

La Vice-Présidente déléguée

Madame Magda TOMASINI
Directrice
INSTITUT NATIONAL D'ÉTUDES
DÉMOGRAPHIQUES
INED
133 BOULEVARD DAVOUT
75980 - PARIS CEDEX 20

LETTRE RECOMMANDEE AVEC AR

Notification de Délibération

Paris, le

0 4 MARS 2020

N/Réf. : SLE/VGR/NDA201001

DEMANDE D'AVIS N°2214563

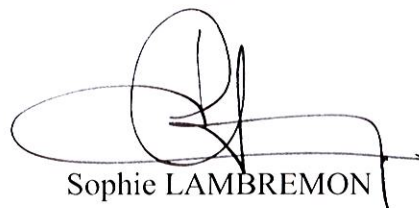
A rappeler dans toute correspondance

Madame,

Vous avez adressé à la Commission nationale de l'informatique et des libertés une demande d'avis portant sur un projet de traitement automatisé de données à caractère personnel nécessaire à la réalisation d'une enquête dénommée « Immigrants chinois à Paris et en région parisienne » (CHIPRE).

Vous trouverez ci-joint l'avis de la CNIL en date du 27 février 2020.

Je vous prie d'agréer, Madame, l'expression de mes salutations distinguées.



Sophie LAMBREMON

RÉPUBLIQUE FRANÇAISE

3 Place de Fontenoy, TSA 80715 - 75334 PARIS CEDEX 07 - 01 53 73 22 22 - www.cnil.fr

Délibération n° 2020-027 du 27 février 2020 portant avis sur un projet de traitement automatisé de données à caractère personnel nécessaire à la réalisation d'une enquête dénommée « Immigrants chinois à Paris et en région parisienne » (CHIPRE)

(demande d'avis n° 2214563)

La Commission nationale de l'informatique et des libertés,

Saisie par l'Institut national d'études démographiques d'une demande d'avis concernant un projet de traitement automatisé de données à caractère personnel nécessaire à la réalisation d'une enquête dénommée « Immigrants chinois à Paris et en région parisienne » (« Chinese immigrants in the Paris Region » ou « CHIPRE ») ;

Vu la convention n° 108 du Conseil de l'Europe pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel ;

Vu le règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE ;

Vu la loi n° 78-17 du 6 janvier 1978 modifiée relative à l'informatique, aux fichiers et aux libertés, notamment le 6° de son article 44 ;

Vu le décret n° 2019-536 du 29 mai 2019 modifié pris pour l'application de la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés ;

Sur la proposition de Mme Anne DEBET, commissaire, et après avoir entendu les observations de Mme Nacima BELKACEM, commissaire du Gouvernement,

Émet l'avis suivant :

1. À titre liminaire, la Commission relève que l'enquête dénommée « Immigrants chinois à Paris et en région parisienne » (CHIPRE) vise à fournir des données permettant une description de la population chinoise en Ile-de-France et à mettre en évidence les aspects cruciaux d'assimilation des immigrants dans le contexte français.

2. Dans ce contexte, elle relève que le traitement projeté poursuit des finalités de recherche scientifique. Elle estime à cet égard que le traitement relève du champ d'application du règlement (UE) 2016/679 du 27 avril 2016 susvisé (ci-après le « RGPD ») et prend acte que l'Institut national des études démographiques (INED) entend se prévaloir de la base licéité mentionnée au e) du 1. de l'article 6 du RGPD, à savoir de « l'exécution d'une mission d'intérêt public ».

3. Dans la mesure où il porte sur des données sensibles au sens de la réglementation, le traitement projeté doit faire l'objet d'un avis préalable de la Commission conformément aux dispositions du 6° de l'article 44 de la loi du 6 janvier 1978 modifiée, sans préjudice de l'obligation de formaliser le cas échéant la création du traitement par

— RÉPUBLIQUE FRANÇAISE —

3 Place de Fontenoy, TSA 80715 - 75334 PARIS CEDEX 07 - 01 53 73 22 22 - www.cnil.fr

un acte réglementaire spécifique. Elle rappelle qu'elle devra être tenue informée et saisie de toute modification substantielle affectant les caractéristiques du traitement et que l'analyse d'impact relative à la protection des données (AIPD) devra faire l'objet d'une mise à jour.

