

**LE «MACHINE LEARNING» À L'ÉPREUVE DES CONTRAINTES DU RGPD : D'UNE  
DIMENSION INDIVIDUELLE A UNE DIMENSION COLLECTIVE DE LA  
PROTECTION DES DONNÉES.**

**THE "MACHINE LEARNING" PUT TO THE TEST OF THE GDPR: FROM AN  
INDIVIDUAL DIMENSION TO A COLLECTIVE DIMENSION OF DATA  
PROTECTION.**

AMIRA EL AIDOUNI

Doctorante à la Faculté des Sciences Juridiques, Economiques et Sociales, CEDOC : Droit  
Economie

FSJES-Souissi, Université Mohammed V de Rabat

Rabat, Maroc

[amira.elaidouni@um5r.ac.ma](mailto:amira.elaidouni@um5r.ac.ma)

BOUZIT MHAMMED

Docteur en sciences juridiques et politiques

FSJES-Souissi, Université Mohammed V de Rabat

Rabat, Maroc

[m.bouzit@um5r.ac.ma](mailto:m.bouzit@um5r.ac.ma)

## **RÉSUMÉ**

L'impact potentiel du Règlement général sur la protection des données (RGPD) sur les programmes de science des données est une question en vogue. Mais il n'y a peut-être pas de question plus importante, ou plus incertaine, que celle de l'impact du règlement sur l'apprentissage automatique (MACHINE LEARNING), en particulier. Compte tenu des récentes avancées dans le domaine de l'apprentissage automatique et de l'intelligence artificielle en général. Dans ce contexte, l'apprentissage automatique est en passe de devenir l'avenir de la science des données en entreprise.

Cet article vise à démystifier cette intersection entre le ML et le RGPD, en se concentrant sur l'importance de protéger les données à caractère personnel.

## **Mots clés :**

MACHINE LEARNING, intelligence artificielle, données à caractère personnel, RGPD, corrélation de données, données de masse.

**ABSTRACT**

The potential impact of the General Data Protection Regulation (GDPR) on data science programs is a hot question. But there is perhaps no more important, or more uncertain, question than the impact of the regulation on machine learning, in particular. Given the latest advancements in machine learning and artificial intelligence in general. In this context, machine learning is fast becoming the future of enterprise data science.

This article aims to demystify this intersection between ML and GDPR, focusing on the importance of protecting personal data.

**Key words :**

MACHINE LEARNING, artificial intelligence, personal data, GDPR, data correlation, big data.

*“La théorie, c’est quand on sait tout et que rien ne fonctionne. La pratique, c’est quand tout fonctionne et que personne ne sait pourquoi”<sup>1</sup>.*

## INTRODUCTION

Aujourd'hui, les machines sont capables de reproduire un comportement humain, mais sans conscience. Plus tard, leurs capacités pourraient croître au point de se transformer en machines dotées de conscience, de sensibilité et d'esprit. Mais cette capacité provient tout d'abord par ce que peuvent appeler les informaticiens : l'intelligence artificielle. À travers laquelle les professionnels en la matière mettent en œuvre un certain nombre de techniques permettant aux machines d'imiter une intelligence réelle (existante et humaine). Aujourd'hui on parle de neurones artificiels, qui peuvent agir dans divers domaines, ces neurones permettent de déduire des informations prédictives à travers l'analyse d'un flux énorme de données à caractère personnel<sup>2</sup>.

Toutes les grandes entreprises dans le monde de l'informatique (Les GAFA)<sup>3</sup> planchent aujourd'hui sur les problématiques de l'intelligence artificielle en tentant de l'appliquer à quelques domaines précis. Cette approche de traitement de lourds calculs au sein de gigantesques bases de données à travers des serveurs, peut prendre un grand nombre de formes: intelligence artificielle faible, le Machine Learning (ML) et le Deep Learning (DL), ces deux derniers sont des intelligences artificielles mais l'inverse n'est pas vrai.

Le Machine Learning appelé l'apprentissage automatique, est une forme d'intelligence artificielle, capable de reproduire un comportement grâce à des algorithmes, eux-mêmes alimentés par un grand nombre de données.

---

<sup>1</sup>Einstein. Voici comment l'intelligence artificielle et son évolution pourraient saisir le droit dans un futur déjà actuel.

<sup>2</sup> Pour plus de détails voir : Vayre Jean-Sébastien, « Les machines apprenantes et la (re)production de la société : les enjeux communicationnels de la socialisation algorithmique », *Les Enjeux de l'information et de la communication*, 2018/2 (N° 19/2), p. 93-111. DOI : 10.3917/enic.025.0093. URL : <https://www.cairn.info/revue-les-enjeux-de-l-information-et-de-la-communication-2018-2-page-93.htm>

<sup>3</sup> L'acronyme GAFA désigne quatre des entreprises les plus puissantes du monde de l'internet (et du monde tout court !) à savoir : Google, Apple, Facebook et Amazon.

L'apprentissage automatique (en anglais machine learning, littéralement « apprentissage machine ») ou apprentissage statistique est un champ d'études de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d' apprendre » à partir de données, c'est-à-dire d'améliorer leurs performances à résoudre des tâches sans être explicitement programmés pour chacune. Plus largement, il concerne la conception, l'analyse, l'optimisation, le développement et l'implémentation de telles méthodes. »<sup>4</sup>

Plusieurs informations personnelles ne sont pas liées à un individu mais à une lignée généalogique, c'est donc tous les membres de sa famille voire de son ethnie qui devraient donner le consentement demandé. Certains ont même proposé la notion de "données collectives pluripersonnelles" qui viserait l'écosystème dans lequel est insérée la personne.

Au regard du développement de l'intelligence artificielle et plus spécialement du Machine learning, on se demande si la notion de protection de données à caractère personnel peut accroître sa dimension pour passer d'une dimension individuelle à une dimension collective.

Le Machine learning (ML) est un concept utilisé pour décrire les systèmes informatiques capables d'apprendre de leurs propres expériences et de résoudre des problèmes complexes dans différentes situations, des capacités que nous pensions auparavant uniques à l'humanité.

Et ce sont les données, souvent personnelles, qui alimentent ces systèmes, leur permettant d'apprendre et de devenir intelligents.

Il est donc important pour nous d'engager la discussion dès maintenant. De quel type de cadre réglementaire avons-nous besoin pour saisir les opportunités offertes par l'IA de manière assurée et juste ? Car nous ne pouvons pas échapper au fait que l'utilisation du (ML) soulève un certain nombre de préoccupations en matière d'éthique, de sécurité, de responsabilité juridique, etc.

Cet article consacrera l'une de ces préoccupations à savoir : ***l'utilisation des données personnelles dans le ML et la question de la vie privée.***

---

<sup>4</sup> Frédéric SUR, Introduction à l'apprentissage automatique, École des Mines de Nancy, p. 11, 2020-2021.

