

Le caractère démocratique de l'usage d'algorithmes dans la décision publique

« (...)la volonté générale est toujours droite et tend toujours à l'utilité publique¹ » affirmait Jean-Jacques ROUSSEAU.

L'usage d'algorithme par les pouvoirs publics amène à reconsidérer les liens entre la volonté générale, subjective et nécessairement humaine et intérêt général, objectif et technique².

Le développement d'algorithme³, permettant le traitement automatisé d'un volume important de données⁴ permet l'identification quasi immédiate de l'intérêt général. Les utilisations d'algorithmes à des fins d'intérêt général sont multiples : en matière sociale, environnementale, fiscale ou de santé publique, dans le domaine des transports ou de la consommation d'énergie par exemple.

Les algorithmes, ainsi utilisés, constituent un tiers entre les citoyens et la définition légale de l'intérêt général, or le régime démocratique vise à permettre une identification aussi directe que possible⁵ de son intérêt⁶.

La question se pose de savoir si les algorithmes, se substituant à la volonté des personnes pour déterminer un intérêt public univoque, menacent le cadre démocratique, ou s'ils constituent au contraire un moyen d'aide à la décision, permettant de déterminer comment atteindre des objectifs d'intérêt publics qui resteraient définis par la nation. Cette question amène à s'interroger sur la capacité du droit à préserver une utilisation des algorithmes respectueuse de la volonté nationale.

Avant de s'interroger sur le caractère démocratique de l'usage d'algorithme dans la décision publique, il importe de rappeler les différentes acceptions de la démocratie numérique. Internet, puissant outil de diffusion de l'information, peut être mis au service de la démocratie de multiples façons. La démocratie numérique désigne à la fois l'utilisation d'internet

¹ *Contrat social*, II, III, in *OC*, III, p. 371.

² La volonté générale sera définie comme la volonté des citoyens qui vise * à l'intérêt commun.

L'intérêt général désigne le bien de tous. Ainsi, la volonté générale est conçue comme le moyen, en démocratie, de déterminer l'intérêt général.

³ La CNIL définit un algorithme comme "une suite finie et non ambiguë d'instructions permettant d'aboutir à un résultat à partir de données fournies en entrée."

⁴ L'administration, lorsqu'elle ouvre des données, rend accessible aux citoyens des informations qu'elle a collectées ou produites.

⁵ Daniel Gaxie, « La démocratie et le citoyen », , *Le Cens caché. Inégalités culturelles et ségrégation politique*, sous la direction de Gaxie Daniel. Le Seuil, 1978, pp. 9-31

⁶ Constitution de 1958, art. 3 : « La souveraineté appartient au peuple qui l'exerce par le biais de ses représentants et par voie de référendum ».

concourant à renforcer l'information et l'engagement citoyen⁷ Internet peut également être utilisé pour identifier la volonté des citoyens. Cette volonté peut être déduite de l'usage quotidien d'internet par les personnes ou sur des plateformes participatives expressément destiné à l'expression des citoyens. Internet peut également être appréhendé comme un espace de délibération publique où se construisent les choix des citoyens.

Concernant ces dernières utilisations d'internet, les algorithmes utilisent la volonté des personnes pour définir l'intérêt général.

Néanmoins, nous nous intéresserons dans cette contribution à la façon dont la prise de décision peut être automatisée en démocratie, et à la façon dont internet assure l'information des citoyens sur la prise de décision publique, concourant ainsi à préserver la transparence de l'action publique⁸.

Après nous être interrogé sur la légitimité démocratique d'un traitement automatisé de données pour identifier l'intérêt général (I), nous rechercherons en quoi l'ouverture des données publiques⁹ assure l'adéquation entre usage d'algorithme dans la décision publique et la volonté générale.

⁷En France, le mouvement de la civic tech s'intéresse aux utilisations d'internet pour promouvoir la démocratie.

⁸ Selon la *Sunlight Foundation*, fondée en 2006 aux États-Unis, la circulation facilitée de l'information permise par les nouvelles technologies doit conduire à une transparence accrue des institutions publiques

⁹ Le site du parlement européen définit les données ouvertes comme les informations que les pouvoirs publics recueillent, produisent ou achètent et qui sont mises à disposition des personnes privés pour être utilisées à d'autres fins.

I. La légitimité démocratique de l'utilisation d'algorithmes pour identifier l'intérêt général

Le droit encadre l'utilisation de données à des fins de recherche de l'intérêt public (A), il s'agit dès lors d'établir si l'utilisation d'algorithme pour traiter ces données s'inscrit dans un cadre démocratique (B).

A. Le cadre juridique de l'utilisation de données à des fins de recherche de l'intérêt public

L'utilisation d'algorithme implique la disponibilité d'un grand nombre de données¹⁰. Les données constituant un moyen privilégié de déterminer comment atteindre un objectif d'intérêt public, la question se pose de savoir comment le droit encadre leur traitement.

Le droit considère l'intérêt général comme une base légale du traitement de données personnelles (1). Le législateur français comme le législateur européen reconnaissent l'existence de « données d'intérêt général » (2).

1. L'intérêt général, base légale permettant le traitement de données

Le droit français et le droit de l'Union européenne reconnaissent l'intérêt général comme base légale justifiant le traitement de données personnelles (a), le pouvoir juridictionnel lorsqu'il interprète l'intérêt général, veille à ce que la poursuite de cet intérêt respecte le principe de proportionnalité (b).

a. La reconnaissance par le droit de l'intérêt général comme base légale justifiant le traitement de données

Le RGPD, dans une optique de protection des données personnelles¹¹ des citoyens européens, prévoit six bases légales pouvant fonder le traitement de données personnelles, parmi lesquelles l'intérêt légitime et l'intérêt public. De même, la loi informatique et liberté autorise elle en France le traitement de données visant à la sauvegarde de l'intérêt public ou d'un intérêt légitime¹².

