

A S S E M B L É E      N A T I O N A L E

X V <sup>e</sup>      L É G I S L A T U R E

# Compte rendu

## **Mission d'information sur l'émergence et l'évolution des différentes formes de racisme et les réponses à y apporter**

– Audition de M. Thomas Dautieu, directeur de la conformité de la Commission nationale de l'informatique et des libertés (CNIL), de Mme Marie Heuzé, juriste au service des affaires régaliennes et des collectivités territoriales, et de Mme Tiphaine Havel, conseillère pour les questions institutionnelles et parlementaires..... 2

Mardi

17 novembre 2020

Séance de 17 heures

Compte rendu n° 47

**SESSION ORDINAIRE DE 2020-2021**

**M. Robin Reda,**  
*Président*



*La mission d'information procède à l'audition de M. Thomas Dautieu, directeur de la conformité de la Commission nationale de l'informatique et des libertés (CNIL), de Mme Marie Heuzé, juriste au service des affaires régaliennes et des collectivités territoriales, et de Mme Tiphaine Havel, conseillère pour les questions institutionnelles et parlementaires.*

*La séance est ouverte à dix-sept heures.*

**M. le président Robin Reda.** Dans le cadre de la mission d'information de la Conférence des présidents sur l'émergence et l'évolution des différentes formes de racisme et les réponses à y apporter, nous continuons, avec la rapporteure, Mme Caroline Abadie, à organiser des auditions qui nous permettent d'entendre un certain nombre d'institutions et d'organismes publics qui œuvrent au quotidien contre le racisme.

Nous avons en particulier pu entendre le Conseil supérieur de l'audiovisuel (CSA), et nous recevons aujourd'hui la Commission nationale de l'informatique et des libertés (CNIL), qui constitue une autorité administrative indépendante ayant une place centrale dans notre démocratie. Ses travaux recouvrent des questions essentielles, en particulier dans un contexte de digitalisation croissante des échanges. Elle est représentée par M. Thomas Dautieu, directeur de la conformité, qui est accompagné de Mme Heuzé et Mme Havel.

M. Dautieu, dans le cadre de cette audition, vous pourrez rappeler le cadre juridique applicable au traitement des données dans notre pays, notamment lorsqu'elles révèlent une appartenance ou une origine ethno-raciale. La question de ces données a souvent été abordée lors des auditions que nous avons menées, non seulement avec des représentants des institutions, mais également des intellectuels, des élus et des associations de terrain. Vous pourrez également revenir, dans votre propos liminaire, sur le rôle et les enjeux de la CNIL, mais également la mission qui incombe à la société et aux institutions en matière de lutte contre toutes les formes de racisme et de discrimination.

**Mme Caroline Abadie, rapporteure.** M. Dautieu, nous vous remercions d'avoir accepté cette audition. Nous avons en effet régulièrement évoqué la CNIL lors des précédentes auditions. Nous avons parlé de statistiques, mais également de baromètres, d'indicateurs, des enquêtes Trajectoires et origines (TeO), et de manière générale de toutes les études qui visent à mesurer les discriminations. Votre présence a donc tout son sens, car elle permettra de discuter des missions de la CNIL, mais également, par exemple, du rapport coécrit par celle-ci et le Défenseur des droits, car nous avons déjà évoqué les questions de l'« enfermement algorithmique » et de la haine en ligne sur lesquelles il porte.

**M. Thomas Dautieu, directeur de la conformité de la CNIL.** Je vous remercie d'accueillir la CNIL pour cette audition. Celle-ci a pour mission de garantir que les fichiers informatiques mis en place tant dans le secteur public que dans le secteur privé respectent le cadre légal, fondé sur la loi du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, dite « loi informatique et libertés », et le règlement général sur la protection des données (règlement (UE) 2016/679 du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, dit règlement « RGPD »). Le premier principe posé par la réglementation (et qui intéresse plus particulièrement votre mission) est que les fichiers doivent être mis à œuvre à des fins licites et légitimes. À ce titre, la CNIL participe à la lutte contre le racisme, dans le sens où, par exemple, un fichier qui serait mis en place à des fins de discrimination en matière

d'accès à l'emploi ou au logement serait, par sa nature même, contraire aux principes de la réglementation.

La loi dite informatique et libertés et le RGPD apportent également des garanties vis-à-vis des données dites « sensibles ». Celles-ci incluent les données relatives à la santé, aux opinions politiques, à l'orientation sexuelle, mais également celles qui révèlent l'origine raciale ou ethnique. Les considérants du RGPD indiquent que « *l'utilisation de l'expression "origine raciale" dans le présent règlement n'implique pas que l'Union adhère à des théories tendant à établir l'existence de races humaines distinctes* ». Dans la loi française, il est question de « *prétendue origine raciale des personnes* ».