## **I- LE CADRE JURIDIQUE ET LA LIMITATION DU MACHINE LEARNING**

Selon Antonio Casilli<sup>5</sup> « il n'y a rien de plus collectif qu'une donnée personnelle ». Antoinette Rouvroy<sup>6</sup>, quant à elle considère que les données personnelles sont en réalité toujours des «données relationnelles». Qu'en est-il de l'arsenal juridique protégeant le Machine Learning ? Et qu'en est-il de la prise de conscience des législateurs de cette dimension collective de la protection des données ?

### **1- L'UTILISATION LIBRE ET NON RÉGULÉE DU MACHINE LEARNING**

L'apprentissage automatique a eu un impact profond surtout dans le secteur de la publicité, permettant une analyse sans précédent des données pour le profilage des consommateurs. Cependant, tout cela pourrait être menacé par les nouvelles lois européennes sur la protection des données, des lois qui devraient forcer une réévaluation presque totale de la façon dont de nombreuses entreprises numériques ont fondé leurs business plan.

#### **1-1 LA NOTION DE GROUPE DANS LA PROTECTION DES DONNÉES PERSONNELLES**

Les spécialistes de la protection de la vie privée ont consacré peu de contributions aux données à caractère personnel des groupes et aux intérêts collectifs dans le traitement des données.

En ce sens, il n'existe pas de notion autonome de vie privée collective, mais seulement une attitude particulière de la vie privée individuelle dans le contexte des groupes.

Cette notion de vie privée de groupe est axée sur le secret et l'intimité et, pour cette raison, elle repose principalement sur le niveau de confiance existant entre les membres d'un groupe. La conséquence est un devoir de confidentialité. La vie privée collective concerne la violation de ce devoir. Néanmoins, cela ne représente pas un changement dans la perspective traditionnelle, qui reste basée sur le droit de l'individu à la vie privée.

---

<sup>5</sup> Antonio Casilli, QUELLE PROTECTION DE LA VIE PRIVÉE FACE AUX ATTAQUES CONTRE NOS LIBERTÉS NUMÉRIQUES ? Colloque La France dans la transformation numérique : quelle protection des droits fondamentaux ?, Conseil d'État, Paris, 6 février 2015.

<sup>6</sup> Antoinette Rouvroy. « Des données et des hommes » droits et libertés fondamentaux dans un monde de données massives. Bureau du comité consultatif de la convention pour la protection des personnes à l'égard du traitement automatisés des données à caractère personnel. Direction Générale Droits de l'Homme et Etat de droit. Strasbourg, p.26, 2016.

A l'aube des années 1980, en l'absence d'instrument juridique en mesure de prendre acte et de sanctionner « les litiges de masse », la doctrine française commence à réfléchir à l'introduction de l'action de groupe en droit français. Cependant, fort des dissensions et alternances politiques, le gouvernement français a longtemps craint l'introduction d'un tel mécanisme<sup>7</sup>.

L'action de groupe en matière de protection des données personnelles fait désormais partie du paysage légal français<sup>8</sup>.

Cet article est placé dans la section 2 du chapitre V de la loi Informatique et libertés qui concerne les droits des personnes à l'égard des traitements de données à caractère personnel. L'action de groupe est donc un nouvel outil permettant aux personnes concernées de garantir leurs droits<sup>9</sup>.

À l'ère du big data, les nouvelles technologies et les analyses prédictives permettent de collecter et d'analyser d'énormes quantités de données pour tenter d'identifier des modèles dans le comportement de groupes d'individus et de prendre des décisions qui affectent la dynamique interne des groupes, avec des conséquences sur les enjeux collectifs des personnes impliquées.

Les questions relatives à la vie privée qui découlent de cette nouvelle situation sont différentes de celles de la vie privée individuelle et de la vie privée collective. Nous ne sommes ni en présence de formes d'analyse qui ne concernent que des individus, ni en présence de groupes au sens sociologique traditionnel du terme, étant donné l'absence de conscience des membres du groupe de faire partie d'un groupe et l'absence d'interactions entre les personnes regroupées par les collecteurs de données<sup>10</sup>.

---

<sup>7</sup>SynthiaTientcheuTcheuko, L'action de groupe : arme efficace contre l'utilisation abusive des données personnelles ? 2020. <https://www.google.com/amp/s/droitdupartage.com/2019/09/30/laction-de-groupe-arme-efficace-contre-lutilisation-abusive-des-donnees-personnelles/amp/>

<sup>8</sup> Publiée au Journal officiel du 19 novembre 2016, la loi de modernisation de la justice du XXI<sup>e</sup> siècle n° 2016-1547 du 18 novembre 2016 crée un nouvel article 43 ter dans la loi n°78-17 du 6 janvier 1978 dite Informatique et libertés, afin de permettre l'introduction d'actions de groupe.

<sup>9</sup>Lexing Alain Bensoussan, Action de groupe et protection des données personnelles, 29/10/2020, <https://www.alain-bensoussan.com/avocats/action-groupe-protection-donnees/2016/12/22/>

<sup>10</sup> Hildebrandt, Mireille. (2008). DefiningProfiling: A New Type of Knowledge?. 10.1007/978-1-4020-6914-7\_2

## 1-2 LE RGPD<sup>11</sup> UN FREIN AU DÉVELOPPEMENT DU MACHINE LEARNING

La réponse à la question de savoir s'il est possible d'utiliser l'IA, et de protéger les données des personnes en le faisant, est oui. C'est à la fois possible et nécessaire afin de sauvegarder les droits fondamentaux de protection des données personnelles. Mais est-ce qu'un arsenal juridique peut freiner l'évolution de l'apprentissage automatique?

La nouvelle réglementation sur la protection des données entrée en vigueur sur le territoire européen en mai 2018<sup>12</sup> a renforcé les droits en matière de vie privée, tout en intensifiant les exigences imposées à ceux qui traitent ces données. Par conséquent les organisations ont une plus grande responsabilité à endosser dans le traitement des données personnelles conformément au règlement, d'où l'austérité des exigences de transparence.

Le RGPD et l'intelligence artificielle sont étroitement liés. Ce règlement touche les deux principaux aspects de l'apprentissage automatique (ML). Premièrement, il renforce la sécurité des données, car l'IA et la confidentialité des données vont toujours de pair. Il pose des obligations strictes aux entreprises qui collectent et traitent toute donnée personnelle.

Le développement de logiciels d'apprentissage automatique et d'intelligence artificielle est étroitement lié au Big Data. La plupart des systèmes basés sur l'IA ont besoin de grands volumes d'informations pour s'entraîner et apprendre. Généralement, les données personnelles font partie de ces ensembles de données d'entraînement. Mieux vaut d'ailleurs avoir un mauvais algorithme avec beaucoup de données qu'un bon algorithme avec peu de données. Mais ces acteurs possèdent les deux. Ce qui signifie que l'impact du RGPD sur le développement de l'IA et de l'apprentissage automatique est inévitable<sup>13</sup>.

Lorsque le GDPR utilise le terme "prise de décision automatisée"<sup>14</sup>, le règlement fait référence à tout modèle qui prend une décision sans qu'un être humain ne participe directement à la décision. Cela peut aller du "profilage" automatisé d'une personne concernée, par exemple en la classant dans des groupes spécifiques tels que "client

---

<sup>11</sup> Rappelons que l'abréviation RGPD désigne le "Règlement général sur la protection des données".

