

Comité de sécurité de l'information Chambre sécurité sociale et santé
--

CSI/CSSS/24/312

DÉLIBÉRATION N° 23/224 DU 7 NOVEMBRE 2023, MODIFIÉE LE 3 SEPTEMBRE 2024, RELATIVE À LA COMMUNICATION DE DONNÉES À CARACTÈRE PERSONNEL PSEUDONYMISÉES DU DATAWAREHOUSE MARCHÉ DU TRAVAIL ET PROTECTION SOCIALE PAR LA BANQUE CARREFOUR DE LA SÉCURITÉ SOCIALE À LA VRIJE UNIVERSITEIT BRUSSEL DANS LE CADRE DU PROJET « ETHNIC DIVERSITY IN FEDERAL PUBLIC SERVICES » (LA DIVERSITÉ ETHNIQUE DANS LES SERVICES PUBLICS FÉDÉRAUX).

Vu la loi du 15 janvier 1990 *relative à l'institution et à l'organisation d'une Banque-carrefour de la sécurité sociale*, en particulier les articles 5 et 15 ;

Vu la demande de la Vrije Universiteit Brussel ;

Vu le rapport de la Banque Carrefour de la sécurité sociale ;

Vu le rapport du président.

A. OBJET

1. La Vrije Universiteit Brussel réalise actuellement (en collaboration avec l'Université Gent et l'Université catholique de Louvain) une étude sur la diversité ethnique dans les services publics fédéraux (dans le cadre du projet « *Ethnic diversity in federal public services* ») et elle souhaite à cet effet avoir recours à des données à caractère personnel pseudonymisées (issues du datawarehouse marché du travail et protection sociale) relatives aux travailleurs salariés des autorités fédérales¹ âgés de 18 à 64 ans.

2.1. L'étude serait réalisée en deux phases. Dans une première phase, la Banque Carrefour de la sécurité sociale communiquerait aux chercheurs des données à caractère personnel pseudonymisées pour environ 10.000 travailleurs salariés des autorités fédérales (*un échantillon*), en vue du développement d'applications spécifiques. Dans une deuxième phase, les chercheurs auraient accès aux mêmes types de données à caractère personnel pseudonymisées de tous les travailleurs salariés des autorités fédérales (*la totalité de la population*), et ce sur un ordinateur sécurisé dans les locaux de la Banque Carrefour de la sécurité sociale et sous la surveillance permanente d'un collaborateur de cette organisation,

¹ Il s'agit des employeurs qui sont repris dans le datawarehouse marché du travail et protection sociale sous un code spécifique pour la variable « pouvoir organisateur », tels l'armée, la police, les services publics, les entreprises publiques, les organismes d'intérêt public, les institutions publiques de sécurité sociale, le parlement et les conseils consultatifs.

pour y appliquer les algorithmes qu'ils ont développés. Les chercheurs pourraient emporter les résultats de leurs actions, uniquement sous forme de données purement anonymes en dehors du bâtiment de la Banque Carrefour de la sécurité sociale (à cet effet, la Banque Carrefour de la sécurité sociale réaliserait, au préalable, une analyse de risque « *small cell* »).

