REPORT

AUTOMATING SOCIETY

2020

ÉDITION FRANÇAISE

ALGORITHM WATCH

BertelsmannStiftung

INFORMATIONS LÉGALES

Automating Society Report 2020

Février 2021

Disponible en ligne à https://automatingsociety.algorithmwatch.org/report2020/france/

Éditeurs
AlgorithmWatch gGmbH
Linienstr. 13
10178 Berlin
Allemagne

Bertelsmann Stiftung
Carl-Bertelsmann-Str. 256
33311 Gütersloh
Allemagne

Rédacteur·trices
Fabio Chiusi
Sarah Fischer
Nicolas Kayser-Bril
Matthias Spielkamp

Rédacteur de l'édition française
Nicolas Kayser-Bril

Traducteur
Charles Robert

Chef de projet
Fabio Chiusi

Coordinateurs des publications
Marc Thümmler

Dessinateurs
Samuel Daveti
Lorenzo Palloni
Alessio Ravazzani

Mis en page
Beate Autering
Beate Stangl

Rédaction additionnelle
Leonard Haas

Date de mise en page: 30 septembre 2020

Table des matières

Vivre dans une société automatisée. Comment les systèmes de prise de décision automatisée se sont-ils généralisés, et que peut-on y faire ?

Par Fabio Chiusi

La date de bouclage de ce rapport était le 30 septembre 2020.
Les développements ultérieurs n'ont pas pu être inclus.

Par une journée nuageuse d'août, à Londres, les étudiant·es étaient en colère. Ils et elles affluaient par centaines sur Parliament Square pour manifester leur colère, leurs pancartes affichant leur soutien pour des alliés inhabituels : leurs professeurs, et une cible plus insolite encore – un algorithme.

Les écoles du Royaume-Uni avaient fermé leurs portes en mars, en raison de la pandémie de COVID-19. Alors que le virus continuait à sévir à travers l'Europe à l'été 2020, les étudiant·es savaient que leurs examens de fin d'année allaient être annulés et leurs évaluations, d'une manière ou d'une autre, modifiées. Ce qu'ils et elles n'auraient pas pu imaginer, cependant, c'est que des milliers d'entre eux allaient se retrouver avec des notes plus basses que prévu.

Les étudiant·es qui manifestaient savaient quel était le responsable, comme le laissaient entendre leurs chants et leurs pancartes : le système de prise de décision automatisée (ADM, pour *automated decision-making*) mis en place par l'Office of Qualifications and Examinations Regulation (Ofqual). Celui-ci prévoyait de produire la meilleure évaluation possible en se basant sur les données disponibles pour les résultats des deux certificats de fin d'études secondaires, le GCSE et le *A level*, de sorte que « la distribution des notes suive un modèle similaire à celui des autres années, afin que les étudiants de cette année ne se retrouvent pas pénalisés par les circonstances ».

Le gouvernement souhaitait éviter l'excès d'optimisme[1] qui aurait résulté du seul jugement humain, d'après ses propres estimations : par rapport aux séries des années précédentes, les notes auraient été trop élevées. Mais à vouloir être « juste, autant que possible, envers les étudiant·es qui n'ont pas pu passer leurs examens cet été », le gouvernement a essuyé un échec spectaculaire, et en cette grise journée d'août, les étudiant·es continuaient d'affluer, de scander des chants et de brandir des pancartes pour exprimer leur besoin urgent de justice sociale. Certain·es étaient désespérés, d'autres s'effondraient en pleurant.

« Arrêtez de nous voler notre avenir », pouvait-on lire sur une pancarte, faisant écho aux manifestations des « vendredis pour l'avenir » des militant·es écologistes. D'autres, en revanche, ciblaient plus spécifiquement les failles du système de notation ADM : « notez mon travail, pas mon code postal », ou encore « nous sommes des étudiant·es, pas des statistiques », dénonçant les résultats discriminatoires du système[2].

Enfin, un chant jaillit de la foule, un chant qui est devenu le symbole de la contestation : « Fuck the algorithm ». Craignant que le gouvernement n'automatise leur avenir de manière opaque et désinvolte, sans tenir compte de leurs compétences et de leurs efforts, les étudiant·es se mirent à crier pour ne pas voir leurs chances être indûment affectées par un code mal conçu. Ils et elles voulaient avoir leur mot à dire, et nous ferions bien de les écouter.

Les algorithmes ne sont ni « neutres », ni « objectifs » ; pourtant, nous avons tendance à penser qu'ils le sont. En vérité, ils ne font que reproduire les préjugés et les croyances de ceux et celles qui les programment et les déploient. Ces individus sont donc, ou du moins devraient être, responsables des bons et des mauvais choix algorithmiques, non pas les « algorithmes » ou les systèmes ADM eux-mêmes. La machine est effrayante, mais le fantôme en son sein est toujours humain. Et les êtres humains, bien plus encore que les algorithmes, sont des machines complexes.

Dans tous les cas, les étudiant·es contestataires n'étaient pas naïfs au point de croire que leurs malheurs étaient exclusivement le fait d'un algorithme. D'ailleurs, ils et elles n'attaquaient pas l'« algorithme » dans un élan de déterminisme technologique : ils et elles étaient motivés par un désir de protection et de promotion de la justice sociale. À cet égard, leur protestation ressemble davantage à celle des luddites. Tout comme ce mouvement ouvrier qui détruisait les métiers à tisser et à tricoter mécaniques au XIX^{e} siècle, ils et elles savent que les systèmes ADM sont purement une histoire de pouvoir, et ne doivent pas être considérés comme une technologie prétendument objective. Alors, ils se mirent à scander « justice pour la classe ouvrière » et à demander la démission du ministre de la Santé, décrivant le système ADM comme étant une preuve flagrante de « classisme ».

Pour finir, les étudiant·es parvinrent à abolir le système qui mettait en danger leur parcours éducatif et leurs chances dans la vie : dans un revirement spectaculaire, le gouvernement britannique abandonna le système ADM, propice aux erreurs, et utilisa les notes prédites par les enseignant·es.

1 « Les travaux de recherche suggèrent que, en estimant les notes que les élèves sont susceptibles d'obtenir, les enseignants ont tendance à être optimistes (quoique pas dans tous les cas) », écrit l'Ofqual, cf. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/909035/6656-2_-_Executive_summary.pdf

2 Cf. le chapitre sur le Royaume-Uni pour plus de détails.

Mais cette histoire ne se résume pas au fait que les manifestant·es ont finalement eu gain de cause. Cet exemple illustre parfaitement comment des systèmes mal conçus, mis en œuvre et supervisés, qui reproduisent les préjugés et la discrimination de leurs créateurs, ne tirent pas parti du potentiel des systèmes ADM, par exemple en matière de comparabilité et d'équité.

Cette contestation révèle, plus clairement que bien d'autres luttes du passé, que nous ne nous contentons plus d'automatiser la société. Nous l'avons déjà automatisée – et quelqu'un l'a enfin remarqué.

/ De l'automatisation de la société à la société automatisée

Lorsque nous avons publié la première édition de ce rapport, nous avions décidé de l'appeler « *L'automatisation de la société* », car les systèmes ADM étaient pour l'essentiel nouveaux, expérimentaux et inexplorés – et surtout, ils étaient l'exception plutôt que la norme.

Cette situation a rapidement évolué. Comme le révèlent clairement les nombreux cas réunis dans ce rapport grâce à notre réseau exceptionnel de chercheur·ses, le déploiement des systèmes ADM a considérablement augmenté en à peine plus d'un an. Ceux-ci touchent désormais toutes sortes d'activités humaines, et plus particulièrement la distribution des services à des millions de citoyen·nes européens – ainsi que leur accès à leurs droits.

L'opacité tenace qui entoure l'utilisation toujours croissante des systèmes ADM nous oblige d'autant plus urgemment à redoubler d'efforts. C'est pourquoi nous avons ajouté quatre pays (l'Estonie, la Grèce, le Portugal et la Suisse) aux 12 que nous avions déjà analysés dans l'édition précédente de ce rapport, portant le total à 16 pays. Bien que cette liste soit loin d'être exhaustive, cela nous permet de brosser un tableau plus large des scénarios d'ADM à travers l'Europe. Compte tenu de l'impact que ces systèmes peuvent avoir sur notre vie quotidienne, et de la profondeur avec laquelle ils remettent en question nos intuitions – voir nos normes et nos règles – sur la relation entre la gouvernance démocratique et l'automatisation, nous pensons qu'il s'agit là d'une initiative essentielle.

Cela s'avère particulièrement vrai dans le contexte de la pandémie de COVID-19, une période au cours de laquelle nous avons assisté à l'adoption (souvent précipitée) d'une multitude de systèmes ADM visant à contribuer à la sécurisation de la santé publique grâce à des outils basés sur des données et à l'automatisation. Nous considérons que cette évolution est si importante que nous avons décidé d'y consacrer un « rapport préliminaire », publié en août 2020 dans le cadre du projet *L'automatisation de la société*.

Même en Europe, les exemples de systèmes ADM déployés sont légion. Songez simplement à certains des cas présentés dans ce rapport, qui viennent s'ajouter aux nombreux cas – de la sécurité sociale à l'éducation, en passant par le système de santé et la justice – que nous avions déjà abordés dans l'édition précédente. Dans les pages qui suivent, et pour la première fois, nous faisons le point sur le développement de ces cas de trois manières. Tout d'abord, par le biais d'articles journalistiques, puis par des sections de recherche cataloguant différents exemples, et enfin, par des bandes dessinées. Ces systèmes ADM sont devenus tellement omniprésents dans nos vies que nous voulions communiquer leur fonctionnement et ce qu'ils *font réellement pour nous*, de manière à la fois rigoureuse et innovante, afin de toucher toutes sortes de publics. Après tout, les systèmes ADM ont un impact sur nous tous·tes.

Ou du moins, ils le devraient. Nous avons pu voir, par exemple, comment un nouveau service automatisé et proactif permet de distribuer les allocations familiales en Estonie. Les parents n'ont même plus besoin de demander ces allocations : dès la naissance, l'État recueille toutes les informations sur chaque nouveau-né et ses parents et les rassemble dans des bases de données. Ainsi, les parents reçoivent automatiquement les prestations auxquelles ils ont droit.

En Finlande, l'identification des facteurs de risque individuels liés à l'exclusion sociale chez les jeunes adultes est automatisée grâce à un outil développé par le géant japonais Fujitsu. En France, les données des réseaux sociaux peuvent être extraites pour alimenter des algorithmes à apprentissage automatique qui sont employés pour détecter la fraude fiscale.

L'Italie teste actuellement la « jurisprudence prédictive ». Cette méthode a recours à l'automatisation pour aider les juges à comprendre les tendances de décisions de justice précédentes sur un sujet particulier. Et au Danemark, le gouvernement a voulu surveiller le clavier et la souris de chaque étudiant·e pendant les examens, ce qui a entraîné – une fois n'est pas coutume – des manifestations d'étudiant·es massives qui ont conduit au retrait du système, du moins pour l'instant.

/ Redressons les torts de l'ADM

En principe, les systèmes ADM sont susceptibles d'améliorer la vie des gens en traitant d'énormes quantités de données, en aidant les personnes impliquées dans des processus décisionnels et en fournissant des applications sur mesure.

En pratique, cependant, nous avons trouvé très peu de cas qui démontraient de manière convaincante un tel impact positif.

Parmi ceux-ci, le système VioGén, déployé en Espagne depuis 2007 pour évaluer les risques dans les affaires de violence domestique, bien qu'il soit loin d'être parfait, affiche des « indices de performance raisonnables » et a contribué à protéger de nombreuses femmes contre la maltraitance à leur encontre.

Au Portugal, un système automatisé centralisé déployé pour dissuader la fraude aux prescriptions médicales, a apparemment réduit la fraude de 80 % en une seule année. En Slovénie, un système similaire utilisé pour lutter contre la fraude fiscale s'est révélé utile pour les inspecteurs, selon les autorités fiscales[3].

3 Cf. le chapitre sur la Slovénie pour plus de détails.

La reconnaissance faciale, qui était pratiquement absente de l'édition 2019, est testée et déployée à un rythme alarmant dans toute l'Europe.

Lorsque l'on regarde l'état actuel des systèmes ADM en Europe, les exemples positifs présentant des avantages évidents se font rares. Tout au long de ce rapport, nous décrivons comment la grande majorité des utilisations a plutôt tendance à nuire aux gens qu'à les aider. Mais pour juger véritablement de l'impact positif et négatif de ces systèmes, nous avons besoin de plus de transparence sur la finalité et de plus de données sur le fonctionnement des systèmes ADM qui sont testés et déployés.

Le message destiné aux responsables politiques ne pourrait être plus clair. Si nous souhaitons vraiment tirer le meilleur parti de leur potentiel tout en respectant les droits fondamentaux et la démocratie, le moment est venu de passer à l'action, de rendre ces systèmes transparents et de remédier aux injustices de l'ADM.

/ La reconnaissance faciale à chaque coin de rue

Différents pays adoptent différents outils. Il y a toutefois une technologie qui est commune dans la plupart d'entre eux : la reconnaissance faciale. Il s'agit sans doute du développement le plus récent, le plus rapide et le plus préoccupant présenté dans ce rapport. La reconnaissance faciale, qui était pratiquement absente de l'édition 2019, est testée et déployée à un rythme alarmant dans toute l'Europe. En à peine plus d'un an depuis notre dernier rapport, la reconnaissance faciale a fait son apparition dans les écoles, les stades, les aéroports et même les casinos. Elle est également utilisée dans des applications de police prédictive, pour appréhender les criminels, pour lutter contre le racisme, et, dans le cadre de la pandémie de COVID-19, pour faire respecter la distanciation sociale, à la fois par le biais d'applications et de systèmes de vidéosurveillance « intelligente ».

Les nouveaux déploiements de l'ADM se poursuivent, alors même que les preuves de leur manque de précision s'accumulent. Et lorsque des difficultés apparaissent, les partisans de ces systèmes essaient simplement de les dissimuler d'une manière ou d'une autre. En Belgique, un système de reconnaissance faciale utilisé par la police est toujours « partiellement actif », bien qu'une interdiction temporaire ait été prononcée par l'Organe de contrôle de l'information policière. Et en Slovénie, l'utilisation de la technologie de reconnaissance faciale par la police a été légalisée cinq ans après avoir fait ses débuts.

Cette tendance, si elle n'est pas remise en cause, risque de normaliser l'idée que l'on peut être surveillé en permanence et de manière opaque, cristallisant ainsi un nouveau statu quo de surveillance de masse généralisée. C'est la raison pour laquelle de nombreux membres de la communauté des libertés civiles auraient souhaité voir une réponse politique beaucoup plus forte de la part des institutions européennes face à cette situation[4].

Même le simple fait de sourire fait désormais partie d'un système ADM employé dans certaines banques en Pologne : plus un·e employé·e sourit, plus sa prime est importante. Et ce ne sont pas seulement les visages qui sont surveillés : en Italie, un système de surveillance sonore a été proposé pour lutter contre le racisme dans les stades de football.

/ Les boîtes noires sont toujours des boîtes noires

En 2015, Frank Pasquale, professeur à la faculté de droit de Brooklyn, disait qu'une société interconnectée basée sur des systèmes algorithmiques opaques était une sorte de « boîte noire ». Cinq ans plus tard, malheureusement, la métaphore reste valable – et elle s'applique à tous les pays que nous avons étudiés dans le cadre de ce rapport, dans tous les domaines : la transparence des systèmes ADM est insuffisante, que ce soit dans le secteur public ou privé. En Pologne, cette opacité est même imposée, comme en témoigne la loi qui a instauré son système automatisé de détection des comptes bancaires utilisés à des fins illégales (« STIR »). En effet, la loi dispose que la divulgation des algorithmes et des indicateurs de risque employés peut entraîner jusqu'à 5 ans de prison.

Bien que nous rejetions catégoriquement l'idée que tous ces systèmes sont intrinsèquement mauvais (nous adoptons plutôt une approche objective et factuelle), il est incontestablement dommageable de ne pas pouvoir évaluer leur fonctionnement et leur impact sur la base de connaissances précises et factuelles. Ne serait-ce que parce que l'opacité entrave sérieusement la collecte des preuves nécessaires pour pouvoir porter un jugement éclairé sur le déploiement d'un système ADM.

Ajoutez à cela la difficulté que nos chercheurs et nos journalistes ont rencontrée pour accéder à des données sur ces systèmes, et vous obtenez un scénario préoccupant pour quiconque souhaite les contrôler et s'assurer que leur déploiement soit compatible avec les droits fondamentaux, l'État de droit et la démocratie.

/ Remettre en question le status quo algorithmique

Que fait l'Union européenne à ce sujet ? Si les documents stratégiques produits par la Commission européenne, sous l'égide d'Ursula Von der Leyen, parlent d'« intelligence artificielle » plutôt que de faire directement référence aux systèmes ADM, ils expriment toutefois des intentions louables : promouvoir et développer une « IA digne de confiance », qui accorde la « priorité aux individus »[5].

Cependant, comme nous le décrivons dans le chapitre consacré à l'UE, l'approche globale de l'UE privilégie l'impératif commercial et géopolitique visant à mener la « révolution de l'IA » plutôt que de s'assurer que ses produits soient conformes aux mécanismes de protection démocratiques, une fois adoptés comme outils politiques.

Cette absence de courage politique, qui est particulièrement évidente dans la décision d'abandonner toute suggestion de moratoire sur les technologies de reconnaissance faciale dans les lieux publics dans les règlements européens sur l'IA, est surprenante. Surtout à une époque où de nombreux États membres se voient confrontés à un nombre croissant de difficultés et de revers juridiques en raison de systèmes ADM déployés à la va-vite qui ont eu un impact négatif sur les droits des citoyens.

Une affaire historique nous vient des Pays-Bas, où des défenseurs des droits civils ont porté devant les tribunaux un système automatisé invasif et opaque censé détecter la fraude aux aides sociales (SyRI), et ont obtenu gain de cause. En effet, ce système a été jugé contraire à la Convention européenne des droits de l'homme par un tribunal de la Haye en février, et a donc été abandonné. Mais l'affaire a également fait jurisprudence : selon l'arrêt de la cour, les gouvernements ont la « responsabilité particulière » de protéger les droits fondamentaux lorsqu'ils mettent en œuvre de tels systèmes ADM. Assurer cette transparence indispensable est considéré comme un élément crucial de cette responsabilité.

Depuis notre premier rapport, les médias et les militant·es de la société civile se sont imposés comme une véritable

4 Comme détaillé dans le chapitre sur l'UE.

5 Cf. le chapitre sur l'UE, et en particulier la section sur le « livre blanc sur l'IA » de la Commission européenne.

Devant le déploiement continu de systèmes d'ADM à travers l'Europe, on est en droit de se demander : le niveau de supervision actuel est-il suffisant ?

force motrice de la responsabilisation vis-à-vis des systèmes ADM. En Suède, par exemple, des journalistes sont parvenus à obtenir la publication du code du système Trelleborg, qui permet de prendre des décisions entièrement automatisées concernant les demandes de prestations sociales. À Berlin, le projet pilote de reconnaissance faciale de la gare de Südkreuz n'a pas réussi à déboucher sur l'application du système dans toute l'Allemagne. Cette issue n'a été possible que grâce à la bruyante opposition des militant·es, si bruyante qu'ils et elles sont parvenus à influencer la position des partis et, en fin de compte, le programme politique des gouvernements.

Les militants grecs d'Homo Digitalis ont pu démontrer qu'aucun vrai voyageur n'avait participé aux essais du système nommé « iBorderCtrl », un projet financé par l'UE qui visait à utiliser l'ADM pour les contrôles aux frontières, révélant ainsi que les capacités de beaucoup de ces systèmes sont souvent surestimées. Dans le même temps, au Danemark, un système de profilage pour la détection précoce des risques associés aux familles et aux enfants vulnérables (le « modèle de Gladsaxe ») a été mis en suspens grâce au travail d'universitaires, de journalistes et de l'Autorité de protection des données (APD) nationale.

Les APD elles-mêmes ont également joué un rôle majeur dans d'autres pays. En France, la CNIL, autorité nationale de protection de la vie privée, a statué qu'un projet de surveillance sonore et un projet de reconnaissance faciale dans des lycées étaient tous deux illégaux. Au Portugal, l'APD a refusé d'approuver la mise en place de systèmes de vidéosurveillance par la police dans les municipalités de Leiria et Portimão, car ceux-ci ont été jugés disproportionnés et auraient constitué « une surveillance et un suivi à grande échelle des personnes, de leurs habitudes et de leur comportement, ainsi qu'une identification des personnes à partir de données relatives à leurs caractéristiques physiques ». Parallèlement, aux Pays-Bas, l'APD néerlandaise a demandé plus de transparence dans les algorithmes prédictifs utilisés par les agences gouvernementales.

Enfin, certains pays ont eu recours à un médiateur pour se faire conseiller. Au Danemark, ces conseils ont permis d'élaborer des stratégies et des orientations éthiques concernant l'utilisation des systèmes ADM dans le secteur public. En Finlande, le médiateur parlementaire adjoint a estimé que les évaluations fiscales automatisées étaient illégales.

Et pourtant, devant le déploiement continu de ces systèmes à travers l'Europe, on est en droit de se demander : ce niveau de supervision est-il suffisant ? Lorsque le médiateur polonais a mis en doute la légalité du système de détection des sourires utilisé dans une banque (et mentionné ci-dessus), sa décision n'a pas empêché un projet pilote ultérieur d'être entrepris dans la ville de Sopot, ni dissuadé plusieurs entreprises de manifester leur intérêt pour l'adoption de ce système.

/ L'inadéquation des audits, de la mise en application, des compétences et des explications

Le militantisme est principalement une démarche réactive. La plupart du temps, les militant·es ne peuvent réagir que si un système ADM est en train d'être testé ou s'il a déjà été déployé. Et le temps que les citoyen·nes mettent sur pied une réponse, il se peut que leurs droits aient déjà été indûment piétinés, même avec les protections cen-

sées être accordées par le droit européen et la législation des États membres. C'est pourquoi il est si important de prendre des mesures proactives pour protéger les droits des citoyen·nes, avant la réalisation de projets pilotes et de déploiements à grande échelle.

