
*Transparence et action administrative à l'ère
de l'intelligence artificielle*

Hugo LAMI, Christian SAMANIEGO et Safia ZARAA

2026-5, pp. 91-115

Pour citer cet article :

Hugo LAMI, Christian SAMANIEGO et Safia ZARAA, « Transparence et action administrative à l'ère de l'intelligence artificielle », *Jurisdoct^oria* 2026-5 [en ligne], *Les Cahiers d'e-Délib* 2026, pp. 91-115

Pour consulter cet article en ligne :

<https://www.jurisdoct^oria.net/2026/07/transparence-et-action-administrative-a-lere-de-lintelligence-artificielle>

©JURISdoct^oria – ISSN 1760-6225

Transparence et action administrative à l'ère de l'intelligence artificielle

HUGO LAMI

Doctorant à l'Université de Lille

CHRISTIAN SAMANEIGO

Étudiant en Master 2 à l'Université de Lille

SAFIA ZARAA

Étudiante en Master 2 à l'Université de Lille

Dès 2018, Cédric Villani interrogeait la transparence des systèmes d'intelligence artificielle et alertait sur le fait qu'une « grande partie des considérations éthiques soulevées par l'IA tiennent à l'opacité de ces technologies. En dépit de leur performance accrue dans de nombreux domaines [...], il est souvent très difficile d'expliquer leurs décisions de manière intelligible par le commun des mortels »¹.

- 2 En 2025, l'intégration stratégique de l'intelligence artificielle (IA) dans les politiques publiques et la généralisation de son utilisation au sein de l'administration publique² offrent l'opportunité de renouveler les interrogations posées par Cédric Villani et de s'interroger sur le concept même de transparence dans le contexte de la gouvernance numérique.
- 3 Dans ce scénario, l'intelligence artificielle, entendue comme « un système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou

¹ C. VILLANI, *Donner un sens à l'intelligence artificielle*, 8 mars 2018.

² Comité interministériel de l'intelligence artificielle, *Faire de la France une puissance de l'IA*, 2025.

virtuels »³, se présente comme un outil ayant le potentiel de transformer la prestation de services publics et la prise de décision administrative grâce à ses capacités de calcul et d'analyse. Toutefois, cette transformation n'est pas sans poser de problèmes. Le caractère ambivalent de cette nouvelle technologie est souligné par la Convention cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit, dont le préambule reconnaît « que les activités menées dans le cadre du cycle de vie des systèmes d'intelligence artificielle peuvent offrir des opportunités sans précédents pour la protection et la promotion des droits de l'homme, de la démocratie et de l'État de droit », mais reste préoccupé « par le fait que certaines activités menées dans le cadre du cycle de vie des systèmes d'intelligence artificielle peuvent porter atteinte à la dignité humaine et à l'autonomie personnelle, aux droits de l'homme, à la démocratie et à l'État de droit ». C'est cette même opposition entre les bénéfices attendus et le risque que représentent les systèmes d'intelligence artificielle (SIA) qui a motivé l'intervention de l'Union européenne et l'édiction d'un Règlement dédié à l'IA (RIA)⁴. Le considérant 6 de ce texte précise que, « compte tenu de l'incidence majeure que l'IA peut avoir sur nos sociétés et de la nécessité de bâtir la confiance, l'IA et son cadre réglementaire doivent impérativement être élaborés dans le respect des valeurs de l'Union ».

- 4 L'un des défis cruciaux est la nécessité de concilier les potentiels apports de cette nouvelle technologie avec la maîtrise humaine de son fonctionnement. Cette conciliation se matérialise, pour le législateur européen, par la mise en œuvre des mesures de transparence par les entreprises qui développent et déploient des systèmes d'IA générative. Ces mesures comprennent la mise en place de systèmes d'évaluation des risques, de cadres d'expérimentations⁵ sous contrôle de la personne publique préalablement à la mise sur le marché des outils d'IA et l'élaboration de métriques pour évaluer les performances, la sécurité et la conformité éthique des systèmes d'IA dans divers contextes et applications.
- 5 La réflexion sur le rôle de l'IA dépasse le cadre économique et politique et pose également des enjeux moraux et éthiques. Si l'IA peut être un outil puissant pour aider les humains à prendre des décisions, il est crucial de se rappeler que les « technologies de l'IA peuvent rendre de grands services à l'humanité et que tous les pays peuvent en bénéficier »⁶ mais qu'elles « soulèvent également des préoccupations éthiques de fond »⁷ et rendent nécessaires « d'assurer la transparence et l'intelligibilité du fonctionnement des algorithmes et des données à partir desquelles ils ont été entraînés ; et leurs éventuelles conséquences sur, entre autres, la dignité humaine, les droits de l'homme et les libertés

³ Règlement européen sur l'intelligence artificielle, art. 3.

⁴ Règlement (UE) 2024/1689 du 13 juin 2024

⁵ En ce sens : La mise en place d'un « bac à sable réglementaire pour l'IA », à savoir « un cadre contrôlé mis en place par une autorité compétente qui offre aux fournisseurs ou aux fournisseurs potentiels de systèmes d'IA la possibilité de développer, de former, de valider et de tester, le cas échéant dans des conditions réelles, un système d'IA innovant, conformément à un plan de bac à sable, pendant une durée limitée et sous surveillance réglementaire » Article 3 RIA.

⁶ UNESCO, *Recommandation sur l'éthique de l'intelligence artificielle*, SHS/BIO/PI/2021/1, 2022.

⁷ *Ibid.*

fondamentales, l'égalité des genres et la démocratie ». Cette approche permettrait de garantir que l'IA est utilisée dans l'administration publique d'une manière qui profite à la société dans son ensemble, en renforçant la réglementation, tout en interrogeant la manière dont nous devrions comprendre les concepts traditionnels tels que la transparence, l'égalité ou la responsabilité.

- 6 La transparence constitue un élément essentiel des dispositifs de contrôle démocratique⁸. Elle se définit au sens commun comme « la qualité d'une institution qui informe complètement sur son fonctionnement, ses pratiques »⁹ et juridiquement comme « le droit du citoyen de savoir ce qu'ont fait, font ou vont faire les autorités administratives »¹⁰. Le principe de transparence trouve son fondement dans l'article 15 de la Déclaration des droits de l'homme et du citoyen de 1789, lequel prévoit que « la Société a le droit de demander compte à tout Agent public de son administration ».
- 7 À ce titre, Jean-Marc Sauvé, ancien vice-président du Conseil d'État, souligne que « la transparence permet la mise en œuvre effective des principes et des valeurs qui fondent l'action publique [...], [qu'] elle renforce ce faisant la confiance des citoyens dans les institutions [et] favorise également l'impartialité et l'objectivité du service public »¹¹.
- 8 Sans transparence, la légitimité de l'administration pourrait être compromise, affaiblissant la confiance des administrés. Ce lien de causalité est mis en exergue par les rapports d'ONG qui utilisent la transparence comme baromètre de la santé démocratique des États¹². De plus, la transparence joue un rôle central dans l'accès à la justice en facilitant la compréhension et la contestation des décisions perçues comme injustes ou discriminatoires.
- 9 Ce principe a connu un développement important dans les années 1980 et s'inscrit dans la volonté de rééquilibrer les rapports entre administration et administrés¹³. Cependant, l'avènement du numérique et des technologies basées sur des algorithmes a introduit de nouveaux défis. Les algorithmes, qui se définissent comme des suites d'étapes permettant d'obtenir un résultat à partir d'informations d'entrée, se déclinent en plusieurs types : les algorithmes fermés, entièrement définis par un code écrit ; les algorithmes apprenants, qui génèrent des opérations à partir de données d'entraînement pour atteindre un objectif ; et les algorithmes d'apprentissage profond, qui utilisent

⁸ B. FRYDMAN, « La transparence un concept opaque ? », *Journal des tribunaux*, 2007/6265, p. 300.

⁹ Centre National de Ressources Textuelles et Lexicales, « Transparence ».

¹⁰ D. TRUCHET cité par E. AUBIN, « La protection constitutionnelle de la transparence administrative », *Nouveaux cahiers du Conseil constitutionnel*, 2018, n° 59, p. 35.

¹¹ Exposé d'ouverture de Jean-Marc SAUVÉ lors du colloque organisé par Transparence International France « Culture du secret contre transparence sans limite : quel équilibre pour garantir l'intérêt général ? » à l'Assemblée nationale le 5 juillet 2011.

¹² Transparency International, indice de la perception de la corruption 2024.

¹³ E. AUBIN, op. cit., p. 35.

des réseaux de neurones artificiels nécessitant un grand volume de données, ce sont ces derniers qui forment le socle des systèmes d'intelligence artificielle.