<sup>12</sup> Le 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 dit règlement général sur la protection des données (RGPD)

<sup>13</sup> Boris Barraud. L'intelligence de l'intelligence artificielle. Boris Barraud. L'intelligence artificielle, Dans toutes ses dimensions, p. 46, L'Harmattan, 2019.

<sup>14</sup> Article 22 du Règlement Général sur la Protection des Données (RGPD)

potentiel" ou "hommes de 40 à 50 ans", à la détermination de l'éligibilité directe d'un demandeur de prêt.

Les deux sections les plus importantes du GDPR pour l'apprentissage automatique sont les articles 13<sup>15</sup> et 22<sup>16</sup>, qui énoncent spécifiquement les droits d'une personne concernée par rapport au traitement automatisé.

Pour l'essentiel, le GDPR n'autorise le profilage et les décisions automatisées qu'avec le consentement exprès du sujet.

Ainsi, même lorsqu'il est autorisé à effectuer un profilage, la prise de décision automatisée doit garantir un traitement équitable et transparent, utiliser des procédures mathématiques et statistiques appropriées, et des mesures doivent être prises pour garantir l'exactitude des données du sujet utilisées dans les décisions.

## **2- ML FACE AUX PRINCIPES FONDAMENTAUX DE LA PROTECTION DES DONNÉES ÉDICTÉS PAR LE RGPD**

Les dispositions du RGPD régissent les devoirs du contrôleur des données et les droits de la personne concernée lorsque des informations personnelles sont traitées. Le RGPD s'applique donc lorsque l'intelligence artificielle utilise des données à caractère personnel, et également lorsqu'elle est utilisée pour analyser ou prendre des décisions concernant des personnes. Dans ce qui suit, nous passerons en revue les principes de la protection des données et les articles du RGPD qui sont particulièrement pertinents pour le développement et l'utilisation de l'intelligence artificielle, puisque tôt ou tard la législation marocaine doit s'aligner au guide européen pour répondre aux questions juridiques qui se multiplient davantage avec les avancées technologiques<sup>17</sup>.

Les règles régissant le traitement des données à caractère personnel reposent sur certains principes fondamentaux. L'article 5<sup>18</sup> du RGPD énumère les principes qui s'appliquent à tout

---

<sup>15</sup> L'article 13 stipule qu'une personne a le droit d'obtenir "la logique sous-jacente, ainsi que l'importance et les conséquences prévues de ce traitement pour lui."

<sup>16</sup> L'article 22 stipule que "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".

<sup>17</sup> Moins d'une année pour se conformer au RGPD, <https://www.cndp.ma/fr/dossiers/reglement.html>

<sup>18</sup> Article 5 du RGPD dispose des: "Principes relatifs au traitement des données à caractère personnel  
1-Les données à caractère personnel doivent être :



traitement de données à caractère personnel. L'essence de ces principes est que les informations personnelles doivent être utilisées de manière à protéger au mieux la vie privée de la personne concernée, et que chaque individu a le droit de décider de la manière dont ces données personnelles sont utilisées.

L'utilisation de données à caractère personnel dans le cadre du développement de l'intelligence artificielle remet en cause plusieurs de ces principes. En résumé, ces principes exigent que les données à caractère personnel soient :

- traitées de manière légale, équitable et transparente (licéité, loyauté, transparence)
- collectés pour des finalités spécifiques, expressément indiquées et justifiées, et non traités d'une manière nouvelle incompatible avec ces finalités (principe de limitation de la finalité)
- adéquates, pertinentes et limitées à ce qui est nécessaire pour atteindre les finalités pour lesquelles elles sont traitées (principe de minimisation des données)
- correct et, si nécessaire, mis à jour (principe d'exactitude)

---

a) traitées de manière licite, loyale et transparente au regard de la personne concernée (licéité, loyauté, transparence);

b) collectées pour des finalités déterminées, explicites et légitimes, et ne pas être traitées ultérieurement d'une manière incompatible avec ces finalités; le traitement ultérieur à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques n'est pas considéré, conformément à l'article 89, paragraphe 1, comme incompatible avec les finalités initiales (limitation des finalités);

c) adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées (minimisation des données);

d) exactes et, si nécessaire, tenues à jour; toutes les mesures raisonnables doivent être prises pour que les données à caractère personnel qui sont inexactes, eu égard aux finalités pour lesquelles elles sont traitées, soient effacées ou rectifiées sans tarder (exactitude);

e) conservées sous une forme permettant l'identification des personnes concernées pendant une durée n'excédant pas celle nécessaire au regard des finalités pour lesquelles elles sont traitées; les données à caractère personnel peuvent être conservées pour des durées plus longues dans la mesure où elles seront traitées exclusivement à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques conformément à l'article 89, paragraphe 1, pour autant que soient mises en œuvre les mesures techniques et organisationnelles appropriées requises par le présent règlement afin de garantir les droits et libertés de la personne concernée (limitation de la conservation);

f) traitées de façon à garantir une sécurité appropriée des données à caractère personnel, y compris la protection contre le traitement non autorisé ou illicite et contre la perte, la destruction ou les dégâts d'origine accidentelle, à l'aide de mesures techniques ou organisationnelles appropriées (intégrité et confidentialité);

2-Le responsable du traitement est responsable du respect du paragraphe 1 et est en mesure de démontrer que celui-ci est respecté (responsabilité)".

- ne pas être conservées sous une forme identifiable pendant une durée excédant celle nécessaire à la réalisation des finalités (principe relatif aux durées de conservation des données)
- traitées de manière à assurer une protection adéquate des données à caractère personnel (principe d'intégrité et de confidentialité)

En outre, le responsable du traitement est garant du respect des principes et doit être en mesure de le prouver (principe de responsabilité). Dans ce qui suit, nous passerons en revue les principaux défis en matière de protection des données liés au développement et à l'utilisation de l'intelligence artificielle. Nous examinons ces défis à la lumière des principes de protection des données les plus pertinents pour l'intelligence artificielle, à savoir les principes de loyauté, de limitation des finalités, de minimisation des données et de transparence.

## **2-1 LE MACHINE LEARNING DOIT RÉPONDRE AU PRINCIPE DE LOYAUTÉ<sup>19</sup>**

Il est facile de penser que l'intelligence artificielle sera capable d'effectuer des analyses plus objectives et donc de prendre de meilleures décisions que les êtres humains. Après tout, l'intelligence artificielle ne sera pas affectée par l'hypoglycémie, par une mauvaise journée ou par le désir d'aider un ami. Et pourtant, les algorithmes et les modèles ne sont pas plus objectifs que les personnes qui les conçoivent et les construisent, et que les données personnelles qui sont utilisées.

Le résultat du modèle peut être incorrect ou discriminatoire si les données rendent une image biaisée de la réalité, ou si elles n'ont aucun rapport avec le domaine en question. Une telle utilisation des données à caractère personnel serait contraire au principe de loyauté. Ce principe exige que tout traitement d'informations personnelles soit effectué dans le respect des intérêts de la personne concernée et que les données soient utilisées conformément à ce qu'elle peut raisonnablement attendre.

