



Journée sur les algorithmes d'aide à la décision publique

Compte rendu de la journée du 4 octobre 2022¹ (Laurent Réveillère²)

Introduction

La Société informatique de France a organisé le 4 octobre 2022 à Bordeaux une journée pour sensibiliser le grand public aux rôles des algorithmes d'aide à la décision publique utilisés par exemple pour l'affectation des étudiants, la sélection des receveurs de greffons cardiaques ou le calcul des impôts. La journée a été l'occasion de faire un état des lieux des résultats scientifiques sur lesquels ces algorithmes se fondent et les recherches actuelles qui se développent pour innover dans ce domaine.

Si une partie des interventions a été consacrée aux défis scientifiques et techniques liés à ces algorithmes, le colloque a aussi été l'occasion de faire une place au droit et donner à entendre à la fois comment le droit pense le rôle et les limites de ces algorithmes, mais aussi comment les institutions publiques appréhendent ces questions dans leurs travaux et leurs décisions. Ceci nous a permis d'aborder plus largement les enjeux sociétaux et les questions d'éthique que soulève l'utilisation de ces algorithmes dans la société.

La journée a été organisée autour d'interventions d'informaticiens et d'informaticiennes, mais aussi d'économistes, de sociologues et de juristes, afin de pouvoir mettre en perspective les défis scientifiques avec les enjeux juridiques et éthiques de ce contexte particulier. Les intervenants lors de cette journée étaient :

1. <https://www.societe-informatique-de-france.fr/les-journees-sif/journee-algospub>.

2. Laurent Réveillère, professeur en informatique à l'université de Bordeaux.

- Thalia Breton, auditrice au Conseil d'État ;
- Jean Cattan, secrétaire général du Conseil national du numérique (CN-Num) ;
- Max Dauchet, membre du Comité éthique et scientifique de Parcoursup (CESP) ;
- Richard Dorent, médecin en cardiologie et maladies vasculaires ;
- Jean-Christophe Filliâtre, chercheur CNRS en informatique ;
- Caroline Gabez, rapporteure générale de la Commission d'accès aux documents administratifs (CADA) ;
- Julien Grenet, chercheur CNRS en économie, économiste et ancien membre du comité chargé de l'évaluation de Parcoursup ;
- Clément Hénin, rapporteur à la Cour des comptes ;
- Claude Kirchner, directeur du Comité national pilote d'éthique du numérique (CNPEN) ;
- Claire Mathieu, chercheuse CNRS en informatique, membre de l'académie des sciences et anciennement chargée de mission Parcoursup ;
- Denis Merigoux, chercheur Inria en informatique et membre du projet Mlang (compilateur utilisé dans le calcul des impôts) ;
- François Pellegrini, vice-président de la Commission nationale de l'informatique et des libertés (CNIL) ;
- Fabien Tarissan, chercheur CNRS en informatique ;
- Noé Wagener, professeur de droit public à l'université Paris-Est Créteil ;
- Célia Zolynski, membre de la Commission nationale consultative des droits de l'Homme (CNCDH).

Programme de la journée

Cette journée a commencé par une introduction de Fabien Tarissan sur les enjeux des algorithmes d'aide à la décision publique puis s'est poursuivie avec sept exposés :

- Claire Mathieu : les algorithmes de Parcoursup ;
- Richard Dorent : un nouveau système de répartition des greffons cardiaques utilisant un algorithme ;
- Denis Merigoux : règles, calcul et politique : investigation des choix de programmation inaperçus pour les aides au logement ;
- Noé Wagener : ambition déçue de la loi Informatique et libertés de régler les « algorithmes publics » ;
- Julien Grenet : Affelnet, APB, Parcoursup, etc. : les algorithmes peuvent-ils présider aux destinées des élèves ? ;
- Jean-Christophe Filliâtre : la vérification déductive avec l'outil Why3 ;
- Clément Hénin : expliquer, justifier et contester le système d'attribution des greffons.

Cette journée a également été l'occasion d'organiser deux tables rondes pour confronter les différents points de vue et débattre de ces enjeux :

- « algorithmes de décision publique : élaboration, évaluation et évolutions », animée par Hugo Gimbert avec la participation de Claire Mathieu, Richard Dorent, Denis Merigoux, Max Dauchet (CESP), Claude Kirchner (CNPEN) et Jean Cattani (CNNum) ;
- « quelles perspectives pour l'acceptabilité des algorithmes ? », animée par Fabien Tarissan avec la participation de Caroline Gabez (CADA), Thalia Breton (Conseil d'État), François Pellegrini (CNIL) et Célia Zolynski (CNCDDH).

Cet article revient sur les témoignages de certains des intervenants.

Témoignage de Claire Mathieu, chercheuse CNRS en informatique, membre de l'académie des sciences et anciennement chargée de mission Parcoursup

Depuis 2018, les formations de l'enseignement supérieur en France utilisent la plateforme Parcoursup pour recruter leurs étudiants. Hugo Gimbert et moi-même, chercheurs au CNRS, avons participé à sa mise en place en tant que chargés de mission au ministère de l'Enseignement supérieur, de la Recherche et de l'innovation (MESRI).

Au moment où j'ai commencé, certains choix étaient déjà faits : comme son prédé-

Parcoursup [est] basé sur l'algorithme de Gale et Shapley..

cesseur APB et comme de nombreuses procédures similaires à l'étranger, Parcoursup serait basé sur l'algorithme de Gale et Shapley³, mais sans hiérarchisation par les candidats avant exécution de l'algorithme.

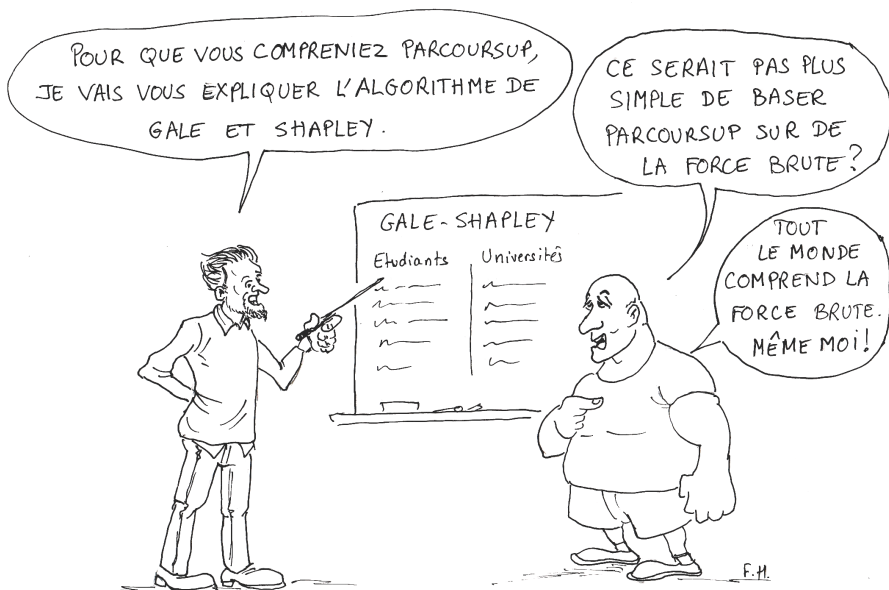
L'algorithme de Gale et Shapley construit un appariement entre des formations et des candidats. Chaque formation a une capacité (nombre de places) et une liste de candidats, qu'elle a classés par ordre de préférence. Chaque élève a une liste de formations auxquelles il a candidaté, formations qu'il a hiérarchisées par ordre de préférence. Dans la version « propositions par les formations » de l'algorithme, un dialogue s'engage entre formations et candidats, alternant propositions par les formations et réponses par les candidats. Chaque formation descend dans son classement, fait des propositions aux candidats suivants de son classement, barre de sa liste ceux qui répondent négativement, et propose les places ainsi libérées aux suivants de son classement. Chaque candidat reçoit des propositions, répond positivement à celle qui est la meilleure dans son classement et négativement aux autres. Le « non » d'un candidat est définitif, mais le « oui » n'est que conditionnel, car il se réserve le droit de

3. https://fr.wikipedia.org/wiki/Algorithme_de_Gale_et_Shapley.

changer d’avis s’il reçoit ultérieurement une proposition d’une formation qu’il préfère — d’où le petit nom de l’algorithme, *deferred acceptance*. Ainsi il peut y avoir une cascade d’allers-retours propositions-réponses. Une fois la situation stabilisée, les « oui » conditionnels deviennent définitifs et définissent l’appariement final.

Dans Parcoursup, les candidats définissent l’ensemble des formations auquel ils candidatent, sans les hiérarchiser explicitement a priori ; les formations font un classement des candidats et le communiquent à la plateforme. Puis le dialogue s’engage entre la plateforme et les candidats, ceux-ci ne révélant (partiellement) leur hiérarchisation que lorsqu’ils ont plusieurs propositions entre lesquelles ils doivent sélectionner leur préférée afin que l’algorithme puisse continuer son exécution.

L’une des demandes du gouvernement était de mettre en place un mécanisme pour aider les candidats défavorisés, un « taux de boursiers ». La loi disait : « *L’autorité académique fixe un pourcentage minimal de bacheliers retenus bénéficiaires d’une bourse nationale de lycée* ». L’autorité académique, c’est le recteur. Imaginons que dans une formation il ait fixé un taux de 5 %. Le texte a deux interprétations possibles : soit que 5 % des propositions (au moins) soient faites à des candidats boursiers, soit qu’à la rentrée 5 % des places (au moins) soient occupées par des boursiers. Avec l’aide de juristes, le groupe de travail a considéré qu’il s’agissait d’une exigence d’opportunités pour les boursiers et non de résultat pour la formation, et a par conséquent privilégié la première interprétation.



