



La Commission de la protection de la vie privée

Délibération STAT n° 01/2010 du 13 janvier 2010

Objet : demande formulée par le groupe de recherche Ghent Institute for Political Studies (Institut d'études politiques de Gand) (associé à l'unité d'enseignement et de recherche de Sciences politiques de l'Université de Gand) afin d'obtenir de la Direction générale Statistique et Information économique la communication de données d'étude codées (provenant de l'enquête EU-SILC, d'ensembles de données en coupe de 2004-2007 et des données longitudinales de 2007) dans le cadre d'une recherche ayant pour sujet "Sociaal Kapitaal in individueel perspectief: een onderzoek naar de invloed van verstedelijking en sociaaleconomische status op de sociale trekingsrechten van individuen" (Capital social dans une perspective individuelle : une recherche sur l'influence de l'urbanisation et du statut socioéconomique sur les droits de tirage sociaux des individus) (STAT/MA/09/030)

La Commission de la protection de la vie privée (ci-après "la Commission") ;

Vu la loi du 4 juillet 1962 *relative à la statistique publique* (ci-après la "loi statistique publique") ;

Vu la loi du 8 décembre 1992 *relative à la protection de la vie privée à l'égard des traitements de données à caractère personnel* (ci-après la "LVP") ;

Vu l'arrêté royal du 13 février 2001 *portant exécution de la loi du 8 décembre 1992 relative à la protection de la vie privée à l'égard des traitements de données à caractère personnel* (ci-après "l'arrêté royal du 13 février 2001") ;

Vu l'arrêté royal du 7 juin 2007 *fixant les modalités relatives à la composition et au fonctionnement du Comité de surveillance Statistique institué au sein de la Commission de la protection de la vie privée* ;

Vu la demande du groupe de recherche Ghent Institute for Political Studies (Institut d'études politiques de Gand) (associé à l'unité d'enseignement et de recherche de Sciences politiques de l'Université de Gand), reçue le 24/11/2009 ;

Vu la demande d'avis technique et juridique adressée au Service public fédéral Économie, PME, Classes moyennes et Énergie (Direction générale Statistique et Information économique) le 14/12/2009 ;

Vu l'avis technique et juridique reçu le 12/01/2010 ;

Vu le rapport du Président ;

Émet, après délibération, la décision suivante, le 13/01/2010 :

I. OBJET DE LA DEMANDE

1. La demande vise à ce que le groupe de recherche Ghent Institute for Political Studies (Institut d'études politiques de Gand, ci-après GhIPS, associé à l'unité d'enseignement et de recherche de Sciences politiques de l'Université de Gand), dénommé ci-après le Chercheur, soit autorisé à obtenir de la Direction générale Statistique et Information économique (ci-après la DGSIE) la communication de données d'étude codées (provenant de l'enquête EU-SILC, d'ensembles de données en coupe de 2004-2007 et des données longitudinales de 2007) dans le cadre d'une recherche ayant pour sujet "Sociaal Kapitaal in individueel perspectief: een onderzoek naar de invloed van verstedelijking en sociaaleconomische status op de sociale trekkingsrechten van individuen".

2. La demande vise également à ce que soit approuvé le contrat de confidentialité qui doit être conclu entre la DGSIE et le Chercheur à la suite de cette communication.

3. Selon l'institution de gestion, la DGSIE, la demande concerne des données en coupe de l'enquête EU-SILC 2004 à 2007 incluse et le fichier longitudinal 2004-2007. La Commission part de ce principe lors de l'examen de la présente demande.

II. EXAMEN DE LA DEMANDE

A. LÉGISLATION APPLICABLE

A.1. Loi statistique publique

4. En vertu des articles 15 et 15*bis* de la loi statistique publique, la DGSIE est habilitée, après autorisation du Comité de surveillance statistique et moyennant un contrat de confidentialité approuvé par ce même comité, à communiquer des données d'étude codées aux destinataires mentionnés dans la loi statistique publique, aux conditions fixées dans cette même loi.