- 10 Toutefois, le fonctionnement des systèmes d'IA est souvent comparé à une boîte noire opaque dont on connaîtrait les informations originelles, les résultats produits, mais où toutes les opérations intermédiaires, y compris le cheminement permettant de parvenir à un résultat, ne seraient pas accessibles. Ce phénomène de boîte noire concerne toutes les parties prenantes à l'IA, y compris leurs concepteurs¹⁴, ce qui a donné naissance à un champ de recherche dédié, l'*Explainable AI* (intelligence artificielle explicable, ou XAI). Ce domaine vise à rendre les systèmes algorithmiques plus transparents en les accompagnant d'explications compréhensibles, fidèles au raisonnement suivi, et en précisant les limites opérationnelles des SIA. Cette opacité soulève la question de la compatibilité entre l'usage croissant de l'IA et les exigences de transparence. Ainsi, bien que la publication des codes sources puisse sembler être une avancée, elle ne permet pas toujours de répondre aux enjeux d'intelligibilité et de motivation des décisions administratives. Comme l'a affirmé le Conseil d'État dans son rapport de 2021 sur l'intelligence artificielle et l'action publique, « l'opacité est destructrice de confiance »¹⁵.
- 11 Face à ces défis, le législateur et l'administration ont cherché à adapter les obligations de transparence aux nouveaux outils numériques. Ainsi, la Loi 2016-1321 du 7 octobre 2016 pour une République numérique (LRN) oblige l'administration à fournir des informations sur le recours à l'IA et les règles de base des algorithmes aux personnes concernées par une décision algorithmique et crée l'article L.311-3-1 du Code des relations entre le public et l'administration (CRPA) qui impose qu'une décision administrative fondée sur un traitement algorithmique mentionne explicitement l'usage d'un tel outil. En outre, le Règlement de l'Union européenne sur l'IA se propose « de répondre aux préoccupations liées à l'opacité et à la complexité de certains systèmes d'IA »¹⁶, notamment en fixant de nouvelles règles, propres aux systèmes d'IA, et susceptibles de favoriser une plus grande transparence entre les fournisseurs et les déployeurs. Enfin, la Commission d'accès aux documents administratifs (CADA) a également affirmé, sous réserve, le principe de communication libre et totale des algorithmes publics¹⁷.
- 12 Le Conseil constitutionnel a lui-même affirmé que le responsable du traitement doit être en mesure, à tout moment, de fournir aux personnes concernées par une décision algorithmique une explication détaillée et intelligible sur la manière dont le traitement a été mis en œuvre, notamment sur le fonctionnement de l'algorithme¹⁸.

¹⁴ Y. BENGIO, *International Scientific Report on the Safety of Advanced AI3*, 21 may 2024, p. 34.

¹⁵ Conseil d'État, *Intelligence artificielle et action publique : construire la confiance, servir la performance*, 31 mars 2022, p. 101.

¹⁶ Règlement sur l'intelligence artificielle, considérant 72.

¹⁷ Voir par ex. : CADA, avis n° 20240340 du 15 février 2024.

¹⁸ Conseil constitutionnel, décision n° 2018-765 DC, *Loi relative à la protection des données personnelles*, 12 juin 2018.

- ¹³ Toutefois, ces efforts visent davantage à rendre les moyens de l'administration transparents, c'est-à-dire le recours à un outil d'IA, qu'à expliciter pleinement le fonctionnement des algorithmes utilisés. En effet, les dispositifs mis en œuvre permettent de comprendre la source de la décision qui sera *in fine* adoptée, mais pas nécessairement le processus ayant conduit à la décision. Plus concrètement, les « règles définissant ce traitement ainsi que les principales caractéristiques de sa mise en œuvre »¹⁹ pourraient être des éléments justificatifs de la solution proposée par le système, mais ne pas permettre de retracer le raisonnement suivi de manière effective.
- ¹⁴ Enfin, des obstacles subsistent. Une transparence totale risquerait parfois de réduire l'efficacité des algorithmes utilisés par l'administration, par exemple en facilitant leur contournement par des fraudeurs. Par ailleurs, l'administration doit concilier l'utilisation de l'IA avec les impératifs d'efficacité. En effet, bien que les administrations adoptent l'intelligence artificielle pour accroître leur efficacité, son utilisation peut paradoxalement ralentir les performances des services publics, soit parce que le système est inefficace et que les décisions prises par lui nécessitent des corrections²⁰, soit parce que la formation des agents ne leur permet pas d'en faire un usage efficient²¹.
- ¹⁵ Si la transparence est indispensable pour garantir la légalité et la compréhension des décisions administratives, son adaptation au fonctionnement spécifique des intelligences artificielles reste un défi. Les autorités publiques et administratives doivent donc, malgré cette opacité inhérente, adapter le concept aux spécificités des intelligences artificielles.
- ¹⁶ Par ailleurs, le marché de l'intelligence artificielle est ainsi constitué que les outils utilisés par l'administration sont souvent développés par des tiers. Il faut ainsi ajouter à la transparence classique entre l'administration et l'administré un niveau supplémentaire de transparence, entre les fournisseurs de systèmes et l'administration. Le Règlement UE sur l'IA précise notamment que le terme de transparence « renvoie au fait que les systèmes d'IA sont développés et utilisés de manière à permettre une traçabilité et une explicabilité appropriées, faisant en sorte que les personnes réalisent qu'elles communiquent ou interagissent avec un système d'IA, que les dépoyeurs soient dûment informés des capacités et des limites de ce système d'IA et que les personnes concernées soient informées de leurs droits ». Les nouvelles modalités d'expression de la transparence sont donc multilatérales. La transparence doit aujourd'hui s'entendre autant dans son application classique, de l'administration vers les administrés, que dans sa nouvelle dimension, des concepteurs de SIA, qui jouent un rôle direct dans la prise de décisions administratives par les outils qu'ils fournissent, à leurs utilisateurs. Cette

¹⁹ CRPA, art. L311-3-1.

²⁰ Voir par exemple : France Info, Linh-Lan DAO, « Erreurs factuelles, absence de modération... Les débuts ratés de l'intelligence artificielle française Lucie, moquée sur les réseaux sociaux », 26 janvier 2025.

²¹ Le rapport sénatorial relatif à l'expérimentation de la vidéosurveillance algorithmique lors des Jeux de Paris relève « le caractère insuffisant de la formation des agents habilités à accéder aux signalements ».

nouvelle dimension est d'ailleurs mise en avant par le RIA, pour qui la transparence doit permettre « que les déployeurs soient dûment informés des capacités et des limites de ce système d'IA »²².

- 17 Cependant, cette adaptation des modes de mise en œuvre de la transparence soulève des interrogations. En effet, en acceptant qu'une partie du processus décisionnel soit inaccessible, peut-on encore parler de transparence ? Comment faire en sorte que le principe de transparence ne devienne pas purement symbolique et soit privé d'effet utile ? Tout ceci traduit le besoin d'un renouvellement de la concrétisation du principe de transparence dans la mesure où il ne s'agit plus uniquement de rendre accessibles les étapes du raisonnement, mais de garantir un certain niveau d'explicabilité, de compréhension pour l'administré autant que pour l'administration.
- 18 L'administration doit ainsi trouver un équilibre entre la protection des droits des administrés et les contraintes imposées par la complexité de l'opacité des algorithmes.
- 19 Dès lors, la transparence ne peut plus s'étendre comme la simple communication des documents administratifs, en l'occurrence codes sources, journaux d'enregistrement ou règles générales de fonctionnement, sans quoi sa finalité serait perdue de vue. Pour cette raison, il semble utile de questionner le sens, le contenu et les modalités de mise en œuvre de ce concept en prenant en compte les spécificités des algorithmes qui constituent aujourd'hui un outil d'aide à la décision pour les administrations et pourraient devenir, demain, un mode de décision courant. Dans ce contexte, il paraît nécessaire d'identifier les conditions dans lesquelles la « boîte noire » pourrait être ouverte. Cela suppose, en premier lieu, de comprendre les spécificités de l'IA qui rendent les systèmes qui l'utilisent opaques (I) et, en second lieu, d'analyser les solutions mises en place pour préserver le sens de la transparence dans un contexte algorithmique (II).

I. LA PRISE EN COMPTE DES SPECIFICITES DE L'IA DANS LA MISE EN ŒUVRE DE LA TRANSPARENCE

- 20 La notion de transparence semble renouvelée par l'intelligence artificielle pour au moins deux raisons. D'une part, cette notion, à l'origine verticale, qui implique que l'administration doit fournir aux destinataires de ses décisions une explication sur les raisons qui l'ont conduite à l'adopter, est aujourd'hui enrichie d'une troisième composante. L'administration n'est plus seulement émettrice de transparence et d'information, elle en est également une destinataire, puisque les fournisseurs des systèmes qu'elle emploie, même s'ils sont placés sous son contrôle, participeront à l'élaboration des décisions administratives. La transparence en matière algorithmique est donc une relation tripartite. Cette asymétrie d'information aboutit au fait que l'information dont l'administré est destinataire est imparfaite (A), par ailleurs, cette asymétrie d'information est doublée d'un phénomène de boîte noire

²² RIA, considérant 27.

qui rend difficile la compréhension, et donc l'explication, des décisions prises sur le fondement des outils d'IA de l'administration pour l'ensemble des parties prenantes (B).

A. Une transparence tripartite asymétrique

- 21 Si la transparence est traditionnellement comprise comme une relation unilatérale de l'administration vers les administrés dans laquelle l'information est rendue accessible, l'intervention de tiers dans le développement d'outils numériques et plus spécifiquement de l'IA et des algorithmes perturbe cet équilibre initial (1). Le développement interne de tels systèmes pourrait constituer un moyen de favoriser les objectifs de la transparence (2).