Comment adapter l'algorithme de Gale et Shapley pour incorporer les taux de boursiers ? C'est simple : il suffit d'insérer une étape de modification de classement. Une fois que la formation a communiqué son classement à la plateforme, mais avant que le dialogue avec les candidats ait commencé, Parcoursup commence par prendre le classement et y faire remonter des boursiers de façon à garantir qu'au moins 5 % des propositions seront faites à des boursiers. Comment le faire, alors qu'on ne sait pas à l'avance combien de propositions seront faites par la formation ? Il suffit de faire remonter les boursiers au minimum pour garantir qu'il y aura au moins un boursier parmi les 20 premiers du classement, au moins deux parmi les 40 premiers, au moins trois parmi les 60 premiers, etc., jusqu'à épuisement du nombre de boursiers ou de candidats. Pour des raisons d'arrondi, le premier doit être un boursier, puis il en faut un deuxième dans les rangs 2 à 21, puis un troisième dans les rangs 3 à 42, etc. Un algorithme de type glouton fait cela facilement.

Cette adaptation semble avoir été relativement bien reçue. En effet, peu de formations oseraient se plaindre que des candidats boursiers aient été remontés, et, les taux étant généralement peu élevés, les candidats non boursiers n'ont été que légèrement descendus.

Une autre demande du gouvernement était de mettre en place un mécanisme pour autoriser, dans une certaine limite, des candidats hors académie à être bien classés. Avant Parcoursup, la plupart des licences avaient pour vocation de servir d'abord leur académie, et *tous* les candidats hors académie étaient classés après *tous* les candidats en provenance de l'académie. Les académies étant plus ou moins riches en universités relativement à la population, à partir du moment où le nombre de places était inférieur au nombre de candidats, à mon avis, cela induisait une injustice de fait entre les académies. Le gouvernement a donc prévu que l'autorité académique (le recteur) puisse, pour chaque formation, autoriser un taux de candidats hors académie.

De nouveau, nous avons interprété cette demande comme portant sur les propositions faites et non sur les inscrits à la rentrée, et la logique a voulu que nous utilisions le même procédé : modifier les classements avant le dialogue avec les candidats. Par exemple, si le taux est de 5 %, la modification garantit que parmi les 20 premiers il y a au maximum un candidat hors académie (classé 20^e pour des raisons d'arrondi), parmi les 40 premiers il y a au maximum deux candidats hors académie (le second classé 40^e), etc. Cette modification a été beaucoup plus critiquée par les formations que la précédente, car elle peut conduire à prendre des candidats qu'elles avaient repérés et mis tout au début du classement, et les faire reculer de dizaines ou de centaines de places. Le travail de classement effectué par les formations ne semblait pas respecté...

Puis, nous nous sommes posé la question de la combinaison de ces deux règles. Que faire si le recteur impose qu'une formation ait au moins 10 % de boursiers et au plus 5 % de candidats hors académie ? Cela peut conduire à des incompatibilités s'il

se trouve que les candidats boursiers sont tous hors académie ! Impossible alors de respecter les deux règles simultanément.

Pour résoudre ce problème, nous avons choisi de ne pas choisir. Notre rôle est d'implémenter la loi, pas de la faire. Nous avons donc fait remonter la question au législateur. La réponse obtenue a été qu'en cas de conflit entre les deux règles, il fallait donner priorité à la règle de taux de boursier minimum. Avec cette réponse, il a été facile de définir un algorithme cohérent pour modifier le classement.

*...l'algorithme se situe entre
la loi et le code...*

Ainsi, on voit que l'algorithme se situe entre la loi et le code ; que c'est la loi qui détermine l'algorithme et non

pas l'algorithme qui détermine la loi, mais que la conception doit se faire dans un dialogue qui permet au législateur de ne pas imposer ce qui serait impossible à mettre en place ! Les spécialistes du droit, de l'algorithmique et de la programmation ont en commun un souci du détail, un souhait d'être guidé par des principes généraux, et une ambition d'exploration exhaustive des possibles ; malgré les différences de vocabulaire, ces points communs rendent le dialogue assez facile.

Nous avons également contribué à la mise en place d'un algorithme pour allouer les places d'internat pour les classes préparatoires, et conçu des simulations pour estimer le temps de convergence de l'algorithme.

Une autre de nos contributions, caractéristique du monde de la recherche, a été la publication de l'algorithme et du code du cœur de Parcoursup. En tant que chercheurs, notre point de vue sur la transparence est que tout ce qui n'a pas explicitement vocation à être confidentiel devrait être public. Nous avons donc argumenté dans ce sens, et cette publication de l'algorithme et du code, obligatoire selon la loi, a eu lieu en mai 2018, au tout début du dialogue avec les candidats, donc des mois avant la date limite légale.

Parcoursup a soulevé de nombreuses critiques. Nombre de candidats se sont retrouvés affectés à des formations autres que celles qu'ils auraient souhaitées, à cause de manque de places dans leurs formations préférées, et d'inadéquation entre leur profil et les profils des formations avec des places disponibles. Même si l'algorithme est perfectible, le cœur du problème n'est pas algorithmique ; ce sont des problèmes de l'Enseignement supérieur.

Témoignage de Richard Dorent, Benoit Audry, Christian Jacquelinet (agence de la biomédecine) et Clément Hénin (Inria)

Avant 2018, la répartition des greffons cardiaques était fondée sur un système séquentiel proposant les greffons, en premier, aux candidats dont le pronostic vital était engagé à court terme, puis aux équipes de transplantation selon une modalité géographique, avec une répartition locale, régionale puis nationale en fonction de

l'établissement où le prélèvement avait lieu. Ce système est celui qui est aujourd'hui en place dans la plupart des pays. Ce système avait un triple inconvénient. Il n'était pas équitable comme le montrait la surmortalité en liste d'attente des patients non prioritaires par rapport aux patients prioritaires. Il ne visait pas l'efficacité, ne prenant pas en compte le risque d'échec après la transplantation ni de façon plus générale l'appariement entre donneur et receveur. Il n'était pas transparent puisque la moitié des greffons étaient alloués aux candidats par les équipes selon des règles locales.

En 2018, un nouveau système est mis en place, dans lequel la répartition des greffons est faite par un algorithme. Cet algorithme a été co-construit par l'agence de la biomédecine et les équipes de transplantation cardiaque. L'utilisation d'un algorithme permet d'avoir une liste d'attente nationale unique, au sein de laquelle tous les candidats sont classés, avec une attribution des greffons aux patients et non plus en partie aux équipes, selon des règles transparentes. Cet algorithme utilise des modèles prédictifs de mortalité en liste d'attente et après la greffe établie avec des modèles de Cox, un modèle géographique gravitaire, qui pondère le classement en fonction de la durée de transport entre les centres de prélèvement et de greffe, et des choix empiriques pour l'appariement entre donneur et receveur, domaine dans lequel les connaissances sont essentiellement le fruit de l'expérience. Le calcul du score commence par le calcul de la probabilité de décès en liste d'attente, l'index de risque cardiaque, index auquel peuvent déroger les catégories de patients dont le pronostic est mal prédit par le modèle. Cet index est ensuite pondéré par la différence d'âge entre donneur et receveur et la durée de transport entre les centres de prélèvement et de greffe. Enfin, les propositions de greffons sont filtrées selon des critères préétablis concernant le groupe sanguin des donneurs et des receveurs, leur différence morphologique et le risque d'échec après la greffe. La performance de l'algorithme fait l'objet d'évaluations régulières. Ces évaluations ont conduit à la modification de l'algorithme à deux reprises depuis son implémentation. Les modifications portaient sur le classement des candidats sous assistance circulatoire mécanique. De plus, le fonctionnement de l'algorithme a été analysé. Pour déterminer le poids de chacune de ses cinq composantes dans le classement final (cf. figure 1), la moyenne de la différence de points entre les candidats classés premiers et la moyenne de tous les autres candidats inscrits en attente entre janvier 2018 et décembre 2020 a été calculée par composante et exprimée en écart type de la moyenne correspondante. Ainsi, il apparaît que le classement final dépend par ordre d'importance décroissante de l'urgence médicale, puis de l'appariement entre donneur et receveur, puis de la géographie et enfin du risque d'échec après la greffe. Une évaluation de l'équité d'accès à la greffe selon le sexe, l'âge, le groupe sanguin et la cause de l'insuffisance cardiaque a été réalisée en utilisant des *Partial Dependence*

... la répartition des greffons est faite par un algorithme.

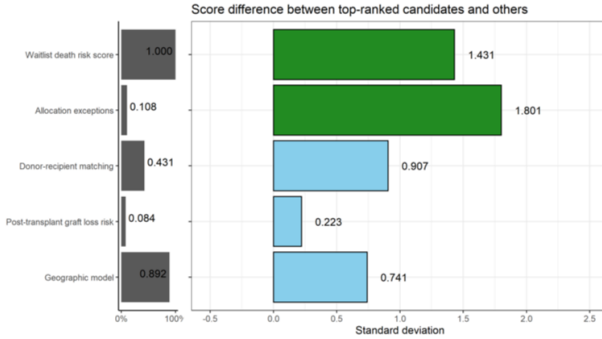


FIGURE 1. Différence de score entre les candidats classés premiers et les autres par composante de l’algorithme exprimée en écart type.