Et pourtant, même dans les pays où une législation préventive est en place, celle-ci n'est pas appliquée. En Espagne, par exemple, toute « action administrative automatisée » est codifiée par la loi, qui prévoit des exigences spécifiques en matière de contrôle de la qualité et de supervision, ainsi que l'audit du système informatique et de son code source. L'Espagne bénéficie également d'une loi sur la liberté de l'information. Cependant, même avec ces lois, il est rare, d'après notre chercheur, que les organismes publics divulguent des informations détaillées sur les systèmes ADM qu'ils utilisent. De même, en France, il existe bien une loi de 2016 qui exige la transparence des algorithmes, mais là encore, en vain.

Même le fait d'ester en justice pour obtenir la transparence d'un algorithme, conformément aux dispositions spécifiques d'une loi sur la transparence des algorithmes, peut ne pas suffire à faire respecter et à protéger les droits des utilisateurs. Comme en témoigne le cas de l'algorithme de Parcoursup, destiné à classer et trier les candidats à l'université en France[6], des exceptions peuvent être prévues à la discrétion du législateur pour dégager une administration de toute responsabilité.

Ce phénomène est particulièrement troublant lorsqu'il vient s'ajouter au déficit généralisé d'aptitudes et de compétences entourant les systèmes ADM dans le secteur public, déploré par de nombreux·ses chercheur·ses. Comment les responsables publics pourraient-ils expliquer ou faire preuve d'une quelconque transparence à l'égard de systèmes qu'ils ne comprennent pas eux-mêmes ?

Récemment, plusieurs pays se sont efforcés de résoudre ce problème. L'Estonie, par exemple, a mis en place un centre de compétences consacré aux systèmes ADM afin de mieux déterminer comment ceux-ci pourraient être utilisés pour développer les services publics et, plus particulièrement, pour guider l'action du ministère des Affaires économiques et des Communications et de la chancellerie d'État pour le développement de l'administration en ligne. La Suisse a également appelé à la création d'un « réseau de compétences » dans le cadre plus large de la stratégie nationale de « Suisse numérique ».

Et pourtant, le manque de culture numérique est un problème bien connu qui touche une grande partie de la population dans plusieurs pays européens. En outre, il est difficile de faire valoir des droits dont on ignore l'existence. Les mouvements de contestation au Royaume-Uni et ailleurs, conjugués à quelques scandales très médiatisés impliquant des systèmes ADM[7], ont certainement sensibilisé la population aux risques et au potentiel de l'automatisation de la société. Mais bien qu'elle soit en hausse, cette prise de conscience n'en est qu'à ses débuts dans de nombreux pays.

Les résultats de notre étude sont clairs : si les systèmes ADM affectent déjà toutes sortes d'activités et de jugements, ils sont encore principalement déployés sans aucune sorte de débat démocratique significatif. Par ailleurs, on observe que dans l'ensemble, les mécanismes d'application et de contrôle – si tant est qu'ils existent – sont à la traîne par rapport au déploiement

La finalité même de ces systèmes n'est pas communément justifiée ou expliquée aux populations concernées, sans parler des bénéfices qu'elles sont censées en retirer. Prenons l'exemple du service proactif « AuroraAI » en Finlande : celui-ci est censé identifier automatiquement certains « événements de la vie », comme le rapportent nos chercheurs finlandais, et dans l'esprit de ses promoteurs, il doit en quelque sorte jouer le rôle d'une « nounou » qui aide les citoyens à répondre à des besoins particuliers de service public pouvant survenir en lien avec certaines circonstances de la vie, par exemple un déménagement, un changement de relations familiales, etc. Selon nos chercheurs, il est vraisemblable que ce système, au lieu de responsabiliser les individus, fasse exactement le contraire, en suggérant certaines décisions ou en limitant les options d'un individu en raison de sa conception et de son architecture.

Il est alors d'autant plus important de savoir ce que le système vise à « optimiser » en termes de services publics : « l'utilisation du service est-elle maximisée, les coûts sont-ils minimisés, le bien-être des citoyen·nes est-il amélioré ? », demandent les chercheurs. « Sur quel ensemble de critères ces décisions se fondent-elles, et qui les choisit ? » Le simple fait que n'ayons pas de réponse à ces questions

6 Cf. le chapitre sur la France

7 Voir la débâcle de l'algorithme « Buona Scuola », cf. le chapitre sur l'Italie.

fondamentales en dit long sur le degré de participation et de transparence qui est admis, même pour un système ADM potentiellement si intrusif.

/ Le piège technosolutionniste

Il existe une justification idéologique globale à tout cela. C'est ce que l'on appelle le « solutionnisme technologique », et c'est un phénomène qui affecte encore sérieusement la façon dont sont développés de nombreux systèmes ADM que nous avons étudiés. Même si cette expression est depuis longtemps dénoncée comme une idéologie fallacieuse qui perçoit chaque problème social comme un « bug » qui nécessite un « correctif » technologique[8], cette rhétorique est encore largement employée, tant dans les médias que dans les milieux politiques, pour justifier l'adoption inconditionnelle de technologies automatisées dans la vie publique.

Lorsqu'ils sont vendus comme des « solutions », les systèmes ADM passent immédiatement dans le domaine décrit par la troisième loi d'Arthur C. Clarke : la magie. Et il est difficile, voire impossible, de réglementer la magie, et plus encore de l'expliquer et de faire preuve de transparence à son égard. On peut voir la main qui se glisse dans le chapeau, et le lapin qui en ressort, mais le processus lui-même est, et *doit rester* une « boîte noire ».

De nombreux chercheurs impliqués dans le projet *L'automatisation de la société* ont dénoncé ce problème comme étant la faille fondamentale dans le raisonnement qui sous-tend nombre des systèmes ADM qu'ils décrivent. Cela implique également, comme le montre le chapitre sur l'Allemagne, que la plupart des critiques de ces systèmes sont présentées comme un rejet total de l'« innovation », dépeignant les défenseurs des droits numériques comme des « néo-luddites ». Non seulement cette attitude ignore la réalité historique du mouvement luddite, qui se préoccupait des politiques du travail et non des technologies en tant que telles, mais surtout, elle menace fondamentalement l'efficacité des mécanismes de supervision et d'application potentiels.

À l'heure où l'industrie de l'« IA » assiste à l'émergence d'un secteur de lobbying « dynamique », notamment au Royaume-Uni, cette tendance risque d'aboutir à des directives de « blanchiment éthique » et à d'autres réponses politiques qui seront inefficaces et structurellement inadaptées pour traiter les implications des systèmes ADM en matière de droits fondamentaux. Cette vision revient en définitive à supposer que nous, humains, devrions nous adapter aux systèmes ADM, bien plus que les systèmes ADM ne devraient être adaptés aux sociétés démocratiques.

Pour contrer ce raisonnement, nous ne devons pas nous abstenir de poser des questions fondamentales : les systèmes ADM peuvent-ils être compatibles avec la démocratie et déployés au profit de la société dans son ensemble, et pas seulement d'une partie de celle-ci ? Il se pourrait que certaines activités humaines – par exemple, dans le domaine de l'aide sociale – ne doivent pas faire l'objet d'une automatisation, ou que certaines technologies, notamment la reconnaissance faciale dans l'espace public, ne doivent pas être encouragées dans une quête sans fin de « leadership technologique », mais plutôt qu'elles soient interdites dans leur ensemble.

Plus encore, nous devons rejeter tout carcan idéologique qui nous empêche de poser de telles questions. Au contraire, ce dont nous avons besoin maintenant, c'est de voir les politiques changer concrètement, afin de permettre un meilleur contrôle de ces systèmes. Dans la section suivante, nous énumérons les principales exigences qui découlent de nos conclusions. Nous espérons qu'elles seront largement débattues pour être enfin mises en œuvre.

Ce n'est qu'à travers un débat démocratique informé, inclusif et étayé par des preuves que nous pourrons trouver le bon équilibre entre les avantages que les systèmes ADM peuvent apporter – et apportent – en termes de rapidité, d'efficacité, d'équité, de prévention et d'accès aux services publics, et les défis qu'ils représentent pour nos droits à tous.

Recommandations politiques

À la lumière des conclusions détaillées figurant dans l'édition 2020 du rapport *L'automatisation de la société*, nous recommandons les interventions politiques suivantes aux décideur·ses du Parlement européen et des parlements des États membres, de la Commission européenne, des gouvernements nationaux, ainsi qu'aux chercheur·ses et

8 Lire Evgeny Morozov (2014), *To Save Everything, Click Here. The Folly of Technological Solutionism*, Public Affairs, https://www.publicaffairsbooks.com/titles/evgeny-morozov/to-save-everything-click-here/9781610393706/

organisations de la société civile (organisations de défense des droits, fondations, syndicats, etc.) et du secteur privé (entreprises et associations professionnelles). Ces recommandations visent à mieux garantir que les systèmes ADM actuellement déployés et ceux qui sont en passe d'être mis en œuvre à travers l'Europe sont effectivement compatibles avec les droits fondamentaux et la démocratie :

1 Accroître la transparence des systèmes ADM

Sans être en mesure de savoir précisément comment, pourquoi et à quelle fin les systèmes ADM sont mis en place, tous les efforts visant à concilier les droits fondamentaux et les systèmes ADM sont voués à l'échec.

/ Établir des registres publics pour les systèmes ADM utilisés dans le secteur public

Nous demandons par conséquent qu'une législation soit adoptée au niveau de l'UE pour obliger les États membres à tenir des registres publics des systèmes ADM utilisés dans le secteur public.

Ces registres devront être assortis de l'obligation légale pour les responsables du système ADM de divulguer et de documenter la finalité du système, de donner une explication du modèle et sa logique sous-jacente ainsi que des informations sur les personnes qui ont développé le système. Ces informations devront être mises à disposition de manière facilement lisible et accessible, y compris sous forme de données numériques structurées basées sur un protocole standardisé.

Les autorités publiques ont la responsabilité particulière de faire la transparence sur les caractéristiques opérationnelles des systèmes ADM déployés dans l'administration publique. Cette nécessité a été mise en évidence par une récente plainte administrative en Espagne, qui fait valoir que « tout système ADM utilisé par l'administration publique devrait être rendu public par défaut ». Si elle est confirmée, cette décision pourrait faire jurisprudence en Europe.

Si les dispositifs de divulgation des systèmes ADM devraient être obligatoires pour le secteur public dans tous les cas, ces exigences de transparence devraient également s'appliquer à l'utilisation des systèmes ADM par des entités privées lorsqu'un système d'IA/ADM a un impact significatif sur un individu, un groupe spécifique ou la société dans son ensemble.

/ Créer des dispositifs d'accès aux données juridiquement contraignants pour soutenir et faciliter la recherche d'intérêt public

Pour accroître la transparence d'un système, il ne suffit pas de divulguer des informations sur sa finalité, sa logique et son créateur, et d'avoir la capacité d'analyser et de tester en profondeur ce qui entre et sort d'un système ADM. Il faut également rendre les données à partir desquelles le système a été entraîné aux chercheur·ses indépendant·es, aux journalistes et aux organisations de la société civile pour encourager la recherche d'intérêt public.

C'est pourquoi nous suggérons de créer des dispositifs d'accès aux données robustes et juridiquement contraignants, explicitement axés sur le soutien et la promotion de la recherche d'intérêt public, dans le respect de la législation sur la protection des données et de la vie privée.

En tirant les leçons des meilleures pratiques aux niveaux national et européen, ces dispositifs à plusieurs niveaux devraient inclure des systèmes de sanctions, de contrôles et de contre-pouvoirs, ainsi que des examens réguliers. Comme l'ont illustré les partenariats de partage de données privées, des préoccupations légitimes ont été exprimées concernant la vie privée des utilisateurs et la possibilité d'une désanonymisation de certains types de données.

Les responsables politiques ont tout intérêt à s'inspirer des dispositifs de partage des données de santé pour faciliter un accès privilégié à certains types de données plus détaillées, tout en veillant à ce que les données à caractère personnel soient protégées de manière adéquate (par exemple, grâce à des environnements d'exploitation sécurisés).

Bien qu'un cadre de responsabilisation efficace nécessite un accès transparent aux données de la plateforme, il s'agit là d'une exigence pour que de nombreuses méthodes d'audit soient également efficaces.

2 Instaurer un cadre de responsabilisation significatif pour les systèmes ADM

Comme l'ont montré les constatations faites en France et en Espagne, même si la transparence d'un système ADM

est exigée par la loi et/ou si des informations ont été divulguées, cela n'entraîne pas nécessairement de responsabilisation. Des mesures supplémentaires sont nécessaires pour garantir que les lois et les normes soient effectivement applicables.

/ Développer et établir des approches pour auditer efficacement les systèmes algorithmiques

Pour que la transparence ait un sens, nous devons compléter la première étape, qui consiste à établir un registre public, par des processus qui contrôlent efficacement les systèmes algorithmiques.

Le terme « audit » est largement utilisé, mais il n'y a pas de consensus sur sa définition. Dans ce contexte, nous entendons par « audit », conformément à la définition de l'ISO, un « processus systématique, indépendant et documenté visant à obtenir des preuves objectives et à les évaluer objectivement afin de déterminer dans quelle mesure les critères d'audit sont remplis ».

Nous n'avons pas encore de réponses satisfaisantes aux questions complexes[9] soulevées par l'audit des systèmes algorithmiques ; cependant, nos recherches indiquent clairement la nécessité de trouver des réponses dans le cadre d'un vaste processus d'engagement des parties prenantes et par des recherches dédiées et approfondies.

Les critères d'audit, tout comme les processus d'audit appropriés, doivent être élaborés selon une approche multipartite qui prenne activement en considération l'effet disproportionné qu'ont les systèmes ADM sur les groupes vulnérables et sollicite leur participation.

Nous demandons donc aux responsables politiques de mettre en place ces processus multipartites afin de clarifier les questions soulevées, et de mettre à disposition des sources de financement visant à permettre la participation des parties prenantes qui ont été jusqu'à présent mal représentées.

Nous demandons également la mise à disposition de ressources adéquates pour soutenir/financer des projets de recherche sur l'élaboration de modèles permettant de contrôler efficacement les systèmes algorithmiques.

/ Soutenir les organisations de la société civile en tant que gardiens des systèmes ADM

Nos conclusions indiquent clairement que le travail des organisations de la société civile est crucial pour lutter efficacement contre l'opacité des systèmes ADM. Par le biais de la recherche et du travail de sensibilisation, et souvent en coopération avec les universitaires et les journalistes, celles-ci sont intervenues à plusieurs reprises dans les débats politiques portant sur ces systèmes au cours des dernières années, veillant dans plusieurs cas à ce que l'intérêt public et les droits fondamentaux soient dûment pris en compte avant et après leur déploiement dans de nombreux pays européens.

Les acteurs de la société civile devraient donc être soutenus en tant que gardiens de l'« automatisation de la société ». En tant que tels, ils font partie intégrante de tout cadre de responsabilisation efficace pour les systèmes ADM.

/ Interdire la reconnaissance faciale qui pourrait équivaloir à une surveillance de masse

Tous les systèmes ADM ne sont pas aussi dangereux les uns que les autres, et une approche de la réglementation basée sur le risque, comme celle de l'Allemagne et de l'UE, reflète bien ce constat. Mais afin d'assurer une responsabilisation réaliste pour les systèmes identifiés comme risqués, des mécanismes efficaces de surveillance et de mise en œuvre doivent être mis en place. Cela est d'autant plus

9 En réfléchissant aux modèles potentiels d'audits algorithmiques, plusieurs questions se posent. 1) Qui/quoi (services/plateformes/produits) doit être audité ? Comment personnaliser les systèmes d'audit en fonction du type de plateforme/service ? 2) Quand un audit doit-il être entrepris par une institution publique (au niveau de l'UE, au niveau national, au niveau local), et quand peut-il être réalisé par des entités/experts privés (entreprises, société civile, chercheurs) ? 3) Comment clarifier la distinction entre l'évaluation de l'impact ex ante (c'est-à-dire au cours de la phase de conception) et ex post (c'est-à-dire en cours d'exploitation) et les défis respectifs ? 4) Comment évaluer les compromis entre les différents avantages et inconvénients de l'auditabilité (par exemple, la simplicité, la généralité, l'applicabilité, la précision, la flexibilité, l'interprétabilité, la confidentialité, l'efficacité d'une procédure d'audit peuvent être en tension) ? 5) Quelles informations doivent être disponibles pour qu'un audit soit efficace et fiable (par exemple, le code source, les données de formation, la documentation) ? Les auditeurs doivent-ils disposer d'un accès physique aux systèmes en cours de fonctionnement pour pouvoir effectuer un audit efficace ? 6) Quelle obligation de produire des preuves est nécessaire et proportionnée pour les vendeurs/prestataires de services ? 7) Comment s'assurer que l'audit est possible ? Les exigences en matière d'audit doivent-elles être prises en compte dans la conception des systèmes algorithmiques (« auditable par construction ») ? 8) Règles de publicité : lorsqu'un audit est négatif et que les problèmes ne sont pas résolus, quel doit être le comportement de l'auditeur, et dans quelle mesure cet échec peut-il être rendu public ? 9) Qui audite les auditeurs ? Comment s'assurer que les auditeurs soient tenus responsables ?

important pour les systèmes présentant un « risque élevé » de violation des droits des utilisateur·trices.

Un exemple crucial qui ressort de nos conclusions est la reconnaissance faciale. Il a été démontré que les systèmes ADM basés sur les technologies biométriques, notamment la reconnaissance faciale, constituent une menace particulièrement grave pour l'intérêt public et les droits fondamentaux, car ils ouvrent la voie à une surveillance massive et sans discernement – d'autant plus qu'ils sont malgré tout déployés à grande échelle et de manière opaque.

Nous appelons à ce que les utilisations publiques de la reconnaissance faciale qui pourraient équivaloir à une surveillance de masse soient interdites de manière décisive jusqu'à nouvel ordre, et de toute urgence, au niveau de l'UE.

Ces technologies peuvent même déjà être considérées comme illégales dans l'UE, au moins pour certaines utilisations, si elles sont déployées sans le « consentement spécifique » des sujets contrôlés. Cette interprétation juridique a été suggérée par les autorités belges, qui ont infligé une amende historique pour les déploiements de la reconnaissance faciale dans le pays.

3 Sensibiliser la population au sujet des algorithmes et renforcer le débat public sur les systèmes ADM

Une plus grande transparence des systèmes ADM ne sera véritablement utile que si ceux qui y sont confrontés, tels que les organismes de réglementation, le gouvernement et les organismes industriels, peuvent gérer ces systèmes et leur impact d'une manière responsable et prudente. En outre, les personnes concernées par ces systèmes doivent être en mesure de comprendre où, pourquoi et comment ces systèmes sont déployés. C'est pourquoi nous devons améliorer la connaissance des algorithmes à tous les niveaux, auprès des acteurs importants ainsi que du grand public, et favoriser des débats publics plus diversifiés sur les systèmes ADM et leur impact sur la société.

/ Établir des centres d'expertise indépendants sur l'ADM

Parallèlement à notre demande d'audit des algorithmes et de soutien à la recherche, nous appelons à la création de centres d'expertise indépendants sur l'ADM au niveau national pour surveiller, évaluer, mener des recherches, rédiger des rapports et fournir des conseils aux gouvernements et à l'industrie en coordination avec les organismes de réglementation, la société civile et les universités sur les implications sociétales et en matière de droits de l'homme de l'utilisation des systèmes ADM. Le rôle général de ces centres sera de créer un système de responsabilisation significatif et de renforcer les capacités.

Les centres nationaux d'expertise doivent impliquer les organisations de la société civile, les groupes de parties prenantes et les organismes chargés de l'application existants tels que les DPA et les organismes nationaux de défense des droits fondamentaux afin de profiter à tous les aspects de l'écosystème et de renforcer la confiance, la transparence et la coopération entre tous les acteurs.

En tant qu'organes officiels indépendants, les centres d'expertise joueraient un rôle central dans la coordination de l'élaboration des politiques et des stratégies nationales relatives à l'ADM et dans le renforcement des capacités (compétences) des agences de réglementation, des gouvernements et des organismes industriels existants afin de répondre à l'utilisation accrue des systèmes ADM.

Ces centres ne devraient pas avoir de pouvoirs réglementaires, mais fournir une expertise essentielle sur la meilleure façon de protéger les droits fondamentaux et de prévenir les dommages collectifs et sociétaux. Ils devraient, par exemple, aider les petites et moyennes entreprises (PME) à remplir leurs obligations d'études d'impact en matière de droits fondamentaux, notamment en inscrivant les systèmes ADM dans le registre public mentionné précédemment.

/ Promouvoir un débat démocratique inclusif et diversifié sur les systèmes ADM

Outre le renforcement des capacités et des compétences des personnes qui déploient les systèmes ADM, il est également essentiel de faire progresser la culture algorithmique du grand public par le biais d'un débat plus large et de programmes diversifiés.

Nos conclusions suggèrent que non seulement les systèmes ADM ne sont pas transparents pour le grand public lorsqu'ils sont utilisés, mais que même la décision de déployer ou non un système ADM à la base est généralement prise sans que le public en soit informé ou n'y participe.

Il est donc urgent d'inclure l'intérêt général dès le début dans la prise de décision sur les systèmes ADM.

Plus généralement, nous avons besoin d'un débat public plus diversifié sur l'impact de l'ADM. Nous ne devons pas nous contenter de laisser la parole à des groupes d'experts, mais rendre la question plus accessible au grand public. Cela signifie qu'il nous faut parler un langage autre que le langage technojudiciaire pour mobiliser le public et susciter son intérêt.

