2022

RAPPORT
« Etude et perspectives de la conférence sur l'intelligence artificielle et la vie privée »

DUMONT TIPHAINE, HERBOCH BENJAMIN, GILLES CAMILLE, DUPIRE LUCIE

MASTER DROIT DU NUMÉRIQUE - Sociologie du numérique

Depuis 2017, la CNIL appelle à la vigilance concernant les évolutions des outils de vidéoprotection. En effet, la multiplication de certains dispositifs de vidéo « augmentée » pose des questions éthiques et juridiques nouvelles. Aujourd'hui, afin d'accompagner leur déploiement dans le respect des droits des personnes, la CNIL soumet un projet de proposition à consultation publique jusqu'au 11 mars 2022. De tels dispositifs ne sont en aucun cas un simple « prolongement » technique des caméras existantes. Ils modifient leur nature même par leur capacité de détection et d'analyse automatisée.

L'intelligence artificielle (IA) désigne tout système capable d'accomplir des tâches d'une manière que nous percevons comme "intelligente". Autrement dit, il s'agit de l'ensemble de techniques, théories et sciences qui simulent les capacités cognitives de l'être humain[1]. Ainsi, l'abstraction, la créativité, la déduction, la résolution de problèmes, la prise de décision ou la capacité d'apprendre, sont associées à l'IA. Une différence doit notamment être faite entre IA dite "forte" et IA dite « faible ». L'IA « forte » serait en capacité de contextualiser des problèmes spécialisés très différents de manière totalement autonome.

Toutefois, selon le Conseil de l'Europe, aucune technologie connue ne permet de démontrer l'existence actuelle d'IA « forte ». Par opposition, l'IA « faible » ou "modérée" n'est en capacité d'être performante que dans son domaine d'entraînement. Quant au Machine Learning, il représente un ensemble de méthodes puissantes qui permettent de créer des modèles prédictifs à partir de données sans avoir été explicitement programmés. Ce type d'intelligence artificielle est mis en cause aujourd'hui.

Se pose alors la question des conséquences d'une généralisation hâtive de l'Intelligence artificielle. Le chercheur postdoctoral Pégny Maël[2], lors de deux matinées d'études organisées par l'Université Paris-Est Créteil, a évoqué les problématiques pouvant exister entre droit au respect de la vie privée et développement des IAs. En effet, le développement de l'Intelligence artificielle a apporté de nombreuses incertitudes quant à son développement éthique en raison du renforcement des inégalités et de l'atteinte à la vie privée que pourrait engendrer son emploi dans différents secteurs tels que notamment la santé, l'assurance ou les services publiques.

Face aux différents enjeux exposés et à la nécessité d'identifier les différents enjeux et ordres juridiques présents en la matière, il est aujourd'hui essentiel de traiter de la notion d'IA et de vie privée.

Dans ce cadre, une exposition liminaire du point de vue de Maël Pégny (I) paraît essentielle afin de fixer le cadre de réflexion. Ainsi, ce cadre permettra dans un second temps d'identifier la nécessité d'une analyse poussée face à l'ère du développement de l'IA (II) afin de les envisager à l'aune des propositions actuelles (III).

I. Une appréhension globalisée du droit au respect de la vie privée et développement des IAs.

Au sein de sa restitution et après avoir exposé le fait que la vie privée est une donnée entretenant des liens importants avec les institutions étatiques, Maël Pégny évoque les enjeux de la collecte de données personnelles, à savoir les questions relatives à la collecte et la formation de connaissances (réelles ou prétendues) sur les individus. L'explosion de la collecte induit l'automatisation de l'analyse. On définit l'intelligence artificielle par son objectif : créer des systèmes informatiques intelligents en exécutant des tâches communément considérées comme intelligentes.

Une fois ces définitions posées, Maël Pégny présente les problématiques relatives à « L'âge d'or de la surveillance[3]" : la surveillance de masse en passant par le capitalisme de la surveillance pour finir par la propagande de masse et le profilage des individus. Enfin ont été abordées les notions de mort de l'anonymat en ligne et les enjeux qu'elle soulève.

S'appliquent à l'IA des prises de position fortes pour interdire certaines pratiques économiques manipulatoires, exploitatrices ou visant au contrôle social. L'IA fait également face à une exigence de transparence, en informant sur l'existence d'une interaction avec un robot ou sur la détection d'émotion. Maël Pégny soulève la question suivante : pourquoi ces restrictions s'appliquent- elles seulement à des systèmes qualifiés d'IA ?

La reconnaissance faciale est l'exemple type d'intelligence artificielle. Elle regroupe deux formes d'outils : la reconnaissance à partir de photographies et la reconnaissance en direct à partir d'images de caméras de vidéo de

[1] Définition donnée par le conseil de l'Europe :https://www.coe.int/fr/web/artificial-intelligence/history-of-ai

[2] Monsieur Pégny Maël est chercheur postdoctoral en Ethique & IA à l'Université de Tübingen

[3] Il s'agit d'une expression employée de Bruce Schneier

surveillance dans l'espace public (FRT). C'est cette seconde forme qui focalise le débat. En effet, elle marque une rupture juridique en appliquant à l'ensemble de la population des pratiques d'identification réservées aux suspects et criminels avérés, et une rupture symbolique : c'est la fin de l'anonymat dans l'espace public. Ainsi, il ne faut pas surestimer l'efficacité de la FRT, qui connaît énormément de dysfonctionnements avec un taux de réussite de 0% sur des tests menés en 2016. Cependant, cela pose un problème de fond majeur : que se passerait-il si la technique fonctionnait bien ? Deux écoles de pensée s'opposent. D'une part, l'école réglementaire qui défend un usage très limité de la reconnaissance faciale dans les cas de terrorisme ou de disparition d'enfant par exemple. D'autre part, l'école prohibitionniste qui défend le fait qu'un tel dispositif ne devrait pas exister. En outre, le développement de ce type d'IA pose deux questions essentielles : la question du développement de la technologie et notamment des bases de données d'entraînement qui doivent être créées, et la question de l'usage de cette technologie et surtout du contrôle institutionnel.

