

**Comité sectoriel de la Sécurité sociale et de la Santé  
Section « Sécurité sociale »**

CSSS/08/155

**DÉLIBÉRATION N° 08/047 DU 2 SEPTEMBRE 2008 RELATIVE À LA COMMUNICATION DE DONNÉES À CARACTÈRE PERSONNEL CODÉES PAR LA BANQUE CARREFOUR DE LA SÉCURITÉ SOCIALE À L'UNIVERSITÉ LIBRE DE BRUXELLES, À LA KATHOLIEKE UNIVERSITEIT LEUVEN ET À L'UNIVERSITÉ CATHOLIQUE DE LOUVAIN, EN VUE D'OPTIMALISER LE PROJET OASIS GRÂCE À L'INTÉGRATION DE DATAMINING ET DE MIEUX DÉLIMITER LE PHÉNOMÈNE DE FRAUDE SOCIALE**

Vu la loi du 15 janvier 1990 *relative à l'institution et à l'organisation d'une Banque-carrefour de la sécurité sociale*, notamment les articles 5 et 15, § 1<sup>er</sup> ;

Vu la demande de l'Université libre de Bruxelles, de la Katholieke Universiteit Leuven et de l'Université catholique de Louvain du 11 août 2008 ;

Vu le rapport d'auditorat de la Banque Carrefour de la sécurité sociale du 20 août 2008 ;

Vu le rapport du président.

**A. OBJET DE LA DEMANDE**

- 1.1.** L'Université libre de Bruxelles, la Katholieke Universiteit Leuven et l'Université catholique de Louvain réalisent actuellement une étude avec un double objectif: d'une part, optimaliser le projet OASIS en y intégrant la technique de datamining et, d'autre part, mieux délimiter le phénomène de fraude sociale.

Dans le cadre de la coopération entre les services d'inspection sociale, il existe actuellement un système d'alerte pour les risques de fraude, c'est le projet OASIS. Les inspecteurs sociaux se basent sur OASIS pour sélectionner les entreprises à

contrôler. Le système fournit également des informations sociales de synthèse par entreprise. Un objectif important d'OASIS consiste à réagir rapidement à des cas de fraude éventuelle. La sélection des entreprises est effectuée sur la base d'avertisseurs qui ont été intégrés dans le système sur base de l'expertise des inspecteurs. Jusqu'à présent il n'existe aucune application (datamining) qui puisse contrôler, valider ou corriger ces avertisseurs. Par ailleurs, les données ont, jusqu'à présent, uniquement été utilisées par les services d'inspection sociale.

La modélisation dynamique, rendue possible par le datamining, permettra une description plus détaillée des (nouveaux) mécanismes de fraude : les services d'inspection, les responsables politiques et les chercheurs intéressés seront à même de mieux cerner le phénomène tel qu'il se produit en Belgique.

En vue de ce datamining systématique, la sélection actuelle de données dans le datawarehouse OASIS devra être confrontée au « cadastre des enquêtes » de l'inspection sociale. Ceci permettra d'optimiser les avertisseurs et les scénarios de fraude existants pour les services de contrôle concernés. Par ailleurs, il convient d'examiner comment une interface pour les divers groupes-cibles peut être développée à partir du datawarehouse existant, c'est-à-dire pour les services publics, les responsables politiques, les scientifiques et le public intéressé. Grâce à un accès différencié au datawarehouse OASIS et à l'interface, les divers groupes-cibles (inspection sociale, services publics fédéraux, universités) peuvent consulter et utiliser les résultats d'enquêtes, les indicateurs et les données qui sont pertinents pour eux.

- 1.2. Le présent projet de recherche vise exclusivement à optimiser la détection d'entreprises potentiellement coupables de fraude par l'agrégation de données au niveau du travailleur en données au niveau de l'entreprise selon la méthode de détermination des alertes actuellement utilisée dans OASIS.

Les données nécessaires à l'exécution de cette étude figurent dans les banques de données fiscales et sociales suivantes : le répertoire des employeurs, la banque de données à caractère personnel DmfA, la banque de données à caractère personnel DIMONA, le fichier « chantiers et sous-traitants » (fichier des chantiers, des maîtres d'ouvrage, des contractants principaux et des sous-traitants), le fichier CDIE (contient des informations sur les dettes éventuelles auprès de l'ONSS ou du fonds de sécurité d'existence concerné), les fichiers « comptes » et « procédures » (le fichier « comptes » contient le solde des cotisations dues à l'ONSS et le fichier « procédures » contient les procédures en cours auprès de l'ONSS), le fichier relatif au chômage temporaire (ONEm), les fichiers de la TVA (ceux-ci contiennent des données concernant les dettes envers la TVA, les déclarations à la TVA, les transactions entre clients et fournisseurs et la signalétique) et le fichier GENESIS (le « cadastre des enquêtes »).