La loi dite informatique et libertés et le RGPD interdisent l'utilisation des données sensibles dans tout type de traitement informatique. Un traitement informatique qui contiendrait des données relatives à l'ethnie ou l'origine raciale supposée serait donc contraire à la loi française et au RGPD, et une disposition du code pénal prévoit jusqu'à 5 ans d'emprisonnement et 300 000 euros d'amende pour l'utilisation de ces données.

Le RGPD interdit également la prise de décision individuelle automatisée (autrement dit, l'utilisation d'algorithmes) qui exploiterait des données sensibles, sauf si les personnes concernées ont donné leur consentement (par exemple, il est possible d'organiser des sites de rencontre fondés sur des données sensibles), ou si un motif d'intérêt public important est engagé (par exemple, il est envisageable de développer des algorithmes qui utilisent des données de santé).

Ces dispositions fixent un cadre juridique, et la CNIL a pour mission d'instruire les plaintes qu'elle reçoit vis-à-vis de l'utilisation des données. Elle dispose d'un pouvoir de contrôle et de sanction, et peut imposer des amendes allant jusqu'à plusieurs dizaines de milliers d'euros et qui sont fonction de la gravité du manquement constaté. Il est à noter que, depuis la mise en place du RGPD, les mécanismes de déclaration auprès de la CNIL ont été abandonnés afin de recentrer les missions de celle-ci sur l'accompagnement des responsables de traitement ainsi que son rôle de contrôle et de sanction.

Je mettrai maintenant l'accent sur deux points. D'une part, il existe des exceptions à l'interdiction de traiter des données sensibles. Ces exceptions ne s'entendent qu'à partir du moment où le fichier concerné répond à une finalité déterminée, explicite et légitime.

La première de ces exceptions concerne les chercheurs (en dehors du domaine de la santé). Ainsi, il est possible, notamment dans le cadre de recherches en sciences sociales, de collecter des données sensibles, y compris celles relatives à la prétendue origine raciale ou ethnique. La finalité du traitement est ici légitime, mais la demande doit répondre à un certain nombre de critères, dont le premier est de disposer du consentement des personnes concernées, à moins que les données en question aient été rendues manifestement publiques par ces dernières, ou que la recherche présente un motif d'intérêt public au sens du code de la recherche (un avis motivé et publié de la CNIL est alors nécessaire). Dans ce dernier cas, cette exception ne vise que la recherche publique. La recherche privée doit disposer du consentement des personnes, ou les données relatives à ces dernières doivent être manifestement publiques.

La deuxième de ces exceptions concerne la santé. La CNIL peut autoriser la collecte de données relatives à l'ethnie, notamment dans le cadre de recherches médicales pour lesquelles il est scientifiquement prouvé que ces données sont importantes, ou dans le cadre

de la pharmacovigilance (des scientifiques peuvent démontrer à la CNIL qu'un médicament est plus efficace pour certaines ethnies).

La troisième des exceptions est le domaine statistique. La loi dite informatique et libertés prévoit une dérogation pour le traitement de données sensibles à des fins statistiques au bénéfice de l'Institut national de la statistique et des études économiques (INSEE) et des services ministériels statistiques, après avis du Conseil national de l'information statistique (CNIS). Les autres traitements statistiques doivent bénéficier d'un des régimes d'exception évoqués ci-dessus pour traiter des données sensibles.

Les statistiques ethniques sont interdites depuis la décision n° 2007-557 DC du 15 novembre 2007 du Conseil constitutionnel. Dans celle-ci, le Conseil indique que « *si les traitements nécessaires à la conduite d'études sur la mesure de la diversité des origines des personnes, de la discrimination et de l'intégration peuvent porter sur des données objectives, ils ne sauraient, sans méconnaître le principe énoncé par l'article 1<sup>er</sup> de la Constitution, reposer sur l'origine ethnique ou la race* ». Un certain nombre de commentateurs de cette décision estiment que le Conseil n'a pas jugé que seules des données objectives (comme la nationalité, qui ne constitue pas une donnée sensible) pouvaient faire l'objet de traitement, mais que cela pouvait également être le cas de données fondées sur un ressenti d'appartenance. Par conséquent, des traitements de recherche en sciences sociales peuvent être engagés sur des données à caractère objectif (nom, origine géographique, nationalité, etc.), mais également des données subjectives (par exemple, la réponse à la question « en pensant à votre histoire familiale, de quelle origine vous sentez-vous ? »), à la condition qu'ils ne visent pas la création d'un référentiel ethno-racial.