Sur le contexte général de réalisation de l'enquête (« CHIPRE »)

4. Le projet d'enquête vise à collecter, à partir d'un échantillon de 800 personnes, des données auprès de personnes nées en Chine, vivant depuis au moins trois mois en Ile-de-France, âgées de plus de 18 ans et de moins de 65 ans au moment de l'enquête, à l'exception des étudiants titulaires d'une bourse du gouvernement chinois (en ce qu'ils sont amenés à retourner en Chine immédiatement à la fin de leurs études).

5. La Commission relève que cet échantillon sera constitué à partir de dix enquêtes « graines », appelé « échantillon de convenance ». Les critères de choix de cet échantillon découleront des conclusions de l'exploration de terrain menée par l'INED préalablement à l'enquête elle-même, qui investigate différents réseaux et s'appuie sur son expérience lors d'autres enquêtes (associations, commerçants, famille).

6. Ce projet d'enquête est scindé en plusieurs volets :

- une enquête exploratoire, qui ne vise pas à recueillir de données à caractère personnel mais à établir une pré-connaissance qualitative de la structuration du réseau ;
- la passation des questionnaires ;
- la passation des registres (« rosters »).

7. Elle relève également que l'INED envisage la passation d'une trentaine d'entretiens qualitatifs *a posteriori*, lesquels sont exclus de la présente saisine. La Commission, qui prend acte que ces entretiens devraient faire l'objet d'un enregistrement audio, rappelle qu'il appartiendra au responsable du traitement de s'assurer de la conformité des traitements de données à caractère personnel mis en œuvre dans ce cadre à la réglementation applicable.

8. Elle relève en outre que pour la réalisation de l'enquête projetée, l'INED va produire plusieurs fichiers, à savoir :

- deux fichiers bruts de collecte : l'un correspondant aux réponses au questionnaire, l'autre contenant les données du « registre » ;
- les fichiers d'étude : correspondant aux fichiers bruts apurés de toute référence à des coordonnées et de données nominatives. Il s'agit alors de fichiers pseudonymisés ;
- le « *fichier de contacts quali* » : ce fichier contient les données des personnes qui ont accepté le principe des entretiens-qualitatifs ;
- le fichier de production et de recherche : fichier pseudonymisé qui sera mis à disposition des chercheurs via le réseau Quetelet ;
- le fichier des entretiens qualitatifs.

9. L'échantillonnage sera réalisé *via* la méthode « *Network Sampling with Memory (NSM)* » : les personnes interrogées sont sollicitées pour lister au maximum six

membres de leur entourage ou de leur réseau de sociabilité correspondant aux critères d'éligibilité de l'enquête et fournir des informations permettant de les identifier partiellement (le « registre » ou « roster »).

10. La Commission relève que les données collectées permettent de reconstituer le réseau social sous-jacent de la population cible et que l'emploi de la méthode « NSM » combine deux modes d'échantillonnage mis en œuvre au moyen d'un algorithme programmé par les services de l'INED :

- le mode « recherche » qui est utilisé pour explorer le réseau, sélectionne prioritairement les membres du réseau de sociabilité des « nœuds-ponts » (répondant ayant un nombre anormalement élevé d'amis et de connaissances au sein de son réseau de sociabilité qui sont listés dans le « registre » et qui n'ont pas été nommés par un autre répondant au sein de l'échantillon) non cités par ailleurs ;
- le mode « liste » qui permet que tous les individus nommés dans l'enquête se rapprochent de la base de sondage constituée par l'ensemble des membres de la composante connectée (en réseau) de la population cible.

11. La Commission observe par ailleurs qu'un chèque-cadeau d'un montant de 15 euros sera remis à la fin de la passation du questionnaire pour dédommager les enquêtés du temps qu'ils y ont consacré ainsi qu'un chèque-cadeau d'un montant de 20 euros en cas de diffusion de l'information relative à l'enquête par le répondant à son réseau de sociabilité. Elle rappelle que la remise des chèques-cadeaux, qui ne saurait être conditionnée à la collecte de données à caractère personnel, a uniquement vocation à dédommager le répondant pour le temps consacré et est sans incidence sur la base légale mobilisée pour le traitement projeté (« *exécution d'une mission d'intérêt public* »).

12. La Commission constate enfin que, ne s'agissant pas d'une enquête obligatoire, l'accord des personnes concernées sera demandé oralement par les enquêteurs avant le début de la passation du questionnaire.