Comme ce fut le cas dans le scandale « Cambridge Analytica » où les utilisateurs de Facebook qui ont répondu au test de personnalité en cause étaient amenés à croire qu'ils opéraient dans le cadre d'une étude universitaire et que le but poursuivi était donc

---

<sup>19</sup>Préface du RGPD, considérant 60.

académique, alors qu'en réalité le but de la récolte des données était commercial et de prospection politique<sup>20</sup>.

Le principe exige également que le responsable du traitement mette en œuvre des mesures visant à prévenir le traitement discriminatoire et arbitraire des données individuelles. La préface du règlement décrit l'utilisation de procédures mathématiques ou statistiques adéquates aux fins du profilage dans ce domaine.

Toutefois, cela ne suffirait pas en soi pour garantir le respect du principe. Le modèle doit également être formé en utilisant des données pertinentes et correctes et il doit apprendre sur quelles données mettre l'accent. Le modèle ne doit pas mettre l'accent sur les informations relatives à l'origine raciale ou ethnique, aux opinions politiques, à la religion ou aux convictions, à l'appartenance à un syndicat, à l'état génétique, à l'état de santé ou à l'orientation sexuelle si cela devait entraîner un traitement discriminatoire arbitraire<sup>21</sup>.

Si l'on soupçonne, ou si l'on prétend, que l'utilisation d'un modèle entraînera des résultats inévitables ou discriminatoires, l'autorité de protection des données peut examiner si le principe de loyauté a été sauvegardé dans le traitement des données à caractère personnel. Ces enquêtes peuvent comprendre un examen de la documentation qui sous-tend la sélection des données, un examen de la manière dont l'algorithme a été développé et la question de savoir s'il a été correctement testé avant son utilisation.

Plus loin, le RGPD, après avoir mentionné l'obligation de ce principe de loyauté, mentionne, sans les attribuer directement au principe, deux obligations qui y sont liées : l'obligation de finalités « déterminées, explicites et légitimes » (c'est le principe de proportionnalité), ainsi qu'un traitement « adéquat, pertinent et limité » des données<sup>22</sup>.

---

<sup>20</sup> Cécile de Terwangne, Les principes relatifs au traitement des données à caractère personnel et à sa licéité, p. 90, 2018.

<sup>21</sup> Préface du RGPD, considérant 71.

<sup>22</sup> Aurélien Bamde, RGPD: le principe de proportionnalité, Droit des personnes, Loi informatique et libertés, RGPD, Déc 14, 2018. <https://aurelienbamde.com/2018/12/14/rgpd-le-principe-de-proportionnalite/#:~:text=Le%20RGPD%20le%20qualifie%20encore%20de%20principe%20de%20minimisation%20des%20donn%C3%A9es.&text=Au%20fond%2C%20la%20r%C3%A8gle%20v%C3%A9hicul%C3%A9e,d%C3%A9faut%2C%20le%20traitement%20est%20illicite>.

Ces nouvelles définitions se passent de commentaire, et permettent déjà de bien mieux cerner le principe de collecte loyale<sup>23</sup>.

## **2-2 LE MACHINE LEARNING DOIT RÉPONDRE AU PRINCIPE DE LIMITATION DE LA FINALITÉ**

De nombreux modèles développés à l'aide de l'intelligence artificielle seront utilisés en relation avec de bonnes causes, comme le diagnostic du cancer. Sommes-nous autorisés à utiliser les données personnelles sans restriction tant que c'est pour une bonne cause ? Le principe de limitation de la finalité signifie que la raison du traitement des données personnelles doit être clairement établie et indiquée lors de la collecte des données. Cela est essentiel pour que la personne concernée puisse exercer un contrôle sur l'utilisation de ses informations<sup>24</sup>.

La finalité du traitement doit également être pleinement expliquée à la personne concernée afin qu'elle puisse choisir en connaissance de cause et de donner ou non son consentement. Pourtant, le développement et l'application de l'intelligence artificielle nécessitent souvent de nombreux types de données à caractère personnel différents, des informations qui, dans certains cas, ont été collectées à d'autres fins. Par exemple, il est possible que les activités d'une personne sur Facebook soient intégrées dans un algorithme qui détermine si elle obtiendra ou non un prêt hypothécaire de la banque.

Un tel recyclage des informations peut être utile et fournir des analyses plus précises que celles qui étaient techniquement réalisables auparavant, mais il peut également être contraire au principe de limitation des finalités. Dans les cas où des données à caractère personnel récupérées précédemment doivent être réutilisées, le responsable du traitement doit examiner si la nouvelle finalité est compatible avec la finalité initiale.

Si ce n'est pas le cas, un nouveau consentement est nécessaire ou la base du traitement doit être modifiée. Dans l'exemple de Facebook évoqué ci-dessus, la personne concernée doit consentir à ce que les informations de Facebook soient utilisées par la banque dans le cadre

---

<sup>23</sup> Le RGPD : nouveau droit de la protection des données personnelles, Zeste de Savoir, 26 avril 2020, <https://zestedesavoir.com/tutoriels/pdf/2529/le-rgpd-nouveau-droit-de-la-protection-des-donnees-personnelles.pdf>

<sup>24</sup> Aurélien Bamdè, RGPD: le principe de finalité, Le Droit dans tous ses états, 2018. <https://aurelienbamde.com/2018/12/14/rgpd-le-principe-de-finalite/#:~:text=L'article%206%2C%202%C2%B0,mani%C3%A8re%20incompatible%20avec%20ces%20finalit%C3%A9s%20%C2%BB.>

de demandes de prêts hypothécaires afin de garantir que le traitement soit effectué dans le respect du principe de limitation de la finalité.

### **2-3 LE MACHINE LEARNING DOIT RÉPONDRE A LA MINIMISATION DES DONNÉES**

Il faut souvent d'énormes quantités de données personnelles pour développer l'intelligence artificielle. D'autre part, le principe de minimisation des données exige que les données utilisées soient adéquates, pertinentes et limitées à ce qui est nécessaire pour atteindre la finalité pour laquelle elles sont traitées. Cela signifie qu'un responsable du traitement ne peut pas utiliser plus de données à caractère personnel que nécessaire et que les informations sélectionnées doivent être pertinentes par rapport à la finalité<sup>25</sup>.

Un des défis du développement de l'IA est la difficulté existante à définir la finalité du traitement car il n'est pas possible de prévoir ce que l'algorithme apprendra. La finalité peut également être modifiée au fur et à mesure que la machine apprend et se développe. Cela remet en cause le principe de minimisation des données car il est difficile de définir quelles données sont nécessaires<sup>26</sup>.

Cependant, la minimisation des données est plus qu'un principe limitant la quantité de détails inclus dans la formation ou dans l'utilisation d'un modèle. Le principe stipule également la proportionnalité, qui limite l'étendue de l'intervention dans la vie privée d'une personne concernée que l'utilisation des données personnelles peut impliquer. Cela peut être réalisé en rendant difficile l'identification des individus contenus dans les bases de données. Le degré d'identification est limité à la fois par la quantité et la nature des informations utilisées, car certains détails en disent plus sur une personne que d'autres. L'utilisation de techniques de pseudonymisation ou de cryptage protège l'identité de la personne concernée et contribue à limiter le degré d'intervention<sup>27</sup>.