On se rend compte de la nécessité de la création de vastes bases de données collectées dans des conditions naturelles pour l'entraînement des modèles. L'approche d'usage réglementé entérine la collecte de données à l'échelle d'une population sans consentement. Maël Pégny alerte quant à la perte de sensibilité contemporaine face à cette collecte et se pose la question de la fidélité à l'esprit originel du droit des données personnelles.

Le développement de la FRT constitue en soi un tournant juridique et le problème du contrôle institutionnel est une question cruciale. Est-on capable de borner l'emploi d'un outil de contrôle social ? La question peut raisonnablement être soulevée au regard de certaines pratiques inquiétantes dans les polices de certains pays démocratiques, notamment aux Etats-Unis qui sont source d'inquiétude comme les identifications de manifestants "Black Lives Matter", ou le programme FBI "COINTELPRO" dont l'objectif fut d'enquêter sur les organisations politiques dissidentes aux Etats-Unis et de perturber leurs activités. Dans la Résolution du Parlement européen du 20 octobre 2020 concernant un cadre pour les aspects éthiques de l'intelligence artificielle, de la robotique et des technologies connexes (2020/2012(INL)), est interdite l'identification biométrique à distance dans l'espace public en temps réel à des fins répressives. Cependant, au regard du nombre d'exceptions, cela ressemble davantage à une autorisation plutôt qu'à une prohibition puisque sont listées toutes les infractions de droit commun.

Toutes ces questions ont un impact conceptuel sur notre ordre légal et notamment sur les données personnelles. C'est une notion extrêmement large et sémantique. L'IA a pour conséquence de déduire des choses des données qu'on lui transmet : "Vous ne pouvez protéger ce que je peux déduire". Ce n'est pas la donnée qu'il faut protéger, mais la connaissance qu'elle contient. Dans ce cadre, Maël Pégny alerte sur la nécessité d'une régulation de l'inférence statistique[4] et notamment sur la validité scientifique des inférences (droit à l'inférence raisonnable) et de la protection de la vie privée contre le contournement statistique. Aujourd'hui, il est possible de procéder à des inférences d'appartenance (déduction de l'appartenance d'un sujet à la base d'entrainement).

Cependant, le problème prépondérant est que le droit des données de l'UE est centré sur les données avant traitement. Une donnée est personnelle si elle permet d'identifier une personne physique. Il y a une nécessité de clarifier le rapport entre traitement et statut de donnée personnelle. Ces problématiques sont mises en lumière par le Groupe de travail "article 29"[5], sur la protection des données. Trois critères des données personnelles peuvent être retenus : contenu, intention, résultat. Avec l'exploitation des données environnementales pour influencer les personnes dans les smart cities[6], tout peut devenir une donnée personnelle par intention ou par résultat.

Enfin, Maël Pégny nous rappelle qu'il ne faut pas penser qu'en termes de barrières mais il s'agit plutôt de se poser la question suivante : comment réorienter l'appareil de connaissance numérique des individus vers des fins saines ? Quelle reconnaissance par les institutions les individus souhaitent-ils ?

En effet, la reconnaissance des individus par les institutions peut être désirée. Ce fut le cas notamment des femmes amérindiennes qui habitaient sur les réserves. Une base centralisée est la première chose qui a été demandée pour leur permettre de jouir des mêmes droits que les femmes américaines, notamment en ce qui concerne les violences sexuelles.

---

[4] L'inférence statistique consiste à induire les caractéristiques inconnues d'une population à partir d'un échantillon issu de cette population.

[5] Le G29 ou Groupe de travail Article 29 sur la protection des données est un ancien organe consultatif de l'Union européenne indépendant sur la protection des données et de la vie privée

[6] Une ville intelligente est une ville utilisant les technologies de l'information et de la communication pour améliorer la qualité des services urbains ou réduire leurs coûts

II. La nécessité d'une analyse poussée face à l'ère du développement de l'IA

A. Une prise en compte de l'éthique dès la conception

1. La nécessité de penser éthiquement l'usage de données personnelles dès la conception

Pour Maël Pégny, l'éthique s'envisage en amont de la conception notamment de l'IA pour y inclure les possibles dangers politiques qu'elle pourrait représenter de par la grande collection des données et leur centralisation ainsi que leur exploitation, qui est le point le plus problématique pour la vie privée. Il faut en outre prendre la mesure de ce qu'est une donnée au travers de son accès, sa qualité et notamment son intégrité.

Force est de constater cependant que la résolution du problème éthique n'en est qu'à ses balbutiements. En témoigne le rapport de la commission Ethique du Ministère fédéral allemand des transports sur les véhicules autonomes et connectés, qui met en exergue l'impossibilité pour l'éthique et le juridique d'être inclus directement dans la programmation, qu'il faut cependant une approche de développement éthique qui parcourt l'ensemble du cycle de vie du logiciel afin d'en améliorer la conception.

