

ÉGALITÉ, INÉGALITÉS, DISCRIMINATIONS

ESSAI DE DIALOGUE INTERDISCIPLINAIRE DES SAVOIRS JURIDIQUES ET QUANTITATIFS

p. 6 AVANT-PROPOS de Jérôme Porta et Christophe Bergouignan

p. 12 Debbie Collier

De l'Apartheid à l'action positive : un aperçu de la loi sur l'égalité au travail en Afrique du Sud

p. 24 Supriya Routh

Une Constitution à deux vitesses. Les travailleurs et l'égalité devant la loi en Inde

p. 36 Fábio Rodrigues Gomes et Roberto Fragale Filho

Le principe d'égalité dans le droit du travail brésilien. Brèves considérations propédeutiques sur l'article 461 de la CLT

p. 48 Laurène Joly

L'égalité à l'aune du handicap

p. 60 Benoît Riandey, Catherine Quantin, Anne-Marie Benoît et Olivier Cohen

Proposition d'un dispositif sécurisé de suivi statistique de la diversité dans les entreprises signataires de la Charte de la diversité

p. 68 Laure Mogueu, Tatiana Eremenko, Xavier Thierry et Rose Prigent

Les familles monoparentales immigrées : des familles doublement pénalisées ?

p. 84 Richard Marcoux

Le travail des enfants au Mali : tentative de mesure et inégalités de genre

p. 90 Dominique Meurs

Mesurer les discriminations : les approches des économistes

JURISPRUDENCE SOCIALE COMPARÉE

THÉMATIQUE : OBLIGATIONS IMPLICITES DE L'EMPLOYEUR ET CRÉATION PRÉTORIENNE :
DES EXEMPLES CONTRASTÉS EN DROIT COMPARÉ

p. 99 Allison Fiorentino

Elena Sychenko (Fédération de Russie)

Nanga Silue (Côte d'Ivoire)

Juan Pablo Mugnolo et Diego Ledesma Iturbide (Argentine)

ACTUALITÉS JURIDIQUES INTERNATIONALES

p. 128 Afrique du Sud

p. 130 Algérie

p. 132 Argentine

p. 134 Australie

p. 136 Autriche

p. 138 Brésil

p. 140 Chili

p. 142 Colombie

p. 144 Espagne

p. 146 États-Unis

p. 148 Fédération de Russie

p. 150 France - DT

p. 152 France - DSS

p. 154 Grèce

p. 156 Italie

p. 158 Lituanie

p. 160 ONU

p. 162 Pays Bas

p. 164 République du Congo
(Congo-Brazzaville)

p. 166 Roumanie

p. 168 Royaume-Uni

p. 170 Serbie

p. 172 Slovénie

p. 174 Turquie

p. 176 Union Européenne - DPS

p. 178 Uruguay

DOSSIER THÉMATIQUE

ÉGALITÉ, INÉGALITÉS, DISCRIMINATIONS
ESSAI DE DIALOGUE INTERDISCIPLINAIRE
DES SAVOIRS JURIDIQUES ET QUANTITATIFS



BENOÎT RIANDEY

Institut national d'études démographiques (INED).

CATHERINE QUANTIN

Département de l'information médicale du CHU de Dijon.

ANNE-MARIE BENOIT

Maison de Sciences de l'Homme Rhône-Alpes.

OLIVIER COHEN

Université Joseph Fourier, Grenoble.

PROPOSITION D'UN DISPOSITIF SÉCURISÉ DE SUIVI STATISTIQUE DE LA DIVERSITÉ DANS LES ENTREPRISES SIGNATAIRES DE LA CHARTE DE LA DIVERSITÉ



ABSTRACT

It appears to be quite difficult to verify compliance of non-discrimination and diversity promotion principles in the enterprises, in terms of recruitment, career management and job stability. Beyond the methodological concerns, it seems indeed particularly sensitive to give full account of these phenomenon on all aspects of career paths and for each enterprise signatory of the French Charter of Diversity. This paper will outline some engineering items in statistics in order to try to address these issues with statistical matching techniques. It will also put into perspective the evolution of the statistical matching techniques perception in connection with the law project for a digital Republic.

