sont au cœur de ses recommandations afin de penser les systèmes dans l'intérêt de l'individu.

Cela renvoie plus largement à la question de l'éthique du numérique, qui n'est pas nouvelle, mais sans cesse nourrie par l'innovation technologique.

Aujourd'hui, Big Data, Intelligence artificielle et algorithmes prédictifs catalysent l'attention. Cela vaut tant pour les activités du secteur privé que celles de l'État. Il suffit pour s'en convaincre de se rappeler les critiques suscitées par la loi renseignement.

(6) Donner un sens à l'intelligence artificielle : pour une stratégie nationale et européenne, sous la présidence de Cédric Villani, mars 2018.

(7) Dominique Cardon, À quoi rêvent les algorithmes. Nos vies à l'heure des big data, Paris, Seuil, La République des idées, 2015.

On retrouve cette volonté d'encadrement dans le rapport de Cédric Villani sur l'intelligence artificielle du 28 mars 2018[6] qui propose la création d'un comité consultatif national d'éthique pour les technologies numériques et l'intelligence artificielle. Cet organe serait chargé d'organiser le débat public, mais aussi de conseiller l'État sur ses choix technologiques.

Cette nouvelle culture qu'on semble chercher à développer, au moins en Europe et particulièrement en France, me fait penser aux mots de Dominique Cardon qui, dans son ouvrage « À quoi rêvent les algorithmes » publié en 2015, avançait cette idée qu'il faut « *guider les rêves des ordinateurs afin qu'ils ne deviennent pas nos cauchemars de demain* » [7]. Je crois sincèrement qu'il ne faut pas diaboliser ces nouveaux outils en y voyant le « Frankenstein » des temps modernes, mais rester lucides et parvenir à des recommandations opérationnelles.

En quoi le droit peut-il apporter des réponses ? Faut-il réguler les algorithmes pour assurer leur loyauté et leur transparence ?

Georgie Courtois : Un des enjeux majeurs de l'intelligence artificielle est la transparence algorithmique.

Le but est de comprendre le processus de décision de l'algorithme. C'est le seul moyen d'obtenir des preuves dans l'hypothèse où la responsabilité algorithmique doit être mise en œuvre. C'est aussi le seul moyen de comprendre quel choix éthique a été fait lors de la conception de l'algorithme.

Ici encore, l'éthique est omniprésente et il est nécessaire pour une meilleure acceptabilité sociale que l'utilisateur d'intelligence artificielle comprenne les fondements du processus de calcul. L'exemple le plus marquant est évidemment celui du « dilemme du Tramway », mis au goût du jour avec la voiture autonome. L'algorithme va-t-il décider de limiter le nombre de pertes humaines ? Va-t-il faire prévaloir un certain type d'individu, par exemple en sauvant un enfant plutôt qu'un adulte ? Ou, va-t-il décider de sacrifier le conducteur du véhicule autonome ?

Ces questions éthiques doivent faire l'objet d'un consensus, pour éviter des dérives, tendant à voir des fabricants de voiture autonome faire payer une option coûteuse garantissant de sauver la vie du conducteur. La régulation aura ici un rôle permettant la transparence et l'uniformisation pour une meilleure acceptabilité sociale.

(8) http://europa.eu/rapid/press-release_IP-18-3362_fr.htm.

À ce titre, la Commission Européenne vient de présenter le 25 avril 2018, une approche européenne sur l'IA visant à stimuler l'investissement et à fixer des lignes directrices en matière d'éthique[8]. Dans ce plan d'action, la Commission évoque les questions de transparence algorithmique, de responsabilité et de sécurité en parlant notamment de la réforme à venir de la directive relative aux produits défectueux au regard des évolutions technologiques, notamment robotique. La Commission Européenne fait également état d'un projet de lignes directrices en matière d'éthique pour l'IA qui sera élaboré sur la base de la charte des droits fondamentaux de l'UE, après une consultation d'experts au sein d'une Alliance pour l'IA. Cela va dans le bon sens pour avoir une règlementation harmonisée et pragmatique.

Nina Gosse : On constate que certaines évolutions récentes vont dans le sens d'une régulation par le droit.

Le Règlement Général relatif à la Protection des Données (RGPD) [9] confirme l'attache ment européen pour la « privacy » et, plus qu'un bouleversement du régime juridique, exprime la volonté de diffuser ces valeurs sur la scène internationale. Plusieurs dispositions trouvent à s'appliquer en matière d'algorithmes, à commencer par l'article 22 du Règlement qui prévoit que toute personne dont les données personnelles sont traitées a le droit, sauf exceptions, de « ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire. »

(9) Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données).

(10) Directive (UE) 2016/680 du Parlement européen et du Conseil du 27 avril 2016 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécution de sanctions pénales, et à la libre circulation de ces données, et abrogeant la décision-cadre 2008/977/JAI du Conseil.

On retrouve ce principe à l'article 11 de la Directive 2016/680 relative aux traitements effectués par les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécution de sanctions pénales[10]. Cet article prévoit, en outre, que tout profilage qui entraîne une discrimination à l'égard des personnes physiques sur la base de données sensibles est interdit.

Ces dispositions sont néanmoins limitées aux algorithmes qui utilisent des données personnelles et concernent seulement leurs effets sur des personnes physiques.
Par ailleurs, on peut douter de leur efficacité en pratique.

(11) Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique.

En réalité, d'autres dispositions peuvent également trouver à s'appliquer. On peut ainsi également citer la Loi Lemaire[11] qui a introduit au sein du Code de la consommation des exigences de loyauté des plateformes et d'information des consommateurs notamment quant au classement ou référencement, au moyen d'algorithmes informatiques, de contenus, de biens ou de services (art. L. 111-7).

L'AUTEURE

Nina Gosse intervient notamment en propriété littéraire et artistique, en propriété industrielle et technologies de l'information et de la communication (internet, audiovisuel, médias, télécoms, etc.). Elle accompagne également ses clients sur des problématiques liées aux données personnelles, à la santé, ainsi qu'en matière d'affaires publiques et réglementaires. Elle dispose par ailleurs d'une expérience en tant que consultante Juridique et Affaires Publiques dans l'audiovisuel et le numérique (août 2014-décembre 2016).

Nina Gosse est titulaire d'un Magistère en droit de la Propriété Intellectuelle et des Nouvelles Technologies (Poitiers, 2014), d'un Master II Recherche en Propriété Intellectuelle (Université de Cambridge, 2013), et est diplômée de l'Université Paris-Sorbonne (ENS Cachan D1 économie, droit et gestion).

L'AUTEUR

Georgie Courtois intervient en conseil et en contentieux en droit des nouvelles technologies (informatique, internet, cryptologie, données personnelles…) et en droit de la propriété intellectuelle (droit d'auteur, droits voisins, marques, dessins et modèles, brevets). Il a développé une connaissance particulière des questions relatives aux innovations technologiques (intelligence artificielle, objets connectés, Big Data, impression 3D, robotique, drones,) ainsi qu'une solide expérience des questions liées à la lutte anti-piratage informatique et à la lutte contre la cybercriminalité. Il enseigne la pratique des procédures judiciaires et du contentieux liés à la Propriété Intellectuelle dans le Master 2 PIDN (Propriété Intellectuelle Droit du Numérique) à la Faculté Jean Monnet, Université de Paris Saclay et intervient en droit des technologies innovantes dans le Mastère Spécialisé DAIM (Droit des Affaires Internationales et Management) de l'ESSEC Business School.

Il est membre fondateur du #HubFranceIA, association française sur l'intelligence artificielle, dont il pilote le groupe sur l'accompagnement juridique, et membre du groupe « Technologie » de l'International Bar Association (IBA).20 novembre 2012,…

Surmonter les risques
de la gouvernance par les algorithmes

Par Enguerand Marique

Aujourd'hui, le recours à la gouvernance algorithmique connaît un essor considérable. Elle peut être comparée à la gouvernance juridique « traditionnelle », pour mieux en comprendre les risques. Cet examen identifie alors les mécanismes mis en place par le droit pour surmonter des problèmes analogues et suggère des solutions permettant de préserver la diversité et la pluralité au sein de nos sociétés démocratiques modernes.

ENGUERAND MARIQUE

Doctorant - chercheur en droit Université Catholique de Louvain

Les programmeurs disposent aujourd'hui d'un pouvoir discrétionnaire dans la rédaction du code de leurs programmes, des algorithmes et dans la manière dont ils envisagent leurs relations avec leurs utilisateurs. Ils disposent d'un pouvoir équivalent à celui d'un législateur, c'est-à-dire dicter des comportements au sein de la communauté[1]. L'architecture informatique assure ensuite l'exécution correcte des opérations.

(1) L. Lessig, « Code Is Law. On Liberty in Cyberspace », *Harvard Magazine*, janv. 2000, disponible sur http://harvardmagazine.com/2000/01/code-is-law-html, consulté le 5 mai 2018.

Cette approche pose problème si elle est appliquée au droit.
Si une règle est traduite en code (informatique), elle devient un processus dont l'issue est certaine et connue, « toute chose étant égale par ailleurs ».

Dans cette perspective, le droit risque de devenir déterministe et inflexible.
Ce serait toutefois oublier que le postulat « toute chose étant égale par ailleurs », que l'on retrouve dans les lois « informatiques », ne vaut pas

en droit : toute personne ou situation est différente.
Il faut donc une marge d'interprétation pour appliquer les lois « juridiques » et proposer des pistes de réflexion pour éviter une gouvernance par le code ou les algorithmes.

Le droit n'est pas une science quantitative

Tout d'abord, grâce aux algorithmes et aux données auxquelles ils s'appliquent, les opérations effectuées deviennent quantifiables. Toutefois, les comportements qui ne sont pas chiffrables ne pourront être pris en considération par les algorithmes[2], d'où les débats relatifs à l'éthique des algorithmes. Si une voiture autonome roule à vive allure et ne peut s'arrêter face à un obstacle, elle peut entrer en collision avec ce dernier ou l'éviter, quitte à rencontrer un autre obstacle. À supposer que ces obstacles soient multiples : une personne âgée, un enfant, une femme enceinte, etc. et les risques de dommage physique identiques, quelle valeur peut-on attribuer à ces différentes options de conduite autonome ?

(2) Virginia Eubanks, *The Digital Dead End : Fighting for Justice in the Information Age*, MIT Press, 2011.

(3) Commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique d'Allistene (Alliance des sciences et des technologies du numérique) http://cerna-ethics-allistene.org/.

Le courant d'analyse économique du droit (« Law & economics ») tente un exercice de quantification et de valorisation mais ne fait actuellement pas l'objet d'un consensus.

Certains proposent de créer des comités d'éthique sectoriels ou nationaux. La CERNA[3] est un exemple en France.

Le droit et la technologie créent des attentes de nature différente

La gouvernance par la loi dans un cadre démocratique permet de répondre en amont à des attentes de justice. En aval, la mise en œuvre des règles juridiques impose en principe la transparence.
La gouvernance algorithmique n'offre pas ces mêmes garanties. La forme de l'algorithme n'est pas débattue mais est sous-traitée à des entreprises de consultance informatique. Seuls les objectifs essentiels d'un traitement algorithmique par l'autorité publique (en cas de justice prédictive ou d'administration automatisée) sont fixés au bout d'un débat. Cette délégation de pouvoir pose en particulier problème lorsque l'algorithme n'est pas rendu public, ce qui pourrait arriver lorsqu'une entreprise craint que la publicité du code permette aux utilisateurs de tromper l'algorithme.

© Fotolia - 92732535 - Subscription

La légitimité des traitements algorithmiques repose comme les décisions juridiques sur une motivation liée à la publicité des règles codées dans les programmes.

Au contraire, l'algorithme applique la règle systématiquement sans fournir le raisonnement.

En pratique, si le Parquet dispose d'un pouvoir d'appréciation discrétionnaire pour classer certaines affaires sans suite, il doit en avertir les victimes, qui peuvent alors forcer les poursuites (art 85 du code de procédure pénale).

En France, le logiciel français *Admission Post-Bac* (APB) en constitue une excellente illustration. Ce logiciel était destiné à optimaliser l'allocation des élèves de terminale entre les filières universitaires contingentées, pour évi-

ter le biais humain qui existait dans le système précédent. Une série de critères étaient pris en considération par APB : premiers choix des candidats, critères sociaux et géographiques, etc. La transcription en code de ces critères, listés dans la loi, a suscité de vives réactions qui ont conduit à des demandes d'accès à l'algorithme auprès de la Commission d'accès aux documents administratifs et, ultérieurement, à une révision majeure du système - renommé pour l'occasion « Parcoursup ».

Le code, en particulier l'intelligence artificielle, engendre une dilution des responsabilités

En 1996, Nissenbaum pointait déjà la dilution des responsabilités en raison de la multitude de personnes impliquées dans la création et la mise en œuvre des programmes, les potentiels 'bugs' informatiques, l'apparente aisance à accuser la technique des défauts souvent attribuables à des personnes et l'absence d'obligation du programmeur de rendre des comptes sur le fonctionnement de son algorithme[4].

(4) Nissenbaum H., "Accountability in a computerized society", *Science and Engineering Ethics*, 1996, 2(2), p. 25-42.

La solution qui était plaidée à l'époque résidait dans davantage de transparence. Les attentes créées par les mécanismes d'intelligence artificielle complexifient cette approche. Les raisons qui justifient la prise d'une décision sont déterminées par l'entraînement de l'algorithme. Ce dernier configure lui-même les règles au regard de situations déjà résolues, éventuellement de façon évolutive. Les informaticiens éprouvent alors des difficultés pour répondre aux demandes de transparence et de systématicité.

Réduire le risque de « faux-positif » ou « faux-négatif »

L'algorithme ne constitue qu'une méthode pour aboutir à un résultat, et ce dernier peut être obtenu avec des traitements algorithmiques différents. La combinaison de plusieurs traitements algorithmiques en parallèle permettra de comparer les résultats. Le cas échéant, des produits de traitements algorithmiques divergents issus d'une même situation seront présumés « faux-positifs »/ « faux-négatifs ». Ils pourront être identifiés et écartés. Dans la gouvernance juridique, ces situations sont résolues par des mécanismes de recours ou d'appel. Un juge d'appel ne considère pas nécessairement que la solution donnée par le juge de première instance est basée sur des faits pertinents et une analyse correcte. Après avoir pris en compte d'autres éléments pertinents, le juge d'appel peut aboutir à la même décision que le premier juge. En l'absence de systématicité et de transparence de l'algorithme, ou d'une multiplicité de traitements algorithmiques pour un même ensemble factuel, un droit d'appel de la décision algorithmique auprès d'un juge doit être garanti. Récemment, un juge d'appel britannique a par exemple mis en avant qu'un logiciel de reconnaissance vocale fonctionnant au moyen d'intelligence artificielle avait un taux d'erreur de 20 % et qu'il en résultait un retrait abusif du visa de 7 000 étudiants.

De nouvelles responsabilités pour réduire les risques ?

Qui est responsable, et dans quelle mesure, si un véhicule autonome choisit *délibérément* d'ignorer la présence d'un piéton, comme dans le récent accident mortel causé par une voiture connectée Uber en Arizona ? Certains proposent de créer une responsabilité objective sur base du *résultat* de la décision qui devrait être conforme aux politiques publiques décidées par le législateur[5] (par exemple : la non-discrimination) ou de permettre à des autorités externes d'accéder confidentiellement aux algorithmes protégés par des secrets d'affaires pour vérifier leur intégrité et leur bien-fondé[6].

(5) Kroll, J. A., Huey, J., Barocas, S., et al., « Accountable algorithms », *University of Pennsylvania Law Review*, 2017, 165(3), p. 633-705.

(6) de Laat P., Algorithmic Decision-Making Based on Machine Learning from Big Data : Can Transparency Restore Accountability ? *Philos. Technol.*, 2017, p. 1-17.

Par ailleurs, la cohabitation des gouvernances juridique et algorithmique nécessite une réflexion particulière pour éviter la création d'un double standard discriminatoire. Ainsi, il ne faudrait pas qu'un piéton victime d'un accident ou de dommages causés par un véhicule autonome soit dans une position juridique plus difficile que s'il avait été confronté à un conducteur automobile traditionnel. Le piéton en question n'a en effet pas choisi entre les deux interlocuteurs et est donc dans une situation comparable.

Conclusion

Il a été brossé dans cet article un aperçu des risques de la gouvernance par l'algorithme-exécutif et par les programmeurs-législateurs. Pour surmonter ces obstacles, le besoin s'exprime de rendre transparent le processus algorithmique, de débattre de la mise en place du code, de multiplier les algorithmes sur un même sujet pour débusquer les résultats « faux-posi-

tifs » ou « faux-négatifs ». Cette approche permettra alors de mettre sur pied des faisceaux de présomptions ou de garantir la diversité, car la dimension humaine des règles ne doit pas être omise. Si d'autres recherches préfèrent une réglementation *ex ante* qui viserait à réguler les algorithmes, prévenir les risques des algorithmes encore non commercialisés en empêchant leur activation sur le marché[7], on peut suggérer un régime de responsabilités *ex post* pour les utilisateurs d'algorithmes qui causeraient un dommage à des tiers ou des cocontractants.