Cette idée pose cependant plusieurs problèmes notamment dans l'idée du développement éthique par essence, puisque les développeur.euses devront obtenir une expertise économique, juridique, sociologique et philosophique ce qui est une attente irréaliste pour Maël Pégny. Une acception plus réaliste serait une délimitation des difficultés de l'opérationnalisation, c'est-à-dire l'absence de limites à la création de l'IA qui est stimulée par la concurrence et le sensationnalisme. Cependant, le numérique et notamment l'IA ne peut plus être pensé uniquement sous un aspect technique. De par l'usage dont le numérique fait l'objet, il est nécessaire d'envisager l'éthique dès la conception du service ou du produit. L'intervenant Maël Pégny avait proposé une charte afin de s'assurer que le produit ou le service était éthique depuis sa conception.

En effet, il évoque dans son écrit *Pour un développement des IAs respectueux de la vie privée dès la conception*[7] le fait de :

- déclarer la finalité de l'usage des données en documentant et justifiant tout écart à la déclaration initiale,
- tester les performances finales du logiciel et limiter le pouvoir prédictif du modèle à ce qui est nécessaire,
- mettre en balance risque de suroptimisation et de perte de performances en prenant en compte la vie privée,
- entraîner si possible son modèle sans avoir recours aux données personnelles, ou d'utiliser des données ayant fait l'objet d'un geste explicite de publication, la mise à jour des données, retrait de publication et l'exercice des droits de rectification, d'effacement et d'opposition,
- déclarer les raisons justifiant l'utilisation des données personnelles ainsi que les mesures prises pour lutter contre la rétro-ingénierie des données, et prendre position sur leur complétude par rapport aux boîtes à outils et aux méthodes d'attaque existantes,
- diffuser en licence libre tous les outils de sécurisation contre la rétro-ingénierie des données,
- mettre le modèle à disposition de tous afin de permettre la vérification publique des propriétés de sécurité si cela n'entraîne pas de faille de sécurité intolérable et le justifier,
- lorsque les mesures de restrictions de la collecte et de lutte contre la rétroingénierie ne sont pas applicables et que la gravité de l'enjeu surpasse les enjeux de vie privée,
- autoriser un modèle encodant des données privées en restreignant strictement l'accès à ce modèle et son emploi pour l'usage ayant justifié l'exception.

[7] Maël Pégny. Pour un développement des IAs respectueux de la vie privée dès la conception. 2021. ffhal-03104692

Cependant, en l'état actuel du droit, la législation a apporté un lot non négligeable de sécurisation et d'éthique du traitement des données qui constitue un commencement à cet idéal de protection qui se dessine au travers de la charte telle que conçue par Maël Pégny.

2. L'état du droit

En effet, le droit de l'Union Européenne envisage la protection des données comme des droits et se réfère aux articles 7 et 8 de la Charte des droits fondamentaux de l'Union, à l'article 16, § 1er, TFUE, ainsi qu'à l'article 8 de la Convention européenne des droits de l'homme et dans la Convention n° 108 du Conseil de l'Europe, le RGPD et le règlement « Vie privée et communications électroniques » de la Communication de 2017 de la Commission européenne remplaçant la directive 2002/58/CE. Cet ensemble de textes législatifs encadrent et protègent la protection de la vie privée et sanctionnent sévèrement ceux qui commettraient un manquement. Par ailleurs, le Règlement Général sur la Protection des Données s'apparente à la Charte qu'imagine Maël Pégny. En effet, ce règlement impose de justifier de la finalité du traitement, la possibilité de retirer son consentement, corriger les données, sécuriser les données par des moyens techniques. Le règlement met également en place un principe de minimisation.

Le principe de minimisation selon la Cnil « *prévoit que les données à caractère personnel doivent être adéquates, pertinentes **et** limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées.* » Maël Pégny dans sa conférence et son écrit émet la critique selon laquelle la finalité peut être ajustée afin d'obtenir de la façon la plus extensive possible les données personnelles, que les libertés accordées à la recherche par exemple sont modifiées au fil de l'eau. A l'occasion d'un traitement, il convient de définir la finalité du traitement, puisque c'est essentiel dans l'exploitation des données personnelles et leur protection. Cette finalité peut en effet être difficile à délimiter si elle concerne la recherche scientifique, celle-ci pouvant être élargie ou précisée.

L'exigence de finalité seule n'est pas suffisante. Cependant force est de constater que pareillement à la Charte de Maël Pégny, le RGPD soumet tout traitement de données à caractère personnel, incluant la recherche scientifique et exige la documentation, la justification ainsi que le consentement du propriétaire des données avant toute modification du traitement ou de sa finalité. Pour la recherche scientifique en matière de santé, certaines dérogations existent. C'est notamment le cas au sujet de la définition de la finalité de ce traitement. Cette définition peut être plus générale s'il n'existe pas de façon plus précise de la délimiter, mais doit demeurer. La CNIL contrôle le traitement des données de façon stricte mais cela n'est évidemment pas le cas de tous les pays, ce qui pose un problème de territorialité des règlements. La limite de la législation se situe au-delà des frontières physiques des pouvoirs législatifs et également au début du domaine plus technique du numérique, notamment dans les algorithmes, que le droit n'a pas encore saisi. Le traitement automatisé ou non automatisé avec ou sans des outils logiciels est tombé sous le coup de la loi. En revanche, il est essentiel de sensibiliser aux questions éthiques dès la conception des produits et services. En effet, une éthique en amont permet aux législateurs, en collaboration avec les experts de la technique numérique, d'élaborer une législation au plus proche des besoins éthiques des outils numériques dans le traitement des données personnelles ainsi que d'étendre la territorialité de la protection des données.

