

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

CSI/CSSS/20/346

DÉLIBÉRATION N° 20/188 DU 1 SEPTEMBRE 2020 RELATIVE À LA COMMUNICATION DE DONNÉES À CARACTÈRE PERSONNEL PSEUDONYMISÉES RELATIVES À LA SANTÉ ISSUES DE L'ENQUÊTE BELGE DE SANTÉ 2013 PAR SCIENSANO À L'HEC LIÈGE (UNIVERSITÉ DE LIÈGE), SMART CITY INSTITUTE, DANS LE CADRE D'UNE RECHERCHE ACADÉMIQUE SUR LA RELATION ENTRE LE DÉVELOPPEMENT DES INITIATIVES SMART CITY DANS UN TERRITOIRE DONNÉ AVEC LA SANTÉ DE SES HABITANTS

Le Comité de sécurité de l'information, chambre sécurité sociale et santé (dénommé ci-après « le Comité ») ;

Vu le règlement (UE) n°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 (règlement général sur la protection des données) ;

Vu la loi du 3 décembre 2017 *portant création de l'Autorité de protection des données*, notamment l'article 114, modifié par la loi du 25 mai ;

Vu la loi du 30 juillet 2018 *relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel* ;

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* ;

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

Vu la loi du 21 août 2008 *relative à l'institution et à l'organisation de la plate-forme eHealth* ;

Vu la demande d'autorisation d'HEC Liège (Université de Liège), Smart City Institute ;

Vu le rapport d'auditorat de la Plate-forme eHealth ;

Vu le rapport de monsieur Bart Viaene ;

Émet, après délibération, la décision suivante, le 1 septembre 2020:

I. OBJET DE LA DEMANDE

1. En 2012, l'Institut scientifique de santé publique (ISP)¹ a été chargé, notamment pour l'autorité fédérale, les Communautés et les Régions, de l'organisation quinquennale d'une enquête nationale de santé au moyen d'une interrogation d'un échantillon de la population belge. Les résultats permettent de déterminer les besoins réels en matière de santé de la population belge, d'établir des rapports entre l'état de santé, certains facteurs (le mode de vie, l'environnement, le statut socio-économique,...) et l'utilisation de soins de santé (préventifs ou curatifs) et de soutenir les décisions politiques en matière de santé publique.
2. L'enquête de santé nationale est effectuée sous la responsabilité de l'ISP qui, pour certains aspects pratiques (en particulier l'extraction de l'échantillon et l'interrogation des personnes de l'échantillon), fait cependant appel aux services de la Direction générale Statistique et Information économique du Service public fédéral Economie, PME, Classes moyennes et Energie en qualité de sous-traitant.
3. L'enquête quinquennale est réalisée, sur base volontaire, auprès d'un échantillon aléatoire pondéré d'au moins dix mille personnes domiciliées en Belgique. Cet échantillon est extrait du registre national des personnes physiques par la Direction générale Statistique et Information économique – 3.500 personnes en provenance de la Flandre, 3.500 en provenance de la Wallonie et 3000 personnes en provenance de Bruxelles (en 2013, 450 personnes des provinces de Namur et de Luxembourg ont été ajoutées, à la demande explicite des autorités compétentes).
4. Les données à caractère personnel, qui sont recueillies au moyen des interviews, sont pseudonymisées par la Direction générale Statistique avant d'être mises à la disposition de l'ISP. Ce codage consiste dans le remplacement du numéro d'identification de la sécurité sociale de chaque intéressé par un numéro d'ordre sans signification. La Direction générale Statistique est la seule à conserver le lien entre les deux numéros.
5. L'ISP procède à un deuxième codage et conserve les données à caractère personnel doublement codées sur un serveur spécifique. Lorsqu'un accès est autorisé pour un tiers, ce dernier reçoit un login et un mot de passe personnalisés lui permettant de télécharger les données de l'ISP.

