



## La Commission de la protection de la vie privée

### Délibération STAT n° 11/2010 du 19 mai 2010

**Objet:** Demande par le DEMO de la communication par la DGSIE de données codées du Registre national extraites des fichiers « état de la population » au 1<sup>er</sup> janvier des années 2006 à 2009 et « mouvements de la population » pour les années 2006 à 2008 (STAT-MA-2010-008)

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)*;

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 vie privée* (ci-après l'AR 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 Centre de recherche en démographie et sociétés (DEMO) de l'Université catholique de Louvain reçue le 15/02/2010;

Vu la demande d'avis technique et juridique adressée au Service public fédéral Economie, PME, Classes moyennes et Energie (Direction générale Statistique et Information économique) le 07/04/2010;

Vu l'avis technique et juridique reçu le 11/05/2010;

Vu le rapport du Président ;

Émet, après délibération, la décision suivante, le 19 mai 2010:

## **I. OBJET DE LA DEMANDE**

1. Le demandeur, le Centre de recherche en démographie et sociétés (DEMO) de l'Université catholique de Louvain, sollicite la communication de la Direction générale Statistique et Information économique (ci-après DGSIE) des données d'étude codées du Registre national extraites des fichiers « état de la population » au 1<sup>er</sup> janvier des années 2006 à 2009 et « mouvements de la population » pour les années 2006 à 2008.
2. La demande porte aussi sur l'approbation du contrat de confidentialité à intervenir entre les parties précitées.

## **II. EXAMEN DE LA DEMANDE**

### **A. LEGISLATIONS APPLICABLES**

#### Loi statistique publique

3. Les articles 15 et 15bis de la loi statistique confient au Comité de surveillance statistique la compétence, d'une part, d'autoriser la communication des données d'étude codées par le DGSIE aux destinataires mentionnés dans la loi, et, d'autre part, l'approbation par ce même Comité du contrat de confidentialité à intervenir entre les parties concernées.
4. L'article 16 de l'arrêté royal du 7 juin 2007 dispose que la Commission est chargée des missions attribuées au Comité de surveillance Statistique jusqu'à l'installation et la nomination des membres de ce Comité.

### La LVP et l'AR du 13 février 2001

5. En vertu des articles 1<sup>er</sup>, § 1<sup>er</sup> et 3, § 1<sup>er</sup> de la LVP et 1<sup>er</sup>, 3<sup>o</sup> de l'AR du 13 février 2001, les données d'étude codées relatives à des personnes identifiées ou identifiables sont des données à caractère personnel dont le traitement est soumis à la LVP et à l'AR du 13 février 2001<sup>1</sup>.

### **B. BASE LÉGALE DE LA DEMANDE**

6. Le demandeur figure au rang des destinataires limitativement énumérés à l'article 15 la loi statistique.
7. En effet, parmi ces destinataires, sont reprises à l'alinéa 1<sup>er</sup>, 4<sup>o</sup> les personnes morales poursuivant un but de recherche scientifique. Le DEMO est un de ces organismes.
8. Le demandeur peut donc introduire la demande d'autorisation pour obtenir la communication des données en question.

### **C. FINALITÉS**

9. Les données à caractère personnel doivent être traitées pour des finalités déterminées, explicites et légitimes (art. 4, § 1<sup>er</sup>, 2<sup>o</sup> de la LVP).
10. Le demandeur sollicite la communication des données d'étude codées en vue de :
  - caractériser l'ensemble des 3000 quartiers et des 262 communes (et de tout autre agrégat) wallonnes à partir de données sociodémographiques actualisées, fiables et strictement comparables;

---

<sup>1</sup> Selon l'article 1<sup>er</sup>, 3<sup>o</sup> de l'arrêté royal du 13 février 2001 exécutant la LVP ainsi que l'Exposé des motifs de la loi du 11 décembre 1998 transposant la Directive 95/46/CE du 24 octobre 1995 du Parlement européen et du Conseil relative à la protection des personnes physiques à l'égard du traitement de données à caractère personnel et à la libre circulation de ces données, 1566/1, 97/98, p. 12, "Sont également considérées comme "données à caractère personnel" les informations codées pour lesquelles le responsable du traitement lui-même ne peut vérifier à quelle personne elles se rapportent, parce qu'il ne possède pas les clés nécessaires à son identification, lorsque l'identification peut encore être effectuée par une autre personne."