Pour ce faire, des programmes détaillés – visant à construire et faire progresser la compétence numérique – doivent également être mis en place. Si nous souhaitons favoriser un débat public informé et créer une autonomie numérique pour les citoyen·nes européens, nous devons commencer par développer et faire progresser la culture du numérique, en mettant l'accent sur les conséquences sociales, éthiques et politiques de l'adoption de systèmes ADM.

UNION EUROPÉENNE Poser les bases de l'avenir de l'ADM en Europe

Alors que les systèmes de prise de décision automatisée (ADM) occupent une place centrale dans la garantie des droits fondamentaux et dans la distribution des services publics en Europe, les institutions de l'Union Européenne sont de plus en plus conscientes de leur rôle dans l'espace publique et privé, aussi bien en termes d'opportunités que de défis.

Par Kristina Penner et Fabio Chiusi

Depuis la publication de notre premier rapport en janvier 2019, et alors même que l'Europe est encore embourbée dans un débat plus global autour de l'intelligence artificielle « digne de confiance », plusieurs institutions, du Parlement européen au Conseil de l'Europe, ont publié des documents visant à donner à l'UE et à l'Europe une orientation en vue de traiter la question de l'ADM au cours des années, voire des décennies à venir.

À l'été 2019, Ursula von der Leyen, nouvelle présidente de la Commission et « techno-optimiste » autoproclamée, s'est engagée à proposer une « législation pour une approche européenne coordonnée sur les implications humaines et éthiques de l'intelligence artificielle » et à « réglementer l'intelligence artificielle (IA) » dans les 100 jours suivant son investiture. En février 2020, la Commission européenne a publié un « livre blanc » sur l'IA contenant « des idées et des actions » – un ensemble de stratégies visant à informer les citoyen·nes et à poser les bases d'une future action législative. Celui-ci plaide également en faveur d'une « souveraineté technologique » européenne : pour reprendre les termes de Von der Leyen, cela se traduit par « la capacité que doit avoir l'Europe de faire ses propres choix, sur la base de ses propres valeurs et en respectant ses propres règles », et devrait « contribuer à faire de nous tous des techno-optimistes ».

Une deuxième initiative fondamentale ayant trait à l'ADM en Europe est le « package législatif » sur les services numériques (DSA, *Digital Services Act*), annoncée dans l'« Agenda pour l'Europe » de Von der Leyen, et censée remplacer la directive sur le commerce électronique en vigueur depuis 2000. Ce package, qui vise à « actualiser nos règles de responsabilité et de sécurité pour les plateformes, services et produits numériques, et à instaurer un marché unique numérique », devrait conduire à des débats fondamentaux sur le rôle de l'ADM dans les politiques de modération de contenu, la responsabilité des intermédiaires et la liberté d'expression de manière générale[10].

Une attention particulière est prêtée aux systèmes ADM dans une résolution approuvée par la Commission du marché intérieur et de la protection des consommateurs du Parlement européen, ainsi que dans une recommandation « sur l'impact des systèmes algorithmiques sur les droits de l'homme » du Comité des ministres du Conseil de l'Europe (une organisation distincte de l'Union Européenne, à ne pas confondre avec le Conseil Européen).

Le Conseil de l'Europe (CdE), en particulier, s'est retrouvé à jouer un rôle de plus en plus important dans le débat

10 Des remarques et recommandations détaillées sur les systèmes ADM dans le contexte de la loi DSA peuvent être trouvées dans les conclusions du projet « Governing Platforms » d'AlgorithmWatch.

DE NOMBREUX OBSERVATEURS PERÇOIVENTUNETENSIONFONDAMENTALE ENTRE LES IMPÉRATIFS ÉCONOMIQUES ET JURIDIQUES DANS LA MANIÈRE DONT LES INSTITUTIONS EUROPÉENNES, EN PARTICULIER LA COMMISSION, FORMULENT LEURS RÉFLEXIONS ET LEURS PROPOSITIONS SUR L'IA ET L'ADM.

politique sur l'IA au cours de l'année passée, et même si son impact réel sur les initiatives réglementaires reste à démontrer, il pourrait bien s'avérer jouer le rôle de « garant » des droits de l'homme. Cette intention ressort clairement de la recommandation intitulée « Décortiquer l'intelligence artificielle : 10 étapes pour protéger les droits de l'homme », rédigée par le Commissaire aux droits de l'homme du CdE, Dunja Mijatović, et dans les travaux du Comité ad hoc sur l'IA (CAHAI) fondé en septembre 2019.

De nombreux observateurs perçoivent une tension fondamentale entre les impératifs économiques et juridiques dans la manière dont les institutions européennes, en particulier la Commission, formulent leurs réflexions et leurs propositions sur l'IA et l'ADM. D'une part, l'Europe souhaite « accroître l'utilisation et la demande de données et de produits et services basés sur les données dans l'ensemble du marché unique », afin de devenir un « leader » des applications commerciales de l'IA, et ainsi de booster la compétitivité des entreprises européennes face à la pression croissante exercée par leurs concurrents aux États-Unis et en Chine. C'est d'autant plus important pour l'ADM, l'hypothèse étant que, grâce à cette économie « data-agile », l'UE pourra « devenir un leader de premier plan pour une société capable de prendre de meilleures décisions grâce aux données – dans les entreprises comme dans le secteur public ». Comme le souligne le livre blanc sur l'IA, « les données sont la pierre angulaire du développement économique ».

D'autre part, le traitement automatisé des données sur la santé, l'emploi et les prestations sociales d'un·e citoyen·ne est susceptible de donner lieu à des décisions aux résultats injustes et discriminatoires. Ce « côté obscur » des algorithmes utilisés dans les processus décisionnels est abordé dans la boîte à outils de l'UE à travers une série de principes. Dans le cas des systèmes à haut risque, des règles doivent garantir que les processus décisionnels automatisés sont compatibles avec les droits fondamentaux et les mécanismes de contrôle démocratiques. Il s'agit d'une approche unique que les institutions européennes qualifient de « centrée sur l'humain », diamétralement opposée à celles appliquées aux États-Unis (guidée par le profit) et en Chine (guidée par la sécurité nationale et la surveillance de masse).

Cependant, des doutes sont apparus quant à la possibilité pour l'Europe d'atteindre ces deux objectifs en même temps. La reconnaissance faciale en est une excellente illustration : même si, comme le montre ce rapport, nous avons désormais des preuves abondantes de déploiements incontrôlés et opaques de cette technologie dans la plupart des pays membres, la Commission européenne ne s'est pas montrée capable d'agir rapidement et de manière décisive pour protéger les droits des citoyens européens. Comme l'ont révélé les fuites du livre blanc de la Commission européenne sur l'IA[11], l'UE était en passe d'interdire « l'identification biométrique à distance » dans les lieux publics, avant de se défiler à la dernière minute et de promouvoir un « grand débat » sur le sujet à la place.

Entre temps, des applications controversées de l'ADM pour les contrôles aux frontières, employant notamment la reconnaissance faciale, sont toujours encouragées dans des projets financés par l'UE.

Politiques et débats

/ La stratégie européenne pour les données et le Livre blanc sur l'IA

Alors que la législation complète promise « pour une approche européenne coordonnée sur les implications humaines et éthiques de l'intelligence artificielle », annoncée par Von der Leyen dans son « Agenda pour l'Europe », n'a pas été mise en œuvre au cours des « 100 premiers jours de son mandat », la Commission européenne a publié une série de documents qui fournissent un ensemble de principes et d'idées qui devraient la guider.

11 https://www.politico.eu/article/eu-considers-temporary-ban-on-facial-recognition-in-public-spaces/

Le 19 février 2020, une « Stratégie européenne pour les données » et un « livre blanc sur l'intelligence artificielle » ont été publiés conjointement, établissant les grands principes de l'approche stratégique de l'UE en matière d'IA (qui inclut les systèmes ADM, bien qu'ils n'y soient pas explicitement mentionnés). Ces principes incluent notamment la « priorité à l'humain » (« une technologie qui fonctionne au service des personnes »), la neutralité technologique (aucune technologie n'est bonne ou mauvaise en soi ; c'est son utilisation qui le détermine) et, bien sûr, la souveraineté et l'optimisme. Comme le déclare Von der Leyen : « Nous souhaitons encourager nos entreprises, nos chercheur·ses, nos innovateur·trices et nos entrepreneur·ses à développer l'intelligence artificielle, et nous voulons que nos citoyen·nes puissent l'utiliser en toute confiance. Nous devons libérer ce potentiel. »

TOUT AU LONG DU DOCUMENT, LES RISQUES ASSOCIÉS AUX TECHNOLOGIES BASÉES SUR L'IA SONT PLUS GÉNÉRALEMENT PRÉSENTÉS COMME DES RISQUES « POTENTIELS », TANDIS QUE LEURS AVANTAGES SONT DÉCRITS COMME BIEN RÉELS ET IMMÉDIATS.

L'idée sous-jacente est que les nouvelles technologies ne doivent pas amener à de nouvelles valeurs. Le « nouveau monde numérique » imaginé par l'administration Von der Leyen doit protéger pleinement les droits fondamentaux et les droits civils. L'« excellence » et la « confiance », termes mis en avant dans le titre même du livre blanc, sont considérées comme les deux piliers sur lesquels un modèle européen de l'IA peut et doit reposer, se différenciant ainsi des stratégies américaine et chinoise.

Cependant, cette ambition est absente des détails du livre blanc. Par exemple, celui-ci propose une approche de la réglementation de l'IA basée sur le risque, dans laquelle la réglementation est proportionnelle à l'impact des systèmes d'IA sur la vie des citoyen·nes. « Pour les cas à haut risque, comme dans le domaine de la santé, de la police ou du transport, peut-on lire, les systèmes d'IA doivent être transparents et traçables, et garantir une supervision humaine ». Les tests et la certification des algorithmes adoptés font également partie des garde-fous devant être mis en place, et devraient devenir aussi répandus que pour les « produits cosmétiques, les voitures et les jouets ». À l'inverse, les « systèmes moins risqués » n'auront qu'à suivre des procédures de labellisation volontaire : « Les opérateurs économiques concernés se verraient alors attribuer un label de qualité pour leurs applications d'IA. »

Mais des critiques ont fait remarquer que la définition même du « risque » dans le document est à la fois circulaire et trop vague, ce qui permet à plusieurs systèmes ADM ayant un impact important de passer à travers les mailles du filet proposé[12].

Les commentaires[13] recueillis lors de la consultation publique, entre février et juin 2020, soulignent à quel point cette idée est controversée. En effet, 42,5 % des réponses conviennent que les « exigences obligatoires » devraient se limiter aux « applications d'IA à haut risque », tandis que 30,6 % doutent d'une telle limitation.

Par ailleurs, il n'existe aucune description d'un mécanisme clair pour faire respecter ces exigences, ni de description d'un processus permettant de s'en rapprocher.

Les conséquences sont immédiatement visibles pour les technologies biométriques, en particulier la reconnaissance faciale. Sur ce point, le livre blanc propose une distinction entre l'« authentification » biométrique, qui est considérée comme ne prêtant pas à controverse (par exemple, la reconnaissance faciale pour déverrouiller un smartphone), et l'« identification » biométrique à distance (comme le déploiement dans les lieux publics pour identifier les manifestants), qui pourrait causer de sérieux problèmes en matière de droits fondamentaux et de protection de la vie privée.

Seuls les cas relevant de cette dernière catégorie seraient problématiques dans le cadre du dispositif proposé par

12 « Pour donner deux exemples : VioGén, un système ADM permettant de prévoir les violences à caractère sexiste, et Ghostwriter, une application permettant de détecter la fraude aux examens, passeraient probablement entre les mailles du filet, de la réglementation, alors qu'elles comportent des risques énormes » (https://algorithmwatch.org/en/response-european-commission-ai-consultation/)

13 « Au total, 1 215 contributions ont été reçues, dont 352 au nom d'une entreprise ou d'une organisation/association commerciale, 406 de la part de citoyens (92 % de citoyens européens), 152 d'institutions universitaires/de recherche, et 73 provenant des pouvoirs publics. Les voix de la société civile étaient représentées par 160 répondants (dont 9 organisations de consommateurs, 129 ONG et 22 syndicats). 72 personnes ont contribué dans la catégorie « autres ». Les commentaires sont parvenus « du monde entier », y compris de pays tels que l'Inde, la Chine, le Japon, la Syrie, l'Irak, le Brésil, le Mexique, le Canada, les États-Unis et le Royaume-Uni. ⁽extrait du rapport de synthèse de la consultation, accessible via le lien suivant : https://ec.europa.eu/digital-single-market/en/news/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence)

l'UE. La FAQ du livre blanc explique qu'« il s'agit de la forme de reconnaissance faciale la plus intrusive, en principe interdite dans l'UE », à moins qu'un « intérêt public substantiel » ne justifie son déploiement.

Le document explicatif affirme que « l'autorisation de la reconnaissance faciale est actuellement l'exception », mais les conclusions de ce rapport contredisent manifestement cette opinion : la reconnaissance faciale semble devenir rapidement la norme. Une version préliminaire du livre blanc qui a fuité semble reconnaître l'urgence du problème, en incluant l'idée d'un moratoire de trois à cinq ans sur le recours à la reconnaissance faciale dans les lieux publics, jusqu'à ce qu'un moyen de les concilier avec les contrôles démocratiques puisse être trouvé, si tant est que cela soit possible.

Juste avant la publication officielle du livre blanc, même la commissaire européenne Margrethe Vestager a préconisé une « pause » de ces applications.

Cependant, immédiatement après l'appel de Mme Vestager, des responsables de la Commission ont ajouté que cette « pause » ne saurait empêcher les gouvernements nationaux d'utiliser la reconnaissance faciale conformément aux règles en vigueur. Pour finir, la version finale du livre blanc a évacué toute mention d'un moratoire et a appelé à un « grand débat européen sur les circonstances spécifiques, le cas échéant, qui pourraient justifier » son utilisation à des fins d'identification biométrique en direct. Parmi celles-ci, le livre blanc évoque notamment la justification, la proportionnalité, l'existence de mécanismes de protection démocratiques et le respect des droits fondamentaux.

Tout au long du document, les risques associés aux technologies basées sur l'IA sont plus généralement présentés comme des risques « potentiels », tandis que leurs avantages sont décrits comme bien réels et immédiats. Cela a conduit de nombreuses organisations de défense des droits fondamentaux[14] à déclarer que la teneur générale du livre blanc laisse entrevoir un revirement inquiétant des priorités de l'UE, faisant passer la compétitivité mondiale avant la protection des droits fondamentaux.

Certaines questions fondamentales sont toutefois abordées dans les documents : par exemple, l'interopérabilité de ces solutions et la création d'un réseau de centres de recherche axés sur les applications de l'IA visant l'« excellence » et le développement des compétences.

L'objectif est « d'attirer plus de 20 milliards d'euros d'investissements dans l'UE par an dans le domaine de l'IA au cours de la prochaine décennie ».

Un certain déterminisme technologique semble également affecter le livre blanc. « Il est essentiel, peut-on y lire, que l'administration publique, les hôpitaux, les services publics et de transport, les autorités de surveillance financière et d'autres secteurs d'intérêt public commencent rapidement à déployer des produits et des services reposant sur l'IA dans leurs activités. Un accent particulier sera mis sur les domaines des soins de santé et du transport, où la technologie est mûre pour un déploiement à grande échelle. »

Toutefois, il reste à voir si cette suggestion d'un déploiement hâtif des solutions ADM dans toutes les sphères de l'activité humaine est compatible avec les efforts de la Commission européenne visant à relever les défis structurels que posent les systèmes ADM en matière de droit et de justice.

/ Résolution du Parlement européen sur l'ADM et la protection des consommateurs

Une résolution, adoptée par le Parlement européen en février 2020, traite plus spécifiquement des systèmes ADM dans le contexte de la protection des consommateurs. Cette résolution fait remarquer à juste titre que « des avancées technologiques rapides ont lieu dans les domaines de l'intelligence artificielle, l'apprentissage automatique, les systèmes complexes fondés sur des algorithmes et les processus de prise de décision automatisés », et que « les applications, possibilités et défis découlant de ces technologies sont nombreux et concernent pratiquement tous les secteurs du marché intérieur ». Le texte souligne également la nécessité d'un « examen du cadre juridique actuel de l'UE » afin de « vérifier qu'il est à même de faire face à l'émergence de l'IA et de la prise de décision automatisée ».

Appelant à une « approche commune de l'Union européenne en matière de développement des processus de prise de décision automatisés », la résolution détaille plusieurs conditions que tout système de ce type devrait posséder

14 Parmi ceux-ci : Access Now (https://www.accessnow.org/cms/assets/uploads/2020/05/EU-white-paper-consultation_AccessNow_May2020.pdf), AI Now (https://ainowinstitute.org/ai-now-comments-to-eu-whitepaper-on-ai.pdf), EDRI (https://edri.org/can-the-eu-make-ai-trustworthy-no-but-they-can-make-it-just/) — et AlgorithmWatch (https://algorithmwatch.org/en/response-european-commission-ai-consultation/).

SI L'« IA » EST EFFECTIVEMENT UNE RÉVOLUTION NÉCESSITANT UN PAQUET DE MESURES LÉGISLATIVES DÉDIÉE, LES ÉLUS VEULENT AVOIR LEUR MOT À DIRE.

pour rester compatible avec les valeurs européennes. Les consommateur·trices devraient être « dûment informé[·e]s » sur l'impact qu'ont les algorithmes sur leur vie, et ils et elles devraient avoir accès à une personne ayant un pouvoir décisionnel afin que ces décisions puissent être vérifiées et corrigées si nécessaire. Ils et elles devraient également être informé·es « lorsque les prix des biens ou services ont été personnalisés sur la base d'une prise de décision automatisée et du profilage du comportement ».

En rappelant à la Commission européenne la nécessité d'une approche fondée sur les risques et soigneusement élaborée, la résolution précise que les mécanismes de protection doivent tenir compte du fait que les systèmes ADM « peuvent évoluer et agir d'une manière qui n'était pas envisagée lors de leur mise sur le marché », et que la responsabilité n'est pas toujours facile à attribuer lorsque le déploiement d'un système ADM entraîne un préjudice.

La résolution fait écho à l'article 22 du RGPD en notant qu'un sujet humain doit toujours être impliqué lorsque « des intérêts publics légitimes sont en jeu », et être responsable en dernier ressort des décisions prises dans le « domaine médical, juridique et comptable, ainsi que dans le secteur bancaire ». Une évaluation « correcte » des risques doit notamment précéder toute automatisation des services professionnels.

Enfin, la résolution liste des exigences détaillées en matière de qualité et de transparence dans la gouvernance des données : parmi celles-ci, « 'importance d'utiliser uniquement des données non faussées et de qualité pour améliorer les résultats des systèmes algorithmiques et renforcer la confiance et l'acceptation des consommateur[·trice]s » ; l'utilisation d'algorithmes « explicables et impartiaux » ; ainsi que la nécessité de « structures de réexamen » permettant aux personnes affectés de « réclamer un réexamen et une correction, par un être humain, des décisions automatisées qui sont définitives et permanentes ».

/ Tirer le maximum du « droit d'initiative » du Parlement européen

Dans son discours d'investiture, von der Leyen a clairement exprimé son soutien à un « droit d'initiative » pour le Parlement européen. « Lorsque cette Assemblée, statuant à la majorité de ses membres, adopte des résolutions demandant à la Commission de soumettre des propositions législatives, a-t-elle déclaré, je m'engage à répondre par un acte législatif dans le plein respect des principes de proportionnalité, de subsidiarité ainsi que de l'accord 'Mieux légiférer'. »

Si l'« IA » est effectivement une révolution nécessitant un paquet de mesures législatives dédiées, censé arriver au cours du premier trimestre 2021, les élus veulent avoir leur mot à dire. Cette volonté, associée à l'intention déclarée de von der Leyen de renforcer leurs capacités législatives, pourrait même déboucher sur ce que Politico a appelé un « moment parlementaire », des commissions parlementaires commençant à rédiger plusieurs rapports différents en conséquence.

Chaque rapport étudie des aspects spécifiques de l'automatisation dans la politique publique qui, bien qu'ils soient

destinés à façonner la législation à venir sur l'IA, sont aussi pertinents pour l'ADM.

Par exemple, dans son *cadre d'aspects éthiques en matière d'intelligence artificielle, de robotique et de technologies connexes*, la Commission des affaires juridiques appelle à la constitution d'une « Agence européenne de l'intelligence artificielle » et, dans le même temps, d'un réseau d'autorités nationales de surveillance dans chaque État membre pour s'assurer que les décisions prises en matière d'automatisation soient et demeurent éthiques.

Dans son rapport intitulé *Dr oits de propriété intellectuelle pour le développement des technologies liées à l'intelligence artificielle*, cette même commission expose sa vision pour l'avenir de la propriété intellectuelle et de l'automatisation. D'une part, le rapport préliminaire indique que « les méthodes mathématiques et les programmes d'ordinateurs ne sont pas brevetables en tant que tels, ils peuvent être intégrés à une invention technique susceptible d'être brevetable », tout en affirmant que, s'agissant de la transparence algorithmique, « la rétro-ingénierie constitue une exception au secret d'affaires ».

Le rapport va même jusqu'à envisager comment protéger « les créations techniques et artistiques générées par l'IA afin d'encourager cette forme de création », imaginant que « certaines œuvres générées par l'IA sont assimilables à des œuvres de l'esprit et que, dès lors, elles pourraient être protégées par le droit d'auteur ».

Enfin, dans un troisième document (*Intelligence artificielle et responsabilité civile*), la Commission détaille une « approche de gestion des risques » pour la responsabilité civile des technologies de l'IA. Selon ce document, « la partie qui est la mieux à même de contrôler et de gérer un risque lié à une technologie est tenue pour strictement responsable, en tant que point d'entrée unique pour les litiges ».