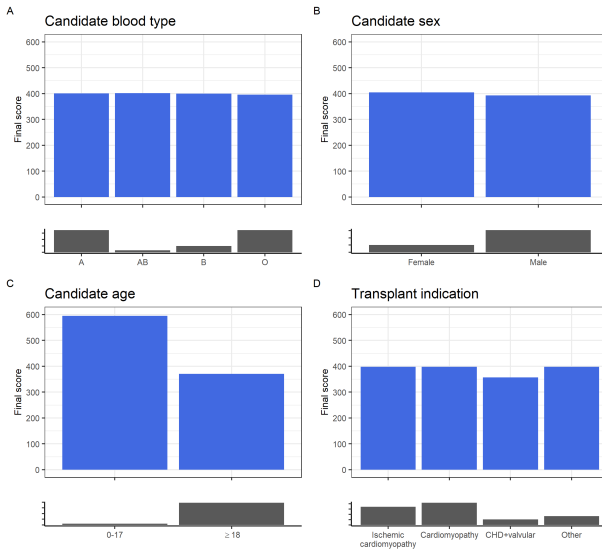


FIGURE 2. Score moyen final par modalité en fonction du groupe sanguin, du sexe, de l’âge et de l’indication de la greffe.

Plots (cf. figure 2). Les histogrammes en gris sous les figures représentent la distribution de la variable dans l’ensemble de la cohorte. Le score final moyen pour les différentes modalités de chacune des quatre variables a été calculé en considérant

que toute la population partageait la modalité indiquée en abscisse de la variable. Par exemple, pour calculer le score moyen des femmes qui apparaît dans la figure, le sexe féminin a été attribué à toute la population. Il apparaît que si le groupe sanguin, le sexe et l'indication de la greffe ont peu d'impact sur le score final, les enfants sont nettement favorisés par rapport aux adultes. Ce résultat était attendu au regard des règles qui avaient été choisies pour l'appariement en âge entre donneur et receveur.

L'utilisation d'un algorithme pour une décision vitale comme la répartition des organes aux candidats en attente de greffe cardiaque a été rendue possible grâce à la volonté de l'agence sanitaire en charge de la répartition des organes, l'agence de la biomédecine, mais aussi grâce à l'adhésion de la communauté médicale impliquée dans cette activité. L'acceptation par les patients eux-mêmes est jusqu'à présent plus implicite qu'explicite. L'algorithme a vocation à évoluer avec le perfectionnement des modèles prédictifs, mais aussi grâce à l'arrivée de nouveaux traitements de l'insuffisance cardiaque terminale.

Témoignage de Noé Wagener, professeur de droit public à l'université Paris-Est Créteil

La loi dite « Informatique et Libertés » du 6 janvier 1978 n'a pas toujours été ce qu'elle est devenue, au sens où ce n'est que très progressivement, selon un processus qui s'est étalé sur trois décennies, qu'elle s'est rétractée dans la seule question de la « protection des données à caractère personnel ». SAFARI⁴, pour le dire plus directement, ne suffit pas à expliquer pourquoi la loi de 1978 a été adoptée, ni comment elle a été mise en œuvre, quand bien même, indiscutablement, c'est ce prisme de lecture, axé sur le seul « fichage », qui s'est imposé. Or, il est possible que ce défaut d'historicisation de l'objet « Informatique et Libertés » — dont témoigne le faible intérêt que continuent de susciter les nombreux cartons sur la loi conservés aux Archives nationales — contribue directement à alimenter les discours aujourd'hui extrêmement répandus qui, surjouant la bifurcation que le droit public de l'informatique serait en train de prendre, soutiennent que, à un droit tout entier concentré sur les données dont le RGPD serait le point d'aboutissement, s'ajouterait *enfin* un « droit des algorithmes » dont le futur règlement européen sur l'intelligence artificielle, actuellement en discussion, serait appelé à devenir le pavillon.

Ce genre de discours n'est ni tout à fait vrai, ni tout à fait faux ; il est caricatural, au sens où, bien évidemment, il n'avait échappé ni aux informaticiens, ni aux juristes, ni au personnel politique des six dernières décennies de réflexion sur les rapports entre l'informatique et les libertés, qu'il y avait un enjeu considérable à réglementer, par-delà la question des données personnelles, la question des traitements algorithmiques *en général*, en particulier lorsque ceux-ci sont utilisés par des administrations. Face à cet étrange pli dans notre mémoire collective, l'ambition des quelques lignes qui

4. Système automatisé pour les fichiers administratifs et le répertoire des individus (SAFARI).

suivent restera minimaliste. Elle consistera seulement à rappeler quelques données historiques simples, que j'articulerai autour des deux idées suivantes : premièrement, la loi de 1978 entendait bien construire *aussi* un « droit des algorithmes » — ou plutôt quelque chose que l'on nommerait aujourd'hui « droit des algorithmes », et qui était d'ailleurs plus ambitieux que les dispositifs que l'on élabore désormais —, et ce, même si la réglementation des « traitements d'informations nominatives » était la priorité des priorités ; or, il apparaît, deuxièmement, que ce droit a non seulement été neutralisé, mais qu'il a été pour partie oublié, au point de permettre à d'étonnants discours de prospérer, comme ceux qui, accompagnant la loi pour une république numérique du 7 octobre 2016, exacerbent aujourd'hui l'innovation que représenterait l'introduction de dispositions juridiques sur les « algorithmes publics ».

Il est frappant, lorsque l'on se plonge dans les archives de la loi de 1978 conservées dans les fonds du ministère de la Justice et du ministère de l'Intérieur aux Archives nationales, de voir à quel point les protagonistes du processus ayant conduit à l'avant-projet de loi — au premier rang desquels Louis Joinet — insistent sur le fait que l'objet premier de la loi n'est pas « les données » ou « les informations » mais « les traitements ». Ce sont ces « traitements » que la loi soumet à quelques « principes généraux », placés en début de texte (selon un ordre bien précis mais qui a été, depuis lors, perdu⁵), et c'est le troisième de ces « principes » qui nous intéresse plus particulièrement : de 1978 à 2004, l'article 3 de la loi de 1978 disposait que « *toute personne a le droit de connaître et de contester les informations et les raisonnements utilisés dans les traitements automatisés dont les résultats lui sont opposés* ».

On peut assez facilement faire la généalogie de ce droit général « de connaître et de contester les informations et les raisonnements utilisés dans les traitements informatisés » : il est l'un des principaux apports de la « commission Informatique et libertés », une commission distincte de la CNIL — qui sera créée quatre ans plus tard — et instituée par le décret du 8 novembre 1974 (*JO* 13 novembre 1974, p. 11403) pour « *proposer au gouvernement [...] des mesures tendant à garantir que le développement de l'informatique, dans les secteurs public, semi-public et privé se réalisera dans le respect de la vie privée, des libertés individuelles et des libertés publiques* ». Cette commission, souvent désignée sous le nom de « commission Tricot » du nom de son rapporteur général, le conseiller d'État Bernard Tricot, opère, en effet, un geste fort quant à la délimitation de ce qu'elle nomme son « *champ à couvrir* » : elle refuse d'enfermer ses réflexions dans la seule question des « informations nominatives » et de la « vie privée », soucieuse, explique-t-elle en ouverture de

5. L'article 1^{er} est certes resté l'article 1^{er}, mais l'article 2 a été partiellement intégré dans l'article 47, après modifications, et l'article 3, dans l'article 119, après modifications aussi.

son rapport, de ne pas donner « à bon compte la conscience d'avoir fait œuvre libérale parce que les secrets et la tranquillité de l'individu et peut-être ceux de l'unité familiale seraient protégés »⁶.

Le constat plus précisément fait par la commission est le suivant : si, effectivement, l'informatique soulève « des risques pour l'avenir » (p. 15 de son rapport final), c'est pour cinq raisons, dont deux seulement sont directement liées à ce que l'on nomme aujourd'hui les « données personnelles ». Certes, « l'informatique permet de conserver les données et les résultats des traitements [...] d'une façon plus massive qu'autrefois » (risque n° 3); certes, aussi, les fichiers sont susceptibles d'être interconnectés, de sorte que s'approche « le jour où, au sein de l'État, chaque fonctionnaire qui détient une parcelle de la puissance publique pourrait tout savoir de chaque homme » (risque n° 5). Mais les risques que l'informatique fait peser sur les libertés vont très au-delà, et ils sont de trois ordres : d'abord, l'ordinateur, dont on dit qu'« il ne se trompe pas », se trompe quand même, notamment parce qu'il existe « bien des raisonnements faux qui peuvent paraître corrects du point de vue logique, bien des déductions inexactes, tendancieuses, moralement ou juridiquement critiquables, devant lesquels l'ordinateur n'aura aucune capacité d'étonnement » (risque n° 1). Ensuite, l'informatique conduit à « une catégorisation des situations et des individus » qui fait que ce serait une « démission [...] de s'en remettre entièrement à elle pour apprécier des situations humaines » (risque n° 2). Enfin, l'informatique « agit dans le sens de l'efficacité technique, mais non dans celui de la liberté », dans la mesure où, en contribuant « à une application très précise de la loi », elle conduit à « une prise en charge accrue de l'individu » (risque n° 4).

La commission Tricot tire de ce constat plusieurs conclusions, et notamment celle selon laquelle il est devenu indispensable de réfléchir aux conditions dans lesquelles on va « pouvoir discuter les « vérités » sorties de l'ordinateur » (p. 80 de son rapport, tome 1), et ce, même si la priorité immédiate reste à « régler le recours aux traitements informatisés nominatifs ». Pour « pouvoir discuter les « vérités » sorties de l'ordinateur », trois mesures phares sont proposées dans le rapport de 1975 (p. 83) :

« La première mesure à prendre consiste à attirer l'attention des responsables de décisions sur la nécessité de ne compter sur l'analyse de système que comme un instrument de travail parmi d'autres et de ne s'en remettre jamais à ses seules conclusions.