- permettre une comparaison aisée entre les caractéristiques de chacune de ces unités spatiales, soit entre elles, soit avec des valeurs agrégées à l'échelle de la région, des provinces, des arrondissements ou de tout autre espace défini. Une comparaison dans le temps est également possible, compte tenu de la disponibilité de l'information;
- proposer un large éventail de tableaux comparatifs et de représentations graphiques et cartographiques directement accessibles permettant d'appréhender au mieux le positionnement de chaque quartier et de chaque commune.

11. Ces finalités ainsi décrites correspondent aux exigences précitées de la LVP.
12. Conformément à l'article 15 de la loi statistique les données codées doivent être communiquées pour des finalités scientifiques ou statistiques. Ces finalités facilitent les prises de décision et aident à la gestion publique.
13. Il ressort de la demande et du projet de contrat de confidentialité que le demandeur peut uniquement utiliser ces données pour l'analyse et l'étude indiquée et l'établissement de statistiques globales et anonymes et que ces données ne peuvent être ni communiquées à des tiers ni utilisées à des fins de contrôle ou de répression.
14. Les exigences de la loi statistique en matière de finalité sont donc respectées.

#### **D. PROPORTIONNALITÉ**

##### Données demandées

15. Il s'agit des données individuelles mais codées du Registre National couvrant la période d'observation allant du 1<sup>er</sup> janvier 2006 au 1<sup>er</sup> janvier 2009. Plus précisément, il s'agit des fichiers « états de la population » pour les 1<sup>er</sup> janvier 2006, 2007, 2008 et 2009 et les fichiers mouvements de la population » pour les années civiles 2006, 2007 et 2008.
- A. Les fichiers « états » fournissent au 1<sup>er</sup> janvier de chaque année, les informations suivantes sur les personnes : mois et année de naissance, lieu de naissance, sexe, nationalité, état civil, la commune de résidence, le secteur statistique de résidence, la situation, la taille et le type de ménage, le type de noyau familial, la position dans le ménage. Comme lors des requêtes de données précédentes, le fichier doit également comprendre sous forme cryptée le numéro d'identification de la personne (IDENPE) et du ménage (IDENT1) ainsi que la rue et le numéro de maison (NRUE et NMAI).

B. Les fichiers « mouvements » concernent toutes les mutations des individus au cours d'une année civile avec le besoin comme variable du type de mutation (naissance, immigration interne, émigration interne, immigration internationale, émigration internationale, décès, migration interne à la commune, changement de nationalité, changement d'état civil,...) de la date (mois et année) de la mutation, de la situation en T1 et T2, de la nationalité, de l'état civil, du sexe, du lieu de naissance. Sont demandés également, sous forme cryptée le numéro d'identification de la personne (IDENPE) et du ménage (IDENT1) ainsi que la rue et le numéro de maison (NRUE et NMAI) dans Situation 1 et Situation 2.

16. Selon la demande, le couplage des fichiers du Registre national entre eux nécessite de disposer des numéros d'identification individuel (IDENPE) et de ménage cryptés (IDENT).

Elle précise que l'identifiant individuel (IDENPE) est indispensable pour pouvoir établir le lien entre les états de la population à différents moments : il est nécessaire de pouvoir suivre chaque individu afin de pouvoir analyser ses mouvements naturels (naissance et décès) et migratoires (entrées, sorties du territoire et mouvements internes) ainsi que ses changements de ménage. Il est également utile pour analyser les changements d'état civil, entre deux moments successifs. Par ailleurs, les chercheurs doivent être en mesure de préciser la raison précise de l'« apparition » ou de la « disparition » de chaque individu au cours d'une année, et cela ne peut se faire qu'en couplant le fichier (Registre national) des mouvements à celui des états de la population.

Quant à l'identifiant ménage (IDENT), le demandeur expose qu'il s'agit d'un numéro séquentiel déterminé par la DGSIE. Cet identifiant est indispensable afin de pouvoir définir la « typologie GédAP » des ménages qui est exploitée dans ses problématiques. La typologie des ménages élaborée prend en considération cet identifiant de ménage et l'adresse des individus qui constituent ce ménage. Elle permet de « coller » au plus près à la réalité des logements, celle proposée par la DGSIE étant trop détaillée.

17. L'avis technique et juridique attire l'attention sur le fait que le demandeur précise la liste des variables en annexe. Cette liste diffère des données demandées dans le questionnaire à plusieurs reprises. Ainsi, le numéro de boîte de l'habitation, les numéros de l'acte de naissance et de l'acte de décès ainsi que la justification légale des changements de nationalité sont demandés dans l'annexe, mais sont absents du formulaire de demande proprement dit. Il est par ailleurs précisé que le numéro de boîte de l'habitation ne doit pas être codé.