(7) Tutt A., "An FDA for algorithms", *Administrative Law Review*, 2017, 69(1), p. 83–123.

Prédiction

et personnalisation algorithmique : des nouveaux outils pour la police

Par Francesca Musiani

Les acteurs du numérique, tout particulièrement les grandes plateformes, recourent de plus en plus à des algorithmes prédictifs, ayant notamment deux objectifs : personnaliser le service rendu à leurs clients ou fournir des éléments utiles à une prise de décision. De par leur nature technique, on a tendance à attribuer à ces dispositifs des qualités d'infaillibilité et d'objectivité alors que, s'agissant de procédures codées pour transformer des données « entrantes » en données « sortantes » sur la base de calculs spécifiques, ils portent souvent inscrites les conséquences des choix de leurs développeurs[1]. Ils peuvent donc tendre à déposséder les individus d'une alternative décisionnelle et à réduire leur libre arbitre, en n'étant plus simplement des mécanismes d'aide à la décision mais de véritables systèmes de décision automatique ou semiautomatique[2].

FRANCESCA MUSIANI

Chargée de recherche au CNRS ISCC (CNRS/Sorbonne Université)

(1) José Van Dijck, "Datafication, dataism and dataveillance : Big Data between scientific paradigm and ideology", *Surveillance & Society*, 12 (2), 2014.

(2) On a abordé plus largement ce sujet dans Francesca Musiani, *Internet et vie privée*, Toulouse : UPPR Editions, 2016.

(3) Élisabeth Grosdhomme-Lulin, *Gouverner à l'ère du Big Data. Promesses et périls de l'action publique algorithmique*, Institut de l'entreprise, 2015.

Les algorithmes prédictifs

Ces algorithmes peuvent avoir une influence sur les droits des consommateurs ainsi que sur la relation de l'usager/citoyen avec les pouvoirs publics. En parallèle du déploiement à large échelle des algorithmes par le secteur privé émerge donc une véritable « action publique algorithmique » [3] destinée, par exemple – comme dans le film *Minority Report*[4] – à anticiper certains comportements en matière de sécurité (on y reviendra) et plus généralement dans le domaine social, tels que la prédiction des

risques de maltraitance, l'aide au diagnostic médical ou l'anticipation des risques de décrochage scolaire.

(4) *Minority Report* est un film américain de SF qui, dans le contexte d'un Washington de 2054, met en scène des êtres humains mutants, les précogs, qui peuvent prédire par leur prescience les crimes à venir.

Quelles nouvelles règles pour l'ère de la « gouvernementalité algorithmique » ?

Se pose donc la question d'intégrer les droits d'information, d'accès et d'opposition, déjà reconnus aujourd'hui, par un encadrement accru des algorithmes employés lors des traitements de données. En France, l'article 10 de la loi Informatique et Libertés interdit déjà qu'une décision, donnant lieu à des effets juridiques à l'égard d'un individu, soit prise « sur le seul fondement d'un traitement automatisé de données destiné à définir le profil de l'intéressé ou à évaluer certains aspects de sa personnalité » [5].

(5) https://www.cnil.fr/fr/loi-78-17-du-6-janvier-1978-modifiee#Article 10

(6) Antoinette Rouvroy et Thomas Berns, « Le nouveau pouvoir statistique », *Multitudes* n° 40, pp. 88-103, 2010.

Toutefois, dans notre ère, qu'on a pu appeler avec pertinence de la « gouvernementalité algorithmique » [6], cette interdiction semble être, de facto, de moins en moins opérante au vu de l'importante asymétrie informationnelle et décisionnelle qui existe entre les concepteurs, les utilisateurs de l'action algorithmique et les personnes qui en font l'objet.

De nouvelles règles sont donc envisagées aux niveaux national et international pour encadrer le fonctionnement des algorithmes prédictifs, notamment : un droit d'opposition au profilage pouvant conduire à des mesures produisant des effets juridiques pour l'individu ou affectant de manière importante ses libertés, intérêts, droits ; une contrainte d'intervention humaine effective sur toute décision prise à l'aide d'un algorithme ; une obligation de transparence sur les types de données personnelles utilisées par l'algorithme et les modalités de son paramétrage, permettant, le cas échéant, des « audit » qui permettraient d'en contester la logique générale ou la véracité des données analysées ; l'interdiction d'utiliser des algorithmes conduisant directement ou indirectement à instaurer une discrimination fondée, entre autres, sur la race, l'origine ethnique, les opinions politiques, la religion, les convictions, l'orientation sexuelle ou l'identité de genre ou encore l'appartenance syndicale.

Les algorithmes prédictifs comme outil de renseignement et d'enquête

Au-delà de leurs implications pour la vie privée, la généralisation de la prédiction et de la personnalisation algorithmiques pose des questions d'envergure quant à leur impact sociétal à moyen terme sur une variété de questions qui incluent la manipulation – lorsque la personnalisation a pour but d'influencer les comportements – mais également le renforcement de certaines formes de discrimination généralisée, l'uniformisation de la société ou encore la réduction de la liberté de choix des individus. Plusieurs de ces questions se posent avec force en ce qui concerne l'utilisation des algorithmes prédictifs dans le domaine du renseignement et de l'enquête policière.

Loi sur le renseignement et « boîtes noires »

En ce qui concerne le premier aspect, un exemple récent particulièrement éclatant a été la révision de l'article L. 851-3 du Code de la sécurité intérieure sur la base de la controversée « loi sur le renseignement » de 2015. Ces nouvelles dispositions autorisent ce qu'on a appelé la détection de « signaux faibles » par la pose de « boîtes noires » chez les opérateurs ; c'est-à-dire que les opérateurs de communications électroniques et les fournisseurs de services peuvent désormais traiter les informations et documents circulants dans les réseaux qu'ils gèrent afin de détecter des liens entre des informations susceptibles de révéler une menace terroriste.

(7) INRIA, Éléments d'analyse technique du projet de loi relatif au renseignement, 30 avril 2015, disponible en ligne à http://www.lemonde.fr/pixels/article/2015/05/13.

Lors de l'annonce de sa première version, cette disposition avait suscité des réactions particulièrement enflammées quant à son opportunité juridique et sa faisabilité

© Gunnar Assmy - 62348117 - Fotolia

L'utilisation d'algorithmes appliqués à des bases de données individuelles, quand elle entraîne des décisions touchant aux libertés individuelles et d'ordre public, requiert un encadrement par le législateur, notamment du contexte politique de sa mise en œuvre.

technique[7]. Le texte définitivement adopté, après de multiples controverses, fait quelques progrès dans l'encadrement des moyens à la disposition des services de renseignement et l'utilisation qu'ils peuvent en faire : notamment, les principes et finalités de la politique publique de renseignement sont posés et la procédure d'autorisation des techniques de recueil est fixée. Cependant, plusieurs acteurs de la société civile, des activistes aux scientifiques, continuent à estimer que cette loi ouvre la voie à une collecte massive et à un traitement généralisé de données personnelles[8] et que la technologie des « boîtes noires » permettant cette collecte n'est pas, en l'état des informations disponibles, susceptible d'un encadrement strict. Le recours à des dispositifs algorithmiques de traitements de données transitant par les réseaux numériques aux fins de détection de « signaux faibles » ou de menaces, quelle que soit la finalité poursuivie, reste donc très problématique.

(8) Antonio Casilli, « Quatre thèses sur la surveillance numérique de masse et la négociation de la vie privée », Rapport du Conseil d'État, 2015, pp. 423-434.

Prédire le crime ? L'exemple de PREDPOL

Au début des années 2010, un certain nombre de villes américaines se sont dotées du logiciel PREDPOL (quasi-acronyme de *predictive policing*). Basé sur des techniques d'élaboration du big data, l'algorithme que sous-tend ce logiciel analyse les archives de la police d'une ville ou d'un territoire spécifique (procès-verbaux, suivis d'arrestations, appels au secours) pour en déduire les endroits où les crimes sont le plus fréquents et ensuite, en arriver à « prédire » les lieux qui sont à surveiller en priorité. La start-up qui développe et distribue ce logiciel a, lors de ses premières implémentations, proclamé son succès : les trois types de délits « couverts » par PREDPOL auraient connu une baisse importante : les cambriolages auraient baissé de 27 % alors que les vols de voiture et dans les lieux publics seraient minorés de 11 % à 15 %. Le service a un coût relativement modeste, environ 30.000 euros, qui a déjà convaincu plusieurs dizaines de villes de l'adopter.

Le scénario proposé par *Minority Report* s'est-il pour autant réalisé de façon efficace ? Plusieurs scientifiques, des informaticiens aux sociologues, émettent des doutes. Si les études techniques de PREDPOL ont été rares – notamment parce que la politique de l'ouverture des données utilisées et du code développé par la start-up laisse fort à désirer – quelques-unes ont quand même pu avoir lieu, en se basant sur quelques jeux de données mis en accès libres par des villes et sur des modèles semblables à PREDPOL. Le chercheur Ismaël Benslimane a ainsi conclu que les résultats de cet algorithme sont loin d'être exceptionnels car la prédiction que les crimes auront lieu majoritairement dans les zones prédites « n'est en fait pas plus performante que la prédiction : « les crimes auront lieu majoritairement dans les zones historiquement les plus criminogènes de la ville[9] ».

(9) Ismaël Benslimane, Étude critique d'un système d'analyse prédictive appliqué à la criminalité : *Predpol®. CorteX Journal*, 2014.

(10) Bilel Benbouzid, « Des crimes et des séismes. La police prédictive entre sciences, techniques et divination », *Réseaux*, n° 6 (206), pp. 95-123, 2017.

(11) Bilel Benbouzid, « De la prévention situationnelle au *prédictive policing* », *Champ pénal/Pénal field* [En ligne], Vol. XII | 2015, DOI : 10.4000/champpenal.9050.

Bilel Benbouzid, spécialiste des implications sociales de la police prédictive, fait également sien l'argument que l'un des principaux problèmes de ce logiciel est qu'il fait trop abstraction du contexte historique et social du crime. PREDPOL réunit un ensemble de variables spatio-temporelles indiquant la forte chance qu'un crime se répète bientôt dans un endroit où il a déjà eu lieu. D'ailleurs, le sociologue fait remarquer que le logiciel se fonde sur un modèle tiré de la sismologie, qui prédit beaucoup moins bien un tremblement de terre que ses répliques[10]. En se concentrant sur la simple répétition du crime, il en vient à conclure automatiquement que le lieu en question a besoin de plus de surveillance et de coercition, au lieu de mettre en place des mécanismes qui élargissent la protection à des zones et des populations particulièrement vulnérables.
Y a-t-il donc un biais politique, sous-tendant PREDPOL et ses descendants, qui renforcerait une vision de la société amplificatrice d'inégalités ?

Conclusions

Il est possible aujourd'hui d'exploiter de façon systématique les données personnelles rendues disponibles sur les réseaux et leur croisement avec d'autres informations, ainsi que de recourir de façon massive à des technologies algorithmiques qui permettraient de prédire la probabilité qu'un fait délictueux ou criminel se produise à un endroit et un moment spécifique. Ces tendances s'accompagnent, en France et ailleurs, d'un déplacement d'un paradigme de surveillance ciblée vers un paradigme de surveillance massive et indifférenciée, dans lequel on puiserait des traces indiquant des possibles comportements déviants[12] – ce qui est illustré de manière controversée par les « boîtes noires » au centre de la loi sur le renseignement. Dans le présent mais surtout dans le futur proche, il appartient au législateur de vérifier que ces innovations se développent dans une utilisation équilibrée et pas dans l'abus des informations et des instruments disponibles ; le risque d'intrusions dans la vie privée, qui seraient disproportionnées par rapport aux objectifs légitimes de prévention de la délinquance et de répression des infractions, mérite une ample réflexion de par son actualité.

(12) Jean-Marc Manach, De la surveillance de masse à la paranoïa généralisée, Bug Brother, janvier 2015, http://bugbrother.blog.lemonde.fr/2015/01/03/de-la-surveillance-de-masse-a-la-paranoia-generalisee/.

L'AUTEURE

Francesca Musiani est chargée de recherche au CNRS, Institut des sciences de la communication (ISCC-CNRS/Sorbonne Université) et chercheuse associée au Centre de sociologie de l'innovation (i3-MINES ParisTech-PSL). Ses travaux portent sur la gouvernance de l'Internet. Elle est l'auteure de Internet et vie privée (Uppr Editions, 2016) et de Nains sans géants. Architecture décentralisée et services Internet (Presses des Mines, 2013 [2015], Prix Informatique et Libertés 2013 de la CNIL). Elle est rédactrice académique pour l'Internet Policy Review et vice-présidente en charge de la recherche de l'Internet Society France.

La Justice prédictive :
Mythe ou réalité ?

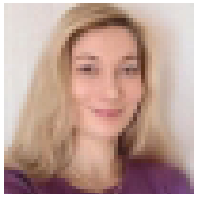
Par Me Cécile Doutriaux

L

(1) Dominique Rousseau Professeur de droit constitutionnel.

« La justice permet la conception d'une société civilisée et pacifiée. Sans justice, c'est la loi de la jungle et des marchés » [1]

Depuis plusieurs années, la justice doit affronter de nombreux enjeux : faire respecter les droits des citoyens, rendre des décisions plus rapidement, tout en étant moins onéreuse. Pour parvenir à ces objectifs, les réformes judiciaires se succèdent, sans toutefois satisfaire les justiciables ni susciter l'adhésion des principaux acteurs : magistrats et avocats.



Me CÉCILE DOUTRIAUX

Avocate
Membre de la Chaire Cyberdéfense des écoles de Saint Cyr - Sogeti - Thalès

Certains affirment aujourd'hui que " la justice est en situation de rupture ", sans que les journées " justice morte " de 2018 n'aient suscité un intérêt médiatique important.

Si l'on examine les reproches adressés à la justice par les justiciables, les deux points majeurs concernent la lenteur et le coût des procédures. Pourtant, depuis de nombreuses années, l'État a mis en place l'aide juridictionnelle (971.181 personnes en ont été bénéficiaires en 2016), les assurances de protection juridique permettent la prise en charge des frais d'un procès et les magistrats, soumis à une politique du chiffre, rendent un nombre important de décisions tous les ans (3.877.819 en 2016 selon le ministère de la justice).

Peu importent en réalité les problèmes auxquels sont confrontés quotidiennement les acteurs du monde judiciaire, l'essentiel pour le pouvoir politique est d'endiguer la générale et profonde insatisfaction des citoyens qui ont désormais de nouvelles attentes vis-à-vis de la justice : ne plus attendre, donner son avis, participer pleinement aux mesures qui les concernent et payer le moindre coût.

(2) Un sondage BVA, publié dans Le Parisien le 16 février 2014, révèle que les trois quarts des Français (75 %) estiment que la justice en France fonctionne mal. Parmi les critiques à son égard : « lente, complaisante, inégalitaire, élitiste, incompréhensible, déshumanisée ».

(3) Selon A. Mendoza-Caminade, « Le droit confronté à l'intelligence artificielle des robots : vers l'émergence de nouveaux concepts juridiques ? », D. 2016, p. 445 s.

(4) La *legaltech* (en français, technologie juridique ou technologie au service du droit), notion issue de l'anglais : *Legal Technology*, fait référence à l'usage de la technologie et de logiciels pour offrir des services juridiques.

En effet, la révolution digitale a permis l'émergence de communautés virtuelles de partage et d'entraide. Constituées autour d'intérêts communs, elles ne reposent pas sur un pouvoir central hiérarchique mais sur une construction commune de règles, ce qui remet en question l'organisation du pouvoir judiciaire, jugé « inégalitaire et élitiste » [2]. La « loi des algorithmes » pourrait incarner ce nouvel état d'esprit et serait « significative du glissement du gouvernement politique, délibéré et vertical, vers la gouvernance mathématique automatique et horizontale ».[3]

La justice, qui n'échappe bien évidemment pas à ces importants changements technologiques et sociétaux, doit évoluer. Ainsi, le projet de loi de programmation et de réforme pour la Justice 2018-2022 a, pour la première fois, envisagé de recourir à des plateformes d'algorithmes pour résoudre certains litiges.

Cette proposition a été incitée par les *legaltech*[4], lesquelles, par un discours économique insistant et offensif, promettent de fournir des outils qui permettraient aux professions juridiques, à partir du traitement d'un volume important de jugements, d'anticiper l'issue d'un procès et d'affiner la stratégie judiciaire. Cette solution alimente l'idée selon laquelle, grâce à l'intelligence artificielle, la décision serait rendue « en toute impartialité », par l'analyse objective d'un volume important de données, ce qui permettrait de réduire l'intervention humaine, subjective et suspectée d'être partiale. Elle permettrait de garantir aux justiciables la confiance, la rapidité et la gratuité des procédures judiciaires et bien évidemment, en cas de succès, cette solution serait étendue à d'autres contentieux.