Le droit de l'Union européenne admet l'exécution d'une mission d'intérêt public comme une des bases légales sur laquelle peut se fonder un traitement de données personnelles. Selon le règlement général sur la protection des données (RGPD)¹³, cette base légale se justifie pour les traitements d'intérêts publics ou les traitements mis en œuvre par les autorités publiques aux fins d'exécuter leurs missions. Le traitement doit être nécessaire à l'exécution de la mission¹⁴.

¹⁰ Alberto Alemanno, « Data for Good: Unlocking Privately-Held Data to the Benefit of the Many », , *European Journal of Risk Regulation* 2 (2018)

¹¹ La CNIL définit une donnée personnelle comme toute information se rapportant à une personne identifiée ou identifiable.

¹² Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés., Art 124.

¹³ Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016, relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données), cons.45.

¹⁴ Le traitement de données concernant non pas les usages d'un service mais son fonctionnement interne peut être conçu comme utile au fonctionnement du service mais non comme indispensable à l'exécution de sa mission.

En droit français, l'intérêt général constitue une des conditions de licéité d'un traitement de données : « lorsque le traitement est nécessaire à l'exécution d'une mission d'intérêt public ¹⁵»

En droit français comme en droit de l'Union Européenne, il n'appartient pas au responsable de traitement de définir l'intérêt public¹⁶. Les traitements, pour se fonder sur cette base légale, doivent porter sur une mission d'intérêt public définie dans des dispositions légales, européennes ou nationale¹⁷.

Ainsi conçu, le traitement de données personnelles s'inscrit dans un cadre strictement démocratique puisqu'il ne permet pas de définir l'intérêt public lui-même, le législateur demeurant seul en mesure de le circonscrire, mais seulement de déterminer quelles sont les conditions de son exécution.

b. La conciliation entre protection des données et l'intérêt public

En mars 2012, le Conseil constitutionnel a considéré que la lutte contre les fraudes par la sécurisation de documents de voyage constituait un motif d'intérêt général¹⁸. Il a néanmoins le même mois censuré le traitement de données personnelles disproportionnées au regard de l'objectif poursuivi¹⁹. En juin 2018, il a admis l'élargissement de l'accès au prétoire comme un motif d'intérêt général²⁰.

¹⁵ Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, art. 5 ; 5.

¹⁶ Le responsable de traitement est la personne morale (entreprise, commune, etc.) ou physique qui détermine les finalités et les moyens d'un traitement, c'est à dire l'objectif et la façon de le réaliser.

¹⁷ RGPD, sup. cons 10 et 45.

¹⁸ Décision n° 2012-652 DC du 22 mars 2012 - Loi relative à la protection de l'identité « Considérant que la création d'un traitement de données à caractère personnel destiné à préserver l'intégrité des données nécessaires à la délivrance des titres d'identité et de voyage permet de sécuriser la délivrance de ces titres et d'améliorer l'efficacité de la lutte contre la fraude ; qu'elle est ainsi justifiée par un motif d'intérêt général ; (...)

¹⁹ Décision n° 2012-652 DC du 22 mars 2012 - Loi relative à la protection de l'identité « Considérant, en second lieu, que la liberté proclamée par l'article 2 de la Déclaration des droits de l'homme et du citoyen de 1789 implique le droit au respect de la vie privée ; que, par suite, la collecte, l'enregistrement, la conservation, la consultation et la communication de données à caractère personnel doivent être justifiés par un motif d'intérêt général et mis en œuvre de manière adéquate et proportionnée à cet objectif. (...) Considérant qu'il résulte de ce qui précède qu'en égard à la nature des données enregistrées, à l'ampleur de ce traitement, à ses caractéristiques techniques et aux conditions de sa consultation, les dispositions de l'article 5 portent au droit au respect de la vie privée une atteinte qui ne peut être regardée comme proportionnée au but poursuivi ; (...).

²⁰ Décision n° 2018-765 DC du 12 juin 2018 - Loi relative à la protection des données personnelles « D'une part, (...) le législateur a entendu permettre la mise en œuvre de traitement de données à caractère personnel relatives aux condamnations pénales, aux infractions par des personnes morales collaborant au service public de la justice (...) D'autre part, (...) le législateur a également entendu ouvrir cette faculté aux personnes victimes ou mises en œuvre dans une procédure pénale, afin de leur permettre de préparer ou de mettre en œuvre un recours en justice. Ce faisant, le législateur a poursuivi des objectifs d'intérêt général. »

Voir aussi Décision n° 2013-684 DC du 29 décembre 2013 - Loi de finances rectificative pour 2013 sur la conciliation entre, d'une part, la lutte contre la fraude fiscale, qui constitue un objectif de valeur constitutionnelle, et, d'autre part, le respect des autres droits et libertés constitutionnellement protégés, au nombre desquels figure le droit au respect de la vie (...).

En mai 2021, le Conseil constitutionnel a précisé les conditions dans lesquelles le traitement de données personnelles était justifié par la lutte contre la propagation de l'épidémie de coronavirus²¹.

2. La reconnaissance législative de données d'intérêt public

Le droit et la gouvernance des données encouragent le partage de données lorsqu'elles sont utiles à l'intérêt général. La loi pose une obligation d'ouvrir certaines données d'intérêt général (a). Le concept d'altruisme des données permet le partage volontaire de données (b).

a. L'obligation légale de partager des données d'intérêt général

La *loi pour une république numérique*²² introduit la notion de données d'intérêt général, pour optimiser l'utilisation des données aux fins d'intérêt général. Cette loi favorise l'exploitation de ces données dans le cadre de la délégation de la gestion d'un service public²³. Des personnes morales de droit privé peuvent se trouver dans l'obligation de transmettre des bases de données dont elles disposent afin d'enrichir les statistiques publiques²⁴.

Cette loi vise, en particulier, à favoriser notamment le développement d'offres d'énergie, d'usages et de services énergétiques²⁵. Les données des administrations, ainsi que des décisions des juridictions administratives et judiciaires, présentant un intérêt économique, social, sanitaire ou environnemental, pourront être réutilisées²⁶.