Par ailleurs, il serait nécessaire qu'obligation soit faite aux spécialistes de la construction et de la manipulation des modèles de faire connaître sans restriction la nature et la source des données qu'ils ont prises en compte, ainsi que toutes les démarches intellectuelles qui ont permis la construction des programmes.

6. « Rapport de la commission Informatique et libertés », Paris, La Documentation française, 1975, p. 19.

Lorsque des décisions intéressant les libertés publiques et privées seraient prises en s'appuyant sur le concours de telles techniques, il serait également fait obligation aux responsables d'en faire la déclaration et de se tenir prêts à répondre aux questions d'experts désignés par l'instance de contrôle ».

Plusieurs des acteurs majeurs de la préparation de la loi du 6 janvier 1978 défendent aussi à titre individuel cette idée d'instituer un régime général des traitements algorithmiques, qui ne s'enfermerait pas dans le seul prisme du « fichage » car le problème est — l'expression revient souvent — plus « fondamental ». C'est le cas de Bernard Tricot qui, lors d'un colloque en 1977, rappelle que « le problème fondamental dans l'avenir est celui des programmes de traitement, des processus de traitement, de ce que l'on appelle les logiciels »⁷ ou, différemment, de Louis Joinet qui, à l'occasion du grand colloque de l'université de Vincennes sur « le nouvel ordre intérieur » (mars 1979), soutient qu'une législation orientée « vers la stricte protection de la vie privée, correspond à une situation d'ores et déjà dépassée, car là n'est pas — ou plus exactement n'est plus — le risque fondamental »⁸.

L'article 3 de la loi de 1978 est le produit direct de cette réflexion. Lorsque le législateur consacre ce « droit de connaître et de contester les informations et les raisonnements utilisés dans les traitements automatisés » en général, c'est-à-dire indépendamment du fait de savoir si les données faisant l'objet des traitements sont à caractère personnel ou non, il opère un geste d'une immense importance, quoique perdu de vue : il extrait la loi hors de la seule approche en termes de *privacy*, pour réfléchir en termes de « contrôle par la transparence », comme on le fait au même moment dans les cercles qui préparent la loi du 17 juillet 1978 sur « la liberté d'accès aux documents administratifs ». Le déplacement qui est ainsi opéré est substantiel : on ne se contente plus d'attribuer à la personne fichée des droits sur son fichage (droit d'opposition, droit d'accès, etc.) — « qui ne sont finalement que des droits très limités, ne permettant aucunement de maîtriser le développement des systèmes d'information »⁹, comme l'écriront quelques années plus tard Henri Delahaie et Félix Paoletti —, mais on cherche à ouvrir les processus de décision en voie d'automatisation, pour les rendre contrôlables. Le rapporteur de la loi à l'Assemblée, le député Jean Foyer, insiste d'ailleurs expressément sur cette relation avec la loi du 17 juillet 1978 : commentant ce qui deviendra l'article 3 de la loi du 6 janvier 1978, il précise que « le droit de regard ainsi reconnu ne se limite pas aux seules informations qui concernent personnellement chaque individu. Elle s'inspire de la même volonté que

7. Bernard Tricot, « Informatique et libertés », dans *Travail social, informatique, libertés, actes du 31^e congrès de l'Association nationale des assistants de service social*, ESF, 1977, cité par André Vitalis, *Informatique, pouvoir et libertés*, Economica, 1981, p. 160

8. Université de Vincennes, *Le nouvel ordre intérieur*, Éditions Alain Moreau, 1980, p. 230.

9. Henri Delahaie et Félix Paoletti, *Informatique et libertés*, La Découverte, coll. Repères, 1986, p. 100.

les mesures tendant, depuis quelques mois, à faciliter l'accès des citoyens aux documents administratifs »¹⁰, que l'on commence précisément à réunir au même moment sous le concept, accueillant mais peu rigoureux, de « transparence ».

Que s'est-il alors passé pour qu'aujourd'hui, et en dépit des choix explicitement faits en 1978 en faveur de la « transparence » des traitements algorithmiques, domine l'idée que l'émergence de règles sur ce sujet serait un phénomène hypercontemporain — pire, même, que ce phénomène serait, pour ce qui concerne les traitements algorithmiques utilisés par les administrations, une avancée propre à la loi du 7 octobre 2016 pour une république numérique ?

Il est tentant, pour répondre à ces questions, de se tourner vers la CNIL, et d'interroger sa responsabilité particulière : si l'article 3 n'a pas connu la destinée qui aurait pu être la sienne, est-ce parce que la commission n'a jamais cherché à l'investir ? Est-ce parce que la commission n'a pas engagé, après 1978, un travail d'élaboration d'une doctrine générale des traitements algorithmiques ? Il est difficile de répondre de manière univoque à une question de ce type. Un point est sûr, néanmoins : la CNIL en a davantage fait, à propos de l'article 3, que ce que son rapport de 2017 sur « les enjeux éthiques des algorithmes et de l'intelligence artificielle »¹¹ en dit — et qui, paradoxalement, semble tout ignorer du travail accompli par la commission durant les décennies quatre-vingt et quatre-vingt-dix.

Dès son premier rapport, celui intitulé « Bilan et perspectives 1978-1980 »¹², la commission fait part de son inquiétude à propos de l'article 3, constatant que la « mise en œuvre du droit de connaître les raisonnements utilisés, c'est-à-dire en fait les traitements, risque d'être encore beaucoup plus délicate que l'exercice du droit d'accès aux informations proprement dites ». S'il est « indispensable », poursuit-elle, « que les détenteurs de traitements et de fichiers soient en mesure d'expliquer en langage clair les procédés et traitement utilisés, et éventuellement d'en faire la preuve ou d'en fournir une trace », cette nécessité « risque de se heurter à des obstacles matériels » (p. 100 du rapport).

Or, cette inquiétude que la commission exprime ne l'empêche pas de rendre, tout au long des années quatre-vingt, des avis mobilisant expressément l'article 3. Le premier de ces avis date, semble-t-il, de 1982. Sollicitée à propos d'un traitement de tirage au sort de certains membres du « conseil supérieur provisoire des universités » (délibération n° 82-182 du 2 novembre 1982), la commission décide, en effet, de

10. Jean Foyer, « Rapport fait au nom de la commission des lois », Assemblée nationale, séance du 4 octobre 1977, Doc. n° 1325, tome II, p. 2.

11. CNIL, *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*, décembre 2017, p. 44-46.

12. Rapport de la Commission nationale de l'informatique et des libertés. Bilan et perspectives 1978-1980, La Documentation française, 1980.

prendre appui sur l'article 3 pour étendre son contrôle au traitement *en soi*, indépendamment des données personnelles qui y sont effectivement traitées. En cela, l'avis est moins anodin qu'il n'y paraît à première vue : si un avis favorable est délivré, c'est à la condition, d'une part, « *que soient mieux précisées les caractéristiques du raisonnement qui confère au traitement son caractère de tirage au sort, et notamment le contenu de la « fonction aléatoire programmée »* », et après avoir relevé, d'autre part, qu'une « commission de contrôle » est instituée en accompagnement de ce traitement, qui dispose « *préalablement, et au moment du tirage au sort, des moyens nécessaires pour tester et contrôler la régularité des opérations et traitements, en particulier [...] les programmes de tri, la liste des critères de tri [ou encore] les listes par section [etc.]* ».

L'avis de 1982, qui tire donc pour la première fois de l'article 3 des conséquences en termes de contrôle, ne représente pas, pour autant, un tournant décisif. Dans bien des avis postérieurs, au contraire, la CNIL préfère en rester à une approche plus molle de l'article 3, se contentant d'exiger, lorsqu'un traitement algorithmique aide à une prise de décision, que les « motifs » de la décision finalement prise soient publicisés (par exemple, délibération n° 85-44 du 15 octobre 1985 portant adoption d'une recommandation relative à la collecte et au traitement d'informations nominatives lors d'opérations de conseil en recrutement) ou qu'une procédure contradictoire accompagne le dispositif (par exemple, délibération n° 86-37 du 18 mars 1986 portant avis relatif à la mise en œuvre du traitement automatisé des infractions aux conditions de travail et de sécurité en matière de transports routiers). Avec l'accélération du déploiement des « systèmes experts » dans les administrations, néanmoins, la commission cherche manifestement à faire mûrir sa doctrine. Dès son quatrième rapport d'activité (1983), d'ailleurs, la CNIL affirmait mesurer l'enjeu entourant le respect, par lesdits « systèmes experts », de l'article 3 : « *s'il est vrai, écrivait-elle alors, que l'informatique est à peu près infallible lorsqu'elle exécute un traitement, la qualité des résultats reste malgré tout absolument dépendante tant de la qualité des raisonnements programmés que des hypothèses de départ, et donc de la qualité d'après laquelle on travaille* » (p. 200), et c'est pourquoi, poursuivait-elle, « *les ouvertures proposées par les projets relatifs à l'intelligence artificielle devraient offrir de multiples sujets de réflexions sur l'application de l'article 3 de la loi du 6 janvier 1978 qui, bien que peu utilisé jusqu'à présent, pourrait devenir fondamental* » (p. 201). C'est dans la lignée de ces réflexions sur le contrôle des systèmes experts qu'à partir de 1986, plusieurs avis isolent spécifiquement l'article 3, en particulier une délibération concernant la mise en œuvre, par le Conseil général de Seine-Saint-Denis, d'un « *traitement automatisé de gestion de l'aide sociale* » (délibération n° 86-03 du 7 janvier 1986) et une autre concernant le « *traitement automatisé d'aide à l'affectation des logements sociaux mis en œuvre par la commune de Guyancourt* » (délibération n° 86-41 du 8 avril 1986). Dans ces deux cas, la commission exige, sur le fondement de l'article 3, que les « barèmes » utilisés soient rendus publics,

