

**Rapport présenté en séance plénière  
le 15 mai 2007**

## **Mesure de la diversité et protection des données personnelles**

Madame Anne Debet

Rapporteur

**Avec le concours de :**

Ines Rogic, Chargée d'instruction

**DOC Projet Diversité Mise à jour : 16/05/2007 08:40 DECO / IRC**

# SOMMAIRE

<b>I. INTRODUCTION</b> .....	<b>4</b>
<b>II. POURQUOI LA CNIL EST-ELLE CONCERNEE : L'IMPORTANCE DES ENJEUX EN MATIERE DE LA PROTECTION DES DONNEES A CARACTERE PERSONNEL</b> .....	<b>5</b>
<b>A. Parce que la mesure de la diversité peut nécessiter le recueil et le traitement de données à caractère personnel</b> .....	<b>6</b>
1. L'utilisation des fichiers de gestion .....	6
2. Les enquêtes par questionnaires .....	8
<b>B. Parce qu'il existe un cadre juridique à respecter</b> .....	<b>8</b>
1. Le respect des conditions générales fixées par la loi informatique et libertés .....	8
2. Le respect des conditions particulières de traitement des données sensibles .....	9
a) La notion de données sensibles .....	9
b) Le régime de protection applicable aux données sensibles .....	12
<b>C. Parce que la mesure de la diversité touche à l'identité humaine et aux libertés individuelles</b> .....	<b>15</b>
<b>III. DES MALENTENDUS A DISSIPER</b> .....	<b>16</b>
1. La CNIL ne réaliserait pas l'importance, pour les entreprises, de mettre en place des outils de lutte contre la discrimination. ....	17
2. La CNIL n'autoriserait pas l'utilisation des fichiers de gestion du personnel pour permettre aux entreprises de mesurer la diversité de leurs effectifs et l'existence éventuelle de discriminations, notamment dans le recrutement .....	17
3. Alors que des pays démocratiques ont déjà un référentiel national « ethno-racial », en France, la CNIL s'y opposerait. ....	18
4. La CNIL représenterait un frein à la recherche. Les chercheurs ne disposeraient pas aujourd'hui de suffisamment d'informations pour pouvoir mesurer la diversité de la société française .....	19
<b>IV. PROPOSITIONS DE LA CNIL</b> .....	<b>24</b>
<b>A. Propositions en l'état actuel de droit</b> .....	<b>24</b>
<b>Recommandation n°1 : ouvrir plus largement aux chercheurs l'accès aux bases de données statistiques et aux fichiers de gestion</b> .....	<b>24</b>
<b>Recommandation n°2 : mesurer la diversité en utilisant les données « objectives » relatives à l'ascendance des personnes (nationalité et/ou lieu de naissance des parents)</b> .....	<b>26</b>
<b>Recommandation n°3 : permettre le développement d'études sur le « ressenti » des discriminations incluant le recueil de données sur l'apparence physique</b> .....	<b>29</b>

<b>Recommandation n°4 : permettre, sous certaines conditions, l'analyse des prénoms et des noms de famille .....</b>	<b>31</b>
<b>Recommandation n°5 : développer le recours à des experts tiers de confiance .....</b>	<b>32</b>
<b>Recommandation n° 6 : garantir la confidentialité et l'anonymat .....</b>	<b>34</b>
<b>Recommandation n°7 : Assurer la transparence par l'information et l'exercice effectif des droits « Informatique et Libertés » .....</b>	<b>35</b>
<b>B. Faut-il modifier la loi ? .....</b>	<b>35</b>
<b>Recommandation n°8 : Assouplir, dans le domaine de la recherche, la dérogation relative au « consentement exprès » (art. 8-II) mais prévoir en contre partie l'autorisation de la CNIL.....</b>	<b>36</b>
<b>Recommandation n°9: Faut-il créer par la loi un référentiel national « ethno-racial » ? .....</b>	<b>37</b>

## I. Introduction

La lutte contre les discriminations et en particulier celles attachées aux origines nationales, ethniques ou raciales des personnes est une priorité pour tous.

Comment s'assurer que l'égalité à la française, celle des articles 1<sup>er</sup> de la déclaration des droits de l'Homme de 1789 et de la Constitution de la Ve République, devienne une réalité ? Comment promouvoir l'égalité des chances et l'égalité des droits, garantir la non-discrimination dans une société française de plus en plus composite ?

La Commission Nationale de l'Informatique et des Libertés comme toute autre institution publique, contribue, dans le champ de compétences qui est le sien, aux actions menées en ce sens. Aussi, doit-elle intervenir dans la définition et la mise en place des outils statistiques de mesure des discriminations et prendre part au débat .

Car, pour lutter contre les discriminations, encore faut-il pouvoir les identifier. Pour les identifier, il paraît naturel de procéder à l'observation statistique des différences, de la diversité. Mais quelle diversité mesurer ? La diversité sociale, « ethnique », religieuse, culturelle... ? Quels critères utiliser pour analyser cette, ces diversité(s) ? Quelles méthodes employer ? Qui peut le faire ? Faut-il d'ailleurs parler de diversité ou, comme le suggère le directeur de l'INED, de données de trajectoires nationales ?

Cette problématique est complexe et délicate.

Complexe, car comme la CNIL a pu le constater, cette « mesure de la diversité » suscite une véritable « effervescence méthodologique <sup>1</sup>», les chercheurs et statisticiens faisant preuve d'une grande imagination en ce domaine et n'étant pas forcément tous d'accord sur les outils à employer.

Délicate, car elle touche à l'essence même de ce qui fait notre identité, à la façon dont on se perçoit et dont on est perçu par les autres et peut contredire le modèle français, égalitaire et universaliste, présenter des risques d'utilisation détournée, étant observé que la notion de race n'a aucune valeur scientifique.

Compte tenu des multiples enjeux soulevés par cette question, le Président de la CNIL, approuvé par ses collègues, a souhaité constituer, en 2005 un groupe de travail, présidé par Madame Anne Debet, professeur des universités et membre de la CNIL.

A la suite des travaux et auditions menés par ce groupe, notre Commission a émis, en juillet 2005, ses recommandations en matière de mesure de la diversité et de lutte contre les discriminations dans le domaine de l'emploi. Or, les positions exprimées par la CNIL ont été parfois mal comprises, certains reprochant même à la Commission de s'opposer au développement des études et enquêtes portant sur ces sujets. En outre, il est apparu que ce sujet suscitait de plus en plus de débats et controverses étant entendu que les employeurs s'accordent pour considérer que les capacités professionnelles constituent le facteur primordial, dans l'activité productive contemporaine, pour évaluer un candidat à l'embauche comme un salarié. La CNIL a donc souhaité approfondir sa réflexion en recueillant le point de vue des différents acteurs concernés : chercheurs, statisticiens, organisations syndicales, représentants des religions, mouvements associatifs, personnalités qualifiées ...

---

<sup>1</sup> Pour reprendre l'expression de P.Simon, chercheur à l'INED

Le groupe de travail a ainsi réalisé plus de soixante auditions entre le 2 novembre 2006 et le 1er février 2007. Parmi ces auditions, celles des 18 et 25 janvier 2007 ont été ouvertes à la presse. C'était une première. Ainsi, les journalistes des médias écrits et audio-visuels, rassemblés dans la salle Médicis du Palais de Luxembourg, ont pu assister aux échanges de points de vue de la Commission avec les personnalités suivantes : MM. Azouz BEGAG, Ministre pour la promotion de l'égalité des chances, Claude BEBEAR, Président de l'Institut Montaigne, Président du Conseil de Surveillance d'AXA et Président du Conseil d'Administration d'IMS-Entreprendre pour la cité, Louis SCHWEITZER, Président de la HALDE (Haute Autorité de Lutte contre les Discriminations et pour l'Egalité), Alain BAUER, criminologue et Président du Conseil d'Orientation de l'Observatoire National de la Délinquance, Richard DESCOINGS, Directeur de l'Institut des Etudes Politiques de Paris, François HERAN, Directeur Général de l'INED (Institut National d'Etudes Démographiques, Dogad DOGOU, président de l'association Africagora).

La Commission remercie très vivement l'ensemble des personnes auditionnées qui ont par leurs contributions, tant orales qu'écrites, et par la formulation de propositions constructives, nourri et enrichi sa réflexion .

Afin d'élargir le débat, la CNIL a souhaité aussi recueillir le point de vue de simples citoyens et a mis en ligne sur son site un questionnaire sur le thème de la mesure de la diversité. Plus de 1000 internautes ont ainsi pu exprimer leurs opinions sur le sujet. Bien entendu, ce « sondage » n'a aucune valeur représentative et n'a d'autre objet que de permettre à chacun d'émettre un avis sur cette question délicate. Les résultats de ce « sondage » figurent en annexe de ce rapport.

A l'issue de ces travaux, la Commission estime aujourd'hui nécessaire de faire part publiquement de sa position et de proposer des pistes de réflexion constructives qui permettront, elle l'espère, de faire progresser la connaissance et, par là même, de mieux lutter contre les discriminations. Il s'agit aussi pour la Commission de mieux expliquer sa position, de dissiper les malentendus qui ont pu résulter d'une mauvaise compréhension de ses positions sur le sujet.

Ces recommandations s'adressent donc à l'ensemble des acteurs concernés : aux décideurs politiques, chefs d'entreprises, syndicats, administrations afin de les guider dans la lutte contre les discriminations, mais aussi aux chercheurs et statisticiens afin de leur faire comprendre que la CNIL est ouverte à leurs préoccupations et reconnaît pleinement la légitimité de leur démarche.

Enfin, et surtout, ces recommandations seront rendues publiques afin que chacun, Français comme étranger, comprenne bien les enjeux, connaisse mieux ses droits comme ses devoirs et concoure ainsi à une meilleure protection des données personnelles .

## **II. POURQUOI LA CNIL EST-ELLE CONCERNEE : L'IMPORTANCE DES ENJEUX EN MATIERE DE LA PROTECTION DES DONNEES A CARACTERE PERSONNEL**

Certains, peu nombreux il est vrai, comprennent mal que la CNIL ait son mot à dire dans ce grand débat national relatif à la mesure de la diversité ou ce qu'on appelle communément les « statistiques ethniques ».

Ce malentendu doit être dissipé. Pourquoi la CNIL intervient-elle sur ce sujet ?

## **A. Parce que la mesure de la diversité peut nécessiter le recueil et le traitement de données à caractère personnel**

L'analyse de la diversité et des discriminations peut nécessiter selon l'objectif de recherche poursuivi, le recueil, le traitement informatique et la conservation pendant un temps donné d'informations sur les personnes. Dès lors, les fichiers de recherche ainsi constitués sont soumis au contrôle de la CNIL.

Autorité administrative indépendante, la CNIL agit dans un cadre précis, celui de la loi « informatique et libertés » du 6 janvier 1978 modifiée en 2004, qui fixe son périmètre de compétences et ses missions.

Garante de la protection des données à caractère personnel et de la vie privée, la CNIL a ainsi pour mission de veiller à la bonne application de la loi dans tous les domaines où des données à caractère personnel sont susceptibles d'être utilisées. Le champ des recherches en sciences humaines n'échappe donc pas à la compétence de la CNIL.

Comme pour toute recherche en sciences humaines, les données sont collectées soit directement auprès des personnes, par voie de questionnaires, par téléphone, ou encore oralement en face à face, soit à partir de l'exploitation de fichiers de gestion ou encore de bases statistiques existantes.

Les données recueillies peuvent permettre d'identifier directement la personne (ex :son nom). Elles peuvent aussi être indirectement nominatives soit parce qu'elles sont indexées par un numéro de correspondance avec l'identité de la personne, soit par corrélation. Ainsi le fait d'enregistrer sur une personne des données telles que :

- son prénom,
- sa nationalité,
- son pays de naissance,
- la nationalité ou le pays de naissance de ses parents ou de ses grands-parents,
- son adresse
- son poste professionnel

est susceptible de l'identifier.

### **1. L'utilisation des fichiers de gestion**

Deux cas de figure peuvent se présenter.

#### **a) L'exploitation des fichiers existants**

L'entreprise, l'université ou l'administration souhaitent mieux connaître la composition de leurs effectifs, faire un bilan des parcours universitaires, des pratiques de gestion des ressources humaines ainsi que des actions « anti-discrimination » mises en place. Elles exploitent alors les données figurant déjà dans les fichiers de gestion : nom de l'étudiant, du candidat ou de l'employé, prénom, nationalité, lieu de naissance, adresse, diplômes, parcours professionnel....

Les fichiers de gestion des ressources humaines constituent ainsi la source « naturelle » des données utilisées pour la constitution de panels d'employés ou pour la réalisation d'analyses patronymiques<sup>2</sup> (Casino, France Télévisions<sup>3</sup>).

Considérée par certains comme un expédient ponctuel, voire comme du « bricolage »<sup>4</sup>, cette méthode revêt, en revanche, pour d'autres chercheurs<sup>5</sup>, un réel intérêt s'agissant notamment de l'analyse des prénoms. Ainsi, pour JF Amadieu<sup>6</sup>, elle « *permet d'observer un vaste ensemble de discriminations avec une seule information. En effet le prénom permet de recueillir les informations sur l'origine sociale d'un individu, sur l'origine géographique et sur une appartenance religieuse supposée. De surcroît, le prénom est par lui-même une variable susceptible de provoquer une discrimination ou une inégalité de chances* ».

Comme le souligne le directeur de l'INED, « *une distinction s'impose. On peut prendre directement pour objet d'étude les patronymes ou les prénoms à consonance étrangère et les utiliser simplement comme repère pour étudier le sort des populations issues de l'immigration, quand il n'existe pas d'autre source disponible* ». La première approche paraît légitime, par exemple, pour étudier les effets d'un patronyme ou d'un prénom sur les chances d'accès à l'emploi (objectif poursuivi par les opérations de testing), il s'agit alors d'explorer les facteurs de discrimination. La deuxième approche, approximative, ne peut constituer qu'un pis-aller en l'absence d'autres sources d'information disponibles.

La méthode est utilisée principalement par les organismes souhaitant détecter le niveau de discrimination en leur sein, soit lors du recrutement, soit lors de l'attribution d'un logement, soit encore dans les parcours scolaires, universitaires et professionnels. Le plus souvent, c'est l'organisme lui-même qui mène l'analyse mais il peut également faire appel à un sous-traitant, généralement un cabinet d'études ou de conseil.