Ainsi, la CNIL a rendu en février 2020 un avis sur une enquête statistique de l'Institut national d'études démographiques (INED) portant sur les mutations sociodémographiques dans les départements et les régions d'outre-mer. La CNIL a observé que ce projet d'enquête visait à collecter des données objectives et ne reposait pas sur le régime ethnique ou prétendu racial des personnes interrogées. Certaines questions, telles que celles relatives aux expériences de discrimination, impliquaient des données subjectives, mais la CNIL a estimé que cette enquête n'avait pas pour objet, même indirectement, de classer les personnes interrogées en fonction soit de leur origine ethnique ou prétendument raciale déclarée, soit d'un référentiel ethno-racial.

En outre, la CNIL a publié en 2007, avant la décision du Conseil constitutionnel, un certain nombre de recommandations destinées à encadrer la mesure de la diversité. Elles sont antérieures au RGPD, mais certaines d'entre elles trouvent encore une pertinence. Elles portent notamment sur l'utilisation de données objectives dans les enquêtes, les études sur le ressenti, la possibilité d'admettre l'analyse des prénoms et des patronymes pour détecter d'éventuelles pratiques discriminatoires, le refus d'un référentiel ethno-racial et la nécessité de garantir la confidentialité et l'anonymat des données par le recours à des techniques d'anonymisation.

La quatrième exception qui permet le traitement des données ethniques ou prétendument raciales concerne les traitements à des fins mémorielles. La loi dite informatique et libertés et le RGPD prévoient des droits pour les personnes qui s'arrêtent au décès de celles-ci. Toutefois, dans certains cas de figures, la CNIL peut décider d'appliquer ces droits aux données relatives à des ancêtres de personnes vivantes. Ainsi, elle a pu se prononcer sur un certain nombre de publications de données, notamment en lien avec la Seconde Guerre mondiale, qui faisaient apparaître des données sensibles. Il a donné aux

personnes la possibilité de s'opposer à la diffusion des données relatives à leurs ancêtres si elles estimaient que cette diffusion pouvait leur causer un préjudice. De plus, cette diffusion devait s'entourer de garanties techniques (par exemple, les données ne devaient pas être collectées de manière automatisée par des moteurs de recherche).

La cinquième et dernière exception concerne le cas où le traitement des données sensibles est nécessaire pour des motifs d'intérêt public importants, après adoption d'un décret en Conseil d'État et avis de la CNIL. Je n'ai pas d'exemple à vous donner qui concerne les données ethniques.

D'autre part, la CNIL se mobilise depuis de nombreuses années en matière de lutte contre le racisme. Elle collabore avec le Défenseur des droits, qui fait partie du collège de 17 commissaires qui constitue la CNIL. Cette collaboration en matière de lutte contre le racisme s'est concrétisée en 2012 par la publication d'un guide à usage des acteurs de l'emploi (portant sur les mesures à mettre en œuvre pour mieux mesurer les éventuelles discriminations et pour progresser s'agissant de l'égalité des chances), en 2015 par la signature d'une convention de partenariat, et en 2020 par l'organisation d'un séminaire portant sur la prévention et la lutte contre les biais discriminatoires des algorithmes.

La CNIL a également engagé des travaux sectoriels sur la lutte contre les discriminations, portant par exemple sur l'accompagnement numérique des personnes en difficulté ou encore l'encadrement, au regard de la loi dite informatique et libertés et du RGPD, des dispositifs d'alerte professionnels (qui permettent notamment de dénoncer des faits de discrimination) et celui des fichiers relatifs à l'accès à la location (qui limite les données pouvant être collectées par les bailleurs, ce qui permet de s'assurer que ceux-ci ne disposent pas de données pouvant conduire à des discriminations).

**M. le président Robin Reda.** Pouvez-vous revenir plus en détail sur les procédures mises en place au regard des opérations de recrutement des employeurs, sachant que ceux-ci peuvent établir par nécessité certains fichiers susceptibles de provoquer des discriminations à l'embauche (je pense, par exemple, aux hôtesse de l'air et au personnel navigant) ? Comment dialoguez-vous avec ces entreprises, et comment, le cas échéant, intervenez-vous auprès d'elles ?