---

<sup>25</sup> Article 5, RGPD, "Principes relatifs au traitement des données à caractère personnel"

<sup>26</sup> Mathias le Masne de Chermont, Adrien Aulas, Hugo Ruggieri, Protection par défaut et minimisation des données... quelles différences ? 22 janvier 2018, <https://aeonlaw.eu/protection-par-defaut-minimisation-des-donnees-queelles-differences/>

<sup>27</sup> Pour plus de détails, voir : Meiko Jensen (Université de Kiel), Cédric Lauradoux (INRIA), Konstantinos Limniotis (HDPa) TECHNIQUES ET MEILLEURES PRATIQUES DE PSEUDONYMISATION, Recommandations sur l'usage des technologies conformément aux dispositions en matière de protection des données et de respect de la vie privée, Agence de l'Union européenne pour la cybersécurité (ENISA), p. 12, NOVEMBRE 2019. [https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices\\_fr](https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices_fr)

Ce principe oblige également les développeurs à examiner minutieusement le domaine d'application prévu du modèle afin de faciliter la sélection des données pertinentes nécessaires à la finalité.

En outre, le développeur doit réfléchir à la manière d'atteindre l'objectif d'une manière qui soit la moins invasive possible pour les personnes concernées. Les évaluations réalisées doivent être documentées, afin de pouvoir être présentées à l'autorité de protection des données en cas d'inspection ou dans le cadre d'une discussion préliminaire.

Bien qu'il soit difficile d'établir à l'avance les informations exactes qui seront nécessaires et pertinentes pour le développement d'un algorithme, il est essentiel que le principe de minimisation des données soit respecté grâce à une évaluation continue des besoins réels. Cela permet non seulement de protéger les droits des personnes concernées, mais aussi de réduire au minimum le risque que des informations non pertinentes conduisent l'algorithme à trouver des corrélations qui, au lieu d'être significatives, sont fortuites et auxquelles il ne faut pas accorder de poids.

La pression pour l'utilisation des données personnelles s'intensifie à mesure que les analyses basées sur l'IA sont utilisées pour promouvoir une efficacité accrue et de meilleurs services. L'autorité de protection des données estime que le principe de minimisation des données devrait jouer un rôle majeur dans le développement de l'intelligence artificielle afin que les droits des personnes concernées soient protégés et que la confiance générale dans les modèles soit maintenue.

Les décisions individuelles automatisées sont des décisions relatives aux personnes qui sont basées sur un traitement automatique. Un exemple en est l'imposition d'une amende sur la base d'une image enregistrée par un radar automatique. Les décisions automatisées sont définies et réglementées à l'article 22 du RGPD. Pour l'essentiel, les décisions individuelles automatisées ne sont pas autorisées. Des exceptions s'appliquent toutefois si la décision automatisée est une condition nécessaire à la conclusion d'un contrat, si elle est autorisée par la loi ou si elle est fondée sur le consentement explicite de la personne concernée.

Le règlement ne définit pas ce qui constitue un consentement explicite par opposition à un consentement ordinaire, mais l'expression indique qu'un geste explicite de la personne concernée est nécessaire. Afin de satisfaire aux exigences du règlement, la décision doit être

fondée uniquement sur un traitement automatisé et produire des effets juridiques ou affecter une personne de manière significative<sup>28</sup>.

Le fait qu'une décision automatisée doit être fondée uniquement sur un traitement automatisé signifie qu'il ne peut y avoir aucune forme d'intervention humaine dans le processus décisionnel. "L'intervention humaine" signifie qu'une personne physique doit avoir procédé à une évaluation indépendante des données à caractère personnel sous-jacentes et être autorisée à réexaminer les recommandations que le modèle a produites<sup>29</sup>.

## **II- LA DIMENSION COLLECTIVE DE LA PROTECTION DES DONNÉES À CARACTÈRE PERSONNEL**

La révolution du Big data multiplie les pratiques de recueil de données dans tous les secteurs professionnels et provoque une nouvelle explosion de la masse de données disponibles. Il faut dire que les données aujourd'hui sont de moins en moins "personnelles". Les données sont plutôt devenues relationnelles. La subjectivation de ce droit devient obsolète par rapport aux caractéristiques relationnelles des données et par rapport à la personnalisation d'un droit de recours. Comment les algorithmes pourront-ils désormais prédire et analyser toutes ces bases de données corrélées ? Quel droit prévoir pour l'intelligence artificielle en général et spécifiquement pour l'apprentissage automatique (ML)?

### **1- PRÉDICTION ET ANALYSE DE DONNÉES BASÉES SUR DES CORRÉLATIONS DE DONNÉES**

L'analyse des corrélations de données a pour but la description de données conjointes. On cherche par cette méthode à donner les liens pouvant exister entre les différentes données ainsi qu'à en tirer des informations qui servent à décrire de façon plus succincte les principales informations contenues dans ces données.

Google utilise par exemple l'outil Google Trends, qui mesure la fréquence des mots ou expressions tapés par les internautes dans leur barre de recherche ; Google Analytics, ou encore Google Correlate qui établit des corrélations entre les données.

---

<sup>28</sup> Wachter, Sandra & Mittelstadt, Brent & Floridi, Luciano. Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation, p. 8, DOI: 10.1093/idpl/ixp005, 2016

<sup>29</sup> Stéphanie Faber, Lignes directrices du G29 sur les « décisions individuelles automatisées » et le « profilage », Protection des données personnelles – RGPD, La Revue Hors-série, p. 8, Janvier 2018.

Ces statistiques permettent de croiser informations et expériences anonymisées et d'adapter son comportement en fonction des résultats, dans une approche prédictive<sup>30</sup>.

Ainsi, la méthode Google rectifie les résultats biaisés des sondages dans lesquels les citoyens mentent encore. En outre, elle permet de mesurer l'impact d'une politique sur la population et de l'ajuster en fonction. À titre d'exemple, Barack Obama, à l'occasion de sa deuxième campagne présidentielle en 2012, son équipe avait testé plusieurs options pour retenir le slogan « Change / We can believe in » et une photo de famille<sup>31</sup>.

En résumé, Seth Stephens-Davidowitz<sup>32</sup> remarque que « les recherches Google constituent la plus importante base de données jamais collectées sur la psyché humaine ». Ce que l'internaute croit relever de sa sphère intime et personnelle, Google le voit.

L'analyse des Big data permet de déduire des informations prédictives à partir de grandes quantités de données afin d'acquérir des connaissances supplémentaires sur les individus et les groupes, qui ne sont pas nécessairement liées aux objectifs initiaux de la collecte de données. En outre, l'analyse regroupe les personnes en fonction de leurs attributs qualitatifs et de leurs habitudes (par exemple, les personnes à faible revenu...) et prédit le comportement futur de ces groupes d'individus.