5. Conformément à l'article 16 de l'arrêté royal du 7 juin 2007, la Commission est, jusqu'à l'installation et la nomination des membres du Comité, chargée des missions attribuées à ce Comité par la loi statistique publique.

A.2. LVP et arrêté royal du 13 février 2001

6. En vertu de l'article 1, § 1 de la LVP et de l'article 1, 3° de l'arrêté royal du 13 février 2001, des données d'étude codées relatives à des personnes physiques identifiées ou identifiables constituent des données à caractère personnel dont le traitement n'est autorisé qu'aux conditions fixées dans la LVP et dans l'arrêté royal du 13 février 2001.

B. BASE JURIDIQUE

7. Le groupe de recherche Ghent Institute for Political Studies (associé à l'unité d'enseignement et de recherche de Sciences politiques de l'Université de Gand) est un des destinataires de données énumérés dans la loi statistique publique, plus précisément un institut scientifique au sens de l'article 15, premier alinéa, 4° de la loi statistique publique.

8. En principe, le Chercheur entre donc en ligne de compte pour être autorisé à recevoir les données demandées.

C. FINALITÉ

9. Les données à caractère personnel doivent être collectées pour des finalités déterminées, explicites et légitimes (article 4, § 1, 2° de la LVP).

10. La demande de données s'inscrit dans le cadre d'une thèse de doctorat ayant pour sujet "Social Kapitaal in individueel perspectief: een onderzoek naar de invloed van verstedelijking en sociaaleconomische status op de sociale trekkingsrechten van individuen". Cette recherche est financée par le Fonds Wetenschappelijk Onderzoek (FWO = Fonds de la Recherche Scientifique). Les données d'étude demandées servent de source à la recherche susmentionnée. Le but du projet est de combler la lacune dans la littérature internationale dans le domaine de la recherche concernant l'influence de l'urbanisation, du statut socioéconomique (SSE) et de l'interaction entre les deux sur le capital social des individus, ce qui se définit traditionnellement comme la capacité des individus à accéder, via leurs relations sociales, à des ressources qui peuvent leur être utiles dans leurs activités quotidiennes. En d'autres termes, le Chercheur veut examiner s'il existe une relation statistique entre la possession d'un capital social individuel d'une part et les indicateurs de statut socioéconomique et d'urbanisation d'autre part. Le Chercheur souhaite également vérifier si l'urbanisation et le SSE n'ont pas un effet d'interaction sur le capital social individuel. Il s'agit en d'autres termes de la symbiose d'une analyse environnementale et d'une analyse de caractéristiques individuelles.

11. Ces finalités répondent aux exigences susmentionnées de la LVP.

12. En vertu de l'article 15 de la loi statistique publique, les données à caractère personnel codées doivent être obtenues pour des finalités statistiques ou scientifiques.

13. Les données sont codées. Le but de cette recherche est purement statistique et scientifique. Les résultats concernent des échelles statistiques agrégées telles que des paramètres de régression. Il n'y a aucune finalité administrative. Selon l'institution de gestion, il n'y a donc pas d'objection sur le plan de la finalité. La Commission adhère à cet avis.

D. DONNÉES

14. Les données d'étude demandées sont les données de ladite enquête EU-SILC (European Union Statistics on Income and Living Conditions). Ces données sont collectées par la DGSIE dans le cadre du Règlement européen (CE) n° 1177/2003 du 16 juin 2003 *relatif aux statistiques communautaires sur le revenu et les conditions de vie (EU-SILC)*. La DGSIE collecte ces données au moyen d'interviews en face-to-face. Les SILC contiennent des informations sur les revenus et les conditions de vie, dont des données sur le capital social individuel.

15. En vue de réaliser la finalité mentionnée au point C, le Chercheur veut disposer de microdonnées codées validées des individus et des ménages, collectées dans les ensembles de données en coupe de 2004, 2005, 2006 et 2007 et les données longitudinales de l'année de référence 2007.