1. Le recours à des développements tiers, source d'opacité

- 22 Le développement de solutions d'IA étant généralement coûteux et long, la personne publique peut avoir recours à des tiers pour la fourniture de ces systèmes. C'est le choix qui a été effectué concernant l'algorithme de vidéosurveillance utilisé lors des Jeux olympiques de Paris 2024. Si le texte de loi prévoyait que « l'État assure le développement du traitement ainsi autorisé », il ajoutait que l'État pouvait aussi en confier le développement à un tiers ou l'acquérir. C'est cette dernière solution qui a finalement été retenue.
- 23 Cette pratique a l'avantage de doter les administrations d'outils prêts à l'emploi, dont elle n'a pas à assumer le coût de développement. Toutefois, le recours aux tiers a pour effet de créer une dépendance de l'administration à l'égard de son fournisseur. Celle-ci n'ayant pas participé à l'élaboration du système, il est possible qu'un certain nombre d'informations lui soient inconnues : soit parce qu'elles ne lui ont pas été communiquées, soit parce qu'elles sont inconnues même du fournisseur. À titre d'exemple, cette asymétrie informationnelle a pu être mise en avant dans le scandale *BriefCam*²³. Dans cette affaire, les forces de sécurité françaises ont acheté, auprès de la société israélienne *BriefCam*, un logiciel d'analyse vidéo capable de trier les données qui lui étaient fournies afin d'isoler les séquences vidéo d'intérêt pour les agents de police. La polémique s'est nouée autour de la fonctionnalité de reconnaissance faciale de ce logiciel. Or, cette fonctionnalité n'était pas intégrée à la solution lors de son achat par l'Administration. Elle a été ajoutée au cours de la durée de vie du SIA, sans que les

²³ G. LIVOLSI, M. DESTAL, C. LE FOLL, « La police nationale utilise illégalement un logiciel israélien de reconnaissance faciale », *Disclose*, 14 novembre 2023.

forces de l'ordre en soient informées par le développeur²⁴, avec comme conséquence de les placer hors-la-loi²⁵.

- ²⁴ Dans ces conditions, la transparence doit s'entendre comme une relation multilatérale dans laquelle l'administration n'est plus maîtresse des horloges ni unique détentrice de l'information, mais occupe tout de même un rôle central : à la fois source et destinataire de la transparence.
- ²⁵ Le Règlement européen sur l'IA a pris acte de cette nouvelle répartition. Il structure le marché de l'IA autour de trois acteurs principaux : le fournisseur, dont le rôle est de développer un outil d'IA, le déployeur, qui est une personne physique ou morale utilisant le SIA hors l'utilisation personnelle ou non professionnelle, et enfin la personne concernée par le résultat produit par le SIA.
- ²⁶ Le règlement fait du fournisseur, entendu comme « une personne physique ou morale, une autorité publique, une agence ou tout autre organisme qui développe ou fait développer un système d'IA ou un modèle d'IA à usage général et le met sur le marché ou met le système d'IA en service sous son propre nom ou sa propre marque, à titre onéreux ou gratuit »²⁶, l'acteur principal de la transparence.
- ²⁷ Dans le même temps, les personnes concernées par les traitements algorithmiques semblent en marge des préoccupations du texte, car, si celui-ci promeut une utilisation de l'IA « garantissant un niveau élevé de protection de la santé, de la sécurité et des droits fondamentaux consacrés dans la Charte des droits fondamentaux de l'Union européenne »²⁷, le corps de son texte ne consacre que deux articles relatifs aux droits des personnes concernées par le traitement. L'article 85 permet ainsi à toute personne « ayant des motifs de considérer qu'il y a eu violation des dispositions du présent règlement [de] déposer des réclamations auprès de l'autorité de surveillance du marché concernée ». L'article 86, quant à lui, prévoit « le droit d'obtenir du déployeur des explications claires et pertinentes sur le rôle du système d'IA dans la procédure décisionnelle et sur les principaux éléments de la décision prise » lorsqu'un système à haut risque « qui produit des effets juridiques ou affecte significativement cette personne de façon similaire d'une manière qu'elle considère comme ayant des conséquences négatives sur sa santé, sa sécurité ou ses droits fondamentaux » a été utilisé.
- ²⁸ La centralité du fournisseur du système d'IA dans la réalisation de la transparence s'apprécie au nombre de dispositions qui lui sont consacrées par le Règlement sur l'intelligence artificielle. La section 3 du Chapitre III du règlement, intitulée « Obligations incombant aux fournisseurs et aux déployeurs de systèmes d'IA à haut risque et à d'autres parties », comporte sept articles relatifs aux obligations des fournisseurs contre deux pour celles incombant aux déployeurs.

²⁴ IGA, IGP, IGG, « Usage de logiciels d'analyse vidéo par les services de la police et de la gendarmerie nationale », 28 octobre 2024, p. 34.

²⁵ CNIL, Décision MED-2024-150, 15 novembre 2024.

²⁶ RIA, art. 3.

²⁷ RIA, considérant 1.

- 29 Ainsi, le texte impose au fournisseur un certain nombre d'obligations qui ont pour objectif de permettre au déployeur de comprendre le système qu'il utilise, c'est par exemple sur lui que repose la responsabilité d'assurer la transparence du SIA au sens du règlement, puisque son article 13 dispose que « la conception et le développement des systèmes d'IA à haut risque sont tels que le fonctionnement de ces systèmes est suffisamment transparent pour permettre aux déployeurs d'interpréter les sorties d'un système et de les utiliser de manière appropriée ». Cette exigence de transparence se matérialise, d'abord, par la mise à disposition, par le fournisseur et à destination du déployeur, d'une documentation technique adéquate. Cette documentation technique doit, par exemple, rendre compte des conditions dans lesquelles le système est fait pour être utilisé : sa destination, son niveau d'exactitude ainsi que ses mauvaises utilisations connues.
- 30 Dans ces conditions, l'administration, en tant que déployeur, serait bien destinataire de la transparence. Le fournisseur doit mettre à sa disposition les informations qui lui permettront, d'abord, de faire une bonne utilisation du système et, le cas échéant, d'écarter ou au moins de questionner les résultats produits par le SIA.
- 31 Une autre preuve du rôle prépondérant du fournisseur dans la mise en place des conditions de la transparence réside dans le fait qu'il lui revient d'assurer la conformité du système au règlement avant son utilisation. Ainsi, tout l'environnement qui entoure le fonctionnement opérationnel de l'outil qu'il fournit est prévu par lui-même en amont de son utilisation. C'est par exemple le cas des systèmes de gestion des risques, prévus par l'article 9 du RIA, qui impose au fournisseur de prévoir les procédures capables de prévenir et gérer les risques induits par son système avant son déploiement.

2. Le développement interne des SIA, possible source de transparence

- 32 Pourtant, l'intervention d'un tiers dans la fourniture de logiciel, cause de l'asymétrie d'information, n'est pas une fatalité. L'État pourrait employer ses moyens à développer des solutions d'IA internes, ce qui lui assurerait un contrôle plus grand sur les conditions d'élaboration et de fonctionnement des SIA qu'il utilise. Le rapport parlementaire relatif à la sécurité lors des Jeux olympiques de Paris recommande d'ailleurs, dans le cas où la vidéoprotection algorithmique venait à être pérennisée, d'envisager « le développement d'une solution publique très encadrée de vidéoprotection algorithmique »²⁸.
- 33 Une telle solution est déjà développée et déployée au sein de certaines administrations à la seule fin d'aider les agents à mieux répondre aux interrogations des administrés. L'IA Albert, développée par

²⁸ Assemblée Nationale, *Rapport d'information sur le bilan des jeux olympiques et paralympiques dans le domaine de la sécurité*, 19 mars 2025.

la DINUM (Direction interministérielle du Numérique), n'a donc pas vocation à édicter des décisions administratives, mais à informer et conseiller les personnes en charge de produire de tels actes.

- 34 Toutefois, et au regard des expérimentations françaises en matière de numérique, il est permis de douter de l'existence d'une solution efficace et pérenne à court terme compte tenu de son coût financier. L'exemple du stockage des données de santé par des prestataires étasuniens illustre la difficulté de l'État français à se doter de solutions numériques qualifiées de « souveraines » : lancé en 2019, le Health Data Hub, est toujours hébergé par l'entreprise Microsoft en 2025. Le ministre de la Santé de l'époque, Olivier Véran, s'était pourtant engagé dès 2020 à ce qu'une solution européenne soit substituée au cloud de Microsoft dans un délai ne dépassant pas deux ans²⁹.
- 35 En tout état de cause, le recours aux tiers et la multilatéralisation de la transparence qu'elle implique, qui s'accompagne d'une opacité inhérente aux SIA évoquée en introduction, rendent nécessaire le renouvellement des conditions d'exercice de la transparence.