Cette approche est adoptée, par exemple, par certaines compagnies d'assurance maladie, qui extraient des informations prédictives sur les risques associés à des segments de clients. On peut citer par exemple, les dossiers clients scorés tout au long du processus d'indemnisation, sur des problématiques de détection de fraude, de calcul de provisions, de recouvrement ou encore pour estimer la complexité des dossiers. Ce procédé se déroule en continu et en parallèle du traitement de la déclaration de sinistre permettant une actualisation permanente de l'évaluation du risque<sup>33</sup>.

---

<sup>30</sup> JurilexBlog, COMMENT FAIRE DES STATISTIQUES COMME GOOGLE ? <https://www.haas-avocats.com/actualite-juridique/comment-faire-des-statistiques-comme-google/>

<sup>31</sup> Big Browser (12 juillet 2017), "Ce que l'analyse des recherches sur Google nous apprend sur la psyché humaine", Le monde, 12 juillet 2017, [https://www.lemonde.fr/big-browser/article/2017/07/11/ce-que-l-analyse-des-recherches-sur-google-nous-apprend-sur-la-psyche-humaine\\_5159228\\_4832693.html](https://www.lemonde.fr/big-browser/article/2017/07/11/ce-que-l-analyse-des-recherches-sur-google-nous-apprend-sur-la-psyche-humaine_5159228_4832693.html)

<sup>32</sup> Ancien scientifique des données chez Google

<sup>33</sup> Benjamin Boulvert, LES ENJEUX DE L'ANALYSE PRÉDICTIONNELLE DANS LE MÉTIER DE L'ASSURANCE, 04.05.2015, <https://www.insurancespeaker-wavestone.com/2015/05/les-enjeux-de-l-analyse-predictive-dans-le-metier-de-l-assurance/>



Dans ces cas, les prédictions basées sur des corrélations n'affectent pas seulement les individus, qui peuvent agir différemment du reste du groupe auquel ils ont été affectés, mais elles affectent également l'ensemble du groupe et le distinguent du reste de la société.

Ces questions ne sont pas nouvelles et peuvent être considérées comme l'effet de l'évolution de diverses technologies de prédiction, dans un contexte caractérisé par un volume accru d'informations disponibles et de puissants logiciels d'analyse. Néanmoins, les formes précédentes de catégorisation et de profilage étaient basées sur quelques variables standards (par exemple le sexe, l'âge, le revenu, l'état civil, le lieu de résidence). Leur capacité prédictive était donc limitée. Aujourd'hui, l'analyse des Big Data utilise des centaines de variables différentes pour déduire des informations prédictives sur des groupes de personnes et, dans de nombreux cas, ces variables concernent des aspects qui ne sont pas clairement liés aux profils finaux créés par les analyses.

Le RGPD sur ce point consacre le droit pour la personne concernée par le profilage de ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire<sup>34</sup>.

La dichotomie entre les individus et les groupes n'est pas nouvelle et a déjà été analysée en ce qui concerne les aspects juridiques des informations personnelles.

Véritablement, la nature inédite sur le plan quantitatif des signaux émis par les individus au travers de leurs diverses communications ou par le biais d'objets connectés pose de nouvelles contraintes quant à leur traitement.

L'analyse et l'interprétation de toute cette masse de données corrélée requiert l'utilisation d'algorithmes particuliers, disposant d'un degré d'autonomie remarquable appuyé par des procédés d'apprentissage et de liaison statistique.

Certes l'IA est la figure de proue, mais le Machine Learning, le Deep Learning et d'autres formes d'IA permettent la corrélation des données à caractère personnel fournies par diverses personnes.

---

<sup>34</sup> Article 22 de la RGPD. Décision individuelle automatisée, y compris le profilage.

## 1-1 LA DIFFICULTÉ DU TRAITEMENT DES DONNÉES DE MASSE

« *Le vice, toujours sombre, aime l'obscurité.* » Bouileau, Épître VIII

L'exploitation de la masse de données numériques produite par les utilisateurs est considérée aujourd'hui comme une opportunité en termes d'innovation économique et stratégique, une forme inédite de création de valeur.

Les données personnelles représentent en effet la matière première à partir de laquelle il est possible d'analyser, de classer les activités, de suivre les comportements et de prédire les centres d'intérêt d'utilisateurs largement dépendants d'un réseau numérique de plus en plus dense et omniprésent. Les informations qui en sont extraites permettent ainsi de délivrer, en temps réel, des offres de plus en plus personnalisées. Ces opérations si rapides et transparentes échappent largement au filtre critique des utilisateurs, peu conscients de la valeur ajoutée de ces traces égrenées au fil de leur trajet numérique, ni de l'étendue de leur utilisation. Le contenu est l'un des plus hétérogènes. Divers contenus sont issus de tout type de plateformes (fixes et mobiles, de l'internet des objets....) et forment une masse en apparence infinie, dont la collecte quasi systématique, est catalysée par un accroissement des capacités de stockage et de traitement de l'information.

Une fois stockées, mémorisées, les données souvent qualifiées de « brutes », apparaissant incohérentes et individuellement anodines, représentent de manière effective des comportements humains complexes et changeants. L'analyse de ces données permettra de connaître a priori les règles qui devraient les régir, en requérant la mise en œuvre de procédés de traitements automatisés élaborés.

Notons que tous les protocoles de navigation permettant de préserver l'anonymat de leurs utilisateurs sont méconnus par les utilisateurs qu'on pourrait qualifier de normaux, ils sont surtout connus par un ensemble de personnes ayant intérêt à vivre caché. Ces derniers utilisent le protocole d' " Onion Router" <sup>35</sup>.

---

<sup>35</sup> La donnée est chiffrée par couches successives – comme un oignon – pour en garantir la confidentialité.

## 1-2 L'ANALYSE DE DONNÉES CORRÉLÉES ET L'ENJEU DE PRÉSERVER L'INTÉRÊT GÉNÉRAL ET LA VIE PRIVÉE DES UTILISATEURS

*« Il s'agit de trouver l'équilibre entre l'intérêt général et la protection des libertés individuelles fondamentales »<sup>36</sup>*

Analyser des données corrélées est un processus très utile, mais ce dernier ne doit pas aller à l'encontre de toute la dynamique actuelle de protection des données à caractère personnel. Si l'objectif à travers l'analyse de données corrélées affiché semble louable, ce processus soulève de nombreuses interrogations au regard des libertés individuelles.

Le droit de l'UE s'oppose à ce qu'une réglementation nationale prévoyant, à des fins de lutte contre la criminalité, une « conservation généralisée et indifférenciée de l'ensemble des données de trafic et de localisation de tous les abonnés et utilisateurs inscrits concernant tous les moyens de communication »<sup>37</sup>. D'où l'intérêt accru de définir ce qui est personnel de ce qui ne l'est pas, et par conséquent faciliter aux algorithmes le recours légal à l'analyse de données corrélées.

Toute donnée ne saurait être qualifiée de « personnelle ». "Toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement, par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres"<sup>38</sup>, est dite donnée personnelle. En outre, certaines données sont qualifiées de sensibles<sup>39</sup> et font l'objet d'un régime plus strict encore. L'utilisation de ces dernières est interdite à l'exception du consentement exprès des personnes concernées ou si justifiée par un intérêt public.