KEY WORDS : Diversity, discrimination, career, statistical matching techniques.

RÉSUMÉ

La vérification du respect des principes de non-discrimination et de promotion de la diversité dans les entreprises, en termes de recrutement, de gestion des carrières et de maintien dans l'emploi, s'avère particulièrement difficile. Au-delà des problèmes méthodologiques propres à la mise en place d'une collecte ponctuelle de données relatives à ces enjeux, il apparaît, en effet, particulièrement délicat de rendre compte de ces phénomènes pour l'ensemble des aspects des parcours professionnels et à l'échelle de chaque entreprise signataire de la charte de la diversité. Cet article propose quelques éléments d'ingénierie statistique pour tenter d'aborder ces enjeux à l'aide des méthodes d'appariement sécurisé et met en perspective l'évolution de la perception des méthodes d'appariement sécurisé dans le cadre du projet de loi pour une République numérique.

MOTS CLÉS : Diversité, discrimination, carrières, appariements sécurisés.

Les responsables des ressources humaines des entreprises se trouvent face à une situation qu'ont connue bien des statisticiens : devoir disposer de certaines statistiques sans être autorisés à collecter les données constitutives de ces statistiques.

De longue date, ce fut la situation de bien des administrations en raison de l'interprétation étroite de l'article 5c de la Convention 108 du Conseil de l'Europe selon laquelle « *Les données personnelles faisant l'objet d'un traitement automatisé sont ... adéquates, pertinentes et non excessives par rapport aux finalités pour lesquelles elles sont enregistrées* ».

Les institutions sociales suspectaient le fait que les familles immigrées ne bénéficiaient pas pleinement de leurs droits aux allocations. Mais, le repérage de ces populations leur était interdit par la loi car la collecte des informations « *ethniques* » nécessaires à cette évaluation était « *excessive par rapport à la finalité* » administrative, le versement des allocations. S'ajoutait l'interdiction constitutionnelle des traitements administratifs différenciés entre citoyens français selon leur origine. Ne restait alors que l'enquête par sondage pour mesurer ces discriminations, malgré sa capacité limitée à mettre en évidence des dysfonctionnements espérés très minoritaires.

La directive européenne du 24 octobre 1995, transcrite en droit français le 6 août 2004, reconnaît cependant la finalité statistique comme compatible avec la finalité initiale de collecte et complète l'article précédent par la mention « *ou pour ses traitements ultérieurs* » (notamment statistiques) (article 6c). La mesure statistique des inégalités d'accès aux droits et celle des discriminations par les administrations n'est donc plus jugée « *excessive* », sans qu'elles aient pour autant fréquemment mis en œuvre des dispositifs de cette nature. De fait, il n'en demeure pas moins que le recueil de données faisant apparaître les origines « *raciales ou ethniques* » est interdit sauf accord exprès de la personne (article 8.II.1), ou collecte par l'Institut national de la statistique et des études économiques (INSEE) et les services

statistiques ministériels après avis du Conseil national de l'information statistique (CNIS) (article 8. II.7), ou s'il est justifié par l'intérêt public et autorisé par la loi (article 8.IV).

Cette évocation relative aux administrations publiques souligne les difficultés de principe auxquelles les directeurs de ressources humaines se voient exposés ainsi que l'importance d'élaborer une solution à la fois sécurisée et reconnue très légitime. C'est l'objet de ce texte qui vise le suivi des parcours professionnels et les discriminations qui sont susceptibles de s'y manifester mais qui peut aussi bien s'appliquer aux administrations comme nous avons déjà eu l'occasion de l'exprimer¹.

Nous nous proposons de montrer qu'il est possible de collecter dans les entreprises, de façon anonyme des informations utiles et interdites aux directions du personnel et de les apparier anonymement à leurs données de gestion pour un traitement sécurisé et donc extérieur à l'entreprise.

À l'occasion de son enquête traitant de la violence familiale², le ministère de la Santé a expérimenté un dispositif de collecte pleinement confidentiel éprouvé à l'étranger³. Nous proposons de s'en inspirer en définissant un cadre institutionnel sécurisé et souple pour de tels traitements qui permettraient de relier les carrières intégrales (comportant phases d'emploi et de chômage) et les caractéristiques des personnes susceptibles d'engendrer des discriminations dans l'espace professionnel.