L'idée a de quoi séduire ! D'autant plus qu'à l'issue de chaque procès, une partie gagne et l'autre perd, ce qui a pour effet de rendre la moitié des justiciables insatisfaits ! Alors, la justice prédictive est-elle un mythe ou une future réalité ?

I. IA et enjeux judiciaires

L'intelligence artificielle consiste à mettre en œuvre un certain nombre de techniques pour permettre aux machines d'imiter une forme d'intelligence humaine. Elle vise à reproduire, à l'aide de machines, des activités mentales, qu'elles soient de l'ordre de la compréhension, de la perception ou de la décision.

Dans le domaine judiciaire, l'intelligence artificielle serait ici utilisée par le biais d'algorithmes de recommandation, employés pour établir des modèles prédictifs à partir d'une quantité importante de jugements, collectés grâce au Big Data et ouverts à tous gratuitement dans le cadre de l'Open Data, prévu aux articles 20 et 21 de la loi n° 2016-1321 du 7 octobre 2016 pour une République numérique.

1. Au niveau technique et juridique

Avant de mettre en œuvre la justice prédictive, certaines précautions légales doivent toutefois être prises.

L'open Data implique la nécessité d'anonymiser les décisions de justice pour respecter la vie privée des justiciables et éviter une condamnation « à perpétuité », puisque les jugements seraient accessibles à tous, sans limitation de durée, par le biais de sites Internet. La majorité des magistrats et des greffiers souhaitent également bénéficier de cette anonymisation, selon le rapport relatif à « l'Open Data des décisions de justice » du 29 novembre 2017, pour éviter tout profilage et toute pression de l'opinion publique sur les réseaux sociaux qui pourrait être exercée à la suite des décisions rendues.

L'anonymisation des décisions de justice, réalisée par la Direction de l'information légale et administrative (DILA), présente un coût important, ce qui explique que le site officiel « https://www.legifrance.gouv.fr » comporte un nombre de décisions réduit (13.360) par rapport au nombre de jugements rendus chaque année (2 630 085 décisions en matière civile et commerciale, 1 200 575 en matière pénale et 231 909 décisions administratives en 2016). L'open Data impliquerait la nécessité d'anonymiser une quantité très importante de données et cette mission serait donc confiée à la Cour de Cassation et au Conseil d'État.

Les legaltech refusent de procéder à cette anonymisation en raison de son coût et affirment qu'il serait possible, grâce à leurs outils numériques, d'établir une carte de France pour mettre en relief les différentes disparités et convergences entre les jugements rendus d'une région à l'autre et d'un magistrat à l'autre.

Pourtant, dans la mesure où les décisions de justice collectées doivent être anonymisées, l'idée selon laquelle il serait possible de déterminer « les orientations » des magistrats quand ils rendent des décisions et la manière dont est traité un même conflit d'une région à l'autre, serait erronée. Cette solution était d'ailleurs illusoire puisque la carrière d'un juge évolue dans le temps, par des changements d'affectation vers différents contentieux, tribunaux et territoires et qu'il est difficile de déterminer son influence individuelle exacte, en cas de jugement rendu en collégialité par plusieurs magistrats.

Aucune décision, pour un même type de litige, n'est identique. Cette situation

aléatoire est souvent incompréhensible et angoissante pour le justiciable qui ferait bien l'économie de l'imprévisibilité judiciaire ! Pourtant, loin d'être l'expression de l'arbitraire et l'illustration de sa partialité, le pouvoir souverain du juge permet de prendre en compte chaque dossier individuellement, chaque personne unique dans son histoire, son éducation, son environnement, son parcours, ses valeurs, ses capacités d'évolution ou de réinsertion.

Les attentes des parties peuvent également se transformer en cours de procédure, soit dans l'apaisement, soit dans la radicalité et il est rare de pouvoir prévoir, par anticipation, quelle sera l'évolution de chacun des justiciables pris individuellement, tant la complexité et l'ambivalence humaine s'expriment devant les tribunaux.

Dans ces conditions, comment paramétrer les algorithmes avec suffisamment de

© Legal computer judge - The lightwriter.

Enfermement logarithmique, pression du réseau social sur le juge, prise en compte d'informations partielles et responsabilité non définie laissent à penser que l'émergence de process algorithmiques dans le traitement judiciaire obèrerait les fondements de l'autorité du juge.

finesse pour prendre en compte la complexité de chaque affaire ?

Une décision prise par un algorithme ou avec l'assistance d'un algorithme présente l'avantage de sembler a priori plus équitable et plus juste qu'une décision humaine. Les legaltech prétendent ainsi que la justice prédictive serait forcément neutre et objective, ce qui permettrait de garantir davantage les droits des citoyens et éliminerait l'arbitraire.

Or, l'intervention humaine serait toujours présente en cas de recours à des plateformes algorithmiques, puisque le paramétrage de l'algorithme, le choix et la pondération des critères et des catégories de données prises en compte pour arriver au résultat recherché, relèvent de la main de l'homme. Les algorithmes seraient en réalité « des opinions exprimées dans du code ». Par conséquent, une justice rendue au moyen de l'intelligence artificielle n'éliminerait pas toute subjectivité ou discrimination.

Enfin, tout l'historique, les faits du litige et tous les arguments longuement développés dans les conclusions des avocats, ne sont pas repris en intégralité dans les jugements. Les décisions de justice résument donc partiellement un litige sans l'exposer intégralement et les outils d'analyse développés actuellement par les legaltech traitent uniquement le dispositif du jugement, c'est-à-dire la solution du litige sans exposer les faits, le raisonnement, ni les arguments ayant conduit à la décision.

Les décisions de justice préconisées, via des plateformes algorithmiques, reposeraient donc sur des données incomplètes et approximatives, donc inexactes, et produiraient des résultats erronés pour trancher un litige.

On peut donc légitimement s'interroger sur la pertinence de développer l'intelligence artificielle dans le domaine judiciaire, sur la base de statistiques partielles.

2. IA et principes directeurs du procès

Avant d'envisager l'application de l'IA à la justice, il est nécessaire de savoir aussi si la mise en place de cette technologie respecterait les principes fondamentaux qui gouvernent les procès.

En effet, tout justiciable doit bénéficier d'un procès équitable. Dans le cadre d'un procès civil et pénal, le juge ne tranche un litige qu'après une libre discussion des prétentions et arguments de chacun des adversaires. Chaque partie au procès a la possibilité de faire valoir son point de vue, de connaître et de discuter les arguments et les preuves de son adversaire, d'échanger avec lui les éléments et les pièces de son dossier, tout au long de la procédure.

Sur le fondement des articles 14 et 15 du Code de Procédure Civile, le juge veille au respect du principe du contradictoire

et soumet à la discussion les arguments soulevés lors des débats.

Le temps de la plaidoirie permet aux avocats d'insister sur les points essentiels d'une affaire et d'attirer l'attention du juge sur les arguments qui leurs paraissent fondamentaux pour trancher le litige et sur les implications que prendrait telle ou telle décision pour la personne qu'ils défendent. Ce moment d'échanges, parfois vifs, permet d'incarner des conclusions écrites et de soulever des interrogations sur les différents arguments échangés par les parties, pour permettre au juge de trancher en toute connaissance de cause.

Face aux arguments inévitablement antagonistes développés par les avocats, les juges se réfèrent à la loi pour rendre une décision mais disposent également d'un « pouvoir souverain » pour trancher les litiges.[5] Ainsi, ils apprécient, en toute indépendance, la valeur probante des éléments de preuve versés aux débats (attestations, procès-verbaux, rapports d'expertise…). Ils peuvent en interpréter la portée, les retenir ou les écarter comme non probants et disposent également d'un pouvoir souverain pour apprécier la connaissance, la compétence, l'intention des parties et leur bonne ou mauvaise foi.

Ces garanties sont fondamentales pour tout État de droit qui revendique des valeurs démocratiques et l'autorité judiciaire veille au strict respect de ces principes.

Le recours à l'intelligence artificielle pourrait difficilement respecter avec rigueur l'ensemble des principes directeurs du procès. En effet, les pièces ne seraient plus examinées, le litige ne serait plus véritablement débattu entre les parties, les juges ne disposeraient plus d'un pouvoir souverain, puisque leur décision serait orientée par le résultat donné par les plateformes algorithmiques, chargées d'anticiper le résultat du procès.

Enfin, si les juges devaient se contenter demain de reproduire ou d'être influencés par la recommandation émise par les algorithmes pour déterminer l'issue d'un litige, cela aurait pour effet de générer des résultats systématiquement identiques pour un même type de litige, ce qui aboutirait à un enfermement algorithmique. Il serait alors impossible de faire évoluer le droit, d'alerter le législateur et la Cour de Cassation sur les problèmes rencontrés sur le terrain judiciaire, dans l'application de la loi et de la nécessité de la faire évoluer.

(5) L'interprétation des contrats relève aussi du pouvoir souverain des juges du fond (sauf dénaturation), mais aussi toutes les situations créatrices ou privatrices de droits : possession, action paulienne, surendettement, déclaration des risques par un assuré, insanité d'esprit pour la nullité d'un testament, connaissance d'un vice affectant la chose vendue, évaluation d'un préjudice, appréciation du bref délai pour agir en rescision de la vente pour vices cachés, le grief causé par l'irrégularité d'un acte de procédure (article 114 du C.P.C.), caractère anormal d'un trouble de voisinage…

En conclusion, l'idée selon laquelle il existerait une inexorable répétition des affaires et donc une possible automatisation des décisions est inexacte pour ceux qui pratiquent le monde judiciaire et ce, même si cela serait parfois plus simple pour les justiciables, les magistrats et les avocats.

II. IA judiciaire : responsabilité et confiance des citoyens

1. La responsabilité

Dans le cadre des jugements qu'il rend, le juge est en principe seul responsable de sa décision.

Un citoyen qui a un doute concernant la neutralité du juge peut demander sa récusation. Si le tribunal est jugé partial, il peut être dessaisi et l'affaire peut être renvoyée devant un autre tribunal, sur le fondement de l'article 342 du Code de Procédure Civile.

Le juge ne peut déléguer ou sous-traiter sa mission de rendre justice et il lui est strictement interdit d'automatiser ses décisions d'où l'impossibilité de se fonder sur un seul barème d'évaluation, comme l'a rappelé la 1ère Chambre Civile de la Cour de Cassation par un arrêt n° 12-25301 du 23 octobre 2013.

De plus, la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés (article 10 al. 2) et le règlement européen sur la protection des données personnelles (article 22) disposent « qu'aucune décision produisant des effets juridiques à l'égard d'une personne ne peut être prise sur le seul fondement d'un traitement automatisé de données ».

Recourir à des machines ne doit pas conduire à une déresponsabilisation des acteurs judiciaires. Or, avant même de prévoir la résolution des litiges via des plateformes algorithmiques, aucune loi n'a règlementé la responsabilité des différents intervenants en matière d'intelligence artificielle. Il est donc à ce jour impossible de savoir, en cas de dommages causés à des tiers par les algorithmes, qui serait responsable et tenu de réparer les préjudices subis.

Aucun régime de responsabilité spécifique n'ayant été envisagé, la seule solution est donc de s'en référer aux régimes de responsabilités existants, à savoir notamment la responsabilité civile et pénale, pour savoir s'ils auraient vocation à s'appliquer à l'intelligence artificielle.

Or, aucun de ces régimes n'est applicable puisque pour être considéré responsable, il faut être doté de la personnalité juridique, c'est-à-dire être sujet et titulaire de droits, ce qui n'est pas le cas d'un robot ou d'une machine.

Ainsi, la responsabilité civile implique un fait de « l'homme » car l'article 1240 du Code Civil dispose que « tout fait quelconque de l'homme, qui cause à autrui un dommage, oblige celui par la faute duquel il est arrivé à le réparer ».

La responsabilité pénale, prévue par l'article 121-1 du Code Pénal, selon lequel « Nul

n'est responsable pénalement que de son propre fait », suppose également une personnalité juridique.

La responsabilité du fait des produits défectueux de l'article 1245 du Code Civil prévoit que « le producteur est responsable du dommage causé par un défaut de son produit, qu'il soit ou non lié par un contrat avec la victime » pourrait convenir davantage.

Mais cela supposerait que les algorithmes soient considérés comme une chose ou un bien meuble, ce qui n'est pas acquis à l'heure actuelle.

Face à cette complexité, certains ont proposé d'attribuer la personnalité juridique aux robots (en leur attribuant un numéro 3, comme pour le numéro 1 ou 2 de la sécurité sociale attribué aux hommes et aux femmes). Cette idée reviendrait à accorder aux algorithmes les mêmes droits que ceux dévolus aux personnes physiques et aux personnes morales (ex : de signer des contrats, de porter plainte…), ce qui n'est pas sérieux.

En réalité, dans la mesure où il est inconcevable d'envisager une totale absence de responsabilité de l'autorité judiciaire, dans le cadre d'une justice basée sur des plateformes algorithmiques, il conviendrait d'élaborer un régime de co-responsabilité.

La difficulté consisterait alors à savoir exactement, dans la chaîne des différents intervenants, qui est véritablement responsable du préjudice subi par le justiciable : le fabricant, la plate-forme, le ministère de la justice et les juridictions qui paramètrent les algorithmes ou le juge qui rend sa décision ?

Si l'on envisage une co-responsabilité entre la plate-forme et le juge, il conviendrait aussi de savoir dans quelle proportion le magistrat saisi de l'affaire s'est conformé ou non à la recommandation émise par la plateforme, ce qui aboutirait à une dilution de responsabilité ou obligerait le juge à se conformer systématiquement à la préconisation émise par la plateforme, pour éviter toute remise en cause.

2. Confiance et adhésion des citoyens

Le projet de loi de programmation et de réforme pour la Justice 2018-2022 déclare avoir pour volonté de « simplifier et de rapprocher la justice du citoyen ».

Il est certain que l'idée selon laquelle tout aléa et toute partialité seraient écartés a de quoi séduire et encore d'avantage si l'on promet une décision rapide ! Quelle serait en effet l'utilité de recourir à un magistrat ou à un avocat, puisque sur la base du Big Data, un volume phénoménal de décisions de justice pourrait permettre d'anticiper la décision du juge ?

La mise en place d'une telle solution, si inhabituelle, suppose toutefois l'adhésion et la confiance des justiciables en cette nouvelle forme de justice. Cette confiance est loin d'être acquise puisque 95 % des citoyens souhaitent que l'humain garde le contrôle sur les algorithmes, car ils craignent avant tout la perte du contrôle humain (53 %), la normativité et l'enfermement algorithmique (56 %) et la collecte disproportionnée de données personnelles (50 %), selon l'étude publiée par la CNIL en décembre 2017[6].

(6) Selon l'étude réalisée par et consultable sur https://www.maddyness.com/2017/12/05/exclu-legal-techs-francaises-tendances-2017/.

Quelle confiance peut-on accorder à ces plateformes ? Cette question nécessite que l'on examine quels sont les acteurs qui proposent ces nouveaux outils numériques.

Ces services sont principalement proposés par les legaltech, qui ne disposent d'aucune compétence juridique et dont le terme « légal » ne correspond à aucune réalité puisque les dirigeants de ces sociétés sont essentiellement issus d'écoles de commerce et ne sont pas juristes.

L'âge moyen des legaltech françaises est de 3 ans pour 69,2 % d'entre elles et si 2/3 ont été créées depuis 2015, une diminution de 40 % a été constatée pour leur création en 2017.

S'il existe environ 75 entreprises legaltech en France, les services proposés ne concernent que très minoritairement la justice prédictive (5,5 %) et la protection des données (4,4 %), étant précisé que ces services sont payants alors qu'en fonction des ressources, l'État prend en charge les frais de justice, pour garantir à chacun le droit de se défendre.

On peut également s'interroger sur le fait que les tribunaux ouvriraient des données sensibles à des sociétés privées chargées de les analyser et de les transformer, afin de les revendre ensuite aux mêmes juridictions et aux justiciables, ce qui aboutirait à une dépendance économique et à une marchandisation de la justice, peu souhaitables.

Concernant la justice prédictive, une expérimentation a été menée en 2017 avec les cours d'appels de Rennes et de Douai en partenariat avec le ministère de la justice et la société *Predictice*, sans toutefois convaincre les juges.

En effet, selon M. Xavier Ronsin, premier président de la cour d'appel de Rennes et ancien directeur de l'École Nationale de la Magistrature, la solution proposée par la legaltech n'a apporté aucune plus-value par rapport à d'autres moteurs de recherche qui permettent déjà une analyse de la jurisprudence et l'outil testé réaliserait davantage une approche statistique et quantitative que qualitative.[7]

(7) Selon le Président E. Macron dans l'interwiew du 31 mars 2018 accordée à Wired sur l'intelligence artificielle et consultable sur le site www.elysee.fr à la page http://www.elysee.fr/interviews/article/interview-du-president-de-la-republique-emmanuel-macron-pour-wired-sur-l-intelligence-artificielle-en-version-francaise/

Enfin, l'ultime problème que posent ces plateformes algorithmiques est le défaut de transparence et d'information quant aux moyens techniques et aux critères utilisés pour mettre en place des recommandations visant à aider, si ce n'est fortement influencer, la prise de décision juridique.