et pour cela, qu'ils fassent l'objet d'une approbation par l'organe délibératif compétent (en l'occurrence, par le conseil municipal ou le conseil général). On avance donc dans la mobilisation de l'article 3 et dans l'exploitation de son potentiel, mais à très petits pas. Une étape supplémentaire est franchie avec une délibération de 1988 sur le Système informationnel de l'Assurance Maladie SIAM (délibération n° 88-31 du 22 mars 1988). Non seulement la CNIL demande que les praticiens et assurés soient effectivement « *informés des conditions d'exercice [du droit de connaître et de contester les informations et les raisonnements utilisés]* » ; mais surtout, elle soutient que « *les thèmes de recherche, critères et raisonnements programmés à l'aide du système SIAM doivent faire l'objet d'un enregistrement systématique de façon à en permettre un contrôle a posteriori* ». Les formules restent fort vagues, certes, mais l'idée d'une obligation de « contrôle » du traitement est désormais bien en place, ce qu'une délibération de 1989 prolongera (délibération n° 89-01 du 10 janvier 1989). Sollicitée à propos d'une « application télématique » de l'Assistance publique de Paris nommée SESAME DIABETE, la CNIL isole en effet à nouveau l'article 3, pour en vérifier le bon respect à partir de deux points de référence qu'elle identifie ainsi : d'une part, « *les patients et médecins utilisateurs pourront avoir accès à un historique des consultations du système comportant les informations et les résultats produits par le système expert* », permettant ainsi « *aux médecins de contrôler la pertinence des réponses fournies* », et en particulier à « *un médecin dûment formé à cet effet* » ; d'autre part, « *la validité des règles du système expert devrait également pouvoir faire l'objet de vérifications périodiques* » (on notera l'étrange conditionnel employé).

Il semble que la CNIL ne soit jamais allée plus loin que cela : les deux délibérations de 1988 (88-31) et 1989 (89-01) représentent les deux fortins les plus avancés de l'article 3 de la loi de 1978. Les conclusions qu'on peut en tirer sont donc mitigées : avec cette obligation d'information, cette obligation d'« enregistrement systématique » des « thèmes de recherche, critères et raisonnements programmés [...] de façon à [...] permettre un contrôle *a posteriori* », ou encore cette obligation de « vérifications périodiques » de « la validité des règles du système expert », on peut dire que la commission esquisse un droit des traitements algorithmiques qui prendrait appui sur l'article 3 ; mais *ce droit se présente d'une manière particulièrement éclatée*. Aucune doctrine véritable ne semble le sous-tendre, au sens où la CNIL ne vérifie pas une liste prédéterminée d'éléments qu'elle aurait elle-même dégagée de l'article 3 au terme d'un important travail interne de réflexion, mais se contente de donner des avis favorables qui constatent que, pour tel ou tel traitement, telle ou telle garantie existe, sans exiger ensuite que, pour tel ou tel autre traitement, telle même garantie soit assurée si elle ne l'est pas. Cette manière de faire est surprenante : elle est sans doute redoutablement pragmatique, mais bien peu stable. Sa fonction, pour le dire plus directement, est davantage d'accompagner le développement des systèmes experts au nom de l'article 3, qu'à veiller au bon respect de cet article.

Quoi qu'il en soit, la mobilisation par la CNIL de l'article 3 décline progressivement dans les années quatre-vingt-dix, jusqu'à disparaître complètement au tournant 2000, au moment même où le recours aux traitements algorithmiques dans les processus décisionnels — dans l'administration et en dehors — explose. Non seulement la commission renonce à investir ce « *droit de connaître et de contester les informations et les raisonnements utilisés dans les traitements automatisés* » que la loi de 1978 avait consacré ; mais elle choisit de ne plus même en parler, alors que quelques années plus tôt, dans son rapport d'activité de 1989, elle pouvait encore se permettre de défendre fermement l'idée que « *la loi du 6 janvier 1978, à partir d'une remarquable évaluation des inconvénients d'une décision par trop automatisée, donne à celui qui a à en pâtir, les moyens de discuter et de contester les vérités sorties par l'ordinateur-oracle* » (p. 21).

La prévision faite par la CNIL en 1983 ne s'est donc pas réalisée : *l'article 3 de la loi du 6 janvier 1978 n'est pas devenu « fondamental »*. Si l'on veut bien se souvenir qu'en 1991, le Comité interministériel de l'informatique et de la bureautique dans l'administration (CIIBA) dressait la liste détaillée des plus de cent systèmes experts en œuvre « dans l'administration française »¹³, alors on peut en conclure qu'il y a, autour de ce « *droit de connaître et de contester les informations et les raisonnements utilisés dans les traitements automatisés* », quelque chose d'un rendez-vous qui a été manqué ; ou qu'on n'a pas voulu honorer.

Dans ces conditions, il est peu surprenant que l'article 3 de la loi de 1978 ait fini par être réécrit au début des années 2000. Lorsqu'en 2004, la France reprend en profondeur la loi du 6 janvier 1978 pour transposer la directive de l'Union européenne du 24 octobre 1995 relative aux données à caractère personnel, elle acte — implicitement, puisque ces choses sont peu débattues — l'échec du projet initial de 1978 de bâtir une grande loi qui, répondant aux défis de « l'informatisation de la société », règle la question des rapports entre informatique et libertés. La loi de 1978 devient ce qu'elle ne voulait pas *seulement* être au départ, c'est-à-dire un texte sur « la protection des données personnelles ».

Dans un tel texte, le « *droit de connaître et de contester les informations et les raisonnements utilisés dans les traitements automatisés* » en général n'a plus sa place, et c'est pourquoi les transformations que la loi du 6 août 2004 opère sur l'article 3 de la loi du 6 janvier 1978 sont spectaculaires. Elles sont spectaculaires sur la forme, d'abord, dans la mesure où l'article 3 perd son statut de disposition principielle placée en tête de la loi, pour glisser bien plus loin dans le texte, à l'article 39. Elles sont spectaculaires sur le fond, surtout, car l'article 3 devenu 39 est désormais rédigé de la manière suivante :

13. Danièle Bourcier et Jean-Paul Costa (dir.), *L'administration et les nouveaux outils d'aide à la décision. Systèmes experts et SIAD*, Editions STH, coll. « Les grands colloques », 1992, p. 200-213.

I.-Toute personne physique justifiant de son identité a le droit d'interroger le responsable d'un traitement de données à caractère personnel en vue d'obtenir :

[...]

5° Les informations permettant de connaître et de contester la logique qui sous-tend le traitement automatisé en cas de décision prise sur le fondement de celui-ci et produisant des effets juridiques à l'égard de l'intéressé. Toutefois, les informations communiquées à la personne concernée ne doivent pas porter atteinte au droit d'auteur au sens des dispositions du livre Ier et du titre IV du livre III du code de la propriété intellectuelle.

Avec la loi du 6 août 2004, on ne substitue donc pas seulement l'expression de « logique qui sous-tend le traitement automatisé » à celle de « raisonnement utilisé » dans ce traitement employée depuis 1978. On fait bien plus que cela. D'abord, le « droit de connaître et de contester » qui avait été accordé à « toute personne » ne l'est plus qu'aux « personnes physiques ». Ensuite, ce droit, qui pouvait être exercé dès qu'on était en présence d'un traitement « dont les résultats sont opposés à une personne », suppose désormais qu'une « décision » ait été prise, d'une part, et qu'elle ait été prise « sur le fondement » d'un traitement automatisé et qui plus est, seulement si ce traitement automatisé produit « des effets juridiques » à l'égard de l'intéressé, d'autre part. Enfin, et surtout, alors que le « *droit de connaître et de contester les informations et les raisonnements utilisés dans les traitements automatisés* » représentait, on l'a vu, le témoignage le plus direct du souci de ne pas donner « *à bon compte la conscience d'avoir fait œuvre libérale parce que les secrets et la tranquillité de l'individu et peut-être ceux de l'unité familiale seraient protégés* » — pour reprendre l'expression employée dans le rapport de 1975 déjà citée —, il ne peut désormais plus être activé que pour des traitements de données à caractère personnel. L'article 3 change, autrement dit, d'objet en 2004, lorsqu'il se transforme en article 39 : d'une disposition qui aurait pu « devenir fondamentale » pour accompagner l'informatisation de la société, comme l'écrivait la CNIL en 1983, il devient l'un des différents « droits des personnes à l'égard des traitements de données à caractère personnel ».

L'idée d'un *régime général des traitements algorithmiques* est donc définitivement abandonnée, au profit d'une mesure de protection des données à caractère personnel. Le déplacement est considérable ; et il l'est d'autant plus que la CNIL n'investit pas véritablement, après 2004, le 5° de l'article 39 (ni, après 2018, le 5° de l'article 119) : hors les deux délibérations de 2018 concernant Parcoursup (délibérations n° 2018-011 du 18 janvier 2018 et n° 2018-119 du 22 mars 2018), presque aucune décision de la commission ne mobilise cette disposition. Le RGPD fait pire encore, d'ailleurs : s'il envisage une obligation, pour le responsable d'un traitement de données à caractère personnel, de fournir à la personne concernée « des informations utiles concernant la logique sous-jacente », ce n'est que dans le cas d'une

« *décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire* ».