Par ailleurs, cette méthode est utilisée dans les analyses des Zones Urbaines Sensibles (ZUS) et les Zones d'Education Prioritaires (ZEP). Engagées par les collectivités locales, associations, organismes publics, ces analyses s'inscrivent dans l'objectif d'ajustement des politiques engagées en faveur des populations de ces zones à travers la mesure de leur impact (relogements, renforcement de la mixité sociale, lutte contre le chômage, réussite scolaire,...).

Elle a également été utilisée par les services des renseignements généraux pour analyser, dans le cadre d'une étude multicritères, le profil des meneurs de violences urbaines.

Deux sortes de données sont utilisées :

---

<sup>2</sup> A cet égard, ainsi que le groupe de travail l'a constaté lors des auditions, la méthode onomastique tend, semble-t-il à se développer car ces données sont le plus facilement accessibles et leur exploitation demande relativement peu de moyens, tant humains que financiers.

<sup>3</sup> L'enquête CASINO a été réalisée auprès des personnels de 4 supermarchés (St Martin d'Hères (38), Vaise (69), Valence (26), ST Etienne (42) ; l'enquête France Télévisions a été réalisée auprès des personnels de France Télévisions en 2006.

<sup>4</sup> F.Héran, P.Simon de l'INED

<sup>5</sup> JF Amadieu, MM Cediey et Foroni (ISM CORUM)

<sup>6</sup> Directeur de l'Observatoire des Discriminations

- Les prénoms / noms, sont extraits du fichier de gestion et classés dans des catégories de type « ethnique » définies par l'organisme lui-même ou son prestataire ;
- Les prénoms uniquement sont extraits du fichier de gestion et ensuite sont classés dans des catégories « discriminant » / « non discriminant »

Or, le recours à aucune des deux n'est scientifique et ne donne de résultat fiable pour la mesure des discriminations. Elles permettent seulement à l'organisme intéressé d'obtenir des informations relatives à l'existence ou non des pratiques discriminatoires.

b) L'enrichissement des fichiers par des indicateurs nouveaux, en particulier sur la nationalité et le lieu de naissance des parents.

Si certains l'envisagent<sup>7</sup>, la Commission n'a été saisie à ce jour d'aucune projet en ce sens.

## 2. Les enquêtes par questionnaires

Les questionnaires intègrent parfois des identifiants (numéros) afin, précisément, de permettre le suivi des réponses données par une même personne à des moments différents (« suivi de trajectoires »). Ils peuvent aussi comporter des indications permettant le repérage des personnes, surtout si l'échantillon concerné est restreint : par exemple, l'indication du poste professionnel (ex : la responsable communication de l'entreprise), du lieu de résidence (ex : commune rurale d'une centaine d'habitants...)

## B. Parce qu'il existe un cadre juridique à respecter

### 1. Le respect des conditions générales fixées par la loi informatique et libertés

Les outils de mesure de la diversité peuvent reposer sur la collecte et le traitement de données permettant, directement ou indirectement, l'identification, même momentanée, des personnes concernées. A ce titre, ils sont soumis à la loi informatique et libertés et doivent donc respecter l'ensemble des principes de protection des données.

Il existe cinq « règles d'or » de protection des données personnelles qui doivent être prises en compte lors de l'élaboration de tout projet de recherche susceptible de faire appel à des données personnelles *a fortiori* lorsqu'elles portent sur des données sensibles (voire infra 2). C'est aussi à la lumière de ces principes clés que la Commission apprécie les projets de recherche et d'enquêtes qui lui sont soumis.

- **le principe de finalité** : les données doivent être recueillies pour des finalités déterminées et légitimes. Par exemple, est-il ou non légitime de vouloir recueillir des données sur les origines raciales des personnes pour analyser, dans une entreprise, d'éventuelles discriminations à l'embauche ?

- **le principe de proportionnalité** : les données collectées doivent être adéquates, pertinentes et non excessives au regard de ces finalités. Ainsi, est-il ou non pertinent de trier

<sup>7</sup> Par exemple, F.Héran (INED)



et de procéder à un traitement des patronymes pour déterminer l'origine ethnique d'une personne ?

- **le caractère limité de la durée de conservation** des données dans un fichier de recherche qui doit être proportionnée aux objectifs poursuivis. L'analyse, à partir du fichier du personnel d'une entreprise, de l'historique des parcours professionnels des salariés, justifie-t-elle que les informations ainsi recueillies soient conservées sous une forme nominative pendant dix ans ?

- **La sécurité des données** : toutes mesures doivent être prises pour assurer la confidentialité des données et éviter leur divulgation . En cas d'enquête par questionnaires est-il prévu, dans le traitement informatique, une séparation entre les coordonnées des personnes et les réponses confidentielles ?

- **le respect des droits des personnes** : tout recueil de données sur des personnes, que ce soit par voie de questionnaires ou par exploitation de fichiers, impose que celles-ci soient informées des conditions d'utilisation de ces données, de leur droit d'obtenir communication de celles-ci, de demander leur rectification, voire leur suppression si elles sont inexactes, et sous certaines conditions de s'opposer à leur traitement. Ainsi, le questionnaire individuel diffusé auprès des salariés d'une entreprise pour connaître leur parcours professionnel, leurs origines sociales, géographiques et leur perception d'une éventuelle discrimination comporte-t-il ou est-il accompagné d'une information claire sur les objectifs de l'enquête, son caractère facultatif, les destinataires des questionnaires, les conditions dans lesquelles les salariés peuvent exercer leur droit d'accès aux informations les concernant ?

## 2. Le respect des conditions particulières de traitement des données sensibles

A l'occasion des nombreux débats organisés ces derniers mois sur le thème de la mesure de la diversité, auxquels la CNIL a pu participer, une question revenait fréquemment sur les « données sensibles » : « Pourquoi la CNIL considère-t-elle que l'origine représente une donnée sensible ? »

Pour répondre à cette question, il convient de définir ce que l'on entend par donnée sensible.

### a) La notion de données sensibles

#### Que dit la loi ?

Au plan européen, l'article 6 de la Convention 108 du Conseil de l'Europe comme l'article 8 de la Directive européenne 95/46 prévoient que les données sensibles relatives à l'origine « raciale [ou ethnique]<sup>8</sup>, les opinions politiques, les convictions religieuses ou autres convictions (...) » ne peuvent être traitées à « moins que le droit interne ne prévoie des garanties appropriées ».<sup>9</sup>

---

<sup>8</sup> Article 8 de la Directive 95/46/CE du Parlement européen et du Conseil du 24 octobre 1995 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données : « 1. Les Etats membres interdisent le traitement des données à caractère personnel qui révèlent l'origine raciale ou ethnique, les opinions politiques, les convictions religieuses ou philosophiques, l'appartenance syndicale, ainsi que le traitement des données relatives à la santé et à la vie sexuelle. »

La loi « informatique et libertés » du 6 août 2004, modifiant la loi du 6 janvier 1978, assure la transposition de la Directive 95/46 dans le droit français.

<sup>9</sup> Article 6 de la Convention 108 du 28 janvier 1981 du Conseil de l'Europe sur la protection des personnes à l'égard du traitement automatisé des données à caractère personnel

Au sens de l'article 8 de la loi du 6 janvier 1978 modifiée en 2004 sont considérées comme sensibles et faisant donc l'objet d'une protection particulière les « *données à caractère personnel qui font apparaître, directement ou indirectement, les origines raciales ou ethniques, les opinions politiques, philosophiques ou religieuses ou l'appartenance syndicale des personnes, ou qui sont relatives à la santé ou à la vie sexuelle de celles-ci*<sup>10</sup> ».

**Quelles sont les données que la CNIL considère comme relevant de l'article 8 de la loi ?**

**- Les données sur les origines raciales ou ethniques :**

Si l'on se réfère aux projets de recherche ou de traitements statistiques qui ont été soumis à la CNIL, il doit être relevé que les données à caractère personnel qui font apparaître directement les origines raciales ou ethniques des personnes sont rarement enregistrées dans les fichiers<sup>11</sup>.

- Les recensements de la population dans les Territoires d'Outre-mer :

Les recensements réalisés en 1983 dans les TOM, en 1989 en Nouvelle-Calédonie, en 1991 et en 2002 à Mayotte comportaient, compte tenu des spécificités des territoires concernés, des questions sur la communauté d'appartenance, la tribu d'appartenance, le statut civil coutumier (ou encore dans le cas de Mayotte, sur la polygamie). Ces recueils de données ont tous été autorisés par décret pris après avis favorable de la CNIL, compte tenu de l'intérêt public présenté.

Le recensement de la population qui devait être réalisé en 2003 en Nouvelle-Calédonie et qui, comme tous les recensements précédents, comportait des questions sur la communauté d'appartenance et la tribu d'origine a été reporté, le Président de la République ayant contesté la pertinence desdites questions. Le recensement qui a finalement eu lieu en 2004 n'a pas intégré ces questions. La mise en œuvre d'une enquête statistique spécifique à la Nouvelle-Calédonie a été décidée par le Congrès de Nouvelle-Calédonie. Cette enquête qui visait à connaître l'appartenance ethnique des personnes a été autorisée par la CNIL le 30 juin 2005.

- Dans certaines recherches médicales, le recueil de telles données permet de déterminer ou de confirmer le cas échéant une corrélation avec la pathologie étudiée<sup>12</sup>. La CNIL et préalablement le comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé<sup>13</sup>, apprécie alors au cas par cas, la pertinence du

---

<sup>10</sup> Il convient de noter que cette énumération des données sensibles reprend celle des critères discriminatoires figurant dans les textes précités, à l'exception du sexe et de l'âge qui sont considérés comme des données d'état civil.

<sup>11</sup> Il doit toutefois être relevé que le développement de la photographie numérique facilite de fait tout type de traitement possible de données sur les apparences physiques des personnes.

<sup>12</sup> Par exemple, la drépanocytose, maladie génétique due à une anomalie de l'hémoglobine qui touche essentiellement l'Afrique intertropicale, certaines régions des Indes, les Antilles, l'Amérique du Sud, les Noirs américains.

<sup>13</sup> Prévu par l'article 54 de la loi du 6 janvier 1978 modifiée, ce comité d'experts, institué auprès du ministère de la recherche, doit pour tout projet de fichier de recherche médicale, émettre un avis sur la méthodologie de la recherche, la nécessité du recours aux données nominatives et la pertinence de celles-ci.

recueil de ces données, en fonction de la finalité de la recherche et de l'argumentaire scientifique présenté.

Il doit être relevé qu'en ce domaine, en l'absence d'un référentiel scientifique reconnu, la question du type de classification retenue est souvent discutée : certaines recherches reposent sur des protocoles internationaux qui utilisent des classifications américaines sujettes à discussion (retenant notamment le type « caucasien » pour caractériser la population blanche). La CNIL a pour sa part longtemps promu une classification des régions d'origine établie par l'OMS mais qui ne semble plus réellement en vigueur aujourd'hui<sup>14</sup>.

**- Les données sur les caractéristiques physiques susceptibles de révéler l'origine raciale ou ethnique:**

Certains fichiers peuvent comporter des données sur les caractéristiques physiques des personnes et en particulier sur leur couleur de peau.

- Tel est le cas des fichiers des renseignements généraux<sup>15</sup> et des fichiers de police judiciaire<sup>16</sup> utilisés en particulier pour les recherches criminelles : ces fichiers compte tenu de l'intérêt public qu'ils présentent ont été autorisés par décret en Conseil d'Etat pris après avis de la CNIL à enregistrer ce type d'informations.

- Dans le cadre d'opérations de recrutement à des emplois reposant sur l'apparence physique (ex mannequins, hôtesse d'accueil, comédiens...) : les fichiers peuvent comporter des données faisant apparaître directement ou indirectement l'origine raciale ou ethnique des personnes. Leur accord exprès est alors recueilli préalablement à l'enregistrement de l'information.

- Certaines recherches médicales (par exemple dans le domaine de la dermatologie) peuvent aussi nécessiter la collecte de données relevant de cette catégorie.

---

<sup>14</sup> Il doit être relevé que la Commission a été saisie dans le passé de plaintes du MRAP et de SOS RACISME sur l'utilisation par les pneumologues d'un spiromètre destiné à mesurer la capacité respiratoire et recueillant l'origine ethnique des patients (notamment sous la forme caucasien, asiatique, africain...). Après explications obtenues notamment auprès de la société française de pneumologie et du constructeur des appareils, il s'avère que cette classification, très sommaire, reposerait sur une caractéristique objective, à savoir que la capacité respiratoire d'une personne originaire d'Asie serait moindre que celle d'un « caucasien », elle-même moindre que celle d'un « africain » ... Il semble que certains logiciels utilisés pour des examens radiographiques comportent aussi des rubriques sur l'origine ethnique...

<sup>15</sup> le décret du 14 octobre 1991 prévoit l'enregistrement des « signes physiques particuliers, objectifs et inaltérables » comme éléments de signalement dans les seuls cas de personnes qui peuvent en raison de leur activité individuelle ou collective porter atteinte à la sûreté de l'Etat ou à la sécurité publique par le recours actif apporté à la violence ainsi que les personnes entretenant des relations directes avec celles-ci .

<sup>16</sup> ainsi le système de traitement des infractions constatées (STIC) mis en œuvre par le ministère de l'intérieur prévoit l'enregistrement de données « sensibles » dans les seuls cas où ces informations résultent de la nature ou des circonstances de l'infraction ou se rapportent à des signes physiques particuliers objectifs et permanents en tant qu'éléments de signalement des personnes dès lors que ces éléments sont nécessaires à la recherche et à l'identification des auteurs d'infractions. La classification utilisée a d'ailleurs récemment été remise en cause dans le cadre du rapport *Fichiers de police et de gendarmerie, Comment améliorer leur contrôle et leur gestion ?*, remis fin 2006 au ministre de l'intérieur par M. Alain Bauer, président du conseil d'orientation de l'observatoire national de la délinquance (La documentation française, Paris, 2007.)

## - L'analyse des patronymes :

La CNIL estime qu'un traitement opéré sur la base de la consonance du nom ou du prénom d'une personne ou encore certains traitements multicritères relèvent de l'application de l'article 8 de la loi du 6 janvier 1978 modifiée .