¹ C. Quantin et B. Riandey, « *Les techniques d'appariements sécurisés. De l'épidémiologie à la démographie OU La démographie dans le sillage des épidémiologistes* », Chaire Quetelet 2006, Louvain la Neuve, Belgique.

² C. Cavalin, « *Événements de vie et santé : principes et méthodes* » in Guilbert Ph. et al, « *Méthodes de sondages* », Dunod, Collection Sciences sup, 2008. Actes du 5^{ème} colloque francophone sur les sondages.

³ Rogers SM et al., (1999) « *Entretiens auto-administrés sur ordonnance et mesure des comportements sensibles* », *Population* n° 2.

I - Une collecte respectant la confidentialité

Les informations constitutives de la diversité sont pour une part des informations personnelles et sensibles qu'on peut ne pas désirer fournir à son employeur ou futur employeur potentiel. Les réticences à se décrire sous cet aspect ne peuvent être dépassées que pour un motif solidaire reconnu très légitime et sans risque de gêne comme de contrepartie négative. Décrire à quelqu'un ses handicaps est blessant sans même évoquer l'utilisation faite de l'information donnée. Les questionnaires enregistrés et auto-administrés sur ordinateur évitent ces réticences : aucun enquêteur n'entend ou ne lit la réponse, ni même ne pose directement la question à la personne concernée. La question ne s'inscrit pas sur l'écran ; la réponse non plus. Reste à assurer l'anonymat du questionnaire et la confidentialité est totale sous réserve que le recoupement des réponses n'identifie pas indirectement la personne.

En la circonstance, comment motiver les répondants pressentis ? Par ailleurs, l'anonymat est-il compatible avec la finalité de la collecte ? Ce n'est pas notre propos de définir quelles questions relatives à la diversité devraient être posées dans les entreprises pour lutter contre les discriminations. Nous n'entrerons pas dans cette analyse relevant davantage de la recherche sociologique que de l'ingénierie statistique ici proposée.

La collecte de ces informations sensibles requiert motivation et confiance. L'entreprise ne saurait de sa seule autorité assurer le climat favorable, même pour qui est sensible à l'exigence de justice qui préside à la démarche et prête confiance au service des relations humaines. On n'imagine pas que la loi puisse rendre la réponse obligatoire – cette obligation serait d'ailleurs vaine – elle peut juste la rendre licite, inciter ou obliger les entreprises à la solliciter, lui apposer une reconnaissance morale.

Un tel questionnaire requiert de façon très publique le label de la HALDE (désormais Défenseur des Droits) et l'approbation de la CNIL pour pouvoir être présenté dans l'entreprise, éventuellement par l'entreprise et plus précisément son responsable de la diversité.

Doit-il être rempli par tous ou seulement par ceux qui se sentent en position de subir une discrimination ? Retenons l'idée que proposer à tous de remplir le questionnaire est une prise de position active pour le respect de la diver-

sité et une condition nécessaire pour que les personnes les plus concernées puissent sans marquage répondre au questionnaire. C'est une forme d'engagement collectif du personnel. Néanmoins, l'assurance d'une possibilité de non réponse discrète est sans doute de nature à éviter des questionnaires fantaisistes perturbateurs et malveillants. L'utilité statistique d'un fort taux de réponse global est une toute autre question que nous examinerons plus loin.

II - Appariement et traitement sécurisé

Si le but unique de ce dispositif était de compter les personnes ressortissant de la diversité, l'analyse isolée de questionnaires anonymes satisferait à son objectif. Mais, ce n'est pas son seul objectif. Ces données doivent être rapprochées de données individuelles de carrière ou de recrutement en sorte de détecter s'il y a des comportements discriminatoires au sein de l'entreprise. Une identification du répondant s'avère donc nécessaire. Une contrainte qui peut paraître contradictoire avec l'exigence d'anonymat que nous avons posée. Précisément pas et c'est l'objet même et l'originalité du dispositif proposé. Le département d'information médicale du CHU de Dijon s'est fait une spécialité de régler ce type de difficulté dans le respect du secret médical, du secret statistique et de la loi Informatique et Libertés¹. Le principe repose sur l'idée que l'identifiant de la personne est crypté de façon irréversible de façon permanente grâce à une clé de chiffrement inchangée. L'algorithme utilisé, le hachage SHA compresse l'information contenue dans la chaîne de caractères identifiante en sorte qu'il n'existe aucune fonction mathématique de retour vers l'identifiant.