Conclusion

Rendre justice est une mission essentielle qui relève d'un pouvoir régalien de l'État et lui permet d'affirmer sa souveraineté sur son territoire.

Il est donc logique que les services régaliens, tels que l'armée, les forces de l'ordre et la justice, incarnés par les ministères des Armées, de l'intérieur et de la justice, soient assurés par l'administration et financés par les contribuables. Leur organisation ne se pose pas en termes de compétitivité économique : faire respecter la loi ne crée pas de richesses à proprement parler mais quelque chose de bien plus essentiel et précieux : la sécurité et le sentiment d'appartenance à une nation. C'est la raison pour laquelle le Conseil Constitutionnel a affirmé que les services publics nationaux, tels que la justice, la défense et la police, sont des services publics insusceptibles de privatisation (CC n° 86-207 DC du 26 juin 1986).

Pourtant, les *legaltech* ne cessent d'inciter les citoyens et les gouvernements successifs à leur confier le traitement du contentieux judiciaire et des pans entiers de contentieux sont désormais exclus des tribunaux avec la déjudiciarisation, qui s'accélère à chaque nouvelle réforme de la justice.

Les algorithmes et l'intelligence artificielle conduisent à une forme de dilution des figures d'autorité traditionnelles.Confier la lourde et difficile tâche de rendre la justice à des sociétés privées, dans le seul objectif de réaliser des économies, est une erreur et si les justiciables sont insatisfaits du système judiciaire actuel, il n'est pas certain qu'ils seront davantage convaincus par la nouvelle justice qui leur est proposée.

La justice prédictive pourrait avoir pour effet de creuser les inégalités sociales et de renforcer la défiance envers les institutions et alors « une chance immense deviendrait un risque immense et cela pourrait totalement mettre à mal notre cohésion nationale et notre vivre-ensemble, ce qui amène à conclure que cette gigantesque révolution technologique est en fait une révolution politique ».[8]

(8) Selon le Président E. Macron dans l'interwiew du 31 mars 2018 accordée à Wired sur l'intelligence artificielle et consultable sur le site www.elysee.fr à la page http://www.elysee.fr/interviews/article/interview-du-president-de-la-republique-emmanuel-macron-pour-wired-sur-l-intelligence-artificielle-en-version-francaise/.

L'AUTEURE

Maître Cécile Doutriaux, avocate, fondatrice du cabinet Juris Défense Avocats, est membre de la Chaire Cyberdéfense & Cybersécurité des écoles de Saint-Cyr SOGETI & THALÈS, de la Réserve Citoyenne Cyberdéfense et auditeur de l'IHEDN. Elle enseigne le droit des technologies de l'informatique et de la communication au Conservatoire National des Arts & Métiers (formation labellisée par la CNIL).

Éthique des algorithmes : attention au trompe-l'œil

Par Me Élisabeth Grosdhomme

L'éthique des algorithmes est un sujet à la mode, vulgarisé par le dilemme du tramway, ce cas d'école utilisé depuis les années 1960 dans les cours d'éthique des universités américaines, légèrement adapté pour traiter des enjeux du véhicule autonome, consistant à interpeler un interlocuteur en lui demandant de décider de la conduite à tenir dans une situation hypothétique où un véhicule autonome, dans des circonstances de circulation inopinées, n'aurait d'autre issue que de sacrifier son passager ou d'écraser tel ou tel piéton alentour (une vieille dame, un groupe d'enfants, une jeune maman, …). C'est évidemment une bonne chose que de s'interroger dès à présent sur les enjeux moraux, et non seulement techniques, des paramètres de programmation introduits dans ces dispositifs qui régiront bientôt de larges pans de notre vie. Pour autant, la focalisation excessive de l'attention sur ce sujet est préjudiciable car elle détourne la réflexion de plusieurs questions préalables d'importance majeure.

ÉLISABETH GROSDHOMME

Directeur général de Paradigmes et caetera.

La question de la pertinence de l'algorithme, de sa pure et simple performance au regard de la tâche à accomplir.

La réalité actuelle des véhicules autonomes, ou plus précisément des accidents qu'ils ont causés, n'est pas du tout celle de choix éthiques mal résolus ; c'est, beaucoup plus prosaïquement, une affaire de logiciels et de procédures défaillants.

La vidéo publiée par la police de Tempe, en Arizona, après le récent accident au cours duquel une voiture autonome de la société Uber a percuté et tué un piéton (une dame qui, à la nuit tombée, dans un endroit non éclairé, traversait la chaussée en dehors

de tout passage piéton), est tout à fait édifiante à cet égard. On y observe d'une part que les capteurs du véhicule n'ont pas détecté le piéton (ni les caméras, ce qui se comprend vu l'obscurité, ni les *lidars*[1], ce qui est plus surprenant), d'autre part que lorsque le piéton, *in extremis*, est devenu visible dans la lumière des phares, le véhicule n'a pas amorcé de freinage d'urgence, enfin que le superviseur censé s'assurer du bon fonctionnement des opérations n'était nullement en train d'accomplir sa mission de surveillance mais, la tête baissée, semblait occupé à tout autre chose, probablement lire ou consulter son téléphone.

(1) La télédétection par laser ou lidar, acronyme de l'expression en langue anglaise « light detection and ranging » ou « laser detection and ranging » (soit en français « détection et estimation de la distance par la lumière » ou « par laser »), est une technique de mesure à distance fondée sur l'analyse des propriétés d'un faisceau de lumière renvoyé vers son émetteur.

Deux ans plus tôt, en 2016, un autre accident fatal, impliquant cette fois une voiture Tesla, avait eu lieu en Floride : la voiture avait heurté un camion qui se trouvait en travers de la route car le logiciel de reconnaissance d'image avait confondu la paroi latérale du camion vu de profil, de couleur claire, avec le ciel, tuant au passage l'ingénieur d'essai qui se trouvait à bord du véhicule.

© futuristic vehicle and graphical user interface(GUI).- Adobe stock

La question de la performance de l'algorithme est distinguable de celle de l'éthique qui vise la transparence des programmations et leur inscription dans un droit positif interdisant des discriminations illégales.

Bien sûr, la technologie va progresser, mais en attendant, il ne faudrait pas que les questions éthiques servent de leurre. Lorsqu'un algorithme, que ce soit pour le guidage des véhicules autonomes, l'attribution d'un logement ou le profilage d'un risque criminel, produit trop de faux-positifs ou de faux-négatifs, il y a certes un problème d'éthique à la clef, mais avant tout un problème de performance, et donc – c'est là le point sensible, donnant lieu à un lobbying actif aux États-Unis notamment – un enjeu de mise en place d'un système d'homologation technique de nature à vérifier la fiabilité des algorithmes avant que leur déploiement ne soit autorisé.

On invoque aussi parfois l'éthique en semblant ignorer qu'il existe d'ores et déjà un cadre de droit positif

Ainsi, par exemple, la discussion du dilemme du tramway, évoqué ci-dessus, débouche généralement sur un débat tendant à savoir si, en cas d'accident inévitable, il vaudrait mieux écraser une vielle dame ou un enfant, un homme bien portant ou malade, etc. Mais ce débat est sans objet dans une démocratie comme la nôtre, où le principe d'égalité des citoyens est inscrit dans la constitution. Il est tout à fait possible qu'un raisonnement utilitariste amène à décider que la jeunesse, à qui il reste davantage d'années à vivre, doit être protégée prioritairement, ou au contraire que la sagesse et l'expérience, accumulées avec les années, ont plus de valeur. Mais peu importe : le principe fondamental d'égalité des citoyens, qui est au cœur de notre État de droit, interdit tout simplement de spéculer sur le fait qu'une vie vaudrait plus qu'une autre.

Il est d'ailleurs symptomatique que la commission d'éthique mise en place par le ministère des transports allemand, qui a remis en juin 2017 ses propositions pour un code d'éthique des véhicules connectés et autonomes, ait conclu en ce sens, prohibant toute distinction fondée sur les caractéristiques individuelles des personnes en cause (âge, sexe, état physique ou mental) de même que toute comparaison ou compensation entre les victimes potentielles.

Là encore, il vaut sans doute la peine de rappeler une évidence : la marge d'appréciation ouverte au débat éthique est strictement bornée par le cadre du droit positif qui, outre l'égalité des citoyens, stipule aussi le respect de la vie privée, la responsabilité du fait des produits et autres principes dont les concepteurs d'algorithmes ne sauraient s'exonérer.

Autre dérive du raisonnement : on en appelle à l'éthique des algorithmes pour se substituer à la défaillance du politique.

C'est tout particulièrement le cas sur les questions de discrimination, domaine exploré notamment par le *think tank* américain ProPublica qui s'efforce de débusquer le « *machine bias* », les biais implicites codés

dans les algorithmes qui viennent renforcer les inégalités présentes dans la société : inégalités entre hommes et femmes dans les algorithmes de recrutement, inégalités entre Noirs et Blancs dans les algorithmes de police préventive, etc …

Le problème est réel, mais il est permis de douter que l'éthique des algorithmes en soit la solution. Autrement dit : il est vrai qu'un algorithme de recrutement à qui l'on demande de sélectionner des candidats pour un poste de direction générale proposera probablement, toutes choses égales par ailleurs, plus de candidats masculins que féminins ; ce résultat reflètera simplement le fait que l'historique de données sur lequel l'algorithme aura été entraîné comportera inévitablement, vu la réalité de notre société, plus d'exemples de directeurs généraux que d'homologues féminines, et aura donc amené l'algorithme à considérer le critère du genre comme corrélé au profil type du titulaire du poste.

Faut-il imputer cet état de fait à l'algorithme ? Manifestement, non. L'algorithme n'est pas responsable des discriminations passées ; il ne fait que les constater.

Faut-il interdire la prise en compte du critère du genre dans le fonctionnement de l'algorithme ? C'est peine perdue : un algorithme intelligent saurait sans doute inférer le sexe des candidats de tout un faisceau d'autres indices présents dans le *curriculum vitae* – parcours de formation, déroulement des étapes de carrière passées, loisirs et centres d'intérêt déclarés, etc ; ou alors il faudrait éliminer tellement d'autres critères que l'algorithme ne pourrait tout simplement plus produire de résultats pertinents.

Serait-il possible de forcer l'algorithme à contrebalancer le poids du passé, afin de ne pas pérenniser les inégalités historiquement constatées ? Oui, il suffirait par exemple de demander à l'algorithme d'inclure dans sa sélection un nombre égal d'hommes et de femmes, charge ensuite à l'humain responsable de la décision finale de faire son choix.

Éviter la perpétuation d'un *statu quo* jugé insatisfaisant est avant tout une question de volonté politique, pas une question d'éthique des algorithmes : ils ne sont qu'un instrument, parmi d'autres, d'une stratégie globale. Qu'il s'agisse d'inégalités entre les sexes, de conditions sociales, d'identités religieuses ou encore d'origines ethniques, discuter du paramétrage des algorithmes risque à nouveau de n'être qu'une distraction du débat principal : avons-nous ou pas une volonté de transformation du *statu quo* et quelle stratégie, quels moyens nous donnons-nous pour ce faire ?

Une fois écartés tous ces trompe-l'oeil, il reste alors les enjeux propres de l'éthique des algorithmes qui, dans un État démocratique, sont essentiellement des enjeux de transparence et de mise en débat

L'essence de l'éthique des algorithmes, c'est l'explicitation de choix implicites qui étaient auparavant peut-être mal formalisés, laissés à l'appréciation de chacun ou au gré des circonstances, et qui doivent désormais être paramétrés, rapportés à des critères bien définis ; des choix autrefois individuels, au cas par cas, et qui doivent désormais s'énoncer en des règles générales.

Autrement dit, l'éthique des algorithmes c'est certes, pour une part, l'injonction faite à leurs concepteurs de traduire dans le paramétrage de leurs outils les principes moraux auxquels ils adhèrent, mais c'est surtout l'impératif d'identifier ces principes aux yeux des tiers et de permettre ainsi leur mise en discussion.

À cet égard, les dix principes pour la prise de décision algorithmique dans le secteur public récemment publiés par le think tank britannique Nesta sont un modèle en la matière. En voici quelques-uns, à titre illustratif :

« 1. Tout algorithme utilisé par une organisation du secteur public devrait être accompagné d'une description de ses fonctions, objectifs et impacts attendus, mis à disposition de ceux qui l'utilisent.

2. Les organisations du secteur public devraient publier les détails décrivant les données avec lesquelles l'algorithme a été entraîné et les hypothèses utilisées à sa création ainsi qu'une évaluation des risques pour atténuer ses biais potentiels.

4. La liste de toutes les données d'entrées utilisées par l'algorithme pour prendre une décision devrait être publiée.

5. Les citoyens doivent être informés lorsque leur traitement a été en partie ou entièrement décidé par un algorithme. »

Et puis surtout celui-ci qui interdit à l'humain de se défausser de sa responsabilité sur la machine :

« 8. Une personne de la direction doit être tenue responsable de toute mesure prise à la suite d'une décision algorithmique. »

L'AUTEURE

Élisabeth Grosdhomme-Lulin (France), est directrice générale de Paradigmes et caetera, société d'études et de conseil qu'elle a fondée en 1998, consacrée à la prospective et à l'innovation. Elle est également administrateur de plusieurs entreprises : Safran (depuis 2011) et Elsan (depuis 2015), et par le passé Société Générale (2003-2013), Ciments Français (2013-2014), Bongrain (2007-2015) et SNCF Réseau (2015-2017) ainsi que l'Institut Polytechnique LaSalle de Beauvais. Elle préside le conseil d'administration de l'ENSCI (École Nationale Supérieure de Création Industrielle), elle a siégé, au cours des années récentes, dans plusieurs commissions gouvernementales.

© Phone - Fotolia.

L'ALGORITHME EST UN OUTIL DE L'ANALYSTE POUR OPTIMISER LE RENSEIGNEMENT OPERATIONNEL ET STRATEGIQUE

L'Intelligence-Led policing ou « police guidée par le renseignement » se base sur la qualité et la contextualisation de l'analyse de la donnée.

Certes, les algorithmes peuvent identifier des relations ou itérations cachées et réaliser avec célérité le formatage de l'information, l'élaboration de graphes mais ils pêchent par leur incapacité à comprendre la donnée substantielle à la criminalité. C'est le fondement de la plus-value de l'action humaine.

Outils de l'expert, les algorithmes peuvent nourrir une analyse qui facilite un processus de décision de haut niveau afin d'identifier les menaces, des vulnérabilités mais aussi planifier l'allocation des forces ou la prise en compte de certains crimes.

Les algorithmes participent donc au processus du renseignement dans une phase qui suit la collecte de la donnée et son traitement. Les recommandations émises passent ensuite par le tamis de l'analyse par des experts dans une phase de contextualisation des faits.

Les algorithmes
et l'Intelligence-Led Policing

Par Daniel Câmara et Nicolas Valescant

Depuis ces dernières années, une transformation a lieu au sein des différentes forces de l'ordre dans le monde. De nouvelles méthodes permettent petit à petit de transposer le modèle de lutte contre la délinquance de la réaction vers la pro-action et l'anticipation. L'Intelligence-Led Policing (ILP), police guidée par le renseignement, fournit des méthodes et des outils aux forces de l'ordre pensés sur la base de l'analyse de données réelles, qui leur permettent d'identifier les priorités et de construire leur action en mettant en place une allocation des ressources adaptée[1]. Toutefois, il convient de ne pas mettre au second plan l'intuition policière, caractéristique importante acquise au contact du terrain. Il faut lui adjoindre la possibilité de prendre des décisions éclairées par l'ensemble du renseignement que nous fournit l'analyse de données.

DANIEL CÂMARA

Capitaine de Gendarmerie. Directeur du département de science des données Service de renseignement criminel de la gendarmerie nationale - SRCGN

NICOLAS VALESCANT

Capitaine de Gendarmerie. Data scientists. Département des sciences de la donnée. Service de renseignement criminel de la gendarmerie nationale - SRCGN

(1) Jerry H. Ratcliffe, Intelligence-Led Policing, ISBN 978-1843923398, Willan, 1 Édition, March, 2008. Jerry Ratcliffe est professeur au Département de justice pénale de l'Université Temple, Philadelphie, PA. Il dirige également le Centre de sécurité et de science criminelle de l'université. Il est membre du Conseil consultatif des sciences pour les programmes du Bureau de la justice du ministère de la Justice des États-Unis, membre du Conseil consultatif de l'Académie nationale du FBI et membre du groupe des Académies nationales des sciences sur les services proactifs.

La nature anticipatrice de l'ILP découle de la synergie créée par le travail commun des échelons de commandement avec les analystes, les premiers incorporant les résultats des études des seconds dans leur processus de décision et de planification.

L'implémentation de méthodes d'ILP a pour effet de refondre les protocoles et politiques « métier ». Les systèmes pyramidaux de renseignement sont repensés et s'y joignent des fonctions plus transverses et coopératives qui ont pour rôle de collecter l'information, de mettre en œuvre des traitements spécifiques et enfin de transmettre rapidement et efficacement ce renseignement aux utilisateurs finaux.