Cette position, exprimée dans une recommandation du 3 décembre 1996 sur l'utilisation des fichiers à des fins politiques a récemment été confirmée par la CNIL, d'une part lors de l'adoption le 9 juillet 2005 de ses premières recommandations sur la mesure de la diversité dans le domaine de l'emploi, d'autre part lors de sa délibération portant refus d'autorisation d'une enquête téléphonique du CRIF. Cette enquête reposait sur la constitution d'un échantillon de sondage à partir d'un tri sur la consonance du nom des intéressés censé faire apparaître leurs appartenances à la communauté juive<sup>17</sup>.

- Enfin, la Commission a parfois considéré que certains traitements à finalité statistique étaient susceptibles, notamment du fait de la nature des informations recueillies, de révéler indirectement l'origine raciale ou ethnique ou la pratique religieuse des personnes.

Tel fut le cas par exemple de l'enquête sur la mobilité géographique et l'insertion sociale de l'INED réalisée en 1992 pour mieux connaître la vie des immigrés et de leurs enfants en France. Au titre des données recueillies, figuraient notamment l'histoire migratoire, la nationalité, la vie matrimoniale (dont la polygamie), la fécondité et la contraception, l'alphabétisation et la maîtrise du français, la pratique religieuse. La CNIL a estimé que certaines données « sensibles » au sens de la loi de 1978 nécessitaient le recueil préalable du consentement des intéressés.

### **Quelles sont les données que la CNIL considère comme ne relevant pas de l'article 8 de la loi ?**

L'adresse, la nationalité et le lieu de naissance ne sont pas considérés par la CNIL comme des données « sensibles » au sens de l'article 8.

En effet, l'information sur le lieu de naissance de la personne fait partie de l'état civil et est considérée comme une donnée « objective ».

La Commission porte cependant une attention particulière au traitement des données relatives à la nationalité et au lieu de naissance dans les fichiers, la pertinence de leur collecte devant être dûment justifiée, au cas par cas, par le responsable du traitement .

### **b) Le régime de protection applicable aux données sensibles**

L'article 8 de la loi du 6 janvier 1978 modifiée interdit le traitement des données sensibles et notamment **"les données qui font apparaître directement ou indirectement les origines**

---

<sup>17</sup> délibération du 2 février 2006

**raciales ou ethniques (...)**, sauf exceptions prévues par la loi. Le non respect de cette interdiction est passible des sanctions pénales prévues au titre de l'article 226-19 du code pénal (5 ans d'emprisonnement et 300 000 euros d'amende). Une telle interdiction traduit la volonté du législateur, tant national qu'européen, de soumettre à un encadrement juridique strict l'enregistrement, dans des fichiers, de données dont une utilisation incontrôlée serait susceptible de porter gravement atteinte aux droits et libertés des personnes, en particulier en raison des risques induits de stigmatisation, de discrimination et d'exclusion des personnes.<sup>18</sup>

La loi prévoit toutefois dix exceptions au principe de l'interdiction de collecte des données sensibles<sup>19</sup>. **Toutes ne sont cependant pas applicables aux traitements de données sensibles mis en œuvre dans le cadre d'études de mesure de la diversité.** En effet, de tels traitements ne sont possibles, sous certaines conditions, que dans les hypothèses suivantes : recueil du consentement exprès de la personne, anonymisation des données, traitements mis en œuvre pour des motifs d'intérêt public.

#### **- Le consentement exprès de l'intéressé (article 8-II,1°)**

Le terme « consentement exprès » signifie un accord explicite et écrit au sens de la jurisprudence du Conseil d'Etat<sup>20</sup>. L'article 8 précise que le consentement de l'intéressé peut permettre de traiter des données sensibles sauf dans le cas où la loi prévoit que l'interdiction de traiter ces données ne peut être levée par le consentement.

A cet égard, le législateur a pu poser des limites au traitement de certaines données comme les données génétiques par les assureurs. Ainsi les articles L 1141-1 à L1141-3 du Code de la Santé publique interdisent aux compagnies d'assurance de prendre en compte les résultats des tests génétiques même si ceux-ci leur ont été transmis par la personne concernée ou avec son accord.

Comme la CNIL l'a souligné en 2005, le recours au consentement soulève des difficultés d'ordre pratique et juridique.

---

<sup>18</sup> Les catégories de "données sensibles" ont toujours fait l'objet d'une protection particulière. Ainsi, l'article 31 de la loi du 6 janvier 1978 dans sa version antérieure à la loi du 6 août 2004, disposait « qu' il est interdit de mettre ou de conserver en mémoire informatisée, sauf accord exprès de l'intéressé, des données nominatives qui, directement ou indirectement, font apparaître les origines raciales ou les opinions politiques, philosophiques ou religieuses ou les appartenances syndicales ou les mœurs de personnes. »

L'article 8 de la loi du 6 janvier 1978, modifiée par la loi du 6 août 2004 reprend quasiment à l'identique le texte de l'article 8 de la directive européenne du 24 octobre 1995 dont la loi du 6 août 2004 assure la transposition. Les « origines ethniques » font leur apparition et la « vie sexuelle » remplace « les mœurs ».

<sup>19</sup> en particulier dans la mesure où la finalité du traitement l'exige, le consentement exprès (article 8-II, 1°), c'est-à-dire, un accord explicite et écrit au sens de la jurisprudence du Conseil d'Etat.

L'anonymisation des données (article 8-III) : les traitements de données sensibles susceptibles de faire l'objet « à bref délai », d'un procédé d'anonymisation reconnu conforme à la loi par la CNIL peuvent être autorisés.

- L'intérêt public (article 8-IV) : le traitement de données sensibles est possible à condition d'être justifié par l'intérêt public et autorisé par la Commission. Cette dérogation figurait dans la loi du 6 janvier 1978. L'alinéa 3 de l'article 31 prévoyait que « pour des motifs d'intérêt public, il peut aussi être fait exception (...) sur proposition ou avis conforme de la Commission par décret en conseil d'Etat. » ;

- Les fichiers de recherche médicale autorisés par la CNIL (article 8-II, 8°).

- L'INSEE et les services statistiques ministériels (article 8-II, 7°) peuvent réaliser des traitements comportant des données sensibles dès lors qu'ils ont été autorisés par la CNIL, dans les conditions prévues par l'article 25 de la loi.

<sup>20</sup> Conseil d'Etat, 5 juin 1987, K.

Sur le plan pratique, le recueil du consentement de chaque employé peut être ressenti comme une contrainte lourde par les entreprises ou les administrations.

Sur le plan juridique, il est difficile de considérer que, dans le cadre des relations de travail, le consentement des salariés soit toujours réellement libre<sup>21</sup>. En outre, le recueil du consentement a pour conséquence de ne soumettre qu'à un simple régime déclaratif des fichiers pourtant sensibles .

### **- L'anonymisation des données à bref délai (article 8-III)**

Les traitements de données sensibles susceptibles de faire l'objet « à bref délai »<sup>22</sup>, d'un procédé d'anonymisation reconnu conforme à la loi par la CNIL peuvent être autorisés. Pour se prévaloir de cette dérogation, trois conditions sont requises :

. le procédé d'anonymisation doit assurer de façon effective l'anonymat des personnes, c'est-à-dire garantir que l'on ne peut associer les données sensibles à une identification quelconque de la personne ;

. le procédé doit être déclaré conforme à la loi par la CNIL. La Commission doit disposer d'une description technique très précise du dispositif envisagé afin d'apprécier sa fiabilité ;

. l'anonymisation doit intervenir à bref délai. La CNIL a, dans le domaine de la santé et le domaine social, autorisé des expérimentations recourant à de tels procédés<sup>23</sup>.

Ce type de procédé peut faire appel à des algorithmes dit de hachage permettant de transcoder les données d'identité propres à une personne<sup>24</sup> en un « numéro d'anonymat » non réversible. Ces procédés ne font cependant pas, pour le moment, l'objet d'une commercialisation et sont utilisés dans des domaines très spécifiques (santé et social).

### **- L'intérêt public (article 8-IV)**

La loi prévoit que le traitement de données sensibles est possible à condition d'être justifié par l'intérêt public et autorisé par la Commission.

La justification de l'*intérêt public* dans un traitement de données à caractère personnel est examinée « au cas par cas ».

---

<sup>21</sup> la Commission rejoint la position du G29 exprimée dans son avis n°8/2001 du 13 septembre 2001 estimant que, dans un contexte salarié, c'est-à-dire une relation de subordination, le consentement des personnes ne peut être considéré comme donné librement.

<sup>22</sup> c'est-à-dire de façon immédiate ou quasi-immédiate et au maximum de l'ordre de quelques minutes.

<sup>23</sup> Délibération n°2004-081 du 9/11/2004 autorisant une expérimentation présentée par la fédération nationale de la mutualité française ayant pour finalité d'accéder sous forme anonymisée aux données de santé figurant sur les feuilles de soin électroniques.

Délibération n°01-054 du 18/10/2000 portant sur l'examen dans le domaine médical de l'outil d'anonymisation FOIN (fonction d'occultation d'informations nominatives) utilisé pour la gestion du SNIIRAM

<sup>24</sup> tel que le procédé de « fonction d'occultation des informations nominatives » (FOIN) utilisé par l'assurance maladie et dans le domaine social (observatoire parisien du RMI) après avis favorables de la CNIL

Peut-on considérer de façon générale que la lutte contre les discriminations présente un intérêt public et dès lors légitime tous les traitements de données sensibles réalisés dans ce cadre ? Autrement dit, peut-on considérer que tous les projets de recherche présentés à la Commission et qui ont pour objectif affiché de mesurer la diversité, de suivre les trajectoires, d'analyser les facteurs de discrimination... revêtent *a priori* tous un intérêt public, et ce quelle que soit la nature publique ou privée de l'organisme responsable de ladite recherche ?

La réponse à cette question est délicate.

En 2006, la CNIL s'est trouvée confrontée à la difficulté d'apprécier si en l'espèce des projets de recherche qui lui étaient soumis relevaient ou non de l'intérêt public. Après discussion<sup>25</sup>, elle l'a finalement admis.

Cette question de la définition de l'intérêt public a été posée lors des auditions. Ainsi, Roxane Silberman, directrice de recherches au CNRS, a souligné que des évolutions législatives sont nécessaires « *sur la façon peu claire dont est située dans la loi la recherche par rapport à la notion d'intérêt public : l'exception d'intérêt public concerne-t-elle la recherche ou pas ?* ».

Ceci démontre, semble-t-il, le caractère parfois peu opérant des régimes dérogatoires introduits par la loi du 6 août 2004 et la nécessité de réfléchir sur une évolution possible de la loi en ce domaine (cf. infra).

#### **- L'INSEE et les services statistiques ministériels (article 8-II,7°)**

Une dernière dérogation est ouverte au bénéfice de l'INSEE et des services statistiques ministériels. Ces derniers peuvent réaliser des traitements comportant des données sensibles dès lors qu'ils ont été autorisés par la CNIL, dans les conditions prévues par l'article 25 de la loi. Ils ne sont alors pas tenus de recueillir le consentement des personnes .

### **C. Parce que la mesure de la diversité touche à l'identité humaine et aux libertés individuelles**

Conformément à l'article 1<sup>er</sup> de la loi informatique et libertés, il appartient à la CNIL, de s'assurer que les recueils et traitements de données sensibles susceptibles d'être mis en œuvre aux fins de mesure de la diversité et d'analyse des discriminations, ne portent atteinte ni à l'identité humaine, ni aux droits de l'homme ni à la vie privée ni aux libertés individuelles ou publiques.

Or, force est de constater que sur ce sujet des statistiques ethniques, de nombreuses voix s'élèvent<sup>26</sup> pour évoquer les risques de détournement, de stigmatisation possibles, certains allant même jusqu'à considérer que ce type de statistiques serait dangereux en soi et réducteur.

---

<sup>25</sup> cf procès-verbal de la séance du 8 juin 2006

<sup>26</sup> Ligue des droits de l'homme, MRAP, SOS Racisme, certaines organisations syndicales

L'identité d'une personne est multiple, comme l'a rappelé Louis Schweitzer. « *La diversité est multifactorielle : elle tient compte du caractère, de la personnalité, de la formation, des convictions ou des idées. Or toute mesure est synonyme de catégorisation et la définition de catégories génère un effet sur la réalité.* »

David Messas, Grand Rabbin de Paris, estime lui que « *les statistiques réduisent la réalité en assimilant l'identité individuelle à des chiffres.* »

Or, comme le souligne Patrick Simon (INED): « *la statistique remet en cause la perspective de passer inaperçu. Ce point est très délicat dans la mesure où si les sondés demandent de la reconnaissance, ils ne souhaitent pas pour autant être identifiés.* »

Le débat ne paraît pas être le même pour les chercheurs d'un côté et les entreprises et associations de l'autre.

Les scientifiques, chercheurs et démographes, estiment que la société française doit mieux se connaître. Pour Patrick Simon (INED), il paraît « *essentiel de faire un état des lieux* » car « *les confusions sont faites entre la société telle qu'elle est réellement et la société telle que l'on voudrait qu'elle soit.* » Sur ce point, Vincent Tiberj (CEVIPOF) relève que « *les statistiques sont essentielles afin d'avoir une idée plus précise de la diversité et des discriminations* ».

Mais si les chercheurs veulent « mesurer », ils ne paraissent pas tous favorables à la création d'une nomenclature nationale « ethno- raciale ». En effet, comme le dit Roxane Silberman<sup>27</sup> en matière de recensement, « *aller jusqu'à la catégorie ethno- raciale paraît prématuré* » tandis que Jean-Luc Richard<sup>28</sup> confirme son « *hostilité à tout enregistrement des origines ethniques* » et Alain Blum<sup>29</sup> met en question leur constructibilité.