Dès lors, l'élaboration de chartes de la part des experts serait une collaboration avec le législateur.

3. L'objet d'une charte développement éthique

La charte telle qu'imaginée par Maël Pégny est une éthique du développement et de son impact. En l'occurrence, cette charte se spécialise sur la protection des données personnelles et de la vie privée, plus précisément la vie privée par la collecte massive de données nécessaire au développement, mais aussi par l'enregistrement, et la possible récupération de données à partir des modèles. Selon Maël Pégny, l'utilisation d'une charte est d'autant plus importante dans les algorithmes et il vise plus spécifiquement les modèles d'apprentissage automatique, qui seront selon lui un « *point de tension entre le désir et le respect de la vie privée et les idéaux de diffusion libre du logiciel et de reproductibilité de la recherche* ». Selon lui, la valeur de l'information privée provient de son caractère restreint, et sa diffusion constituerait une « dégradation et non une ouverture à un épanouissement collectif ». Dès lors, elle se doit de rester privée autant que possible.

L'utilisation d'une charte telle que celle de Maël Pégny serait également une prise de conscience, un engagement à titre personnel de conformité avec l'état actuel du droit, les enjeux éthiques et les questions de sécurité. La charte est un moyen de communication des principes éthiques sur lequel les modèles sont créés. C'est un début de stratégie, une politique d'orientation dans la façon d'utiliser les données personnelles et une collaboration avec le législateur qui n'a pas l'expertise de saisir les enjeux dans les domaines techniques numériques mais qui pourrait se saisir des chartes qui prendraient en mesure l'éthique et ainsi s'approprier ces valeurs pour les traduire en règles contraignantes et impératives.

B. L'éthique de la surveillance

1. La surveillance étatique

Selon Bruce Schneider, nous sommes dans un "*âge d'or de la surveillance"*, les institutions n'ayant jamais eu autant de données sur les individus ni de moyens de surveillance continue. Cette surveillance n'est pas uniquement marketing comme le souligne S. Zuboff en parlant de capitalisme de la surveillance. En ce sens, la Commission européenne dans une note blanche du 19 février 2020 *Une approche européenne axée sur l'excellence et la confiance* met en garde contre « la mise en place de barrières et d'exigences lourdes qui peuvent constituer un obstacle à l'innovation ». Au lieu de cela, ces 14 ministères suggèrent que l'Europe « devrait se tourner vers des solutions juridiques non contraignantes telles que l'autorégulation, la labellisation volontaire et d'autres pratiques » de ce type. L'anonymat tend également à décroître du fait de l'archivisme et de la capacité d'identifier par croisement de données massives, de la surveillance dans l'espace public. Il y a deux écoles par rapport à cette surveillance. La première est celle de la réglementation que décrit l'ouvrage *The partial Line* qui défend un usage limité de la surveillance pour le terrorisme par exemple ou des disparitions, position que choisit l'UE. La seconde, est celle de la prohibition, c'est l'idée que cela ne devrait pas exister.

La surveillance est encadrée par l'Union Européenne au travers de textes législatifs, que conteste Maël Pégny. L'Union Européenne afficherait une ambition de donner une définition des applications de l'IA à haut risque (santé, sécurité, droits fondamentaux) et l'identification de caractéristiques de la technologie (opacité, complexité, dépendance aux données, comportement autonome). Selon lui, on cherche à créer des bacs à sable pour tester des règlements.

L'Union européenne interdit la mise en place de « techniques subliminales agissant sur l'inconscient », l'exploitation de vulnérabilités de groupes vulnérables spécifiques (handicapés, enfants), pouvant causer un préjudice psychologique ou physique et enfin la prise en compte de certains excès des autorités publiques (sauves usages militaires, coopération internationale de services): notation sociale fondée sur l'IA à des fins générales et exige une transparences sur les interactions avec un robot, détection d'émotions, associations avec des catégories sociales sur la base de données biométriques, trucages vidéo ultra réalistes. Cette surveillance embarque également le développement de LFRT, un impact conceptuel sur la notion de DCP, la fin du paradigme de protection par restriction d'accès, remet en cause la distinction donnée-programme/algorithme, de se pencher sur le rapport traitement et données.

L'Union Européenne et notamment la Cour Européenne des Droits de l'Homme ainsi que la CJUE se sont pourtant emparées de ces questions, avec opposition entre ces deux institutions, la Cour de Justice étant plus protectrice en l'occurrence des droits fondamentaux. Elles se reposent sur l'article 8 de la Convention Européenne de Sauvegarde des Droits de l'Homme qui protège la vie privée. Il est cependant apparu que les institutions nationales et européennes sont assez permissives au niveau de la surveillance de masse sous couvert de protection antiterroriste en témoigne un arrêt rendu par la CJUE Télé 2 du 21 décembre 2014 qui détermine la nécessité d'encadrer strictement la conservation des données et leur accès, et détermine le fait d'organiser une conservation générale des données comme incompatibles avec les articles 7 et 8 de la Convention de Sauvegarde des Droits de l'Homme. Il convient également de permettre plus de transparence sur les algorithmes, droit mis en place par le RGPD mais très rarement exercé, et les administrations sont réticentes à donner l'algorithme utilisé en témoigne les scandales autour de la mise en place de Parcoursup.