Des principes importants concernant l'utilisation de l'ADM dans le système de justice pénale se trouvent dans le rapport de la Commission des libertés civiles, de la justice et des affaires intérieures intitulé *L'intelligence artificielle en droit pénal et son utilisation par les autorités policières et judiciaires dans les affaires pénales*. Après avoir dressé une liste détaillée d'utilisations réelles et actuelles de l'« IA » – qui sont en réalité des systèmes ADM – par les forces de police[15], la Commission « estime qu'il est nécessaire de créer un régime clair et équitable pour l'attribution de la responsabilité juridique des conséquences négatives potentielles de ces technologies numériques avancées ».

Le rapport détaille ensuite certaines de ses dispositions : pas de décisions entièrement automatisées[16], une explication des algorithmes qui soit « intelligible pour les utilisateur[·trice]s », une « étude d'impact obligatoire sur les droits fondamentaux [...] de tout système d'IA à des fins répressives ou judiciaires » avant son déploiement ou son adoption, plus « un audit périodique obligatoire de tous les systèmes d'IA utilisés par les services répressifs et judiciaires pour tester et évaluer les systèmes algorithmiques une fois qu'ils sont en service ».

Le rapport préconise également un moratoire sur les technologies de reconnaissance faciale destinées aux forces de l'ordre, « jusqu'à ce que les normes techniques puissent être considérées comme pleinement respectueuses des droits fondamentaux, que les résultats obtenus soient non discriminatoires et que la confiance du public soit assurée quant à la nécessité et à la proportionnalité du déploiement de ces technologies ».

L'objectif est de renforcer à terme la transparence globale de ces systèmes, et de conseiller aux États membres d'avoir une « compréhension complète » des systèmes d'IA adoptés par les services répressifs et judiciaires, et – sur le modèle d'un « registre public » – de détailler « les types d'outils utilisés, les types de criminalité auxquels ils s'appliquent et les entreprises dont les outils sont utilisés ».

La Commission de la culture et de l'éducation et la Commission de la politique industrielle travaillent également sur leurs propres rapports à l'heure où nous écrivons ces lignes.

---

15 À la page 5, le rapport indique : « Les applications de l'IA utilisées par les services répressifs comprennent des applications telles que les technologies de reconnaissance faciale, la reconnaissance automatisée des plaques d'immatriculation, l'identification du locuteur, l'identification vocale, les technologies de lecture labiale, la surveillance auditive (par exemple, des algorithmes de détection des coups de feu), la recherche et l'analyse autonomes de bases de données identifiées, la prévision (police prédictive et analyse des zones sensibles), les outils de détection des comportements, les outils autonomes d'identification de la fraude financière et du financement du terrorisme, la surveillance des médias sociaux (scrapping et collecte de données pour les connexions de minage), les dispositifs de capture d'identité d'abonnés mobiles internationaux (IMSI) et les systèmes de surveillance automatisés intégrant différentes capacités de détection (comme la détection des battements de cœur et les caméras thermiques). »

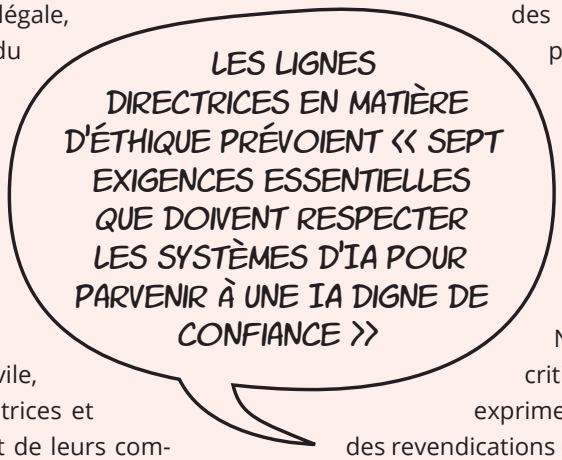
16 « Dans les contextes judiciaires et répressifs, la décision finale doit toujours être prise par un être humain. » (p. 6)

Toutes ces initiatives ont abouti à la création d'une Commission spéciale sur l'« intelligence artificielle à l'ère du numérique » (AIDA), le 18 juin 2020. Composée de 33 membres et créée pour une durée initiale de 12 mois, celle-ci « analysera la future incidence » de l'IA sur l'économie de l'UE, et « en particulier sur les compétences, l'emploi, les technologies de la finance, l'éducation, la santé, les transports, le tourisme, l'agriculture, l'environnement, la défense, l'industrie, l'énergie et l'administration en ligne ».

/ Groupe d'experts de haut niveau sur l'IA et AI Alliance

En 2018, le Groupe d'experts de haut niveau (GEHN) sur l'IA, un comité composé de 52 experts, a été constitué par la Commission européenne afin de soutenir la mise en œuvre de la stratégie européenne sur l'IA, d'identifier les principes à respecter pour parvenir à une « IA digne de confiance », et, en tant que comité directeur de l'AI Alliance, afin de créer une plateforme multipartite ouverte (qui comptait plus de 4 000 membres au moment de la rédaction de ce rapport) qui permettra d'élargir la contribution aux travaux du GEHN.

Après la publication de la première ébauche des *Lignes directrices en matière d'éthique pour une IA digne de confiance* en décembre 2018, document auquel ont répondu plus de 500 contributeurs, une version révisée a été publiée en avril 2019. Celle-ci propose une « approche centrée sur l'humain » pour parvenir à une IA légale, éthique et robuste tout au long du cycle de vie du système. Toutefois, il ne s'agit encore que d'un cadre volontaire sans recommandations concrètes et applicables en matière d'opérationnalisation, de mise en œuvre et d'application.



Les organisations de la société civile, de protection des consommateur·trices et de défense des droits ont fait part de leurs commentaires et ont appelé à la transposition des lignes directrices en droits concrets[17]. Par exemple, l'association à but non lucratif de défense des droits numériques Access Now, membre du GEHN, a demandé instamment à la Commission européenne de clarifier les modalités selon lesquelles les différentes parties prenantes pourront tester, appliquer, améliorer, approuver et faire respecter cette « IA digne de confiance », tout en reconnaissant la nécessité de déterminer les lignes rouges de l'Europe.

Dans un éditorial, deux autres membres du GEHN ont affirmé que le groupe avait « travaillé pendant un an et demi, tout cela pour que ses propositions détaillées soient essentiellement ignorées ou simplement mentionnées à titre indicatif » par la Commission européenne, qui a rédigé la version finale[18]. Ils ont également fait valoir que, puisque le groupe était initialement chargé d'identifier les risques et les « lignes rouges » de l'IA, les membres du groupe ont souligné que les systèmes d'armement autonomes, de notation des citoyens et d'identification automatisée des individus par la reconnaissance faciale étaient des applications de l'IA à éviter. Cependant, les représentants de l'industrie, qui dominent le comité[19], sont parvenus à faire supprimer ces principes avant que le projet ne soit publié.

Ce déséquilibre qui vise à mettre en évidence le potentiel de l'ADM plutôt que les risques qu'elle comporte s'observe également dans le deuxième rapport du groupe. Dans son document *Recommandations de politique et d'investissement pour une IA digne de confiance en Europe*, publié en juin 2019, le GEHN propose 33 recommandations visant à « guider l'IA digne de confiance vers la durabilité, la croissance et la compétitivité, ainsi que vers l'inclusion – tout en renforçant l'autonomie, les avantages et la protection des êtres humains ». Ce document est principalement un appel à encourager l'adoption et l'expansion de l'IA dans les secteurs privé et public en investissant dans des outils et des applications destinés à « aider les populations vulnérables » et à ne « laisser personne de côté ».

Néanmoins, et malgré toutes les critiques légitimes, ces deux directives expriment encore des préoccupations et des revendications critiques concernant les systèmes de prise de décision automatisée. Par exemple, les *Lignes directrices en matière d'éthique* prévoient « sept exigences

17 Par exemple, le Bureau européen des unions de consommateurs (BEUC) : insérer lien vers la référence

18 Mark Coeckelbergh et Thomas Metzinger : Europe needs more guts when it comes to AI ethics, https://background.tagesspiegel.de/digitalisierung/europe-needs-more-guts-when-it-comes-to-ai-ethics

19 Le groupe était composé de 24 représentants d'entreprises, 17 universitaires, 5 organisations de la société civile et 6 autres membres, dont l'Agence des droits fondamentaux de l'Union européenne.

essentielles que doivent respecter les systèmes d'IA pour parvenir à une IA digne de confiance ». Le document offre ensuite des conseils pour la mise en œuvre pratique de chacune de ces exigences : supervision et contrôle humains, robustesse et sécurité techniques, confidentialité et gouvernance des données, transparence, diversité, absence de discrimination et équité, bien-être sociétal et environnemental, et responsabilité.

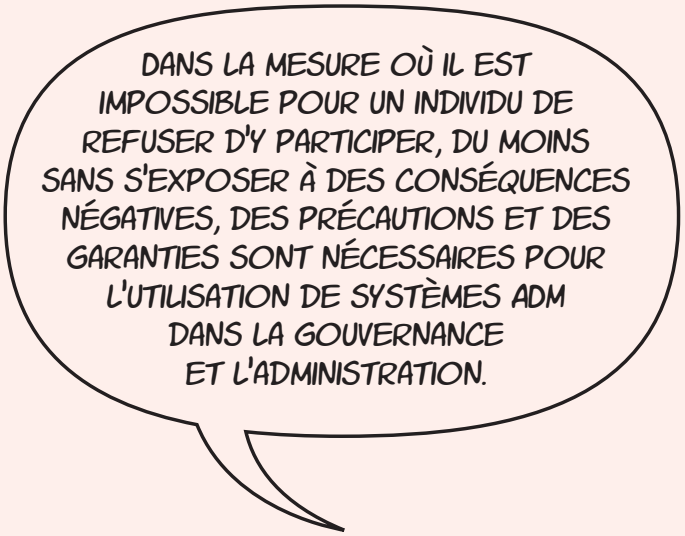
Ces directives prévoient également un projet pilote concret, appelé « liste d'évaluation pour une IA digne de confiance », qui vise à rendre ces principes généraux opérationnels. Le but est de faire adopter celle-ci « lors du développement, du déploiement ou de l'utilisation de systèmes d'IA » et de l'adapter « aux cas d'utilisation spécifiques dans lesquels le système est appliqué ».

La liste aborde de nombreuses questions liées au risque d'atteinte aux droits fondamentaux par les systèmes ADM. Celles-ci incluent l'absence de supervision et de contrôle humains, les problèmes de robustesse et de sécurité techniques, l'incapacité à éviter les traitements injustes ou à fournir un accès égal et universel à ces systèmes, et l'absence d'accès significatif aux données qui y sont introduites.

Contextuellement, la liste incluse dans les lignes directrices fournit des questions utiles pour aider ceux et celles qui déploient des systèmes ADM. Par exemple, elle préconise « une évaluation de l'incidence potentielle sur les droits fondamentaux » là où il pourrait y avoir un impact négatif sur ceux-ci. Elle demande également si « des mécanismes spécifiques de contrôle et de surveillance » ont été mis en place dans le cas de « systèmes d'autoapprentissage ou d'IA autonome », et si des processus existent « pour assurer la qualité et l'intégrité de vos données ».

Les remarques détaillées portent également sur des questions fondamentales pour les systèmes ADM, telles que leur transparence et leur explicabilité. Ces questions incluent « dans quelle mesure les décisions et donc les conséquences du système d'IA peuvent-elles être comprises ? » et « dans quelle mesure la décision du système influence-t-il les processus décisionnels de l'organisation ? ». Ces questions sont particulièrement pertinentes pour évaluer les risques posés par le déploiement de tels systèmes.

Afin d'éviter les partis pris et les effets discriminatoires, les lignes directrices préconisent « des processus de supervision permettant d'analyser et de traiter la finalité, les contraintes, les exigences et les décisions du système de



manière claire et transparente », tout en exigeant la participation des parties prenantes tout au long du processus de mise en œuvre des systèmes d'IA.

De plus, les recommandations en matière de politique et d'investissement prévoient de déterminer les lignes rouges par le biais d'un « dialogue institutionnalisé sur la politique d'IA avec les acteurs concernés », y compris des experts de la société civile. En outre, elles appellent à interdire la notation à grande échelle des individus au moyen de l'IA comme défini dans les lignes directrices en matière d'éthique, et à fixer des règles très claires et strictes pour la surveillance à des fins de sécurité nationale et d'autres fins prétendument d'intérêt public ou national. Cette interdiction inclurait les technologies d'identification biométrique et le profilage.

En ce qui concerne les systèmes de prise de décision automatisée, le document indique également que « pour parvenir à mettre en œuvre une IA digne de confiance, il sera essentiel de définir clairement si, quand et comment l'IA peut être utilisée aux fins de l'identification automatique de personnes », avertissant que « toute forme de notation des citoyen[·ne]s peut entraîner la perte de [leur] autonomie et mettre en péril le principe de non-discrimination», et « ne doit être utilisée que si elle se justifie clairement et lorsque les mesures sont proportionnées et équitables ». Il souligne en outre que « la transparence ne peut ni empêcher la discrimination, ni garantir l'équité ». Cela signifie qu'il doit être possible pour un·e citoyen·ne de ne pas participer à un mécanisme de notation, idéalement sans que cela ne lui porte préjudice.

D'une part, le document reconnaît que « certaines applications d'IA sont certes susceptibles d'apporter des avantages considérables aux individus et à la société, mais qu'elles peuvent également avoir des incidences négatives, y compris des incidences pouvant s'avérer difficiles à anti-

ciper, reconnaître ou mesurer (par exemple, en matière de démocratie, d'état de droit et de justice distributive, ou sur l'esprit humain lui-même) ». D'autre part, le groupe affirme cependant qu'il faut éviter toute réglementation inutilement prescriptive.

En juillet 2020, le GEHN sur l'IA a également présenté la version finale de sa liste d'évaluation pour une intelligence artificielle digne de confiance (ALTAI), compilée après un processus pilote ayant impliqué 350 partenaires.

La liste, qui est entièrement volontaire et sans aucune portée réglementaire, vise à traduire en actions les sept exigences énoncées dans les lignes directrices en matière d'éthique du GEHN sur l'IA. L'intention est de fournir à tous ceux qui souhaitent mettre en œuvre des solutions d'IA compatibles avec les valeurs de l'UE – par exemple, les concepteurs et développeurs de systèmes d'IA, les statisticien·nes, les responsables ou spécialistes des marchés publics et les responsables juridiques/de la conformité – une boîte à outils d'autoévaluation.

/ Conseil de l'Europe : comment protéger les droits fondamentaux dans le cadre de systèmes ADM

Parallèlement au Comité ad hoc sur l'intelligence artificielle (CAHAI), établi en septembre 2019, le Comité des ministres du Conseil de l'Europe[20] a publié un document substantiel et probant.

Envisagée comme un instrument normatif, sa *Recommandation aux États membres sur l'impact des systèmes algorithmiques sur les droits de l'homme* décrit[21] les « défis importants » qui se posent avec l'émergence et notre « recours croissant » à l'égard de ces systèmes, et qui sont pertinents « pour les sociétés démocratiques et l'État de droit ».

Ce texte, qui a fait l'objet d'une période de consultation publique avec des commentaires détaillés d'organisations de la société civile, va au-delà du livre blanc de la Commission européenne en ce qui concerne la protection des valeurs de l'UE et des droits fondamentaux.

La recommandation analyse de manière approfondie les effets et les configurations changeantes des systèmes algorithmiques (annexe A) en se concentrant sur toutes les étapes du processus qui entrent dans la fabrication d'un algorithme, c'est-à-dire l'acquisition, la conception, le développement et le déploiement continu.

Bien qu'elle suive généralement l'approche de l'« IA centrée sur l'humain » des lignes directrices du GEHN, la recommandation définit les « obligations des États » (annexe B) ainsi que les responsabilités des acteurs du secteur privé (annexe C). Par ailleurs, la recommandation ajoute des principes tels que l'« autodétermination informationnelle »[22], énumère des suggestions détaillées pour des mécanismes de responsabilisation et des recours efficaces, et exige des évaluations d'impact sur les droits fondamentaux.

Bien que le document reconnaisse clairement le « potentiel important des technologies numériques pour faire face aux défis sociétaux et pour favoriser une innovation et un développement économique socialement bénéfiques », il invite également à la prudence. Ceci afin de garantir que ces systèmes ne perpétuent pas délibérément ou accidentellement « les inégalités raciales, de genre et autres disparités au sein de la société et au sein de la population active, qui n'ont pas encore été éliminées de nos sociétés ».

Au contraire, les systèmes algorithmiques devraient être utilisés de manière proactive et sensible pour résoudre ces déséquilibres, et « accorder une attention particulière aux besoins et aux voix des groupes vulnérables ».

Mais surtout, la recommandation identifie le risque potentiellement plus élevé pour les droits fondamentaux lié à l'utilisation de systèmes algorithmiques par les États membres pour la prestation de services et de politiques

20 Le Conseil de l'Europe est à la fois « un organe gouvernemental où les approches nationales des problèmes européens sont discutées sur un pied d'égalité et un forum permettant de trouver des réponses collectives à ces défis ». Son travail inclut « les aspects politiques de l'intégration européenne, la sauvegarde des institutions démocratiques et de l'État de droit et la protection des droits de l'homme – en d'autres termes, tous les problèmes qui nécessitent des solutions paneuropéennes concertées ». Bien que les recommandations adressées aux gouvernements des États membres ne soient pas contraignantes, le Comité peut, dans certains cas, demander aux gouvernements de l'informer des suites données par eux à ces recommandations (Art. 15b des statuts). Les relations entre le Conseil de l'Europe et l'Union européenne sont définies dans (1) le *Recueil des textes régissant les relations entre le Conseil de l'Europe et l'Union européenne*, et (2) le *Mémorandum d'accord entre le Conseil de l'Europe et l'Union européenne*.

21 Sous la supervision du Comité directeur sur les médias et la société de l'information (CDMSI) et préparé par le Comité d'experts sur la dimension droits de l'Homme du traitement automatisé des données et de différentes formes d'intelligence artificielle (MSI-AUT).

22 « Les États doivent veiller à ce que la conception, le développement et le déploiement continu de systèmes algorithmiques permettent aux personnes d'être informées à l'avance du traitement des données (y compris de ses objectifs et de ses résultats possibles) et de contrôler leurs données, notamment grâce à l'interopérabilité », peut-on lire à la section 2.1 de l'annexe B.

publiques. Dans la mesure où il est impossible pour un individu de refuser d'y participer, du moins sans s'exposer à des conséquences négatives, des précautions et des garanties sont nécessaires pour l'utilisation de systèmes ADM dans la gouvernance et l'administration.

La recommandation aborde également les conflits et les difficultés potentielles découlant des partenariats public-privé dans un large éventail d'utilisations.

La recommandation exhorte ainsi les gouvernements des États membres à abandonner les processus et refuser d'utiliser un système ADM si « son opacité empêche toute supervision ou tout contrôle par un être humain » ou si les droits fondamentaux sont menacés ; et à déployer des systèmes ADM si et seulement si la transparence, la responsabilité, la légalité et la protection des droits humains peuvent être garantis « tout au long du déploiement ». Par ailleurs, le suivi et l'évaluation de ces systèmes doivent être « constants », « inclusifs et transparents », et comporter un dialogue avec toutes les parties prenantes concernées, ainsi qu'une analyse de l'impact environnemental et des autres externalités potentielles affectant « les populations et leurs environnements ».

À l'annexe A, le Conseil de l'Europe donne également une définition des algorithmes à « haut risque », dont les autres organismes pourront s'inspirer. Plus précisément, la recommandation explique que « l'expression « à haut risque » est employée en référence à l'utilisation de systèmes algorithmiques dans des processus ou des décisions susceptibles d'avoir des conséquences graves pour les individus, ou dans des situations où l'absence de solutions de rechange engendre une probabilité particulièrement élevée d'atteinte aux droits de l'homme, notamment en introduisant ou en amplifiant des inégalités distributives. ».

Le document, qui n'a pas nécessité l'unanimité des membres pour être adopté, est non contraignant.

/ Réglementation du contenu terroriste en ligne

Après une longue période de progrès laborieux, un règlement visant à empêcher la propagation de contenus terroristes en ligne a gagné du terrain en 2020. Si le règlement adopté devait inclure des outils automatisés et proactifs pour la reconnaissance et le retrait de contenus en ligne, ceux-ci relèveraient probablement de l'article 22 du RGPD.

Comme le souligne le Contrôleur européen de la protection des données (CEPD) : « étant donné que les outils automatisés, tels qu'envisagés par la proposition, pourraient non seulement conduire au retrait et à la conservation de contenus (et de données connexes) concernant la personne les ayant téléchargés, mais aussi, en fin de compte, à des poursuites pénales à son encontre, ces outils affecteraient de manière significative cette personne, auraient une incidence sur son droit à la liberté d'expression et présenteraient des risques importants pour ses droits et libertés », et, par conséquent, relèveraient de l'article 22(2).

En outre, et surtout, ce règlement nécessiterait des garanties plus substantielles que celles que la Commission prévoit actuellement. Comme l'explique le groupe de défense des droits numériques European Digital Rights (EDRi) : « Le règlement sur les contenus terroristes proposé doit être réformé en profondeur pour être à la hauteur des valeurs de l'Union et pour protéger les droits et libertés fondamentaux de ses citoyen·nes. »

Une première vague de critiques virulentes de la proposition initiale, émanant de groupes de la société civile et de commissions du Parlement européen (PE), y compris des avis et des analyses de l'Agence des droits fondamentaux de l'Union européenne (FRA), de l'EDRi, ainsi qu'un rapport critique conjoint de trois rapporteurs spéciaux des Nations unies, ont mis en évidence les menaces pesant sur le droit à la liberté d'expression et d'information, le droit à la liberté et au pluralisme des médias, la liberté d'entreprise et les droits à la vie privée et à la protection des données personnelles.

Les aspects critiqués comprennent notamment une définition trop vague du contenu terroriste, le champ d'application du règlement (qui couvre actuellement les contenus à caractère éducatif et journalistique), l'appel susmentionné à des « mesures proactives », le manque de supervision judiciaire efficace, l'insuffisance des obligations de signalement pour les services répressifs, et l'absence de garanties pour « les cas où il y a des motifs raisonnables de penser que les droits fondamentaux sont affectés » (EDRi 2019).