Peut-être le plus fascinant est-il encore ailleurs, ceci dit : il n'est même pas sûr que toutes ces évolutions se soient faites de manière bien consciente. On en veut pour preuve le rapport de la CNIL de 2017 intitulé *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*. Dans ce rapport extrêmement documenté, la CNIL pointe en ces termes « *les limites de l'encadrement juridique actuel* » en matière d'algorithmes (p. 46) : les dispositions actuellement applicables, écrit-elle, « *ne concernent les algorithmes que dans la mesure où ils utilisent pour fonctionner des données à caractère personnel et où leurs résultats s'appliquent directement à des personnes. C'est notamment le cas de la loi Informatique et libertés, la seule parmi celles évoquées qui vise directement les algorithmes (mentionnés comme « traitement automatisés de données à caractère personnel »)*. Or, bien des algorithmes n'utilisent pas de données à caractère personnel. [...] On pourrait imaginer [...] un algorithme établissant les menus des cantines scolaires selon certains critères (optimisation du coût des denrées, qualité nutritionnelle, etc.) et qui pourrait être utilisé à l'échelle d'un pays. Un tel algorithme, sans traiter de données personnelles, serait susceptible d'avoir des impacts sociaux et économiques du fait même de son échelle de déploiement. Or, la loi n'a jusqu'ici pas pris en compte cette dimension nouvelle ». C'est passer un peu vite sous le tapis le fait que si « *la loi n'a jusqu'ici pas pris en compte cette dimension nouvelle* », c'est précisément parce que... cette « dimension » lui en a été retirée en 2004. Un rapport collectif rendu mi-2019 par les élèves de la promotion Molière de l'ENA à la demande de la mission Etalab (direction interministérielle du numérique et du système d'information et de communication de l'État) et intitulé *Éthique et responsabilité des algorithmes publics* témoigne du même oubli (p. 3) : « *Les premières normes applicables aux algorithmes publics* », est-il écrit, « *sont issues de la loi relative à l'informatique, aux fichiers et aux libertés du 6 janvier 1978 [...]. Ces obligations se sont néanmoins révélées insuffisantes, avec l'expansion de l'usage des algorithmes par les administrations pour la prise de décisions individuelles. En effet, elles ne permettraient pas de garantir une complète information du citoyen, dans la mesure où de nombreux traitements algorithmiques publics sont alimentés par des données qui ne sont pas toujours à caractère personnel. Par ailleurs, les décisions individuelles faisant intervenir des traitements algorithmiques peuvent également concerner des personnes morales, lesquelles ne bénéficiaient pas des dispositions favorables de la loi du 6 janvier 1978. La mise en place d'un cadre juridique spécifique à l'utilisation d'algorithmes par les pouvoirs publics s'est donc imposée afin de répondre à une demande croissante de transparence et d'explicabilité* ».

Présenter le débat en ces termes, c'est s'étonner qu'en matière d'algorithmes, la loi ne contienne pas ce qu'on a choisi de lui retirer, tout en passant sous silence le fait

qu'on a fait ce choix. On a beau jeu, dans ces conditions, de pointer l'extrême nouveauté des dispositions relatives aux « algorithmes publics » de la loi du 7 octobre 2016 pour une république numérique. Le nouvel article L. 311-3-1 du Code des relations entre le public et l'administration — qui dispose qu'« *une décision individuelle prise sur le fondement d'un traitement algorithmique comporte une mention explicite en informant l'intéressé* », avant d'ajouter que « *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* » — est certes une avancée par rapport à ce que la loi de 2004 a fait de l'article 3 de la loi de 1978. Cela ne doit pas pour autant conduire à faire perdre de vue que *le dispositif de 2016 reste très en retrait par rapport à celui de 1978*, sur les trois points centraux suivants : d'abord, ce nouveau droit de connaître que la loi de 2016 consacre ne peut être actionné qu'en cas de « décisions administratives individuelles » ; ensuite, il ne concerne jamais que les algorithmes utilisés par les administrations, qu'il s'agisse de personnes publiques ou de personnes privées en charge d'une mission de service public ; enfin, la loi de 2016 admet un nombre très important de dérogations au droit de connaître, telles que « *le secret des délibérations du Gouvernement et des autorités responsables relevant du pouvoir exécutif* », « *la sûreté de l'État, la sécurité publique, la sécurité des personnes* », mais aussi, et peut-être surtout tant la pratique montre que cette dérogation est désormais massivement mobilisée par les administrations, « *la sécurité des systèmes d'information des administrations* ».

Si l'on veut bien, enfin, se souvenir que cette « transparence des algorithmes publics » introduite en 2016 a été un argument majeur pour libéraliser le recours aux décisions administratives entièrement automatisées — c'est-à-dire prises sur le *seul* fondement d'un traitement automatisé, ce que l'article 2 de la loi de 1978 interdisait rigoureusement —, alors ce n'est pas caricaturer que de résumer le mouvement qui s'est opéré entre 2004 et 2018 de la manière suivante : on a autorisé en France un recours massif aux décisions administratives intégralement automatisées (2018) quelques années après avoir réduit considérablement les garanties initialement envisagées pour les individus en matière de traitements algorithmiques (2004), tout en communiquant autour du fait qu'on les avait accrues (2016). L'élaboration d'un droit des traitements algorithmiques — qui est désormais une urgence — mérite, c'est certain, un débat plus informé et de meilleur niveau.

Témoignage de Max Dauchet, membre du CESP

Tout français a déjà entendu parler de Parcoursup, et dans la plupart des familles cette plateforme a fait l'objet de discussions enflammées lors des repas de famille. Parcoursup est le souffre-douleur concentrant tous les mécontentements concernant l'accès à l'enseignement supérieur. Les flèches viennent de tous bords, de toutes les classes sociales avec des arguments parfois contraires. Pourtant Parcoursup n'est

qu'une plateforme chargée de la mise en œuvre pratique de la loi relative à l'orientation et la réussite des étudiants. On peut être ou non d'accord quant aux choix qui ont guidé la loi, il est légitime d'en débattre, mais Parcoursup ne fait que traduire le mieux possible la loi.

Laissons ici de côté les débats sur les capacités d'accueil des formations, il s'agit d'un important sujet de société, mais qui ne concerne pas la plateforme : le rôle de celle-ci se borne à faire « les meilleures propositions possibles » aux candidats pour des capacités d'accueil données.

Beaucoup d'opinions et de témoignages révèlent une incompréhension plus ou moins grande du processus de proposition de formations aux candidats. Malgré les améliorations réalisées chaque année en matière de rapidité, transparence, équité, justice sociale, les enquêtes IPSOS commanditées par le ministère montrent plutôt une érosion de la confiance. Parcoursup est perçue comme une boîte noire où l'on rentre des vœux et qui sort mystérieusement des propositions aux candidats, et les algorithmes sont accusés de tous les maux. Ce que l'on ne comprend pas inspire toujours crainte et suspicion.

Le terme *algorithme* est évoqué dans Parcoursup à deux propos. Le premier à travers les mal nommés algorithmes locaux, qui désignent simplement l'usage d'une feuille de tableur pour classer les candidats en fonction de leurs notes et de l'appréciation des commissions d'examen des vœux. Le débat fait rage sur le manque de transparence entre la part d'automatique — ici un préclassement des candidats sur critères chiffrés — et la part d'appréciation humaine sur examen des dossiers. Ce débat est légitime, mais hors des propos de la journée qui nous a réunis à Bordeaux, car il n'interroge en rien l'algorithmique.

Il en va autrement de l'algorithme d'appariement des vœux des candidats et de leur classement par les formations. La question de l'appariement de l'offre (ici l'offre de formation) et de la demande (ici les vœux des candidats) est une question tout à fait générale que l'on retrouve partout. Julien Grenet a rappelé dans son exposé qu'il n'y a pas de solution « idéale » au problème, mais que l'algorithme utilisé par Parcoursup est un des meilleurs possibles. Il s'agit en effet d'une adaptation du fameux algorithme de Gale et Shapley. Cet algorithme est utilisé par toutes les plateformes dans le monde comparables à Parcoursup. Il a le mérite d'être simple à mettre en œuvre, d'avoir de bonnes propriétés et d'avoir un caractère intuitif (l'exemple historique est celui des mariages stables). Pourtant cet algorithme n'est pas si simple. Il a valu à Gale le prix Nobel 2012 d'économie (Shapley ne pouvait le recevoir car il était mort quelques années auparavant) et la preuve de ses propriétés est astucieuse. Cet algorithme de Parcoursup a fait l'objet d'un effort exemplaire de transparence, avec l'accès en données ouvertes à la vérification de sa conformité et des éléments de preuve de son code. Cependant cette transparence n'entraîne pas pour autant la confiance, car pour comprendre il faut une bonne culture algorithmique.

On touche ici le titre de ce bref article. Les algorithmes deviennent omniprésents dans la société, de façon apparente ou souvent enfouie, non perceptible directement. Les interactions avec la société en font des objets sociotechniques complexes et nouveaux. La plateforme Parcoursup en est une bonne illustration. Il faut souligner l'initiative novatrice de l'État d'avoir accompagné la création de Parcoursup de celle d'un comité éthique et scientifique. Ce comité recommande l'enseignement de l'algorithme Gale et Shapley.

L'école publique obligatoire avec son apprentissage de la lecture et de l'écriture du français a été un grand progrès démocratique qui a permis à tous de participer aux débats de société de façon éclairée, à voter en connaissance de cause et à exprimer et porter des opinions de façon intelligible.

Dans la société numérique où nous entrons, l'acculturation à l'algorithmique relève du même enjeu démocratique. Les choses progressent, lentement, mais progressent, dans notre dispositif éducatif. Il faut approfondir la démarche. Quand pour toutes les familles, un proche ou un ami maîtrisera assez l'algorithme de Gale et Shapley pour l'expliquer aux autres, les débats sur Parcoursup y gagneront.