- 2.2. Par ailleurs, une version brouillée des données à caractère personnel serait créée et communiquée aux chercheurs². Il s'agit d'un troisième ensemble de données (en plus du premier ensemble de données de l'échantillon et du deuxième ensemble de données de la population complète). Par cette communication, la Banque Carrefour de la sécurité sociale souhaite tester la technique du brouillage de données et acquérir de l'expérience pratique en la matière. L'organisation développe actuellement un système de brouillage et souhaite fréquemment l'utiliser à l'avenir. En effet, la procédure précitée (constituée de deux phases) n'est pas parfaite, à savoir que toutes les situations ou catégories possibles ne sont pas disponibles dans l'échantillon relativement limité (surtout les cas rares font parfois défaut). De cette manière, les chercheurs développent des applications qui ne couvrent pas toutes les situations ou catégories et ils doivent corriger ces applications au cours de la deuxième phase, dans les locaux de la Banque Carrefour de la sécurité sociale, sur la base de leurs constatations dans les fichiers de population. À cet égard, le brouillage est susceptible d'offrir une solution. En effet, de plus gros fichiers peuvent être communiqués de sorte que certaines situations ou catégories rares sont désormais présentes. Toutefois, celles-ci sont représentées au moyen de données ne reflétant pas la situation réelle. Les informations brouillées constituent, par conséquent, aussi une valeur ajoutée pour les chercheurs.
3. De la même manière, il serait procédé au traitement de données à caractère personnel pseudonymisées, d'une part, des travailleurs salariés des autres employeurs publics³ (tels les communautés, les régions, les provinces et les communes) et, d'autre part, des travailleurs salariés des employeurs privés. Dans une première phase, la Banque Carrefour de la sécurité sociale transmettrait aux chercheurs des données à caractère personnel pseudonymisées d'un échantillon des intéressés (environ 10.000 travailleurs salariés des employeurs publics non-fédéraux et des employeurs privés), en vue du développement d'applications spécifiques. Dans une deuxième phase, les chercheurs accéderaient aux mêmes types de données à caractère personnel pseudonymisées de l'ensemble des travailleurs salariés des deux catégories précitées, dans les circonstances précisées ci-avant (sur un ordinateur sécurisé au sein des locaux de la Banque Carrefour de la sécurité sociale et sous la surveillance permanente d'un collaborateur de cette organisation). Pour ces données à caractère personnel également, une version brouillée qui sera communiquée aux chercheurs serait créée.

² Le brouillage est une technique par laquelle des données réelles sont remplacées par des données qui ne représentent pas une situation réelle. Les valeurs réelles peuvent être remplacées par des valeurs fictives ou il peut y avoir une permutation de valeurs (par exemple, le remplacement du domicile Anvers par le domicile Hasselt). Certaines conditions peuvent aussi être imposées (par exemple, les valeurs doivent se situer dans une certaine fourchette, certaines valeurs sont interdites, ...). Certaines conditions relatives à la relation entre certaines variables peuvent aussi être imposées de sorte à conserver certaines rapports (par exemple, une valeur X de la variable A va toujours de pair avec une valeur Y de la variable B).

³ Il s'agit des employeurs qui sont repris dans le datawarehouse marché du travail et protection sociale sous un code non-fédéral pour la variable « pouvoir organisateur ».

4. Par intéressé – travailleur salarié des autorités fédérales, travailleur salarié d'un autre employeur public ou travailleur salarié d'un employeur privé– les données à caractère personnel suivantes seraient traitées (pour chaque année de la période 2010-2021): le sexe, la classe d'âge, la région du domicile, la nationalité actuelle, la première nationalité, la première nationalité des parents, la nationalité d'origine⁴, le pays de naissance, la date d'inscription au registre national, la date d'obtention de la nationalité, le contexte d'immigration⁵, le niveau de formation et le domaine d'études⁶, le statut du travailleur, le code NACE et le type d'employeur d'après la variable « pouvoir organisateur ». La nationalité / le pays (hormis quelques exceptions) seraient indiqués en faisant référence à une des 30 classes applicables. La date d'inscription au registre national et la date d'obtention de la nationalité sont mises à la disposition de la Vrije Universiteit Brussel en faisant référence à l'année applicable.
5. La Vrije Universiteit Brussel conserverait, en tant que responsable du traitement, les données à caractère personnel pseudonymisées reçues au cours de la première phase jusqu'à la fin du projet « *Ethnic diversity in federal public services* », c'est-à-dire jusqu'au 31 décembre 2024. L'organisation détruira donc les données à caractère personnel pseudonymisées reçues au plus tard le 31 décembre 2024. Les résultats de son étude seront uniquement publiés sous forme anonyme.

B. EXAMEN

Compétence du Comité de sécurité de l'information

6. En vertu de l'article 5 de la loi du 15 janvier 1990 *relative à l'institution et à l'organisation d'une Banque-carrefour de la sécurité sociale*, la Banque Carrefour de la sécurité sociale recueille des données à caractère personnel auprès des institutions de sécurité sociale, les enregistre, procède à leur agrégation et les communique aux personnes qui en ont besoin pour la réalisation de recherches pouvant être utiles à la connaissance, à la conception et à la gestion de la protection sociale.
7. En vertu de l'article 15 de la loi du 15 janvier 1990 *relative à l'institution et à l'organisation d'une Banque-carrefour de la sécurité sociale*, toute communication de données à caractère personnel par la Banque Carrefour de la sécurité sociale ou une autre institution de sécurité sociale doit faire l'objet d'une délibération préalable de la chambre sécurité sociale et santé du Comité de sécurité de l'information⁷.