Trois de ces banques de données contiennent des données au niveau du travailleur (personnes physiques), à savoir la banque de données à caractère personnel DmfA,

la banque de données à caractère personnel DIMONA et le fichier des chômeurs de l'ONEm. Les informations seront agrégées jusqu'au niveau de l'employeur / de l'entreprise par les fournisseurs des données de base, c'est-à-dire par les gestionnaires des banques de données à caractère personnel concernées. Ensuite, ces données de base seront enregistrées dans le datawarehouse marché du travail et protection sociale et seront couplées au niveau de l'employeur / entreprise. Le numéro d'identification de l'employeur / entreprise sera codé avant de transmettre les données au chercheur.

Il y a lieu de remarquer que certains employeurs / certaines entreprises concerné(e)s sont des personnes physiques.

**1.3.** Les données issues du répertoire des employeurs sont les suivantes :

- l'enregistrement ONSS + numéro de contrôle (codé),
- le rôle linguistique administratif,
- le rôle linguistique de l'employeur,
- la date d'affiliation (année et mois),
- la date de radiation d'un employeur auprès de l'ONSS (année et mois),
- le code de l'entité (arrondissement),
- la date de mise à jour (année et mois),
- le numéro TVA (codé),
- le code "forme juridique" de l'entreprise,
- l'abréviation de la forme juridique ou sa description,
- le type d'employeur,
- la date d'attribution de la catégorie (année et mois),
- la catégorie ONSS - l'activité économique de l'employeur (trois éléments numériques),
- la catégorie ONSS - date de radiation (année et mois),
- la catégorie ONSS - la catégorie initiale (trois éléments numériques),
- la catégorie ONSS - la nouvelle catégorie (trois éléments numériques),
- le code NACE (l'activité économique) (cinq éléments numériques),
- le code du lieu du siège d'exploitation (arrondissement),
- le code régionalisation (indication de la présence de succursales),
- le décret linguistique,
- le code importance (catégories de dimension),
- le code apprenti,
- le matricule de l'enregistrement traité (codé),
- le matricule de destination (codé),
- le matricule de destination - date de transfert (année et mois),
- le matricule de destination - code reprise (codé),
- le matricule d'origine (codé),
- le matricule d'origine - date de transfert (année et mois),
- le matricule d'origine - code reprise (codé),
- le numéro du secrétariat social (codé),
- le code répertoire juridique,

- le code postal de l'employeur (arrondissement),
- la présence ou non d'un mandataire ou d'un curateur.

**1.4.** Les données issues de la banque de données à caractère personnel DmfA sont les suivantes :

- l'année et le trimestre de la déclaration,
- le matricule ONSS codé,
- la catégorie d'employeur.

Variables calculées au niveau de l'entreprise :

- le nombre d'ouvriers à temps partiel,
- le nombre de travailleurs déclarés par catégorie,
- le nombre de jours rémunérés par catégorie,
- le nombre de jours rémunérés par catégorie (uniquement ouvriers),
- le nombre d'ouvriers avec le code travailleur 10, 14, 15, 22, 26, 27, 35,
- le nombre de jours assimilés par type et par catégorie,
- le salaire des ouvriers,
- le nombre d'ouvriers,
- la moyenne d'heures par semaine du travailleur,
- le nombre d'heures de la prestation,
- le numéro de la commission paritaire.

**1.5.** Les variables au niveau de l'entreprise issues de la banque de données à caractère personnel DIMONA sont :

- le nombre de NISS (le nombre de personnes physiques),
- le nombre de travailleurs entrants,
- le nombre de travailleurs sortants.

**1.6.** Les variables au niveau de l'entreprise en provenance de l'ONEm sont :

- le nombre de jours de chômage temporaire pour le type P01 (raisons économiques),
- le nombre de jours de chômage temporaire pour le type P02 (intempéries),
- le nombre de travailleurs en chômage temporaire par type.

**1.7.** Les données en provenance de la TVA sont, outre le numéro de TVA codé :

Données relatives à la déclaration TVA :

- la date d'inscription (année et mois),
- la nature de la déclaration,
- le régime,
- le code demande de remboursement,

- le montant de la déclaration,
- la période (années et mois concernés),
- le numéro du bureau compétent (codé),
- le code NACE (cinq éléments numériques).

Données relatives aux dettes envers la TVA :

- le montant des dettes de TVA,
- le numéro CRTI (centres régionaux de traitement de l'information),
- le bureau de recette de la TVA (codé).