Le Chercheur souhaite obtenir les données longitudinales de l'année de référence 2007 car cela permet de distinguer les effets de sélection des effets causatifs. Les ensembles de données en coupe 2004-2006 devront fournir suffisamment d'éléments statistiques pour détecter des effets significatifs. Puisque le Chercheur recherche le lien entre le capital social individuel et le statut socioéconomique des personnes, il a besoin d'une large gamme d'indicateurs sociodémographiques et socioéconomiques, dont des informations concernant la formation scolaire, le revenu, la propriété, la situation professionnelle, etc. Des informations sur le secteur statistique et/ou ce qu'on appelle la lettre statistique (agrégat de plusieurs secteurs statistiques) sont nécessaires pour servir d'input à l'analyse environnementale dans le modèle multiniveaux (voir ci-après la description des méthodes d'analyse).

16. Concrètement, les variables demandées sont les suivantes :

- toutes les données de revenus, tant au niveau du ménage qu'au niveau individuel :
 - * toutes les variables dites PY et HY ;
 - * les variables I47 et I50 et les variables de contrôle (revenu du travail) et leur complément : I57 à I61, I62 à I66, I69 à I86, I87 à I96 ;
 - * les variables I98, I99, I102, I103 à I114, I115 à I117, I118 à I120, I184 à I186 ;
 - * les variables I121 ;
 - * les variables I129 à I132 ;
 - * les variables H64 à H93 et H98 à H101 ;
 - * toutes les variables d'encadrement ou complémentaires qui sont nécessaires à l'interprétation des variables susmentionnées. Par exemple, des variables qui traduisent combien de mois un revenu a été obtenu (I52, etc.), certains filtres (par exemple I97) et certains poids (PB040) ;

- toutes les informations socioéconomiques, tant au niveau du ménage qu'au niveau individuel :
 - * toutes les variables PB, PE et PL ;
 - * les variables I6 à I40 et I157 à I162 ;

- un certain nombre de données contextuelles (si disponibles – varient de vague en vague), tant au niveau individuel qu'au niveau du ménage :
 - * les variables PH et I133 à I139 ;
 - * les variables I162 à I172 et I173 à I183 ;
 - * toutes les variables HH, HB et HS ;
 - * les variables H5 à H49 et H50 à H63 ;

- les deux fichiers d'enregistrement D-file et R-file.

17. Selon l'institution de gestion, tant les données en coupe pour les années 2004 à 2007 incluse que le fichier longitudinal 2004-2007 sont demandés. Les analyses envisagées par le Chercheur sur les données en coupe sont de nature statique et différent de l'analyse dynamique (sur l'impact des transitions) qu'il prévoit pour l'ensemble de données longitudinales.

E. PROPORTIONNALITÉ

E.1. Quant à la nécessité d'obtenir des données codées

18. Le Chercheur ne peut recevoir les données à caractère personnel codées que si un traitement de données anonymes n'offrirait pas la possibilité de réaliser les finalités statistiques ou scientifiques visées (article 4 de la LVP).

19. L'étude et l'analyse visées ici ne peuvent pas supporter la 'perte de données' d'une éventuelle anonymisation par la DGSIE (par exemple en reprenant les données demandées dans des tableaux indiquant des totaux).

20. La nécessité de pouvoir disposer de données non agrégées est clairement indiquée et la Commission reconnaît dès lors le besoin des données à caractère personnel codées demandées pour les finalités visées. Une communication d'informations simplement anonymes ne peut pas suffire dans ce cas.

21. Les finalités justifient donc le traitement de données à caractère personnel codées.

E.2. Quant à la quantité de données

22. Il ressort de la demande qu'à cet égard, la preuve en la matière est fournie en fonction de l'ensemble des données demandées et par catégorie de données. En effet, le SSE en tant que concept multiple n'est pas uniquement défini par la situation des revenus des personnes mais également par la formation scolaire dont elles ont bénéficié, le travail qu'elles exercent, etc. Cela explique la demande de données de revenus détaillées, de données socioéconomiques et de données sur le contexte du ménage.