Dans ce contexte, ce sont surtout les revendications de reconnaissance de l'identité dite « noire », qui s'entendent de plus en plus fréquemment. Pour Patrick Lozès,<sup>30</sup> la mesure de la diversité, avec l'introduction de catégories ethniques, « *n'a d'autre but que de favoriser l'intégration nationale et diminuer les discriminations* ». Cette réflexion, que partage Carole Da Silva<sup>31</sup>, est poursuivie par Dogad Dogoui<sup>32</sup> qui estime que « *on ne peut apporter un remède efficace à un mal qui ronge la société [la discrimination] si on ne sait pas précisément de quoi on parle.* »

### **III. DES MALENTENDUS A DISSIPER**

Sur la base des propositions formulées par le groupe de travail qu'elle avait constitué, la CNIL, en juillet 2005, a estimé nécessaire de rendre publiques ses premières

---

<sup>27</sup> Directrice de recherches au CNRS

<sup>28</sup> Maître de conférences à l'Université de Rennes 1, Docteur en démographie économique, ancien chargé de mission au Commissariat général du Plan, Ancien Vice-Président de l'Université de Brest.

<sup>29</sup> INED

<sup>30</sup> Président du Conseil représentatif des associations noires (CRAN)

<sup>31</sup> Directrice-Fondatrice de l'Association pour Favoriser l'Intégration Professionnelle (AFIP)

<sup>32</sup> Président du Club Africagora



recommandations à destination des employeurs privés et publics qui, dans le cadre de leur politique de lutte contre les discriminations, auraient à utiliser des données à caractère personnel.

Toutefois, il apparaît que ses recommandations n'ont pas toujours été bien comprises car depuis leur publication, elle est régulièrement interpellée sur ce point par les entreprises et les chercheurs, notamment. C'est pour cette raison que la Commission estime nécessaire de dissiper quelques malentendus qui circulent....

### **1. La CNIL ne réaliserait pas l'importance, pour les entreprises, de mettre en place des outils de lutte contre la discrimination.**

#### **Qu'avons-nous dit ?**

Dans ses recommandations de 2005, la CNIL a estimé que « *les objectifs de lutte contre les discriminations poursuivis par les employeurs publics et privés sont légitimes au regard de l'intérêt public en cause reconnu par les dispositions du code du travail et du code pénal.* »

Consciente des efforts déployés et des moyens investis dans la lutte contre la discrimination, notamment au niveau des entreprises, la Commission ne remet bien entendu pas en question la nécessité de mesurer l'impact de ces efforts ou des politiques mises en oeuvre. Elle ne méconnaît pas non plus l'importance de la recherche en sciences humaines et des statistiques pour non seulement mieux connaître les différentes composantes de notre société mais aussi et surtout analyser en profondeur les facteurs de discrimination.

### **2. La CNIL n'autoriserait pas l'utilisation des fichiers de gestion du personnel pour permettre aux entreprises de mesurer la diversité de leurs effectifs et l'existence éventuelle de discriminations, notamment dans le recrutement.**

#### **Qu'avons-nous dit ?**

La CNIL estime que l'utilisation, à des fins de mesure statistique de la diversité, des données enregistrées dans les fichiers de gestion des ressources humaines existants et dûment déclarés auprès de la CNIL ne pose pas de difficulté de principe au regard de la loi « informatique et libertés » dès lors bien entendu que cette exploitation s'effectue dans le respect des règles de protection des données, s'agissant en particulier de la confidentialité et de l'information des personnes sur les conditions d'utilisation des données les concernant et sur leurs droits.

En revanche, s'il s'agit d'identifier l'origine raciale ou ethnique des personnes, le traitement de leur nationalité ou de leur adresse n'est pas pertinent. Ainsi, le fait d'être de nationalité française n'est en aucune façon révélateur de l'origine d'une personne, pas plus que le fait d'habiter aux Minguettes ou ailleurs.

En outre, la CNIL a considéré qu'il n'y avait pas lieu de procéder à « *l'enregistrement dans les fichiers de gestion des ressources humaines de la nationalité d'origine d'un employé ou d'un candidat à un emploi ayant la nationalité française, de la nationalité ou du lieu de*

*naissance de ses parents ou de ses grands- parents.* » Autrement dit : il n'est pas nécessaire de connaître l'origine des parents d'un candidat à l'emploi dans l'appréciation de sa candidature et la décision de le convoquer, ou non, à un entretien d'embauche.

### **3. Alors que des pays démocratiques ont déjà un référentiel national « ethno-racial », en France, la CNIL s'y opposerait.**

#### **Qu'avons-nous dit ?**

La loi informatique et libertés permet sous certaines conditions (cf supra) le traitement de données sur les origines raciales ou ethniques . La CNIL a exprimé ses réserves à l'égard du recueil et du traitement, dans des catégories prédéfinies, de données relatives à l'origine raciale ou ethnique réelle ou supposée des employés ou des candidats à un emploi en l'état de l'information statistique disponible et en particulier de l'absence de référentiel national de typologies ethno-raciales et par voie de conséquence de résultats statistiques fiables aux niveaux national et local qui pourraient servir aux employeurs comme base de comparaison fiable (indicateurs nationaux ou par bassin d'emploi). En conséquence, la CNIL a recommandé dans ses conclusions de juillet 2005 que « *les données relatives à l'origine raciale ou ethnique réelle ou supposée des employés ou des candidats à un emploi dans le cadre de la mesure de la diversité des origines ne soient pas recueillies* ».

#### **Qu'en est-il à l'étranger ?<sup>33</sup>**

En Grande-Bretagne, la question de l'appartenance ethnique a été introduite à l'occasion du recensement de 1991. Le débat autour de la définition des catégories a été porté par la sous-commission parlementaire à l'immigration et aux *Relations Raciales*, les minorités ethniques, les statisticiens et la Commission pour l'Égalité Raciale (CRE). De 1991 à 2001, on est passé de 9 catégories initiales à 16 catégories actuelles, structurées dans cinq groupes : « blanc », « métis », « asiatique ou Anglais d'Asie », « noir ou Anglais noir » et « Chinois ou autre peuple ethnique ». Ces catégories représentent donc, pour la Grande-Bretagne, son référentiel « ethno-racial », sa grille d'analyse. Elle est appliquée de manière identique au niveau de la statistique publique et au niveau de l'entreprise et de l'établissement.

Les Etats-Unis enregistrent des données relatives aux origines ethniques ou à la « race » depuis le premier recensement en 1790. Au fil des années, ces catégories ont changé en fonction des débats politiques et transformations juridiques. Le dernier recensement, en 2000, abordait 14 catégories de « races » et 4 catégories spécifiques pour l'« origine » hispanique.

Par ailleurs, les entreprises et établissements ont à leur disposition, depuis 1996, un standard de référence, le Standard Form 100 ou Employer Information Report EEO-1, promulgué par l'*Equal Employment Opportunity Commission* (EEOC) et l'*Office of Federal Contract Compliance* (OFFC). Les catégories figurant sur ces formulaires sont donc officiellement fixées.

---

<sup>33</sup> Les informations ici présentées sont tirées de la publication du Centre d'Analyse Stratégique, *Les statistiques « ethniques » : premiers éléments de cadrage*, Pierre-Yves Cusset, CAS, octobre 2006

Au Canada, le recensement aborde les questions sur l'origine ethniques, les « minorités visibles », les groupes autochtones ou la religion dans son questionnaire long adressé à 20% de la population. A titre d'exemple, la question sur l'origine ethnique comptait, en 2001, 25 propositions de catégories.

Dans les entreprises et établissements, la grille de classification des minorités visibles, entre autres groupes de population, est fixée par la loi, en partenariat avec *Statistique Canada*. Les catégories sont identiques dans tous les fichiers de collecte.

#### **4. La CNIL représenterait un frein à la recherche. Les chercheurs ne disposeraient pas aujourd'hui de suffisamment d'informations pour pouvoir mesurer la diversité de la société française .**

##### **Qu'avons-nous dit ?**

Depuis sa création, la CNIL travaille étroitement et en bonne intelligence avec les chercheurs, les statisticiens et démographes. A l'occasion des discussions intervenues avec les chercheurs, la Commission a su faire comprendre ses préoccupations en recevant les arguments qui lui étaient opposés. Un dialogue fructueux a ainsi pu être établi, ce qui a été confirmé par les représentants de la statistique publique lors de leurs auditions par le groupe de travail.

S'agissant des nombreuses enquêtes statistiques menées notamment par l'INSEE ou par l'INED, la CNIL admet que les personnes soient questionnées sur leur nationalité ou leur mode d'acquisition de la nationalité française.

Dans un certain nombre d'enquêtes démographiques, sont collectés la date et le lieu de naissance, la nationalité, la date d'arrivée en France, le pays de naissance des parents. Ces éléments permettent d'obtenir une connaissance sur les origines des personnes, notamment sur les immigrés.

En dehors des enquêtes spécifiques, telle l'enquête « Mobilité Géographique et Insertion Sociale » (MGIS), citée plus haut, les statisticiens peuvent obtenir, à l'occasion d'enquêtes démographiques, des informations sur les origines des personnes, qui ne relèvent pas en tant que telles des dispositions de l'article 8 de la loi du 6 janvier 1978 modifiée. Il s'agit des données d'état civil sur la date et le lieu de naissance, la nationalité, la date d'arrivée en France, le pays de naissance des parents. Tous ces éléments permettent de connaître les origines des personnes, notamment des immigrés.<sup>34</sup>

---

<sup>34</sup> Définition du Haut Conseil à l'intégration : "Est immigrée, une personne née étrangère à l'étranger et venue s'installer en France, qu'elle ait ou non la nationalité française."

Tableau des enquêtes présentées à la CNIL

Enquête	Année	Réalisation	Données	Observations CNIL
Recensements généraux de la population dans les TOM <sup>35</sup>	1983 - TOM 1991 - Nouvelle Calédonie 2002 - Mayotte	INSEE	communauté d'appartenance, la tribu d'appartenance, le statut civil coutumier, la polygamie (Mayotte)	<p>La question sur l'appartenance ethnique des personnes, compte tenu des caractéristiques socio-démographiques propres aux Territoires d'Outre-Mer, était utile à la finalité du recensement et sa mise en mémoire répondait ainsi à un motif d'intérêt public au sens de l'alinéa 3 de l'article 31 de la loi du 6 janvier 1978.</p> <p>La Commission a pris en considération, pour émettre ses avis, le fait que les ethnies étaient consacrées par certains textes et que la connaissance de l'appartenance ethnique était utile pour définir les politiques sociales et économiques à mener.</p>
L'enquête sur la mobilité géographique et l'insertion sociale (MGIS) de l'INED <sup>36</sup>	1992	INED, INSEE,	la nationalité, la vie matrimoniale (dont la polygamie), la fécondité et la contraception, les enfants, l'alphabétisation et la maîtrise du français, la scolarité et les études, la vie professionnelle, la formation professionnelle, le logement, les revenus, les loisirs, la vie sociale et la pratique religieuse.	<p>Prenant en compte l'enjeu des problèmes posés par l'intégration des populations étrangères en France, la CNIL a considéré que les données recueillies, en dépit de ce qu'elles touchaient de près à la vie privée des personnes interrogées, étaient "adéquates, pertinente et non excessives" au regard de la finalité de l'enquête.</p> <p>Le choix de la nationalité des populations soumises à l'enquête et les informations collectées sur ces personnes relatives à la polygamie, l'asile politique, la pratique religieuse et la langue maternelle sont des données qui ne peuvent être recueillies, sauf dérogation prévue par la loi, qu'avec l'accord exprès des intéressés.</p>

<sup>35</sup> délibération n° 83-12 du 18/01/1983 portant avis sur la mise en œuvre du RGP dans les TOM ; délibération n° 89-02 du 10/01/1989 portant avis sur la mise en œuvre du RGP en Nouvelle-Calédonie ; délibération n° 91-027 du 02/04/1991 portant avis sur la mise en œuvre du RGP à Mayotte ; délibération n° 95-116 du 17/10/1995 portant avis sur le projet de décret présenté par l'INSEE, portant application des dispositions de l'article 31 alinéa 3 de la loi du 6/01/1978 au traitement mis en œuvre à l'occasion du RGP en Nouvelle-Calédonie ; délibération n° 02-012 DU 14/03/2002 portant avis sur le projet de décret, présenté par l'INSEE portant application des dispositions de l'article 31 alinéa 3 de la loi du 6/01/1978 au traitement réalisé à l'occasion du RGP à Mayotte en 2002

<sup>36</sup> délibération n° 92-057 du 9 juin 1992 portant avis sur la mise en œuvre par l'INED d'une "enquête nationale sur la mobilité géographique et l'insertion sociale"

Enquête	Année	Réalisation	Données	Observations CNIL
<b>Recensement Général de la Population (RGP)<sup>37</sup></b>		INSEE	<b>date et lieu de naissance</b> : avec l'indication du pays de naissance et l'année d'arrivée en France pour les personnes nées à l'étranger, <b>nationalité</b> : nationalité française avec distinction entre <b>nationalité française à la naissance/ou par acquisition</b> <b>Depuis 1962</b> : mention de la nationalité à la naissance <b>indication de leur nationalité pour les étrangers</b>	La CNIL n'a jamais remis en cause cette collecte dans le cadre du recensement qui a justement pour finalité la mesure de la population.
<b>La mise en place du tronc commun des enquêtes<sup>38</sup></b>	1996	INSEE	le pays de naissance, la nationalité avec la mention français de naissance et français par acquisition	<p>Pour répondre aux recommandations du Haut Conseil à l'intégration désireux que l'intégration des immigrés en France soit mesurée de manière régulière.</p> <p>Toutes les enquêtes statistiques auprès des ménages comportent donc ces éléments qui permettent de répartir les immigrés selon leur pays d'origine.</p>

<sup>37</sup> délibérations n°81-003 du 10 mars 1981, n°89-010 du 14 février 1989, n°96-110 du 17 décembre 1996, n°98-023 du 24 mars 1998, n°98-099 du 8 décembre 1998, n°03-068 du 18 décembre 2003

<sup>38</sup> délibération n° 96-40 du 07 mai 1996. La Commission a accepté les propositions de l'INSEE, estimant « qu'il était utile de disposer de données statistiques pour mesurer l'intégration des personnes et ainsi ne pas laisser la porte ouverte à des interprétations et argumentations erronées. »

Enquête	Année	Réalisation	Données	Observations CNIL
« Enquête famille » <sup>39</sup>	Depuis 1982	INED, INSEE	Depuis 1989, langues d'usage au sein de la famille ; Le pays de naissance des parents.	La CNIL a donné un avis favorable à la mise en œuvre de l'enquête après de longues discussions avec ses concepteurs. Elle n'a pas estimé que les questions sur le pays de naissance des deux parents de la personne interrogée ainsi que sur la langue maternelle étaient susceptibles de faire apparaître l'appartenance raciale ou ethnique des personnes, admettant les arguments des chercheurs qui indiquaient que "l'existence de plusieurs ethnies par pays et de plusieurs pays par ethnie – comme c'est le cas en Afrique, brouillait les pistes et ôtait toute validité à l'équation pays de naissance = ethnie = race". <sup>40</sup> Ces données n'ont donc pas été assimilées à des données sensibles au sens de la loi de 1978.
L'enquête « histoire de vie » <sup>41</sup>	2003	INSEE, INED	Langues, discriminations vécues, la nationalité et le pays de naissance des parents	La Commission a émis un avis favorable à la mise en œuvre de l'enquête, considérant que les données étaient pertinentes au regard de la finalité poursuivie. Il s'agissait, en effet, de décrire, hiérarchiser et analyser les différents types de liens sociaux qui permettent aux individus de s'intégrer dans la société française.
« Intégration des deuxièmes générations » ( enquête européenne « TIES ») <sup>42</sup>	2006/2007	INED	Habitudes de vie ; <b>Constitution de l'échantillon ; tri par consonance des patronymes</b>	« L'enquête envisagée par l'INED est justifiée par l'intérêt public dans la mesure où elle doit permettre de mesurer l'intégration des secondes générations turques et marocaines et contribuer à remédier ainsi à l'insuffisance actuelle de données statistiques dont souhaitent disposer les pouvoirs publics pour définir et mettre en œuvre des politiques en matière d'intégration à l'attention de ces populations, tant au niveau national qu'euro péen ».

