

<b>Comité de sécurité de l'information Chambre sécurité sociale et santé</b>
----------------------------------------------------------------------------------

CSI/CSSS/21/164

**DÉLIBÉRATION N° 21/090 DU 7 DÉCEMBRE 2021 RELATIVE À LA COMMUNICATION DE DONNÉES À CARACTÈRE PERSONNEL PSEUDONYMISÉES PAR LA BANQUE CARREFOUR DE LA SÉCURITÉ SOCIALE AU CENTRUM VOOR DEMOGRAFIE, FAMILIE EN GEZONDHEID DE L'UNIVERSITÉ D'ANVERS EN VUE D'UNE ÉTUDE DES PARCOURS DE VIE DE FRÈRES ET SŒURS EN BELGIQUE**

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 loi du 3 décembre 2017 *relative à la création de l'Autorité de protection des données*, en particulier l'article 114 ;

Vu la loi du 5 septembre 2018 *instituant le comité de sécurité de l'information et modifiant diverses lois concernant la mise en œuvre 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*, en particulier l'article 97 ;

Vu la demande de l'Université d'Anvers;

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

Vu le rapport de monsieur Bart Viaene.

**A. OBJET**

1. Le « Centrum voor Demografie, Familie en Gezondheid » qui fait partie de la Faculté des sciences sociales de l'Université d'Anvers réalise pour l'instant une étude sur les parcours de vie de frères et sœurs en Belgique. Il souhaite vérifier dans quelle mesure les frères et sœurs ont une influence sur leur trajet de vie mutuel et quels mécanismes jouent un rôle à cet égard. Il tiendrait à cet effet aussi compte des caractéristiques socioéconomiques des personnes concernées et vérifierait dans quelle mesure les caractéristiques des parents ont une influence. L'étude serait réalisée en deux phases. Au cours d'une première phase, la Banque Carrefour de la sécurité sociale transmettrait des données à caractère personnel pseudonymisées d'un groupe limité de la population cible aux chercheurs, en vue du développement d'algorithmes spécifiques (en outre, plusieurs déformations seraient aussi réalisées, par exemple en attribuant systématiquement au sein de la population l'âge ou d'autres caractéristiques d'une personne concernée à une autre personnes concernée, à titre de *scrambling* (brouillage)). Au

cours de la deuxième phase, les chercheurs auraient accès aux mêmes types de données à caractère personnel pseudonymisées d'un groupe plus large de la population cible, sur un ordinateur sécurisé dans les bâtiments de la Banque Carrefour de la sécurité sociale, sous surveillance permanente, pour y appliquer les algorithmes qu'ils ont développés, et ils pourraient emporter les résultats, uniquement sous la forme de données purement anonymes, en dehors du bâtiment de la Banque Carrefour de la sécurité sociale, après réalisation d'une analyse de risques « small cells » par cette dernière.

2. La Banque Carrefour de la sécurité sociale extrairait donc dans le datawarehouse marché du travail et protection sociale un échantillon de personnes ayant atteint en 1998 l'âge respectif de 18, 23 et 28 ans et ayant au moins un frère ou une sœur biologique qui séjourne aussi en Belgique. Ils seraient suivis dans le temps, en même temps que leurs parents biologiques respectifs et les alliés des frères et sœurs. Les parents et les enfants seraient à cet égard toujours étudiés dans une composition de ménage complète.
3. Seuls les ménages composés de deux enfants au moins sont retenus. L'échantillon total comprend trois échantillons parallèles spécifiques (comprenant chacun quelque dix mille ménages). Dans chaque échantillon, l'âge de l'enfant aîné au 31 décembre 1998 joue un rôle primordial (respectivement 18, 23 et 28 ans). Tous les enfants biologiques des parents de cet enfant aîné seront suivis. Des données à caractère personnel pseudonymisées d'environ 400.000 personnes au total seraient donc traitées à travers le temps.
4. En ce qui concerne le premier échantillon (les personnes âgées de 18 ans), la procédure suivante serait appliquée. Tout d'abord, toutes les personnes âgées de 18 ans au 31 décembre 1998 ayant la nationalité belge et leur père et mère biologiques respectifs sont identifiés. Ensuite, l'affiliation est déterminée (situation au 31 décembre 2008, l'enfant aîné est alors âgé de 28 ans). Il est ensuite déterminé quelles personnes concernées ont un frère ou une sœur biologique et un échantillon d'environ dix milles individus est extrait. Finalement, tous les frères et sœurs biologiques à la fin de la période d'observation (31 décembre 2019) sont identifiés. Les mêmes étapes seraient également parcourues pour le deuxième échantillon (les personnes âgées de 23 ans) et pour le troisième échantillon (les personnes âgées de 28 ans). Le résultat est un échantillon total d'environ trente mille individus répartis en trois catégories d'âge. Au cours de la première phase, les chercheurs recevraient des données à caractère personnel de trois pour cent de tout échantillon partiel (comprenant environ trois cents ménages). Au cours de la deuxième phase, ils traiteraient des données à caractère personnel de tout échantillon partiel (comprenant environ dix mille ménages) dans le bâtiment de la Banque Carrefour de la sécurité sociale.
5. Le demandeur fait observer qu'il souhaite suivre divers acteurs de manière longitudinale. Il s'agit de l'individu échantillonné même, du père biologique, de la mère biologique et de leurs enfants, quel que soit le ménage dont ils font partie. Ce dernier groupe constitue le point d'attention de l'étude, étant donné que le parcours de vie de frères et de sœurs doit être étudié. Pour chaque membre de ce groupe de frères et sœurs, les partenaires et enfants respectifs sont aussi suivis, mais uniquement dans la mesure où ils font partie du ménage de ce membre. Toute personne qui est suivie pendant la période complète ou non, recevrait un numéro d'ordre unique (sans signification) et un numéro de ménage unique (sans signification).