L'émergence de nouveaux modes d'action, à la fois terroristes et de criminalité conventionnelle, justifient la nécessaire transformation organisationnelle des forces de l'ordre. Dans leur activité quotidienne, lors de contacts avec la population et les acteurs de la société civile, les gendarmes mènent un travail de recueil du renseignement dont il convient d'exploiter la richesse. La pertinence du renseignement, que l'analyse de données pourra fournir, est intimement liée à la qualité de la donnée remontée. C'est pour cette raison qu'il faut fiabiliser la remontée d'information directe du terrain afin de pouvoir la normaliser et la traiter automatiquement. L'analyse du renseignement doit être intégrée au processus de planification pour que soient prises en compte les spécificités et les évolutions de la criminalité et de la lutte contre le terrorisme.

Les techniques de l'ILP sont particulièrement efficaces dans la lutte contre la criminalité organisée, mais elles sont également transposables à l'ensemble du spectre de la criminalité. Si cette évolution du processus de prise de décisions aura pour effet de transformer les forces de l'ordre, elle se fera au service de l'efficacité dans le cadre de la prévention et de la répression des actes criminels et délictueux, raisons primaires de l'engagement de chaque gendarme et policier.

Que signifie exactement le terme Intelligence de l'Intelligence-led policing ?

La signification anglo-saxonne d'intelligence diffère de son sens français, c'est pourquoi sa transcription en français est Police guidée par le renseignement. Le concept de renseignement ne recoupe pas uniquement l'information et la collecte de données, mais il est la somme à la fois de ces données et de l'utilisation qui en est faite par l'analyse. Si un accès à la donnée est nécessaire, il n'est toutefois pas suffisant pour parler d'une véritable production de renseignement. La donnée constitue le niveau le plus basique. En renseignement criminel, elle peut concerner les statistiques de criminalité, les bases de données d'auteurs, les rapports de police, etc.

L'information n'est que de la donnée mise en contexte, ce qui accroît sa pertinence.

(2) OSCE Guidebook Intelligence-Led Policing, ISBN 978-3-903128-04-0, TNTD/SPMU Publication Series Vol. 13, Vienna, June 2017.

La connaissance est de l'information interprétée et comprise. Le renseignement est de la donnée soit une somme d'informations et de connaissances évaluées, analysées et présentées sous une forme intelligible grâce à laquelle une prise de décision sera facilitée. Bien plus complète qu'une simple agrégation de données[2], la production de renseignement nécessite l'expérience de terrain et le travail d'analystes en renseignement.

Principe de fonctionnement

(3) J. H. Ratcliffe, 'Intelligence-led policing', Trends and Issues in Crime and Criminal Justice, 248, pp. 6, 2003.

Les études de Ratcliffe[3] ont montré, par la création d'un modèle conceptuel simple, comment l'ILP pouvait contribuer à la résolution d'une affaire criminelle. En appliquant le modèle des trois « I » (Interpréter, Influencer, Impacter), présenté en figure 1, les analystes interprètent de façon active l'environnement afin de mettre en évidence à la fois les acteurs en jeu dans l'affaire et leurs liens. Dans l'idéal, l'analyste devrait analyser les points de faiblesse du processus criminel et ainsi les moyens de le perturber. La nature de la relation entre les différentes parties prenantes est indiquée par le sens des flèches.

L'analyse doit influencer les décideurs, qui par leurs prises de décision et leurs actions vont avoir sur l'environnement criminel un impact. Ce dernier sera interprété par la suite et pris en compte dans l'analyse de l'environnement. Le terme décideur regroupe un spectre très large qui va du chef opérationnel, en charge des moyens à engager sur le terrain pour faire face au risque, jusqu'au législateur. Cette approche nécessite une recherche proactive de nouvelles données et sources de la part des analystes, afin de pouvoir aider les enquêteurs et décideurs à comprendre l'environnement criminel tel qu'il est sur le terrain.

La plus-value de l'ILP dans ce modèle se situe au sein de la flèche « Impact », qui peut se décliner sur de nombreux niveaux, pas seulement opérationnels. Vouloir réduire la criminalité sur le long terme peut nécessiter des actions autres que purement policières. Par exemple, cela peut être une interaction avec un constructeur automobile pour résoudre une vulnérabilité sur ses véhicules. Dans ce cas spécifique, l'action sera d'informer et d'expliquer le problème au décideur, en l'occurrence le constructeur, qui pourra à la lumière des informations fournies réviser le mécanisme en cause. Si les arguments sont assez percutants pour que le constructeur soit convaincu et agisse dans ce sens, cela contribuera à une réduction du nombre de faits.

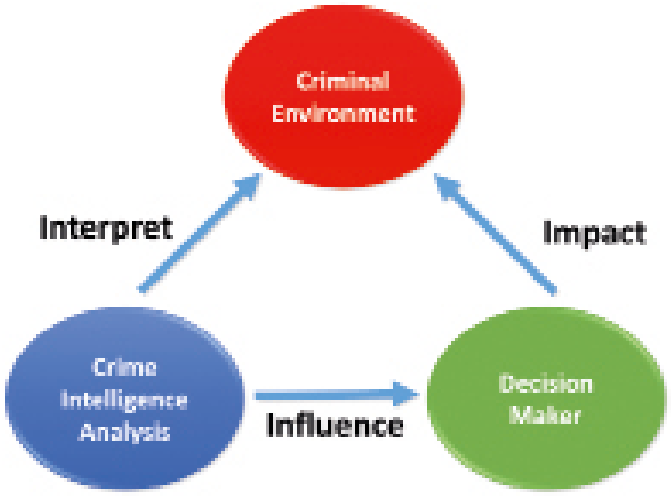


Figure 1 : Modèle des trois « I » de Ratcliffe 5.

Où et quand ce concept a-t-il été développé ?

Même si certaines initiatives antérieures existaient, il est communément admis que *l'Intelligence-Led Policing* tel qu'elle est connue aujourd'hui a été développé au début des années quatre-vingt-dix au Royaume-Uni. Elle a été créée dans le but de diminuer le temps passé par les forces de police à répondre aux faits délictueux et criminels, pour se concentrer sur le ciblage des auteurs afin de ne pas seulement réprimer les faits mais de faire cesser leur commission. L'intégration de l'ILP a été accélérée après les attentats du 11 septembre 2001 car il est alors apparu plus clairement que le partage d'information, principalement entre les différents échelons et services de police, était fondamental pour répondre aux nouvelles menaces du XXIe siècle.

Quelle est la relation entre *Intelligence-Led Policing* et les algorithmes ?

Il convient avant tout de définir le concept d'algorithme qui est un processus ou un enchaînement de règles prédéfinies et ordonnées suivies pour résoudre un problème spécifique. C'est particulièrement le cas dans le domaine de l'informatique où toute information doit être transformée en un certain nombre de règles compréhensibles par la machine qui les suivra pour effectuer son traitement. Tout processus composé d'un ensemble précis d'instructions visant à obtenir un résultat, comme l'est une recette de cuisine, est par nature un algorithme.
Le mot précède en effet de loin l'invention des ordinateurs, ses racines remontant aux travaux de mathématiciens du IXe siècle. Le concept moderne d'algorithme, aujourd'hui appliqué à l'informatique, date de 1928 quand David Hilbert présenta le problème de la décision. L'arrivée des premiers ordinateurs ne s'est produite que de nombreuses années plus tard.

L'Intelligence-Led policing se base sur l'analyse des données, qui peuvent être massives. Ainsi, si les idées et concepts principaux de l'ILP sont indépendants de l'utilisation de la puissance de calcul que fournit l'informatique, en pratique il s'avère que pour être utile le processus doit être automatisé. Les inférences et analyses doivent être traduites au sein d'algorithmes compréhensibles et exécutables par la machine. Les algorithmes peuvent identifier des relations ou itérations cachées au sein de la masse de données et fournir à l'analyste humain une information utile et plus condensée que la donnée brute. Si l'informatique excelle pour formater l'information, utiliser la donnée, élaborer des graphes ou même déceler une relation cachée au sein de milliers de rapports de police, elle n'est cependant pas efficace pour comprendre la donnée présentée. Les analystes doivent ainsi focaliser leur attention sur la plus-value de l'action humaine et laisser aux algorithmes la tâche de traiter des milliers de sources d'information qu'ils présentent sous une forme ergonomique et compréhensible.

Le cercle du renseignement est-il compatible avec *l'Intelligence-Led policing* ?

Ces deux concepts sont parfaitement compatibles et cela constitue la base du travail d'analyse. Le cercle du renseignement est un processus interactif et itératif qui guide le rythme de la production de renseignement. Il est constitué de 5 étapes : l'Expression des besoins, la Collecte, le Traitement, l'Analyse et la Diffusion.

L'expression des besoins est une étape clé, qui aide les analystes à concentrer leurs efforts. La collecte demande une masse de travail importante et délicate car les analystes doivent à la fois collecter une somme suffisante de données pour couvrir tous les aspects du sujet traité sans surcharge par une collecte trop large et inefficace. Une difficulté supplémentaire est l'accès aux données, parfois très ardu. La problématique de la collecte d'information est toutefois bien connue des forces de l'ordre qui y dédient d'importantes ressources. La transformation de l'information collectée en prise de décision éclairée dépend du processus d'analyse. Elle a pour résultat une estimation plus fiable des événements passés ou futurs et de l'impact possible de la décision[4]. Le traitement consiste à prioriser et référencer l'information collectée. Lors du traitement, les analystes doivent éliminer l'information inutile, redondante ou incorrecte et classer la donnée. Cette organisation permet de faciliter l'identification de relations entre les différentes entités.

(4) Marilyn Peterson, Intelligence-Led Policing : The New Intelligence Architecture, U.S. Department of Justice Office of Justice Programs Bureau of Justice Assistance, NCJ 210681, September 2005.

Les cibles et orientations très diverses du renseignement nécessitent des productions tout aussi diverses, qui dépendent de ce que l'analyste veut mettre en avant. Les analyses stratégiques et opérationnelles sont les plus courantes ; elles diffèrent notamment par l'horizon visé.

L'analyse stratégique a pour objectif une efficacité à moyen ou long terme. Elle vient nourrir le processus de décision de haut niveau, les décideurs politiques et la planification pouvant par exemple préconiser un changement d'allocation des forces ou modifier la prise en compte de certains crimes. Son objectif, d'un point de vue plus statistique que tourné vers des identités précises, est avant tout d'identifier les menaces clés, les vulnérabilités, les risques mais également les opportunités d'action.

L'analyse opérationnelle vise un impact fort à court terme. Elle aide le management opérationnel des forces de l'ordre dans leur action quotidienne et leurs enquêtes. Elle peut cependant comporter des informations nominatives sur des suspects ou criminels connus. La mise en œuvre de traitements informatiques visant à assister la production de ce type d'analyses peut ainsi nécessiter une déclaration CNIL. Quelle que soit la nature de la production de renseignement, elle doit être claire et précise afin d'optimiser sa plus-value dans le processus de décision (cf. note 2).

Le processus de renseignement criminel est complexe, conduit à différents niveaux et ne se limite pas au cycle du renseignement. C'est un travail continu d'identification et d'analyse de menaces émergentes.

Comment les algorithmes peuvent-ils aider ce processus ?

Les algorithmes peuvent assister l'analyste à chacune des étapes de *l'Intelligence-Led policing*. La figure 2 présente un certain nombre de fonctionnalités possibles grâce à l'application de techniques d'intelligence artificielle ou de *machine-learning* au sein du cycle du renseignement. Les algorithmes peuvent aider les analystes à recentrer leur travail sur de nouvelles hypothèses mises en évidence par des éléments difficiles à déceler sans leur aide. L'analyse statistique des faits passés peut aider à évaluer l'efficacité de certaines contre-mesures mises en œuvre.
Ainsi l'expérience condensée des faits et analyses passées optimise l'impact des actions présentes et futures.

La prochaine génération d'applications de renseignement devra apprendre automatiquement de ses erreurs et réussites. L'informatique et les algorithmes sont d'importants alliés dans la lutte contre la criminalité. Ils vont prendre une place de plus en plus importante dans le processus de prise de décision des forces de l'ordre, qui ne peuvent pas se permettre d'ignorer les avancées technologiques, notamment en raison de l'explosion de la donnée qu'il convient

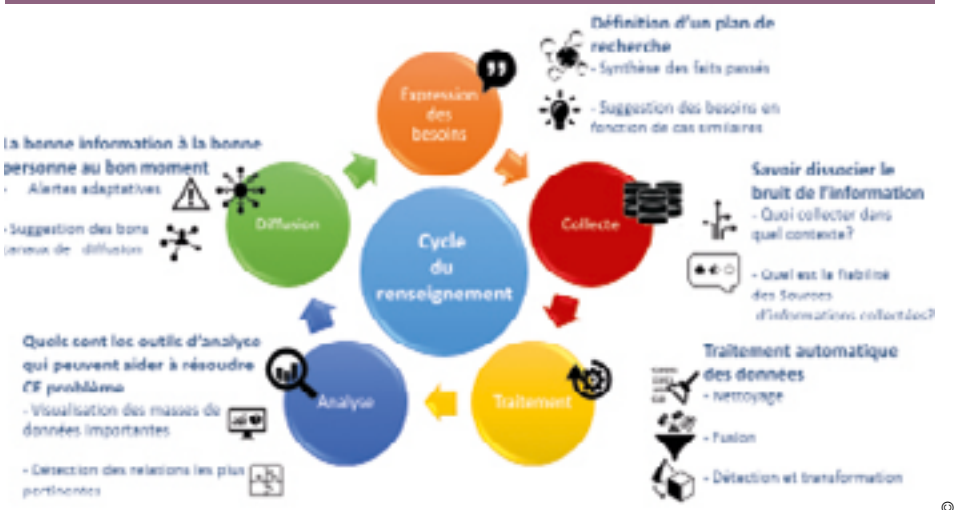


Figure 2. Comment l'intelligence artificielle peut aider dans le cercle du renseignement.

d'analyser pour apporter une réponse pertinente. La criminalité évoluant également à un rythme soutenu, les criminels et leurs réseaux utilisant les nouvelles technologies en détournant parfois leur usage principal à leur avantage, les forces de l'ordre doivent avoir la capacité de déceler rapidement ces évolutions et d'agir efficacement.

L'AUTEUR

Le Capitaine Nicolas Valescant est chercheur au service central du renseignement criminel. Sa mission est de contribuer au développement des capacités d'analyse de données du service en vue de son exploitation au profit du renseignement criminel. Notamment engagé sur le projet de l'analyse décisionnelle, il y applique des méthodes d'intelligence artificielle et d'analyse des séries temporelles en vue de comprendre et anticiper les schémas criminels. Nicolas Valescant est titulaire d'un Master en commerce et en économétrie à Kedge Business School et Queen's University.

L'AUTEUR

**Daniel Câmara est chercheur au service central du renseignement criminel où il applique les théories des réseaux sur l'analyse et la prédiction de patterns de criminalité. Il a travaillé par le passé comme chercheur à Telecom ParisTech, l'INRIA, EURECOM, et à l'UFM. Il a également été professeur au sein de l'université pontificale catholique du Minas Gerais, Brésil. Ses domaines de recherches principaux comprennent les réseaux sans fil, les systèmes distribués, la qualité logicielle et les algorithmes d'intelligence artificielle.
Le Dr. Câmara est titulaire d'un doctorat en télécommunications de Telecom ParisTech et un autre d'informatique de l'Université Fédérale du Minas Gerais, au Brésil.**

© Blueboy2014 - Adobe stock.

LA GENDARMERIE NATIONALE A RÉORIENTE SA PRATIQUE DU RENSEIGNEMENT CRIMINEL

L'adoption, par le SRCGN, de la police guidée par le renseignement (Intelligence Led-Policing) s'est accompagnée de l'intégration des échelons territoriaux à un recueil ciblé et coopératif du renseignement par une transversalité des pratiques et de la constitution d'une *task force* au sein du SRCGN constituée d'analystes capables de générer ou de paramétrer des outils algorithmiques et d'assurer une interprétation fiable de leurs recommandations.

Ce service rejoint en cela une orientation internationale qui prend en compte la nécessité d'acquérir les outils de traitement de données massives et complexes afin d'en tirer des phénomènes itératifs ou des signaux faibles. Ils peuvent être ainsi intégrés à la chaîne du renseignement et accroître l'efficacité des enquêteurs ou des décideurs.

La protection des données à caractère personnel fait que la transparence sur ces processus « métiers » doit s'inscrire dans le respect d'un encadrement juridique, harmonisé au niveau européen, et le souci de conserver l'analyste comme le garant de la pertinence des traitements, l'algorithme n'étant qu'un outil.

Vers une modernisation

des métiers de l'analyse de la criminalité : la place des algorithmes et les atouts de l'*intelligence led policing*

Par Clara Bader

Conscients de la nécessité d'appréhender la criminalité dans toute sa complexité de manière proactive et non plus seulement réactive, les forces de l'ordre et de renseignement ont peu à peu adopté les procédés de l'intelligence led policing (ILP). Pour faire face à la masse de données accessibles, notamment avec le Big Data, ces méthodes de travail sont accompagnées de réflexions sur l'importance d'utiliser des outils basés sur l'intelligence artificielle.