Cette loi permet à l'Etat d'accéder à diverses bases de données non spécifiquement mentionnées dans les textes mais dont l'intérêt public se conçoit aisément : transport ferroviaire SNCF, l'Ifremer, l'Office national des forêts, la Comédie-Française, l'Établissement français du sang, Observatoire des loyers. La mise à disposition de l'Etat de ces données permet l'amélioration du travail législatif.

Le partage volontaire de données d'intérêt général est promu par la gouvernance des données.

b. L'altruisme des données

L'Union européenne a introduit le concept d'altruisme des données²⁷, selon lequel le possesseur de données, personne physique ou morale, doit être incité à partager les

²¹ Décision n° 2021-819 DC du 31 mai 2021 Loi relative à la gestion de la sortie de crise sanitaire.

²² LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique_

²³ Ibid. art 17.

²⁴ Ibid. art. 19.

²⁵ Ibid. art 23 et 24.

²⁶ Ibid.art 6 modifiant le code des relations entre le public et l'administration, Art. L. 312-1-1.

²⁷ Règlement (UE) 2022/868 du Parlement européen et du Conseil du 30 mai 2022 portant sur la gouvernance européenne des données et modifiant le règlement (UE) 2018/1724 (règlement sur la gouvernance des données) (Texte présentant de l'intérêt pour l'EEE)

données²⁸ qu'il estimerait utile à l'intérêt général²⁹, sans contrepartie. Le *Règlement sur la gouvernance des données (data governance act (dga))* complète les règles du marché numérique et les outils disponibles pour faciliter l'accès aux données et leur réutilisation, en particulier par le secteur public, cela dans le plein respect des valeurs européennes sur la protection des libertés individuelles.

Dans le domaine de la santé, ou les données personnelles, à caractères sensibles, bénéficient d'une protection particulière, le développement de certains espaces de données dépend largement de ce concept³⁰.

La mise à disposition de l'Etat de bases de données vise à déterminer comment accomplir une mission d'intérêt général démocratiquement identifiée de la façon la plus efficace possible. Cette réflexion sur les conditions d'exécution d'une mission de service public peut avoir une incidence sur la définition de l'intérêt public, en particulier lorsque plusieurs missions de service public entrent en concurrence³¹. Ces bases de données peuvent être utilisées par les représentants pour déterminer quel intérêt privilégier. Le traitement algorithmique des données pourrait être employé pour privilégier une mission de service public parmi d'autres. Il est loisible de s'interroger sur les conditions d'un usage démocratique des algorithmes.

²⁸ A titre d'exemple, l'Urssaf développe des partenariats statistiques dans l'objectif d'optimiser l'échange de données et d'encourager le partage d'informations au service d'une meilleure compréhension de l'environnement économique et des besoins de ses publics. Plus de 200 conventions de partenariat ont été signées avec une grande variété d'acteurs tels que collectivités locales, préfectures, conseil économique et social, Insee, Directe, Banque de France, Agence de développement économique.

²⁹ Il vise des causes d'intérêt général telle que la santé, la lutte contre le changement climatique, l'amélioration de la mobilité et des services publics, et plus largement tout domaine de recherche scientifique.

³⁰ Anne-Sophie Hulin, *Les données de santé au service du bien commun ? L'exemple de l'espace européen des données de santé*, 29 mars 2023, Observatoire international sur les impacts sociétaux de l'IA et du numérique.

³¹ Sur la pluralité d'intérêt général : Jean-Marie Pontier, « Bien commun et intérêt général », *Les Cahiers Portalis*, vol. 4, no. 1, 2017, pp. 33-52.

B. La conciliation entre prise de décision automatisée et expression de la volonté générale

La légitimité démocratique implique à la fois que les citoyens participent à la prise de décision et que les décisions adoptées visent effectivement à l'intérêt général³². L'utilisation d'algorithmes par les personnes publiques, quoique renforçant ce second type de légitimité, pourrait limiter le premier. Une absence totale de participation humaine à la prise de décision constituerait une rupture avec le régime démocratique qui implique que l'intérêt général soit déterminé par la volonté générale. En cela, la démocratie implique la délibération, faute de quoi elle serait supplantée par la technocratie, privilégiant l'efficacité technique.

Cette nécessité de maintenir une participation humaine pour demeurer dans un système démocratique est consacrée par la loi. L'article 1er de la loi Informatique et Libertés pose comme premier horizon que « L'informatique doit être au service de chaque citoyen [...], qu'elle ne doit porter atteinte ni à l'identité humaine, ni aux droits de l'homme, ni à la vie privée, ni aux libertés individuelles ou publiques³³.

L'utilisation d'algorithmes demeure démocratique tant qu'ils demeurent un simple outil permettant de poursuivre des fins d'intérêt général déterminées dans le cadre d'un débat démocratique. Ils pourraient être détournés de cette visée par un manque de maîtrise par les représentants du processus algorithmique de prise de décision (1) ou par une volonté d'exercer un pouvoir dépassant le cadre démocratique (2). Enfin, les algorithmes pourraient nuire au processus délibératif, ce que le RPGD peine à éviter (3).

1. Les effets sur la démocratie d'une méconnaissance de l'absence de neutralité des algorithmes

Les représentants nationaux, lorsqu'ils se réfèrent à des algorithmes sans discernement(a), peuvent se trouver dupes par les auteurs de ces algorithmes, poursuivant d'autres fins, économiques en particulier(b).

a. L'absence de neutralité des algorithmes

Un algorithme demeure le reflet des données d'apprentissage qui lui ont été fournies pour l'entraîner. L'ensemble des citoyens ne se limitant pas à l'ensemble des internautes, les algorithmes ne peuvent rendre compte de la volonté de tous les citoyens³⁴. La Commission nationale de l'informatique et des libertés (CNIL) a documenté certains biais associés à l'utilisation d'algorithmes³⁵.

³² Sur les légitimités démocratiques : Rosanvallon, Pierre. « La question de la légitimité démocratique : l'exemple de la Justice », *Après-demain*, vol. 30,nf, no. 2, 2014, pp. 5-6.

³³ Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.