La permanence de cet identifiant haché permet de rapprocher les données relatives à la même personne, mais dans l'anonymat total. Le hachage avec une autre clé sécurisée empêche tout rapprochement entre l'identifiant haché obtenu et l'identifiant initial directement ou indirectement nominatif.

¹ C. Quantin, B. Gouyon, F.A. Allaert, O. Cohen, « *Méthodologie pour le chaînage de données sensibles tout en respectant l'anonymat : application au suivi des informations médicales* ». Dossier spécial « panel et appariements sécurisés », *Courrier des statistiques* n°113-114, juin 2005 et *Journal de la Société française de statistique*, vol 146 n°3, 2005 coordonnés par B. Riandey et MH Gensbittel.

L'identifiant du monde du travail est le NIR (numéro d'inscription au répertoire, plus connu sous l'appellation numéro INSEE, alias numéro de sécurité sociale) validé par un nombre clé². Les extraits de fichiers de carrière à apparier aux questionnaires diversité seraient donc identifiés par le NIR haché par une clé C1 spécifique à l'entreprise.

Le répondant du questionnaire diversité s'identifierait par son NIR (qui n'apparaîtrait même pas à l'écran), immédiatement haché par la même clé C1 en sorte de permettre l'appariement, et immédiatement ré-haché avec une clé C2 en sorte que le service des relations humaines ne puisse pas identifier la personne par hachage des NIR du fichier du personnel avec C1. Les questionnaires assortis de leurs identifiants cryptés par double hachage seraient alors transférés sans risque à un prestataire statistique. Celui-ci hacherait une troisième fois l'identifiant avec une clé C3 pour éviter tout risque de collusion avec l'équipe informatique de l'entreprise. En temps opportun, le prestataire déchiffrerait le questionnaire (mais pas l'identifiant) pour entreprendre les analyses statistiques. Quand il recevrait le fichier de carrière, le prestataire hacherait deux fois l'identifiant (avec les clés C2 puis C3) et apparierait alors les informations relatives à la même personne dans les deux fichiers.

III - Précautions d'usage du Numéro INSEE

Le NIR véhicule beaucoup de craintes et de fantasmes tant son utilisation incontrôlée serait potentiellement génératrice de dangers pour les libertés individuelles. Dans le dispositif proposé, le NIR n'est durablement stocké que dans l'esprit de la personne concernée. Il est immédiatement crypté par le premier hachage dès qu'il est contrôlé par son nombre-clé. Les informaticiens peuvent prendre les précautions de rigueur pour éviter son piratage avant hachage.

D'un point de vue légal, la CNIL précise bien qu'un NIR haché n'est pas un NIR. Cela l'est d'autant moins lorsque

² Ce nombre-clé, compris entre 1 et 97 est la différence à 97 du reste de la division du numéro de sécu par 97, le plus grand nombre premier à deux chiffres. C'est donc « la preuve par 97 » de l'exactitude du numéro de sécurité sociale. En d'autres termes, ce nombre clé détecte 96 erreurs sur 97... comme autrefois la preuve par 9 de nos multiplications ou divisions détectait 8 erreurs de calcul sur 9.

ce NIR est doublement haché. De fait, il n'est pas une donnée personnelle car il est strictement anonyme ; à ce titre il n'est pas assujéti à l'obligation du décret en Conseil d'État après avis motivé de la CNIL qu'impose l'article 27 de la loi Informatique et Libertés modifiée en 2004 pour l'utilisation en clair. Cette remarque ne nie pas pour autant l'avantage porté par l'examen du Conseil d'État au dispositif proposé tant l'aval des autorités supérieures de la République apporte de confiance dans les nouveaux dispositifs mis en œuvre. Une validation formelle et plutôt générale par le Conseil d'État des garanties apportées par les techniques de hachage favoriserait par ailleurs beaucoup la diffusion de cette sécurité informatique, y compris dans d'autres domaines.