CLARA BADER

Capitaine de Gendarmerie. Adjointe au chef du département recherches. Division du renseignement. Service Central du Renseignement Criminel (SCRC). Pôle judiciaire de la gendarmerie nationale

Si ces derniers font polémique, c'est leur maîtrise par les analystes et leur encadrement juridique qui est le cœur des inquiétudes. Or, en préconisant la planification et l'organisation de dossiers d'analyse pour limiter la collecte de données, l'ILP semble une réponse adaptée à ces résistances et s'intègre dans la droite file du Règlement Général sur la Protection des Données annoncé par l'Union européenne pour un nouveau traitement des données personnelles.

Du rapprochement judiciaire à l'*intelligence led policing* : un changement de paradigme au sein de la gendarmerie nationale

En 2015, la gendarmerie nationale a adopté le concept de renseignement criminel, avec l'ambition de mieux exploiter l'information liée aux activités criminelles, pour guider son action. Elle s'est dotée ainsi d'une chaîne du renseignement criminel à trois niveaux (départemental, régional et national) et a transformé à l'échelon national le Service Technique de Recherches judiciaires et Documentation

(STRJD) en Service Central de Renseignement Criminel (SCRC). Ce changement de paradigme s'accompagne d'une réflexion en profondeur sur les métiers et les outils en suivant le modèle de l'*intelligence led policing (ILP)*.

Traditionnellement le travail de rapprochement judiciaire s'inscrit dans une temporalité courte qui est celle de l'enquête judiciaire, dont l'action de collecte des informations est limitée dans le temps et par le cadre juridique de l'infraction ciblée. Le renseignement criminel, dont le rapprochement judiciaire est une étape, se positionne sur un cycle long de plusieurs années, afin de couvrir l'évolution de phénomènes criminels et d'étudier des groupes criminels organisés dont les méthodes sont transverses.

Les informations sont récoltées au cours de multiples étapes : en enquête pré-judiciaire mais également sur le temps de l'enquête, de l'incarcération, en post-pénal et en renseignement administratif. La connaissance agglomérée permet dès lors d'adopter une attitude pro-active en diffusant des orientations tant opérationnelles, tactiques que stratégiques.

Le SCRC s'est doté d'une capacité de recherches et d'innovation sur l'intelligence artificielle.

Les analystes étant confrontés à la diversité des sources, des formats et à la multiplication des données collectées sur plusieurs années, le SCRC a fait le choix de recruter des ingénieurs et des *data scientist* pour développer des outils répondant aux besoins des analystes (détection de phénomènes, enrichissement du renseignement, pérennisation de l'information, etc.).

Ainsi, dans le cadre des besoins de l'analyse opérationnelle, des outils sont développés pour identifier des images, réaliser des environnements numériques sur des individus ou encore croiser des données téléphoniques pour effectuer des recoupements. L'identification automatique d'entités (numéro de téléphone, nom et prénom, numéro d'immatriculation, etc.) est également un des défis majeurs pouvant permettre de diminuer le temps d'intégration des données qui est effectué pour l'heure de manière manuelle par les analystes.

Pour ce qui est de l'étude des processus criminels, afin de tirer profit des milliers de documents judiciaires qui arrivent chaque jour au service central, les ingénieurs du SCRC étudient des algorithmes permettant de détecter de nouveaux phénomènes et d'en évaluer l'ampleur.
Ces méthodes de *machine learning*, dont la base d'apprentissage est paramétrée par les *data scientist*, en coordination avec des analystes expérimentés, devraient permettre de proposer une sélection moins aléatoire et plus homogène.

FOCUS
VERS UNE MODERNISATION DES MÉTIERS DE L'ANALYSE DE LA CRIMINALITÉ : LA PLACE DES ALGORITHMES ET LES ATOUTS DE L'INTELLIGENCE LED POLICING

© Concept of machine learning to improve artificial intelligence - Semisatch

La détection de phénomènes et de signaux faibles à partir de métadonnées est l'antichambre de l'expertise des analystes. Les analystes doivent en conséquence être capables de paramétrer les algorithmes pour obtenir des recommandations pertinentes.

En effet, la détection effectuée par des analystes dépend de leur sens de l'observation, de leur expérience mais également de leur jugement (biais cognitifs), ce qui réduit l'opportunité de rendre visibles des phénomènes que d'autres pourraient trouver pertinents.

Cependant, malgré le respect de la borne PAC (Probablement Approximativement Correct), ce qui signifie que la plupart des choix sont bons à 99 %, ce type de pré-analyse par des algorithmes basés sur des méthodes de *machine learning* ne fait pas consensus. En effet la pertinence des propositions amenées par ces algorithmes va dépendre de la masse d'informations à partir de laquelle ils auront pu apprendre. Dans un premier temps, celle-ci ne sera pas suffisante comparée aux recherches qui peuvent être faites sur le Big data. Il paraît donc indispensable que les analystes effectuent un suivi très rapproché des paramétrages et puissent réorienter les algorithmes jusqu'à l'obtention de résultats satisfaisants.

Des questionnements partagés au niveau international par les forces de police et les services de renseignement

La mise en place de systèmes de police basés sur le renseignement (*intelligence led policing*) est un mouvement international qui a commencé il y a plus de 20 ans au Royaume-Uni, en grande partie en réaction à l'absence de proactivité de la police judiciaire traditionnelle. Instaurée en Amérique du Nord puis dans de nombreux pays d'Europe, la pratique de cette méthode de travail a soulevé de multiples problématiques constructives, notamment avec l'avènement du Big Data. De fait, la montée du terrorisme et le développement d'outils de grande puissance par des services comme la *National Security Agency* (NSA) ayant considérablement complexifié la tâche des analystes, un mouvement de réflexion international s'est imposé sur le choix des méthodes d'analyse et les outils associés.

Confrontés aux mêmes écueils techniques, les services de renseignement criminel débattent depuis plusieurs années des opportunités de choisir des plates-formes d'analyse externalisées, solutions complètes développées par des sociétés privées, ou de développer en interne des briques logicielles conçues sur mesure pour les analystes des forces de police (approche combinée).

Ces réflexions internes à chaque force de police étant communes, de nombreux pays européens ressentent le besoin d'harmoniser les pratiques et d'échanger sur les outils appropriés. Des analystes experts se rencontrent ainsi dans diverses instances afin de partager leurs réflexions sur les méthodes du renseignement criminel et les questionnements que cela soulève en termes techniques et juridiques. Beaucoup se sont concentrés dans un premier temps sur le choix des outils mais les analystes ont cependant très vite pris conscience qu'il était indispensable au préalable de définir les « processus métier » et les objectifs associés afin d'aboutir à un choix éclairé et juridiquement viable.

Les possibilités techniques offertes par les algorithmes étant infinies, l'Homme devant appréhender les potentialités d'analyse qui en découlent, il est essentiel de définir les besoins en amont pour que l'ensemble des analystes apprécient la montée en compétence que ces outils représentent.

De fait, le passage à un modèle inspiré de l'*intelligence led policing* implique de multiplier les types d'analyses afin qu'elles soient directement ciblées et utiles au destinataire : analyses relationnelles, géographiques, financières, quantitatives, etc. Cela demande un saut qualitatif de ces analyses, qui doit être appuyé par des logiciels permettant d'attaquer des données complexes et d'aider l'analyste à sélectionner les données pertinentes susceptibles de faire l'objet d'un traitement.

Dès lors, cela ne doit pas être décorrélé de la mise en place de formations pour que les analystes soient en mesure d'évoluer et surtout d'adhérer à ces changements. L'expérience autrichienne a ainsi mis en avant l'importance de l'adhésion des praticiens. Si un outil n'est pas compris, testé et approuvé par les analystes, une résistance au changement peut faire qu'un outil ne sera jamais utilisé.

Une exploitation liée à un cadre juridique approprié, garant du contrôle et de la légitimité du travail des forces de l'ordre

La première étape du cycle du renseignement criminel selon l'ILP exige de planifier les sujets à analyser, de définir le périmètre d'étude ainsi que les objectifs de l'analyse. Validées par les différentes chaînes hiérarchiques, les thématiques de travail et les méthodes de traitement de l'information associées doivent faire consensus. Il s'agit de travailler par dossiers d'analyse, comme cela est effectué au sein de l'agence européenne Europol, dont les *Analysis Project* rentrent dans les grandes priorités définies pour cinq ans (EMPACT) par les États membres.

Par la suite, afin de ne pas être confronté à une multiplication de la masse d'informations, l'analyste doit définir un plan de collecte, identifier ses sources d'informations et se donner un cadre temporel. L'ensemble de ce cycle du traitement de l'information permet d'avoir un regard critique sur les finalités de l'analyse et le cadre juridique dans lequel il s'inscrit et de collecter uniquement les données nécessaires.

D'autre part, le traitement des données collectées doit faire l'objet de plusieurs mesures de contrôle par l'analyste : évaluation de la fiabilité des informations selon une grille prédéfinie, effacement des données non utilisées, classification des données sensibles, etc.

Enfin, pour ces projets d'analyse, les outils doivent être choisis pour leur adéquation avec les analyses et les résultats recherchés, dans une logique de pondération et de contrôle. L'analyste doit être maître de l'orientation des analyses en permanence et évaluer la pertinence des résultats.

Avec ces méthodes de travail, les processus d'analyse des forces de l'ordre vont donc faire gagner en efficacité comme en transparence.

Vers un changement de l'encadrement légal des algorithmes en Europe

Afin d'harmoniser juridiquement le traitement des données personnelles par les algorithmes, l'Union européenne a apporté une définition commune dans le cadre du Règlement Général sur la Protection des Données (RGPD / règlement n°2016/679). Applicable à partir du 25 mai 2018, sa transposition en France va nécessiter une refonte de la loi informatique et libertés

(n° 78-17 du 6 janvier 1978) ainsi qu'un changement d'orientation dans la façon d'aborder cette question au sein de la police et la justice. Si auparavant il suffisait de faire une déclaration préalable d'outils traitant de données personnelles, cette nouvelle directive exige d'être transparent sur l'ensemble du cycle d'utilisation des données à l'échelle d'un service.
La réflexion sur les processus métiers, les objectifs et les périmètres des analyses préconisée par l'intelligence *led policing* va ainsi permettre au SCRC de rentrer dans le cadre de cette nouvelle directive européenne.

Il est également intéressant de se pencher sur le concept de la « protection de la vie privée dès la conception » (*privacy by design*), une solution permettant aux logiciels basés sur l'intelligence artificielle de garantir la protection des données personnelles. Ce concept impose la maîtrise de l'ensemble du cycle de l'information et de son utilisation par l'outil, dès sa conception, en définissant les objectifs de l'utilisation des données.

À titre d'exemple, ce concept a été intégré au sein du logiciel développé par la société autrichienne *KIVU Technologies* de Robert WESLEY. Basés sur l'exploitation des métadonnées des communications sur les réseaux sociaux, ses algorithmes proposent des analyses relationnelles et détectent des individus potentiellement intéressants pour les analystes.

Afin de garantir le respect des libertés privées et des données personnelles, le logiciel a été conçu pour crypter immédiatement les identités des individus. Si une demande d'accès aux données est déposée, une autorité indépendante génère trois clés de déchiffrement : une pour les forces de police, une pour les représentants de la justice et une pour les autorités de contrôle parlementaire. Ces trois clés sont nécessaires ensemble pour toute consultation qui, consignée, laisse une trace numérique. Ce processus extrêmement rapide, quelques microsecondes, est automatisé à partir de critères bien définis.

Si les algorithmes offrent des potentialités infinies d'aide à l'analyse, il devient évident que leur encadrement juridique est aussi incontournable qu'une transparence sur les processus « métiers » des services de renseignement criminel et leurs objectifs de travail. Le terme étant d'atteindre un degré élevé de connaissance des groupes criminels pour adopter des mesures pro-actives, soutenues par les instances de régulation et de contrôle.

FOCUS
VERS UNE MODERNISATION DES MÉTIERS DE L'ANALYSE DE LA CRIMINALITÉ : LA PLACE DES ALGORITHMES ET LES ATOUTS DE L'INTELLIGENCE LED POLICING

L'AUTEURE

La capitaine Clara Bader est adjointe au chef du département recherche, division du renseignement au sein du Service Central du Renseignement Criminel (SCRC) du Pôle judiciaire de la gendarmerie nationale. Elle a travaillé à la mise en place d'un département spécialisé en renseignement criminel (renseignement opérationnel et tactique) et dispense des formations internes à la gendarmerie sur le renseignement criminel. Titulaire d'un master de géoéconomie et d'intelligence stratégique (Institut de relations internationales et stratégiques - IRIS) et d'un master de géopolique (Institut Français de Géopolitique, elle a acquis auparavant une expérience lors d'enquêtes de terrain à Johannesburg (Afrique du Sud) pour l'Institut Français de Géopolitique. Elle y rencontre des acteurs sur le terrain dans les bidonvilles ce qui lui permet de pratiquer une analyse des violences urbaines et des conflits sociaux et politiques autour des migrants africains. Un stage de veille et d'analyse en Intelligence Stratégique à GEOS complète un poste d'analyste au sein de la Direction Générale de la Sécurité Intérieure (DGSI) où elle œuvrait comme analyste en contre-ingérence économique - suivi des investissements étrangers avec la Direction Générale du Trésor.

© Scott Maxwell - Adobe stock.

UN DOUBLE IMPÉRATIF DE VIGILANCE ET DE LOYAUTÉ POUR LES TRAITEMENTS A CARACTÈRE PERSONNEL

L'autonomie de certains logiciels, la sensibilité du traitement des données à caractère personnel et la crainte d'une substitution d'un pouvoir décisionnel au détriment de l'homme impose de légiférer. Deux principes fondateurs doivent guider le développement des algorithmes et de l'IA : la loyauté et la vigilance. Une attention particulière doit être apportée à la phase critique et autonome du machine learning ou auto-apprentissage qui peut être réalisé sans assistance humaine. Le conseil d'État a récemment recommandé l'établissement d'un droit dédié à cette question et en ce sens, le Règlement Général sur la Protection des Données apporte une définition européenne commune à la notion de donnée à caractère personnel tout permettant une limitation de sa portée quand les informations touchent à la sécurité nationale, la défense nationale ou encore la sécurité publique.

Le défi consiste à définir ce qui peut ou doit être communiqué aux tiers. La délivrance intelligible par l'éditeur des mécanismes de l'algorithme ou l'explication de son mécanisme devrait permettre à une personne touchée par les recommandations de l'algorithme d'en saisir la logique.

Les raisons d'un droit
dédié aux algorithmes

Par Paul Brenaut

Les algorithmes occupent une place de plus en plus importante au sein de notre société. La quantité de données à laquelle nous sommes confrontés montre les limites d'un traitement entièrement humain et incite à recourir à ce type de mécanismes afin d'en tirer une interprétation optimale.

PAUL BRENAUT

Capitaine de Gendarmerie. Analyste. Service Central de Renseignement Criminel de la Gendarmerie Nationale (SCRCGN)

Qu'ils soient dédiés à la sécurité, à la santé ou encore à l'économie, les algorithmes interrogent sur la manière dont sont traitées nos données, en particulier lorsqu'il s'agit de données à caractère personnel. Bien que couverts, selon la Commission Nationale de l'Informatique et des Libertés (CNIL), par la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés (ci-après loi informatique et libertés), les algorithmes traitant des données à caractère personnel ont récemment fait l'objet d'une recommandation du Conseil d'État d'établir un droit afin
de les appréhender spécifiquement.
Le caractère autonome de certains logiciels et la crainte de voir l'humain s'éloigner de son pouvoir/devoir décisionnel sont les principaux moteurs de ce qui apparaît comme un besoin de légiférer[1]. Plus précisément, l'utilisation d'algorithmes prédictifs à l'égard des individus a conduit l'institution à émettre quelques recommandations d'encadrement. Ces dernières envisagent d'assurer l'effecti-

(1) Rapport annuel « Le numérique et les droits fondamentaux » du Conseil d'État de 2014, p. 237 http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/144000541.pdf.

vité de l'intervention humaine dans la prise de décision, de veiller à la non-discrimination, de mettre en place des garanties de procédure et de transparence et enfin de développer le contrôle des résultats produits par les algorithmes.

L'absence du terme « algorithme » au sein de la loi informatique et libertés de 1978 incite donc le législateur à davantage encadrer cette notion. Il est vrai que l'évolution technologique nécessite une adaptation juridique et une définition claire et adaptée aux termes techniques associés. C'est en ce sens que le Règlement Général sur la Protection des Données (règlement n° 2016/679) applicable à partir du 25 mai 2018 a apporté une définition commune uniformisée au sein de l'Union européenne à la notion de donnée à caractère personnel.