<sup>39</sup> délibération n° 94-020 du 1<sup>er</sup> mars 1994

<sup>40</sup> extrait de l'article précité de F. Héran – juillet 2004

<sup>41</sup> demande d'avis n° 827.528 AT n° 025346 du 29 janvier 2003

<sup>42</sup> délibération n°2006-170 du 27 juin 2006

Par ailleurs, au cours de l'année 2006, la Commission s'est prononcée sur trois enquêtes prévoyant le recueil de données sensibles. Il s'agit des traitements du Conseil représentatif des institutions juives de France (CRIF), de l'Institut National des Etudes Démographiques (INED) et de l'Observatoire Régional des Etudes Supérieures (ORES) qui ont fondé leur méthodologie d'enquête sur l'analyse des patronymes et des prénoms.

- Pour l'INED, l'objectif de l'enquête était d'analyser l'intégration des descendants d'immigrés marocains et turcs, le critère étant donc l'origine des personnes. L'enquête n'a d'ailleurs finalement été réalisée que sur les descendants d'immigrés turcs. Le tri patronymique était effectué, selon des critères linguistiques, pour en déduire une origine nationale supposée des ascendants des personnes sélectionnées (et non leur origine « ethnique ») et ainsi déterminer les personnes susceptibles de participer à l'enquête, étant entendu qu'une enquête téléphonique auprès des personnes devait permettre de vérifier si elles répondaient aux critères fixés par le protocole d'enquête et d'obtenir alors leur accord.

La Commission a autorisé ce traitement, prenant en compte l'intérêt public de l'enquête et le fait que le consentement écrit des personnes concernées serait recueilli.

- Pour l'Observatoire Régional des Etudes Supérieures (ORES), il s'agissait d'étudier les parcours des étudiants inscrits dans l'enseignement supérieur du Nord-Pas de Calais en analysant notamment leurs prénoms et le cas échéant leurs noms.

L'étude devait être menée sur un échantillon obtenu après le croisement d'une base de 150 000 étudiants et d'une base de 10 à 13 000 prénoms d'origine arabe ou musulmane. Pour les prénoms d'origine incertaine et d'usage « mixte », le nom de la personne devait également être pris en compte.

Quoique reconnaissant que l'objectif poursuivi revêtait bien un intérêt public, la CNIL, dans un premier temps, n'a pas autorisé la mise en œuvre de ce traitement en raison de l'absence d'information des personnes concernées.

Toutefois, à la suite des échanges avec la CNIL, l'ORES a pris en compte l'ensemble des arguments de la Commission et a modifié le projet afin de garantir l'anonymat des étudiants en basant son étude uniquement sur le traitement des prénoms. Une démarche d'information des étudiants et de sensibilisation des chefs d'établissement en matière d'« informatique et libertés » a été également prévue.

- En revanche, l'enquête du CRIF n'a pas été autorisée par la CNIL.

Le CRIF avait souhaité constituer une liste des personnes supposées appartenir à la communauté juive afin de « *mesurer, objectiver et analyser l'état de son opinion en France* ».

Dans sa délibération du 2 février 2006, la Commission a considéré que l'obstacle résidait dans « *la constitution de l'échantillon des personnes à interroger qui repose uniquement sur un tri sur la consonance de leur nom et que ce tri a pour objet même de faire apparaître l'appartenance, réelle ou supposée, des intéressés à la communauté juive* ». Par ailleurs, la Commission a estimé que « *le postulat selon lequel les personnes ayant un nom figurant dans le « Guide des patronymes juifs » appartiennent à la communauté juive est contestable et*

*souligne les risques liés à une sélection de ce type ». Elle a donc considéré que « le traitement de données ne répondait pas aux conditions posées par l'article 6 de la loi aux termes duquel les données doivent être « adéquates, pertinentes et non excessives au regard des finalités pour lesquelles elles sont collectées », d'autant que les objectifs poursuivis par l'enquête, dont **la légitimité n'est pas en cause**, semblent pouvoir être atteints en modifiant la méthode de constitution de l'échantillon des personnes à interroger. »*

## **IV. PROPOSITIONS DE LA CNIL**

A la suite des auditions auxquelles le groupe de travail a procédé, la Commission propose d'examiner les pistes de réflexion suivantes, qui, pour certaines d'entre elles, ont d'ailleurs été suggérées par plusieurs personnes auditionnées. Elles peuvent être regroupées en deux catégories :

- celles qui peuvent être décidées à droit constant (A);
- celles qui impliquent une modification de la loi (B).

### **A. Propositions en l'état actuel du droit**

Il est dans le rôle de la Commission de proposer des pistes pour améliorer l'utilisation des outils de connaissance actuellement disponibles, afin qu'ils soient plus largement ouverts au monde de la recherche, tout en les enrichissant de données plus pertinentes. Ceci suppose, en contrepartie, que les chercheurs respectent pleinement les « règles du jeu » en matière de protection des données personnelles : c'est pourquoi la CNIL considère indispensable de rappeler un certain nombre de recommandations méthodologiques en la matière.

### **Recommandation n°1 : ouvrir plus largement aux chercheurs l'accès aux bases de données statistiques et aux fichiers de gestion**

#### **a. L'accès aux bases statistiques publiques**

L'INSEE, les services statistiques ministériels et les organismes de recherche publique (CNRS, INED...) disposent de gisements de données considérables qui permettent déjà de constituer une « photographie » fidèle et détaillée de la société française, de ses flux migratoires, de ses pratiques d'intégration ainsi que de ses diversités culturelles, sociales, professionnelles...

Lors des auditions, un certain nombre de chercheurs ont cependant fait valoir la complexité des procédures d'accès aux bases de données de l'INSEE et notamment à l'Echantillon Démographique Permanent (EDP). En particulier, il a été souligné que « *l'EDP n'est utilisé que par cinq ou six chercheurs français alors qu'en Grande Bretagne 150 personnes travaillent sur son équivalent britannique* »<sup>43</sup>. Or, il s'agit de données particulièrement riches et variées et selon les interlocuteurs auditionnés, insuffisamment exploitées.

---

<sup>43</sup> Jean-Luc Richard



Tout en reconnaissant que la politique de diffusion des données de l'INSEE était plus restrictive que dans d'autres pays, le directeur général de l'INSEE a cependant souligné que l'institut s'est engagé, depuis quelques mois, dans une démarche de plus grande ouverture en mettant à disposition des chercheurs des fichiers statistiques plus riches (fichiers intermédiaires).

**La Commission estime que les bases des données statistiques publiques doivent être ouvertes plus largement au monde de la recherche dès lors bien entendu que cet accès et les exploitations de ces données s'opèrent dans des conditions garantissant la protection des données et l'anonymat des personnes.**

A cet égard, la création de **centres d'accès sécurisés**, comme l'a suggéré le directeur général de l'INSEE Jean-Michel Charpin, citant le cas d'exemples étrangers en la matière, constitue une piste qui doit être approfondie. Il existe déjà, en France, un centre, « le Réseau Quételet » (dénommé comité de concertation pour les données en sciences humaines et sociales) qui a été créé par un décret du 12 février 2001<sup>44</sup>.

Comme l'a souligné le directeur général de l'INSEE, il est « *envisageable que l'INSEE aille plus loin, à l'instar de certains pays qui ont mis en place des **centres d'accès sécurisé** (Etats-Unis, Allemagne, Canada, etc.) permettant à des chercheurs d'accéder à des données encore plus riches. Cette orientation est intéressante mais requiert des financements relativement importants.* »

**La Commission soutient cette proposition et estime que le financement nécessaire à la réalisation du projet devrait pouvoir être trouvé sur le plan national, vu l'intérêt général évident que revêt la création d'un tel centre.**

- Il convient sans doute aussi de procéder à **un réexamen d'ensemble des conditions générales de diffusion des données issues du recensement**, telles que prévues par le décret du 5 juin 2003 relatif au recensement de la population. A cet égard, des propositions de modification, allant dans le sens d'une plus grande accessibilité, font actuellement l'objet de réflexions entre la CNIL et l'INSEE.

**Rappel des conditions actuelles de diffusion des données du recensement de la population :**

---

<sup>44</sup> il a notamment pour mission de s'informer des activités de production et d'utilisation des données; de déterminer les ensembles de données dont l'intérêt scientifique justifie le recueil, la préservation et la diffusion pour la recherche, de formuler des propositions en vue de la production de grandes enquêtes utiles à la recherche.

- le niveau d'agrégation des résultats se présentant sous la forme de fichiers détails (c'est-à-dire de questionnaires individuels comportant pour tout élément d'identification la zone géographique dans laquelle est située l'adresse de la personne avant répondu) est de 50 000 habitants.
- les résultats statistiques présentés sous la forme de fichiers tableaux (simples comptages avec ventilation selon différents critères), lorsqu'ils ne comportent pas de données sensibles (données relatives à la nationalité ou aux migrations sont disponibles au niveau de la commune ou au niveau infra communal d'un quartier fixe d'environ 2000 habitants (dits IRIS 2000) ou 5000 habitants (IRIS 5000). Des tableaux avec variables sensibles sont diffusés pour les communes de plus de 5000 habitants et pour des zones infra communales fixes regroupant environ 6000 personnes.
- Moyennant la signature d'une licence d'usage avec l'INSEE, les collectivités territoriales et leurs regroupements, les administrations et les établissements publics ayant une mission de création ou de gestion de service public peuvent obtenir, à l'îlot, des tableaux sans donnée sensible et des fichiers détails sur les logements.

Ainsi, il serait envisagé de modifier les modalités de diffusion de certaines variables (**nationalité, pays de naissance, pays de résidence cinq ans auparavant, année d'arrivée en France**) qui ne sont actuellement communicables que sous la forme de tableaux standards. L'Insee propose aussi que l'indicateur « immigré »<sup>45</sup> ne fasse l'objet d'aucune restriction de diffusion, contrairement à ce que prévoient les modalités actuelles de diffusion du recensement général de la population de 1999.

#### **b. L'accès aux fichiers de gestion**

Les fichiers de gestion, qu'il s'agisse des fichiers de personnel, d'élèves, d'étudiants... comportent déjà un certain nombre de données (ex : nom, prénom, nationalité, lieu de naissance, adresse) susceptibles d'être utilisées à des fins statistiques pour mesurer la diversité, suivre les trajectoires et évaluer les politiques de lutte contre les discriminations.

**En ce domaine également, la Commission estime que ces fichiers, qui constituent une source substantielle d'information pour les chercheurs peuvent leur être ouverts plus largement dès lors bien entendu que ces exploitations s'opèrent dans des conditions et selon une méthodologie rigoureuse garantissant la protection des données et l'anonymat des personnes (cf. ci-après).**

**Recommandation n°2 : mesurer la diversité en utilisant les données « objectives » relatives à l'ascendance des personnes (nationalité et/ou lieu de naissance des parents)**

<sup>45</sup> valant 0 si la personne est française de naissance ou née en France et 1 sinon.

Dans ses recommandations du 5 juillet 2005, la CNIL a estimé que les données relatives à la nationalité ainsi qu'au lieu de naissance des parents pouvaient être utilisées pour mesurer la diversité au sein des entreprises et, plus largement, dans la société française.

Certes ces données ont sans doute une utilité limitée dans le temps, dans la mesure où comme le souligne P. Simon (INED) pour les descendants des enfants d'immigrés, l'analyse de telles données n'aura sans doute plus grand sens.... Toutefois, elles semblent constituer, pour les chercheurs, des éléments intéressants et aujourd'hui suffisants pour déterminer indirectement l'origine des personnes.

A cet égard, l'étude sur la mesure de la diversité menée par P. Simon et M. Clément auprès de plus de 1300 personnes<sup>46</sup> a montré que respectivement 85 % des personnes ayant déclaré des ascendants nés dans la région du Maghreb se considéraient aussi au moins en partie de cette origine, ce qui semble montrer, aux yeux des chercheurs, la « fiabilité » de la donnée relative au pays de naissance des parents en tant qu'indicateur indirect de l'origine des personnes.

Selon les chercheurs et statisticiens, il semble également que le fait de poser des questions sur l'origine et la nationalité des parents voire des grands-parents soit bien accepté par la société française, alors que celle-ci ne semble pas prête à recourir à des critères reposant sur les origines raciales ou ethniques<sup>47</sup>.