³⁴ . À l'aide d'un logiciel baptisé Adfisher, ils ont créé 17 000 profils dont ils ont ensuite simulé la navigation sur le Web afin de mener une série d'expériences. Ils ont ainsi constaté que les femmes se voyaient proposer des offres d'emploi moins bien rémunérées que celles adressées à des hommes, à niveau similaire de qualification et d'expérience.

³⁵CNIL, Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle synthèse du débat public animé par la CNIL dans le cadre de la mission de réflexion éthique confiée par la loi pour une république numérique.

Les représentants politiques ne peuvent avoir conscience de l'ensemble des biais pouvant résulter de l'usage d'algorithmes. Ils peuvent ne pas évaluer à leur juste mesure le caractère partial des algorithmes. Or ces biais pourraient être exploités à des fins idéologiques ou économiques³⁶.

b. L'utilisation d'algorithmes à des fins privées

Les données personnelles sont particulièrement valorisées lorsqu'elles sont utilisées par des plateformes de création de services de technologies civiques. Les réseaux sociaux, qui sont des entreprises à vocation commerciale, ciblent les profils à des fins publicitaires. Or ces mêmes réseaux sont utilisés pour recueillir des données à des fins politiques³⁷.

La question demeure de savoir si les objectifs commerciaux poursuivis par ces plateformes sont compatibles avec des visées politiques .

2. L'utilisation de données à des fins privés par les représentants

Les algorithmes peuvent être employés par les décideurs politiques à des fins électorales ou pour conserver le pouvoir.

Les données disponibles sur internet³⁸ sont exploitées pour organiser les campagnes électorales et cibler les électeurs indécis³⁹ en leur envoyant des messages personnalisés⁴⁰.

L'usage à des fins politiques des données collectées lors d'une démarche de consultation ou pour la réalisation d'un service public représente un risque de détournement des données personnelles. Cet usage permettrait à des élus d'accroître leurs chances de réélections⁴¹. Enfin, la diffusion de données personnelles erronées ou invérifiables sur un candidat à la veille des élections peut être utilisée pour le décrédibiliser.

Le règlement général sur la protection des données (RGPD)⁴² ne permet pas de prévenir le traitement à des fins non démocratiques de données librement cédées⁴³.

³⁶ Un biais nous pousse vers des informations anxiogènes.

³⁷ Voxe.org (d'abord connue pour ses comparateurs de programmes politiques) développe sur Messenger un chatbot de « décryptage » de l'actualité et des conseils « pour s'engager »

³⁸ Réseaux sociaux, données publiques d'historique des bureaux de vote et socio-démographiques de l'INSEE, données personnelles recueillies sur le terrain : emails, numéros de téléphone...

³⁹ Le recueil massif de données dans le cadre des campagnes politiques a été mis à jour lorsque Cambridge Analytica, lors des Présidentielles américaines de 2016, a utilisé illégalement les données de 87 millions d'utilisateurs de Facebook.

⁴⁰ L'observatoire de la vie citoyenne de la CNIL observe ces pratiques.

⁴¹ . À la rentrée 2018, un président de Région utilise une base de données de lycéens pour leur souhaiter par SMS une « excellente rentrée », ce qui a déclenché des plaintes à la CNIL donnant notamment lieu à un rappel des règles à respecter. En septembre 2019, l'autorité de protection des données britannique, la ICO (*Information Commissioners's Office*), s'inquiète de la possible utilisation par le Premier Ministre Boris Johnson des données collectées par les sites en .gov.uk

⁴² *Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016, relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données), chap. 2.*

Le RGPD pose certains principes (la transparence, le consentement, la proportionnalité et la finalité), et reconnaît des droits pour les individus (information, accès, modification, opposition, effacement).

⁴³ Le consentement constitue l'une des bases juridiques sur lequel peut se fonder un traitement de données personnelles.

Si, en démocratie, un jeu de données peut être utilisé pour accéder au pouvoir, les algorithmes peuvent également être utilisés pour contrôler la population dans des systèmes autoritaires.

Les technologies utilisées pour favoriser la participation citoyenne peuvent être employées par le pouvoir à des fins de surveillance. S'informer sur les citoyens permet de les contrôler. Les réseaux⁴⁴ qui permettent la mobilisation citoyenne, peuvent être exploités par des gouvernements qui déjouent les codes de sécurité, pour obtenir l'identité des protestataires⁴⁵.

Les algorithmes peuvent nuire à la démocratie lorsqu'ils se substituent à l'expression de la volonté générale mais aussi lorsqu'ils entravent la délibération publique.

3. Les effets sur la délibération de l'usage d'algorithmes dans les médias sociaux

Le ciblage d'informations par algorithmes dans les médias sociaux peut entraver le bon déroulement de la délibération publique.

Les algorithmes peuvent nuire à la mise en œuvre de règles de la délibération publique⁴⁶ Selon le concept de bulle de filtre⁴⁷ « les algorithmes aident les gens à s'entourer de médias qui soutiennent ce qu'ils pensent déjà ». Les citoyens se retrouvent enfermés dans un système de pensée car ils n'ont connaissance par le biais d'internet et des réseaux que des arguments corroborant leurs préjugés⁴⁸.

La démocratie représentative implique l'information des citoyens sur les données traitées. Cette dernière rend possible l'intégration du traitement algorithmique de données au débat démocratique⁴⁹. Il s'agit, dès lors, d'établir comment les citoyens doivent être informés des décisions prises à l'aide d'algorithmes.

⁴⁴ Le forum LIHKG et la messagerie Telegram en particulier.

⁴⁵ Nocetti, Julien. « Réseaux sociaux et révoltes sociétales. Les réseaux de la colère », Thierry de Montbrial éd., *Ramses 2021. Le grand basculement*, Institut français des relations internationales, 2020, pp. 216-219.

⁴⁶ *Espace public et démocratie délibérative : un tournant*, Paris, Gallimard, coll. « NRF Essais », 2023

⁴⁷ Eli Pariser, *The Filter Bubble: What the Internet Is Hiding from You*, Penguin Press (New York, mai 2011) (ISBN 978-1594203008).