⁴ La nationalité d'origine est déterminée en fonction de la nationalité actuelle, de la première nationalité et de la première nationalité des parents.

⁵ Le contexte d'immigration est déterminé en fonction de ces mêmes données à caractère personnel en combinaison avec le pays de naissance, la date d'inscription au registre national et la date d'obtention de la nationalité.

⁶ En fonction de la source d'information, il peut aussi s'agir de notions comparables, telles que le niveau d'études, le type d'études, le domaine d'études, la filière, la catégorie d'études et le code diplôme.

⁷ Le Comité de sécurité de l'information constate que les chercheurs souhaitent non seulement obtenir des données à caractère personnel pseudonymisées, mais également des données anonymes (tableaux). Dans la

Licéité du traitement

8. En vertu de l'article 6 du 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*, le traitement n'est licite que si, et dans la mesure où, au moins une des conditions mentionnées à cet article est remplie.
9. La communication des données à caractère personnel pseudonymisées précitées du datawarehouse marché du travail et protection sociale par la Banque Carrefour de la sécurité sociale à la Vrije Universiteit Brussel, pour l'exécution d'une étude relative à la diversité ethnique au sein des services publics fédéraux, est licite au sens de l'article 6, 1, alinéa 1^{er}, e), puisqu'elle est nécessaire à l'accomplissement d'une tâche d'intérêt public.

Principes relatifs au traitement des données à caractère personnel

10. En vertu du 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*, les données à caractère personnel doivent être collectées pour des finalités déterminées, explicites et légitimes et elles ne peuvent pas être traitées ultérieurement d'une manière incompatible avec ces finalités (limitation des finalités), elles doivent être adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées (minimisation des données), elles doivent être conservées sous une forme permettant l'identification des personnes concernées pendant une durée n'excédant pas celle nécessaire au regard des finalités pour lesquelles elles sont traitées (limitation de la conservation) et elles doivent être traitées de façon à garantir une sécurité appropriée des données à caractère personnel, y compris la protection contre le traitement non autorisé ou illicite et contre la perte, la destruction ou les dégâts d'origine accidentelle, à l'aide de mesures techniques ou organisationnelles appropriées (intégrité et confidentialité).

Limitation de la finalité

11. La communication de données à caractère personnel pseudonymisées poursuit une finalité déterminée, explicite et légitime, à savoir l'analyse de la diversité ethnique au sein des services publics fédéraux dans le cadre du projet « *Ethnic diversity in federal public services* ». Les minorités ethniques semblent sous-représentées dans les échelons supérieurs des autorités fédérales. Les chercheurs souhaitent identifier les facteurs susceptibles de promouvoir ou de freiner la diversité ethnique au sein des autorités fédérales (au cours des premières phases de recrutement et pendant les phases ultérieures en matière de carrière et de promotion).

mesure où cette communication est effectuée selon les dispositions de la délibération n° 18/140 du 6 novembre 2018, le Comité de sécurité de l'information ne doit pas se prononcer à cet égard.