¹ Depuis le 1er avril 2018, l'Institut scientifique de Santé publique (ISP) et le Centre d'Étude et de Recherches Vétérinaires et Agrochimiques (CERVA) ont fusionné pour créer le nouveau centre fédéral de recherche Sciensano (Arrêté royal du 28 mars 2018 portant exécution de la loi du 25 février 2018 portant création de Sciensano, en ce qui concerne le siège social, la gestion et le fonctionnement, ainsi que l'adaptation de divers arrêtés concernant les prédécesseurs légaux de Sciensano).

6. La section Santé du comité sectoriel² a formulé une recommandation positive concernant le traitement de données à caractère personnel dans le cadre de l'enquête nationale de santé (recommandation n° 12/03 du 20 novembre 2012).
7. L'AIM a réalisé en juillet 2015 une analyse de risque "*small cell*" sur l'ensemble des données à caractère personnel codées obtenues lors de l'enquête nationale de santé organisée en 2013, afin d'exclure la possibilité de réidentification des intéressés. Le comité a reçu le rapport de cette analyse.

II. COMMUNICATION D'UNE SÉLECTION DE DONNÉES À CARACTÈRE PERSONNEL PSEUDONIMISÉES RELATIVES À LA SANTÉ

8. HEC Liège (Université de Liège), Smart City Institute, souhaite obtenir une sélection de données à caractère personnel relatives à la santé pseudonymisées issues de l'enquête de santé 2013.
9. L'objectif du présent projet est d'étudier la relation entre le développement des initiatives Smart City dans un territoire donné avec la santé de ses habitants. Un volume considérable de ressources privées et publiques consacrées aux initiatives "Smart City".³ Cependant, il n'y a pas d'études empiriques sur les effets des smart cities sur la vie des citoyens. Ce projet interprète les initiatives des villes intelligentes comme le résultat d'une expérience naturelle (quasi-expérience) visant à déterminer si les citoyens des smart cities ont une meilleure santé, physique et mentale.
10. À cette fin, le projet utilisera des informations sur les initiatives smart city dans les communes belges, obtenues d'une enquête réalisée par l'Smart City Institute de l'Université de Liège qui montre qu'aucune initiative smart city n'a été développée dans les communes belges avant 2013. Ensuite, ces informations seront utilisées pour créer des groupes de traitement et de contrôle sur base des communes de résidences des répondants de l'Enquête de Santé belge. Finalement, des méthodes statistiques du propensity score et les différences-in-différences (cette dernière sera faite une fois que les données 2018 seront disponibles) seront utilisées pour obtenir les estimations des effets des smart city sur la santé des répondants.
11. Les données à caractère personnel relatives à la santé pseudonymisées suivantes sont communiquées par Sciensano au demandeur :
 - données relatives aux interviews (4 variables) ;
 - données démographiques (5 variables) ;
 - données relatives aux caractéristiques des logements (2 variables) ;

² Maintenant appelée la chambre sécurité sociale et santé du comité de sécurité de l'information.

³ Les estimations publiées dans le « 2016 World Cities Report of the United Nations Human Settlements Programme » indiquent que le marché des smart cities s'élevait à 506,8 milliards de dollars américains en 2012 et atteindrait 1,3 billion de dollars américains en 2019.

- données relatives à l'utilisation d'un proxy (2 variables) ;
- données relatives au niveau d'étude (2 variable) ;
- données relatives à l'emploi (6 variables) ;
- données relatives aux revenus des individus (1 variables) ;
- données relatives au logement (3 variables) ;
- données relatives à la perception de la santé (5 variables);
- données relatives à la santé mentale (5 variables) ;
- données relatives à la qualité de vie liée à la santé (4 variables) ;
- données relatives à la consommation de boissons alcoolisées (2 variables) ;
- données relatives au tabagisme (à l'exclusion du tabagisme passif) (6 variables) ;
- données relatives à l'activité physique (4 variables) ;
- données relatives à l'accessibilité financière des services de santé (4 variables) ;
- données relatives à la santé et environnement, exposition à la fumée de tabac (12 variables) ;
- données relatives à la violence (1 variables) ;
- données relatives à la santé sociale (4 variables).