B. La transparence à l'épreuve de « la boîte noire »

- 36 Si le mode de fonctionnement des algorithmes est souvent obscur, la transparence demeure un principe fondamental de l'action administrative. Dans cette perspective, il convient d'abord d'analyser la transparence comme principe fondamental des décisions administratives (1), avant d'envisager son rôle central en tant que garantie du principe de non-discrimination dans les décisions algorithmiques (2).

1. La transparence comme principe fondamental des décisions administratives

- 37 Classiquement, la transparence administrative se traduit par l'accès des citoyens aux documents administratifs et par la publication d'informations sur les processus décisionnels³⁰. L'objectif de l'introduction de ce principe était de mettre fin au traditionnel secret administratif, lequel a été remis en question par le biais de différentes réformes au cours des années 1970³¹. À cet égard, la loi du 17 juillet 1978 garantit le libre accès aux documents administratifs, tandis que la loi du 11 juillet 1979 impose la motivation des actes administratifs et améliore les relations entre l'administration et le public, en permettant aux citoyens d'être informés des motifs des décisions administratives défavorables les concernant. Cette approche vise à garantir un contrôle démocratique et à renforcer la confiance entre l'Administration et le public.

²⁹ J. HOURDEAUX, « Health Data Hub : Véran s'engage à retirer l'hébergement à Microsoft d'ici " deux ans " », *Mediapart*, 22 novembre 2020.

³⁰ J. MARCHAND, « Réflexion sur le principe de transparence », *Revue de droit public*, 2014, pp. 677 à 703

³¹ B. DELAUNEY, *Le secret administratif*, Titre VII, n° 10, 2023.

- ³⁸ Cependant, à l'ère des décisions automatisées, la définition traditionnelle apparaît désormais insuffisante. Les systèmes d'intelligence artificielle induisent une opacité algorithmique où les décisions reposent sur des modèles complexes, souvent incompréhensibles pour l'administration, les administrés et, parfois, pour les concepteurs eux-mêmes. Il semblerait donc que les mécanismes d'accès aux informations tels que nous les connaissons aujourd'hui ne permettent plus de répondre aux attentes des citoyens, qui demandent désormais transparence et, parallèlement, explicabilité des décisions administratives adoptées. En effet, la mise à disposition des codes sources est « loin d'avoir réglé les craintes des citoyens quant à la manière dont ils sont administrés »³². Ainsi, même si les codes sources sont rendus accessibles, cela ne suffit pas à apaiser les inquiétudes des citoyens concernant la manière dont leurs données ou informations sont utilisées pour la prise de décision.
- ³⁹ La transparence ne peut donc se limiter à rendre publiques des données, mais doit inclure des dispositifs capables d'expliquer les logiques sous-jacentes des algorithmes³³ et d'assurer aux destinataires une certaine loyauté. Ainsi, l'article 22 du règlement général sur la protection des données (RGPD), prévoit que dans le cas d'une décision automatisée, le responsable du traitement doit mettre « en œuvre des mesures appropriées pour la sauvegarde des droits et libertés et des intérêts légitimes de la personne concernée ». Ces mesures appropriées doivent permettre à la personne concernée « d'exprimer son point de vue et de contester la décision ». Une partie de la doctrine défend donc l'existence d'un droit à l'explication, qui va bien plus loin qu'un simple droit à l'information, concrétisation du principe de transparence. Dans l'arrêt de la Cour de justice de l'Union européenne du 7 décembre 2023³⁴ concernant les décisions individuelles automatisées, l'avocat général défendait l'obligation imposée par une lecture conjointe des articles 15 et 22 du RGPD de délivrer des « explications suffisamment détaillées sur la méthode utilisée pour le calcul du score et les raisons qui ont conduit à un résultat déterminé »³⁵.
- ⁴⁰ Avec l'émergence des algorithmes, le principe de transparence est désormais indissociable des principes d'explicabilité et de loyauté. Ces principes sont intrinsèquement liés, en ce qu'une décision ne peut être perçue comme transparente que si elle est compréhensible – il s'agit donc du droit à l'explicabilité – et loyale, uniquement si les critères sur lesquels se fonde l'algorithme, et, par conséquent l'Administration, sont identifiables par les destinataires. En effet, l'objectivité supposée des algorithmes est souvent mise en avant pour justifier leur utilisation, les présentant comme des outils impartiaux, exempts des biais cognitifs et émotionnels propres aux êtres humains. Pourtant, cette perception occulte une réalité essentielle : les algorithmes ne sont pas neutres, ils sont le reflet

³² D. MERIGOUX, M. ALAUZEN, J. BANULS *et al.*, « De la transparence à l'explicabilité automatisée des algorithmes : comprendre les obstacles informatiques, juridiques et organisationnels », *INRIA Paris*, R-953, 2024, p. 3.

³³ *Ibid.*

³⁴ CJUE, *SCHUFA Holding (Scoring)*, aff. C-634/21, 7 décembre 2023.

³⁵ Conclusions de l'avocat général M. PRIIT PIKAMAE, présentées le 16 mars 2023, aff. C-634/21.

des choix de conception et des données sur lesquelles ils reposent³⁶. Dès lors, la transparence ne peut se limiter à la simple divulgation des codes sources, mais doit également permettre de comprendre les logiques sous-jacentes à leur fonctionnement afin d'assurer l'explicabilité et la loyauté des décisions administratives³⁷.

- ⁴¹ Ce principe est toutefois variable en fonction du créateur de l'algorithme, du type d'algorithme utilisé, de la complexité de celui-ci, mais également de l'expertise des destinataires³⁸. Bien que l'algorithme puisse être créé par l'Administration, dans les faits, cette dernière dispose rarement des ressources et de l'expertise nécessaires à la réalisation d'un tel travail. C'est pour cette raison, comme dit précédemment, qu'elle externalise très souvent à des entités privées la création de ces algorithmes. Pour autant, l'Administration et le fournisseur doivent être capables d'expliquer le raisonnement permettant d'arriver à la solution. Cette externalisation soulève de nouveaux enjeux. En effet, dans ce cadre, le principe de transparence se heurte aux droits de la propriété intellectuelle, le code source étant protégé par le droit d'auteur³⁹, il ne peut être librement divulgué en ce que sa communication pourrait porter atteinte aux droits des concepteurs, limitant ainsi l'accès à l'information. Il en résulte une tension entre la nécessité de transparence et les impératifs de protection de la propriété intellectuelle.

2. La transparence comme garantie du principe de non-discrimination dans les décisions algorithmiques

- ⁴² Pour que soient respectés le principe de transparence et, par conséquent, les principes de loyauté et d'explicabilité, encore faut-il s'assurer que l'algorithme, qu'il soit créé par l'Administration ou par un tiers, ne comporte pas de biais discriminatoires. Selon Jean-Bernard Auby⁴⁰, l'enjeu résiderait dans le contrat. Il reviendrait aux cocontractants et, notamment, à l'Administration de veiller au contrôle de la création de l'algorithme, afin que celui-ci n'intègre pas des biais discriminatoires⁴¹. C'est la connaissance de l'Administration tout au long du processus d'élaboration qui lui permettra de comprendre le fonctionnement de l'algorithme et ainsi de pallier les effets néfastes des boîtes noires⁴². Par ailleurs, la doctrine propose, notamment pour les algorithmes présentant des risques de discrimination, qu'ils fassent l'objet de contrôles, par des organes externes – on peut imaginer par une

³⁸ C. RICHARD, « Dans la Boîte Noire des Algorithmes. Comment Nous Nous Sommes Rendus Calculables », *Revue du Crieur*, 2018

³⁷ *Ibid.*

³⁸ J.-B. AUBY, « Algorithmes publics. Utilisation d'algorithmes dans l'Administration », *JurisClasseur administratif*, Lexis 360 Intelligence, 2025.

³⁹ Article L112-2 du Code de la propriété intellectuelle.

⁴⁰ J.-B. AUBY, 2025, *op. cit.*

⁴¹ *Ibid.*

⁴² *Ibid.*

autorité administrative indépendante – durant leur élaboration notamment dans le cadre d'audits ⁴³. L'enjeu concernant les algorithmes auto-apprenants est tout autre. En effet, ces algorithmes « évoluent sans intervention humaine directe en s'enrichissant de ce qu'il[s] découvre[nt] dans les contextes dans lesquels on l'applique »⁴⁴. Il appartient à l'Administration de garantir la transparence et l'explicabilité tout au long du cycle de vie de l'algorithme, en intégrant des mécanismes de contrôle pour veiller à ce que les évolutions internes respectent le principe de transparence et ses corollaires que sont les principes d'explicabilité et de loyauté⁴⁵. La doctrine indique que l'explication devrait d'abord être demandée au déployeur, non au fournisseur du système d'IA. L'obligation d'explicabilité pèserait donc sur le déployeur de l'IA, en l'espèce l'administration et non sur son concepteur. Pour autant, si l'explicabilité pèse effectivement sur l'Administration, encore faut-il qu'elle soit en mesure de comprendre l'algorithme qu'elle utilise. Le concepteur est donc indirectement concerné par cette obligation de transparence.