IV - Secret statistique

Ces précautions suffisent-elles à garantir le respect du secret statistique ? Il semble raisonnable de le penser. Cependant, un fichier sans identifiant peut permettre d'identifier certaines personnes par recoupement d'informations. De même, un tableau statistique se référant à de très petits groupes (petites PME, services d'une entreprise) peut conduire à divulguer des informations personnelles, on parle alors de données indirectement nominatives.

Deux précautions permettent de contrer ces risques : la première consiste à limiter la précision de l'information contenue dans les fichiers pour en réduire le caractère indirectement identifiant ; la seconde réside dans l'organisation institutionnelle et juridique du prestataire chargé de l'appariement et de l'exploitation des données ainsi obtenues.

Rappelons que la loi autorise l'accès à des données personnelles (directement ou, comme dans le dispositif proposé, potentiellement indirectement nominatives) pour une finalité statistique. Les statisticiens sont attachés au respect du secret statistique, l'un des fondements de leur code éthique, sans même évoquer la dissuasion apportée par le code pénal. Néanmoins, on doit veiller à éviter au prestataire un détail d'information excessif par rapport au traitement à opérer. Il ne s'agit pas ici de déterminer, a priori, la description exacte des informations à lui transmettre. D'autant que ces informations ne sauraient être standardisées dans la mesure où la structure et la complexité des carrières potentielles varie selon la taille de l'entreprise ou du groupe, ce qui appelle une analyse de la part des experts du métier en coopération avec le prestataire.

On devra s'intéresser aussi à la précision du diagnostic apporté à l'entreprise. Si le diagnostic est favorable (du type d'une mention « rien à signaler : RAS »), aucun risque de divulgation ne se présente, que l'entreprise accueille ou non des représentants de la diversité. Si le diagnostic est une mise en garde ou une mise en cause, la question se pose, mais c'est un problème général à traiter : peut-on se contenter de dire qu'il y a un problème, sans permettre d'identifier le problème, la question dépasse le cadre du secret statistique.

V - Population enquêtée

Nous n'avons pas encore évoqué le public appelé à répondre au questionnaire recueillant les variables associées à la diversité, qui constitue le cœur du dispositif. Cette question recevra une réponse différente selon qu'il s'agit seulement de détecter des discriminations ou de mesurer la diversité en elle-même. L'idée serait que les entreprises qui ont signé la charte pour la diversité soient tenues à mettre en œuvre ce dispositif. On pourrait cependant proposer un cadre plus large.

Quoi qu'il en soit, pour le seul diagnostic de discriminations internes à l'entreprise, le dispositif est fait pour apporter des variables individuelles indicatrices de la diversité qui seront greffées (appariées) aux enregistrements anonymisés extraits des fichiers de carrière. Du strict point de vue de l'analyse statistique relative à ce diagnostic, il ne serait pas nécessaire que les salariés ne considérant pas comme potentiellement victimes de discriminations liées à leur origine remplissent le questionnaire. Un objectif plus large du questionnaire et la politique de communication liée à ce dispositif pousse cependant à proposer le questionnaire à tout le personnel, même si un taux de réponse très partiel des personnes non concernées n'induit pas de biais dans l'analyse.

Dans le cadre des procédures de recrutement, l'ensemble des candidats devraient répondre au questionnaire recueillant les variables associées à la diversité, par exemple en parallèle du premier contact en face à face. En revanche, le dispositif proposé ne permet pas de contrôler la sélection initiale des candidatures. En ce sens, il est complémentaire du CV anonyme. Pôle emploi est l'exemple même de l'administration qui devrait systématiquement mettre en place un tel dispositif et encourager son développement.

S'il s'agit de mesurer la diversité dans les entreprises ou dans la population, on doit alors soumettre le questionnaire à un échantillon représentatif de la population, éventuellement dans le cadre de la statistique publique.

VI - Le prestataire : un hébergeur de données sensibles ?