III. COMPETENCE

12. Conformément à l'article 42, § 2, 3°, de la loi du 13 décembre 2006 portant dispositions diverses en matière de santé, toute communication de données à caractère personnel relatives à la santé requiert, sauf les exceptions prévues, une autorisation de principe de la chambre sécurité sociale et santé du comité de sécurité de l'information.
13. Le comité estime dès lors qu'il est compétent pour se prononcer sur cette communication de données à caractère personnel relatives à la santé.

IV. EXAMEN DE LA DEMANDE

A. ADMISSIBILITÉ

14. Le traitement de données à caractère personnel relatives à la santé est en principe interdit, et ce conformément au prescrit de l'article 9, §1er 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 (RGPD).
15. Selon l'article 9, §2, j), du RGPD, cette interdiction ne s'applique pas lorsque ce traitement est nécessaire à des fins de recherche scientifique ou à des fins statistiques, conformément à l'article 89, §1er, sur la base du droit de l'Union ou du droit d'un Etat membre qui doit être proportionné à l'objectif poursuivi respecter l'essence du droit à la protection des données et prévoir des mesures appropriées et spécifiques pour la sauvegarde des droits fondamentaux et des intérêts de la personne concernées.

16. A la lumière de ce qui précède, le comité de sécurité de l'information est par conséquent d'avis qu'il existe un fondement admissible pour le traitement des données à caractère personnel pseudonymisées relatives à la santé envisagé.

B. FINALITÉ

17. Le traitement de données à caractère personnel est uniquement autorisé pour des finalités déterminées, explicites et légitimes.⁴
18. Au vu des objectifs du traitement tels que décrits ci-dessus, le comité considère que le traitement des données à caractère personnel envisagé poursuit bien des finalités déterminées, explicites et légitimes.

C. PROPORTIONNALITÉ

19. Conformément à l'art. 5, b) et c) du RGPD, les données à caractère personnel doivent être adéquates, pertinentes et non excessives au regard des finalités pour lesquelles elles sont obtenues et pour lesquelles elles sont traitées ultérieurement.
20. Les données à caractère personnel doivent être conservées sous une forme permettant l'identification des personnes concernées uniquement pendant la durée nécessaire à la réalisation des finalités pour lesquelles les données à caractère personnel sont traitées.
21. Le principe de proportionnalité implique que le traitement doit en principe être réalisé au moyen de données anonymes. Cependant, si la finalité ne peut être réalisée au moyen de données anonymes, des données à caractère personnel pseudonymisées peuvent être traitées. Vu la nécessité de réaliser des analyses détaillées à partir de ces données, les chercheurs ont besoin d'avoir accès à des données pseudonymisées afin d'être en mesure de réaliser ces analyses qu'ils ne pourraient pas réaliser à l'aide de données anonymes. Cette finalité justifie donc le traitement de données à caractère personnel pseudonymisées.
22. Le comité de sécurité de l'information estime que les données à caractère personnel qui seraient transmises au demandeur sont effectivement de nature pseudonymisées puisque le numéro d'identification utilisé pour les membres d'un ménage est codé une première fois par la Direction générale de la Statistique et une deuxième fois, spécifiquement pour le projet, par l'ISP.
23. Les données pseudonymisées seront conservées pendant une période de trois ans. La durée est nécessaire pour effectuer l'analyse statistique des données, la rédaction des rapports scientifiques, la présentation des résultats consolidés lors de colloques et conférences scientifiques, la soumission à des revues scientifiques, les possibles révisions, et l'acceptation pour publication.

⁴ Art. 5 point 1 b) RGPD.