⁴⁸ Déjà ac presse papier amplifie

⁴⁹ Anne-Marie Le Pourhiet, « Participation et citoyenneté : progrès démocratiques ou gadgets constitutionnels ? », *Revue française de droit constitutionnel*, vol. 116, no. 4, 2018, pp. 753-766.

II. L'ouverture des données publiques comme moyen d'assurer l'adéquation entre usage d'algorithme dans la décision publique et la volonté générale

Nous rechercherons par quel moyen le droit garantit un usage démocratique des algorithmes, selon lequel l'automatisation du traitement des données serait conçu comme une aide à l'interprétation par les représentants de la volonté générale. L'ouverture des données publiques, quoique nécessaire à assurer la transparence en démocratie (A), ne suffit pas toujours à permettre leur exploitation par les citoyens (B).

A. L'ouverture des données comme moyen de garantir la transparence du processus démocratique

L'ouverture des données constitue un moyen de renforcer la croissance économique, l'innovation ou l'efficacité des services publics⁵⁰ mais aussi un gage de démocratie. Après avoir rappelé comment le droit organisait en pratique l'ouverture des données publiques (1), nous montrerons en quoi elle apparaît désormais comme un prérequis en démocratie (2).

1. La consécration par le droit de l'ouverture des données publiques

La *loi pour une république numérique* a substitué à l'ouverture à la demande des données un principe d'ouverture des données d'intérêt général⁵¹. Ces informations, une fois rendues publiques, peuvent librement être utilisées à d'autres fins.

L'ouverture des données publiques s'inscrit dans une tradition de transparence de l'action publique et de partage des informations détenues par la puissance publique⁵². La loi sur la liberté d'accès aux documents administratifs et de la réutilisation des informations publiques⁵³ adoptait une conception large du droit à l'information des citoyens.

Désormais, des portails d'*open data*⁵⁴ permettent le libre téléchargement de jeux de données en vue d'un traitement. Ces données peuvent donc être exploitées à des fins scientifiques, techniques ou communicationnelles⁵⁵.

⁵⁰Le rapport sur la politique publique de la donnée, des algorithmes et des codes sources, rendu par la commission présidée par Eric Bothorel, met en avant une série d'arguments en faveur d'une plus grande ouverture des données. Parmi ceux-ci, « un levier économique pour la création de nouveaux services ou d'entreprises innovantes, en particulier dans les secteurs de l'intelligence artificielle, un gain en performance et en réactivité dans des situations d'urgence grâce au recueil d'informations ; une meilleure évaluation des politiques publiques grâce au croisement de données issues de différentes administrations. »

⁵¹ LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique, JORF n°0235 du 8 octobre 2016.

⁵² En octobre 2016, la France a pris la présidence du Partenariat pour le gouvernement ouvert (PGO - « *Open Government Partnership* »).

⁵³ Loi n° 79-587 du 11 juillet 1979 relative à la motivation des actes administratifs et à l'amélioration des relations entre l'administration et le public.

⁵⁴ La plateforme data.gouv.fr permet aux services publics de publier des données et à la société civile de les enrichir, modifier, interpréter en vue de coproduire des informations d'intérêt général. Il est le premier site au monde à proposer aux usagers d'enrichir et d'améliorer les données disponibles, et également d'en déposer de nouvelles.

⁵⁵ Il devient ainsi possible de créer des applications ou des graphiques, de vérifier de hypothèses.

En matière environnementale par exemple, mettre les données d'intérêt général au service de l'environnement et de la gouvernance de la transition écologique⁵⁶ constituait l'une des mesures pour un agenda national et européen sur un numérique responsable⁵⁷. L'expression « Open data » est d'ailleurs issue de la recherche publique environnementale. Ce concept visait à partager à moindre coût les données pour favoriser une réflexion globale tout en favorisant la participation⁵⁸.

Le RGPD consacre les principes de Légalité, équité et transparence⁵⁹, imposant ainsi le traitement ouvert et équitable des informations traitées⁶⁰.

2. L'intérêt démocratique de l'ouverture des données publiques

En démocratie, l'ouverture des données constitue un moyen d'information des citoyens. Elle permet à ces derniers de se déterminer de façon éclairée (a) et d'assurer le contrôle de l'action publique (b).

a. L'ouverture des données, condition de la transparence de la décision publique

L'ouverture des données constitue un moyen d'observation par les citoyens de l'activité des services publics. La transparence de l'État auprès de ses citoyens⁶¹, permet de maintenir la confiance des citoyens envers les pouvoirs publics⁶². Si l'information des citoyens sur les fondements de la décision publique apparaît comme une nécessité démocratique, l'émergence et le développement d'algorithmes dans la décision publique amène à redéfinir les conditions de sa mise en œuvre. La démocratie n'exclut pas une part d'expertise réservée aux représentants nationaux⁶³. Néanmoins, en l'absence d'algorithme, les représentants, exploitent directement les données pour se prononcer dans l'intérêt général. Dans cette

⁵⁶Annie Blandin-Obernesser ; Sophie Flak ; Hervé Pillaud ; Thomas Landrain ; Marylou Leroy ; Philippine Régniez ; Méneould Michaud de Brisis Conseil National du Numérique, Feuille de route sur l'environnement et le numérique., 9 juillet 2020, 160p.

⁵⁷ LOI n° 2021-1755 du 23 décembre 2021 visant à renforcer la régulation environnementale du numérique par l'Autorité de régulation des communications électroniques, des postes et de la distribution de la presse

⁵⁸Goncalves, D., Rufat, S. (2016). Open data et droit de la donnée : les collectivités à l'épreuve des réglementations européennes. *Cybergeo : European Journal of Geography* [En ligne], Science et Toile, document 787, mis en ligne le 17 août 2016

⁵⁹ Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016, relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données), art. 12.

⁶⁰Le principe de responsabilité à l'égard des usagers peut également impliquer une ouverture des données. L'accès aux données permet en particulier d'invoquer la responsabilité de l'administration.