Sur les finalités du traitement

13. L'enquête CHIPRE a pour objectifs principaux de :

- remédier au manque de données quantitatives détaillées sur les Chinois nés en Chine et résidant en France et ainsi constituer une base de données quantitatives permettant d'enrichir les connaissances sur les Chinois vivant à Paris et en région parisienne ;
- mesurer l'hétérogénéité de cette population ;
- explorer les liens entre les populations chinoises d'origines géographiques différentes en Chine et les populations non-chinoises, de même que le rôle des réseaux de sociabilité dans la recherche d'emploi et l'adaptation à la société d'accueil ;
- tester l'existence d'un modèle d'adaptation sélective des membres de cette population notamment en fonction de la région d'origine en Chine ;
- étudier les déterminants et perceptions de discriminations éventuelles intra et extra-communautaires (au sein de la société française).

14. La Commission, qui prend acte des précisions apportées selon lesquelles le projet d'enquête ne poursuit que des finalités de recherche scientifique, considère que les finalités poursuivies par le traitement projeté sont déterminées, explicites et légitimes, conformément aux dispositions du b) du 1. de l'article 5 du RGPD.

Sur les données traitées

15. En premier lieu, la Commission observe que les finalités poursuivies impliquent la collecte de nombreuses données à caractère personnel de natures très variées auprès des personnes enquêtées dont certaines relèvent de la catégorie des données dites « sensibles » au sens de l'article 9 du RGPD, notamment des données concernant la santé et des données qui révèlent les origines ou les convictions religieuses.

16. Elle rappelle que, dans la décision n° 2007-557 DC du 15 novembre 2007, le Conseil constitutionnel a jugé que *« si les traitements nécessaires à la conduite d'études sur la mesure de la diversité des origines des personnes, de la discrimination et de l'intégration peuvent porter sur des données objectives, ils ne sauraient, sans méconnaître le principe énoncé par l'article 1^{er} de la Constitution, reposer sur l'origine ethnique ou la race »*.

17. La Commission observe que le projet d'enquête, qui vise principalement à collecter des données objectives, ne repose pas sur la prétendue origine raciale ou ethnique des personnes interrogées. Si certaines questions telles que celles figurant dans les sections relatives aux emplois passé et actuel et celles relatives aux expériences de discrimination impliquent la collecte de données subjectives, elle estime que cette enquête n'a pas pour objet, même indirectement, de classer les personnes interrogées en fonction soit de leur prétendue origine ethnique ou raciale déclarée, soit d'un référentiel ethno-racial. Dans ces conditions, la Commission considère que les questions posées ne sont pas contraires à la décision précitée du Conseil Constitutionnel.

18. En deuxième lieu, elle relève qu'il est prévu de collecter auprès des répondants de l'enquête principale des données relatives à de nombreuses personnes de leur entourage, à la fois des proches (par exemple les enfants) mais également six personnes de leur réseau de sociabilité, notamment de contact, afin que l'INED puisse étendre l'échantillon de l'enquête et décrire la nature du réseau de sociabilité du premier répondant.

19. Si la Commission prend acte des raisons invoquées au soutien d'une telle collecte, à savoir l'absence d'une base de sondage nationale permettant de tirer aléatoirement les personnes appartenant à la population cible, elle considère au regard de la finalité poursuivie par le traitement projeté, que seules les données d'identité, de contacts et de relation avec le premier répondant sont nécessaires pour que l'enquêteur contacte directement la personne concernée dans le délai maximal d'un mois à compter de la collecte auprès du premier répondant.

20. Ces éléments rappelés, la Commission estime, qu'il reviendra à l'INED de contacter le membre de réseau de sociabilité du premier répondant pour l'informer, y compris, lorsque l'enquêteur devra lui indiquer qu'il n'est pas éligible à l'échantillonnage.

Auquel cas, les données à caractère personnel le concernant devront être supprimées à l'issue de la communication.

21. La Commission prend par ailleurs acte que le répondant ne recueille pas le consentement des membres de son réseau de sociabilité au traitement de leurs données par l'INED, mais uniquement leur accord à être contactés par un enquêteur dans le cadre de la méthodologie scientifique retenue. Il résulte de ces éléments qu'un tel accord ne saurait être à un consentement des membres de son réseau au traitement de leurs données à caractère personnel, au sens du RGPD.