6. En ce qui concerne les individus échantillonnés, les pères, les mères, les frères et les sœurs, les données à caractère personnel pseudonymisées suivantes sont traitées par trimestre (période 1998-2019). Les dates sont toujours limitées à l'année et au mois dans lesquels elles tombent. Les montants sont toujours répartis en classes adéquates.

*Caractéristiques de la personne et du ménage:* le numéro d'ordre unique de la personne concernée et du chef de ménage, l'indication selon laquelle la personne concernée est ou n'est pas un individu échantillonné, le nombre de membres du ménage, la relation au chef de ménage, la position au sein du ménage LIPRO, le type de ménage, le sexe, l'année/le trimestre de la naissance et du décès, la position socio-économique, la province du domicile, la distance par rapport au domicile des autres acteurs, la nationalité actuelle et la première nationalité (en classes), l'état civil et la cohabitation légale.

*Prestation professionnelle et régime de travail:* les prestations en tant que travailleur salarié (le code d'importance, le code travailleur, la fonction, le code NACE, le code de prestation, le type de prestation, le pourcentage de travail à temps partiel, le nombre de jours et d'heures à prester suivant le contrat, le type de contrat, le caractère public/privé de l'employeur et le numéro d'ordre unique) et les prestations en tant que travailleur indépendant (la profession, la catégorie de cotisation, la qualité et la période applicable ainsi que la mention du mois et de l'année d'affiliation et de suppression).

*Volume de travail et salaire:* le nombre de jours à temps plein normalement rémunérés, le nombre de jours à temps partiel normalement rémunérés, le nombre de jours équivalents temps plein en ce compris les jours assimilés, le nombre de jours équivalents temps plein à l'exclusion des jours assimilés, le nombre (moyen) d'heures par semaine du travailleur de référence, le nombre de jours par semaine du régime de travail du travailleur, le salaire brut (en classes), la date de début (année et mois) et la date de fin (année et mois) de paiement des allocations familiales, la qualité dans le dossier des allocations familiales (enfant bénéficiaire, attributaire ou allocataire), le salaire journalier (en classes) et le revenu provenant de l'activité indépendante (en classes).

*Réduction des prestations:* le statut de travail à temps partiel avec garantie de revenus, le montant de l'allocation (en classes), la situation à la fin du mois (travail à temps partiel avec garantie de revenus ou non), le nombre de jours indemnisés, le code secteur NACE en cas d'interruption de carrière ou de crédit-temps, le type de contrat, le statut, l'allocation (en classes), l'activité complémentaire, la durée et la raison de l'interruption de carrière ou du crédit-temps, le régime applicable et le nombre de jours indemnisés.

*Chômage:* le montant de l'allocation de chômage (en classes), le montant de l'allocation journalière (en classes), la situation à la fin du mois (en chômage ou non), le statut, le nombre de jours indemnisés, la durée du chômage, le statut relatif au chômage temporaire, le type de chômage temporaire, le montant de l'allocation en cas de chômage temporaire (en classes) et le nombre de jours indemnisés en cas de chômage temporaire.

*Invalidité:* la date de début (années et mois), la date de fin (année et mois), la date de début de la maladie (année et mois), la profession en cas d'invalidité, le cumul éventuel des allocations, l'applicabilité au dernier jour du trimestre, le nombre de jours indemnisés (en

classes), le type d'allocation, le régime applicable, le montant de l'indemnité d'invalidité (en classes), la date de début de la période de paiement (année et mois) et la date de fin de la période de paiement (année et mois).