⁶¹ La gestion de l'épidémie en constitue un exemple. Les indicateurs sont rendus publics et les actions sont motivées par la donnée : confinement, déconfinement, évolution des règles de vaccination dépendent des données publiques des services hospitaliers

⁶² Ainsi, Sarah Labelle et Jean-Baptiste Le Corf (2012) ont montré que les politiques d'ouverture des données s'accompagnent de discours qui visent le politique à travers la qualité des processus démocratiques, l'administratif à travers une amélioration des relations entre administration et administrés, et la technique à travers le soutien à l'innovation.

⁶³ Jean-Michel Belorgey, « Elus, citoyens et experts : distribution des rôles », *Revue Projet*, vol. 295, no. 6, 2006, pp. 51-56.

Marc-Antoine Gavray, « La démocratie selon Platon ou de l'ignorance collective érigée en principe politique », *Cahiers philosophiques*, vol. 160, no. 1, 2020, pp. 23-38.

perspective, la participation des citoyens à la décision publique est de même nature que la prise de décision par les représentants de la nation ; un raisonnement humain constitue le lien entre données et la décision adoptée.

L'usage d'algorithme pour identifier l'intérêt général redéfinit la nature de la participation des citoyens à la décision publique. Cette dernière se fonde en effet sur un volume, une variété et une quantité des données sans précédent. Les algorithmes constituent un intermédiaire nouveau et spécifique dans le processus décisionnel. Ils ne sont pas sans incidence sur la délibération publique⁶⁴.

b. L'utilisation de données dans le cadre d'un débat démocratique

Pour éviter que les données ne soient utilisées comme argument d'autorité ou pour reproduire sans recul la réalité, les citoyens doivent être en mesure de les intégrer comme éléments de délibération.

L'ouverture des données permet d'intégrer l'usage de données au débat public⁶⁵. Les citoyens peuvent vérifier l'exactitude des données mentionnées par leurs représentants. D'autre part un débat public peut porter sur la pertinence du choix de données traitées. Il s'agit d'éviter « de se saisir d'éléments non publiés pour déformer la réalité »⁶⁶

La simple vérification des données apparaît insuffisante à assurer la démocratisation de l'espace public des données.

La représentativité d'un échantillon dans un contexte donné peut être discutée. La référence à des données dans le cadre d'un débat dont les partis prenants seraient en mesure de « lire les données » permet d'en limiter un usage biaisé. Il s'agit par l'ouverture des données publiques de réduire le clivage entre experts et non experts. Internet devient un outil permettant à chacun de s'approprier les éléments pris en compte dans la décision publique. Une réflexion approfondie sur l'usage de donnée présuppose une acculturation de chacun à divers aspects de la démocratie numérique telles que les statistiques⁶⁷.

⁶⁴ Capelle, C., Lehmans, A., Liqueste, V. (2017). De la visibilité à la médiation : l'open data et ses usages en éducation. Colloque international « Big data et visibilité en ligne : un enjeu pluridisciplinaire de l'économie numérique », Novembre, Fort de France, Martinique. <halshs01633284>

⁶⁵ Anne Lehmans. « Les réinventions de la démocratie à l'aune de l'ouverture des données : du discours de la participation aux contraintes de la gouvernance », *Les Enjeux de l'information et de la communication*, vol. 19/2, no. 2, 2018, pp. 135-146.

⁶⁶ CNIL, Cahier air2021 - Entre partage et protection, quelle éthique pour l'ouverture des données

⁶⁷ La présentation de jeux de données, en fonction de l'usage visé, nécessite des opérations complexes. Ainsi, l'éducation à la citoyenneté numérique représente une dimension de l'éducation à la citoyenneté qui vise à apprendre aux élèves à travailler, vivre et partager dans des environnements numériques de manière positive. Voir aussi Clément Mabi « Quel(s) numérique(s) pour la démocratie ? », *Cahiers de l'action*, vol. 57, no. 1, 2021, pp. 89-100.

B. Les conditions d'une possible exploitation des données publiques ouvertes par les citoyens

L'accessibilité des données ne suffit pas à permettre aux citoyens de les utiliser dans une visée démocratique⁶⁸. Encore faut-il que les citoyens soient informés de leur existence et en mesure d'en analyser l'usage par les pouvoirs publics.

Après avoir examiné comment la législation relative à l'ouverture des données publiques rend possible l'exploitation par les citoyens de données ouvertes (1), nous évaluerons les effets d'une absence de normalisation sur l'interopérabilité (2) puis dans quelle mesure le droit peine à imposer la lisibilité des données ouvertes (3).

1. La législation permettant l'exploitation par les citoyens de données ouvertes

La loi pour une république numérique consacre la lisibilité des données publiques⁶⁹.

Sans dresser un inventaire de l'ensemble des conditions nécessaires à rendre accessibles et intelligibles des données, nous examinerons dans quelle mesure le droit rend possible l'exploitation des données publiques ouvertes par les citoyens⁷⁰.

La protection des données personnelles, quoique consacrant des droits individuels, impacte l'exercice de droits politiques de citoyens. Le devoir d'information des personnes concernées favorise la prise en compte des données par les citoyens(a) alors que l'anonymisation des données en limite la vérifiabilité (b).

a. L'information des personnes concernées par le traitement de données

La connaissance par les citoyens de l'utilisation des données pour guider la décision publique apparaît nécessaire à une utilisation démocratique des algorithmes dans un cadre démocratique.

La loi pour une république numérique impose l'information de toute personne intéressée par un traitement algorithmique⁷¹. L'information des personnes concernées par le traitement de données, si elle contribue à l'information de citoyens relative à l'existence d'un traitement de données, ne suffit pas à assurer l'information de l'ensemble des citoyens nationaux.

⁶⁸ « L'ouverture reste une utopie si les conditions organisationnelles et culturelles de leur usage ne sont pas réunies. Même si les données sont accessibles, sur les plans technique et juridique, le « public » visé par l'ouverture doit connaître leur existence d'une part, être capable de les utiliser d'autre part. »

⁶⁹ Art. L. 300-4.-du code des relations entre le public et l'administration.