**M. Thomas Dautieu.** La mission de la CNIL est double. D'une part, elle doit expliquer aux entreprises la loi et le RGPD de manière concrète. Cette activité correspond à l'activité de production de droit souple de la CNIL. D'autre part, elle porte une mission d'enquête, sur la base notamment de plaintes. Ainsi, elle reçoit plus de 14 000 plaintes par an, dont une partie importante concerne le monde du travail.

Le recrutement est de plus en plus automatisé. Des fichiers sont donc mis en place. L'un des grands principes de la loi informatique et libertés est que les données collectées doivent être pertinentes. Les entreprises doivent donc uniquement collecter les données nécessaires à leurs opérations de recrutement. Une recommandation de la CNIL existe s'agissant de ces dernières. Toutefois, elle date de 2004, à une époque où le monde informatique était très différent, notamment s'agissant des algorithmes (qui aident de plus en plus au recrutement, sans pour autant remplacer la prise de décision humaine). Or, les biais algorithmiques, qui ne sont pas nécessairement volontaires, peuvent générer des discriminations. Par conséquent, la CNIL a lancé un groupe de travail sur l'application de la loi dite informatique et libertés aux algorithmes en matière de recrutement. Dans ce cadre, elle

procède à des auditions de concepteurs d'algorithmes, de syndicats, d'organismes de recrutement, etc., afin d'adapter cette recommandation de 2004.

**Mme Caroline Abadie, rapporteure.** Le cadre qui s'applique aux statistiques et à la recherche semble relativement souple. Je vous remercie pour ces précisions, car nos auditions nous avaient laissé l'impression qu'en dehors de l'enquête TeO de l'INED, les possibilités offertes pour observer les discriminations étaient limitées.

En ce qui concerne les algorithmes, le rapport rendu par la CNIL avec le Défenseur des droits indique que « *le plus souvent, c'est la combinaison de plusieurs critères neutres qui peut emporter des effets discriminatoires* ». Pourriez-vous présenter des exemples de ce principe qui ne concernent pas le recrutement ?

**M. Thomas Dautieu.** Dans le cadre de l'accès au logement, les bailleurs ont accès à des dossiers informatiques de plus en plus complets, même si la loi limite les données qui peuvent être collectées. En combinant l'ancienne adresse d'une personne candidate à un nouveau logement, sa nationalité et son niveau de revenu, il est possible de provoquer certaines discriminations.

Cependant, le réel danger est qu'un algorithme qui n'est pas programmé pour discriminer certaines personnes parvienne, par son fonctionnement autonome, à des formes de discrimination, en favorisant certains critères sur d'autres. Pour cette raison, la CNIL insiste sur la nécessité de vérifier régulièrement si les algorithmes remplissent bien la fonction pour laquelle ils ont été conçus, sans créer de biais.

**Mme Caroline Abadie, rapporteure.** Merci pour votre exemple, qui est très clair. L'orientation des lycéens repose aujourd'hui sur des algorithmes au travers de Parcoursup. Avez-vous été amenés à étudier le fonctionnement de ce dernier ?

**M. Thomas Dautieu.** L'affectation des lycéens se fonde sur un double mécanisme, au niveau national avec Parcoursup, et au niveau local des établissements d'enseignement supérieur, qui peuvent choisir ou non d'utiliser des algorithmes.

Nous n'avons pas audité l'ensemble des algorithmes mis en place par les établissements d'enseignement supérieur, mais nous avons été saisis de plaintes sur le sujet. Des parlementaires se sont également interrogés sur un éventuel désavantage des lycéens d'origine rurale par rapport à ceux habitant les centres-villes. La CNIL préconise (et cette préconisation a été confirmée par le Conseil d'État) de permettre à tous les lycéens d'accéder au code source de tous les algorithmes utilisés par les établissements d'enseignement supérieur. Celui de Parcoursup a été rendu public. Il peut donc faire l'objet d'un audit, et il est possible de vérifier quels critères sont utilisés, et quelle est leur pondération. Ainsi, comme je l'indiquais, la notion d'auditabilité des algorithmes est essentielle.

**Mme Caroline Abadie, rapporteure.** Cette notion est-elle pour autant suffisante ? Certains rapports indiquent que le lycée dans lequel le futur étudiant a obtenu son baccalauréat, combiné à un autre critère en apparence neutre, peut créer des discriminations.

**M. Thomas Dautieu.** L'auditabilité des algorithmes est indispensable. Il serait très risqué que les professionnels qui utilisent ces algorithmes ne soient pas en mesure de les remettre en cause. Le premier champ de l'utilisation de l'intelligence artificielle est celui de la santé. Elle sert notamment à détecter des pathologies. Il serait dangereux que les

professionnels de santé soient contraints par les résultats fournis par les algorithmes, sans pouvoir vérifier ce qui a conduit la machine à préconiser, par exemple, un traitement en particulier. La capacité des professionnels à questionner les algorithmes et les données me paraît essentielle, sans quoi un trop grand pouvoir sera donné à ces dispositifs.