*Prépension et pension*: le montant de la prépension (en classes), la situation à la fin du mois, le statut, le nombre de jours indemnisés en cas de prépension (en classes), la date de début de la pension (année et mois), le type de pension, la situation administrative ou juridique de la pension, le code avantage, la nature de l'avantage, le montant brut de la pension (en classes), la situation fiscale, le nombre d'enfants à charge et le nombre d'autres personnes à charge.

*Situation vis-à-vis du centre public d'action sociale*: la réglementation applicable, le type d'aide du centre public d'action sociale compétent, la date de début du paiement (année et mois), la date de fin du paiement (année et mois), la catégorie du bénéficiaire du revenu d'intégration sociale, la catégorie du bénéficiaire de l'aide financière et le montant remboursé par l'Etat au centre public d'action sociale (en classes).

*Carrière selon SIGEDIS (depuis le début de la carrière jusqu'en 2019)*: l'année, le code, la source des informations, le numéro d'ordre unique de l'employeur, le nombre d'heures par semaine du travailleur de référence (en classes), la rémunération (en classes), le nombre de jours prestés et de jours assimilés (en classes), le nombre de jours équivalents temps plein prestés et assimilés (en classes), la date de début de la période (année et mois) et la date de fin de la période (année et mois).

*Situation en tant que demandeur d'emploi et niveau de formation*: la catégorie de demandeur d'emploi, le nombre de mois d'inscription, la date d'inscription (année et mois), le niveau de formation, le code du diplôme, la date du diplôme (année et mois), le type d'enseignement secondaire obligatoire, la catégorie, le degré et le type de formation, le niveau d'étude, le domaine d'étude, la catégorie d'étude, le programme d'étude, la forme d'enseignement, la délivrance du diplôme à la fin ou non de l'année académique et le type d'organisation de délivrance.

*Variables supplémentaires (dérivées/construites)*: l'indicateur LWI (le *low work intensity*, la faible intensité de travail), le revenu sur base annuelle (en classes), le revenu brut (en classes), le nombre de jours prestés par année, le nombre de jours prestés par année convertis en équivalents temps plein, le nombre de jours assimilés par année, le nombre de jours assimilés par année convertis en équivalents à temps plein et le revenu payé pour les jours prestés sur base annuelle (en classes).

*Variables de l'échantillon*: la variable dummy pour la mère biologique, la variable dummy pour le père biologique, la variable dummy pour les enfants (frères et sœurs), le numéro d'ordre unique de la mère biologique, le numéro d'ordre unique du père biologique et l'indication de l'âge de l'enfant aîné au 31 décembre 1998 (18 ans pour le premier échantillon, 23 ans pour le deuxième échantillon, 28 ans pour le troisième échantillon). Par variable dummy, on entend en l'espèce: l'indication que la personne est ou non (oui/non) un type d'acteur déterminé (père biologique, mère biologique, ...).

7. Pour les membres du ménage des acteurs mentionnés au point 6 qui étaient âgés de 16 ans ou plus dans la période 1998-2019, les mêmes données à caractère personnel pseudonymisées seraient mises à la disposition pour l'ensemble de la période de l'étude. Pour les membres du ménage des acteurs mentionnés au point 6 qui étaient âgés de moins de 16 ans dans la période 1998-2019, seules quelques données à caractère personnel pseudonymisées seraient mises à la disposition, à savoir l'année de naissance, le trimestre de naissance, le sexe et la relation à la personne de référence. Ces personnes ne doivent pas faire l'objet d'un suivi longitudinal spécifique. Ces informations sont uniquement nécessaires tant que ces personnes font partie du ménage des acteurs précités.
8. Il s'agit d'une communication ponctuelle de données à caractère personnel pseudonymisées. Les chercheurs souhaitent conserver les données à caractère personnel d'une sélection de tout échantillon partiel qui ont été mises à la disposition au cours d'une première phase, jusqu'au 31 décembre 2027 au plus tard. Au cours de la deuxième phase, ils ont accès aux données à caractère personnel de tout échantillon partiel, dans les bâtiments de la Banque Carrefour de la sécurité sociale, sous la surveillance d'un collaborateur de ce dernier. Toutefois, par la suite, ils peuvent uniquement disposer de données purement anonymes.

## **B. EXAMEN**

### Compétence du Comité de sécurité de l'information

9. 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.
10. 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 section sécurité sociale et santé du Comité de sécurité de l'information.

### Licéité du traitement

11. 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.
12. La communication de données à caractère personnel pseudonymisées du datawarehouse marché du travail et protection sociale par la Banque Carrefour de la sécurité sociale au « Centrum voor Demografie, Familie en Gezondheid » de l'Université d'Anvers est par

conséquent licite puisqu'elle est nécessaire à l'accomplissement d'une mission d'intérêt public, au sens de l'article 6, 1, alinéa 1<sup>er</sup>, e), du RGPD.