Art 3 loi pour une république numérique. La loi pour une république numérique prévoit que « toute mise à disposition effectuée sous forme électronique en application du présent livre se fait dans un standard ouvert, aisément réutilisable et exploitable par un système de traitement automatisé

⁷⁰ Samuel Goeta, Instaurer des données, instaurer des publics : une enquête sociologique dans les coulisses de l'open data. Sociologie. Télécom ParisTech,2016.

⁷¹ « Art. L. 311-3-1.-Sous réserve de l'application du 2° de l'article L. 311-5, une décision individuelle prise sur le fondement d'un traitement algorithmique comporte une mention explicite en informant l'intéressé. Les règles définissant ce traitement ainsi que les principales caractéristiques de sa mise en œuvre sont communiquées par l'administration à l'intéressé s'il en fait la demande.

Le droit à l'information prévu par le RGPD⁷² constitue un droit individuel et non politique. Il tend à limiter toute exploitation abusive de données personnelles⁷³. Ainsi, il ne permet pas d'informer les citoyens sur l'utilisation d'algorithmes dans une visée démocratique puisqu'elle ne permet pas d'informer les citoyens n'étant pas personnellement visés par le traitement de leurs données. Or un contrôle efficace par les citoyens de l'usage de données publiques nécessiterait l'information de chacun d'entre eux de l'ouverture des données traitées par l'Etat à des fins publiques.

Les individus doivent être informés de leurs droits et pouvoir soulever leur éventuelle violation.

Si l'information des usagers de l'existence d'un traitement de données les concernant constitue un moyen pour ces derniers de se saisir d'un jeu de données, l'anonymisation des données pourrait au contraire limiter l'utilisation de données.

b. Le contrôle de données anonymisées

La *loi pour une république numérique* dispose que certaines données doivent être anonymisées⁷⁴. L'anonymisation⁷⁵ permet d'exploiter des données personnelles dans le respect des droits et libertés des personnes, en particulier lorsque la date au-delà de laquelle des données personnelles peuvent être conservées est échu⁷⁶.

L'exigence d'anonymisation est renforcée par l'ouverture des données dans la mesure où les usages des données publiées demeurent inconnus. L'exigence de secret des communications individuelles est d'ailleurs souvent associée à la recherche de transparence⁷⁷.

Néanmoins, l'anonymisation des données, bien qu'indispensable à la protection des données personnelles, peut limiter le contrôle de leur usage par les pouvoirs publics. En effet, les données, une fois anonymisées deviennent difficilement vérifiables. De plus, la nécessité d'anonymisation des données peut limiter leur utilisation. Elle implique en effet d'éviter que des informations ne soient recoupées en vue d'identifier leur auteur.

⁷² Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016, relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données), chap. 2.

Le RGPD pose certains principes (la transparence, le consentement, la proportionnalité et la finalité), et reconnaît des droits pour les individus (information, accès, modification, opposition, effacement),

⁷³ En septembre 2019, Facebook suspend 69 000 applications qui recueillaient abusivement des données personnelles de la plateforme

Kate Conger, Gabriel J.X. Dance and Mike Isaac, Facebook's Suspension of 'Tens of Thousands' of Apps Reveals Wider Privacy Issues, New-York Times, 20 septembre 2019, <https://www.nytimes.com/2019/09/20/technology/facebook-data-privacy-suspension.html>

⁷⁴ « Art. L. 312-1-2.-Sauf dispositions législatives ou réglementaires contraires, lorsque les documents et données mentionnés aux articles L. 312-1 ou L. 312-1-1 comportent des mentions entrant dans le champ d'application des articles L. 311-5 ou L. 311-6, ils ne peuvent être rendus publics qu'après avoir fait l'objet d'un traitement permettant d'occulter ces mentions.

⁷⁵ Selon la CNIL, « l'anonymisation est un traitement qui consiste à utiliser un ensemble de techniques de manière à rendre impossible, en pratique, toute identification de la personne par quelque moyen que ce soit et de manière irréversible. »

⁷⁶ Le RGPD limite la durée de conservation des données.
RGPD, supr. Cons. 45.

⁷⁷ « privacy for the weak and transparency for the powerful »

2. Les difficultés à imposer juridiquement l'intelligibilité des données

Pour que l'ouverture des données soit opérationnelle, l'intelligibilité des données doit s'ajouter à leur visibilité et à leur lisibilité. Or, il semble difficile d'imposer la lisibilité de données. A des difficultés techniques à imposer l'ouverture des sources (a), s'ajoute une raison juridique dans la mesure où la source des algorithmes utilisés pour traiter les données constitue la propriété de leur auteur (b). Enfin, l'ouverture de certaines données est limitée par les secrets légaux et la sécurité nationale (c).

a. La nécessité de repenser le format des données dès leur conception pour les rendre intelligibles

Les données dont l'usage n'a pas été envisagé en amont constitue une boîte noire. L'ouverture des données amène à s'interroger sur leur cycle de vie dès leur conception. Leur utilisation doit être prise en compte dès leur collecte et doit donc être définie en amont⁷⁸. L'utilisabilité des données, définie par la norme comme « le degré selon lequel un produit peut être utilisé, par des utilisateurs identifiés, pour atteindre des buts définis avec efficacité, efficience et satisfaction, dans un contexte d'utilisation spécifié » devient déterminante⁷⁹.

L'utilisabilité des données par des acteurs n'ayant pas participé à leur traitement implique une redéfinition, dès le recueil des données, des formats d'interopérabilité⁸⁰. En effet, lorsque les pratiques sont distinctes, il devient impossible de permettre la publication de jeux de données de qualité et universellement utilisables. Par conséquent, le partage des données implique la standardisation de leur traitement.

La directive INSPIRE de la Direction générale de l'environnement de la Commission européenne datant de 2007⁸¹, « vise à établir en Europe une infrastructure de données géographiques pour assurer l'interopérabilité entre bases de données et faciliter la diffusion, la disponibilité, l'utilisation et la réutilisation de l'information géographique en Europe »⁸². L'intervention du législateur européen a permis l'émergence d'outils tels le Géocatalogue et le Géoportail. L'obligation des informaticiens d'harmoniser leur pratique rend exploitables les données. Les licences qui déterminent les conditions juridiques de réutilisation des données sont essentielles.