24. Le comité de sécurité de l'information dit que ce délai de conservation est raisonnable et précise que les données à caractère personnel pseudonymisées devront être détruites au plus tard le 31 décembre 2023.
25. Le demandeur déclare que le traitement des différentes données à caractère personnel pseudonymisées demandées est nécessaire pour les raisons suivantes :

Module 1:	Description des données dans ce module	Données de l'interview (« Interview related information »)
	La raison exacte pour laquelle ces données sont nécessaire	Identification numérique des individus et des logements
Module 2:	Description des données dans ce module	Données démographiques (« Demographic information »)
	La raison exacte pour laquelle ces données sont nécessaire	Caractéristiques socioéconomiques que peuvent affecter la santé
Module 3:	Description des données dans ce module	Caractéristiques des logements (« Household characteristics»)
	La raison exacte pour laquelle ces données sont nécessaire	Caractéristiques du logement que peuvent affecter la santé
Module 4:	Description des données dans ce module	Information sur le répondant (« Information on use of proxy»)
	La raison exacte pour laquelle ces données sont nécessaire	Identification numérique des répondants
Module 5:	Description des données dans ce module	Éducation (« Education»)
	La raison exacte pour laquelle ces données sont nécessaire	Caractéristiques socioéconomiques que peuvent affecter la santé
Module 6:	Description des données dans ce module	Emploi (« Employment»)

	La raison exacte pour laquelle ces données sont nécessaire	Caractéristiques socioéconomiques que peuvent affecter la santé
Module 7:	Description des données dans ce module	Revenues des individus (« Income»)
	La raison exacte pour laquelle ces données sont nécessaire	Caractéristiques socioéconomiques que peuvent affecter la santé
Module 8:	Description des données dans ce module	Logement (« Housing»)
	La raison exacte pour laquelle ces données sont nécessaire	Caractéristiques du logement que peuvent affecter la santé
Module 9:	Description des données dans ce module	Perception de la santé (« Perceived health»)
	La raison exacte pour laquelle ces données sont nécessaire	Variables de (perception de la) santé
Module 12:	Description des données dans ce module	Stress et bien-être (« Mental health»)
	La raison exacte pour laquelle ces données sont nécessaire	Variables de santé (mentale)
Module 14:	Description des données dans ce module	Qualité de vie (« Health related quality of life»)
	La raison exacte pour laquelle ces données sont nécessaire	Variables de (perception de la) santé
Module 15:	Description des données dans ce module	Consommation de boissons alcoolisées (« Consumption of alcoholic beverages»)
	La raison exacte pour laquelle ces données sont nécessaire	Styles de vie que peuvent affecter la santé
Module 16:	Description des données dans ce module	Consommation de tabac (« Tobacco»)
	La raison exacte pour laquelle ces données sont nécessaire	Styles de vie que peuvent affecter la santé

Module 18:	Description des données dans ce module	Activité physique (« Physical activity»)
	La raison exacte pour laquelle ces données sont nécessaire	Styles de vie que peuvent affecter la santé
Module 33:	Description des données dans ce module	Accès aux soins de santé (« Financial accessibility of health services»)
	La raison exacte pour laquelle ces données sont nécessaire	Accès aux soins de santé
Module 35:	Description des données dans ce module	Santé et environnement, Exposition à la fumée de tabac (« Health and environment, housing, passive smoking»)
	La raison exacte pour laquelle ces données sont nécessaire	Caractéristiques de l'environnement que peuvent affecter la santé
Module 37:	Description des données dans ce module	Violence (« Violence»)
	La raison exacte pour laquelle ces données sont nécessaire	Caractéristiques de l'environnement que peuvent affecter la santé
Module 38:	Description des données dans ce module	Vie sociale (« Social health»)
	La raison exacte pour laquelle ces données sont nécessaire	Caractéristiques de l'environnement que peuvent affecter la santé

En résumé, puisque le projet de recherche vise à étudier la relation entre le développement des initiatives Smart City dans un territoire donné avec la santé de ses habitants, ils ont besoin de variables qui permettent d'identifier des individus et leur lieu de résidence (Modules 1 et 4) ainsi que de variables relatives à leur santé (Modules 12 et 14). Par ailleurs, ils nécessitent aussi de variables de contrôle, c'est-à-dire les facteurs qui peuvent affecter la santé de l'individu (en dehors les initiatives des villes intelligentes) telles que les caractéristiques (socioéconomiques) des individus (Modules 2, 5, 6 et 7) et leur environnement (Modules 35, 37 et 38) et logements (Modules 3 et 8), ainsi que leurs styles de vie (Modules 15, 16 et 18) et l'accès au soin de santé (Module 33).