### Principes en matière de traitement de données à caractère personnel

- 13.** 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 (principe de la limitation des finalités), elles doivent être adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées (principe de la 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 (principe de la 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 (principe d'intégrité et de confidentialité).

### Limitation de la finalité

- 14.** Le traitement des données à caractère personnel pseudonymisées par l'Université d'Anvers en tant que responsable du traitement vise une finalité légitime, à savoir la réalisation d'une étude sur les parcours de vie de frères et sœurs en Belgique.

### Minimisation des données

- 15.** 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 communiquées en données à caractère personnel non pseudonymisées.
- 16.** Les données à caractère personnel à communiquer (par trimestre de la période 1998-2019) sont pertinentes, adéquates et non excessives par rapport à la finalité précitée. Elles se limitent, par personne concernée, principalement à plusieurs caractéristiques de la personne, de son ménage et de sa formation, à des informations relatives à ses prestations professionnelles, au régime de travail, au volume de travail et au salaire et à la situation spécifique au niveau de la sécurité sociale (prestations réduites, chômage, invalidité, prépension, pension et intervention du centre public d'action sociale). Au cours d'une première phase, la Banque Carrefour de la sécurité sociale réalise par ailleurs plusieurs déformations des données à caractère personnel, par exemple en brouillant les données de la sorte qu'il n'est plus possible de déterminer les données originales sur la base des données traitées.

17. Le numéro d'identification de la sécurité sociale de chaque intéressé est remplacé par un numéro d'ordre unique sans signification. Les caractéristiques personnelles sont limitées et sont réparties en classes. Les dates sont limitées à l'année et au mois dans lesquels elles tombent. Les montants sont répartis en classes adéquates.
18. Les caractéristiques demandées de la personne et du ménage paraissent nécessaires pour différentes raisons. Elles permettent de reconstruire la composition de ménage (cette dernière est en effet essentielle pour l'étude de la formation et de la dissolution des relations). La composition tant du ménage parental que des ménages des frères et sœurs constitue l'essence de l'étude. L'âge des divers acteurs (les parents biologiques, les frères et sœurs et leurs enfants) est important pour le calcul de variables dérivées, telles la différence d'âge entre les partenaires et l'âge des parents à la naissance de leur premier enfant. Le secteur statistique du domicile des personnes concernées n'est pas communiqué en tant que tel mais est uniquement utilisé pour déterminer la distance entre le domicile des différents acteurs (en effet, la proximité constituerait un facteur important de la « transmission » intergénérationnelle et intragénérationnelle). Par ailleurs, les chercheurs souhaitent comparer les positions socio-économiques des individus échantillonnés et des partenaires ainsi que vérifier leur impact sur les décisions relatives au marché du travail et sur les dynamiques de famille.
19. Le comportement sur le marché du travail joue, selon les chercheurs, un rôle important au niveau de l'influence réciproque entre frères et sœurs. Afin de vérifier dans quelle mesure la position du travail influence la formation et la dissolution de relations, ils souhaitent obtenir une idée précise de la position des divers acteurs sur le marché du travail, en ce compris le régime de travail, le volume de travail, les revenus professionnels (structure salariale et évolution salariale), la situation en matière de sécurité sociale (et les indemnités susceptibles de compléter le revenu) et la formation. Par ailleurs, la transition par trimestre occupe une place centrale (l'étude vise à expliquer dans quelle mesure des changements sur le marché du travail ont ou n'ont pas une influence sur la formation et la dissolution de relations). Les informations (limitées) relatives à l'employeur servent uniquement à corriger les différences au niveau des revenus en fonction du secteur d'occupation et de la taille de l'organisation.

#### Limitation de la conservation

20. Les données à caractère personnel reçues au cours de la première phase de l'étude sont détruites par l'Université d'Anvers dès qu'elles ne sont plus nécessaires à la réalisation de la finalité précitée et au plus tard le 31 décembre 2027. Ce délai de conservation peut uniquement être prorogé au moyen d'une délibération du Comité de sécurité de l'information.

#### Intégrité et confidentialité

21. Le demandeur met tout en œuvre pour éviter une identification des personnes concernées et s'abstient de toute tentative visant à convertir les données à caractère personnel pseudonymisées reçues au cours de la première phase 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 les résultats du traitement uniquement sous une forme qui ne permet pas d'identifier les personnes concernées.

22. Lors du traitement des données à caractère personnel, il 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 réglementation relative à la protection de la vie privée, en particulier 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* 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 le traitement de données à caractère personnel pseudonymisées par le *Centrum voor Demografie, Familie en Gezondheid* de l'Université d'Anvers, en vue de l'étude des parcours de vie de frères et sœurs en Belgique, tel que décrit dans la présente délibération, est autorisé moyennant le respect des mesures de protection des données qui ont été définies.

Bart VIAENE  
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).