Il est à noter que l'immense majorité des données produites par les individus apparaît inoffensive, insignifiante, et ne semblerait pas justifier a priori une intégration dans le régime de la donnée personnelle, force est de constater que les algorithmes d'inférence

<sup>36</sup> Mme Guimberteau, avocate affiliée au cabinet FTPA.

<sup>37</sup> Cour de justice de l'Union Européenne, dans son arrêt du 21 décembre 2016 (aff. C-203/15) affirme bien que les réglementations nationales ne peuvent prévoir qu'une conservation ciblée des données de connexion.

<sup>38</sup> L'article 2 de la loi française informatique et des libertés du 6 janvier 1978.

<sup>39</sup> Ces données sont définies aux articles 8 de la loi du 6 janvier 1978 et 9 §1 du RGPD

permettent aujourd'hui d'extraire des caractéristiques individuelles de ces traces anonymes qui peuvent servir à leur tour à identifier et contrôler les individus.

La transformation de la trace en traces anonymes est un mécanisme qui permet de préserver cet équilibre tant convoité entre intérêt général et vie privée. La transformation de la trace s'opère au travers de multiples croisements d'informations.

Enfin, notons que l'utilisateur doit être mis au courant si tout responsable de traitement étudie ses données, d'ailleurs il est mentionné dans le RGPD que toute utilisation des données doit être précédée d'un consentement actif de l'individu. De plus, même si sémantiquement parlant, une « analyse de données corrélées », veut dire qu'il ne faut pas cibler une certaine catégorie de données, le responsable du traitement (institution publique ou autre) ne peut collecter que les données dont l'enquête a besoin. C'est-à-dire qu'en plus d'informer, il faut cibler les informations. Mais en analysant la situation actuelle et avec tous les algorithmes existants, force est de constater que ce mécanisme de ciblage n'est en l'état pas applicable.

La mise en place d'un arsenal juridique assez coriace et cohérent, est donc incontournable pour assurer un contrôle technique à priori. D'ailleurs le règlement européen du 27 avril 2016 et la loi pour une République numérique sont des exemples récents parfaits, qui viennent étendre les dispositions de la loi informatique et liberté. Ce cadre juridique doit cependant être analysé au regard des spécificités des traitements prédictifs et de l'analyse des données corrélées.

## **2- LE MONDE JURIDIQUE ET L'INTELLIGENCE ARTIFICIELLE (MACHINE LEARNING)**

L'intelligence artificielle commence à se faire connaître aussi bien à l'échelle internationale qu'à l'échelle nationale. Il est à noter que l'intelligence artificielle n'est pas pour le moment encadrée juridiquement, ou pour le moins il n'est pas fréquent de trouver des lois ou codes dédiés qu'à l'intelligence artificielle. L'utilisation de colossaux flux de données est totalement fluide et non régulée de manière « effective ». C'est aussi le cas pour le Machine Learning et le Deep Learning qui sont à ce jour méconnus par les législateurs. Plusieurs enjeux juridiques dans le domaine de l'intelligence artificielle et du Machine Learning existent et il serait temps de les prendre en considération.

## **2-1- L'INTELLIGENCE ARTIFICIELLE UN OUTIL EFFICACE POUR LE CHANGEMENT DU TRAVAIL JURIDIQUE ET JUDICIAIRE**

L'intelligence artificielle commence à avoir un réel impact sur les décisions de justice et sur les nouveaux amendements. Mais il est aussi essentiel que l'intelligence artificielle soit au service du pouvoir juridique et judiciaire<sup>40</sup>. Cette dernière (IA) devrait soutenir le travail des juristes et des tribunaux en garantissant une meilleure qualité de la justice et en respectant les principes juridiques fondamentaux. Les calculs algorithmiques doivent être effectués à partir d'une base de données garantissant une meilleure qualité de la justice, tout en respectant les jalons du Droit<sup>41</sup>.

Tout traitement de données juridiques et judiciaires passe en instaurant des systèmes dits d'intelligence artificielle ou des méthodes dérivées des sciences des données améliorant le fonctionnement de la justice, sa transparence et l'application du droit et la cohérence de la jurisprudence.

Insérer l'intelligence artificielle et le Machine Learning dans le fonctionnement législatif et judiciaire ne permettra que leur avancement. Le but étant de développer une réflexion d'ensemble, en adoptant petit à petit de nouveaux outils. Il est fondamental de prendre le soin d'insérer ses derniers, en prenant en considération les divers inconvénients de l'IA.

L'influence des algorithmes sur le droit est telle que la plupart des juristes sont (ou seront bientôt) touchés et donc concernés par les nouvelles « lois » produites par les traitements automatisés de données<sup>42</sup>.

## **2-2 LA SECURITE JURIDIQUE ET L'AIDE A LA PRISE DE DÉCISION**

La sécurité juridique est l'une des premières attentes des divers acteurs, d'une part la sécurité juridique des justiciables et ensuite de tout acteur agissant de manière directe ou circonstancielle dans la prise de décision.

La rationalisation des décisions judiciaires permet une meilleure uniformisation et par conséquent l'obtention d'une meilleure sécurité juridique. Selon J. Boulouis« *La formule*

---

<sup>40</sup> Discours du ministre de la Justice de Lettonie et le Président de la CEPEJ.

<sup>41</sup> Sacha Gaillard, L'INTELLIGENCE ARTIFICIELLE ET L'EXERCICE DU DROIT, 25 mars 2019, <https://www.village-justice.com/articles/intelligence-artificielle-exercice-droit,31053.html>

<sup>42</sup> Boris Barraud, Le droit en datas : comment l'intelligence artificielle redessine le monde juridique, Revue Lamy droit de l'immatériel, p.1, 2019. <https://hal.archives-ouvertes.fr/hal-02445023/document>

*sonne en effet comme une sorte de redondance, tant il paraît évident qu'un droit qui n'assurerait pas la sécurité des relations qu'il régit cesserait d'en être un. Imagine-t-on un droit qui organiserait l'insécurité, ou même qui la rendrait possible ?* ». La sécurité juridique est une partie intégrante du droit, celle-ci sonne parfaitement comme une tautologie<sup>43</sup>.

Chaque juge s'adapte à chaque situation, tout législateur s'adapte s'accommoder aux conjonctures, par conséquent il y a divers présupposés qui ont une incidence volontaire ou involontaire sur la prise de décision (l'influence médiatique, les préjugés, les circonstances et pressions à l'échelle internationale, le respect des conventions, le témoignage de diverses parties liées à un procès, l'égoïsme...). La variation des décisions ne fait que compliquer la tâche de l'introduction l'IA dans ce processus juridique et judiciaire gouvernés par l'aléa, la collégialité et une multitude de jurisprudences.

L'insertion de l'IA dans le système judiciaire et juridique permettra d'englober les attentes des citoyens, justiciables, juges, législateur et gouvernement. Elle permettra d'avoir une justice garantissant l'identification de solution partielle gouvernée par la fiabilité, légalité et la certitude.