La nécessité que l'appariement et le traitement soient menés à l'extérieur de l'entreprise par un prestataire apparaît comme évidente. Certains ont rapporté l'idée que cette mission pourrait être confiée à des organismes publics comme l'INSEE, un service statistique ministériel ou l'Institut national d'études démographiques (INED).

Nous ne plaidons pas du tout en ce sens. Bien sûr, la loi de juin 1951 relative à la statistique publique apporte des garanties spécifiques quant au respect du secret statistique. Cependant, le législateur peut absolument spécifier ces garanties pour d'autres institutions et c'est ce que nous proposons. Ce n'est pas la fonction de ces établissements de réaliser un travail à façon pour les entreprises. Ils seraient vite dépassés sous la tâche. Tout au plus, il pourrait leur revenir de récupérer les résultats pour opérer une synthèse statistique, mais à un niveau global. Il s'agirait alors davantage d'utiliser les données pour rendre compte des phénomènes à une échelle plus large ou pour évaluer le fonctionnement du dispositif et non pour le mettre en œuvre comme nous le proposons ici. Par ailleurs, le fait que l'opérateur ne soit pas un service public pourrait apporter de la souplesse dans cette mise en œuvre.

Nous proposons de rapprocher cette situation de celle qui a prévalu le 9 août 2004 quand la loi de santé publique a défini le statut d'hébergeur de données de santé, qui s'est trouvé au cœur du dispositif du Dossier Médical Personnel. Il s'agissait là de la seule fonction de base de données ; celle-ci figure dans le dispositif ici proposé, les fichiers de gestion n'étant accueillis que de façon très temporaire. Il s'y ajouterait ici une fonction de traitement informatique et statistique. Ces hébergeurs de données de santé sont probablement de bons candidats pour le dispositif ici défini puisqu'ils apportent des garanties de sécurité pour des données au moins aussi sensibles. De plus, ils commencent à acquérir une expérience dans le champ des questions extérieures au domaine strictement médical. Ainsi, l'hébergeur de données de santé HC Forum s'est

engagé sur l'Observatoire de l'enfance en danger, question sociale elle aussi très sensible. Il semble, par ailleurs, justifié d'élargir le champ des prestataires à des entreprises qui seraient par la loi dotées d'un statut plus étroit d'hébergeur de données sensibles mentionnant particulièrement les obligations du secret statistique et les peines afférentes à son non-respect. VII - Agrément des hébergeurs de données et coordination du développement de leur activité Le statut d'hébergeur de données (sociales) sensibles nous semble devoir être défini par la loi en sorte que son agrément fonde la confiance des différents acteurs intervenant avec l'hébergeur dans le dispositif (salariés et candidats au recrutement, entreprises et établissements). Ce n'est pas tant qu'il traitera de données très risquées, compte tenu de l'anonymisation antérieure des données, mais un engagement au secret statistique calqué sur la loi de juin 1951 apparaît une nécessité éthique. Par ailleurs, la rigueur du dispositif informatique de collecte, de transfert et de cryptage des données, installé dans l'entreprise (ou validé) par l'hébergeur est le fondement de la sécurité du dispositif.

Ce métier d'hébergeur est un nouveau métier de la statistique à organiser et à coordonner. Ce peut être dans un cadre associatif ou de façon plus contrainte. Il doit bénéficier d'une veille mathématique, en particulier en cryptographie, et surtout d'un lien méthodologique étroit avec les professionnels des ressources humaines.

VIII - Un dispositif proposé au COMEDD

Né d'une approche pluridisciplinaire, le dispositif ici proposé a été présenté lors des auditions organisées par le Comité pour la Mesure et l'Évaluation de la Diversité et des Discriminations (COMEDD). Au-delà de la description du dispositif qui pourrait être mis en œuvre nous avons particulièrement insisté sur le rôle potentiellement protecteur des libertés publiques des appariements sécurisés par hachage irréversible du NIR pour favoriser la confiance du public dans le dispositif à mettre en place. Dans cette optique, le COMEDD pourrait inciter la CNIL et le Conseil d'État à prendre publiquement position en faveur de ce dispositif afin que cet appui puisse être présenté aux entreprises et aux personnes à l'occasion de la mise en œuvre éventuelle de la collecte, ce qui permettrait de lever les doutes pouvant émerger quant à la finalité strictement statistique et la confidentialité du dispositif. Le COMMED, en relation avec le Défenseur des Droits, pourrait proposer au Gouvernement