23. Selon l'avis juridique et technique, le capital social – la principale variable indépendante de cette recherche – doit être considéré en tant que donnée multidimensionnelle. La composition de ce résultat et les nombreuses variables indépendantes qui peuvent l'influencer rendent les données demandées proportionnelles aux finalités de la recherche. La Commission constate en effet que, étant donné la présente mission de recherche, les variables concernent des thèmes pertinents des questionnaires de SILC et conclut que les données demandées constituent des données essentielles et donc adéquates, pertinentes et non excessives, conformément à l'article 4, § 1, 3° de la LVP.

E.3. Quant au délai de conservation des données

24. Les données à caractère personnel ne peuvent pas être conservées pendant une durée excédant celle nécessaire à la réalisation des finalités pour lesquelles elles sont obtenues ou pour lesquelles elles sont traitées ultérieurement (article 4, § 1, 5° de la LVP).

25. Le délai de conservation des données s'étend jusque décembre 2013 inclus. L'institution de gestion considère qu'un délai de conservation de quatre ans est acceptable pour clôturer correctement la recherche et la Commission se rallie à cet avis.

26. Une fois passé ce délai, les données et sauvegardes doivent être complètement détruites par le Chercheur. Il n'est pas permis de continuer à utiliser les données d'étude codées plus longtemps pour les mêmes finalités, sauf prolongation consentie. Si les finalités sont déjà atteintes avant l'échéance de ce délai, les données et sauvegardes doivent être détruites par le Chercheur avant ce terme, c'est-à-dire immédiatement après la réalisation des finalités.

F. DÉCLARATION

27. Avant de procéder à un ou plusieurs traitements, automatisés en tout ou en partie, des données codées demandées en vue de réaliser les finalités envisagées, le Chercheur doit en faire la déclaration auprès de la Commission.

G. SÉCURITÉ

G.1. Conseiller en sécurité de l'information

28. D'après les documents transmis par le Chercheur, il apparaît que ce dernier dispose d'un conseiller en sécurité de l'information.

G.2. Politique de sécurité

29. Il ressort des documents transmis par le Chercheur que ce dernier dispose d'une politique de sécurité.

30. D'après le formulaire d'évaluation en matière de sécurité qui accompagnait la demande de communication des données et le contrat de confidentialité, on peut établir que les 14 mesures en matière de sécurité généralement recommandées par la Commission lors du traitement de données à caractère personnel sont réalisées, selon le Chercheur : un conseiller en sécurité a été désigné et toutes les autres questions ont également obtenu une réponse positive avec des explications complémentaires sur les mesures prises. Seul le document 'informatieveiligheidsbeleid Universiteit Gent' (politique de sécurité de l'information de l'Université de Gand) auquel il est fait référence n'a pas été joint au dossier. Toutefois, selon l'institution de gestion, les mesures prises, telles que décrites dans le document d'évaluation en matière de sécurité, témoignent d'une politique de sécurité suffisante et de qualité et la Commission se rallie à cet avis.

G.3. Personne physique responsable

31. L'identité de la personne physique responsable a été communiquée. Celle-ci est personnellement responsable du respect de toutes les obligations concernant l'exécution de la loi statistique publique, de la LVP, de leurs arrêtés d'exécution, de toute autre disposition légale ou réglementaire visant la protection de la vie privée, des dispositions de la présente décision de la Commission et des dispositions du contrat de confidentialité.

32. Cette personne exercera un contrôle effectif de l'utilisation licite des données fournies.

33. Les mesures dont il est question aux points G.1. à G.3. inclus, qui doivent garantir la protection et la sécurité des données d'étude transmises, comme l'exigent l'article 16 de la LVP et l'article 15 *bis* de la loi statistique publique, sont efficaces si elles sont contrôlées et suivies de façon stricte dans la pratique.