Le CEPD souligne que ces « garanties appropriées » devraient inclure le droit de bénéficier d'une intervention humaine et le droit d'obtenir une explication de la décision prise par des moyens automatisés (EDRi 2019).

Bien que les garanties suggérées ou exigées figurent maintenant dans le rapport préliminaire du Parlement euro-

péen sur la proposition, il reste à voir qui pourra retenir son souffle le plus longtemps avant le vote final. Lors des trilogues à huis clos entre le PE, la nouvelle CE et le Conseil de l'UE (qui a pris ses fonctions en octobre 2019), seules des modifications mineures sont encore possibles, selon un document qui a fuité.

Contrôle et réglementation

/ Premières décisions sur la conformité des systèmes ADM avec le RGPD

« Bien qu'il n'y ait pas eu de grand débat sur la reconnaissance faciale lors des négociations relatives au RGPD et à la directive "police-justice", la législation a été conçue de manière à pouvoir s'adapter au fil de temps en fonction de l'évaluation des technologies. [...] L'heure est venue pour l'UE, alors qu'elle débat du caractère éthique de l'IA et de la nécessité d'une réglementation, de déterminer si, le cas échéant, la technologie de reconnaissance faciale peut être autorisée dans une société démocratique. Si la réponse est "oui", alors seulement pourrons-nous nous pencher sur la question de savoir comment, et quelles garanties et responsabilités mettre en place. » – Wojciech Wiewiórowski, CEPD.

«Les dispositifs de reconnaissance faciale sont particulièrement intrusifs et présentent des risques majeurs d'atteinte à la vie privée et aux libertés individuelles des personnes concernées. » – (CNIL 2019)

Depuis la publication du dernier rapport *L'automatisation de la société* d'AlgorithmWatch, nous avons vu les premiers cas d'amendes et de décisions liées à des violations de la réglementation prononcées par les autorités nationales de protection des données (APD) en se basant sur le RGPD. Les études de cas suivantes illustrent toutefois les limites du RGPD dans la pratique en ce qui concerne l'article 22 relatif aux systèmes ADM, et montrent qu'il laisse le soin aux autorités de protection de la vie privée d'évaluer les affaires au cas par cas.

En Suède, un projet pilote de reconnaissance faciale, réalisé dans une classe pendant une période limitée, a été jugé contraire à plusieurs dispositions du règlement sur la protection des données (en particulier les articles 2(14) et 9(2) du RGPD) (Comité européen de la protection des données 2019).

Un cas similaire en France a été suspendu par la Commission nationale de l'informatique et des libertés (CNIL), qui a fait part de son inquiétude lorsque deux lycées ont envisagé d'introduire la technologie de reconnaissance faciale en partenariat avec la société américaine Cisco. L'avis de la CNIL est non contraignant et la procédure suit son cours[23].

L'autorisation préalable des autorités de protection des données n'est pas nécessaire pour réaliser de tels essais, le consentement des utilisateurs étant généralement considéré comme suffisant pour traiter leurs données biométriques. Et pourtant, ce n'était pas le cas en Suède, en raison d'un déséquilibre des pouvoirs entre le gestionnaire des données et les personnes concernées. Au lieu de cela, une évaluation d'impact adéquate et une consultation préalable avec l'APD ont été jugées nécessaires.

Le Contrôleur européen de la protection des données (CEPD) l'a confirmé :

« Le consentement doit être explicite ainsi que libre, informé et spécifique. Pourtant, il est clair qu'une personne ne peut pas refuser ni même accepter de se soumettre lorsqu'elle a besoin d'accéder à des espaces publics couverts par une surveillance par reconnaissance faciale. [...] Enfin, la conformité de la technologie avec des principes tels que la minimisatio n des données et l'obligation de protection des données dès la conception est plus que douteuse. La technologie de reconnaissance faciale n'a jamais été totalement précise, ce qui peut avoir de graves conséquences pour les personnes faussement identifiées, qu'il s'agisse de criminels ou autre. [...] Il serait toutefois malavisé de se focaliser uniquement sur les questions de protection de la vie privée. Il s'agit fondamentalement d'une question éthique dans une société démocratique. » (CEPD 2019)

Access Now a commenté :

« Alors que de plus en plus de projets de reconnaissance faciale se développent, nous constatons déjà que le RGPD offre des garanties utiles en matière de droits fondamentaux qui peuvent être opposées à la collecte et à l'utilisation illégales de données sensibles telles que les données biométriques. Mais le battage irresponsable et souvent infon-

23 Voir le chapitre sur la France (Kayalki 2019)

dé autour de l'efficacité de ces technologies et de l'intérêt économique sous-jacent risque de conduire les gouvernements centraux et locaux ainsi que les entreprises privées à tenter de circonvenir la loi. »-

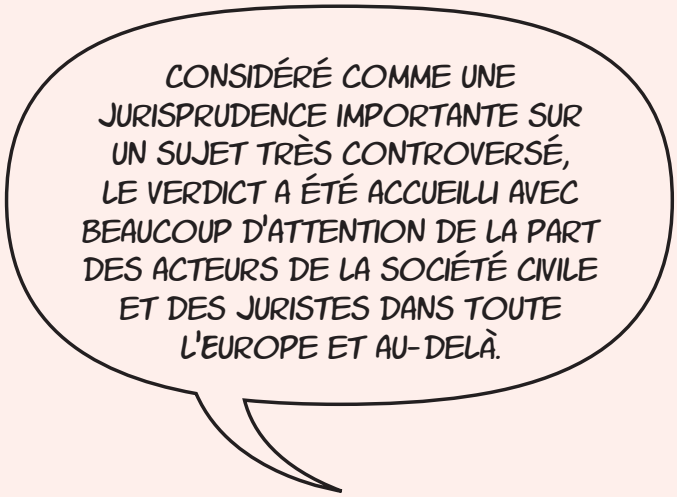
/ La reconnaissance faciale automatique utilisée par la police du sud du Pays de Galles jugée illégale

Au cours de l'année 2020, le Royaume-Uni a été témoin de la première application très médiatisée de la Law Enforcement Directive[24] concernant l'utilisation des technologies de reconnaissance faciale par la police dans les espaces publics. Considéré comme une jurisprudence importante sur un sujet très controversé, le verdict a été accueilli avec beaucoup d'attention de la part des acteurs de la société civile et des juristes dans toute l'Europe et au-delà[25].

L'affaire a été portée devant les tribunaux par Ed Bridges, un homme de 37 ans de Cardiff, qui a déclaré que son visage avait été scanné sans son consentement alors qu'il faisait ses courses de Noël en 2017, ainsi que lors d'une manifestation pacifique contre la prolifération des armes un an plus tard.

Le tribunal a d'abord affirmé la légalité de l'utilisation de la technologie de reconnaissance faciale automatisée (« RFA ») par la police du sud du Pays de Galles, la déclarant légale et proportionnée. Mais la décision a été contestée en appel par Liberty, un groupe de défense des droits civiques, et la Cour d'appel d'Angleterre et du Pays de Galles a décidé de casser le verdict du tribunal de première instance et de déclarer la technologie illégale le 11 août 2020.

En statuant contre la police du sud du Pays de Galles sur trois des cinq motifs, la Cour d'appel a constaté qu'il existait des « lacunes fondamentales » dans le cadre normatif existant autour de l'utilisation de la RFA, que son déploiement ne satisfaisait pas au principe de « proportionnalité » et, par ailleurs, qu'une évaluation adéquate de l'impact sur la protection des données (DPIA) n'avait pas été effectuée, omettant de multiples étapes cruciales.



Le tribunal n'a toutefois pas jugé que le système produisait des résultats discriminatoires sur la base du sexe ou de la race, car la police du sud du Pays de Galles n'avait pas recueilli suffisamment de données pour émettre un jugement à ce sujet[26]. Cependant, le tribunal a jugé nécessaire d'ajouter la remarque suivante : « La RFA étant une technologie nouvelle et controversée, nous espérons que toutes les forces de police qui ont l'intention de l'utiliser à l'avenir voudront bien s'assurer que tout ce qui peut raisonnablement être fait a été fait pour que le logiciel utilisé soit dépourvu de tout préjugé racial ou sexiste. »

Suite à l'arrêt de la cour, Liberty a exhorté la police du sud du Pays de Galles et d'autres forces de police de renoncer à l'utilisation des technologies de reconnaissance faciale.

L'ADM en pratique : gestion et surveillance des frontières

Alors que la Commission européenne et ses partenaires débattaient de la nécessité de réglementer ou d'interdire les technologies de reconnaissance faciale, des essais approfondis de ces systèmes étaient déjà en cours dans toute l'Europe.

24 La Law Enforcement Directive, en vigueur depuis mai 2018, « porte sur le traitement des données à caractère personnel par les gestionnaires de données à des «fins répressives» - qui ne relève pas du champ d'application du RGPD » https://www.dataprotection.ie/en/organisations/law-enforcement-directive

25 Décision rendue le 4 septembre 2019 par la Haute Cour de Cardiff dans l'affaire Bridges v. the South Wales Police (High Court of Justice 2019)

26 La police a affirmé qu'elle n'avait pas accès à la composition démographique de l'ensemble des données de formation pour l'algorithme adopté, « Neoface ». La Cour note que « le fait demeure, cependant, que la police du sud du Pays de Galles n'a jamais cherché à s'assurer, directement ou par le biais d'une vérification indépendante, que le logiciel en l'espèce ne présente pas un biais inacceptable fondé sur la race ou le sexe ».

L'UE FINANCE À HAUTEUR DE 8 MILLIONS D'EUROS CE PROJET VISANT À DÉVELOPPER « DES MÉTHODES AMÉLIORÉES POUR LA SURVEILLANCE DES FRONTIÈRES »

Cette section met en évidence un lien crucial et souvent passé sous silence entre la biométrie et les systèmes de gestion des frontières de l'UE, en montrant clairement comment des technologies susceptibles de produire des résultats discriminatoires risquent d'être appliquées aux personnes – par exemple, celles en situation de migration – qui souffrent déjà le plus de la discrimination.

/ Reconnaissance faciale et utilisation des données biométriques dans les politiques et les pratiques de l'UE

Au cours de l'année passée, la reconnaissance faciale et d'autres types de technologies d'identification biométrique ont suscité beaucoup d'intérêt de la part des gouvernements, de l'UE, de la société civile et des organisations de défense des droits, en particulier dans le domaine de l'application de la loi et de la gestion des frontières.

En 2019, la Commission européenne a chargé un consortium d'organismes publics de « dresser la carte de la situation actuelle de la reconnaissance faciale dans les enquêtes criminelles de tous les États membres de l'UE », dans le but d'« avancer vers un échange potentiel de données faciales ». Elle a demandé au cabinet de conseil Deloitte de réaliser une étude de faisabilité sur l'extension du système d'images faciales Prüm. Prüm est un système qui connecte les bases de données d'ADN, d'empreintes digitales et d'immatriculation des véhicules à l'échelle européenne pour permettre des recherches à l'échelle de plusieurs pays. La crainte est qu'une base de données paneuropéenne des visages de ses citoyens puisse être utilisée pour instaurer une surveillance omniprésente, injustifiée ou illégale.

/ Systèmes de gestion des frontières sans frontières

Comme rapporté dans la précédente édition du rapport *L'automatisation de la société*, la mise en œuvre d'un système de gestion des frontières global, interopérable et intelligent dans l'UE, initialement proposé en 2013 par la Commission, suit son cours. Bien que les nouveaux systèmes annoncés (EES, ETIAS[27], ECRIS-TCN[28]) n'entreront en fonctionnement qu'en 2022, le règlement Entry/Exit System (EES) a déjà introduit pour la première fois dans la législation européenne les images faciales comme éléments d'identification biométrique et l'utilisation de la technologie de reconnaissance faciale à des fins de vérification[29].

L'agence des droits fondamentaux de l'Union européenne (FRA) a confirmé ces changements : « Le traitement des images faciales devrait être introduit de manière plus systématique dans les systèmes informatiques utilisés à grande échelle dans l'UE à des fins d'asile, d'immigration et de sécu-

27 ETIAS (EU Travel Information and Authorisation System) est le nouveau système d'exemption de visa pour la gestion des frontières de l'UE développé par eu-LISA. « Les informations soumises lors de la demande seront automatiquement traitées vis-à-vis des bases de données existantes de l'UE (Eurodac, SIS et VIS), des futurs systèmes EES et ECRIS-TCN, ainsi que des bases de données pertinentes d'Interpol. Cela permettra de vérifier à l'avance les risques potentiels en matière de sécurité, d'immigration illégale et de santé publique. » (ETIAS 2019)

28 Le système européen d'information sur les casiers judiciaires - ressortissants de pays tiers (ECRIS-TCN), qui doit être développé par eu-LISA, sera un système centralisé de recherche de concordance/non-concordance visant à compléter la base de données existante des casiers judiciaires de l'UE (ECRIS) sur les ressortissants de pays tiers condamnés dans l'Union européenne.

29 EES entrera en service au premier trimestre 2022, suivi par le système ETIAS d'ici la fin de l'année 2022 – qui devrait « changer la donne dans le domaine de la justice et des affaires intérieures (JAI) ».

rité [...] une fois que les mesures juridiques et techniques nécessaires auront été prises. »

Selon Ana Maria Ruginis Andrei, de l'Agence européenne pour la gestion opérationnelle des systèmes d'information à grande échelle au sein de l'espace de liberté, de sécurité et de justice (EU-LISA), cette nouvelle architecture d'interopérabilité étendue a été « assemblée afin de forger le moteur idéal pour lutter avec succès contre les menaces à la sécurité intérieure, pour contrôler efficacement l'immigration et pour éliminer les angles morts en matière de gestion de l'identité ». En pratique, cela consiste à « conserver les empreintes digitales, les images faciales et autres données personnelles de 300 millions de ressortissants extraeuropéens, en fusionnant les données de cinq systèmes distincts » (Campbell 2020).

/ ETIAS : contrôles de sécurité automatisés aux frontières

Le système européen d'information et d'autorisation concernant les voyages (ETIAS), qui n'est pas encore entré en fonction à l'heure où nous rédigeons ce rapport, utilisera différentes bases de données pour automatiser les contrôles de sécurité numériques des voyageurs extraeuropéens (ceux qui n'ont pas besoin de visa) avant leur arrivée en Europe.

Ce système permettra de recueillir et d'analyser des données pour la « vérification avancée des risques potentiels en matière de sécurité ou d'immigration illégale » (ETIAS 2020). Son objectif est de « faciliter les contrôles aux frontières ; d'éviter les lenteurs bureaucratiques et les retards pour les voyageurs lorsqu'ils se présentent aux frontières ; et d'assurer une évaluation coordonnée et harmonisée des risques des ressortissants de pays tiers » (ETIAS 2020).

Ann-Charlotte Nygård, chef de l'Unité d'assistance technique et de renforcement des capacités de la FRA, distingue deux risques spécifiques concernant l'ETIAS : « Premièrement, l'utilisation de données qui pourraient entraîner une discrimination involontaire de certains groupes, par exemple si un demandeur est issu d'un groupe ethnique particulier présentant un risque élevé d'immigration illégale ; deuxièmement, l'évaluation du risque de sécurité sur la base de condamnations passées dans le pays d'origine. Certaines de ces condamnations antérieures pourraient être considérées comme déraisonnables par les Européens, comme les condamnations des personnes LGBT dans certains pays. Pour éviter cela, [...] les algorithmes doivent être vérifiés pour s'assurer qu'ils ne sont pas discriminatoires, et ce type de vérification doit impliquer des experts de différents domaines » (Nygård 2019).

/ iBorderCtrl : reconnaissance faciale et évaluation des risques aux frontières

iBorderCtrl était un projet impliquant les agences de sécurité de Hongrie, de Lettonie et de Grèce qui visait à « permettre des contrôles aux frontières plus rapides et plus approfondis des ressortissants de pays tiers franchissant les frontières terrestres des États de membres de l'UE ». iBorderCtrl utilisait une technologie de reconnaissance faciale, un détecteur de mensonges et un système de notation pour alerter la police aux frontières s'il jugeait qu'une personne était potentiellement dangereuse ou si son droit d'entrée paraissait douteux.

Le projet iBorderCtrl a pris fin en août 2019, et ses résultats – pour une mise en œuvre potentielle du système à l'échelle de l'UE – sont contradictoires.

Bien qu'il soit nécessaire de « déterminer jusqu'où le système ou une partie de celui-ci sera utilisé », la page « Résultats » du projet évoque « la possibilité d'intégrer les fonctionnalités similaires du nouveau système ETIAS et d'étendre les capacités liées à la procédure de passage des frontières là où les voyageurs se trouvent (bus, voiture, train, etc.). »

Toutefois, les modules auxquels il est fait référence ne sont pas spécifiés, et les outils ADM qui ont été testés n'ont pas fait l'objet d'une évaluation publique.

Dans le même temps, la page FAQ du projet confirme que le système qui a été testé n'est pas considéré comme « actuellement adapté au déploiement à la frontière [...] en raison de sa nature de prototype et de l'infrastructure technologique au niveau de l'UE ». Cela signifie que « des développements supplémentaires et une intégration au sein des systèmes existants de l'UE seraient nécessaires pour une utilisation par les autorités frontalières. »

En particulier, si le consortium iBorderCtrl a pu démontrer, en principe, le fonctionnement de cette technologie pour les contrôles aux frontières, il est également clair que les contraintes éthiques, légales et sociétales doivent être résolues avant tout déploiement à grande échelle.

/ Projets Horizon2020 associés

Plusieurs projets ultérieurs se sont focalisés sur le test et l'élaboration de nouveaux systèmes et de nouvelles technologies pour la gestion et la surveillance des frontières, dans le cadre du programme Horizon2020. Ceux-ci sont listés sur le site CORDIS de la Commission européenne, qui fournit des informations sur toutes les activités de recherche bénéficiant du soutien de l'UE qui s'y rapportent.

Le site montre que 38 projets sont actuellement en cours dans le cadre du programme/thème « H2020-EU.3.7.3. – Renforcer la sécurité par la gestion des frontières » de l'Union européenne. Son programme parent – « Sociétés sécurisées – Protection de la liberté et de la sécurité de l'Europe et de ses citoyens », qui est doté d'un budget global d'environ 1,7 milliard d'euros et finance 350 projets, prétend combattre « l'insécurité, qu'elle soit due à la criminalité, à la violence, au terrorisme, aux catastrophes naturelles ou d'origine humaine, aux cyberattaques ou aux atteintes à la vie privée, et à d'autres formes de troubles socioéconomiques qui touchent de plus en plus les citoyens » par le biais de projets visant principalement à développer de nouveaux systèmes technologiques basés sur l'IA et l'ADM.

Certains projets qui sont déjà achevés et/ou dont les applications sont déjà utilisées – par exemple, FastPass, ABC4EU, MOBILEPASS et EFFISEC – ont tous examiné les exigences en matière de « contrôles aux frontières automatisés (ABC) intégrés et interopérables », de systèmes d'identification et de portails « intelligents » à différents postes frontaliers.

TRESSPASS est un projet en cours qui a débuté en juin 2018 et prendra fin en novembre 2021. L'UE finance le projet à hauteur de près de 8 millions d'euros, et les coordinateurs d'iBorderCtrl (ainsi que de FLYSEC et XP-DITE) visent à « exploiter les résultats et les concepts mis en œuvre et testés » par iBorderCtrl et à « les développer pour créer une solution de sécurité multimodale pour le passage des frontières basée sur les risques, dans un cadre juridique et éthique solide » (Horizon2020 2019).

Le projet a pour objectif de transformer les contrôles de sécurité aux frontières en passant de l'ancienne stratégie, déclarée obsolète, « basée sur des règles » à une nouvelle stratégie « basée sur les risques ». Cette stratégie inclut l'application de technologies biométriques et de capteurs, d'un système de gestion basé sur les risques et de modèles pertinents pour évaluer l'identité, les possessions, les capacités et les intentions des individus. Elle vise à permettre des contrôles par le biais de « liens avec les systèmes existants et des bases de données externes telles que VIS/SIS/PNR » et recueille des données de toutes les sources susmentionnées à des fins de sécurité.

Un autre projet pilote, FOLDOUT, a débuté en septembre 2018 et se terminera en février 2022. L'UE finance à hauteur de 8 millions d'euros ce projet visant à développer « des méthodes améliorées pour la surveillance des frontières » afin de lutter contre l'immigration illégale, en mettant l'accent sur « la détection des personnes à travers un feuillage dense dans des climats extrêmes », en combinant « divers capteurs et technologies et en les fusionnant intelligemment en une plateforme de détection intelligente efficace et robuste » pour suggérer des scénarios de réaction. Des essais sont en cours en Grèce, en Finlande et en Guyane française.

MIRROR, pour *Migration-Related Risks caused by misconceptions of Opportunities and Requirement* (risques liés à la migration causés par des idées fausses sur les opportunités et les exigences), a débuté en juin 2019 et se poursuivra jusqu'en mai 2022. L'UE contribue à hauteur d'un peu plus de cinq millions d'euros à ce projet, qui vise à « comprendre comment l'Europe est perçue à l'étranger, détecter les divergences entre l'image et la réalité, repérer les cas de manipulation des médias et développer ses capacités à contrer ces idées fausses et les menaces pour la sécurité qui en découlent ». Sur la base d'une « analyse de la menace spécifique à la perception, le projet MIRROR combinera des méthodes d'analyse automatisée de textes, de supports multimédias et de réseaux sociaux avec des études empiriques » en vue de « développer des connaissances technologiques et des idées pratiques, [...] validées de manière approfondie avec les agences frontalières et les décideurs politiques, par exemple par le biais de projets pilotes. »

Parmi les autres projets déjà achevés, mais qui méritent d'être mentionnés, on peut citer Trusted Biometrics under Spoofing Attacks (TABULA RASA), qui a débuté en novembre 2010 et s'est achevé en avril 2014. Il visait à analyser « les faiblesses des logiciels d'identification biométrique dans le cadre de leur vulnérabilité au spoofing, diminuant ainsi l'efficacité des dispositifs biométriques » » Un autre projet, Bodega, qui a débuté en juin 2015 et s'est terminé en octobre 2018, a examiné comment mettre à profit « l'expertise du facteur humain » dans le cadre de « l'introduction de systèmes de contrôle aux frontières plus intelligents comme les portails automatisés et les systèmes de libre-service basés sur la biométrie ».