⁴³ Si le principe de transparence tend à être adapté aux nouvelles exigences induites par l'utilisation de l'IA, force est de constater que la reconnaissance de cette obligation a pris du temps. En témoignent les algorithmes dits « locaux ». Ces algorithmes (tels que Parcoursup) utilisés par les établissements d'enseignement supérieur pour le classement des candidatures et la répartition de vœux ont été progressivement soumis à des obligations de transparence. Ces évolutions trouvent leur source dans la saisine de la juridiction administrative par l'UNEF (Union Nationale des Étudiants de France) en 2019 afin de demander l'annulation de la décision implicite de refus de l'Université des Antilles de communiquer son algorithme local. Si le tribunal administratif avait fait droit au recours de l'UNEF, le Conseil d'État, saisi en cassation, a cassé ce jugement au motif que « que seuls les candidats sont susceptibles de se voir communiquer les informations relatives aux critères et modalités d'examen de leurs candidatures ainsi que les motifs pédagogiques qui justifient la décision prise »⁴⁶.

⁴⁴ Toutefois, le juge de cassation note l'évolution du droit positif intervenue postérieurement au jugement de première instance selon laquelle « chaque établissement est désormais tenu de publier les critères généraux encadrant l'examen des candidatures par les commissions d'examen des vœux en application de l'article D. 612-1-5 du code de l'éducation dans sa rédaction résultant du décret du 26 mars 2019 ». Cette décision intervient après l'avis du Défenseur des droits, qui avait formulé des recommandations contraires. En effet, l'autorité administrative indépendante avait dénoncé l'absence de transparence et appelé à des modifications législatives visant à « rendre publiques toutes les informations relatives au traitement, y compris algorithmique, et à l'évaluation des dossiers des

⁴³ Ibid.

⁴⁴ P. TIFINE, *L'utilisation des algorithmes dans la procédure d'inscription dans l'enseignement supérieur – Approche juridique*, Projectics / Proyéctica / Projectique, 2023.

⁴⁵ J-B. AUBY, 2025, *op. cit.*

⁴⁶ Conseil d'État, n°427916, 12 juin 2019.

candidats par les commissions locales des établissements d'enseignement supérieur en amont du processus de leur affectation dans les formations de premier cycle de l'enseignement supérieur, afin d'assurer la transparence de la procédure et de permettre aux candidats d'effectuer leurs choix en toute connaissance de cause »⁴⁷.

- ⁴⁵ Les conditions de communication des codes sources de Parcoursup ne feront explicitement mention des traitements algorithmiques qu'en 2021, puisque depuis cette date et jusqu'à aujourd'hui, l'article D 612-1-5 du code de l'éducation dispose que sont publiés « sous la forme de rapport, des critères en fonction desquels les candidatures ont été examinées l'année précédente et précisant, le cas échéant, dans quelle mesure des traitements algorithmiques ont été utilisés pour procéder à cet examen ».
- ⁴⁶ Ainsi, après le cas de Parcoursup, une autre affaire a récemment attiré l'attention sur les mêmes problématiques de transparence et de biais algorithmiques : celle des algorithmes utilisés par la CNAF. En effet, le Conseil d'État a été saisi le 15 octobre 2024 au sujet des algorithmes de la CNAF (Caisse Nationale des Allocations Familiales) témoigne également des questionnements en matière de transparence que soulève le recours aux décisions automatisées au sein de l'Administration. Dans cette affaire, l'algorithme utilisé pour déterminer celles et ceux pouvant faire l'objet d'un contrôle est critiqué pour son opacité et ses potentiels biais discriminatoires illégaux⁴⁸. L'absence de transparence sur les critères utilisés par l'algorithme empêche les citoyens de comprendre comment leurs droits sont évalués. En effet, après avoir obtenu l'accès au code source d'une ancienne version de l'algorithme utilisé, les associations ont révélé que certains critères augmentaient un « score de suspicion ». Ce score représente le risque qu'un allocataire soit coupable d'une fraude, et incluait des facteurs tels que de faibles revenus, le chômage, le bénéfice du revenu de solidarité active (RSA) ou de l'allocation adulte handicapé (AAH)⁴⁹. En réponse, elles ont demandé à la CNAF de cesser l'utilisation de cet algorithme, dénonçant son manque de transparence et ses possibles biais discriminatoires. Cette affaire met en lumière l'importance de garantir la transparence des décisions administratives en ce que l'utilisation de ces technologies, en raison des biais qu'elles peuvent reproduire, fragilise l'égal accès des citoyens aux services publics. En effet, le principe d'égalité devant le service public garantit à tous les citoyens placés dans une situation similaire d'être traités de manière similaire⁵⁰. Ce principe a pour corollaire celui de non-discrimination, prévu à l'article 14 de la Convention européenne des droits de l'Homme.

⁴⁷ Défenseur des droits, Décision 2019-021 du 18 janvier 2019 relative au fonctionnement de la plateforme nationale de préinscription en première année de l'enseignement supérieur (Parcoursup), 18 janvier 2019.

⁴⁸ A. SENECA, « Algorithme de ciblage antifraude dans les CAF : des associations saisissent le Conseil d'État », *Le Monde*, 16 octobre 2024.

⁴⁹ H. PAULIAT, « L'intelligence artificielle et les risques de rupture d'égalité dans l'accès aux droits et aux services publics », *RFDA*, 5 mars 2025.

⁵⁰ Conseil d'État, 10 mai 1974, publié Lebon, n° 88032 88148.

- ⁴⁷ Le 15 janvier 2026, la CNAF a pris la décision de publier le code source de cet algorithme⁵¹, une explication détaillée de son fonctionnement⁵² ainsi qu'une Charte éthique⁵³ relative à l'utilisation d'outils algorithmiques. Le Conseil d'État n'a pas encore rendu sa décision sur l'affaire des algorithmes de la CNAF, mais celle-ci pourrait avoir des répercussions importantes sur la manière dont les Administrations devront utiliser les technologies automatisées pour prendre des décisions. Il lui appartiendra notamment de déterminer le caractère discriminatoire des 17 variables sur lesquelles se fonde le système pour produire ses résultats⁵⁴.
- ⁴⁸ Ainsi, la transparence administrative, traditionnellement perçue comme une relation unilatérale entre l'Administration et les administrés, s'est transformée en une relation impliquant un nouvel acteur : les fournisseurs et développeurs des systèmes d'IA. Cette évolution a mis en lumière de nouvelles problématiques, notamment celle de l'opacité des algorithmes. Dès lors, garantir une transparence effective ne peut se limiter à la simple diffusion d'informations : elle suppose également la mise en place d'un cadre juridique permettant d'assurer l'explicabilité et la loyauté des systèmes d'IA. L'exemple de l'algorithme de la CNAF, dont la communication à l'utilisateur ne lui permet pas d'infirmer ou de confirmer le caractère discriminatoire est, de ce point de vue, parlant. C'est dans cette perspective que se construit actuellement un cadre adapté à la réalisation de la transparence algorithmique.

II. UN ENCADREMENT NORMATIF EN COURS DE CONSTRUCTION

- ³¹ Partant du constat des nombreux obstacles à la conciliation d'une transparence traditionnelle et du fonctionnement d'un SIA, et du choix assumé des administrations de faire de l'IA un nouvel outil administratif, un nouveau cadre normatif relatif à la conciliation de ces deux impératifs se construit. Ce nouvel encadrement s'illustre notamment par la réglementation européenne relative à l'intelligence artificielle, qui organise une typologie des SIA fondée sur le risque et adapte l'encadrement de ces systèmes en conséquence (A), mais qui laisse perplexe sur les conditions concrètes de transparence et notamment sur le contenu du droit à l'explicabilité qu'elle garantit (B).

⁵¹ CAF, *Modèle de datamining « données entrantes »*, 15 janvier 2026.

⁵² CAF, *présentation détaillée du « datamining données entrantes »* 2026, janvier 2026.

⁵³ CAF, *La Charte éthique sur l'usage des algorithmes et de l'intelligence artificielle de la Cnaf*, 15 janvier 2026.

⁵⁴ La Quadrature du Net, « Algorithme discriminatoire de notation de la CNAF : 10 nouvelles organisations se joignent à l'affaire devant le Conseil d'État », 20 janvier 2026.

A. Des obligations proportionnées au risque que représente les SIA

- 49 Pour faire face aux défis posés par les SIA, la réglementation européenne adopte une approche fondée sur le risque que représentent les usages des SIA (1), cette approche se veut protectrice des droits fondamentaux tout en préservant des possibilités d'innovation (2).

1. Transparence et classification : l'objectif d'une régulation basée sur le risque

- 50 Devant les progrès considérables des systèmes d'IA, le législateur européen a fait le choix de les classer selon le risque qu'ils représentent et d'adapter en conséquence les obligations auxquelles ils sont soumis. Dans ce sens, le législateur a défini le risque à l'article 3 du RIA comme « la combinaison de la probabilité d'un préjudice et de la sévérité de celui-ci ». Le règlement européen sur l'IA fixe, en son article 6, la grille permettant de classer, ou non, un système comme représentant un risque élevé. Ainsi, un SIA sera à haut risque si :

- Il « est destiné à être utilisé comme composant de sécurité d'un produit couvert par la législation d'harmonisation de l'Union ». C'est par exemple le cas des SIA destinés à être des composants de sécurités des ascenseurs⁵⁵ ou de certains jouets destinés à des enfants de moins de 14 ans⁵⁶.
- Il est « utilisé dans un domaine et pour une finalité listée par l'annexe III du Règlement » (par exemple, les « systèmes d'IA destinés à être utilisés par les autorités publiques ou en leur nom pour évaluer l'éligibilité des personnes physiques aux prestations et services d'aide sociale essentiels » seront classés à haut risque).