2. Les algorithmes discriminatoires

La raison pour laquelle les algorithmes peuvent être discriminatoires est l'utilisation de données biaisées. Le biais le plus fréquent serait le manque de représentativité des données mobilisées dans un système d'apprentissage automatique si celui-ci est la traduction de pratiques et comportements. Par exemple, dans les données d'emploi les femmes sont moins représentées et occupent des filières de métiers ainsi que des postes à rémunérations moindres. L'algorithme pourrait en déduire que les femmes sont moins productives que les hommes et ne vont pas vers des postes à responsabilité et accentuerait le recrutement des hommes sur cette base. Quand bien même des caractéristiques neutres en apparence seraient utilisées, cela peut avoir des effets discriminatoires comme l'a souligné le Défenseur des droits dans sa décision Parcoursup qui a mis en relief que l'utilisation de la donnée de l'établissement d'origine compromettait la mixité sociale et l'ascension des jeunes défavorisés.

Les algorithmes avec des biais pourraient également remettre en cause la liberté d'expression. Maël Pégny prend en exemple la loi visant à lutter contre les contenus haineux sur internet adoptée par l'Assemblée Nationale le 13 mai 2020 destinée à retirer certains contenus haineux sous 24 h des réseaux sociaux, des plates-formes collaboratives et des moteurs de recherche. L'utilisation de procédés algorithmiques sur la base de mots clés pourraient entraîner une censure excessive et devenir contreproductive.

On peut concevoir des débuts de solutions en soutenant la recherche afin de développer les études de mesures ainsi que des méthodologies de prévention des biais. Cumulativement, il serait opportun de réaliser des études d'impacts afin d'anticiper les effets discriminatoires des algorithmes, et cela a par ailleurs été initié dans un tableau d'exemples théorique de l'impact de l'IA et des décisions automatisées *Examples of theoretical assessment of harm and significant impact of AI or automated decisions* de 2020. Des organisations et institutions commencent à s'intéresser à la prévention des discriminations dans les algorithmes, en témoignent les travaux de l'institut international ombudsman en partenariat avec la CNIL : *Algorithme : prévenir l'automatisation des discriminations, qui a eu lieu pendant la crise sanitaire.* Il faut également citer le rapport de l'Agence des droits fondamentaux de l'Union européenne qui ouvre la discussion de l'éthique et des droits fondamentaux *Bien préparer l'avenir : l'intelligence artificielle et les droits fondamentaux* du 14 décembre 2020.

C. Les approches juridiques et philosophiques

1. L'inapplicabilité de la dichotomie données-logiciels en matière de modèles d'apprentissage automatique

Conséquence juridique prééminente des problèmes de sécurité des modèles, la dichotomie données-logiciels pose aujourd'hui des problèmes d'éparpillement du droit en matière de données. En effet, l'avènement des modèles d'apprentissage automatique (Mode Machine Learning) ont ceci de particulier que les données sont encodées comme un logiciel et ainsi, sont soumises à un régime différent. Bien que la dichotomie posait déjà problème avant les modèles d'apprentissage automatique dans la mesure où les traitements étaient déjà possiblement invertibles[8], la différence de régime au sein de la législation européenne entre données brutes et données traitées fut renforcée. Du fait du caractère fondamentalement indéterminé et évolutif des concepts juridiques ou comme l'énonce le philosophe H.Hart de la « texture ouverte du droit »[9], le raisonnement de cette science molle autour d'un possible régime juridique en matière d'IA pose de nombreux problèmes face aux ambiguïtés sémantiques et syntaxiques des différents textes juridiques. Face au droit de la protection des données personnelles, les droits de la propriété intellectuelle et le secret commercial s'exercent de façon prépondérante.

[8] Invertible ou inversible vient du mot « invertible » et plus précisément « inverstible matrix ». En mathématiques et plus particulièrement en algèbre linéaire, une matrice inversible est une matrice carrée A pour laquelle il existe une matrice B de même taille n avec laquelle les produits AB et BA sont égaux à la matrice identité.

[9] Charnock, Ross (2005), Jugements lexicologiques: définition lexicale et "texture ouverte" dans l'interprétation juridique, dans Greenstein, Rosalind, La langue, le discours et la culture en anglais du droit, Publications de la Sorbonne : Paris, p. 33-55

Toutefois, la réelle question à se poser est la suivante : quel est l'enjeu de la définition d'un régime juridique sûr et stable ? En effet, l'application distributive de différents régimes juridiques s'exercent dans différents domaines sans poser aujourd'hui de réelles difficultés. Cependant, en matière d'IA, l'enjeu étant de protéger les données notamment les données personnelles, la fixation d'un régime juridique, bien que pouvant évoluer avec le temps, est aujourd'hui primordiale. Il existe des interférences avec la dichotomie données-logiciels comme l'énonce le Professeur Michael Veale[10] au sein de son ouvrage *Algorithms that remember: model inversion attacks and data protection law,*. C'est notamment le cas des modèles susceptibles d'une attaque par inférence d'appartenance[11] ou par inversion portant sur des données pseudonymisées ou soumises à un autre traitement cryptographique. Une telle analogie vient faire exploser la barrière entre données traitées et données brutes. Il convient tout de même de noter que l'état actuel du droit n'est pas étranger à cet éclatement. En effet, les données pseudonymisées ont bien fait l'objet d'un traitement, qui plus est d'un traitement explicitement conçu pour protéger les droits des sujets de données. Toutefois, celles-ci sont soumises au même statut juridique que les données personnelles dont elles sont tirées et constituent déjà une exception au cadre juridique général des données traitées. Cette problématique autour des données personnelles est notamment appréhendée par le Règlement du Parlement européen et du Conseil 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 (ci-après « RGPD ») notamment par la technique de probabilité de ré-identification[12], partie présente de l'identification de la notion de « caractère personnel » des données[13]