Références :

Access Now (2019) : Comments on the draft recommendation of the Committee of Ministers to Member States on the human rights impacts of algorithmic systems https://www.accessnow.org/cms/assets/uploads/2019/10/Submission-on-CoE-recommendation-on-the-human-rights-impacts-of-algorithmic-systems-21.pdf

AlgorithmWatch (2020) : Our response to the European Commission's consultation on AI https://algorithmwatch.org/en/response-european-commission-ai-consultation/

Campbell, Zach/Jones, Chris (2020) : Leaked Reports Show EU Police Are Planning a Pan-European Network of Facial Recognition Databases https://theintercept.com/2020/02/21/eu-facial-recognition-database/

CNIL (2019) : French privacy regulator finds facial recognition gates in schools illegal https://www.biometricupdate.com/201910/french-privacy-regulator-finds-facial-recognition-gates-in-schools-illegal

Coeckelbergh, Mark/Metzinger, Thomas(2020) : Europe needs more guts when it comes to AI ethics https://background.tagesspiegel.de/digitalisierung/europe-needs-more-guts-when-it-comes-to-ai-ethics

Comité de protection des données (2020) : Law enforcement directive https://www.dataprotection.ie/en/organisations/law-enforcement-directive

Comité des ministres (2020) : Recommendation CM/Rec(2020)1 of the Committee of Ministers to Member States on the human rights impacts fo algorithmic systems https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154

Comité européen de la protection des données (2019) : Facial recognition in school renders Sweden's first GDPR fine https://edpb.europa.eu/news/national-news/2019/facial-recognition-school-renders-swedens-first-gdpr-fine_en

Commissaire aux droits de l'homme (2020) : Unboxing artificial intelligence: 10 steps to protect human rights https://www.coe.int/en/web/commissioner/-/unboxing-artificial-intelligence-10-steps-to-protect-human-rights

Commission des affaires juridiques (2020) : Rapport préliminaire : With recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies https://www.europarl.europa.eu/doceo/document/JURI-PR-650508_EN.pdf

Commission des affaires juridiques (2020) : Artificial Intelligence and Civil Liability https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf

Commission des affaires juridiques (2020) : Rapport préliminaire : On intellectual property rights for the development of artificial intelligence technologies https://www.europarl.europa.eu/doceo/document/JURI-PR-650527_EN.pdf

Commission des libertés civiles, de la justice et des affaires intérieures (2020) : Rapport préliminaire : On artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters https://www.europarl.europa.eu/doceo/document/LIBE-PR-652625_EN.pdf

Commission européenne (2018) : White paper: On Artificial Intelligence - A European approach to excellence and trust https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

Commission européenne (2020) : A European data strategy https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

Commission européenne (2020) : Shaping Europe's digital future – Questions and Answers https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_264

Commission européenne (2020) : White Paper on Artificial Intelligence: Public consultation towards a European approach for excellence and trust https://ec.europa.eu/digital-single-market/en/news/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence

Commission européenne (2018) : Security Union: A European Travel Information and Authorisation System - Questions & Answers https://ec.europa.eu/commission/presscorner/detail/en/MEMO_18_4362

Contrôleur européen de la protection des données (2019) : Facial recognition: A solution in search of a problem? https://edps.europa.eu/press-publications/press-news/blog/facial-recognition-solution-search-problem_de

Delcker, Janosch(2020) : Decoded: Drawing the battle lines — Ghost work — Parliament's moment https://www.politico.eu/newsletter/ai-decoded/politico-ai-decoded-drawing-the-battle-lines-ghost-work-parliaments-moment/?utm_source=POLITICO.EU&utm_campaign=5a7d137f82-EMAIL_CAMPAIGN_2020_09_09_08_59&utm_medium=email&utm_term=0_10959edeb5-5a7d137f82-190607820

EDRi (2019) : FRA and EDPS: Terrorist Content Regulation requires improvement for fundamental rights https://edri.org/our-work/fra-edps-terrorist-content-regulation-fundamental-rights-terreg/

ETIAS (2020) : European Travel Information and Authorisation System (ETIAS) https://ec.europa.eu/home-affairs/what-we-do/policies/borders-and-visas/smart-borders/etias_en

ETIAS (2019) : European Travel Information and Authorisation System (ETIAS) https://www.eulisa.europa.eu/Publications/Information%20Material/Leaflet%20ETIAS.pdf

Groupe d'experts de haut niveau sur l'intelligence artificielle (2020) : Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment https://ec.europa.eu/digital-single-market/en/news/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

Horizon2020 (2019) : robusT Risk basEd Screening and alert System for PASSengers and luggage https://cordis.europa.eu/project/id/787120/reporting

High Court of Justice (2019) : Bridges v. the South Wales Police https://www.judiciary.uk/wp-content/uploads/2019/09/bridges-swp-judgment-Final03-09-19-1.pdf

Hunton Andrew Kurth (2020) : UK Court of Appeal Finds Automated Facial Recognition Technology Unlawful in Bridges v South Wales Police https://www.huntonprivacyblog.com/2020/08/12/uk-court-of-appeal-finds-automated-facial-recognition-technology-unlawful-in-bridges-v-south-wales-police/

Kayalki, Laura (2019) : French privacy watchdog says facial recognition trial in high schools is illegal https://www.politico.eu/article/french-privacy-watchdog-says-facial-recognition-trial-in-high-schools-is-illegal-privacy/

Kayser-Bril, Nicolas (2020) : EU Commission publishes white paper on AI regulation 20 days before schedule, forgets regulation https://algorithmwatch.org/en/story/ai-white-paper/

Leyen, Ursula von der (2019) : A Union that strives for more - My agenda for Europe https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf

Leyen, Ursula von der (2020) : Paving the road to a technologically sovereign Europe https://delano.lu/d/detail/news/paving-road-technologically-sovereign-europe/209497

Leyen, Ursula von der (2020) : Shaping Europe's digital future https://twitter.com/eu_commission/status/1230216379002970112?s=11

Leyen, Ursula von der (2019) : Opening Statement in the European Parliament Plenary Session by Ursula von der Leyen, Candidate for President of the European Commission https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_19_4230

Nygård, (2019) : The New Information Architecture as a Driver for Efficiency and Effectiveness in Internal Security https://www.eulisa.europa.eu/Publications/Reports/eu-LISA%20Annual%20Conference%20Report%202019.pdf

Parlement européen (2020) : Artificial intelligence: EU must ensure a fair and safe use for consumers https://www.europarl.europa.eu/news/en/press-room/20200120IPR70622/artificial-intelligence-eu-must-ensure-a-fair-and-safe-use-for-consumers

Parlement européen (2020) : On automated decision-making processes: ensuring consumer protection and free movement of goods and services https://www.europarl.europa.eu/doceo/document/B-9-2020-0094_EN.pdf

Police du sud du Pays de Galles (2020) : Automated Facial Recognition https://afr.south-wales.police.uk/

RGPD (Art 22) : Automated individual decision-making, including profiling https://gdpr-info.eu/art-22-gdpr/

Sabbagh, Dan (2020) : South Wales police lose landmark facial recognition case https://www.theguardian.com/technology/2020/aug/11/south-wales-police-lose-landmark-facial-recognition-case

Valero, Jorge (2020) : Vestager: Facial recognition tech breaches EU data protection rules https://www.euractiv.com/section/digital/news/vestager-facial-recognition-tech-breaches-eu-data-protection-rules/

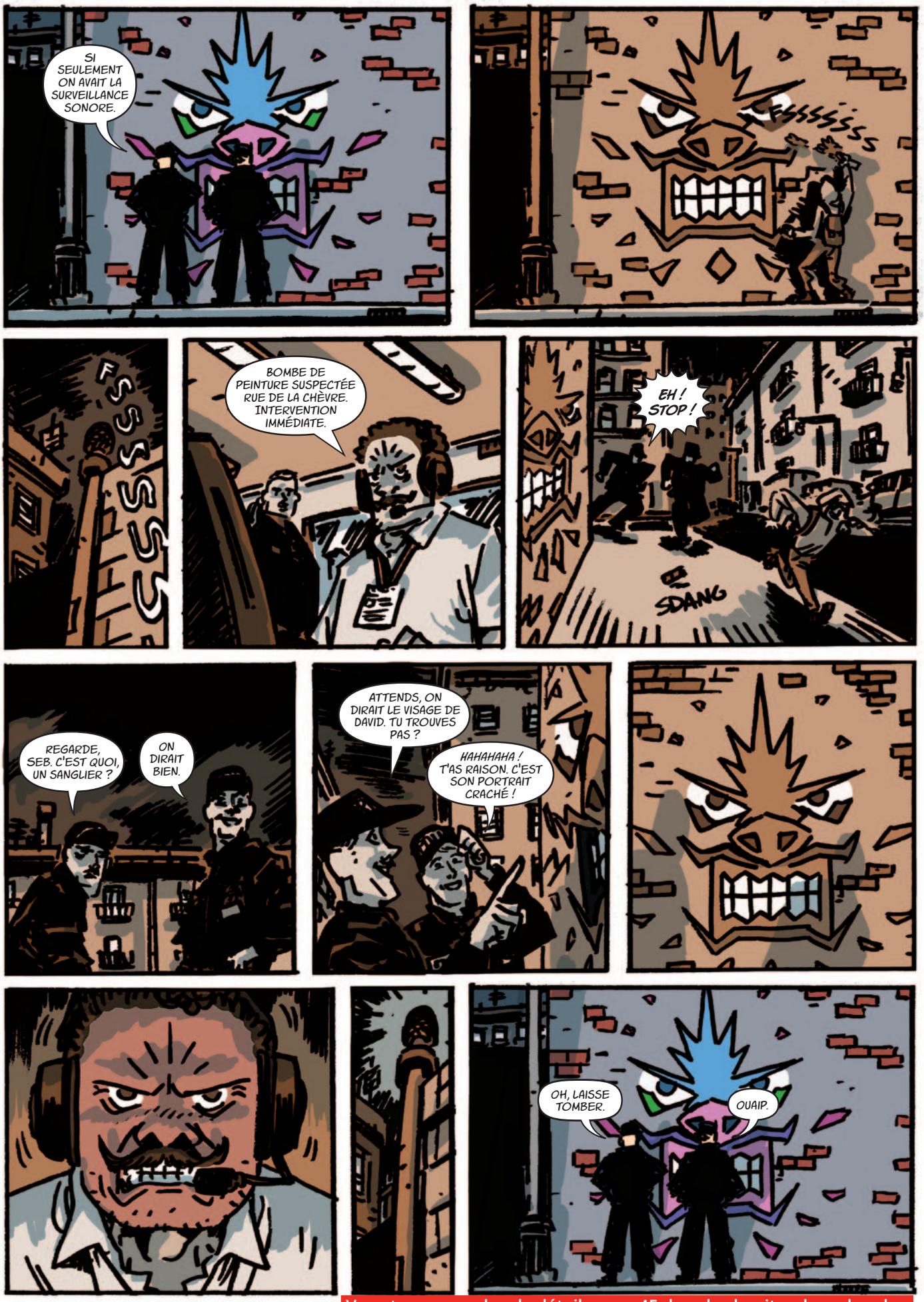
FRANCE

ARTICLE

PAGE 38

RECHERCHE

PAGE 44



Vous trouverez plus de détails page 45 dans le chapitre de recherche, sous l'intertitre "Automatisation de la surveillance dans les grandes villes".

Entre santé et contrôle : Deux siècles de gestion des données des patient·es

Le « Health Data Hub » se prépare à offrir les données de santé des Français·es aux startups qui en feront la demande. C'est la dernière étape d'un projet de centralisation des données de santé qui a débuté il y a 200 ans, oscillant toujours entre soins et contrôle, mais plutôt vers le contrôle.

Par Nicolas Kayser-Bril

En 1876, au Congrès d'Hygiène de Bruxelles, le médecin français Olivier du Mesnil fut impressionné par le système d'information sanitaire de la capitale belge. Là bas, écrit-il, les médecins envoient immédiatement une copie du certificat de décès au bureau central d'hygiène. Chaque semaine, le personnel du bureau produit une carte de la ville où chaque décès est marqué d'une épingle, avec des couleurs différentes pour chaque maladie contagieuse, et la montre au bourgmestre. Cette technologie, écrit M. du Mesnil, « permet à l'administrateur le plus étranger aux recherches scientifiques d'exercer un contrôle incessant sur l'état sanitaire de la population1. »

M. du Mesnil utilise l'exemple belge pour souligner le retard de la France. Comme son voisin du nord, la France était à l'époque un pays qui s'industrialisait rapidement. Les chemins de fer accéléraient le transport des personnes et des animaux, et les usines regroupaient grand nombre de travailleurs, entassés dans des logements insalubres et précaires.

Source gallica.bnf.fr/ Bibliotheque de France.

Collecte de données, 1930

/ Contrôler le choléra

Les principaux bénéficiaires de cette combinaison étaient les bactéries, qui pouvaient sauter rapidement d'hôte en hôte sur de grandes distances. Deux épidémies de choléra, en 1832 et 1854, firent plus de 100 000 morts chacune[2]. La bactérie responsable de la maladie ne fut identifiée qu'en 1884, mais les gouvernements européens comprirent bien avant qu'ils avaient besoin de collecter des données pour endiguer la progression de la maladie, par exemple pour décider des politiques de confinement. Après la première épidémie de choléra, le gouvernement français commença

à recueillir des informations sur les causes possibles du choléra : le mauvais air, la nourriture pourrie et même « l'ignorance[3] ».

Malgré sa létalité, le choléra était loin d›être le principal tueur du XIXe siècle. La dysenterie et la tuberculose étaient largement plus dangereux, mais étaient limitées aux couches les plus pauvres de la société[4]. Voir le gouvernement agir en priorité contre le choléra – qui tuait parmi toutes les classes sociales – était pour beaucoup la preuve qu'il ne s'intéressait à la santé des pauvres uniquement lorsqu'elle avait un impact sur celle des riches. Les mesures gouvernementales de lutte contre le choléra, y compris la collecte de données, se heurtèrent à la méfiance d'une large part de la population[5].

/ Police sanitaire

Au début du XXe siècle, la santé était une question d'ordre public plutôt que de bien-être. Jusqu'à la première guerre mondiale, la santé était du ressort du ministère de l'intérieur. Pour suivre et prédire la propagation de la tuberculose dans les grandes villes, les autorités introduisirent un « casier sanitaire » pour chaque bâtiment, calqué sur le casier judiciaire de chaque individu. Les travailleurs devaient présenter les deux pour obtenir un emploi. (Le casier sanitaire a été interrompu vers 1908. Le personnel hospitalier trouvait que les données agrégées au niveau de l'immeuble n'étaient pas très utiles[6].)

Un changement de perspective se produisit dans les premières décennies du XXe siècle. D'une part, les besoins de l'armée évoluèrent. L'ampleur des combats devint telle qu'il ne suffisait plus d'enrôler une partie d'une classe d'âge. Il fallait enrôler tous les jeunes hommes disponibles, et il fallait qu'ils soient en bonne santé. Le niveau de santé général de la population acquit alors une importance militaire. Par ailleurs, l'eugénisme, une discipline pseudo-scientifique

> UNE TELLE COLLECTE DE DONNÉES METTAIT EN DANGER LA CONFIDENTIALITÉ DES ÉCHANGES ENTRE LE MÉDECIN ET SES PATIENT ES, MAIS LEUR PRINCIPALE PRÉOCCUPATION CONCERNAIT SANS DOUTE LA POSSIBLE DIMINUTION DE LEUR STATUT.

qui prétendait améliorer une population en en extirpant ses membres en mauvaise santé, gagnait en popularité. La santé et l'hygiène devinrent des objectifs politiques en elles-mêmes. On leur donna leur propre ministère en 1920[7].

/ Monopole du savoir

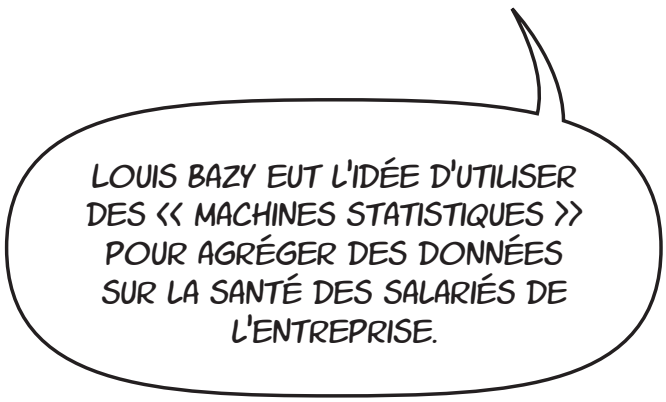
Les statistiques de la santé, qui ne s'intéressaient jusqu'alors qu'aux épidémies et au contrôle des pauvres, tournèrent leur attention vers le bien-être. La Société des Nations (SDN), une organisation internationale nouvellement créée et fondamentalement optimiste, était le fer de lance du mouvement. Dans de nombreux pays, la SDN commanda des enquêtes sur la santé et le bien-être.

Tous les médecins français n'étaient pas enthousiasmés par cette idée[8]. Pour certains d'entre eux, une telle collecte de données mettait en danger la confidentialité des échanges entre le médecin et ses patient·es. Leur principale préoccupation concernait sans doute la possible diminution de leur statut. À l'époque, les médecins étaient les dépositaires ultimes des connaissances médicales. La transmission d'informations à l'État était considérée comme une dévolution de pouvoir. Les médecins étant presque entièrement financés par leurs patients, ils n'avaient aucune raison de coopérer avec des systèmes qui semblaient menacer leur statut.

En tout état de cause, l'ambition de collecter de données de santé pour améliorer le bien-être de la population était limitée à une fraction des contribuables français·es. Dans les colonies, la santé était toujours considérée comme un facteur de production, à optimiser dans la mesure où elle rendait les plantations et les mines plus rentables. Jusqu'en 1946, les colonies françaises employaient au plus quatre statisticiens dans leur ensemble. Il est peu probable que les données de santé fussent pour eux une priorité[9].

/ Aiguilles à tricoter

Parmi les médecins, certains virent dans les données structurées une opportunité. En 1929, Louis Bazy, chirurgien consultant pour la compagnie ferroviaire Paris-Orléans, eut l'idée d'utiliser les « machines statistiques » de son



employeur pour agréger des données sur la santé des salariés de l'entreprise. Il conçut un système où l'affliction de chaque employé malade était codée sur une carte perforée, avec ses données personnelles. La machine à statistiques pouvait traiter 400 fiches par minute et, à l'aide de la « machine à tabuler », fournissait des informations sur la propagation d'une maladie et des corrélations entre les variables. Les applications de sa méthode à la médecine et à la recherche était infinies, écrivit M. Bazy[10].

Tous les médecins n'avaient pas accès à ces machines à calculer. *Le Mouvement Sanitaire*, un magazine pour médecins progressistes, publia en 1933 un article à leur attention, expliquant comment traiter les cartes perforées avec des aiguilles à tricoter[11]. Malgré ces efforts de vulgarisation, il est peu probable que ces premières tentatives de médecine informatisée aient gagné de nombreux adeptes.

Les cartes perforées de M. Bazy

/ La tendance à la centralisation

Pendant la seconde guerre mondiale, le gouvernement français fit de gros efforts pour implémenter ses politiques eugénistes. Il créa pour cela un Institut national d'hygiène (INH). À partir de 1942, l'institut démarra une collecte de données à grande échelle pour suivre les effets de la répression gouvernementale contre l'alcoolisme et les maladies vénériennes. L'INH construisit également un référentiel central d'informations sur 35 000 patient·es atteint·es de cancer[12]. Après la guerre, l'INH s'agrandit et continua de surveiller la santé du pays (il devint l'Institut national de la santé et de la recherche médicale, Inserm, en 1964).

Peu après, le gouvernement d'après-guerre offrit une assurance maladie à tous·tes les citoyen·nes. Avec elle arriva le numéro de sécurité sociale donné à la naissance, qui reste immuable jusqu'à la mort. Le fait de disposer d'un identifiant unique pour chaque citoyen·ne relança le vieux rêve de gouvernance par les nombres, où les décisions, forcément rationnelles, seraient prises uniquement sur la base de données.

En France comme dans d'autres pays du bloc occidental, la planification centrale était considérée comme une nécessité. Le gouvernement estima devoir collecter des données complètes sur la morbidité (c'est-à-dire sur les maladies affectant la population). Une première tentative pour forcer les médecins hospitaliers à remplir des formulaires après chaque intervention puis à les envoyer à une autorité centrale échoua en 1945. Une autre tentative fut faite en 1958 et une autre en 1972. Comme dans les années 1930, les médecins évitèrent d'effectuer ces nouvelles tâches. Ils critiquaient la méthodologie, se plaignaient de la charge de travail supplémentaire et déclaraient qu'ils n'y voyaient aucun avantage pour eux[13].

/ Numérisation

Tout changea dans les années 1980. Une nouvelle tentative de centralisation des données de morbidité commença en 1982. Au début de la décennie suivante, tous les hôpitaux transmettaient des données à une autorité centrale.

Ce succès – pour l'État – est peut être lié à l'environnement économique des années 1980. La faible croissance économique a incité le gouvernement à réduire les dépenses de santé. Malgré les réticences initiales des médecins, plusieurs ministres de la Santé firent pression et rendirent le nouveau système obligatoire en 1991[14].