- 51 Dans ce contexte, la gouvernance numérique exige la mise en œuvre de mesures de transparence solides visant à renforcer la confiance du public dans l'utilisation de ces technologies. Ces mesures comprennent l'atténuation des risques associés à l'utilisation de systèmes d'IA à haut risque en établissant un cadre réglementaire pour leur développement et leur application.

- 52 La nécessité de telles mesures est d'autant plus évidente lorsqu'on considère que « les systèmes d'IA pourraient avoir un impact négatif sur la santé et la sécurité des personnes »⁵⁷. Dans ce sens, le règlement sur l'intelligence artificielle met en avant l'importance « que les risques de sécurité susceptibles d'être générés par un produit dans son ensemble en raison de ses composants numériques, y compris les systèmes d'IA, soient dûment prévenus et atténués »⁵⁸. En effet, le législateur européen a insisté sur le fait que « l'ampleur de l'impact négatif du système d'IA sur les droits fondamentaux protégés par la Charte revêt une importance particulière lorsqu'il s'agit de classer un système d'IA

⁵⁵ Directive 2014/33/UE du Parlement européen et du Conseil du 26 février 2014 relative à l'harmonisation des législations des États membres concernant les ascenseurs et les composants de sécurité pour ascenseurs.

⁵⁶ Directive 2009/48/CE du Parlement européen et du Conseil du 18 juin 2009 relative à la sécurité des jouets.

⁵⁷ Règlement sur l'intelligence artificielle, considérant 47.

⁵⁸ *Ibid.*

dans la catégorie des systèmes à haut risque »⁵⁹. Le législateur a donc fait le choix d'attacher à certains usages de l'IA le caractère risqué par nature, notamment dans les domaines de la biométrie, l'éducation, l'emploi, les prestations publiques, la gestion des migrations et l'administration de la justice. Une approche qui, bien que nécessaire, pourrait s'avérer insuffisante face à l'évolution rapide de ces technologies.

- 53 Cette dynamique était déjà observable avant l'adoption du Règlement sur l'intelligence artificielle, comme en témoigne la classification de certains systèmes à haut risque. Le système espagnol VioGén pour la prévention des violences de genre est un exemple d'application de l'IA à haut risque depuis son utilisation en 2007, qui met en évidence à la fois les avantages et les préoccupations en matière de protection des droits fondamentaux⁶⁰. En effet, « lorsque le dépôt de plainte est effectué – par la victime, un membre de sa famille, un témoin ou un agent – il est immédiatement enregistré dans le système VioGén (...) Sur la base des résultats combinés des questions et de l'enquête policière, le système détermine l'un des cinq niveaux de risque suivants : non évalué, faible, moyen, élevé et extrême »⁶¹. Ce système, utilisé par les forces de police, est employé dans le but d'évaluer le niveau de risque auquel pourrait être confrontée une victime de violence de genre au moment du dépôt de sa plainte. Un tel fonctionnement a soulevé des préoccupations quant aux droits fondamentaux, puisque les mesures de protection en faveur de la victime sont conditionnées par l'évaluation du risque effectuée par le système. En réponse aux préoccupations exprimées, le gouvernement a annoncé des modifications de l'algorithme visant à réduire « le risque d'erreur dans l'évaluation réalisée »⁶² qui impliquent notamment une prise en charge prioritaire par l'algorithme des cas les plus sensibles (nouveau dépôt relatif à un même cas, présence d'enfants...) ou la disparition de la catégorie « risque non apprécié » qui pouvait laisser penser qu'une potentielle victime n'encourait pas un risque particulier alors qu'il signifiait que sa requête n'avait pas été traitée.
- 54 Des recherches ont déjà exploré en 2023 le concept de systèmes d'IA à haut risque et leur réglementation. Ainsi, le Règlement sur l'intelligence artificielle introduit un cadre fondé sur le risque pour régir l'IA, avec des exigences strictes pour les systèmes à haut risque qui pourraient avoir un impact sur les droits fondamentaux⁶³ en intégrant l'évaluation et la gestion des risques. Pour faciliter l'évaluation des risques et la conformité aux réglementations, les chercheurs ont développé des

⁵⁹ Règlement sur l'intelligence artificielle, considérant 48.

⁶⁰ E. FALLETTI, « Human Rights Protection and High-Risk AI Systems: The Spanish Model in Gender-Based Violence Prevention », SSRN, 2022.

⁶¹ Ministère de l'intérieur Espagnol, *El Sistema VioGén supera los seis millones de valoraciones de riesgo a víctimas de violencia de género*, 19 mai 2023.

⁶² La Moncloa, *Interior diseña un nuevo modelo de respuesta policial a la violencia de género*, 15 janvier 2025.

⁶³ J.H GOLPAYEGANI, PANDIT, D. LEWIS, « To Be High-Risk, or Not To Be—Semantic Specifications and Implications of the AI Act's High-Risk AI Applications and Harmonised Standards », *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*, 2023.

paramètres de description (tels que AIRO “Risk Ontology”) qui aident à identifier les systèmes d’IA à haut risque, à documenter les informations sur les risques et à réaliser des évaluations d’impact basées sur le Règlement sur l’intelligence artificielle et les normes harmonisées⁶⁴. Cela étant, avec le développement des technologies IA, le législateur européen a établi dans le chapitre III du Règlement un cadre spécifique pour les exigences applicables aux systèmes d’intelligence artificielle. Il incombe donc aux fournisseurs de mettre en place un système de gestion des risques, de respecter les règles relatives aux données personnelles et de fournir une documentation technique complète. Les déployeurs doivent recevoir des documents d’information leur permettant de comprendre les résultats produits par le système et d’identifier clairement le fournisseur⁶⁵. En conséquence, conformément aux règles de transparence, le cadre européen exige que les documents fournis contiennent des informations sur les capacités et les limites du système, ainsi que sur les risques associés et les instructions destinées à interpréter les résultats du système.

2. Concilier transparence et protection des droits fondamentaux dans une régulation basée sur le risque

- ⁵⁵ Face aux défis que présentent les SIA, le manuel du Conseil de l’Europe⁶⁶ recommande des mesures de transparence et de contrôle pour ces systèmes, notamment la possibilité pour les personnes concernées de demander des explications sur le processus décisionnel ; l’inscription des systèmes d’IA utilisés dans des registres publics ; le maintien d’un accès non numérique aux services publics pour les personnes vulnérables, voire le droit de solliciter un contrôle humain des décisions⁶⁷.
- ⁵⁶ Ainsi, la plupart des chartes et recommandations émises au niveau international préconisent de développer des systèmes d’IA responsables, centrés sur l’humain et respectueux des droits fondamentaux⁶⁸. Les systèmes d’IA à haut risque présentent des risques significatifs pour les utilisateurs, la société ou l’environnement, en particulier dans des domaines critiques, tels que la santé, les transports, l’application de la loi et l’emploi⁶⁹. Néanmoins, la classification de ces systèmes d’IA selon les catégories de risque définies par le règlement sur l’IA montre, en Allemagne, que sur 514 cas de SIA analysés, seuls 30 % des systèmes d’IA seraient encadrés par la réglementation européenne

⁶⁴ D. GOLPAYEGANI, H. J. PANDIT, D. LEWIS, « AIRO: An Ontology for Representing AI Risks Based on the Proposed EU AI Act and ISO Risk Management Standards », *International Conference on Semantic Systems*, 2022.

⁶⁵ L. GYRNBAUM, « IA ACT : une définition des systèmes d’IA et une régulation entre protection et soutien à l’innovation », *La semaine juridique : édition générale* n° 38, 2024.

⁶⁶ Conseil de l’Europe, *l’administration et vous*, mai 2024.

⁶⁷ G. MARTI, L. CLUZEL-METAYER et S. MERABET, *Droit et Intelligence artificielle*, Lexis 360 Intelligence, 2020.

⁶⁸ Voir par exemple : UNESCO, *Recommandation sur l’éthique de l’intelligence artificielle*, 2022 ; Conseil de l’Europe, *Convention-cadre du Conseil de l’Europe sur l’intelligence artificielle et les droits de l’homme, la démocratie et l’État de droit*, 2024.

⁶⁹ N. DÍAZ-RODRÍGUEZ, J. DEL SER, M. COECKELBERGH et al., « Connecting the dots in trustworthy Artificial Intelligence », *Inf. Fusion*, vol. 99, 2023.

comme étant à haut risque et environ 7 % relèveraient de l'obligation de transparence⁷⁰. En France, certaines associations déplorent le manque de précision de la catégorisation des SIA. À ce titre, Amnesty International, dans le cadre du recours formé devant le Conseil d'État relatif à l'algorithme utilisé par la CNAF, a pu mettre en exergue le manque de clarté de la notion de « notation sociale »⁷¹ telle que définie par la réglementation.