Nécessité de la communication de données codées

18. Le Chercheur ne peut recevoir les données à caractère personnel codées que si un traitement de données anonymes n'offre pas la possibilité de réaliser les finalités statistiques ou scientifiques qu'il poursuit (article 4 de la LVP).
19. Il ressort implicitement de la demande que seule, la communication de données codées peut permettre de réaliser la recherche.

Nécessité de la communication de chaque donnée

20. L'avis juridique et technique contient plusieurs remarques quant à la proportionnalité de certaines données demandées.
21. Plus fondamentalement, la DGSIE attire l'attention de la Commission sur les risques non négligeables d'identification indirecte des individus résultant de l'ampleur des données demandées et des possibilités de couplages entre fichiers « mouvements » et fichiers « états » sur base des identifiants individuels et de ménage codés.

En effet, l'identifiant individuel codé rend possible le couplage des caractéristiques de l'individu reprises dans les fichiers « états » avec les caractéristiques des événements reprises dans les fichiers « mouvements ». Par ailleurs, l'identifiant de ménage codé permet de lier ces informations avec les caractéristiques des individus membres des ménages auquel l'individu d'étude a successivement appartenu. Ainsi, on pourrait relier chaque individu avec ces cohabitants, conjoints, enfants... Au final, on aboutit à la possibilité de reconstitution pour un même individu d'une histoire de vie détaillée. Les risques d'identification indirecte sont renforcés par le détail des variables demandées. Ainsi pour un même individu, on connaîtra le sexe, le mois et année de naissance, les nationalités successives pendant la période d'étude, les états civils successifs, les lieux de résidence successifs au secteur statistique près, la date éventuel du décès ou de l'émigration au mois près, l'évolution de toutes ces caractéristiques pour tous les membres du ménage.

En outre, l'accès aux variables « numéro de rue sous forme codée », « numéro de maison sous forme codée » et « numéro de boîte de l'habitation sous forme non codée » permettraient au demandeur d'identifier les personnes habitant la même rue ou le même bâtiment ou le même logement. Cela augmenterait significativement le risque d'identification indirecte. Le fait de vouloir identifier les personnes habitant la même rue ou le même bâtiment ne fait pas partie des

objectifs de la recherche et identifier les personnes habitant le même logement peut être effectué sur base de l'appartenance au même ménage.

22. Les numéros de l'acte de naissance et de l'acte de décès sont demandés dans l'annexe. Couplée avec la variable relative à la commune de naissance/décès et à l'année de naissance/décès (présents dans la demande), chacune de ces informations permet une identification directe des individus. En effet, par année et commune, le numéro de l'acte de naissance/décès est unique.
  
23. Considérant que l'article 15 de la loi statistique publique soumet à autorisation du Comité de surveillance la communication de données d'étude codées c'est-à-dire des données à caractère personnel qui ne peuvent être mises en relation avec une personne physique identifiée ou identifiable que par l'intermédiaire d'un code. Par ailleurs, ce même article dispose in fine que les caractéristiques qui permettent d'identifier le déclarant sont supprimées et munies d'un code, avant d'être communiquées, afin que le responsable de la recherche ne puisse pas raisonnablement identifier le déclarant à l'aide ces données. Qu'il ressort clairement de l'avis technique et juridique émis par la DGSIE que vu l'ampleur des données demandées et des possibilités de couplage entre fichiers « mouvements » et fichiers « états », il y a des risques non négligeables d'identification indirecte, sinon directe des personnes. La Commission fait sienne cette analyse et considère dès lors que les données demandées contiennent des informations qui peuvent être mises en relation avec des personnes physiques identifiées ou identifiables.

### **III. Décision générale**

24. 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**

25. La Commission considère que la communication au Chercheur spécifié des données d'étude codées visées ne peut être autorisée en vue de la finalité exposée en raison des motifs explicités aux points 21 à 23.

**PAR CES MOTIFS,**

**La Commission n'autorise pas** la Direction générale Statistique et Information économique à communiquer les données d'étude codées dont question aux conditions précitées au Centre de recherche en démographie et sociétés (DEMO) de l'Université catholique de Louvain.

L'approbation du contrat de confidentialité est sans objet.

Pour l'Administrateur e.c.,

Le Président,

(sé) Patrick Van Wouwe

(sé) Willem Debeuckelaere