L'effort de collecte de données visait avant tout à contrôler et réduire les coûts. En connaissant le nombre de procédures effectuées par chaque hôpital et les montant d'argent reçus, le ministère de la Santé pouvait classer leurs performances, au moins en termes financiers. Les hôpitaux qui dépensaient trop étaient rappelés à l'ordre. Dans les années 2000, le gouvernement alla plus loin et introduisit la tarification à l'acte. Une crise cardiaque de niveau 1 vaut ainsi 1 205,57 €, et passe à 1 346,85 € si le patient décède dans les deux jours[15]. Chaque intervention effectuée par les médecins est codée selon une classification stricte et les hôpitaux sont payés en conséquence par la sécurité sociale.

Pour parcourir la liste de plus de 6 000 procédures, les hôpitaux embauchent des consultants externes pour « optimiser » leurs pratiques de codage. Comme l'expliquait AlgorithmWatch en mai 2019, l'optimisation du code n'est rien de moins qu'une « triche généralisée » pour maximiser les revenus, selon un sociologue de la santé de l'université de Lille 2.

/ Problèmes de qualité

La France ayant un système d'assurance maladie obligatoire et un payeur unique, les données de morbidité peuvent être reliées à la consommation de médicaments, dès lors qu'un médicament est remboursé par la sécurité sociale. Pour environ 99% de la population, la Caisse nationale d'assurance maladie dispose d'une information complète sur

les procédures hospitalières et les médicaments prescrits depuis le début des années 1990[16].

Cet ensemble de données, unique au monde, a permis aux chercheur·ses de trouver des corrélations entre médicaments et morbidité. C'est ainsi que le benfluorex (vendu sous le nom de Mediator) a été lié à une valvulopathie cardiaque, conduisant au retrait du médicament en 2009.

Cependant, les informations sur les procédures hospitalières collectées dans ces bases de données sont de nature comptable, et non médicale. L'optimisation du codage des procédures nuit grandement à la qualité des données, car les hôpitaux ont interpréter la réalité de manière à augmenter leur recettes. Personne ne sait exactement à quel point les données sont mauvaises, car très peu d'études sont faites. Une étude menée en 2011 a par exemple montré que, sur une procédure spécifique, le mauvais code était utilisé huit fois sur dix[17].

/ FOMO

Malgré cette performance épouvantable, en 2019, le gouvernement français a fait pression pour construire une base de données encore plus grande, appelée « Health Data Hub ». Cédric Villani, un mathématicien qui a dirigé la stratégie d'intelligence artificielle du président Emmanuel Macron, a écrit dans un rapport parlementaire que le vrai risque pour la santé serait de « ne pas s'ouvrir à l'IA[18] ».

Depuis 2004, le gouvernement a poussé tous·tes les Français·es à ouvrir un dossier de santé électronique (DSE). Après un démarrage lent, le DSE sera activé pour tous·tes par défaut en 2021 et devrait, à terme, être connecté au Health Data Hub[19].

La CNIL a critiqué le projet, citant ses objectifs trop vagues. En effet, les données du Hub pourront être utilisées par n'importe quelle entreprise qui en fait la demande, pourvu que son application soit « d'intérêt public », ouvrant la porte à des finalités entièrement commerciales. Les critiques ont également souligné que les données personnelles stockées dans le Hub sont pseudonymisées mais pas agrégées, si bien qu'elles peuvent être désanonymisées[20].

/ Relations toxiques

Un médecin qui souhaitait être cité uniquement comme Gilles a déclenché une « grève des données » lors du lancement officiel du Health Data Hub en décembre 2019. Lui et d'autres ont appelé leurs collègues à cesser de remplir les formulaires qui alimentent le Hub. Selon Gilles, depuis les années 80, la France est passée « d'une santé qui guérit à une santé qui compte », pointant du doigt les systèmes de gestion des coûts. Il ne voit aucun avantage dans la nouvelle base de données, affirmant que la recherche n'en avait pas besoin. « Le logiciel nous prive que du temps qui devrait être consacré aux soins des patients », ajoute-t-il.

Même s'il refuse de donner des chiffres sur le succès de la grève, la colère de Gilles est largement partagée. En janvier 2020, plus de mille médecins ont démissionné de leurs fonctions administratives, citant la tarification à l'acte comme étant l'un des plus gros problèmes[21].

Par ailleurs, le concepteur du Health Data Hub a quitté son poste peu après la mise en route du projet pour rejoindre une entreprise du secteur privé spécialisée dans la vente de données sur la santé. Il n'y voit aucun conflit d'intérêts[22].

/ Rendre la population gouvernable

Vu sue le temps long, la principale innovation du Health Data Hub est que, pour la première fois, un gouvernement français a utilisé un nom anglais pour un projet officiel[23]. La logique qui a conduit à sa création continue directement un projet vieux de près de deux siècles. Deux siècles d'efforts du gouvernement français pour recueillir des données sur la santé de ses citoyen·nes, pour rendre sa population plus « lisible » et plus gouvernable.

Personne ne sait ce que le Health Data Hub apportera, mais l'histoire donne quelques indications. Le système d'information que Bruxelles a mis en place dans les années 1870, que M. du Mesnil admirait tant, fonctionnait peut-être. La ville ne subit aucune grande épidémie jusqu'à la grippe espagnole de 1918.

Cela dit, sur la même période, aucune grande ville de France ne subit non plus d'épidémie. En revanche, l'espérance de vie à Bruxelles, par rapport à la campagne belge et aux autres grandes villes, a diminué entre 1885 et 1910[24]. Collecter des données sur la santé des citoyen·nes ne va pas toujours de pair avec l'amélioration de leur santé réelle.

Notes

1 Du Mesnil, O. L'Exposition et le Congrès d'Hygiène et de Sauvetage de Bruxelles en 1876. *Annales d'Hygiène Publique*, 1877, p. 11.
2 Bourdelais, P., Raulot, J. Y. & Demonet, M. La marche du choléra en France: 1832 et 1854. *Annales. Histoire, Sciences Sociales*, 33(1), 1978, p. 125-142.
3 Drouet, S. Santé et environnement en France: deux siècles de «symbiose» juridique (1802-2002). *Revue juridique de l'Environnement*, 28(3), 2003, p. 324
4 Sur la tuberculose, voir Bello, S., Signoli, M. & Dutour, O. Analyse de l'évolution de la mortalité par tuberculose du XVIIIe au XXe siècle dans quatre grandes villes françaises. *Médecine et Maladies Infectieuses*, 30(5), 2000, 275-283.
5 Sur les attitudes de la populations aux politiques de lutte contre le choléra, voir Morris, R. J. *Cholera 1832: the social response to an epidemic*. Taylor & Francis, 1976. Je n'ai pas trouvé de monographie similaire sur la France mais n'ai pas de raison de penser que la situation y fut différente.
6 Fijalkow, Y. Mesurer l'hygiène urbaine en épargnant les propriétaires: Le casier sanitaire des maisons de Paris. *Les Annales de la Recherche Urbaine*, 53(1), 1991, p. 73-78.
7 Drouet, S., op. cit.
8 Pour une discussion du problème entre docteurs, voir Le Secret Médical et la Collaboration des Médecins-Praticiens à l'Hygiène Sociale, *Le Mouvement sanitaire : organe officiel de l'Association des médecins hygiénistes français*, 1927, p. 626-630.
9 Le Service Colonial des Statistiques. *Bulletin de la Statistique* Générale de France, octobre-novembre 1945, 377-379.
10 Bazy, L. L'emploi des machines pour l'obtention des statistiques médicales. *La Presse Médicale*, 18 janvier 1933, p. 105-106.
11 R.H.H, L'emploi des machines pour l'obtention des statistiques médicales [Critique]. *Le Mouvement Sanitaire*, 1933, p. 121-122.
12 Picard, J. F. Aux origines de l'Inserm: André Chevallier et l'Institut national d'hygiène. *Sciences sociales et santé*, 21(1), 2003, 5-26.
13 Chaperon, J., Villard, J. M. & Riou, F. L'information médicale enjeu de la gestion hospitalière. *Politiques et management public*, 6(2), 1988, p. 35-46.
14 Sur les débuts du PMSI, voir Mossé, P. La rationalisation des pratiques médicales, entre efficacité et effectivité. *Sciences sociales et santé*, 16(4), 1988, 35-60.
15 Arrêté du 6 mars 2019 fixant pour l'année 2019 les éléments tarifaires mentionnés aux I et IV de l'article L. 162-22-10 du code de la sécurité sociale.
16 Bezin, J., Duong, M., Lassalle, R., Droz, C., Pariente, A., Blin, P. & Moore, N. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. *Pharmacoepidemiology and Drug Safety*, 26(8), 2017, p. 954–962.
17 Chantry, A. A., Deneux-Tharaux, C., Bal, G., Zeitlin, J., Quantin, C. & Bouvier-Colle, M. H. Le programme de médicalisation du systeme d'information (PMSI)–processus de production des données, validité et sources d'erreurs dans le domaine de la morbidité maternelle sévere. *Revue d'épidémiologie et de santé publique*, 60(3), 2012, p. 177-188.
18 Rapport n° 401 (2018-2019) de MM. Gérard Longuet, sénateur et Cédric Villani, député, fait au nom de l'Office parlementaire d'évaluation des choix scientifiques et technologiques, déposé le 21 mars 2019.
19 Cuggia, M., Polton, D., Wainrib, G. & Combes, S. Health Data Hub, Mission de Préfiguration. 2019. p.65.
20 Hourdeaux, J. «Health Data Hub»: le méga fichier qui veut rentabiliser nos données de santé. *Mediapart*, 2019.
21 Hôpital : «Nous avons pris la décision inédite et difficile de démissionner». *Libération*, 2020.
22 Foucart, S. & Horel, S. Données de santé : conflit d'intérêts au cœur de la nouvelle plate-forme. *Le Monde*, 2019.
23 Cette affirmation demande à être nuancée. "FranceConnect", en 2015, utilisa deux mots français dans un ordre étrange, et "Make our Planet Great again", en 2017, n'était qu'un appel à projets. "Choose France", début 2019, était le slogan d'un autre programme.
24 Eggerickx, T. & Debuisson, M. La surmortalité urbaine: le cas de la Wallonie et de Bruxelles à la fin du XIXe siècle (1889-1892). *Annales de démographie historique*, Jan. 1990, p. 23-41.

Rercherche

Par Nicolas Kayser-Bril

Contexte

Au sein des institutions et des grandes entreprises françaises, beaucoup se considèrent comme des pionniers de l'intelligence artificielle. Certes, ils et elles prétendent rarement être les meilleurs au monde – et pourtant presque tous·tes se plaignent de ne pas l'être. Ce paradoxe semble indiquer que les Français·es considèrent que leur place naturelle se trouve dans le peloton de tête.

Cet état d'esprit donne lieu à des situations cocasses, comme la fois où le gouvernement a baptisé sa stratégie d'intelligence artificielle « IA pour l'humanité » (sans demander son avis au reste de l'humanité, évidemment). Malgré cette pointe d'arrogance, les Français·es ont plusieurs arguments à l'appui de leurs prétentions. Pour commencer, la France sera sans doute le premier pays européen à disposer d'une base de données centralisée de l'identité biométrique de tous ses citoyens, permettant ainsi aux services publics d'offrir (ou d'exiger) l'identification par reconnaissance faciale. Deuxièmement, presque tous les ministères et les grandes villes ont achevé leurs campagnes de numérisation, et beaucoup d'entre eux utilisent maintenant leurs vastes bases de données pour implémenter divers algorithmes. Enfin, les autorités fiscales françaises sont les participants les plus actifs au système européen d'échange automatisé d'informations. Au cours de l'année 2018, celles-ci ont envoyé et reçu des informations sur près de 3 millions de contribuables, plus que tout autre pays. Par ailleurs, la France est le pays qui dispose du plus de moyens législatifs pour garder les algorithmes sous contrôle, bien que la loi soit appliquée de manière inégale, voire pas du tout.

Certaines entreprises du secteur privé, comme Thales ou Idemia (anciennement Morpho-Safran), sont des leaders mondiaux dans le domaine de la biométrie. Si vous avez déjà franchi une frontière internationale, il y a de fortes chances que vous ayez été soumis aux outils automatisés d'Idemia, qui gère le système d'information sur les visas de l'espace Schengen et le programme PreCheck de la TSA aux États-Unis. Ces entreprises font activement pression sur les agent·es de la fonction publique à tous les niveaux pour qu'ils et elles soutiennent leur travail. Cela prend souvent la forme de projets de surveillance menés dans le cadre d'initiatives de « villes intelligentes ». Cependant, cela mène parfois à prendre des mesures disproportionnées, comme ces écoles où des centaines d'enfants sont obligés de se soumettre à l'identification biométrique pour accéder aux cantines.

Dans ce pays où le simple fait de mesurer la discrimination raciale est un délit passible de cinq ans de prison, il existe peu d'études sur les biais algorithmiques. Malgré cette limitation, plusieurs organismes de surveillance, tant au sein de l'administration que dans la société civile, documentent régulièrement les nombreux cas où la prise de décision automatisée porte atteinte à la liberté des citoyen·nes français·es. Cependant, la liste croissante des cas où l'opinion ou les conclusions de ces organismes sont ignorées par les autorités jette le doute sur la durabilité du modèle français, du moins pour ce qui est d'une prise de décision automatisée et démocratique.

Catalogue des nouvelles applications de l'ADM

/ ALICEM

Bien qu'il ne s'agisse pas d'un système de prise de décision automatisée en soi, Alicem est un programme gouvernemental qui devrait permettre à tout secteur du système administratif français d'offrir une identification par reconnaissance faciale. Initialement prévu pour début 2020, il a été mis en veilleuse après un tollé public fin 2019.

Dans le cadre de ce programme, les citoyens enregistrent les caractéristiques biométriques de leur visage à l'aide d'une application pour smartphone. Ils peuvent ensuite utiliser l'application pour effectuer les procédures administratives qui exigent actuellement de se rendre en personne dans une agence gouvernementale.

Alicem pourrait fonctionner comme une base de données centralisée des caractéristiques biométriques faciales de ses citoyens (le gouvernement insiste bien sur le fait qu'aucune donnée biométrique n'est stockée). Même si, officiellement, il n'est pas prévu d'utiliser cette base de données à tout autre fin, cela ouvre potentiellement la voie à un large éventail d'applications, notamment la reconnaissance globale des visages par les caméras de vidéosurveillance présentes sur la voie publique.

Début septembre 2019, un expert en sécurité a révélé que les développeurs de l'application avaient publié une partie du code sur Stack Overflow (un forum pour développeurs informatiques), publié des vidéos privées du projet sur YouTube, et avaient également rendu les serveurs de préproduction du projet librement accessibles en ligne. Autant de signes de procédures de sécurité pour le moins lacunaires, a fortiori pour une base de données biométriques officielle. Comme dans le cas d'Aadhaar, la célèbre base de données biométriques indienne victime de fuites, il est presque certain que les données détenues par Alicem finiront par tomber entre les mains de criminels.

Après la publication par Bloomberg d'un rapport sur le projet en octobre 2018, la levée de boucliers a conduit le gouvernement à retarder le lancement de l'application. Son lancement est prévu pour 2020, mais aucune date officielle de lancement n'a été rendue publique.

/ Automatisation de la surveillance dans les grandes villes

À Saint-Étienne (175 000 habitants), la ville prévoyait de mettre en place des microphones afin de détecter automatiquement les bruits suspects. Le projet devait être mis en œuvre dans un quartier pauvre, en coordination avec un système de vidéosurveillance et un drone autonome équipé d'une caméra. Il était prévu d'enregistrer tous les sons « suspects », notamment les coups de feu, mais aussi les perceuses électriques, les aérosols et les sifflets.

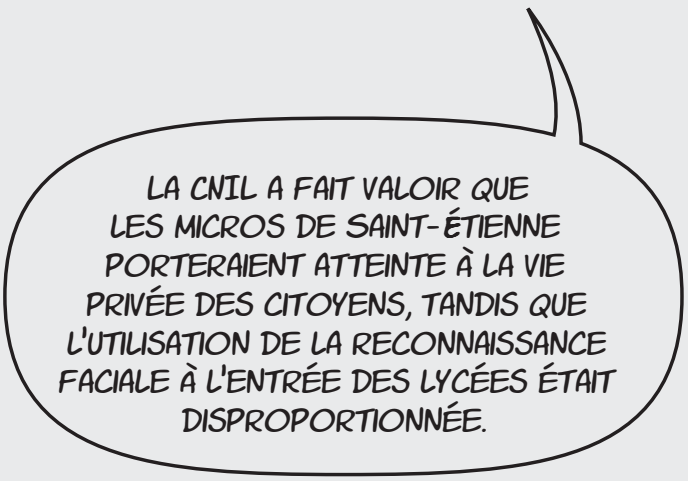
Un autre plan, cette fois-ci dans les deux plus grandes villes du sud du pays, Nice et Marseille, suggérait d'introduire dans certains lycées la reconnaissance faciale à l'entrée du bâtiment. En vertu de ce plan, les élèves auraient dû se soumettre à un contrôle facial avant de pouvoir entrer dans leur lycée. Ce programme devait être financé et mis en œuvre par Cisco, une société américaine.

Cependant, ces deux projets ont été suspendus après avoir été jugés illégaux par la CNIL, l'autorité française de protection des données, qui a fait valoir que les microphones de Saint-Étienne porteraient atteinte à la vie privée des citoyens, tandis que l'utilisation de la reconnaissance faciale à l'entrée des lycées était disproportionnée.

/ Le score cœur

Début 2018, après avoir pris des mesures similaires pour les patient·es souffrant de problèmes hépatiques et rénaux, les hôpitaux français ont introduit un « score cœur » pour les patient·es ayant besoin d'une transplantation cardiaque. Auparavant, les patient·es en attente d'une transplantation cardiaque étaient divisés en deux catégories, « urgence » et « super urgence », selon la gravité de leur état. Cependant, un examen des pratiques passées a révélé qu'un·e patient·e sur quatre classé en « super urgence » pour une greffe du cœur n'était pas réellement à haut risque. En revanche, un tiers des patient·es à haut risque n'étaient pas classés dans la catégorie « super urgence ».

Le système actuel calcule un score à chaque fois qu'un nouveau cœur est disponible pour une transplantation. L'algorithme qui calcule le score est transparent et relativement compréhensible (les jeunes, par exemple, reçoivent plus de points, et lorsque la compatibilité entre le donneur et le receveur potentiel est moindre, cela se traduit par moins de points), et le score est accessible aux médecins en ligne. En



fonction de ce score, les médecins choisissent la personne qui recevra l'organe. La Cour des comptes a fait l'éloge du nouveau système dans un rapport de 2019. Toutefois, elle a averti que les données des patients étaient souvent erronées car le personnel hospitalier ne tenait pas les dossiers numériques à jour. Le rapport indique que, dans un hôpital, jusqu'à un·e patient·e sur quatre fait l'objet de données erronées. D'après les auditeurs, ces manquements compromettent l'objectivité et l'acceptation du système.

/ L'apprentissage automatique contre la fraude fiscale

Le ministère des Finances a introduit un amendement dans la loi de finances 2020 qui lui permettra de récupérer des données sur les réseaux sociaux, les sites web de petites annonces ou les plateformes de vente aux enchères afin de détecter la fraude fiscale. Les autorités fiscales pourront, par exemple, comparer le train de vie d'une personne sur Instagram avec ses déclarations d'impôts. L'apprentissage automatique devrait permettre de déceler les malversations à partir d'un jeu de données de formation. Cependant, la CNIL a vivement critiqué ce plan.

Le projet de loi laissera le programme fonctionner pendant les trois prochaines années. Plusieurs députés ont fait appel devant le Conseil constitutionnel, dans l'espoir d'obtenir l'annulation de la mesure. Mais, dans une décision finale, le Conseil a jugé que le projet était légal.

L'automatisation à Pôle emploi

L'agence nationale pour l'emploi a numérisé chaque procédure en vue d'accélérer le traitement des dossiers. Si cette démarche s'est avérée bénéfique pour les situations les plus simples, elle a créé de véritables goulets d'étranglement pour les cas complexes. Les documents numérisés, par exemple, sont traités par des sous-traitants qui les classent parfois de manière inappropriée, ce qui entraîne des décisions erronées. En fin de compte, les travailleurs sociaux se sont plaints d'une augmentation de leur charge de travail, car ils et elles doivent reprendre à la main une grande partie des processus automatisés.

/ « Health Data Hub »

Début décembre 2019, le gouvernement a créé une structure juridique pour héberger le « Health Data Hub » (HDH), la plateforme des données de santé. Le HDH est une plateforme qui recueillera toutes les données produites par le système de santé publique, et mettra ces données à la disposition des start-ups et des entreprises qui ont un projet « d'intérêt public ». Ce projet met en œuvre la stratégie d'IA définie par le président Macron en 2018, dont la santé est l'un des quatre piliers. Les hôpitaux publics ne sont pas satisfaits de cette initiative, car elle leur enlève les données qu'ils conservent de manière indépendante depuis des années. Le fait que la plateforme soit partiellement hébergée sur Microsoft Azure, une solution de cloud computing, fait craindre que des données sensibles soient partagées avec des tiers étrangers. Toutes les informations partagées sur le « Health Data Hub » sont censées être anonymes, mais comme elles sont susceptibles d'être partagées dans un format non agrégé, une réidentification reste possible.

/ Discours de haine sur les réseaux sociaux

La loi « Avia », largement calquée sur la loi allemande *NetzDG*, a été adoptée par le Parlement français au début de l'année 2020. Cette loi aurait forcé les réseaux sociaux à supprimer tout contenu identifié comme un discours haineux dans les 24 heures. Compte tenu de la vitesse requise, ces retraits auraient dû être effectués automatiquement.

La loi prévoyait également d'étendre les pouvoirs de censure de la police. Dans le cadre de la législation actuelle, les policiers peuvent censurer tout site web « faisant l'éloge du terrorisme » après avoir notifié le propriétaire du site 24 heures à l'avance. La nouvelle loi aurait réduit ce délai de notification à une heure. La police française interprète le terme « terrorisme » de manière très large, invoquant ce terme contre les groupes de gauche, les écologistes et de nombreuses communautés arabes ou musulmanes, quelles que soient leurs intentions.