### **La Commission a examiné dans quelle mesure les données sur l'ascendance pourraient être utilisées de manière complémentaire dans les fichiers suivants .**

#### **1) Dans des enquêtes adossées au recensement et éventuellement dans le recensement**

La question sur la nationalité et/ou le lieu de naissance des parents, a été jusqu'à présent réservée à des enquêtes spécifiques (enquête Mobilité géographique et insertion sociale (MGIS) en 1992 ; enquête Etude de l'Histoire familiale (EHF) en 1999, associées au recensement...). L'INSEE envisage de mener au second semestre 2008 une enquête intitulée « Trajectoires et origines sur les immigrés et descendants d'immigrés ». L'échantillon d'enquête sera constitué à partir des informations issues du recensement, de l'échantillon démographique permanent<sup>(48)</sup> et des actes de naissance de l'état civil en particulier pour obtenir les lieux de naissance des parents. Certains chercheurs (notamment F. Heran et P. Simon de l'INED) proposent que la question sur la nationalité et/ou le lieu de naissance des parents figure dans le recensement, ce qui permettrait d'obtenir des statistiques précises et exhaustives sur l'origine des personnes. L'INSEE pourrait donc être invité à retenir de telles propositions. Les données obtenues pourraient être utilisées par les entreprises et les administrations, sous réserve que les statistiques ainsi diffusées respectent l'anonymat .

---

<sup>46</sup> enquête anonyme qui consistait notamment à répondre à des questions sur le pays de naissance des parents et des grands-parents, l'origine par aire géographique ou pays , l'appartenance à une nomenclature « ethno-raciale », cf. article paru dans « Population et Sociétés » n° 425 juillet-aout 2006

<sup>47</sup> cf. notamment l'audition de Vincent Tiberj (CEVIPOF).

<sup>48</sup> Cette méthode d'appariement a déjà été utilisée lors de l'enquête « Mobilité géographique et insertion sociale » de 1992-1993, qui s'intéressait aux 20-29 ans nés en France d'au moins un parent né en Algérie, en Espagne ou au Portugal.

**La Commission estime que l'intégration de questions sur la nationalité et le lieu de naissance des parents peut être admise dans le cadre d'enquêtes spécifiques adossées au recensement dès lors que l'objet de celles-ci le justifie. Elle considère qu'il peut aussi être envisagé que de telles questions puissent être posées dans le cadre du recensement de la population dès lors que toutes précautions méthodologiques auraient été prises pour garantir la protection des données et que l'acceptabilité de ces questions aurait été préalablement testée.**

## **2) Enquêtes au sein des entreprises et des administrations**

La réalisation d'enquêtes par voie de questionnaires pourrait être envisagée au sein des entreprises, comme le suggérait la CNIL dans ses recommandations du 5 juillet 2005. Dans le cadre d'un programme national de lutte contre les discriminations, dont les modalités seraient validées par la HALDE et par une instance d'expertise statistique (en lien avec le CNIS), des questions sur l'ascendance (nationalité et/ ou lieu de naissance des parents) pourraient alors être posées afin d'établir des statistiques dans le strict respect des conditions posées par la CNIL .

Un « audit diversité » pourrait aussi avoir lieu tous les ans au moins dans les grandes entreprises (ou celles qui ont adhéré à la charte de la diversité) éventuellement sous le contrôle de la Halde qui examinerait avec la CNIL, les modalités de réalisation de l'enquête et, seule, cette fois, les résultats. Les statistiques obtenues pourraient être comparées aux données statistiques du recensement propres à la région et au bassin d'emploi selon des modalités de diffusion à définir.

**La Commission considère que l'intégration de questions sur la nationalité et le lieu de naissance des parents peut aussi être admise dans le cadre d'enquêtes par questionnaires menées auprès des entreprises et des administrations, dès lors que ces enquêtes s'inscrivent dans le cadre d'un programme national de lutte contre les discriminations, dont les modalités seraient validées par la HALDE et par une instance d'expertise statistique (en lien avec le CNIS), et dès lors que toutes précautions méthodologiques sont prises pour garantir la protection des données.**

## **3) L'intégration des données sur l'ascendance des personnes dans les fichiers des usagers du service public**

Il est proposé par certains chercheurs de faire figurer au sein des fichiers administratifs, selon des modalités préservant l'anonymat des personnes, des données objectives relatives à l'ascendance (lieu de naissance et/ou nationalité des parents)<sup>49</sup>.

Ainsi, les fichiers de gestion des administrations publiques tels que les fichiers des caisses d'allocations familiales, de l'assurance maladie, les fichiers de l'éducation nationale, de la fonction publique... , intégreraient, de façon systématique (de manière obligatoire ?) mais à des fins exclusivement statistiques, des variables descriptives relatives au pays de naissance

---

<sup>49</sup> cf. notamment les auditions de F.Héran, directeur de l'INED et de P.Simon (INED).

des intéressés et de leurs parents, de leur nationalité actuelle et antérieure, de leur date d'installation en France. De telles données devraient être qualifiées de « données de trajectoire ou de parcours », et non « statistique d'origine » ou « statistique ethnique ». L'enregistrement de ces données de trajectoire devrait faire l'objet de solides garanties de confidentialité ( par exemple en recourant à des techniques d'anonymisation) afin que les informations ne soient accessibles qu'aux services statistiques ministériels et aux chercheurs habilités.

**La Commission estime que l'intégration, dans les fichiers de gestion des administrations, de données sur l'ascendance des personnes, ne peut être envisagée dans la mesure où elle impliquerait l'enrichissement, obligatoire, des fichiers administratifs par des données supplémentaires, sensibles, qui ne serviraient donc pas strictement à la gestion et où elle serait donc susceptible de comporter, au moins aux yeux des personnes concernées, le risque d'une utilisation détournée de ces informations à d'autres fins, ce qui serait de nature à altérer la confiance des usagers à l'égard de l'administration.**

La Commission considère en conséquence qu'il est préférable de privilégier la voie des enquêtes par questionnaires auprès des personnes, à l'égal de ce qui est proposé pour les entreprises. Ces enquêtes pourraient, pour certaines, reposer sur la constitution de panels permettant ainsi de suivre les trajectoires des personnes tant sur le plan scolaire, que professionnel, social ou encore dans le domaine de la santé.

### **Recommandation n°3 : permettre le développement d'études sur le « ressenti » des discriminations incluant le recueil de données sur l'apparence physique**

Comme l'a souligné Claude Bébéar, lors de son audition, *« le cœur du problème est de savoir de quelle manière les personnes se ressentent et de quelle façon elles sont ressenties par les autres »*.

Pour la plupart des personnes auditionnées, il est particulièrement important aujourd'hui de pouvoir mesurer la réalité de la discrimination vécue, ce qui suppose d'abord de demander aux personnes ce qu'elles pensent de leur propre discrimination, comment elles la subissent au quotidien...

Toutes s'accordent aussi sur la nécessité, en ce domaine, d'une démarche prudente et progressive, fondée sur le volontariat et l'auto déclaration<sup>50</sup>.

Dès lors, et ainsi que le proposent notamment les directeurs de l'INED et de l'INSEE, ne conviendrait-il pas de développer des enquêtes ad hoc par sondage, sur échantillon aléatoire, menées par la statistique publique, en particulier dans le cadre de programmes évalués par le CNIS ? Ces enquêtes permettraient de poser des questions sensibles susceptibles d'identifier

---

<sup>50</sup> y compris de ceux qui se déclarent en faveur de l'élaboration d'un référentiel « ethno-racial » ; cf. notamment le point de vue de Dogad Dogoui, président du club Africagora qui souligne la nécessité de respecter le choix des individus.

les obstacles à l'intégration et les facteurs de discrimination, tels que la couleur de la peau, ou encore toutes les autres caractéristiques visibles, audibles ou perceptibles<sup>51</sup>.

Faut-il, comme le suggère le directeur général de l'INSEE, s'en tenir à des questions ouvertes du type : « *pensez-vous que le refus d'accès à un logement, à un prêt, à un emploi, etc. soit lié à votre origine ou à votre couleur de peau ?* », « *de quelle(s) origine(s) vous diriez-vous ?* » sans énumération et tout en précisant que plusieurs réponses sont possibles?

Certains, tels le CRAN, proposent, avec précaution<sup>52</sup> de se référer d'ores et déjà à des typologies en usage dans la vie ordinaire<sup>53</sup>, en prévoyant une rubrique « autre » et une rubrique « ne souhaite pas répondre » pour préserver la liberté des personnes interrogées.

Pour P.Simon, il pourrait être envisagé d'élaborer deux batteries de questions qui seraient libellées différemment, l'une recueillant l'auto perception des personnes interrogées, l'autre recueillant ce que cette personne pense être la perception d'elle par autrui. Toutes précautions devraient être prises (par exemple sous la forme de réponses graduées et de la possibilité pour la personne de ne pas répondre) pour préserver l'auto détermination et la liberté de chacun et éviter tout risque de « catégorisation » ou pour reprendre l'expression de P.Simon (INED), prévenir « *tout risque de cristallisation d'une nomenclature ethno-raciale qui conforterait les stéréotypes raciaux* ».

**La Commission estime que des enquêtes sur le « ressenti » des discriminations, incluant le recueil d'informations sur les caractéristiques physiques, pourraient être réalisées dès lors que :**

- 1) celles-ci seraient conduites dans le cadre de la statistique publique ;**
- 2) les questionnaires auraient fait l'objet d'une validation scientifique (par exemple par le CNIS ou une autre instance scientifique indépendante : cf. ci-après) ;**
- 3) des tests préalables auraient été menés de façon à mesurer le degré d'acceptabilité des questions qui devraient être formulées sous forme de questions ouvertes<sup>54</sup> ;**
- 4) le caractère facultatif de ces questions serait expressément rappelé ;**
- 5) des précautions méthodologiques rigoureuses auraient été prises (cf ci-après).**

---

<sup>51</sup> cf. enquête histoires de vie- construction des identités menée en 2003 par l'INSEE et l'INED et le projet d'enquête de l'INSEE et de l'INED sur les trajectoires et origine des immigrés et leurs descendants.

<sup>52</sup> Le CRAN (Conseil Représentatif des Associations Noires) souligne ainsi que ces rubriques peuvent être discutées et doivent faire l'objet d'un large débat au sein de la nation. Cf [www.lecran.org](http://www.lecran.org)

<sup>53</sup> telles que, notamment, blanc, noir, arabo-berbère, asiatique

<sup>54</sup> Lors de son audition, D.Sabbagh, directeur de recherche au CERI a ainsi souligné qu'en France il serait opportun de pouvoir disposer de données précises quant à l'ampleur de la concordance entre les résultats produits par l'auto identification ethno-raciale des personnes et ceux produits par l'observation externe. Selon lui, il conviendrait de commanditer des enquêtes en la matière pour savoir dans quelle mesure les identités raciales sont cristallisées dans la société française. Il suggère aussi que d'autres études soient menées afin d'évaluer, d'une part, la légitimité du *principe* de la classification « ethno-raciale » des personnes par les pouvoirs publics à des fins anti discriminatoires, d'autre part, la légitimité des *différentes modalités possibles d'application de ce principe* – c'est-à-dire l'acceptabilité de telle ou telle catégorie éventuellement proposée – auprès, d'une part, des premiers intéressés, soit les membres des « minorités visibles », d'autre part, de l'opinion publique dans son ensemble.

**De telles enquêtes devraient être soumises à autorisation de la CNIL dès lors qu'elles ne reposeraient pas sur une collecte de données réalisée de façon anonyme.**

### **Recommandation n°4 : permettre, sous certaines conditions, l'analyse des prénoms et des noms de famille**

**L'analyse des noms et prénoms, aux fins de classement dans des catégories «ethno-raciale» n'est pas pertinente en raison du manque de fiabilité de cette méthode et du risque de stigmatisation qui pourrait néanmoins en découler .**

En effet, ainsi que la CNIL l'a rappelé en 2005 et comme cela a été souligné à plusieurs reprises lors des auditions<sup>55</sup>, on ne peut établir de lien fiable entre ces données et l'appartenance à une catégorie raciale ou ethnique déterminée, qu'il faudrait d'ailleurs au préalable pouvoir définir. Le port de tel ou tel prénom reflète avant tout le choix des parents qui peut certes faire référence indirectement à une appartenance sociale, culturelle ou religieuse, mais qui correspond le plus souvent à un choix personnel ( parfois influencé par des phénomènes de mode)<sup>56</sup>. Le fait de porter tel ou tel nom de famille ne peut non plus être considéré comme révélateur de l'appartenance à une catégorie « ethno-raciale » donnée. La francisation des noms, le choix des femmes de porter le nom de leur époux, la transmission du nom du père... sont autant de raisons qui conduisent à rejeter l'utilisation de la méthode patronymique comme critère de mesure de la diversité.

Au delà, une telle méthode, si elle était utilisée aux fins de classement « ethnique », comporterait un risque grave de stigmatisation et d'atteinte à l'identité humaine. Reposant sur le postulat, contestable, que tel nom = telle origine ethnique – ce qui supposerait alors l'élaboration d'un référentiel « ethno-racial » (cf. *infra*).-, les annuaires publics, fichiers de gestion... seraient ainsi susceptibles d'être traités, sur ces critères, à des fins notamment discriminatoires.

---

<sup>55</sup> cf. notamment les auditions de C.Bébéar, de L.Schweitzer, président de la HALDE qui évoque en particulier l'utilisation, par un nombre de plus en plus important de familles d'origine maghrébine, de prénoms à consonance celte ou anglo-saxonne . « Ceci prouve que l'indicateur du prénom est artificiel. Il en est de même en ce qui concerne les noms de famille ». cf. également R.Silberman du CNRS : « La méthode patronymique est-elle une alternative ? C'est certainement utile – cela a été utilisé aux Etats-Unis – mais présente des limites : les changements volontaires de nom, les stratégies des parents en matière de choix de prénoms pour leurs enfants, les mariages avec des effets différents pour l'homme et la femme, la multiplication des mariages mixtes avec des effets très différents selon l'origine des conjoints, tout cela introduit des biais très importants et rend l'usage de cette méthode parfois aventureux. Si cela permet de mobiliser des fichiers à moindre frais, encore faut-il mesurer les biais et par ailleurs c'est finalement plus déloyal vis à vis du citoyen que des demandes précises portant sur la filiation et l'origine. ».

<sup>56</sup> Selon P.Simon, la méthode du repérage par le prénom présente des limites méthodologiques importantes : « près du tiers des parents d'origine maghrébine donnent des prénoms « français » ou « internationaux » à leurs enfants ; tandis que des parents qui ne sont pas d'origine maghrébine peuvent choisir des prénoms à consonance arabe, comme Nadia, par exemple ».