Minimisation des données

12. Les chercheurs ne sont pas en mesure de réaliser la finalité au moyen de données anonymes, étant donné qu'ils doivent pouvoir suivre la situation de personnes individuelles pendant un certain temps. Ils s'engagent contractuellement à mettre en œuvre tous les moyens possibles pour éviter une identification des personnes concernées. En toute hypothèse, il leur est interdit d'entreprendre toute action visant à convertir les données à caractère personnel pseudonymisées en données à caractère personnel non pseudonymisées.
13. Dans la première phase de l'étude, les données à caractère personnel portent sur quelque 10.000 travailleurs salariés des autorités fédérales et sur quelque 10.000 travailleurs salariés des employeurs publics non-fédéraux et des employeurs privés. Le numéro d'identification de la sécurité sociale de chaque intéressé est remplacé par un numéro d'ordre unique sans signification. Les caractéristiques personnelles proprement dites sont limitées et sont réparties en classes. Les dates sont limitées à l'année.
14. Pour déterminer la diversité ethnique au sein des autorités fédérales, les chercheurs de la Vrije Universiteit Brussel ont besoin de l'origine de toutes les personnes concernées. Celle-ci est déterminée en fonction de la nationalité actuelle, de la première nationalité et de la première nationalité des parents. Le contexte d'immigration de l'intéressé est déterminé en fonction de son origine, de son pays de naissance, de l'année d'inscription au registre national et de l'année d'obtention de la nationalité.
15. Outre la classe de l'origine et la classe du contexte d'immigration, les variables précitées qui servent de base à cette information (en classes appropriées et dates limitées à l'année) sont également mises à la disposition des chercheurs puisque l'étude se focalise également sur l'opérationnalisation de l'ethnicité ou de l'origine ethnique. Les chercheurs souhaitent en effet pouvoir dresser la carte des différents scénarios des intéressés et les analyser.
16. Le niveau de formation et le domaine d'études, le sexe, la classe d'âge, la région du domicile, le statut du travailleur, le code NACE et le type d'employeur sont des variables susceptibles d'avoir un impact sur la situation des intéressés. Les chercheurs utiliseraient ces données à caractère personnel notamment pour comparer la situation des différents groupes.
17. Les données à caractère personnel pseudonymisées à communiquer au cours de la première phase semblent pertinentes et non excessives par rapport à la finalité précitée. Au cours de la deuxième phase, les chercheurs appliquent les applications qu'ils ont développées à l'aide des données à caractère personnel pseudonymisées reçues précédemment à l'ensemble de la population, dans un environnement sécurisé auprès de la Banque Carrefour de la sécurité sociale et sous la surveillance d'un de ses collaborateurs. Seuls ces résultats, sous forme de données purement anonymes, peuvent quitter les locaux de la Banque Carrefour de la sécurité sociale. En outre, une version brouillée des données à caractère personnel est créée et communiquée aux chercheurs (voir supra, le point 2.2.).

Limitation de la conservation

18. Les données à caractère personnel pseudonymisées seront détruites par les chercheurs dès qu'elles ne sont plus nécessaires à la réalisation de la finalité précitée et ce au plus tard le 31 décembre 2024. Cette durée de conservation peut, le cas échéant, uniquement être prolongée par une décision explicite en la matière de la chambre sécurité sociale et santé du Comité de sécurité de l'information.

Intégrité et confidentialité

19. Le demandeur met tout en œuvre pour éviter une identification des personnes concernées et s'abstient, à tous les égards, de toute tentative visant à convertir les données à caractère personnel pseudonymisées reçues de la part de la Banque Carrefour de la sécurité sociale au cours de la première phase de l'étude en données à caractère personnel non-pseudonymisées. Par ailleurs, il ne communique, en aucun cas, ces données à caractère personnel pseudonymisées à des tiers. Il publie, en outre, les résultats du traitement qu'il a réalisé uniquement sous une forme qui ne permet d'aucune façon d'identifier les assurés sociaux concernés.
20. Pour le surplus, le demandeur tient compte de la loi du 15 janvier 1990 *relative à l'institution et à l'organisation d'une Banque-carrefour de la sécurité sociale* et de toute autre disposition réglementaire relative à la protection de la vie privée, en particulier 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* et de la loi du 30 juillet 2018 *relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel*.

Par ces motifs,

la chambre sécurité sociale et santé du comité de sécurité de l'information

conclut que la communication de données à caractère personnel pseudonymisées du datawarehouse marché du travail et protection sociale par la Banque Carrefour de la sécurité sociale à la Vrije Universiteit Brussel, dans le cadre du projet « *Ethnic diversity in federal public services* », telle que décrite dans la présente délibération, est autorisée moyennant le respect des mesures de protection des données qui ont été définies.

Les modifications de cette délibération, approuvées par le comité de sécurité de l'information le 3 septembre 2024, entrent en vigueur le 18 septembre 2024.

Michel DENEYER
Président

Le siège de la chambre sécurité sociale et santé du Comité de sécurité de l'information est établi dans les bureaux de la Banque Carrefour de la sécurité sociale, à l'adresse suivante: Quai de Willebroeck 38 - 1000 Bruxelles (tél. 32-2-741 83 11).