26. Le comité estime que la demande est proportionnelle.

D. TRANSPARENCE

27. Le responsable du traitement de données à caractère personnel, collectées à des fins déterminées, explicites et légitimes ou l'organisation intermédiaire doit, préalablement au codage de données, en principe communiquer certaines informations à la personne concernée.
28. Lors de l'organisation de l'enquête de santé, les ménages sélectionnés reçoivent une lettre d'invitation et un dépliant informatif expliquant ce qu'est l'enquête de santé, le type de questions qui seront posées lors de l'interview et les institutions susceptibles d'utiliser ces données. Il est également précisé que la participation à cette enquête est libre. Les intéressés sont informés du fait que leurs données à caractère personnel seront traitées de manière codée à des fins de recherche scientifique.
29. Le comité estime que la demande répond aux exigences de transparence.

E. MESURES DE SÉCURITÉ

30. Le demandeur doit prendre toutes les mesures techniques et organisationnelles nécessaires à la protection des données à caractère personnel. Ces mesures doivent garantir un niveau de protection adéquat, compte tenu, d'une part, de l'état de la technique en la matière et des frais qu'entraîne l'application de ces mesures et, d'autre part, de la nature des données à protéger et des risques potentiels.
31. Pour garantir la confidentialité et la sécurité du traitement de données, tout organisme qui conserve, traite ou communique des données à caractère personnel est tenu de prendre des mesures dans les onze domaines d'action suivants liés à la sécurité de l'information: politique de sécurité; désignation d'un délégué à la protection des données; organisation et aspects humains de la sécurité (engagement de confidentialité du personnel, information et formations régulières du personnel sur le thème de la protection de la vie privée et sur les règles de sécurité); sécurité physique et de l'environnement; sécurisation des réseaux; sécurisation logique des accès et des réseaux; journalisation, traçage et analyse des accès; surveillance, revue et maintenance; système de gestion des incidents de sécurité et de la continuité (systèmes de tolérance de panne, de sauvegarde, ...); documentation.
32. Les données à caractère personnel relatives à la santé peuvent uniquement être traitées sous la surveillance et la responsabilité d'un professionnel des soins de santé. Même si cela n'est pas strictement requis, le comité estime qu'il est préférable de traiter de telles données sous la responsabilité d'un médecin, ce qui est le cas en l'espèce.
33. Le comité rappelle que lors du traitement de données à caractère personnel, le professionnel des soins de santé ainsi que ses préposés ou mandataires sont soumis au secret.

34. Il est interdit d'entreprendre toute action visant à convertir les données à caractère personnel pseudonymisées qui ont été communiquées en données à caractère personnel non pseudonymisées.
35. Le comité de sécurité de l'information constate qu'une analyse de risques "*small cell*" a été réalisée en 2015 sur l'ensemble des données à caractère personnel codées de la banque de données de l'enquête de santé 2013⁵.

Par ces motifs,

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

conclut que la communication des données à caractère personnel telle que décrite dans la présente délibération est autorisée moyennant le respect des mesures de protection de la vie privée qui ont été définies, en particulier les mesures en matière de limitation de la finalité, de minimisation des données, de limitation de la durée de conservation des données et de sécurité de l'information.

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.

⁵ Voir à ce sujet la recommandation n° 11/03 du 19 juillet 2011 de la section santé du Comité sectoriel de la sécurité sociale et de la santé relative à la note du Centre fédéral d'expertise des soins de santé portant sur l'analyse *small cell* de données à caractère personnel codées provenant de l'Agence intermutualiste.