Table ronde sur l'explicabilité des algorithmes, animée par Fabien Tarissan, chercheur CNRS en informatique

La SIF a souhaité conclure cette journée par une dernière table ronde permettant de prolonger les éléments de réflexion qui ont été proposés précédemment. L'occasion de s'interroger, avec différents représentants et représentantes d'instances publiques, sur la question de l'explicabilité des dispositifs algorithmiques d'aide à la décision, dans une perspective d'acceptabilité de ces approches par la société.

Pour cela, commençons par donner la parole à Caroline Gabez, qui est rapporteure générale de la Commission d'accès aux documents administratifs (CADA), autorité administrative indépendante bien connue de la communauté des informaticiens et des informaticiennes pour l'avis qu'elle a rendu en 2016 autour de l'accès au code source qui régissait l'affectation des néo-bacheliers dans le supérieur. Outre cet avis, quelle est la position générale de la CADA sur cette question de transparence administrative ?

Nous terminons ce tour de table par Célia Zolynski, professeure de droit privé à Paris Panthéon Sorbonne et personnalité qualifiée de la Commission nationale consultative des droits de l'homme (CNCDH). Comme son nom l'indique, la CNCDH concentre son travail sur la question des droits de l'homme, ce qui l'a amené à s'interroger sur les effets potentiellement discriminatoires des algorithmes. Deux avis récents sur le sujet, celui de 2022 intitulé « l'impact de l'intelligence artificielle sur les droits fondamentaux », mais aussi l'étude de 2021 sur « la lutte contre la haine en ligne » font le lien avec notre journée puisqu'elles mettaient en évidence le besoin, à la fois de transparence dans les systèmes des algorithmes, mais aussi d'explicabilité ou l'intelligibilité.

Témoignage de Caroline Gabez, rapporteuse générale de la CADA

Le régime juridique des algorithmes utilisés par l'administration est marqué par des exigences de transparence administrative plus ou moins fortes, sur lesquelles je vous propose de revenir, à partir de l'exemple de Parcoursup.

Le droit d'accès aux documents administratifs, qui est constitutionnellement garanti, est un gage de transparence administrative. La Commission d'accès aux documents administratifs (CADA) est l'autorité administrative indépendante chargée de se prononcer sur le caractère communicable des documents administratifs. Le principe est celui du libre accès, sous réserve des secrets protégés.

Dans un avis du 23 juin 2016 (n° 20161989), la CADA a ainsi considéré que les fichiers informatiques constituant l'algorithme de la plateforme d'admission post-bac dite « APB », à laquelle a succédé Parcoursup, constituaient un document administratif, librement communicable à toute personne.

Le code des relations entre le public et l'administration (CRPA) consacre un régime de droit commun de l'accès aux documents administratifs pris sur le fondement d'un traitement algorithmique, ce que nous verrons dans un premier temps. Nous constaterons ensuite que s'agissant de Parcoursup, la transparence est assurée par des dispositions spécifiques du code de l'éducation qui enrichit le régime de droit commun en ce qui concerne l'algorithme de la plateforme nationale, tout en le restreignant, pour les traitements algorithmiques locaux.

Le CRPA n'impose pas expressément la communication des algorithmes publics, seuls les codes sources étant visés dans la liste des documents administratifs communicables. Le législateur impose cependant des exigences particulières de transparence à l'administration dans l'utilisation des algorithmes. En premier lieu, les administrations ont l'obligation de publier en ligne les règles définissant les principaux traitements algorithmiques utilisés, lorsque ces traitements fondent des décisions individuelles (article L312-1-3 du CRPA). Un second procédé prévoit que les administrés ont le droit d'être informés qu'une décision administrative les concernant a été prise sur le fondement d'un traitement algorithmique (article L. 311-3-1 du CRPA). Par ailleurs, s'ils en font la demande, ils ont le droit d'obtenir communication des règles définissant le traitement, des principales caractéristiques de sa mise en œuvre et de la finalité poursuivie. La CADA est compétente pour connaître des refus opposés à des demandes formulées en application de ces dispositions, qu'elle considère comme le droit commun de l'accès aux documents administratifs pris sur le fondement d'un traitement algorithmique. Ces textes ne sont toutefois assortis d'aucune sanction, ce qui en limite la portée effective. En outre, la CADA n'est que très peu saisie de ces questions. Des dispositions législatives particulières peuvent déroger à ce cadre général. Le cas de Parcoursup en fournit une bonne illustration.

Parcoursup est la plateforme numérique nationale de préinscription dans une formation de premier cycle de l'enseignement supérieur, créée en 2018¹⁴. Contrairement à certaines idées reçues, la plateforme ne procède pas au classement des candidatures, qui relève de la compétence de chaque établissement d'enseignement. Au niveau national, la plateforme Parcoursup gère les réponses à apporter aux candidats en fonction des décisions prises par les établissements grâce à un traitement algorithmique d'affection, aussi appelé « algorithme national ». Il établit, pour chaque formation, l'ordre « d'appel » dans lequel les propositions d'admission sont adressées aux candidats. Cet ordre traduit le classement pédagogique opéré par les établissements et tient compte d'autres critères, tels que les quotas boursiers. À la différence du CRPA, le code de l'éducation (paragraphe II de l'article L. 612-3) lève toute ambiguïté au sujet de la communication de l'algorithme utilisé pour le fonctionnement de la plateforme. Il prévoit en effet que la communication du code source de la plateforme s'accompagne de celle d'un cahier des charges *et de l'algorithme* de traitement. Ces éléments — qui permettent notamment à chacun de vérifier que le fonctionnement de la plateforme est conforme au droit — ont été mis en ligne par le ministère de l'Enseignement supérieur, le 21 mai 2018. La CADA s'est déclarée compétente pour se prononcer sur leur communication, dans un avis du 6 septembre 2018 (n° 20182093) et, plus récemment, le 13 janvier 2022 (n° 20213847).

Si les règles d'accès posées par le code de l'éducation sont sur ce point plus riches que celles du CRPA, il en va différemment des documents administratifs se rapportant aux algorithmes locaux. Une procédure de tri des candidatures est mise en œuvre au niveau de chaque établissement. Pour ce faire, les établissements réunissent des commissions d'examen des vœux, chargées de proposer les réponses faites aux candidats. Ces commissions peuvent utiliser des algorithmes, dits « locaux », pour classer les candidatures. Les établissements sont libres de recourir ou non à des outils algorithmiques et, s'ils y recourent, d'utiliser le logiciel d'aide à la décision mis à leur disposition par le ministère de l'Enseignement supérieur ou tout autre logiciel de leur choix. Le code de l'éducation a instauré un régime d'accès particulier, qui ne relève pas des attributions de la CADA. Ces dispositions prévoient que les obligations du CRPA sont réputées satisfaites par la délivrance, aux candidats qui en font la demande, d'une information portant sur les critères et modalités d'examen de leurs candidatures ainsi que des

Les règles définissant les principaux traitements algorithmiques locaux n'ont pas à être communiquées

14. Cette plateforme permet aux personnes qui souhaitent accéder à une formation du premier cycle de l'enseignement supérieur de formuler des vœux d'inscription et de répondre aux propositions d'admission qui leur sont faites par les établissements dispensant ces formations.

motifs pédagogiques qui justifient la décision prise. Ces dispositions visent à protéger le secret des délibérations des équipes pédagogiques. C'est le choix du législateur. Il n'en demeure pas moins qu'il s'agit d'une dérogation qui ne dit pas son nom. Dans un avis du 10 janvier 2019 (n° 20184400), la CADA n'a pu que déduire de ces dispositions que les règles définissant les principaux traitements algorithmiques locaux n'ont pas à être communiquées, tout en relevant que le code de l'éducation n'interdit pas que les établissements en assurent une diffusion spontanée. Le Conseil d'État a confirmé cette analyse, dans une décision du 12 juin 2019, université des Antilles. Le Conseil constitutionnel a, pour sa part, considéré que ces dispositions sont conformes à la Constitution (décision n° 2020-834 QPC du 3 avril 2020). Il a toutefois formulé une réserve d'interprétation allant dans le sens d'une plus grande transparence. À la lumière de cette décision, dans un avis du 10 septembre 2020, n° 20201766, la CADA a dégagé les principes de communication suivants :

- les candidats ont accès, sur le fondement du code de l'éducation, aux informations relatives aux connaissances et compétences attendues pour la réussite dans la formation, ainsi qu'aux considérations en fonction desquelles les établissements apprécieront les candidatures. Ces informations sont portées à leur connaissance via la plateforme Parcoursup. Les tiers peuvent aussi y avoir accès, sur demande, en application du CRPA ;
- à l'issue de la procédure de sélection, chaque candidat a également le droit d'obtenir la communication par l'établissement, des informations relatives aux critères et modalités d'examen de sa candidature, ainsi que des motifs pédagogiques justifiant la décision ;
- afin de garantir un droit d'accès aux tiers à l'issue de la procédure, chaque établissement est enfin tenu d'élaborer et de publier un rapport exposant les critères en fonction desquels les candidatures ont été examinées et précisant, le cas échéant, *dans quelle mesure des traitements algorithmiques ont été utilisés* pour procéder à cet examen. Saisie d'une demande en ce sens, la CADA peut inviter les établissements à se conformer à cette obligation.

Ces dispositions imposent aux établissements de rendre compte de leurs décisions dans une certaine mesure. Mais elles ne sauraient remplacer l'information prévue par le droit commun, qui permet de connaître le degré et le mode de contribution du traitement algorithmique à la prise de décision, les données traitées et leur source, les paramètres du traitement et le cas échéant leur pondération, appliquée à la situation de l'intéressé, ainsi que les opérations effectuées par le traitement. Une certaine opacité demeure donc, tranchant avec la promotion générale de transparence des documents se rapportant aux algorithmes. Au final, si la transparence est une valeur importante dans nos sociétés contemporaines, celle des algorithmes reste à la maîtrise du législateur, ainsi que, comme dans le cas de Parcoursup, des administrations.