(2) Article L311-3-1 du Code des relations entre le public et l'administration : « Sous réserve de l'application du 2° de l'article L. 311-5, une décision individuelle prise sur le fondement d'un traitement algorithmique comporte une mention explicite en informant l'intéressé. Les règles définissant ce traitement ainsi que les principales caractéristiques de sa mise en œuvre sont communiquées par l'administration à l'intéressé s'il en fait la demande. »

Jusqu'à présent, peu de textes font état de dispositions claires visant à encadrer les algorithmes. La loi du 7 octobre 2016 pour une République numérique prévoit un droit pour les personnes faisant l'objet de décisions individuelles prises sur le fondement d'un traitement algorithmique de s'en faire communiquer le fonctionnement[2]. La solution de communication du fonctionnement des algorithmes, si elle fait office de garde-fou, soulève néanmoins quelques difficultés. Deux apparaissent comme évidentes. La première relève de la divulgation en elle-même qui pourrait poser des enjeux de violation de propriété intellectuelle ou d'intérêt national. Dans un esprit similaire, le RGPD laisse, par exemple, une marge de manœuvre aux différents États membres en leur permettant de limiter la portée des obligations concernant l'utilisation de données (y compris les décisions individuelles automatisées) pour des domaines définis tels que la sécurité nationale, la défense nationale ou encore la sécurité publique[3]. D'où un second défi qui tient à ce qui doit être communiqué : l'algorithme en lui-même ou l'explication de son mécanisme qui permettrait au profane de saisir la logique d'un procédé difficilement compréhensible.

(3) Article 23 du RGPD.

C'est ainsi que la loi pour une République numérique a confié à la CNIL la mission de conduire une réflexion sur les enjeux éthiques et les questions soulevées par l'évolution des technologies numériques[4]. Dans le rapport annuel de la commission consacré aux algorithmes et à l'intelligence artificielle, sorti en décembre 2017, deux principes fondateurs doivent guider le développement des algorithmes et de l'IA : il s'agit de la loyauté et de la vigilance.

(4) Article 59 2° e) de la loi du 7 octobre 2016 pour une République numérique.

Le principe de la loyauté dégagé par le Conseil d'État en 2014 suppose que les personnes concernées doivent être informées de manière loyale sur le traitement de leurs données à caractère personnel et leur confère notamment un droit d'accès. Selon le Conseil d'État, « la loyauté consiste à assurer de bonne foi le service de classement ou de référencement sans chercher à l'altérer ou à le détourner à des fins étrangères à l'intérêt des utilisateurs ».

Le principe de vigilance apparaît dans le rapport de la CNIL de décembre 2017 et viserait à surveiller l'aspect instable que peuvent potentiellement prendre les algorithmes, en particulier en ce qui concerne le *machine learning* ou auto-apprentissage qui apporte une autonomie d'adaptation à la machine, c'est-à-dire sans assistance humaine.

La place de l'humain, seul décideur, est justement un enjeu éminemment important. Elle est d'ailleurs consacrée à la fois par l'article 10 de la Loi Informatique et Libertés de 1978 mais également plus récemment par l'article 22 du RGPD qui tend à harmoniser les législations des États membres en ce qui concerne la protection des données personnelles. L'algorithme ne doit rester qu'un outil d'aide à la décision[5]. Il ne doit en aucun cas se substituer à l'Homme en tant que décideur. C'est en ce sens que récemment les tests de la justice prédictive dans les Cours d'appel de Rennes et de Douai ont alimenté des débats autour de l'avenir du rôle du juge. En effet, si ce dernier conserve l'obligation de rendre sa décision et son autorité souveraine de juger le fond de l'affaire, le poids plus important que prend la machine dans la justice semble échauder certaines personnes. Pourtant, il ne s'agit ici que d'un logiciel fournissant au juge des éléments permettant d'anticiper l'issue du litige (uniquement pour les affaires relevant du droit civil). Si les algorithmes permettent d'analyser la jurisprudence relative à l'affaire en cours afin de rechercher les solutions déjà établies pour des cas similaires, le juge quant à lui n'est pas soumis au logiciel et conserve son appréciation souveraine.

(5) Article 22 al. 1 du RGPD « La personne concernée a le droit de ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire ».

En matière de sécurité et notamment d'outils développés dans le cadre du renseignement criminel, l'enjeu est également de rassurer la population quant à la finalité de ces derniers et sur la façon dont ils fonctionnent. Les fantasmes, naissant autour de termes tels que « prédictif » et alimentés de surcroît par nombres de fictions décrivant un monde ultra sécuritaire, doivent être compensés par une pédagogie adaptée. Les outils dits « prédictifs » ou d'analyse décisionnelle ne reposent pas sur une prescience mais donnent un éclaircissement au décideur afin que ce dernier puisse opérer un choix.
En effet, les données collectées, puis analysées à l'aide des algorithmes offrent une vision globale sur une situation ou un problème à trancher. Ces applications indiquent ainsi la tendance la plus probable en fonction des cas déjà identifiés.

Pourtant, certaines voix soulignent le fort pouvoir de persuasion des algorithmes, affichant un résultat précis, et le risque pour le décideur d'être tenté d'opter systématiquement pour celui-ci. Loin d'être une boule de cristal, les outils prédictifs développés ont néanmoins le mérite de prendre en considération une quantité importante de données que l'humain ne peut plus se permettre à l'heure actuelle de traiter seul.
À charge pour lui, en restant le véritable décideur, de ne pas se reposer exclusivement sur ces supports et de considérer les résultats, ni plus ni moins, comme une éventualité.

L'AUTEUR

Le capitaine Paul Brenaut est depuis 2016, analyste au Service Central de Renseignement Criminel de la Gendarmerie Nationale (SCRCGN). Il est également consultant au Centre des Hautes Études du Ministère de l'Intérieur (CHEMI) où il participe à l'élaboration de modules de e-formation sur la prévention de la radicalisation. Il est titulaire d'un Master 2 - Usages sociaux du droit et communication juridique de l'Université Paris 2 Panthéon-Assas.

© Fotolia - 61280368

LA RESPONSABILITÉ DU MÉDECIN NE PEUT ÊTRE DÉLÉGUÉE À UN ALGORITHME PRÉDICTIF

Juridiquement, un algorithme prédictif est considéré comme un dispositif médical. À ce titre, il doit fournir un haut niveau de sécurité au praticien. Sa place dans la cinématique thérapeutique doit être appréciée au regard du régime de la responsabilité.

Parmi les problématiques que soulèvent son emploi, on relève la détermination du caractère impératif de ses recommandations et d'une consultation préalable comme acte de diagnostic. Le droit ne milite pas pour une soumission du praticien aux conclusions d'un algorithme. Même s'il doit tenir compte des avancées que procurent ces dispositifs, en termes de gestion de données et de techniques médicales, le praticien de santé engage sa responsabilité au regard de son diagnostic et du protocole thérapeutique qu'il a engagé.

Il faut évoquer la nécessaire information du patient et l'obtention de son consentement éclairé lorsque le traitement de son cas comprend une phase algorithmique. Il défendra ses droits en cas de résultat thérapeutique déficient en attaquant le producteur du logiciel, le praticien ou l'établissement de santé selon une stratégie judiciaire adaptée et pragmatique.

Quelle responsabilité
pour le médecin face aux algorithmes prédictifs en santé ?

Par Nathalie Nevejans

Largement négligée, la question de la responsabilité du médecin face aux algorithmes prédictifs se pose pourtant de manière urgente à l'heure où ces technologies sont en voie de déploiement dans l'ensemble du secteur de la santé et pourraient faire évoluer le rôle de l'humain dans la décision médicale.

L'une des difficultés principales en matière d'algorithme prédictif est celle de la responsabilité des humains qui y recourent. Ce problème délicat est rarement abordé par la doctrine[1]. Afin de faire avancer la réflexion, il est intéressant de concentrer le débat sur le secteur de la santé où les algorithmes prédictifs jouent un rôle de plus en plus important. En effet, ils permettent, par exemple, d'analyser une grande quantité de données afin d'anticiper des événements potentiellement mortels dans le secteur de la réanimation (*WAVE Clinic Platform aux USA*), ou encore de personnaliser les traitements et les stratégies thérapeutiques (*Watson* d'*IBM* en oncologie). Plusieurs questions doivent alors se poser.

NATHALIE NEVEJANS

Maître de conférences en droit (Faculté de droit de Douai). Membre du Comité d'éthique du CNRS (COMETS)

(1) V. cpdt, L. Mazeau, « Intelligence artificielle et responsabilité civile : le cas des logiciels d'aide à la décision en matière médicale », *Rev. Prat. Prosp. Innov.*, avril 2018, dossier 6.

Quelle est la nature juridique des algorithmes prédictifs en santé ?

Il repose sur les concepteurs de l'algorithme prédictif l'obligation de respecter les règles de mise sur le marché des dispositifs médicaux afin de fournir un matériel sûr. Le logiciel peut s'analyser en un dispositif médical s'il est destiné par le fabricant à être utilisé spécifiquement à des fins diagnostiques ou thérapeutiques (art. L. 5211-1, al. 1er, du Code de la santé publique (CSP)). Ainsi, selon

(2) « Le logiciel "dispositif médical" à l'ANSM », 3 décembre 2014, PDF p. 2, ANSM [en ligne], http://ansm.sante.fr/var/ansm_site/storage/original.pdf

l'Agence nationale de sécurité du médicament et des produits de santé, « un logiciel isolé capable d'accomplir par lui-même une finalité médicale relèvera du statut de dispositif médical »[2]. De même, le Règlement 2017/745 du 5 avril 2017 relatif aux dispositifs médicaux s'applique indiscutablement aux algorithmes prédictifs, puisque l'art. 2.1 englobe dans la définition du dispositif médical tout logiciel destiné par le fabricant à être utilisé pour des fins de prédiction d'une maladie. Le considérant 19 du Règlement précise d'ailleurs que « les logiciels spécifiquement destinés par le fabricant à une ou plusieurs des fins médicales visées dans la définition de la notion de dispositif médical, constituent, en soi, des dispositifs médicaux ».

Quelle est la valeur contraignante de l'algorithme prédictif pour le médecin ?

(3) « Éléments pour élaborer une aide à la prise de décision partagée entre patient et professionnels de santé », Fiche méthodologique, mars 2018, HAS [en ligne], https://www.has-sante.fr/portail/jcms/c_2838959/fr/elements-pour-elaborer-une-aide-a-la-prise-de-decision-partagee-entre-patient-et-professionnel-de-sante.

Même si aucune solution n'est généralisable, il convient de réfléchir à la place que l'on souhaite donner à ce nouvel instrument en médecine. En effet, l'algorithme prédictif en santé consiste, en principe, en un outil technologique d'aide à la décision du médecin en cours de déploiement. Or, dans le cadre de la prise de décision partagée concernant une question de santé individuelle, il existe déjà des aides à la décision destinées aux patients, leur permettant, notamment, de considérer les différentes options et d'être informés sur les risques et les bénéfices, qui passent par des supports numériques (vidéos, outils multimédias interactifs, forums, …)[3].

Toutefois, cette appréhension de l'algorithme en tant qu'outil d'aide à la décision pourrait évoluer, posant alors la question de la force contraignante de sa solution en santé. En effet, au regard de la qualité de l'outil, de sa pertinence ou encore de son taux de réussite, certains pourraient, par exemple, considérer que le recours à l'algorithme prédictif serait obligatoire pour le médecin, voire affirmer que l'analyse dégagée devrait s'imposer au professionnel de santé. Cette évolution des mentalités engendrerait un véritable bouleversement de la nature de cet outil d'aide à la décision. On pourrait alors aboutir à un transfert de la décision médicale vers l'algorithme prédictif, de sorte que la relation médecin-patient pourrait perdre son sens et que le professionnel de santé serait déresponsabilisé en cas de dommage.

Cette voie ne semble pas acceptable d'un point de vue déontologique et plus largement éthique. En effet, l'art. 32 du Code de déontologie médicale (C. déont.) dispose que « dès lors qu'il a accepté de répondre à une demande, le médecin s'engage à assurer personnellement au patient des soins […] ». Or, les soins seraient-ils encore personnels si un algorithme décidait à la place du professionnel de santé ? En effet, jusqu'à aujourd'hui, même si de nombreux examens et analyses passent par le biais de technologies, comme la radiographie, seul l'aspect matériel du soin est délégué à la machine, de sorte que le médecin en conserve toujours la maîtrise intellectuelle. De même, le soin reste également « personnel » lorsqu'un praticien se fait assister par un robot chirurgical, puisque c'est le chirurgien qui accomplit le geste, même s'il est réalisé in fine par la machine. C'est donc bien l'homme disposant de l'intentionalité qui maîtrise l'intervention, de sorte que la machine est reléguée à sa place d'outil. Si l'algorithme prédictif avait une valeur obligatoire pour le médecin, on assisterait alors à un changement de paradigme de la nature du soin qui serait vidé de son humanité.

À notre sens, la solution de l'algorithme prédictif ne doit pas s'imposer au médecin. Il est essentiel que cette technologie conserve le rôle d'outil d'aide à la décision, même si elle devient centrale lors des soins. Dès lors que le médecin garde sa place dans le soin, il est alors pertinent de s'interroger sur ses obligations et sa responsabilité.

Quelles obligations pour le médecin en matière d'algorithme prédictif face au patient ?

(4) Cass. 1re civ., 28 janvier 2010, n° 09-10.992, *JurisData* n° 2010-051304 ; *Resp. civ. et assur.*, 2010, comm. 85, comm. Ch. Radé.

(5) Contrairement, par exemple, à la télémédecine (art. R. 6316-2 du CSP).

En dehors d'une situation d'urgence, le médecin serait tenu d'informer le patient d'un recours à l'algorithme prédictif concernant l'acte médical, en application des dispositions des articles L. 1111-2 et L. 1111-4 du CSP, afin de recueillir son consentement, faute de quoi il engagerait sa responsabilité délictuelle[4]. Plus encore, même

L'algorithme de santé ne peut être qu'une aide à la décision même s'il intègre les dernières avancées en matière médicale.

© Doctor using modern computer with Medical record diagram - WrightStudio

si aucun texte spécifique n'existe encore en la matière[5], l'information du patient devrait s'étendre au procédé, c'est-à-dire au rôle joué par l'algorithme prédictif dans la décision médicale, comme on peut le déduire de l'art. 35 du C. déont. qui impose au médecin de donner au patient « une information loyale, claire et appropriée sur son état, les investigations et les soins qu'il lui propose ».

Quelle responsabilité pour le médecin en cas de dommage causé au patient à la suite de l'utilisation de l'algorithme prédictif ?

Tout dépendra de savoir si le dommage causé au patient est dû à une défaillance de l'algorithme ou de l'appareil qui l'héberge. En effet, selon l'art. L. 1142-1, I, du CSP, le professionnel ou l'établissement de santé n'est responsable des conséquences dommageables d'actes de prévention, de diagnostic ou de soins qu'en cas de faute, sauf s'agissant d'un matériel constituant un dispositif médical. Aussi, si le dommage était la conséquence d'un défaut de sécurité du dispositif médical, le régime de la responsabilité du fait des produits défectueux des articles 1245 à 1245-17 du Code civil (C. civ.) pourrait s'appliquer et permettre à la victime de poursuivre le producteur de l'algorithme prédictif pendant trois ans, à l'exclusion du professionnel ou de l'établissement de santé. Alors même que l'Union européenne[6] et la France[7] analysent les logiciels comme des produits au sens de l'art. 1245-2 du C. civ., le Règlement relatif aux dispositifs médicaux, dans son art. 2.1, devrait balayer l'argument selon lequel le logiciel non incorporé dans un appareil ne serait pas un produit. En effet, ce texte classe les logiciels, même utilisés seuls, parmi les dispositifs médicaux. Au demeurant, que le droit écarte ou non les algorithmes prédictifs de la catégorie des produits, la victime aurait toujours la faculté de préférer agir dans le secteur public contre l'hôpital pendant dix ans sur le fondement de la responsabilité sans faute[8] ou de poursuivre dans le secteur privé l'établissement ou le professionnel de santé libéral pendant dix ans sur le fondement de la responsabilité pour faute de l'art. L. 1142-1 du CSP[9].

(6) Rép. de Lord Cockfield au nom de la Com. eur., 15 novembre 1988, *JOCE* n° C 114/42, 8 mai 1989.

(7) Rép. Min. n°15677, *JOAN* Q, 24 août 1998, p. 4728.

(8) CE, 4 octobre 2010, n° 327449, *CHU de Besançon c/ Thomas Dutreux, JurisData* n° 2010-017829, concernant la responsabilité d'un hôpital en raison des brûlures causées par un matelas chauffant.

(9) Casse. 1re ci., 20 mars 2013, n° 12-12.300, *JurisData* n° 2013-004818, concernant la conception et la délivrance par un chirurgien-dentiste d'un appareillage.