Les débats entourant la notion de la patrimonialisation de la donnée[14] sont aussi au cœur de notre sujet. En effet, la notion de la propriété par le droit de la propriété intellectuelle et le secret d'affaire est envisagée sous le spectre d'une propriété privée par le travail et l'investissement financier, en opposition avec le droit de la protection des données, véritable avancée vers un droit fondé sur le rapport à soi. Ainsi reprenant les termes de M. Pégny : « lorsque le droit des données personnelles se voit donner prééminence sur la propriété intellectuelle ou le secret des affaires, l'intuition que la vie privée constitue un droit fondamental écrasant le droit de propriété, ou constituant un droit de propriété privilégié » est encensée. En d'autres termes, « je suis le propriétaire de mes données, même si c'est vous qui travaillez avec et sur elles ».

Outre l'inapplicabilité de la dichotomie données-logiciels, celle des données anonymes/personnelles doit être envisagée.

2. L'inapplicabilité de la dichotomie données personnelles-anonymes en matière de modèles d'apprentissage automatique

Selon le Considérant 26 du RGPD les données anonymes sont les données dépourvues de référence à une personne physique, ou traitées de telle manière que les personnes physiques ne soient plus identifiables. En effet, la difficulté pour toute entreprise est de garantir que les données collectées, ayant fait l'objet d'une anonymisation ne peuvent plus, par recoupement, contribuer à la ré-identification d'une personne physique. La notion d'identification directe employée au sein du Considérant 26 est définie comme « l'identification par le nom propre, éventuellement secondée par une information distinguant les homonymes, tandis que l'identification indirecte est définie par une combinaison unique d'identifiants permettant de singulariser l'individu au sein d'un groupe ». Comme l'énonce le Working Party 29[15] (G29), aujourd'hui remplacé par le Comité Européen sur la Protection des Données[16] lors de l'Opinion 4/2007 On the concept of personal data[17] : « une donnée n'est anonyme que lorsque l'anonymisation est

[10] Michael Veale, Professeur au Departement des sciences, technologies, ingénierie et politique publique de l'Université du Collège de Londres

[11] La confidentialité différentielle vise à ne pas dévoiler vers l'extérieur l'appartenance à la base même d'une entité qu'elle contient.

[12] Mesure pour déterminer si un ensemble de données est personnel ou non dans les dernières interprétations

[13] L'impact du RGPD sur les innovations en matière d'IA, Béatrice DELMAS-LINEL, Grégoire DUMAS – Le Big Data et le Droit - e éd. - Février 2020 (Thèmes et commentaires)

[14] Essai de prospective juridique sur les modes de valorisation des données : fonds informationnel, droit à la portabilité et dividende de la données, Valérie-Laure BENABOU, Professeur à l'Université Aix-Marseille et Paris-Saclay, Le Big Data et le Droit - e éd. - Février 2020 (Thèmes et commentaires)

[15] Le G29 ou Groupe de travail Article 29 sur la protection des données est un ancien organe consultatif de l'Union européenne indépendant sur la protection des données et de la vie privée

[16] Le CEPD est institué par le règlement général sur la protection des données (articles 68 à 76). Il a pris la suite du Groupe de l'article 29 (le G29).

[17] Opinion 4/2007 on the concept of personal data, 01248/07/EN WP 136, adopté le 20 juin 2007, Groupe de travail Article 29 sur la protection des données

irréversible, c'est-à-dire quand il n'est pas possible de retrouver l'identité d'une personne physique à partir de cette donnée avec les moyens existants de la technologie. Cette possibilité de ré-identification peut en outre être comprise dans un sens technique absolu, ou non dans un sens relatif aux moyens à la disposition du contrôleur de données ».

Ainsi, comme l'énonce le Professeur Nadezdha Purtova[18], une grande partie des difficultés éprouvées quant à cette définition provient de sa souplesse mais aussi de l'introduction au fil du temps de modalités supplémentaire en raison des nombreuses évolutions technologiques, continuant à modifier les conditions sous lesquelles une personne peut être identifiable[19]. L'avancement technologique remettant sans cesse en question les contours de la définition de donnée anonyme, un mouvement de datafication[20] discrète du concept de donnée personnelle fait aujourd'hui son œuvre, entraînant une dépendance contextuelle du statut de données personnelle à l'évolution de l'état de l'art technologique.

Dans ce cadre, les conditions d'objectif et de résultat donnent une portée potentiellement très ample au concept de donnée personnelle. Cette vaste extension de la notion peut être justifiée par le fait que les données dont le contenu fait directement référence à un individu ne sont pas les seules à pouvoir causer un tort informationnel.