⁷⁸ Le volume, la diversité et l'hétérogénéité des données nécessitent un pilotage dans le traitement de l'infrastructure technique et logicielle ainsi que dans la description et la mise en forme des données en vue de leur publication, dans une logique de formatage comme condition du partage. Le pilotage doit intégrer la valeur des données au regard des usages projetés, de la transformation de la donnée en information, dans une démarche de qualification qui rompt avec les routines organisationnelles et qui nécessite de la part des agents une expertise dans la représentation des bases de données à l'aune de leurs usages externes. Le format informatique, la licence, l'organisation de la base de données, conditionnent l'utilisabilité et la réutilisabilité des données ouvertes et les installent dans une logique de normalisation.

⁷⁹ Norme ISO 9241-11

⁸⁰ Anne Lehmans. L'horizon d'une culture de la donnée ouverte : de l'utopie aux pratiques de gouvernance des données. Revue COSSI : communication, organisation, société du savoir et information, 2018. fahal-01770335

⁸¹ Directive 2007/2/CE du Parlement européen et du Conseil du 14 mars 2007 établissant une infrastructure d'information géographique dans la Communauté européenne (INSPIRE)

OJ L 108, 25.4.2007, p. 1–14

⁸² La plateforme PIGMA en constitue une application.

L'obligation juridique de partager à l'identique se heurtant aux intérêts commerciaux des entreprises, les collectivités territoriales renoncent à imposer des licences trop exigeantes⁸³.

Pour légitimer la démocratie en rendant intelligible les choix techniques dissimulés dans les algorithmes permettant la recherche de l'intérêt général, il s'agit de permettre l'ouverture des données mais également l'ouverture des sources.

b. Le caractère non obligatoire de l'ouverture des sources des algorithmes

L'ouverture du code de la plateforme et des données de la participation tend à assurer l'adaptabilité des traitements algorithmiques. Or l'ouverture de la source d'un algorithme se heurte à la nécessité de protéger le droit d'auteur de son concepteur. Ce conflit entre ouverture de la source dans une perspective démocratique et droit d'auteur doit être résolu dès la conception d'algorithmes car chacune de ces positions implique un choix technologique distinct. Pour les grandes plateformes de réseaux sociaux, l'opacité est la norme. De nombreuses entreprises, soucieuses de conserver la maîtrise de leur architecture technique et le modèle économique lui étant associé, privilégient des solutions propriétaires. L'ouverture des sources est en revanche privilégiée par les acteurs de technologies civiques, afin d'obtenir la confiance de leurs usagers.

c. La non-applicabilité du principe d'ouverture aux données publiques touchant à la défense nationale ou aux secrets légaux

La loi pour une république numérique prévoit que l'ouverture des données «ne concerne ni les informations personnelles, ni celles touchant à la sécurité nationale, ni celles couvertes par les différents secrets légaux»⁸⁴. L'article L. 311-5 CRPA, exclut de la communicabilité des données publiques celles qui sont couvertes par le « secret des délibérations du gouvernement et des autorités responsables relevant du pouvoir exécutif » (2o a) ou par le « secret de la défense nationale » (2o b), celles qui concernent « la conduite de la politique extérieure de la France » (2o c) ou encore qui sont relatives « à la sûreté de l'État, à la sécurité publique ou à la sécurité des personnes » (2o d). La mission de sécurité nationale, consistant notamment en la protection de la nation contre les agressions armées, constitue une part importante de ces informations non communicables⁸⁵. Il s'agit ainsi d'éviter que l'ouverture de données publiques, entraînant la divulgation de données sensible, ne puisse faciliter la commission d'actes de terrorisme⁸⁶

⁸³ La qualité des métadonnées est cruciale pour contextualiser des jeux de données.

Charlotte Maday, « Qualité des données et Open data », *I2D – Information, données & documents*, p58.

⁸⁴ LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique, JORF n°0235 du 8 octobre 2016.

⁸⁵ Cette mission vise « l'ensemble des menaces et des risques susceptibles d'affecter la vie de la Nation, notamment en ce qui concerne la protection de la population, l'intégrité du territoire et la permanence des institutions de la République » l'article L. 1111-1 du code de la défense

⁸⁶ Bertrand Warusfel, « Enjeux et limites de l'ouverture des données en matière de sécurité et de défense », *Revue française d'administration publique*, vol. 167, no. 3, 2018, pp. 551-564

Pour conclure, les algorithmes constituent un instrument privilégié pour identifier l'intérêt général. Si l'on considère, dans une perspective démocratique, que seule la volonté générale est à même de saisir l'intérêt général, il faut permettre l'ouverture des données publiques et du code source des algorithmes utilisés par les décideurs publics, ceci afin de permettre aux données d'être intégrées à la délibération publique. Devenus lisibles, les algorithmes ne se substituent pas à la volonté des citoyens mais constituent pour eux une aide à la décision.

Le RGPD pose un principe de transparence et de responsabilité des usagers. Bien que la loi pour une république numérique ait substitué un principe d'ouverture des données d'intérêt général à l'ouverture à la demande des données ; l'ouverture des données publiques utilisées traitées par algorithmes à des fins d'identification de l'intérêt général demeure en pratique à la discrétion des représentants. En outre, pour que l'ouverture des données publiques et des codes sources des algorithmes permette un contrôle efficient de l'action publique, il s'agit d'en assurer l'intelligibilité. Or la lisibilité des algorithmes se heurte à des obstacles techniques, culturels et juridiques.

Le respect du principe démocratique amène à rechercher comment rendre exploitables les données publiques d'intérêt général. Sur un plan purement juridique, de nombreuses questions demeurent. En premier lieu, il faut définir comment rendre intelligible ces données, en particulier comment concilier lisibilité des données et protection des données personnelles des intéressés. Il faudrait ensuite déterminer dans quelle mesure l'ouverture des données publiques pourrait devenir contraignante.

Au-delà de l'utilisation des données publiques par les citoyens dans le cadre du débat public, il reste à déterminer quelles pourraient être les influences de leur ouverture sur l'équilibre des pouvoirs.