Il convient d'ajouter qu'en l'absence de défaillance de l'algorithme prédictif, le médecin qui s'appuie sur ce dernier pour orienter son diagnostic, évaluer les risques ou encore personnaliser un soin, par exemple, reste maître de la décision médicale dont lui ou son établissement sera responsable sur le fondement de la responsabilité pour faute de l'art. L. 1142-1 du CSP. En effet, en tant que simple aide à la décision, l'algorithme prédictif livre un élément – peu importe qu'il soit prépondérant – parmi d'autres à la réflexion du professionnel de santé. Ce dernier doit évaluer la valeur de la réponse de son outil technologique à la lumière de son savoir. La voie proposée au patient est imputable au seul médecin et non à la machine, car le praticien n'est pas là pour simplement entériner la solution de l'algorithme prédictif. Sinon n'y aurait-il pas une nouvelle forme de violation de l'art. 5 du C. déont. selon lequel « le médecin ne peut aliéner son indépendance professionnelle sous quelque forme que ce soit » ?

Quelle responsabilité pour le médecin en cas de dommage causé au patient en l'absence de recours ou d'application des solutions de l'algorithme prédictif ?

Les situations peuvent être des plus variées. Le médecin peut refuser de recourir à un algorithme prédictif, ou encore s'écarter de la voie qu'il propose pour recourir à ses propres lumières. Le patient victime d'un dommage pourrait-il attaquer en responsabilité le médecin ou l'établissement de santé au motif que sa solution n'aurait pas eu les mêmes bénéfices que ceux indiqués par l'outil d'aide à la décision ?

A priori, la réponse devrait être négative, sinon cela reviendrait à dire que la solution de l'algorithme prédictif était obligatoire et que le médecin ne disposait d'aucune marge de manœuvre face à la technologie. Toutefois, les tribunaux admettraient probablement que si l'algorithme tenait compte des dernières techniques médicales[10], notamment les plus récentes[11],le médecin ou l'établissement de santé pourrait alors engager sa responsabilité pour faute pour ne pas avoir fondé ses soins sur les données acquises ou actuelles de la science en application des articles L. 1110-5 du CSP et 32 du C. déont. D'ailleurs, le médecin qui refuserait ou négligerait les voies proposées par l'algorithme prédictif pourrait violer également l'art. 33 du C. déont., qui encourage le médecin à élaborer son diagnostic en recourant au besoin aux méthodes scientifiques les mieux adaptées. En cas de dommage, cette attitude pourrait alors s'analyser en une perte de chance pour le patient de guérir ou d'éviter la mort[12].

(10) ass. 1re civ., 13 mai 1959, Bull. civ. 1959, I, n° 240.

(11) Cass. 1re civ., 27 octobre 1970, Bull. civ. 1970, I, n° 283.

(12) Cass. 1re civ., 10 janvier 1990, Bull. civ. 1990, I, n° 10.

L'AUTEURE

Nathalie Nevejans est maître de conférences HDR en droit privé à la faculté de droit de Douai et membre du Centre de recherche en droit éthique et procédures (EA n° 2471). Auteure de nombreux articles, participant à des manifestations provenant tant du monde académique que du secteur industriel, elle est l'une des spécialistes en France du droit et de l'éthique de la robotique et des technologies émergentes. Elle est également membre du Comité d'éthique du CNRS (COMETS).

© Control room of railway - Chungking - #54216714

LA SNCF : UNE STRATÉGIE VERS LA *DATA COMPANY*

La SNCF a intégré dans sa stratégie que la donnée était une valeur d'entreprise. En conséquence, elle a mis en place des structures d'expertises autour de la Big Data, de l'Intelligence artificielle et de l'Internet des Objets pour aider les différentes entités du groupe à s'approprier ces technologies, pouvoir exprimer les besoins par métiers stratégiques et cerner les usages permis par l'exploitation de la donnée. Les principaux enjeux de l'entreprise se déclinent en termes de sécurité, de service aux clients et de performance opérationnelle. La SNCF a constitué un écosystème d'open innovation qui devient un observatoire des usages possibles et de la valeur perçue des données.

La gouvernance de la donnée est pilotée dans le respect des normes légales et réglementaires ce qui a été l'occasion de clarifier les conditions de partage des données à l'intérieur du groupe. La SNCF a souhaité se positionner comme un tiers de confiance sur la question des données personnelles et la traduire par de nouveaux services directement utiles aux voyageurs pour démontrer cette posture.

SNCF : de l'algorithme
à la *data company*

Par Sébastien Pialloux

À l'ère du Digital, la capacité à exploiter la data est devenue un indicateur clé de la mesure de la valeur d'une entreprise. La SNCF génère un patrimoine de données absolument phénoménal. Avec ses 30 000 km de voies, 17 000 trains et 4 millions de voyageurs par jour, elle dispose d'un gisement de valeur particulièrement riche. Consciente de ce formidable atout, la SNCF a défini la data comme un des piliers de l'accélération de sa transformation digitale amorcée en 2015. L'industrialisation de premiers cas d'usage a permis de voir l'impact des algorithmes de datascience. Son enjeu est dorénavant d'évoluer vers une « data company ». En passant d'une approche par la technologie à une stratégie Data, en accompagnant la nécessaire adaptation des cultures et organisations, et en transformant les contraintes réglementaires en opportunités.

SÉBASTIEN PIALLOUX

Directeur Data & IoT. Direction générale e.SNCF. SNCF – Direction Digital

Centrer sa stratégie data sur les enjeux de SNCF plutôt que sur la donnée

La donnée reste un moyen sans être une fin en soi. Pour autant, la stratégie data de la SNCF s'est construite progressivement. À partir de 2015, nous avons tout d'abord mis en place des centres d'expertises autour de la Big Data, de l'Intelligence artificielle et de l'Internet des Objets (des "Fabs") pour aider les différentes entités du groupe à s'approprier ces technologies. Depuis, la soixantaine de cas d'usage traités a permis :

- d'appréhender la variété des données industrielles (base de données, texte, vidéo, photo, ...) mais aussi l'hétérogénéité de la qualité et de l'accessibilité de ces données au sein des systèmes d'information ;

- de prendre conscience de l'étendue des métiers impactés par l'utilisation de la donnée. Ainsi, les algorithmes d'études sémantiques (dits algorithme "NLP" – Natural Language Processing) ont été utilisés sur des activités aussi diverses que les prestations juridiques, la maintenance prédictive de la végétation ou la sécurité ferroviaire par l'analyse automatisée des nombreux écrits ;

- de définir une vision réaliste des structures et ressources à mettre en place.

En parallèle, nous avons créé des *assets*[1] technologiques et scientifiques.
Par exemple, des réservoirs de données (des "*datalakes*") facilitent l'accès de chacun des EPICs (*Experimental Control system*) de SNCF à ses propres données. Mais nous avons aussi mis en place des équipes d'experts de sujets de pointe comme la datascience, les capteurs IoT[2] ou encore la cybersécurité.

(1) Composant d'un système d'information qui possède une valeur pour l'entreprise et qui doit être géré.

(2) IoT = Internet of Things ou Internet des objets. Désigne des capteurs pouvant transmettre des données via des réseaux sans fil en disposant d'une autonomie pouvant aller jusqu'à plusieurs années.

Le champ des possibles s'est alors naturellement ouvert. La définition d'une stratégie data nous permet dorénavant de prendre le recul nécessaire pour éviter d'être noyés par le trop-plein de données disponibles.
Elle est la rencontre entre les questions métiers stratégiques et les usages permis par cette donnée.

© Cyber laser target on a speed concept - stefanocar_75

La valeur de la donnée, dans un cadre légal, résulte de la conjonction de l'évolution des métiers de la SNCF et des besoins des usagers, tout en prenant en compte la performance opérationnelle, les cultures et les organisations mouvantes de l'entreprise.

Cette stratégie s'ancre tout d'abord dans les principaux enjeux de l'entreprise : la sécurité, le service aux clients, la performance opérationnelle et nos collaborateurs.
Puis, nous isolons les questions majeures pour lesquelles la data serait pertinente.
Nous regardons en parallèle la facilité d'accès aux données : il est ainsi plus simple de concevoir des algorithmes de maintenance prédictive sur la dernière génération de trains nativement connectés que sur des TGV d'anciennes générations pauvres en capteurs. Enfin, nous sommes particulièrement attentifs à résoudre en priorité les irritants de nos clients ou de nos collaborateurs pour faciliter la mise en œuvre des projets et l'appropriation du sujet data par les utilisateurs.

Nous croisons ensuite deux grilles d'analyse pour identifier les usages de la donnée. La première vise à identifier le but recherché : savoir ce qu'il s'est passé, comprendre pourquoi cela s'est passé, prédire ce qu'il pourrait se passer et prescrire ce qu'il pourrait être fait. (Notons que, si les perspectives du prédictif sont séduisantes, le « savoir » et le « comprendre » sont souvent des prérequis sous-estimés).
La deuxième sert à définir si l'on cherche à faire mieux (via des outils d'aide à la décision), à faire différemment (en inventant de nouveaux services ou des process de production) ou à monétiser[3].

(3) Adaptation de la classification proposée par Bernard Marr dans son livre « Data Strategy ».

La SNCF se concentre pour l'instant sur les deux premières dimensions, mais n'exclut pas le troisième cas pour des données industrielles uniquement.

Adapter ses modes de fonctionnement pour ancrer la data company sur le terrain

Amener le groupe SNCF vers une « data company », c'est aussi faire évoluer une solide culture d'entreprise et des organisations porteuses d'exigences fortes en matière de performance opérationnelle pour transporter, chaque jour, des millions de passagers en toute sécurité.

L'enjeu est tout d'abord d'incorporer des marqueurs culturels de l'économie numérique :

- la vitesse : le rythme des innovations sur la data est tellement élevé que ce qui est valable aujourd'hui a de fortes chances d'être obsolète dans 18 mois ;

- l'itération plutôt que la planification : compte tenu de la vitesse, savoir pivoter et exécuter rapidement est plus important que concevoir des plans data à 3 ans ;

- la probabilité vérifiée plutôt que la certitude paralysante : d'une part, la donnée n'a

pas besoin d'être exhaustive ou parfaite pour rendre le service qu'on attend d'elle. D'autre part, alors qu'elle apporte une plus grande « prédictibilité », la data fait aussi passer d'une notion de certitude à une notion de probabilité. Le Big Data implique alors de changer notre relation à l'exhaustivité. Il vaut mieux un projet qui aide à mieux prédire rapidement qu'un projet qui intégrerait tous les paramètres pour tout prédire mais arriverait trop tard.

L'évolution vers la data company se traduit également par un équilibre constant à trouver sur des enjeux organisationnels. Par exemple :

- mutualisation versus décentralisation : trouver un équilibre entre un socle commun d'expertises et de technologies pour capitaliser sur des assets rares et chères, et la décentralisation de l'usage pour favoriser une appropriation rapide de la data par le terrain ;

- silos versus transversalité : un groupe large gagne en efficacité quand il est organisé en business unit plus agile. Néanmoins,la data prend généralement toute sa valeur quand elle est transversalisée. Il faut donc veiller à trouver un bon équilibre pour développer cette valeur sans pénaliser l'agilité.

Tirer parti des opportunités générées par le cadre réglementaire sur la donnée

Ensuite, le cadre réglementaire, vécu parfois comme une contrainte nécessaire, peut aussi devenir un atout pour la SNCF en concentrant, tout d'abord, notre innovation sur la valeur ajoutée que nous devons à nos clients. Le RGPD acte en effet que les données qu'ils nous confient leur appartiennent. La SNCF souhaite donc se positionner comme tiers de confiance sur ces données personnelles. Leur traitement doit alors se traduire par de nouveaux services directement utiles aux voyageurs pour démontrer cette posture de tiers de confiance. C'est par exemple le cas pour les données de géolocalisation de l'application SNCF. Elles nous ont permis de concevoir des algorithmes prédictifs afin de faciliter l'utilisation de cette application par des propositions de destinations personnalisées. Cette stratégie volontariste autour de l'expérience de voyage est différenciante par rapport à certains GAFAM, dont l'utilisation de la donnée conduit notamment à l'amélioration du ciblage publicitaire.

Par ailleurs, une attention particulière a été portée à la facilitation du partage de la donnée dans le respect du cadre juridique en vigueur. Ainsi, les textes réglementaires obligent la SNCF à préciser sa gouvernance sur la donnée. Ils permettent par ricochet de clarifier les rôles et conditions de partage de l'ensemble des données à l'intérieur du groupe. Le RGPD conduit à nommer un DPO (Data Protection Officer) mais pose en creux la question des autres rôles et responsabilités autour de la donnée. C'est le cas des Chief Data Officers qui animent l'écosystème data d'une business unit, des "data owners" qui portent la responsabilité des données, ou encore de "data stewards" qui veillent à la qualité et à l'intégrité des données. Autre exemple : les lois et décrets régissant la séparation entre SNCF Réseau et SNCF Mobilités ou encore la Loi pour une République Numérique (dite loi "Lemaire") conduisent la SNCF à cartographier et classifier ses différentes données. De ce fait, ils l'amènent à objectiver, en interne, la pertinence réglementaire de tel ou tel cloisonnement sur la donnée. Cela permet de facto d'accélérer une transversalité légale de toutes les données, quand le flou généré par l'absence de gouvernance était plutôt de nature à freiner par principe tout partage de données.

Enfin, les lois pour la croissance, l'activité et l'égalité des chances économiques (dites lois "Macron") et la loi « Lemaire » ont conforté la pertinence de notre politique d'Open data débutée en 2010. Avec plus de 200 "jeux" de données SNCF ouverts, nous avons constitué un écosystème d'open innovation autour de la donnée qui devient un observatoire des usages possibles et de la valeur perçue de nos données. La mise en place de l'API (Application Programming Interface), en 2015, nous a permis de découvrir d'autres utilisations de nos data, au-delà des applications de mobilité, telles que l'information dans des lieux publics et touristiques, l'optimisation énergétique, les robots etc. Le choix de la diffusion de données par une API (et non uniquement par téléchargements) offre en effet une connaissance fine des usages et permet d'engager la discussion avec les utilisateurs et de mesurer la valeur perçue. Cette expérience révèle en outre à quel point la frontière est ténue entre l'enjeu de créer un écosystème d'innovation autour de nos données et celui de protéger notre patrimoine immatériel. La ligne est réglementaire, elle est aussi celle de la protection commerciale et industrielle.

L'AUTEUR

Directeur Data & IoT de la Direction Digital de SNCF, Sébastien Pialloux est responsable des centres d'expertise sur la Big Data, l'Intelligence Artificielle et l'Internet des Objets. Manager confirmé du secteur des transports et du Digital, il a travaillé pendant 10 ans chez Air France KLM à la Direction du Programme (planning des vols), puis à la Direction Marketing de la division eCommerce. En 2015, il rejoint la SNCF pour monter la pratice Data.

NOTES

NOTES

NOTES

NOTES

NOTES

DIRECTEUR DE LA PUBLICATION

Général de division **Philippe GUIMBERT**

RÉDACTION

Directeur de la rédaction :
Général d'armée (2S) **Marc WATIN-AUGOUARD,**
directeur du centre de recherche de l'EOGN

RÉDACTEUR EN CHEF

Colonel (ER) **Philippe DURAND**

MAQUETTISTE PAO

Maréchal des logis-chef **Anne PELLETIER**
SDG

COMITÉ DE RÉDACTION

• Général de corps d'armée **Christian RODRIGUEZ,**
major général de la Gendarmerie nationale
• Général de corps d'armée **Thibault MORTEROL,**
Commandant des écoles de la Gendarmerie nationale
• Général de division **Philippe GUIMBERT,**
Conseiller communication du directeur général
de la Gendarmerie nationale - chef du Sirpa-gendarmerie
• Lieutenant-colonel **Jean-Marc JAFFRÉ,**
Directeur-adjoint au centre de recherche de l'EOGN

COMITÉ DE LECTURE

• Général d'armée **David GALTIER,**
Inspecteur général des armées – gendarmerie
• Général de corps d'armée **Christian RODRIGUEZ**
Major général de la Gendarmerie nationale
• Général de corps d'armée **Thibault MORTEROL,**
Commandant des écoles de la Gendarmerie nationale
• Général de corps d'armée **François GIERÉ,**
Directeur des opérations et de l'emploi
• Général de division **Philippe GUIMBERT,**
Conseiller communication du directeur général
de la Gendarmerie nationale - chef du Sirpa-gendarmerie
• Colonel **Laurent VIDAL,**
délégué au patrimoine – DGGN
• Lieutenant-colonel **Édouard EBEL,**
département gendarmerie au sein
du service historique de la Défense

DÉPOT LÉGAL

Raison sociale de l'éditeur :
CREOGN, avenue du 13e Dragons,
77010 Melun cedex
Général (2S) Watin-Augouard
Imprimerie : SDG - 11 rue Paul Claudel
87000 Limoges
Juillet 2018
ISSN 1243-5619

Adobe stock

Gendarmerie et patrimoine

Une institution multiséculaire comme la gendarmerie nationale, a traversé les formes politiques, les transformations de la société et des populations dont ses personnels sont issus. Ce numéro dépasse le patrimoine immobilier et mobilier de la gendarmerie nationale, qui est le témoin de ses savoirs faire et être, pour exprimer l'apport de la gendarmerie au patrimoine national dans les domaines littéraire, architectural, de la conservation de la mémoire, d'un art mécanique ou par la préservation de vieux métiers.