- ⁵⁷ Les difficultés relatives à la classification des systèmes à haut risque s'illustrent dans le rapport du comité d'évaluation de la vidéoprotection algorithmique publié en janvier 2025. Si, intuitivement, ce système présente un risque élevé, en ce qu'il a vocation à participer aux missions de maintien de l'ordre, la qualification n'est pas unanime. Au terme de l'analyse des porteurs du rapport, le texte défend le fait que ce système, qui n'autorise par ailleurs pas la reconnaissance faciale, « paraît devoir être regardé comme destiné à “améliorer le résultat d'une activité humaine préalablement réalisée”. Il ne relèverait donc pas de la qualification d'IA à haut risque ». Cette interprétation souligne la complexité de la qualification des systèmes d'IA au regard des définitions réglementaires.
- ⁵⁸ Si le Règlement sur l'IA, autant que la loi française, encadrent l'usage de l'intelligence artificielle, il convient de s'intéresser aux conditions dans lesquelles le droit à l'explicabilité des décisions produites par un SIA ou un algorithme, s'exerce concrètement.

B. Une définition du contenu du droit à l'explicabilité de l'administré en cours d'élaboration

- ⁵⁹ Reste à déterminer, pour l'administré, ce que recouvre le droit à l'explication des décisions algorithmiques (1). Ce droit recouvre des réalités différentes, que l'on se place au niveau européen ou français (2).

1. Un contenu difficilement identifiable au regard de la réglementation européenne

- ⁶⁰ Le règlement sur l'IA prévoit en son article 86 un droit à l'explication des décisions individuelles. Ce droit s'applique aux décisions individuelles, produisant des effets juridiques ou similaires, fondées sur les résultats d'une IA classée à haut risque, à l'exception de celles intervenant dans le domaine des infrastructures critiques.
- ⁶¹ L'article est moins clair quant à la qualité de l'information fournie à la personne concernée. Le texte énonce que la personne « a le droit d'obtenir du déployeur des explications claires et pertinentes sur

⁷⁰ M. P. HAUER, T. D. KRAFFT, A. SESING-WAGENPFEIL, « Quantitative study about the estimated impact of the AI Act », *arXiv*, 2023.

⁷¹ Amnesty international, « France. Les autorités doivent cesser d'utiliser l'algorithme discriminatoire du système de protection sociale », 16 octobre 2024.

le rôle du système d'IA dans la procédure décisionnelle et sur les principaux éléments de la décision prise ».

- 62 La personne concernée a donc le droit de connaître la finalité du SIA et son degré d'implication dans la prise de décision finale. Toutefois, le texte est moins explicite concernant le fonctionnement même du SIA. En effet, « les principaux éléments de la décision prise » peuvent couvrir une importante diversité de réalités : ces principaux éléments sont-ils ce que l'on pourrait assimiler à la motivation classique des décisions administratives ? Sont-ils, au contraire, des éléments propres au fonctionnement du ou des SIA qui sont intervenus dans le processus décisionnel ? Si oui, quels sont les principaux éléments de ces systèmes ?
- 63 Les considérants du règlement n'apportent que peu d'indices quant à la détermination des « *principaux éléments de la décision prise* ». Ponctuellement, le texte renseigne sur les conséquences des lacunes en matière de droit à l'explication des décisions de justice, en précisant que « l'exercice de droits fondamentaux procéduraux importants, tels que le droit à un recours effectif et à accéder à un tribunal impartial, ainsi que les droits de la défense et la présomption d'innocence, pourrait être entravé, en particulier lorsque ces systèmes d'IA ne sont pas suffisamment transparents, explicables et documentés »⁷².
- 64 Toutefois, de manière sectorielle, le règlement s'avère plus précis. C'est particulièrement le cas en ce qui concerne les SIA utilisés dans des procédures judiciaires. Dans ce domaine, il détermine que « les conséquences de l'utilisation des outils d'IA sur les droits de la défense des suspects ne devraient pas être ignorées, en particulier la difficulté d'obtenir des informations utiles sur le fonctionnement de ces outils et partant, la difficulté de saisir la justice pour contester leurs résultats, en particulier pour les personnes physiques faisant l'objet d'une enquête »⁷³. Le règlement adopte ici une vision finaliste de l'explication des décisions fondées sur l'IA, le volume et la qualité des informations communiquées devant permettre l'exercice des droits et libertés fondamentaux garantis. Le droit à l'explication des décisions recouvre donc non seulement la motivation classiquement admise, mais également les règles de fonctionnement des outils d'IA utilisés, au moins dans certains cas.
- 65 Si cette précision permet d'avancer d'un pas dans la définition du contenu du droit à l'explication des décisions individuelles, restent à déterminer les règles de fonctionnement d'un système pouvant être utiles à la mise en cause des décisions produites par un système. Sur ce point, le règlement est définitivement silencieux.
- 66 Peut-être est-il alors possible de déterminer un socle commun d'information ayant une incidence potentielle sur la prise de décision algorithmique. Il est par exemple, courant que des SIA présentent des biais qui trouvent leur source dans les données ayant servi à leur entraînement : un outil de police

⁷² RIA, considérant 59.

⁷³ *Ibid.*

prédictive entraîné dans des quartiers défavorisés pourra être tenté de faire du milieu social une variable déterminante dans le risque du passage à l'acte. Dans ce cadre, les jeux de données d'entraînement et de test pourraient constituer des informations « utiles sur le fonctionnement » d'un système dans la mesure où ils pourraient permettre de trouver la source du biais.

- 67 Dès lors, il pourrait être intéressant d'étendre la notion de donnée traitée aux données utilisées dans le développement des SIA. Les données d'entraînement, de test et de validation pourraient en effet offrir des prises à l'administré dans la compréhension des systèmes ayant participé à l'élaboration de la décision administrative dont il est l'objet.
- 68 Finalement, le Règlement s'avère plus loquace lorsqu'il s'agit de régir les relations entre les fournisseurs et les déployeurs des SIA que les rapports entre le déployeur et les personnes concernées par le traitement.

2. *Un contenu clairement défini par le droit français*

- 69 Tout d'abord, le droit français pose un principe général de publicité des règles de fonctionnement des systèmes algorithmiques utilisés par l'administration. L'article L312-1-3 du code du CRPA dispose que « sous réserve des secrets protégés en application du 2° de l'article L. 311-5, les administrations [...] publient en ligne les règles définissant les principaux traitements algorithmiques utilisés dans l'accomplissement de leurs missions lorsqu'ils fondent des décisions individuelles ».
- 70 Le même code consacre également un droit individuel à l'explication des décisions fondées sur un traitement algorithmique, puisque son article L 311-3-1 dispose qu'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 ». Ces éléments sont précisés par décret⁷⁴. Doivent être communiqués sur demande de l'administré :
- Le degré et le mode de contribution du traitement algorithmique à la prise de décision ;
 - Les données traitées et leurs sources ;
 - Les paramètres de traitement et, le cas échéant, leur pondération, appliqués à la situation de l'intéressé ;
 - Les opérations effectuées par le traitement.

⁷⁴ Décret n° 2017-330 du 14 mars 2017 relatif aux droits des personnes faisant l'objet de décisions individuelles prises sur le fondement d'un traitement algorithmique.

- ⁷¹ Il est ici important de noter que ce droit à l'explicabilité, tel que conçu par le CRPA, ne couvre pas le droit d'accès aux codes sources utilisés par l'administration. Comme a pu le relever la CADA, ces dispositions législatives « concernent uniquement les traitements algorithmiques fondant des décisions individuelles et ne consacrent pas un droit d'accès à l'algorithme lui-même, mais seulement à certaines informations le concernant »⁷⁵. Pourtant, et contrairement à une idée communément admise, les codes sources, en dépit de leur complexité, peuvent apporter une réelle plus-value dans la compréhension des systèmes algorithmiques. En effet, la complexité du code source des algorithmes n'est pas une limite insurmontable. En France, certains dysfonctionnements d'algorithmes utilisés par les administrations ont pu être mis en évidence par le travail associatif et journalistique. C'est par exemple le cas de l'algorithme de la CNAF, évoqué précédemment.
- ⁷² Les codes sources sont donc un vecteur de transparence que l'on pourrait qualifier de déléguée dans la mesure où ils supposent l'intervention d'un tiers qualifié. Si ces documents ne sont pas couverts par la notion d'explication des décisions algorithmiques, ils sont néanmoins, en principe, accessibles.
- ⁷³ Dans les faits, les codes sources constituent bien des documents administratifs au sens du droit français⁷⁶. Or, l'article L 311-1 du CRPA dispose que « sous réserve des dispositions des articles L. 311-5 et L. 311-6, les administrations mentionnées à l'article L. 300-2 sont tenues de publier en ligne ou de communiquer les documents administratifs qu'elles détiennent aux personnes qui en font la demande, dans les conditions prévues par le présent livre ».
- ⁷⁴ Si l'initiative de la demande d'accès au code source relève de l'administré, il est à noter que le droit français fait de la publication des « documents qu'elles (les administrations) communiquent en application des procédures prévues au présent titre, ainsi que leurs versions mises à jour »⁷⁷, une obligation légale.
- ⁷⁵ C'est dans ce cadre qu'un certain nombre d'algorithmes publics ont été rendus accessibles. Il est ici possible de citer en exemple les outils servant au calcul de l'impôt : un grand nombre de codes sources, ainsi que leurs mises à jour, ont été mis en ligne par l'administration⁷⁸.
- ⁷⁶ À l'exception des réserves, qui tiennent essentiellement à des secrets couverts par la loi, les codes sources des algorithmes publics sont donc accessibles. Pour autant, il est possible de questionner la stratégie visant à faire reposer sur l'administré l'initiative des démarches conduisant à leur publication.