G.4. Séparation des autres traitements

34. Le Chercheur doit séparer le présent traitement des données dont il est ici question pour les finalités indiquées des autres traitements de données à caractère personnel qu'il contrôle éventuellement.

G.5. Interdiction de décodage

35. Le Chercheur doit s'engager contractuellement à mettre en œuvre tous les moyens possibles afin d'éviter que soit retrouvée l'identité des personnes auxquelles les données à caractère personnel communiquées se rapportent.

G.6. Interdiction de couplage

36. Le Chercheur ne peut pas tenter de coupler les données à caractère personnel obtenues à des données à caractère personnel qui lui ont déjà été transmises en application d'autres autorisations.

G.7. Confidentialité

37. Le Chercheur s'engage à respecter la confidentialité des données d'étude et à veiller à ce que ces dernières ne soient utilisées que par les membres de son personnel en vue de l'exécution de la recherche visée. En l'occurrence, il s'agirait d'un doctorant du GhIPS cité nommément.

H. AUTRES CONDITIONS D'UTILISATION

H.1. Diffusion des résultats

38. Les documents stipulent que les données qui peuvent être publiées se limiteront à des tendances générales et à des corrélations macrosociologiques.

39. Le Chercheur doit en effet veiller à ce qu'après l'analyse et l'utilisation des données, les résultats publiés restent anonymes et globaux de sorte que les données individuelles ne puissent pas être identifiées directement ou indirectement à partir de ces résultats.

40. Les résultats ne peuvent donc être diffusés que sous une forme globale et anonyme.

41. Au minimum deux semaines avant la diffusion, le Chercheur doit d'ailleurs soumettre la publication envisagée à la DGSIE.

H.2. Objectif scientifique, normes scientifiques et méthodes d'analyse

42. Pour le GhIPS, destinataire au sens de l'article 15, premier alinéa, 4° de la loi statistique publique, se pose l'exigence de décrire les méthodes d'analyse qui seront utilisées lors de la recherche et de prouver que le projet de recherche répond aux normes scientifiques en vigueur.

43. En ce qui concerne les méthodes d'analyse qui seront utilisées, le Chercheur précise qu'au moyen d'une analyse factorielle exploratoire (exécutée en SPSS), on vérifiera combien de dimensions empiriques peuvent être distinguées dans le concept mesuré 'capital social individuel'. Le modèle de mesure obtenu sera ensuite testé avec une analyse factorielle confirmatoire (exécutée en Amos). La valeur sur la variable 'capital social individuel' qui est attribuée à une personne interrogée sera le résultat d'une somme pondérée des valeurs des différents indicateurs de capital social qui sont reprises dans le modèle de mesure. Les poids sont définis par les charges des éléments sur les facteurs. Une analyse multiniveaux permettra de vérifier la relation entre le capital social individuel d'un côté et les indicateurs de SSE et l'urbanisation de l'autre côté. Cette technique permet également d'analyser les données longitudinales dans un seul modèle, dans lequel le moment de mesure (établi chez un individu) est considéré comme le niveau le plus bas de l'analyse. Les autres niveaux sont l'individu, le ménage et la commune.

44. En ce qui concerne les normes scientifiques en vigueur auxquelles la recherche doit répondre, le Chercheur précise que le projet de recherche pour lequel les données seront utilisées est financé via un partenariat avec le FWO. Le fait que la proposition de recherche ait été approuvée au FWO montre déjà que son but est considéré comme de grande valeur par la communauté scientifique. La concurrence pour un financement au FWO témoigne aussi de cette valeur. Les résultats de la recherche doivent non seulement conduire à une thèse pour l'obtention du titre de doctorat mais seront également publiés dans la littérature spécialisée nationale et internationale. De cette manière, le projet de recherche doit contribuer à la connaissance scientifique de la répartition du pouvoir et du bien-être en général et du capital social en particulier.