Cependant, ces dispositions fondamentales de la loi ont été invalidées par le Conseil constitutionnel en juin, la définition de discours de haine étant extrêmement vague, incluant notamment « l'éloge du terrorisme » et « la publication de contenu pornographique accessible aux mineurs ».

Comment l'ADM a-t-elle évolué au cours de l'année dernière ?

/ Une poignée d'algorithmes de l'administration sont désormais ouverts, mais beaucoup d'autres restent fermés

La « loi pour une République numérique » de 2016 dispose que tout·e citoyen·ne peut demander à voir les règles sous-jacentes d'un algorithme utilisé pour prendre une décision. Cependant, depuis l'entrée en vigueur de la loi en 2017, seule une poignée d'algorithmes a été rendu publics, mais l'administration est convaincue que la transparence finira par s'imposer.

La situation pourrait bien changer, car la loi dispose qu'à partir du 1er juillet 2020, toute décision prise par l'administration sur la base d'un algorithme fermé sera considérée comme nulle.

Un chercheur a analysé les demandes d'accès à l'information contenant le mot « algorithme » dans les archives de la Commission d'accès aux documents administratifs (CADA). Sur les 25 demandes qu'il a analysées (entre 2014 et 2018), il a pu constater que les Français·es ignoraient apparemment leur droit à connaître les règles régissant les algorithmes du secteur public.

/ Le retour peu populaire des radars automatiques

Selon le gouvernement, les Gilets jaunes ont détruit plus de 3 000 des 4 500 radars automatiques français. Pour les propriétaires de voitures mécontents, ces radars étaient un symbole des abus de pouvoir du gouvernement.

Les radars contrôlent automatiquement la vitesse de tous les véhicules, et les nouvelles versions de ces radars peuvent déceler d'autres comportements illégaux, comme l'utilisation d'un téléphone au volant. Ces dispositifs décriés sont de retour alors que le gouvernement prévoit de déployer 6 000 nouvelles unités améliorées d'ici la fin 2020.

/ Deux organismes de surveillance en moins, un de plus

Algotransparency et La Data en Clair, deux organisations de surveillance et de presse qui figuraient dans le précédent rapport *Automating Society*, ont cessé leurs activités en 2018. Les deux organisations n'ont donné aucune raison pour expliquer cet arrêt.

En août 2019, La Quadrature du Net, aux côtés de 26 autres organisations de la société civile, a lancé *Technopolice*, un observatoire des initiatives de « villes intelligentes ». Le groupe appelle à une « résistance méthodique et continue » contre ce qu'il considère une mise en œuvre de l'État de surveillance.

/ Bob Emploi

Bob Emploi, un service en ligne qui prétendait pouvoir réduire le chômage en mettant automatiquement en relation les demandeurs d'emploi et les offres d'emploi, n'a pas réussi à percer en 2019. Environ 50 000 comptes ont été créés au cours de l'année, alors que la France compte plus de 3 millions de chômeurs. Si le projet est toujours en cours de développement, les prétentions de ses algorithmes de mise en relation ont disparu du site officiel. Celui-ci ne propose plus désormais que de fournir des données pour « accompagner » les demandeur·ses d'emploi dans leurs recherches.

/ Parcoursup – Sélection des étudiant·es

Un syndicat d'étudiant·es a intenté un procès au gouvernement pour obtenir le code source d'un algorithme qui trie les candidats à l'université. Un tribunal de première instance a accédé à sa demande, mais cette décision a ensuite été cassée par le Conseil d'État. Bien que l'administration soit tenue, en vertu de la loi, de fournir les détails de tout algorithme utilisé pour prendre une décision qui affecte la vie d'un citoyen (conformément à la loi pour une République numérique mentionnée ci-dessus), le législateur a prévu une exemption pour la sélection à l'université. Les juges ont fondé leur décision exclusivement sur cette exemption.

Principales conclusions

La France continue à s'essayer avec enthousiasme à la prise de décision automatisée, tant au niveau national que local. Des organismes de surveillance, tels que la Commission nationale de l'informatique et des libertés (CNIL) et La Quadrature du Net, sont actifs et visibles dans les débats publics, mais leur pouvoir reste limité, notamment en raison des tensions politiques croissantes. Lorsque la CNIL a déclaré illégal le programme de reconnaissance faciale en région Provence-Alpes-Côte d'Azur, l'homme fort local, Renaud Muselier, a déclaré sur Twitter que la CNIL était une organisation « poussiéreuse », qui plaçait la sécurité des étudiants « en dessous de son idéologie ».

Le gouvernement a mis en œuvre certaines composantes de sa stratégie d'intelligence artificielle. Parmi les quatre piliers proposés en 2018, le transport, la sécurité et la santé ont reçu un soutien sous forme de financement ou de législation accélérée. Nous n'avons trouvé aucune mesure liée à l'ADM et au quatrième pilier, l'environnement.

Malgré une situation politique très instable, les questions relatives à l'ADM et à la collecte de données font toujours l'objet d'un débat public et politique. Des articles sur Parcoursup ont fait la une des journaux au début de l'année 2019. À l'automne, le programme de collecte massive de données par l'administration fiscale, baptisé « Big Brother Bercy », du nom du siège du ministère des Finances, a provoqué une vague de protestations suffisamment forte pour que certains députés proposent des amendements au projet.

De même, un projet qui prévoyait d'obliger tous les citoyens à utiliser la reconnaissance faciale sur une application mobile pour accéder aux services d'e-gouvernement, appelé Alicem, a été suspendu après avoir été dénoncé par des journaux et d'autres détracteurs comme étant disproportionné. Bien que plusieurs organisations de la société civile aient activement travaillé sur la question, il a fallu attendre la publication d'un article dans le média étasunien Bloomberg pour que le tollé prenne réellement forme.

Ce manque de stabilité politique risque d'empêcher les propositions et l'expérience françaises d'alimenter le débat à travers l'Europe. Par exemple, M. Villani, député du parti présidentiel, a rédigé la stratégie d'IA du pays en 2018. Cependant, après avoir déclaré sa candidature à la mairie de Paris contre le candidat de son propre parti, il est devenu un adversaire politique du président Macron.

Références :

Alderson, Eliot (2019). Various tweets, http://archive.is/dxafE http://archive.is/cuveX

Algorithmes: les administrations jouent-elles la transparence ? (2019) in: *Acteurs Publics*. [online] https://www.acteurspublics.fr/webtv/emissions/numerique-public/algorithmes-les-administrations-jouent-elles-la-transparence

BBC (2018) Aadhaar: 'Leak' in world's biggest database worries Indians, in: BBC [online], http://archive.is/aWjPG

Berne, Xavier (2019): Le Conseil d'État s'oppose à la communication des « algorithmes locaux » de Parcoursup, in: *Next INpact*, [online] https://web.archive.org/web/201906131606l2/https://www.nextinpact.com/news/107971-le-conseil-detat-soppose-a-communication-algorithmes-locaux-parcoursup.htm

Cellard, Loup (2019). Les demandes citoyennes de transparence au sujet des algorithmes publics [research note]. http://www.loupcellard.com/wp-content/uploads/2019/07/cellard_note_algo_public.pdf

Fouquet, Hélène (2019). France Set to Roll Out Nationwide Facial Recognition ID Program, in: *Bloomberg* [online], https://archive.is/ZyZOf

Guedj, Léa (2019): Pôle Emploi dématérialisé: le casse-tête des travailleurs précaires, in: *France Inter*, [online] https://web.archive.org/web/20191103130035/https://www.franceinter.fr/pole-emploi-dematerialise-le-casse-tete-des-travailleurs-precaires

Guide du Score Cœur (2018), Agence de la biomédecine, [online] https://www.agence-biomedecine.fr/IMG/pdf/guide_score_coeur_v2.pdf

Hourdeaux, Jérôme (2019): La Cnil juge illégale la reconnaissance faciale à l'entrée des lycées, in: *Médiapart*, [online] https://web.archive.org/web/20191030190441/https://www.mediapart.fr/journal/france/281019/la-cnil-juge-illegale-la-reconnaissance-faciale-l-entree-des-lycees

Hourdeaux, Jérôme (2019): «Health Data Hub»: le méga fichier qui veut rentabiliser nos données de santé, in: *Médiapart*, [online] https://web.archive.org/save/https://www.mediapart.fr/journal/france/221119/health-data-hub-le-mega-fichier-qui-veut-rentabiliser-nos-donnees-de-sante?onglet=full

La Quadrature du Net (2019): Mouchards et drones à Saint-Etienne : le maire veut étouffer le débat, in: *Technopolice*, [online] https://archive.is/A7zfc

La Voix Du Nord (2019): Budget 2020: le Conseil constitutionnel censure très partiellement la collecte de données [online] https://archive.is/xWsUB

Rees, Marc (2019):« Collecte de masse » : la CNIL critique le mégafichier de Bercy, in: *Next INpact*, [online] https://web.archive.org/web/20191119134909/https://www.nextinpact.com/news/108248-collecte-masse-cnil-critique-megafichier-bercy.htm

Rees, Marc (2019): La controverse #BigBrotherBercy se poursuit à l'Assemblée nationale, in: *NextINpact*, [online] https://archive.is/dEaaO

Rees, Marc (2020): Cyberhaine: mais que prévoit la proposition de loi Avia ?, in: *NextINpact*, [online] https://archive.is/MRMNz

Saviana, Alexandra (2019). Lancement de la reconnaissance faciale en France: mais qu'allons-nous faire dans cette galère ? in: *Marianne* [online], https://archive.is/tW4p7

Sécurité sociale 2019 (2019), Cour des Comptes, [online] https://www.ccomptes.fr/system/files/2019-10/RALFSS-2019-08-politique-des-greffes.pdf

Tesquet, Olivier (2019): Des micros dans la rue : la CNIL tire les oreilles (intelligentes) de Saint-Etienne, in: *Télérama*, [online] https://archive.is/aiKtF

Untersinger, Martin (2020): La disposition phare de la loi Avia contre la haine en ligne censurée par le Conseil constitutionnel, in: *Le Monde*, [online] https://archive.is/GvHlS

"Vitesse, téléphone, ceinture: qu'est-ce que le nouveau radar «tourelle» ? " in: *L'Express* [online] https://web.archive.org/web/20190826051258/https://www.lexpress.fr/actualite/societe/vitesse-telephone-ceinture-qu-est-ce-que-le-nouveau-radar-tourelle_2070713.html

L'équipe

/ Beate Autering

Mise en page et conception graphique

Beate Autering est une graphiste freelance. Elle est diplômée en design et dirige le studio beworx avec Tiger Stangl. Ensemble, il et elle créent des dessins, des graphiques et des illustrations et fournissent également des services d'édition d'images et de postproduction. Parmi leurs clients figurent iRights, mdsCreative, Agentur Sehstern, Patrimoine mondial de l'UNESCO et visitBerlin.

/ Fabio Chiusi

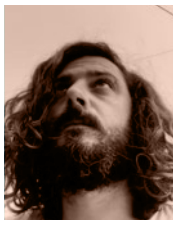
Rédacteur en chef du rapport, auteur de de l'introduction et du chapitre sur l'UE

Photo: Julia Bornkessel

Fabio Chiusi travaille chez AlgorithmWatch en tant que corédacteur et chef de projet pour l'édition 2020 du rapport L'automatisation de la société. Après une décennie dans le journalisme technologique, il a commencé à travailler comme consultant et assistant de recherche dans le domaine des données et de la politique (Tactical Tech) et de l'IA dans le journalisme (Polis LSE). Il a coordonné le rapport Persuasori Social sur la réglementation des campagnes politiques sur les réseaux sociaux pour le projet PuntoZero, et a travaillé en tant que collaborateur technico-politique au sein de la Chambre des députés du Parlement italien pendant la législature actuelle. Fabio est chercheur au Nexa Center for Internet & Society à Turin et professeur adjoint à l'université de Saint-Marin, où il enseigne le journalisme et les nouveaux médias, l'édition et les médias numériques. Il est l'auteur de plusieurs essais sur la technologie et la société. Le dernier en date, « Io non sono qui. Visioni e inquietudini da un futuro presente » (DeA Planeta, 2018), est actuellement en cours de traduction en polonais et en chinois. Il écrit également en tant que journaliste spécialisé dans la politique technologique pour le blog collectif ValigiaBlu.

/ Samuel Daveti

Auteur de bandes dessinées



Samuel Daveti est un membre fondateur de l'association culturelle Double Shot. Il est l'auteur du roman graphique en langue française Akron le guerrier (Soleil, 2009), et le coordinateur du volume anthologique Fascia Protetta (Double Shot, 2009). En 2011, il est devenu membre fondateur du collectif de bandes dessinées autoproduites Mammaiuto. Samuel a également écrit Un Lungo Cammino (Mammaiuto, 2014 ; Shockdom, 2017), qui fera l'objet d'un film réalisé par la société de médias Brandon Box. En 2018, il a écrit The Three Dogs, illustré par Laura Camelli, qui a remporté le prix Micheluzzi au Napoli Comicon 2018 et le prix Boscarato du meilleur webcomic au Festival de la bande dessinée de Trévise.

/ Sarah Fischer

Rédactrice du rapport

Sarah Fischer est chef de projet pour le projet « Éthique des algorithmes » à la Bertelsmann Stiftung, où elle est principalement responsable des études scientifiques. Elle a précédemment travaillé comme post-doctorante dans le cadre du programme « Confiance et communication dans un monde numérisé » à l'université de Münster, où elle s'est concentrée sur le thème de la confiance dans les moteurs de recherche. Dans ce même groupe de formation à la recherche, elle a obtenu son doctorat avec une thèse sur la confiance dans les services de santé sur Internet. Elle a étudié les sciences de la communication à l'université Friedrich Schiller de Iéna, et est le coauteur des articles « Where Machines can err. Sources of error and responsibilities in processes of algorithmic decision making » et « What Germany knows and believes about algorithms ».

/ Leonard Haas

Rédacteur adjoint

Leonard Haas travaille comme assistant de recherche chez AlgorithmWatch. Il est notamment responsable de la conception, de la mise en œuvre et de la maintenance de l'inventaire mondial des directives éthiques pour l'IA. Il est étudiant en master dans le domaine des sciences sociales à l'Université Humboldt de Berlin et détient deux licences de l'Université de Leipzig en humanités numériques et en sciences politiques. Ses recherches portent sur l'automatisation du travail et de la gouvernance. En outre, il s'intéresse à la politique des données d'intérêt public et aux luttes du travail dans l'industrie technologique.

/ Graham Holliday

Relecteur

Photo: Josh White

Graham Holliday est un rédacteur, auteur et professeur de journalisme indépendant. Il a occupé plusieurs postes à la BBC pendant près de deux décennies et a été correspondant de Reuters au Rwanda. Il travaille comme rédacteur pour les émissions Parts Unknown et Roads & Kingdoms de CNN – le journal international de la correspondance étrangère. Ses deux premiers livres, publiés par feu Anthony Bourdain, ont fait l'objet de recensions dans le New York Times, le Los Angeles Times, le Wall Street Journal, le Publisher's Weekly, le Library Journal et sur la radio NPR, entre autres médias.

/ Nicolas Kayser-Bril

Rédacteur de l'édition française, auteur du chapitre sur la France

Photo: Julia Bornkessel

Nicolas Kayser-Bril est un journaliste spécialiste des données qui travaille pour AlgorithmWatch en tant que reporter. Il a été le pionnier des nouvelles formes de journalisme en France et en Europe et est l'un des plus grands experts dans le domaine du datajournalisme. Il intervient régulièrement dans des conférences internationales, enseigne le journalisme dans des écoles de journalisme françaises et dispense des formations dans les salles de rédaction. Journaliste et développeur autodidacte (ainsi que diplômé en économie), il a commencé par développer de petites applications interactives basées sur des données pour le journal Le Monde à Paris en 2009. Il a ensuite constitué l'équipe de datajournalisme d'OWNI en 2010, avant de cofonder et de gérer Journalism++ de 2011 à 2017. Nicolas est également l'un des principaux contributeurs au Guide du datajournalisme, l'ouvrage de référence pour la vulgarisation du datajournalisme dans le monde.

/ Lorenzo Palloni

Auteur de bandes dessinées

Lorenzo Palloni est un dessinateur de bandes dessinées, auteur de plusieurs romans graphiques et webcomics, écrivain primé, et l'un des fondateurs du collectif d'auteurs de bandes dessinées Mammaiuto. Il travaille actuellement sur des romans pour les marchés français et italien. Lorenzo est également professeur d'écriture de scénarios et de storytelling à la Scuola Internazionale di Comics di Reggio Emilia (École internationale de bande dessinée de Reggio Emilia).

/ Kristina Penner

Autrice du chapitre sur l'UE

Kristina Penner est la conseillère exécutive d'AlgorithmWatch. Ses recherches portent sur les systèmes de sécurité sociale, la notation sociale et les impacts sociétaux de l'ADM, ainsi que sur la durabilité des nouvelles technologies d'un point de vue holistique. Son analyse du système de gestion des frontières de l'UE s'appuie sur son expérience antérieure en matière de recherche et de conseil sur le droit d'asile. Son expérience inclut également des projets sur l'utilisation des médias dans la société civile et le journalisme sensible aux conflits, ainsi que l'implication des parties prenantes dans les processus de paix aux Philippines. Elle est titulaire d'un master en études internationales/recherche sur la paix et les conflits de l'université Goethe de Francfort.

/ Alessio Ravazzani

Auteur de bandes dessinées

Alessio Ravazzani est un graphiste éditorial, dessinateur et illustrateur qui collabore avec les plus prestigieux éditeurs de bandes dessinées et de romans graphiques en Italie. Il est auteur au sein du collectif Mammaiuto, dont il est membre depuis sa fondation.

/ Matthias Spielkamp

Rédacteur du rapport

Photo: Julia Bornkessel

Matthias Spielkamp est cofondateur et directeur exécutif d'AlgorithmWatch. Il a témoigné devant plusieurs commissions du Bundestag allemand sur l'IA et l'automatisation. Matthias est membre du conseil d'administration de la section allemande de Reporters sans frontières et des conseils consultatifs de la Stiftung Warentest et du Whistleblower Network. Il a été membre de ZEIT Stiftung, de Stiftung Mercator et de l'American Council on Germany. Matthias a fondé le magazine en ligne mobilsicher.de, qui traite de la sécurité des appareils mobiles et compte plus de 170 000 lecteurs par mois. Il a écrit et édité des livres sur le journalisme numérique et la gouvernance de l'Internet et a été nommé l'un des 15 architectes bâtissant un avenir axé sur les données par Silicon Republic. Il est titulaire d'une maîtrise de journalisme de l'université du Colorado à Boulder et d'une maîtrise de philosophie de l'université libre de Berlin.

/ Marc Thümmler

Coordinateurs des publications

Photo: Julia Bornkessel

Marc Thümmler est responsable des relations publiques et de la sensibilisation chez AlgorithmWatch. Il est titulaire d'un master en études des médias, a travaillé comme producteur et monteur pour une société cinématographique, et a géré des projets pour la Deutsche Kinemathek et l'organisation de la société civile Gesicht Zeigen. En plus de ses fonctions principales chez AlgorithmWatch, Marc a participé à la campagne de financement et de crowdsourcing OpenSCHUFA, et il a coordonné le premier numéro du rapport L'automatisation de la société, publié en 2019.

ORGANISATIONS

/ AlgorithmWatch

AlgorithmWatch est une organisation de recherche et de plaidoyer à but non lucratif qui s'engage à surveiller et analyser les systèmes de prise de décision algorithmique ou automatisée (ADM) et leur impact sur la société. Si l'utilisation prudente des systèmes ADM peut profiter aux individus et à la société, elle comporte par ailleurs de grands risques. Afin de protéger l'autonomie humaine, les droits fondamentaux et afin maximiser le bien public, nous considérons qu'il est crucial de que les systèmes ADM soient responsables devant les institutions démocratiques. L'utilisation de systèmes ADM qui affectent de manière significative les droits individuels et collectifs doit être publique, de manière claire et accessible. Les individus doivent également être en mesure de comprendre comment les décisions sont prises et de les contester si nécessaire. Par conséquent, nous travaillons à permettre aux citoyen·nes de mieux comprendre les systèmes ADM et de développer des moyens de parvenir à une gouvernance démocratique de ces processus – avec un mélange de technologies, de réglementations et d›institutions de contrôle appropriées. Avec cela, nous nous efforçons de contribuer à une société juste et inclusive et de maximiser les avantages des systèmes ADM pour la société au sens large.

https://algorithmwatch.org/en/

ALGORITHM WATCH

/ Bertelsmann Stiftung

La Bertelsmann Stiftung œuvre pour la promotion sociale et l'inclusion pour tous. Elle s'est engagée à faire progresser cet objectif grâce à des programmes visant à améliorer l'éducation, à façonner la démocratie, à faire progresser la société, à promouvoir la santé, à dynamiser la culture et à renforcer les économies. Par ses activités, la Bertelsmann Stiftung veut encourager les citoyens à contribuer au bien commun. Fondée en 1977 par Reinhard Mohn, cette fondation à but non lucratif détient la majorité des actions de Bertelsmann SE & Co. KGaA. La Bertelsmann Stiftung est une fondation privée non partisane.

Avec son projet « Ethics of Algorithms », la Bertelsmann Stiftung examine de près les conséquences de la prise de décision algorithmique dans la société dans le but de s'assurer que ces systèmes sont utilisés au service de la société. L'objectif est d'aider à informer et à faire progresser les systèmes algorithmiques qui facilitent une plus grande inclusion sociale. Cela implique de s'engager pour ce qui est le mieux pour une société plutôt que pour ce qui est techniquement possible – afin que les décisions informées par les machines puissent servir au mieux l'humanité

https://www.bertelsmann-stiftung.de/en

BertelsmannStiftung