⁷⁵ CADA, Avis 20226179, Caisse Nationale des Allocations Familiales, Séance du 15/12/2022.

⁷⁶ L'article L300-2 du CRPA dispose que « sont considérés comme documents administratifs, [...] quels que soient leur date, leur lieu de conservation, leur forme et leur support, les documents produits ou reçus, dans le cadre de leur mission de service public, par l'État, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission. Constituent de tels documents notamment les dossiers, rapports, études, comptes rendus, procès-verbaux, statistiques, instructions, circulaires, notes et réponses ministérielles, correspondances, avis, prévisions, codes sources et décisions ».

⁷⁷ CRPA, art. L312-1-1.

⁷⁸ Direction générale des finances publiques, « ouverture des données publiques de la DGFIP », *impots.gouv.fr*, 24 août 2018.

Les codes sources publics ne pourraient-ils pas, dès leur publication, et sans que l'administré n'ait à entamer des démarches, être rendus publics ?

- ⁷⁷ La procédure actuelle crée le sentiment d'un secret inutile et donc d'une opacité volontaire à l'égard des systèmes utilisés par l'administration. Reste la possibilité, pour l'administration, de rendre publics ces codes de sa propre initiative. Par une circulaire, l'ancien Premier ministre, Jean Castex, avait promu cet état d'esprit volontaire qui implique « en outre, un renforcement de l'ouverture des codes sources et des algorithmes publics, ainsi que de l'usage du logiciel libre et ouvert »⁷⁹. En ce sens, il demandait alors aux préfets d'intégrer dans leurs relations contractuelles des clauses favorables à l'ouverture des données publiques et des codes sources des algorithmes.
- ⁷⁸ Toutefois, quatre ans après ce vœu pieux, force est de constater que le processus d'ouverture des codes sources reste difficile. Tout d'abord, les voies par lesquelles ces ouvertures sont consenties sont disparates, d'un ministère à l'autre, mais également d'une personne publique à l'autre.
- ⁷⁹ Ainsi, l'exemple des efforts fournis par la mairie de Nantes⁸⁰ est remarquable : la ville a recensé les algorithmes qu'elle utilise, ce qui est une obligation posée par la Loi pour une République numérique, et a accompagné cette publicité, non seulement des principales règles régissant le fonctionnement de ces outils, mais également de l'entièreté de leur code source et de « logigrammes » qui consistent en une représentation graphique du fonctionnement de l'algorithme sous forme d'arbre de décision. Ainsi, ils sont plus complets que les simples règles et principes fondamentaux du fonctionnement de l'algorithme et moins complexes que les codes sources bruts qui nécessitent des compétences techniques avancées.
- ⁸⁰ D'un autre côté, la réticence de certaines administrations freine le processus de contrôle démocratique des outils numériques utilisés par l'administration. L'exemple de la CNAF est parlant en la matière. L'association à la source des inquiétudes quant au risque potentiel de discrimination s'est heurtée au refus de communication du code source par l'administration. L'argument avancé par la CNAF pour motiver ce refus était la présence de « mentions qui si elles étaient communiquées pourraient donner des indications aux fraudeurs pour tromper l'algorithme »⁸¹. Cet argumentaire n'a pas totalement convaincu la CADA qui après une longue procédure a finalement ouvert le code source de l'algorithme de la CNAF en distinguant « les variables et les coefficients du modèle de datamining actuel et d'autre part, ceux se rapportant aux versions antérieures du modèle en estimant que si les premiers sont au nombre des éléments dont la divulgation porterait effectivement atteinte à la politique de lutte contre la fraude sociale mise en œuvre par la CNAF ainsi que plus généralement, à

⁷⁹ Circulaire n°6264/SG du 27 avril 2021 relative à la politique publique de la donnée, des algorithmes et des codes sources.

⁸⁰ Open Data Nantes Métropol, Consultation des algorithmes de Nantes Métropole, [en ligne].

⁸¹ La Quadrature du Net, *Réponse CAF demande communication code source algorithme de profilage des allocataires à des fins de contrôle*, [en ligne].

l'efficacité des contrôles, il n'en est pas de même des seconds, utilisés dans les modèles antérieurs et sur le fondement desquels ne reposent plus les contrôles actuels et à venir »⁸².

- ⁸¹ L'argumentaire développé en premier lieu par la CNAF mérite l'attention : celui-ci utilise les mêmes ressorts que ceux développés dans le domaine répressif. La transparence serait un frein à l'efficacité des outils déployés par l'administration, leur connaissance permettrait de tromper le système et leur communication serait donc impossible. L'exemple des « boîtes noires des services de renseignement », qui permettent aux services secrets français d'intercepter, grâce à un algorithme, les données de connexion suspectes⁸³ en est un exemple : les critères de détection tout autant que le code source ne peuvent être communiqués au risque de voir les personnes visées par le traitement s'adapter pour le contourner⁸⁴. Or, si cette idée venait à être acceptée, une large part des algorithmes publics pourraient en bénéficier : *Quid* des algorithmes d'affectation des bacheliers ? Connaître leur fonctionnement pourrait permettre à certains d'en tordre les règles. *Quid* des systèmes de calcul des prestations sociales qui pourraient encourager la fraude ?

CONCLUSION

- ⁸² Ainsi, le fonctionnement et les conditions dans lesquelles l'administration s'approvisionne en SIA sont des sources d'opacité. Dans ces conditions, si le principe de transparence reste inchangé, les conditions de sa mise en œuvre doivent s'adapter aux spécificités de l'IA. L'objectif principal demeure : permettre aux citoyens de savoir ce que fait l'administration. Pour cela, l'information doit être à la fois accessible et compréhensible. Toutefois, l'émergence de l'intelligence artificielle et des « boîtes noires » qu'elle génère souligne la nécessité d'adapter la mise en œuvre du principe de transparence. La complexité des codes sources et l'intervention de nouveaux acteurs expliquent pourquoi les modalités de réalisation de la transparence doivent évoluer.
- ⁸³ Si l'intelligence artificielle permet d'accroître l'efficacité de l'action administrative, elle engendre également des risques liés aux biais algorithmiques et à l'opacité des décisions automatisées. Dès lors, deux options se dessinent : rejeter l'usage de ces outils ou les encadrer rigoureusement afin de préserver les principes fondamentaux du droit administratif. Le choix de la seconde voie semble s'imposer, nécessitant une adaptation du concept de transparence aux spécificités de l'IA.
- ⁸⁴ En définitive, la transparence administrative ne peut plus se limiter à une simple diffusion d'informations. Elle doit s'adapter aux nouvelles réalités, en veillant à ce que les décisions automatisées restent accessibles et compréhensibles par les administrés.

⁸² CADA, Conseil 20225787 - Séance du 15/12/2022

⁸³ Articles L851-1 et suivants du code de la sécurité intérieure.

⁸⁴ M. UNTERSINGER, « Loi sur le renseignement : la « boîte noire » reste obscure », *Le Monde*, 1^{er} avril 2015.

RESUME :

L'essor de l'intelligence artificielle transforme profondément l'action administrative, notamment en apportant de nouveaux outils permettant d'optimiser la prise de décision administrative. Toutefois, cette évolution soulève des interrogations autour du principe de transparence. L'opacité des algorithmes, souvent comparés à des « boîtes noires », complexifie la compréhension des décisions, aussi bien pour les citoyens que pour les administrations elles-mêmes. La transparence administrative, historiquement conçue comme une communication directe entre l'administration et les administrés, devient aujourd'hui une relation tripartite impliquant également les fournisseurs d'IA. Cette nouvelle configuration génère une asymétrie d'information qui remet en cause la capacité de l'administration à pleinement maîtriser les outils qu'elle utilise, avec des risques de biais discriminatoire. Face à ces évolutions, le droit tente d'adapter les conditions de mise en œuvre de la transparence, notamment en matière d'explicabilité et de loyauté.

SUMMARY:

The rise of artificial intelligence is profoundly transforming administrative action, particularly by providing new tools that help optimize administrative decision-making. However, this development also raises important concerns regarding the principle of transparency. The opacity of algorithms, often described as “black boxes,” makes administrative decisions more difficult to understand, both for citizens and for public administrations themselves. Administrative transparency, historically conceived as a direct relationship between public authorities and citizens, has now evolved into a tripartite relationship that also involves AI providers. This new configuration creates an information asymmetry that challenges the administration's ability to fully control and understand the tools it relies on, while also increasing the risk of discriminatory bias. In response to these developments, the law is seeking to adapt the conditions under which transparency is implemented, particularly with regard to explainability and fairness.