45. Selon l'institution de gestion, le Chercheur envisage aussi bien des analyses statiques que dynamiques et longitudinales. Le projet d'effectuer une analyse environnementale (via ce qu'on appelle le 'multi-level modelling') contribuera également à la qualité scientifique de cette recherche. La méthodologie proposée à ce sujet satisfait aux normes scientifiques actuelles les plus élevées.

46. Pour autant que la Commission ait pu en juger, le caractère scientifique tant de la finalité que de la méthodologie du projet de recherche ne peut en effet pas être contesté.

H.3. Contrôle

47. Le Chercheur accepte expressément que des représentants de la Commission aient, à chaque instant et sans mise en demeure préalable, accès aux locaux et à l'infrastructure informatique où les données communiquées sont conservées, pour contrôler l'exécution des dispositions de la décision

qu'elle a prise, des dispositions de la loi statistique publique et de ses arrêtés d'exécution ainsi que des dispositions du contrat de confidentialité.

48. Sur simple demande, la Commission peut obtenir un accès à d'autres locaux et à d'autres systèmes TIC afin de contrôler si aucune violation des dispositions de sa décision, des dispositions de la loi statistique publique et de ses arrêtés d'exécution et des dispositions du contrat de confidentialité n'est commise.

I. LE CONTRAT DE CONFIDENTIALITÉ

49. Les données d'étude sont communiquées au Chercheur en vertu d'un contrat de confidentialité qu'il conclut avec la DGSIE.

50. Le contrat de confidentialité, qui est inséré en annexe de la demande de communication des données, fixe les conditions auxquelles les données peuvent être transmises par la DGSIE et utilisées par le Chercheur.

51. Le contrat de confidentialité contient au moins les mentions légalement obligatoires telles que définies à l'article 15*bis* de la loi statistique publique, dont la durée du contrat de confidentialité qui, aux yeux de la Commission, est conclu jusqu'à fin décembre 2013. Cela ne signifie aucunement qu'au terme de ce délai contractuel, la confidentialité des données elles-mêmes peut être rompue. Elle doit dès lors être respectée de manière illimitée dans le temps.

52. Les dispositions contractuelles relatives à la vie privée et à la confidentialité figurant dans le contrat de confidentialité sont reprises dans la présente décision de la Commission, ce qui permet ainsi également à des personnes étrangères au contrat de confidentialité de s'adresser à la Commission qui peut donc contrôler le respect des conditions auxquelles les données peuvent être utilisées par le Chercheur.

III DÉCISION GÉNÉRALE

53. Lors du traitement des données à caractère personnel obtenues, le Chercheur doit tenir compte de la LVP, de la loi statistique publique, de leurs arrêtés d'exécution et de toute autre disposition légale ou réglementaire de protection de la vie privée, des dispositions de la présente décision de la Commission et des dispositions du contrat de confidentialité qu'il aura conclu avec la DGSIE.

IV. DÉCISION SPÉCIFIQUE

54. La Commission estime que :

- la présente demande concerne des données en coupe de l'enquête EU-SILC 2004 à 2007 incluse et le fichier longitudinal 2004-2007 ;
- la communication par la DGSIE de ces données d'étude codées au GhIPS est autorisée en vue de la finalité du point C (recherche statistique et scientifique) ;
- la durée de conservation des données et donc la durée du contrat de confidentialité sont limitées à fin décembre 2013, délai au terme duquel la confidentialité des données elles-mêmes doit être respectée de manière illimitée dans le temps.

PAR CES MOTIFS,

la Commission,

autorise la DGSIE à communiquer au GhIPS les données à caractère personnel susmentionnées, aux conditions précitées ;

approuve le contrat de confidentialité y afférent, aux conditions précitées.

Pour l'Administrateur e.c.,

Le Président,

(sé) Patrick Van Wouwe

(sé) Willem Debeuckelaere