Témoignage de Célia Zolynski, professeur à l'université Paris 1 Panthéon-Sorbonne, personnalité qualifiée au sein de la CNCDH

La Commission nationale consultative des droits de l'homme (CNCDH) a débuté ses travaux consacrés à l'impact de l'intelligence artificielle sur les droits humains en juin 2021, menée 35 auditions de différents experts, issus du monde académique, de représentants de la société civile, du gouvernement et d'instances internationales, et eu des échanges nourris entre les différents membres du groupe de travail dont les horizons variés ont beaucoup apporté à notre réflexion.

Le contexte paraissait justifier que la CNCDH s'autosaisisse sur le sujet de l'impact des algorithmes d'aide à la décision sur les droits fondamentaux. Tout d'abord, en raison du recours de plus en plus important à ces systèmes dans le cadre de processus décisionnels pouvant impacter chaque individu, un groupe d'individus voire l'ensemble de la société. Ensuite dans la mesure où ceux-ci peuvent être source d'atteintes majeures aux droits fondamentaux, notamment le respect de la vie privée et la protection des données à caractère personnel, la non-discrimination, mais également l'accès à la justice, l'accès aux droits sociaux, ou encore le respect de la dignité humaine.

Cette étude s'est appuyée sur les nombreux travaux déjà consacrés à ce sujet essentiel et particulièrement ceux de la CNIL menés depuis son étude de décembre 2017, ou encore du défenseur des droits, notamment consacrés aux risques de discriminations résultant des biais algorithmiques et du recours aux outils biométriques. Il convient d'ajouter les travaux menés dans le cadre du groupe ad hoc sur l'IA du Conseil de l'Europe (CAHAI), chargé d'examiner les possibilités de mettre en place un cadre juridique relatif à l'IA et depuis, le projet de convention dont le contenu a été discuté au sein du « comité sur l'IA » institué à cet effet en janvier 2022 pour garantir le développement, la conception et l'application de l'intelligence artificielle, fondés sur les normes du Conseil de l'Europe en matière de droits de l'homme, de démocratie et de droits de l'homme. Au niveau international, de nombreuses instances ont également alerté sur l'impact de l'IA sur les droits fondamentaux, dont l'UNESCO, dans le cadre de sa recommandation sur l'éthique de l'IA publiée en novembre 2021.

Cette réflexion est également très opportune dans le contexte de l'adoption prochaine du règlement de l'UE sur l'intelligence artificielle (*AI Act*). Ce texte doit en effet, comme le précise son considérant n° 5, fixer un premier cadre juridique régional pour « *favoriser le développement, l'utilisation et l'adoption de l'intelligence artificielle dans le marché intérieur, tout en respectant un niveau élevé de protection des intérêts publics, tels que la santé et la sécurité et la protection des droits fondamentaux, tels que reconnus et protégés par le droit de l'Union* ». Cette réforme en cours est un préalable nécessaire à la prise en compte des droits fondamentaux.

Néanmoins, cette proposition de règlement sur l'IA, en l'état des discussions, ne prévoit pas de garanties propres à assurer le respect effectif des droits fondamentaux. Par conséquent, l'avis adopté du 7 avril 2022 par la CNCDH invite les pouvoirs publics à promouvoir un encadrement juridique ambitieux pour garantir le respect des droits fondamentaux. Il s'agirait alors de retenir une approche fondée sur les droits de l'Homme, qui implique de placer la personne humaine en position de définir ses besoins et, par conséquent, de soutenir le développement d'une IA au service de l'humain et de son autonomie. Parmi les 19 recommandations formulées dans cet avis, outre la définition des « lignes rouges », autrement dit les usages de l'IA à proscrire en raison de leur caractère contraire aux valeurs de l'Union dont les droits fondamentaux, plusieurs recommandations prônent de faire peser sur les utilisateurs, et non pas seulement sur les concepteurs ou fournisseurs d'un système d'IA, des exigences pour garantir le respect des droits fondamentaux. D'autres soulignent la nécessité de reconnaître de nouveaux droits aux personnes ayant fait l'objet d'une décision impliquant un algorithme.

Dans la perspective de cette journée, je reviendrai sur deux points importants. En premier lieu, les travaux qui ont été menés ont mis en exergue la nécessité qu'au delà des obligations à imposer au fournisseur du système d'IA, saisi notamment pour la proposition de règlement IA, il est essentiel de conférer un rôle majeur à l'utilisateur de ce système. Par utilisateur, entendons ici toute personne utilisant sous sa propre autorité un système, notamment les autorités et administrations publiques, agences ou autres organismes. À cet égard, l'avis préconise d'imposer une évaluation de l'impact du recours à ce système sur les droits fondamentaux et de mener une étude d'impact pour le responsabiliser. Cela permettrait d'identifier les risques et les mesures d'atténuation à mettre en œuvre, s'agissant particulièrement des effets discriminatoires et d'imposer également à l'utilisateur de respecter un principe de supervision du système tout au long de son cycle de vie dès lors que des atteintes aux droits fondamentaux peuvent survenir après sa prise en main par l'utilisateur. L'avis préconise en outre de faire dialoguer l'utilisateur avec les parties prenantes en amont du recours au système d'IA, par exemple avec les représentants du personnel ou des groupes de personnes concernées par les décisions prises sur le fondement d'un système d'IA en cas de risque d'atteinte aux droits fondamentaux. Cela pourrait éviter certains choix inappropriés s'agissant du recours à un système d'IA. Par ailleurs, afin de permettre ce dialogue avec l'ensemble des parties prenantes, et parce que les systèmes d'IA touchent l'ensemble de la population, il est fondamental tout d'abord de développer des outils d'information et de formation accessibles à tous, notamment les agents des administrations publiques et représentants du personnel, et d'inclure les enjeux techniques, politiques et sociétaux de l'IA dans les programmes d'éducation à la citoyenneté. Ensuite, il paraît essentiel pour les pouvoirs publics d'organiser des débats publics sur la question afin d'informer les citoyens du fonctionnement de ces systèmes, de leur finalité ainsi que pour les mettre en capacité de

se positionner à l'égard des orientations nationales en la matière.

Le second point concerne la nécessité de réintroduire l'humain dans le processus de décision automatisée pour les décisions individuelles. Au niveau de l'utilisateur, l'avis met l'accent sur la nécessité, pour l'ensemble des traitements algorithmiques produisant des effets sur les droits des personnes, d'imposer une intervention humaine de manière plus générale, à des degrés divers, selon le domaine considéré. Il paraît également essentiel de former et d'informer de façon claire, complète, et compréhensible l'intervenant — le médecin, le juge, les agents administratifs — sur les caractéristiques de l'algorithme utilisé (technologie à l'origine de sa conception, le type de données utilisées pour sa modélisation, les paramètres de fonctionnement et la pondération des critères retenus par le concepteur de l'algorithme, la fiabilité, etc.). Cette information relative aux « rouages de la machine » favoriserait une prise de distance vis-à-vis du système utilisé, et ainsi une réduction du biais cognitif d'automatisation. Par ailleurs, il convient de ne pas faire peser de contrainte particulière sur ces personnes lorsqu'elles s'écartent de l'indication algorithmique. S'agissant de la « personne concernée », l'avis recommande notamment de consacrer de nouveaux droits à la personne exposée à un système algorithmique d'aide à la décision, en particulier un droit à être informé de l'intervention d'un système algorithmique dans la prise de décision, y compris pour les systèmes utilisés par les acteurs privés. L'avis recommande encore de mener une réflexion en ce qui concerne le droit d'être informé des modalités, du processus de la prise de décision, ce qui rejoint le principe d'explicabilité ; ce point est essentiel en ce qu'il garantit l'effectivité du droit au recours à l'encontre des décisions individuelles fondées sur un algorithme.

Synthèse de la journée et remerciements

Par cette journée, nous espérons avoir pu mettre en perspective les défis scientifiques avec les enjeux juridiques et éthiques liés aux algorithmes d'aide à la décision publique. Plus de 150 personnes se sont inscrites pour participer à cette journée sur place à Bordeaux et 115 personnes se sont connectées pour suivre la retransmission en direct. Les enregistrements vidéos des différents intervenants sont accessibles sur le site web de la SIF. Cette journée a pu être organisée grâce au soutien de nos partenaires : université de Bordeaux¹⁵, ministère de l'Enseignement supérieur et de la Recherche¹⁶, Bordeaux métropole¹⁷, Talents du numérique¹⁸, département Sciences de l'ingénierie et du numérique de l'université de Bordeaux¹⁹,

15. <https://www.u-bordeaux.fr>.

16. <https://www.enseignementsup-recherche.gouv.fr/fr>.

17. <https://www.bordeaux-metropole.fr>.

18. <https://talentsdunumerique.com>.

19. <https://www.u-bordeaux.fr/recherche/structures-de-recherche/departements-de-recherche/SIN>.

fédération de recherche sciences informatiques, humaines et sociales du CNRS²⁰, laboratoire bordelais de recherche en informatique²¹, Bordeaux INP²², fondation Anthony Mainguené²³, Inria²⁴, région Nouvelle-Aquitaine²⁵ et Institut des sciences sociales du politique²⁶.



20. <https://www.printemps.uvsq.fr/la-federation-de-recherche-sihs-sciences-informatiques-humaines-et-sociales-1>.

21. <https://www.labri.fr>.

22. <https://www.bordeaux-inp.fr/fr>.

23. <https://www.fondation-anthonymainguene.org>.

24. <https://www.inria.fr/fr>.

25. <https://www.nouvelle-aquitaine.fr>.

26. <https://isp.cnrs.fr>.