Enfin, l'interprétation maximaliste du droit des données mènerait à son inapplicabilité en raison de son approche par la catégorisation, au détriment de celle par la réglementation de l'usage des connaissances déduites sur les personnes concernées. Selon une autre perspective critique défendue par les Professeurs Wachter et Mittelstadt au sein de de l'article A right to reasonable inference[21], le droit actuel n'offre qu'une protection imparfaite dans un monde où la distinction nette entre données et logiciel ne peut plus constituer les critères de constitution des législations existantes et futures, devant au contraire être recentré sur la notion d'inférence raisonnable[22]. Une approche par l'usage et non le résultat serait, selon eux, la solution. L'applicabilité d'une telle approche aurait pour mérite de dissoudre les problèmes de catégorisation posés, dont les modèles peuvent être considérés à la fois comme des données et comme des logiciels au titre du droit actuel.

III. De fructueuses perspectives de conciliation entre développement et éthique

A. La législation Canadienne en matière de gestion de la politique sur les services, le numérique et la prise de décision automatisée

On peut voir actuellement qu'au Canada, la mise en oeuvre d'audits incluant les enjeux de discriminations est obligatoire pour les institutions publiques depuis le 1er avril 2020 dans le cadre de la Loi sur la gestion des finances publiques, la Politique sur les services et le numérique[23] et la Directive sur la prise de décision automatisée[24]. Il fut notamment mis en œuvre une place afin d'accompagner les administrations dans leurs analyses d'impact nommé IEA (en français, Évaluation de l'Incidence Algorithmique).

Concrètement, la directive permet de déterminer le niveau de l'évaluation de l'incidence par quatre niveaux d'incidence[25]: probable, vraisemblablement modérée, vraisemblablement élevée, vraisemblablement très élevée.

Sachant que les décisions de niveau IV mèneront souvent à des effets irréversibles et permanents.

---

[18] Professeur à l'Institut de droit, de technologie et de société de l'Université de Tilburg

[19] Nadezhda Purtova (2018) The law of everything. Broad concept of personal data and future of EU data protection law, 2. Definition of personal data in EU data protection law: flexibility, adaptability, and uncertainty Law, Innovation and Technology, 10:1, 40-81, DOI: 10.1080/17579961.2018.1452176, https://www.tandfonline.com/doi/full/10.1080/17579961.2018.1452176

[20] La Datafication est un néologisme et un buzzword utilisé pour désigner la montée en puissance incontournable de la donnée (data) dans les pratiques marketing et d'autres domaines d'activité de l'entreprise.

[21] Wachter, Sandra and Mittelstadt, Brent, A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI (October 5, 2018). Columbia Business Law Review, 2019(2), Available at SSRN: https://ssrn.com/abstract=3248829

[22] Nécessaire pour combler l'écart de responsabilité actuellement posé par les " déductions à haut risque ", c'est-à-dire les déductions tirées des analyses de Big Data portant atteinte à la vie privée ou à la réputation ou ayant un faible impact sur la vie privée, à la réputation, ou une faible vérifiabilité dans le sens où celles-ci sont prédictives ou fondées sur des opinions alors qu'elles sont utilisées dans des décisions importantes.

[23] Article 4.4.2.4 de la Politique sur les services et le numérique

[24] Directive sur la prise de décisions automatisée en vigueur le 1er avril 2012 et devant être respectée au plus tard le 1er avril 2020 https://www.tbs-sct.gc.ca/pol/doc-fra.aspx?id=32592&section=html

[25] Annexe A

Ensuite, le niveau d'incidence permettra de déterminer les exigences d'évaluation[26] : examens par les pairs, avis, maillon humain de la prise de décisions, exigences en matière d'explication, formation, planification des mesures d'urgence et, approbation de l'exploitation du système.

B. Le projet parlementaire européen en matière d'IA

Le 21 avril 2021, la Commission européenne a présenté une proposition de règlement ayant pour finalité d'établir des règles harmonisées en matière d'intelligence artificielle[27].

S'inscrivant dans le prolongement de la publication par la Commission européenne d'un Livre Blanc sur l'intelligence artificielle[28] visant à définir les différentes options stratégiques en matière de promotion de l'IA ainsi que ses risques associés, la proposition poursuit quatre objectifs :

- veiller à la sûreté des systèmes d'IA sur le marché de l'Union ainsi qu'au respect des droits fondamentaux et les valeurs de l'Union ;
- assurer la sécurité juridique pour faciliter l'investissement et l'innovation ;
- faciliter le développement d'un marché unique et prévenir sa fragmentation par l'utilisation d'IA sûres et dignes de confiance.

Au travers de ce projet de règlement, la Commission a cherché à trouver un équilibre entre expansion du marché unique européen de l'IA et adoption d'un cadre réglementaire conciliant les différents droits et intérêts en jeu. Afin de construire une Europe adaptée à l'ère du numérique, la Commission a souhaité proposer de nouvelles règles et actions en faveur de l'excellence et de la confiance dans l'intelligence artificielle. Dans ce cadre, l'adoption d'exigences minimales fut proposée afin de concilier les risques liés à l'IA tout en adoptant un cadre flexible pouvant s'adapter aux futurs développements technologiques.

La proposition de la Commission prévoit notamment d'établir un cadre réglementaire différencié en fonction des risques présentés pour les utilisations. Ainsi, une distinction est faite entre :

- les utilisations interdites car présentant des risques inacceptables,
- les utilisations réglementées car présentant des risques élevés et,
- les utilisations soumises à des obligations de transparence car présentant certains risques de manipulation.