d'instituer, par la loi ou par décret, le statut d'entreprise d'hébergeur de données sensibles, s'inspirant de façon simplifiée du statut d'hébergeur de données médicales. Ce statut imposerait l'obligation de compétence de l'entreprise en matière de sécurité informatique et d'appariements sécurisés et de démo-statistique des ressources humaines en entreprise. Un décret devrait régir la procédure d'agrément de ces entreprises et leur suivi.

Nous avons aussi rappelé l'importance de la rédaction des questionnaires de recueil des variables indicatrices de la diversité. Ainsi, le COMMED pourrait inciter le Défenseur des Droits à définir des priorités en matière d'observation des discriminations d'embauche ou de déroulement de carrière, et dans cette logique veiller à la rédaction et au test de protocoles et questionnaires adaptés à cette mesure, à la présentation à donner au processus de collecte.

Le COMMED a retenu une large part du dispositif proposé questionnaire sensible validé par le Défenseur des Droits et la CNIL, traitement extérieur par une entreprise agréée. Par contre, elle met en avant une solution simplificatrice : c'est à la date de l'analyse que le salarié compléterait son extrait de carrière par le questionnaire du Défenseur des Droits aussitôt adressé à l'hébergeur de données sensibles. L'appariement sur le NIR haché en devient superflu. Le questionnaire du Défenseur des Droits serait donc administré aux salariés toujours présents dans l'entreprise. On serait donc exposé au biais d'échantillonnage des enquêtes rétrospectives : le questionnaire du Défenseur des Droits ferait défaut pour les salariés sortis de l'entreprise par décès, retraite, démission ou licenciement. La validité de cette procédure suppose que les sorties d'une entreprise soient indépendantes de tout processus de discrimination, ce qui constitue une hypothèse pour le moins problématique.

IX - Les appariements sécurisés, des dispositifs d'avenir

Bien que le volet du dispositif relatif aux appariements sécurisés n'ait pas été retenu par le COMEDD, il nous semble important de rappeler le potentiel qu'il recèle pour la statistique, et ce, bien au-delà de la seule proposition faite au COMEDD. Un dispositif basé sur les mêmes principes que celui proposé pourrait, en effet, être adapté à d'autres administrations assurant des services et pres-

tations aux populations, comme les Caisses d'Allocations Familiales. On signale d'ailleurs, que dans son rapport de l'année 2000, la CNIL recommandait au Centre de recherche pour l'étude et l'observation des conditions de vie (CREDOC) d'y recourir pour construire l'Observatoire du RMI de la ville de Paris. En interprétant cette recommandation dans une perspective plus large on peut conclure que ce qui était bon pour l'observation du RMI, il y a quinze ans, ne peut que l'être pour l'ensemble de la statistique publique. C'est la raison pour laquelle nous préconisons la généralisation de ce type de dispositifs bien au-delà des quelques domaines où ils ont déjà été mis en œuvre (éducation nationale, VIH-SIDA,...). Pour autant, l'intérêt exprimé par les cadres de l'INSEE ne s'est pas encore matérialisé, sans doute parce que l'institut est autorisé à utiliser le NIR en clair (non crypté) ce qui implique peut-être une moindre urgence à mettre en place des appariements sécurisés.

Dans le passé, les services statistiques étaient inévitablement tributaires des seules données de gestion. Aujourd'hui, les administrations pourraient stocker dans leurs bases de gestion ces informations sensibles cryptées dès leur saisie par l'intéressé et uniquement déchiffrables par des statisticiens extérieurs. Ce serait une belle avancée de l'utilité sociale de la statistique et de la technologie du secret statistique. Le projet de loi pour une République numérique transmis au Conseil d'État le 6 novembre 2015 va d'ailleurs clairement dans ce sens en s'appuyant sur de nombreuses contributions soutenues par les experts de la statistique et de la recherche publique.