**Mme Caroline Abadie, rapporteure.** Le champ d'investigation du racisme est particulièrement large. Au-delà du travail sur les préjugés, l'éducation et la mémoire, nous devons également explorer la question de la chaîne pénale dans nos travaux. Souhaitez-vous soulever des points que nous n'avons pas encore abordés et qui pourraient nous aider à construire notre rapport, par exemple sur les réseaux sociaux, la haine en ligne, le recrutement ou le logement ?

**M. Thomas Dautieu.** À titre personnel, je suis convaincu que l'un des principaux champs à creuser est celui de l'utilisation des données massives par les algorithmes. La CNIL travaille sur le sujet depuis quelques mois. Des études américaines ont montré que lorsque les algorithmes n'étaient pas régulièrement audités et calibrés, ils pouvaient dévier vers des formes de racisme. Ainsi, un *chatbot* développé par Microsoft s'est, après son lancement, rapidement mis à tenir des propos inacceptables.

Nous devons travailler sur la puissance de la machine, et la foi en celle-ci, d'autant qu'elle est alimentée par des données que nous produisons, même si l'exploration de ce champ prendra de nombreuses années. Votre mission doit s'attacher à cette question, qui est essentielle à l'ère numérique dans laquelle nous vivons.

**M. le président Robin Reda.** Je poserai pour terminer deux actions d'actualité. La première porte sur la reconnaissance faciale, qui est évoquée dans le cadre de la lutte contre le terrorisme, mais dont on imagine facilement les difficultés de mise en œuvre pratique et juridique au regard des préconisations de la CNIL sur la vidéoprotection, ainsi que les risques qu'elle peut faire peser, certains publics pouvant être plus surveillés que d'autres dans l'espace public. Cette question est directement liée à la lutte contre les discriminations et les racismes, au regard des débats déjà existants sur les contrôles policiers « au faciès ». Un groupe de travail de la CNIL réfléchit-il sur ce sujet ?

Ma seconde question concerne le domaine électoral. Les États-Unis, dont l'approche ouvertement communautariste est très différente de la nôtre, tendant à segmenter les campagnes électorales et les votes en fonction de l'appartenance ethnique. En France, cette tendance pourrait également constituer une tentation pour certains, surtout si de nouveaux mécanismes de vote (comme le vote par correspondance) venaient à entrer en jeu. Avez-vous observé des tentations, voire des tentatives d'exploiter des fichiers électoraux à des fins communautaristes ou ouvertement ethniques ?

**M. Thomas Dautieu.** Depuis quelques années, nous constatons un développement, non pas des dispositifs de reconnaissance faciale, mais d'une volonté de multiplier ceux-ci, pour des motifs en partie sécuritaires.

La CNIL s'intéresse à cette question, et a publié l'an dernier une note d'analyse et de positionnement sur la reconnaissance faciale. Celle-ci ne constitue pas une évolution de la vidéoprotection, mais marque une rupture, car elle utilise des données particulièrement sensibles, les données biométriques. En effet, s'il est possible de changer de carte bleue, il est impossible de refaire son visage. Ces données sont donc extrêmement personnelles.

Les systèmes de reconnaissance faciale ne peuvent alors être mis en place que pour un motif extrêmement important, et avec toutes les garanties nécessaires en matière de protection des données. Nous sommes de plus en plus régulièrement saisis par des collectivités locales et des gestionnaires de transports publics sur des questions liées à des caméras intelligentes ou de systèmes de reconnaissance faciale. La CNIL apporte une attention toute particulière au sujet, et a appelé à de nombreuses reprises à engager un débat au Parlement.

S'agissant des opérations électorales, vous avez raison. La CNIL constate le développement de logiciels aidant à réaliser de la prospection électorale, notamment par voie électronique. Elle a produit un certain nombre de documents sur ce thème, mais n'a pas observé de tentative d'utiliser des données à des fins de segmentation ethnique, raciale ou religieuse du vote. À ma connaissance, elle n'a pas été saisie de plaintes ni n'a reçu d'information en ce sens, mais le numérique est en train de pénétrer la sphère de la propagande électorale, les données étant de plus en plus accessibles aux partis politiques et aux candidats.

**M. le président Robin Reda.** Nous vous remercions.

*La séance est levée à dix-sept heures quarante.*