La proposition engloberait en outre tout « logiciel développé à l'aide d'une ou de plusieurs des techniques et approches énumérées à l'annexe I et qui peut, pour un ensemble d'objectifs définis par l'homme, générer des résultats tels que du contenu, des prédictions, des recommandations ou des décisions influençant les environnements avec lesquels ils interagissent ».

L'annexe I contient ainsi une liste de formes d'IA regroupant les approches d'apprentissage automatique (« machine learning »), les approches basées sur la logique et la connaissance, et les approches statistiques.

Point réellement innovant, la proposition aurait aussi pour objectif l'établissement de codes de conduite contraignants par les fournisseurs de systèmes d'IA ne présentant pas de risques élevés ainsi que la création d'un Conseil européen de l'intelligence artificielle (« European Artificial Intelligence Board »), composé de représentants des États membres et de la Commission.

Enfin, le Parlement européen et les États membres devront adopter la proposition de règlement dans le cadre de la procédure législative ordinaire afin que le règlement devienne applicable dans toute l'UE dans un délai de 24 mois suivant son entrée en vigueur.

---

[26] Annexe B

[27] Cette proposition fait suite aux trois résolutions adoptées par le Parlement européen le 20 octobre 2020[1] en matière d'IA, et des travaux en cours du Conseil de l'Europe.

[28] livre blanc de la Commission européennehttps://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_fr.pdf

C. Le manuel sur la protection des données et la confidentialité pour les développeurs d'IA en Inde

L'Inde exploite le potentiel de la technologie et prévoit de devenir une économie numérique de 1 000 milliards d'euros en 2025. Avec ses près de 5 millions d'ingénieurs en technologie, l'Inde est l'un des plus grands réservoirs de talents informatiques au monde. L'adoption de lignes directrices pour le développement responsable de l'IA est donc très importante, devant notamment passer par l'éducation des développeurs et des architectes de solutions en matière de bonnes pratiques afin de répondre aux attentes de la société civile et donc des utilisateurs sur le long terme.

L'intelligence artificielle s'imposant comme une technologie fondamentale qui alimente de nombreuses solutions électroniques de sociétés indiennes, les risques d'abus de cette technologie ont engendré une nécessité d'élaborer des politiques inclusives mais aussi des initiatives axées sur les praticiens afin d'intégrer de façon proactive de pratiques éthique et de confidentialité by design dans les solutions d'IA.

Dans ce cadre, le 15 juillet 2021 le Conseil de la sécurité des données Indienne[29] en collaboration avec l'Association allemande pour la coopération internationale (Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH)[30], la fondation Digital India[31] et le Koan Advisory Group a publié un manuel pour aider les développeurs à élaborer des solutions d'IA qui tiennent compte des considérations éthiques et de la protection de la vie privée[32].

Ce guide principalement destiné aux développeurs d'IA qui sont déjà familiarisés avec les processus d'apprentissage automatique (ML), comme les start-ups en phase de démarrage, permettra de limiter les problèmes pouvant survenir ultérieurement et engendrer des complications obligeant notamment les sociétés et starts up à abandonner la conception du produit ou de l'application par sa vision interdisciplinaire.

Outre les explications simples sur l'éthique de l'IA, il comprend : des listes de contrôle pour les développeurs à différents points d'intervention, des bonnes pratiques, et des exemples de défis auxquels les développeurs peuvent être confrontés.

Fait important, ce manuel évalue notamment l'impact du projet de loi sur la protection des données personnelles, 2019 (projet de loi PDP Indien)[33] ainsi que de la réglementation actuelle en Inde[34].

---

[29] Le DSCI est un organisme industriel de premier plan en matière de cybersécurité et de protection des données en Inde, créé par NASSCOM®, qui s'engage à rendre le cyberespace sûr, sécurisé et fiable grâce à ses divers programmes et initiatives.

[30] Est un prestataire de services dans le domaine de la coopération internationale pour le développement durable et de l'éducation internationale. Son expertise diversifiée est demandée dans le monde entier - par le gouvernement allemand, les institutions de l'Union européenne, les Nations unies, le secteur privé et les gouvernements d'autres pays. Nous travaillons avec des entreprises, des acteurs de la société civile et des institutions de recherche, favorisant une interaction réussie entre la politique de développement et d'autres domaines politiques et champs d'activité. Notre principal commanditaire est le ministère fédéral allemand de la coopération économique et du développement (BMZ).

[31] DIF est un groupe de réflexion politique qui promeut l'inclusion numérique, la cybersécurité, la fabrication mobile, la consommation domestique, les produits logiciels et les villes intelligentes

[32] https://www.dsci.in/content/privacy-handbook-for-ai-developers

[33] Le projet de loi sur la protection des données personnelles, 2019 a été présenté à Lok Sabha par le ministre de l'électronique et des technologies de l'information, M. Ravi Shankar Prasad, le 11 décembre 2019. Le projet de loi vise à assurer la protection des données personnelles des individus, et établit une autorité de protection des données à cet effet.

[34] Les données personnelles sont protégées par l'article 21 de la Constitution indienne, qui garantit à chaque citoyen le droit à la vie privée en tant que droit fondamental. La loi sur les technologies de l'information (IT Act) de 2000 contient quelques dispositions qui traitent des crimes associés aux données personnelles. Les sections pertinentes de la loi sont les sections 43A, 72 et 72A.

Table des matières