Bijlage - Health Interview Survey 2013 – List of variables selected for the project:

1. Interview related information

id_anon	Identification number of respondent	Num
hh_cluster	Household cluster	Num
wfin	Weight of individual within the sample	Num
provw	Stratum	Num

2. Demographic information

age5	Age group (5 categories)	Num
HC04	Gender	Num
HC05	Civil status	Num
HC06_1	Nationality (3 categories)	Num
prov*	Province of residence	Num

3. Household characteristics

hhtype2	Household composition	Num
nbr_per	Number of household members	Num

4. Information on use of proxy

NR02	Respondent of interview	Num
NR03	Reason why a proxy was used	Num

5. Education

ET_1	Highest educational level within the household	Num
ET_3	Highest diploma	Num

6. Employment

EM01	A paid job at this moment	Num
EM02	Current non employment status	Num
EM03	Ever had paid job	Num
EM04	Currently (last) employed as employee or self-employed	Num
EM05	Fulltime of parttime employed	Num
EM08	Absence from work due to health problems	Num

7. Income

IN_1	Reported equivalent household income (Belgian weighted quintiles)	Num
------	---	-----

8. Housing

LO01	Description of the dwelling	Num
LO02	Housing tenure	Num
LO03	Monthly rent	Num

9. Perceived health

SH01	Subjective health	Num
SH01_1	Good subjective health	Num
SH01_2	Bad subjective health	Num
SH02	Chronic condition	Num
SH03	Long term limitation	Num

12. Mental health

SL_1	Eating disorder	Num
SL_2	Depressive disorder	Num
SL_3	Anxiety disorder	Num
SL_4	Sleeping disorder	Num
WB_1	Mean GHQ-12 score of psychological distress	Num

14. Health related quality of life

QL01	Problem in mobility	Num
QL_2	Health-related quality of life score	Num
QL06	Global health score (VAS)	Num
QL06_1	Global health score (VAS)	Num

15. Consumption of alcoholic beverages

AL_1	Problematic alcohol consumption (CAGE questions at 2+ cut-off)	Num
AL_35	Average number of drinks per week (7 days)	Num

16. Tobacco (excluding passive smoking)

TA06	Current smoking status	Num
TA06_1	Smoking habits (daily, occasional, former, never smoker)	Num
TA07_8	Mean number of filtered cigarettes smoked per day	Num

TA07_9	Mean number of non-filtered cigarettes smoked per day	Num
TA07_10	Mean number of tobacco cigarettes (all types) smoked per day	Num
TA07_11	Mean number of nicotine e-cigarette smoked per day	Num

18. Physical activity

PA01	Days vigorous physical activity last week	Num
PA03	Days moderate physical activity last week	Num
PA_01	Mean time devoted to at least moderate physical activity per day	Num
PA08_1	Leisure time physical activity	Num

33. Financial accessibility of health services

AC_1	Absolute monthly expenses for health care	Num
AC_2	Relative monthly expenses for health care	Num
AC_3	Perception of health care expenses regarding total income	Num
AC_4	Postponement of medical care consumption for financial reason	Num

35. Health and environment, housing, passive smoking

HE01_1	Environmental nuisance in the neighbourhood: at least one of the listed conditions is a serious problem	Num
HE0103_1	Environmental nuisance in the neighbourhood: accumulation of rubbish (binary)	Num
HE0201_1	Annoyance at home: air pollution (binary)	Num
TA2001_1	Exposure to tobacco smoke indoors: at home	Num
TA2002_1	Exposure to tobacco smoke indoors: at work	Num
HE03	Someone smokes inside home every day or almost every day	Num
LO04	Number of bedrooms in the household	Num
LO04_2	overcrowded households	Num
LO05	Unable to keep household warm in the winter	Num
LO05_1	Unable to keep household warm in the winter	Num
LO06	Humidity problems in the household	Num
LO07	Mold problems in the household	Num

37. Violence

TR01	Victim of burglary, robbery, or armed robbery, of verbal or psychological violence (eg. insults, threats, isolation), of physical violence (eg. being	Num
------	---	-----

	pushed, being beaten) or of sexual violence (eg. exhibitionisme, rape) in the past 12 months	
--	--	--

38. Social health

SO_1	Appreciation social contacts (binary)	Num
SO_2	Frequency of social contacts (binary)	Num
SO_4	Quality of social support (binary)	Num
SO02	Frequency of social contacts	Num