**En revanche, le recours au prénom ainsi que, le cas échéant au nom de famille pour détecter d'éventuelles pratiques discriminatoires dans le parcours scolaire, universitaire ou professionnel, à l'exclusion de tout classement dans des catégories « ethno-raciales », peut constituer un indicateur intéressant sur le plan statistique dès lors que les conditions suivantes sont réunies :**

- **l'analyse de ce seul critère ne saurait être suffisant pour apprécier la discrimination ;** il convient de procéder à une analyse réellement objective (multicritères) des parcours ou des trajectoires de vie **prenant en compte l'ensemble des autres facteurs discriminants** (ex : sexe, âge, lieu de résidence, CSP...) **ou susceptibles d'expliquer la différence de parcours** (ex : diplômes, compétences, ancienneté et expérience professionnelle ... ) ;
- **le recours au prénom et au nom de famille peut constituer un repère pour étudier les trajectoires de populations issues de l'immigration, quand il n'existe pas d'autre source disponible ;**
- de telles études doivent être conduites selon une méthodologie rigoureuse :
  - **l'analyse des noms et des prénoms doit seulement permettre un classement de ces données en catégories « potentiellement discriminant » / « non-discriminant » ;**
  - **les fichiers de gestion (fichiers de personnel, fichiers d'élèves, d'étudiants, ...), les annuaires, professionnels et d'anciens élèves ne peuvent être utilisés qu'à des fins statistiques de suivi des trajectoires et d'évaluation des politiques de lutte contre les discriminations ;**
  - **des précautions doivent être prises pour assurer la confidentialité des données traitées et garantir l'anonymat des résultats (notamment en recourant à des tiers de confiance [cf ci-après], en prévoyant une taille d'échantillon suffisamment importante et en ne diffusant que des résultats statistiques) ;**
  - **les personnes concernées (salariés, étudiants, représentants légaux des élèves...) doivent être clairement informées de la finalité de l'étude, des conditions de sa réalisation et de leur droit de s'opposer éventuellement au traitement de leurs données.**

### **Recommandation n°5 : développer le recours à des experts tiers de confiance**

Afin de garantir à la fois le sérieux scientifique des études et la protection de la vie privée des personnes, la Commission recommande aux entreprises, établissements scolaires, universités et de façon générale à tout organisme qui souhaiterait réaliser en son sein une mesure de la diversité des personnels, des élèves... de faire appel à une expertise extérieure indépendante qui soit également un « tiers de confiance » .

#### **- Sur l'expertise**

Dans un domaine aussi sensible, il importe de garantir, sur un plan scientifique, le sérieux des études entreprises ainsi que leur impartialité . Quelles garanties prévoir ? Il paraît difficile,



dans le domaine de la recherche de définir, par exemple, des procédures d'agrément des organismes habilités à procéder à des études sur la mesure de la diversité et l'analyse des discriminations.

Confronté à cette difficulté dans le domaine de la recherche médicale, le législateur, à l'initiative de la CNIL (cf. recommandation du 19 février 1985) l'a, par la loi du 1<sup>er</sup> juillet 1994<sup>57</sup>, résolue en créant un comité scientifique chargé d'évaluer la pertinence des projets de recherche, l'utilité du recours aux données nominatives et la méthodologie employée, l'avis du comité étant ensuite communiqué à la CNIL.

Cette piste mériterait d'être creusée. Elle pourrait se traduire le cas échéant par une proposition de modification de la loi de 1978 (cf ci-après).

### **- Sur le tiers de confiance**

**S'inspirant de la procédure suivie en matière de recherche médicale (cf. article 55 de la loi de 1978 modifiée), il pourrait être préconisé que les données personnelles (qu'il s'agisse des questionnaires d'enquête ou de données issues de fichiers de gestion) soient reçues et traitées :**

- 1) s'il est fait appel à un prestataire externe (organisme de recherche), par un responsable de la recherche désigné à cet effet ;**
- 2) si l'exploitation des données se fait en interne, par un responsable de la recherche désigné à cet effet par le responsable du traitement ;**
- 3) Ces personnes s'engageraient contractuellement à respecter la confidentialité des données et veilleraient au respect de la finalité statistique des traitements. et[ S'il s'agit d'une personne physique, interne à l'entreprise, elle devrait être désignée pour ses compétences professionnelles et son intégrité personnelle. Une formation pratique devrait lui être assurée par l'employeur.**
- 4) Il serait également possible, sous réserve du respect des mêmes conditions, de confier ce rôle du tiers de confiance à une « cellule diversité » interne à l'organisme au sein de laquelle la veille sur la légalité des procédures pourrait être assurée par le Correspondant Informatique et Libertés, institué par la loi du 6 août 2004.**

Des clauses de confidentialité devraient être signées avec l'ensemble des prestataires externes (cabinets de conseil, d'études, de recherche...) susceptibles d'avoir accès aux bases de données de gestion.

**Ce tiers de confiance devrait aussi s'assurer que :**

---

<sup>57</sup> cette loi , un des volets du triptyque bioéthique voté en 1994 , a ainsi complété la loi du 6 janvier 1978 par un chapitre spécifique consacré aux fichiers de recherche médicale.

- 1) **les traitements seront réalisés dans un cadre confidentiel, c'est-à-dire par un nombre limité de personnes spécialisées et dans un environnement informatique sécurisé.**
- 2) **les résultats de l'enquête seront produits sous une forme statistique agrégée, de façon à garantir l'anonymat des personnes concernées**
- 3) **les fichiers de données individuelles constitués pour la réalisation de l'étude (échantillons, réponses) seront effacés à l'issue de l'enquête ou archivés dans des conditions préservant leur confidentialité.**

### **Recommandation n° 6 : garantir la confidentialité et l'anonymat**

L'anonymat est garanti à la fois par l'absence de données directement nominatives (le nom) et par l'absence de données permettant d'identifier indirectement la personne qui y répond (via un numéro renvoyant à l'identité de la personne, la désignation d'un poste particulier ou par recoupement d'informations telles que l'adresse, les date et lieu de naissance... ).

Or bon nombre de fichiers constitués dans le cadre de la recherche et des statistiques sont indirectement nominatifs. Le recueil de données identifiantes sur les personnes constitue un moyen de distinguer un cas par rapport à un autre, d'éliminer les doubles enregistrements d'autant plus fréquents que les informations sont susceptibles de provenir de sources d'information différentes, de rassembler et de vérifier sur une personne les renseignements obtenus, de suivre des trajectoires individuelles.

**Dès lors donc que le recours à des données identifiantes est justifié, il importe que toutes ces précautions soient prises pour préserver la confidentialité des données et garantir l'anonymat. Ainsi les statistiques produites doivent être suffisamment agrégées pour éviter tout risque de réidentification.**

La CNIL a toujours favorisé la mise en place et le développement de mesures de sécurité spécifiques, tout particulièrement pour les fichiers de recherche portant sur des données sensibles telles que la santé: chiffrement (cryptage), séparation des données d'identité et des renseignements sensibles, recours aux techniques d'anonymisation à la source des données d'identité<sup>58</sup>.

Elle estime que de telles mesures doivent aussi être développées dans le domaine de la recherche en sciences humaines, tout particulièrement lorsque des données sensibles sur l'origine des personnes sont recueillies.

**En conséquence elle estime que les pouvoirs publics doivent encourager, notamment dans le domaine de la statistique publique, un recours beaucoup plus systématique aux techniques de chiffrement et d'anonymisation.**

---

<sup>58</sup> Reposant sur l'utilisation d'algorithmes dits de hachage.

## **Recommandation n°7 : Assurer la transparence par l'information et l'exercice effectif des droits « Informatique et Libertés »**

Dès lors que des enquêtes sont réalisées par voie de questionnaires auprès des personnes ou par exploitation des fichiers de gestion et que des données directement ou indirectement nominatives sont recueillies, les personnes concernées doivent être parfaitement informées des objectifs poursuivis, des conditions de réalisation de l'enquête, (en particulier des destinataires des données), du caractère facultatif de celle-ci, ainsi que de leurs droits d'opposition, d'accès et de rectification qui peuvent être exercés tant que les données utilisées permettent leur identification.

La Commission constate que cette information, pourtant essentielle, est trop souvent négligée. Or, dans un domaine aussi sensible que la mesure de la diversité et la lutte contre les discriminations, elle constitue, au delà du respect de la règle de droit, un facteur clé pour assurer l'adhésion et la participation pleine et entière de chacun.

C'est la raison pour laquelle elle recommande de procéder à cette information non seulement lors du lancement de l'enquête mais également au cours de celle-ci.

De même elle préconise une consultation des instances représentatives du personnel en cas d'enquêtes menée dans le domaine de l'emploi.

De façon plus générale, il apparaît souhaitable que le lancement des enquêtes nationales soit annoncée publiquement de façon à sensibiliser la population sur ces questions .

### **B. Faut-il modifier la loi ?**

Le cadre juridique actuel est-il adapté ? Faut-il modifier le corpus législatif? Force est de constater que cette question<sup>59</sup> qui figurait dans le questionnaire transmis aux différentes personnes auditionnées a suscité relativement peu de réactions. Les quelques observations formulées<sup>60</sup> ont essentiellement porté sur la nécessité d'assouplir, dans le domaine de la recherche, les dispositions de l'article 8 de la loi informatique et libertés et sur l'opportunité ou non de définir par voie législative un référentiel ethno-racial.

---

<sup>59</sup> La question était la suivante : faudrait-il modifier le corpus législatif pour y intégrer les actions relatives à la « mesure de la diversité » ?

- 1) la Constitution (art. 1<sup>er</sup>, ...)
- 2) la loi « Informatique et Libertés » (art.8)
- 3) la loi sur la statistique publique (n° 51-711 du 07/06/1951)
- 4) la loi sur la lutte contre les discriminations (n°2001-1066 du 16/11/2001)
- 5) la loi pour l'égalité des chances (n°2006-396 du 31/03/2006)
- 6) le Code du travail (L.122-45,...)
- 7) le Code pénal (art.225-1,...)
- 8) créer une nouvelle loi pour la promotion de la diversité ?

<sup>60</sup> cf. audition de Roxane Silberman, directeur de recherches au CNRS

## **Recommandation n°8 : Assouplir, dans le domaine de la recherche, la dérogation relative au « consentement exprès » (art. 8-II) mais prévoir en contre partie l'autorisation de la CNIL**

A la suite des auditions réalisées, la Commission n'estime pas qu'une refonte fondamentale de l'article 8 serait utile. En effet, le principe d'interdiction de collecte et de traitement de données sensibles n'a pas été remis en cause. Toutefois, des propositions d'assouplissement et de précision, des dérogations ont été suggérées.

La nécessité de recueillir le consentement exprès (c'est-à-dire écrit) des personnes pour procéder au recueil et au traitement des données sensibles, hors les cas de dérogation, si elle permet certes « d'échapper » à l'autorisation de la CNIL apparaît aux yeux de certains chercheurs comme un dispositif trop lourd. De surcroît, le consentement exprès est de nature à induire des réactions de rejet, les personnes hésitant à formaliser leur consentement par une signature, craignant que cet acte ne les engage. Comme l'a indiqué Mme Silberman directeur de recherche au CNRS, lors de son audition, la loi est trop restrictive sur ce point, certes avec de bonnes intentions, et l'on a multiplié les freins à la réalisation d'enquêtes sur ces sujets sensibles et donc aussi à la lutte contre la discrimination. Elle a également estimé que *« des évolutions seront nécessaires sur le consentement exprès lui-même, mais également, en lien avec cette question, sur la façon peu claire dont est située dans la loi la recherche par rapport à la notion d'intérêt public : l'exception d'intérêt public concerne-t-elle la recherche ou pas »*.

**Faut-il assouplir la loi sur ce point ? Il est vrai qu'actuellement les chercheurs – et la CNIL - ont peu de choix : soit il est procédé au recueil du consentement exprès et l'enquête est alors soumise à simple déclaration, la CNIL n'ayant (quasiment) plus de droit de regard, soit la CNIL considère qu'elle revêt un intérêt public et elle est soumise à autorisation sans que le consentement soit alors requis<sup>61</sup>.**

**Cette situation n'est pas satisfaisante. On ne peut en effet considérer, comme il a été dit, que le consentement de la personne constitue réellement une garantie pour elle, en particulier lorsqu'il s'agit d'enquêtes réalisées dans les entreprises, dans la mesure où la relation de travail est de nature à « biaiser » le consentement. Par ailleurs, peut-on considérer que le critère de l'intérêt public - difficile à manier, comme la CNIL a maintes fois eu l'occasion de le constater- constitue un critère réellement opérant ?**

**Ne conviendrait-il pas dans le domaine de la recherche et des statistiques de soumettre l'ensemble de ces traitements à un régime d'autorisation (à l'instar de ce qui se fait pour la recherche médicale) et ne plus exiger le consentement exprès des personnes ?**

**Il s'agirait d'appliquer à l'ensemble des fichiers de recherche comportant des données sensibles le régime applicable aux traitements de données de santé ayant pour fin la recherche dans le domaine médical. Ces traitements sont soumis à autorisation de la CNIL, les personnes concernées disposant du droit de s'opposer (sans avoir à justifier de raisons légitimes) au traitement de leurs données (cf. article 56)**

---

<sup>61</sup> mis à part le régime dérogatoire dont bénéficient depuis la loi de 2004 l'INSEE et les services statistiques des ministères.

Il doit en outre être noté que l'article 54 de la loi du 6 janvier 1978 modifiée en 2004 prévoit une procédure spécifique d'autorisation dans le cadre de laquelle la CNIL se prononce sur le traitement en question suite à une appréciation technique par un comité consultatif. En effet, le Comité Consultatif sur le Traitement de l'Information en Matière de Recherche dans le Domaine de la Santé (CCTIRS), institué auprès du ministre chargé de la recherche et composé de personnes qualifiées ayant des compétences variées, « émet un avis sur la méthodologie de la recherche au regard des dispositions de la loi [« informatique et libertés »], la nécessité du recours à des données à caractère personnel et la pertinence de celles-ci par rapport à l'objectif de la recherche. » (art. 54, al.1)

« La mise en œuvre du traitement de données est ensuite soumise à l'autorisation de la Commission nationale de l'informatique et des libertés, qui se prononce dans les conditions prévues à l'article 25 »(art.54, al.4). Ce dernier vise la procédure de l'autorisation.