22. En troisième lieu, la Commission relève qu'il est envisagé que l'enquêteur indique des observations relatives à l'attitude de l'enquêté (hostile, impatiente et agitée) et l'acceptabilité du registre (combien de coordonnées de ses amis données ; demande des justifications sur la raison pour laquelle on demande le registre etc.) dans un « *Questionnaire qualité* ». Au regard des finalités poursuivies par le traitement projeté, elle estime qu'une telle collecte n'apparaît pas justifiée.

Sur les droits des personnes

23. La Commission relève qu'une information sera délivrée, en français ou en chinois, aux personnes concernées, dans le formulaire de consentement à la participation de l'enquête, et de manière publique *via* un mini-site web disponible en français et en chinois. Elle prend acte que le formulaire de consentement fait mention de la possibilité de refuser de répondre à certaines questions.

24. La Commission rappelle que l'information délivrée doit être conforme aux mentions prévues à l'article 13 du RGPD et, s'agissant des personnes dont les données sont collectées indirectement, aux dispositions de l'article 14 du RGPD et de l'article 79 de la loi n° 78-17 du 6 janvier 1978 modifiée. En particulier, elle demande que les supports précités soient modifiés afin de mentionner la base légale du traitement ou encore les durées de conservation des données.

24. En outre, ces mêmes supports indiquent que les réponses seront anonymisées et qu'aucune utilisation des données ne permettra d'identifier les personnes concernées. À cet égard, la Commission rappelle que conformément à l'avis 05/2014 du groupe de travail « *de l'article 29* » sur la protection des données sur les techniques d'anonymisation, l'anonymisation est un processus technique irréversible garantissant que les données ne puissent plus permettre l'identification des personnes que ce soit par individualisation, corrélation ou inférence.

25. Dans la mesure où il n'est pas établi qu'un procédé d'anonymisation sera effectivement mis en œuvre dans le cadre du traitement projeté, elle demande que les supports précités soient modifiés sur ce point.

26. Les autres droits des personnes n'appellent pas d'observation particulière.

Sur les destinataires des données

27. La Commission, qui prend acte de l'engagement de l'INED selon lequel seul « *le cloud de l'INED* » hébergera les données personnelles collectées, rappelle qu'il

convient de mettre en œuvre les mesures de sécurité nécessaires au traitement des données par des chercheurs situés sur le territoire des États-Unis au regard du choix du responsable de traitement de ne pas organiser un transfert des données hors de l'Union européenne.

28. La Commission demande, conformément aux articles 13 et 14 du RGPD, que l'ensemble des catégories de destinataires, notamment la communauté scientifique *via* des réseaux de diffusion, apparaisse dans les informations délivrées aux personnes concernées.

Sur les durées de conservation

29. Il résulte des précisions apportées par l'INED que plusieurs durées de conservation sont susceptibles d'être appliquées. La Commission estime que la durée de conservation ne saurait être supérieure à cinq ans à compter de la dernière publication scientifique dont l'INED indique qu'elle interviendra au plus tard en 2023.

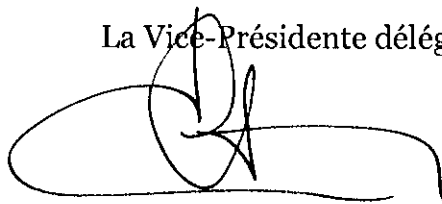
Sur les mesures de sécurité

30. De manière générale, la Commission prend acte qu'une formation spécifique sera délivrée aux enquêteurs afin de les sensibiliser à la réglementation applicable en matière de protection des données. De la même manière, elle relève que le formulaire de recueil de données sera rempli directement sur la base cible en connexion distante, ce qui est de nature à constituer une bonne pratique.

31. Par ailleurs, si la Commission relève qu'il est prévu qu'un engagement de non téléchargement des données soit souscrit entre l'INED et les membres de l'équipe de recherche situés aux États-Unis, elle estime – au regard de l'ensemble des éléments portés à sa connaissance – que des mesures de sécurité complémentaires devraient être mises en place afin de répondre à l'exigence de sécurité prévue à l'article 32 du RGPD. La Commission rappelle, en tout état de cause, que cette obligation nécessite la mise à jour des mesures de sécurité au regard de la réévaluation régulière des risques. À cet égard, la Commission rappelle qu'il conviendra d'apporter une attention spécifique à la réévaluation des mesures de sécurité dans le cadre de la mise à jour régulière de l'analyse d'impact.

Pour la Présidente

La Vice-Présidente déléguée

A stylized, handwritten signature in black ink, consisting of a large, flowing loop followed by a series of smaller, connected strokes.

Sophie LAMBREMON