## CONCLUSION

Il ne fait aucun doute que l'intelligence artificielle et le Machine Learning sont en train de changer le paradigme de protection des données à caractère personnel. Tous les niveaux sont en train d'être bouleversés par ces technologies. L'adoption de plus en plus répandue de l'IA et du ML dans divers secteurs semble être davantage une évolution à long terme. Ces deux technologies sont en perpétuelle évolution. Leur utilisation nécessite de s'engager dans une démarche constructive, en se basant non plus sur l'intelligence artificielle mais plutôt en se projetant sur ce qu'on peut appeler une "intelligence collective" (ce qui permettra de préserver l'équilibre entre intérêt général et intérêt personnel).

Il est extrêmement important de garantir une transparence totale dans ce processus d'analyse des données, mais cette garantie s'avère difficile compte tenu des nouveaux

---

<sup>43</sup> J. Boulouis, « Quelques observations à propos de la sécurité juridique », in *Du droit international au droit de l'intégration. Liber amicorum : Pierre Pescatore*, NomosVerlag, 1987, p. 53, cité par J.-G. Huglo, Dossier : Le principe de sécurité juridique, *Cah. Cons. const.* 2001, n° 11.

enjeux. Pour répondre au manque de transparence des processus, l'intérêt doit être porté sur le côté technique et ce en l'examinant en détail.

En effet c'est en posant un regard critique sur les liens entre différents responsables du traitement des données à caractère personnel, personnes concernées, structures publiques, règlements et lois existantes, divers algorithmes, les modalités d'implémentation de ces derniers et la variété de ses utilisations, qu'une appréhension des conséquences sociales, éthiques et juridiques pourra émerger.

Ce qui permettra à l'avenir l'utilisation d'un processus algorithmique capable de démontrer la légalité, l'équité et la transparence d'une décision prise par un humain ou une machine. Cet objectif ne pourra être atteint que si les exigences légales en matière de protection des données fluctuent.

**BIBLIOGRAPHIE**

**Antonio Casilli (6 février 2015)**, QUELLE PROTECTION DE LA VIE PRIVÉE FACE AUX ATTAQUES CONTRE NOS LIBERTÉS NUMÉRIQUES ? Colloque La France dans la transformation numérique : quelle protection des droits fondamentaux ?, Conseil d'État, Paris.

**Aurélien Bamdé (2018)**, “RGPD: le principe de finalité”, Le Droit dans tous ses états,

**Aurélien Bamdé (2018)**, le principe de proportionnalité, Droit des personnes, Loi informatique et libertés, RGPD.

**Antoinette Rouvroy (11 janvier 2016)**, « des données et des hommes » droits et libertés fondamentaux dans un monde de données massives. Bureau du comité consultatif de la convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel. Direction Générale Droits de l'Homme et Etat de droit. Strasbourg.

**Benjamin Boulvert (04.05.2015)**, “LES ENJEUX DE L'ANALYSE PRÉDICTIVE DANS LE MÉTIER DE L'ASSURANCE”, <https://www.insurancespeaker-wavestone.com/2015/05/les-enjeux-de-lanalyse-predictive-dans-le-metier-de-lassurance/>

**Big Browser (12 juillet 2017)**, “COMMENT FAIRE DES STATISTIQUES COMME GOOGLE ?” JurilexBlog <https://www.haas-avocats.com/actualite-juridique/comment-faire-des-statistiques-comme-google/>

**Big Browser (12 juillet 2017)**, “Ce que l'analyse des recherches sur Google nous apprend sur la psyché humaine”, Le monde, 12 juillet 2017, [https://www.lemonde.fr/big-browser/article/2017/07/11/ce-que-l-analyse-des-recherches-sur-google-nous-apprend-sur-la-psyche-humaine\\_5159228\\_4832693.html](https://www.lemonde.fr/big-browser/article/2017/07/11/ce-que-l-analyse-des-recherches-sur-google-nous-apprend-sur-la-psyche-humaine_5159228_4832693.html)

**Boris Barraud (2019)**, “L'intelligence de l'intelligence artificielle. dans toutes ses dimensions”, L'Harmattan,

**Boris Barraud (2019)**, “Le droit en datas : comment l'intelligence artificielle redessine le monde juridique”, Revue Lamy droit de l'immatériel

**Cécile de Terwangne (2018)**, “Les principes relatifs au traitement des données à caractère personnel et à sa licéité”

**Frédéric SUR (2020-202)**, “Introduction à l'apprentissage automatique”, École des Mines de Nancy,

**J. Boulouis (2001)**, « Quelques observations à propos de la sécurité juridique », in Du droit international au droit de l'intégration. Liber amicorum : Pierre Pescatore, NomosVerlag, 1987, p. 53, cité par J.-G. Huglo, Dossier : Le principe de sécurité juridique, Cah. Cons. const, n° 11.



**Hildebrandt, Mireille. (2008),**DefiningProfiling: A New Type of Knowledge?. 10.1007/978-1-4020-6914-7\_2

**Lexing Alain Bensoussan (29/10/2020),** Action de groupe et protection des données personnelles. <https://www.alain-bensoussan.com/avocats/action-groupe-protection-donnees/2016/12/22/>

**Mathias le Masne de Chermont, Adrien Aulas, Hugo Ruggieri (22 janvier 2018),** “Protection par défaut et minimisation des données...quelles différences?” , <https://aeonlaw.eu/protection-par-defaut-minimisation-des-donnees-quelles-differences/>

**Meiko Jensen (Université de Kiel), Cédric Lauradoux (INRIA), Konstantinos Limniotis (HDPa) (NOVEMBRE 2019),** “TECHNIQUES ET MEILLEURES PRATIQUES DE PSEUDONYMISATION”, Recommandations sur l’usage des technologies conformément aux dispositions en matière de protection des données et de respect de la vie privée, Agence de l’Union européenne pour la cybersécurité (ENISA), [https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices\\_fr](https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices_fr)

**Sacha Gaillard (25 mars 2019),** “L’INTELLIGENCE ARTIFICIELLE ET L’EXERCICE DU DROIT”, <https://www.village-justice.com/articles/intelligence-artificielle-exercice-droit,31053.html>

**Stéphanie Faber (Janvier 2018),** “Lignes directrices du G29 sur les « décisions individuelles automatisées » et le « profilage »”, Protection des données personnelles – RGPD, La Revue Hors-série.

**SynthiaTientcheuTcheuko (2020),** “L’action de groupe : arme efficace contre l’utilisation abusive des données personnelles ? “

**Vayre Jean-Sébastien (2018),** « Les machines apprenantes et la (re)production de la société : les enjeux communicationnels de la socialisation algorithmique », Les Enjeux de l'information et de la communication, 2018/2 (N° 19/2).

**Wachter, Sandra &Mittelstadt, Brent &Floridi, Luciano (2016),** “Why a Right to Explanation of AutomatedDecision-MakingDoes Not Exist in the General Data Protection Regulation”, DOI: 10.1093/idpl/ipx005,

**Le 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 dit règlement général sur la protection des données (RGPD).