Ainsi qu'il a été indiqué précédemment (cf. recommandation sur les experts tiers de confiance), il pourrait être envisagé de s'inspirer de cette procédure pour l'élargir aux traitements des données sensibles ayant pour fin la recherche et les statistiques (et notamment la mesure de la diversité). Un comité consultatif, composé de statisticiens, sociologues, démographes, d'experts en étymologie des noms propres, etc., pourrait être créé à cet effet. Il apprécierait la pertinence du traitement, de la méthode utilisée, des données collectées, etc.

Son avis serait ensuite adressé à la CNIL qui disposerait ainsi de l'expertise nécessaire pour donner ou refuser son autorisation.

A la différence de la procédure mise en place pour la recherche médicale, il pourrait être envisagé que le recours au Comité ne soit pas systématique mais laissé à l'initiative des chercheurs ou de la CNIL .

**La Commission estime qu'il conviendrait de modifier des dispositions de la loi informatique et libertés et en particulier de ses articles 8 et 25 afin d'assurer une meilleure protection des données personnelles sensibles recueillies et traitées dans le cadre des études sur la mesure de la diversité et des discriminations, garantir le sérieux scientifique de ces études et harmoniser les régimes de formalités préalables applicables aux traitements de données sensibles mis en oeuvre à des fins de recherche.**

### **Recommandation n°9: Faut-il créer par la loi un référentiel national « ethno-racial » ?**

En 2005, la CNIL, à l'issue d'une première série de consultations, avait émis de fortes réserves sur la création d'une nomenclature nationale de catégories « ethno-raciales », estimant que le principe de la création d'un tel référentiel relevait du législateur. Le groupe de travail avait ainsi observé que le débat était ouvert au sein de la communauté statistique nationale et européenne et avait conclu à l'opportunité de rejoindre les positions de prudence de l'INSEE et de l'INED, favorables au maintien du droit en l'état, tout en étant conscient des limites d'une solution reposant sur la seule analyse de l'origine nationale des personnes.

Deux ans après cette première prise de position y-a-t-il lieu pour la CNIL d'évoluer ?

Quels sont les principaux points de vue exprimés ?

- L'ensemble des partenaires sociaux consultés qu'il s'agisse des organisations syndicales de salariés<sup>62</sup> ou du MEDEF<sup>63</sup> expriment de vives réserves sur l'établissement d'un tel référentiel<sup>64</sup>.

- De même, la plupart des chercheurs et statisticiens<sup>65</sup>, experts<sup>66</sup> et personnalités<sup>67</sup> auditionnés, y sont également hostiles, soulignant le risque accru de discrimination qui pourrait résulter d'une telle « catégorisation ». Un tel référentiel choquerait la sensibilité d'une partie importante de la population. Il figerait les représentations collectives et individuelles sur la base de fondements incertains et contingents au risque d'accentuer les clivages, les préjugés, voire les discriminations au lieu d'aider à les réduire. A cet égard, le Président de la HALDE rappelle la possibilité de réaliser des enquêtes approfondies qui étudient les distinctions et les discriminations dont peuvent faire l'objet certaines catégories de populations, telle l'enquête « Trajectoires et origine des migrants et de leurs descendants » de l'INSEE et de l'INED ( et à laquelle la HALDE participe).

- De nombreuses personnes auditionnées estiment aussi que les référentiels existant à l'étranger ne constituent pas des exemples très probants. Ils résultent d'un assemblage peu cohérent d'un grand nombre de critères qui mêlent couleur de peau, origine géographique, nationalité, héritage culturel, etc., Le référentiel britannique est jugé inadapté et inefficace car multipliant à l'excès les catégories et ne disant finalement rien de la réalité de la discrimination vécue. A été aussi soulignée<sup>68</sup> la difficulté de prendre en compte, dans ce type de grille, le métissage beaucoup plus important en France qu'au Royaume Uni.

- En revanche, certains chercheurs et experts<sup>69</sup> et les associations de défense des minorités visibles (tel le CRAN, Africagora, l'AFIP<sup>70</sup>) y seraient plutôt favorables. De telles

---

<sup>62</sup> CGT, CFDT, FO

<sup>63</sup> C.Kopp, directrice des ressources humaines du groupe ACCOR, soulignant en particulier le risque de communautarisation de la société française qui pourrait en résulter.

<sup>64</sup> la CFDT se déclarant toutefois ouverte à des expérimentations en ce domaine.

<sup>65</sup> cf. notamment le directeur de l'INSEE

<sup>66</sup> cf. notamment les représentants de l'Agence Nationale pour la cohésion sociale (ACSE)

<sup>67</sup> En particulier, L.Schweitzer, président de la HALDE, A. Bauer, président du conseil d'orientation de l'observatoire national de la délinquance, JP Dubois, président de la Ligue des Droits de l'Homme.

<sup>68</sup> cf. notamment C. Da Silva de l'AFIP.

<sup>69</sup> Yazid Sabeg (CS Communication), D.Sabbagh, directeur de recherche au CERI ; ce dernier estime qu'il serait opportun de pouvoir disposer de données précises sur à l'ampleur de la concordance entre les résultats produits par l'auto identification ethno-raciale des personnes et ceux produits par l'observation externe. Il conviendrait de commanditer des enquêtes en la matière pour savoir dans quelle mesure les identités sont cristallisées dans la société française D'autres études devraient être menées afin d'évaluer, d'une part, la légitimité du *principe* de la classification ethno raciale des personnes par les pouvoirs publics à des fins anti discriminatoires, d'autre part, la légitimité des *différentes modalités d'application de ce principe* – c'est-à-dire l'acceptabilité de telle ou telle

statistiques seraient de nature à favoriser l'intégration nationale et la diminution des discriminations. Si la réalisation de ces études est encouragée, ces associations préconisent que les réponses soient facultatives, favorisent l'auto déclaration et soient recueillies selon une procédure garantissant l'anonymat. Les associations en question se prononcent d'ailleurs en faveur de l'intégration des questions ethno raciales facultatives au sein du recensement. .

Pour certains<sup>71</sup>, les personnes interrogées doivent avoir la possibilité d'auto déclarer leur origine, à partir de quatre ou cinq groupes de population prédéfinis. A cette fin, il a été souligné qu'il serait alors nécessaire d "*homogénéiser la terminologie employée*". le changement passera nécessairement par un certain nombre de traumatismes, étant entendu que la décision ultime appartiendra au législateur.

### **Quelles conclusions tirer de cette diversité de points de vue ?**

**Force est de constater que sur ce sujet, les auditions n'ont pas permis de dégager un réel consensus en faveur de l'élaboration d'un tel référentiel, tout au contraire. Les personnes auditionnées sont dans leur grande majorité hostiles à une telle nomenclature ou, à tout le moins, pour certaines, partisans d'une démarche prudente et progressive. En tout état de cause, l'intervention d'un cadre législatif en la matière apparaît, aux yeux des personnes auditionnées, prématuré<sup>72</sup>.**

**Parallèlement, il semble qu'une certaine évolution des esprits se fasse jour sur cette question délicate**, comme en témoignent la multiplication des débats et des tribunes, les prises de positions de certaines personnalités politiques, les résultats de sondages et enquêtes enfin. Ainsi, si l'on se réfère aux résultats du sondage Eurobaromètre réalisé en juin-juillet 2006 auprès de 25 000 Européens (dont un millier vivant en France) 78 % de la population accepterait, en France, de communiquer des informations anonymes concernant son origine ethnique dans le cadre du recensement, mais parce que la question disait : « si cela pouvait aider à lutter contre les discriminations »<sup>73</sup>.

Comme l'a montré l'enquête exploratoire de l'INED sur la mesure de la diversité réalisée en novembre 2005, l'intégration de questions « ethno-raciales » dans les enquêtes scientifiques est acceptée sans réserve par 72% des personnes interrogées, 68% étant favorables à ce que de telles questions soient posées dans le cadre du recensement. Toutefois, les immigrés et leurs descendants directs semblent exprimer plus de réticences que la moyenne à répondre à des

---

catégorie proposée – auprès, d'une part, des membres des « minorités visibles », d'autre part, de l'opinion publique. Il serait de toute façon nécessaire de **passer par la voie législative**, la mise en place d'un dispositif de collecte de données sur la diversité « ethnique » de la population française représentant une rupture dont on ne peut nier l'importance.

<sup>70</sup> association créée en 2002, afin d'aider les jeunes diplômés issus des « minorités visibles » à trouver un emploi répondant à leurs attentes et au niveau de leurs compétences.

<sup>71</sup> C. Da Silva, de l'association AFIP

<sup>72</sup> Le directeur de l'INED, citant comme exemple, la nomenclature des catégories socioprofessionnelles, estime d'ailleurs qu'une nomenclature élaborée par la statistique publique n'a pas systématiquement vocation à devenir une référence juridique ; elle est avant tout une nomenclature de travail à géométrie variable susceptible d'être modifiée.

<sup>73</sup> Résultats corroborés par ceux de l'enquête mesure de la diversité de l'INED

questions « ethno-raciales » (12 %) alors que 96 % ne voient pas de difficultés à répondre à des questions sur l'ascendance ou l'origine géographique<sup>74</sup>. En revanche, les personnes immigrées sont majoritairement hostiles à l'intégration, dans les fichiers du personnel ou administratifs, de questions sur l'origine en particulier avec la nomenclature « ethno-raciale »<sup>75</sup>.

Ces résultats, très partiels, doivent être étayés par d'autres études. Comme le souligne P.Simon, (INED) ils doivent aussi être interprétés dans le contexte actuel français, où les références à l'origine sont trop souvent utilisées à des fins de stigmatisation.

Risques de renforcement des stéréotypes, de stigmatisation<sup>76</sup>, de communautarisme, classification incertaine, non scientifique, réductrice, approximative... autant de raisons, bonnes ou mauvaises, qui expliquent les réticences actuelles et qui justifient une attitude mesurée sur ce sujet.

La plupart des chercheurs interrogés sont d'ailleurs très prudents et même ceux qui seraient le plus enclins à recourir à une classification ethno-raciale reconnaissent la nécessité de démarches exploratoires pour tester l'acceptabilité de ce type de questions.

**Dès lors, l'idée d'une classification nationale « ethno-raciale » qui s'imposerait dans toute la statistique publique n'est pas aujourd'hui acquise.**

Au surplus il doit être rappelé que sur le plan législatif, deux projets d'amendements visant respectivement à obliger les entreprises et organismes publics à utiliser un cadre de référence<sup>77</sup> pour mesurer la diversité des origines au sein de leurs effectifs et à introduire dans le recensement une question sur les phénotypes ont été présentés au Sénat en février 2006 dans le cadre de l'examen du projet de loi sur l'égalité des chances.

Ces textes ont finalement été retirés, sur demande du Gouvernement, celui-ci estimant que des études fondées sur les noms de famille, l'origine géographique, le quartier d'habitation, ou des critères sociaux, pouvaient être conduites et qu'en la matière des problèmes juridiques se posaient tant au regard de la Constitution que de la loi informatique et libertés<sup>78</sup>.

---

<sup>74</sup> Il n'en est pas de même des « noirs » ce qui infirme sur ce point l'idée selon laquelle les minorités les plus exposées aux discriminations se défieraient le plus des catégorisations.

<sup>75</sup> 1/3 des personnes enquêtées se déclarent hostiles à l'enregistrement des catégories ethno-raciales, les rejets culminant chez les personnes se considérant comme arabe ou berbères (52 %).

<sup>76</sup> A été notamment cité le cas des statistiques ethniques de la délinquance.

<sup>77</sup> Ce référentiel aurait été défini conjointement par la HALDE, l'INSEE, l'INED.

<sup>78</sup>Extraits de l'intervention du ministre du travail, M Gérard Larcher : « plutôt que de recourir à une classification générale, il est possible de solliciter les organismes de recensement et d'études démographiques pour qu'ils utilisent pleinement les données autorisées, en particulier les noms de famille, l'origine géographique, le quartier d'habitation, ou des critères sociaux, pour suivre et appuyer l'évolution vers une plus grande égalité. Il est également possible que chaque entreprise définisse, en accord avec ses personnels, les meilleurs indicateurs d'évolution de ses effectifs ».

Le Gouvernement avait également estimé que, « en la matière, des problèmes de nature juridique se posent. Le recueil automatisé de données relatives à l'origine ethnique ou raciale des personnes, porte atteinte aux principes posés tout à la fois par l'article 1<sup>er</sup> de la Constitution, l'article 1<sup>er</sup> de la Déclaration des droits de l'homme et du citoyen et l'alinéa 5 du préambule de la Constitution de 1946.



Pour ces différentes raisons, il semble nécessaire de maintenir la position dégagée en 2005 et de rappeler que la Commission est d'avis que la décision de création d'une nomenclature nationale de catégories ethniques, qui serait utilisable, de façon obligatoire, en particulier pour les statistiques publiques et pour le recensement, appartiendrait au Législateur sous le contrôle du Conseil Constitutionnel.

Une telle classification devrait en outre être établie par les autorités scientifiques compétentes ainsi que par les représentants de la société civile.

**La Commission émet de fortes réserves sur la création d'une nomenclature nationale de catégories « ethno-raciales » et estime, en tout état de cause, que la décision de création d'une telle nomenclature, si elle devait être utilisée, de façon obligatoire, en particulier pour les statistiques publiques et pour le recensement, appartiendrait au législateur sous le contrôle du Conseil Constitutionnel.**

---

*Souvenons-nous, en 1991, le Conseil constitutionnel a déclaré inconstitutionnelle la mention de « peuple corse ».*

*Au surplus, les dispositions proposées ici ne seraient pas conformes à la loi. Je rappelle en effet les termes de l'article 8 de la loi du 6 janvier 1978 : « Il est interdit de collecter ou de traiter des données à caractère personnel qui font apparaître, directement ou indirectement, les origines raciales ou ethniques, les opinions politiques, philosophiques ou religieuses ou l'appartenance syndicale des personnes, ou qui sont relatives à la santé ou à la vie sexuelle de celles-ci ».*