Données relatives aux clients et aux fournisseurs :

- le numéro TVA du client (codé),
- le numéro TVA du fournisseur (codé),
- le montant.

Données relatives à la signalétique :

- le statut,
- la date de suite régime fiscal (année et mois),
- la date de fin (année et mois),
- l'indication de faillite,
- le bureau de contrôle compétent (codé).

**1.8.** Les données des fichiers « comptes » et « procédures » sont les suivantes :

- l'enregistrement ONSS + le numéro de contrôle (codé),
- le solde du compte en cents,
- le code expliquant le solde du compte,
- le numéro de la procédure,
- la date de clôture de la procédure (année et mois).

**1.9.** Les données des fichiers « chantiers et sous-traitants » sont les suivantes :

Données relatives aux chantiers :

- le numéro d'identification du chantier dans 30 bis + numéro de contrôle (codé),
- la date de signature du contrat (année et mois),
- la date de début des travaux,
- la date de fin des travaux (année et mois),
- la date de fin du chantier,
- le solde de crédit du chantier,
- le solde débiteur du chantier,
- le montant total prévu pour le chantier.

Données relatives au maître d'ouvrage :

- le numéro d'identification ONSS + numéro de contrôle (codé),
- le numéro TVA (codé),
- le numéro d'enregistrement (codé),
- le numéro d'enregistrement ONSS + numéro de contrôle (codé),
- le numéro BCE (codé),
- le pays si le siège se situe à l'étranger,
- la date d'inscription comme maître d'ouvrage (année et mois).

Données relatives au contractant principal :

- le numéro d'identification ONSS + numéro de contrôle (codé),
- le numéro TVA (codé),
- le numéro d'enregistrement (codé),
- le numéro d'inscription ONSS + numéro de contrôle (codé),
- le numéro BCE (codé),
- le pays si le siège se situe à l'étranger,
- la date d'inscription comme contractant principal (année et mois).

Données relatives au sous-traitant :

- le numéro d'identification ONSS + numéro de contrôle (codé),
- le numéro TVA (codé),
- le numéro d'inscription (codé),
- l'ancien numéro d'identification (codé),
- le nouveau numéro d'identification (codé),
- le numéro BCE (codé),
- le pays si le siège se situe à l'étranger.

**1.10.** Les données du fichier CDIE sont les suivantes :

- le numéro d'inscription ONSS + numéro de contrôle de l'employeur (codé),
- le numéro TVA de l'employeur (codé),
- l'indication selon laquelle l'employeur est enregistré ou non,
- la date de validité de l'attestation C30bis en ce qui concerne l'enregistrement,
- l'indication selon laquelle il y a une retenue ou non,
- la date de validité de l'attestation C30bis en ce qui concerne la retenue (année et mois),
- le fait d'être en règle ou non vis-à-vis de l'ONSS,
- le fait d'être en règle ou non vis-à-vis du fonds de sécurité d'existence,
- la date à laquelle une décision a été « forcée » (année et mois).

**1.11.** Les données issues du « cadastre des enquêtes » (GENESIS) sont les suivantes :

- le numéro de référence de l'enquête,
- le service d'inspection concerné,

- le motif de l'enquête,
- l'instance qui a demandé l'enquête,
- la date d'ouverture de l'enquête,
- la date de clôture de l'enquête,
- le bureau régional concerné,
- le NISS codé de l'employeur (personne physique),
- l'année de naissance de l'employeur (personne physique),
- le sexe de l'employeur (personne physique),
- la nationalité de l'employeur (personne physique),
- les dispositions sur lesquelles l'enquête est fondée,
- le résultat de l'enquête.

**1.12.** Les données demandées portent sur tous les employeurs dans l'ensemble des secteurs. Il s'agit d'environ 250.000 entreprises. Toutes les données sont demandées sur la base de chiffres mensuels ou trimestriels avec un maximum de cinq ans, à compter des données les plus récentes qui soient disponibles (soit la période 2002-2006, soit la période 2003-2007). Les données de la banque de données GENESIS sont demandées depuis la création de la banque de données « cadastre des enquêtes » jusqu'à la date de fourniture des données.

La Banque Carrefour de la sécurité sociale se charge de la collecte et du couplage des données, ainsi que du codage des numéros d'identification des entreprises avant de transmettre les données à l'Université libre de Bruxelles, à la Katholieke Universiteit Leuven et à l'Université catholique de Louvain.

**1.13.** Les données à caractère personnel seraient conservées jusqu'au 31 décembre 2010, après quoi elles seraient détruites.

## **B. EXAMEN DE LA DEMANDE**

**2.1.** En vertu de l'article 5, § 1<sup>er</sup>, 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 sécurité sociale.

**2.2.** Les données à communiquer portent uniquement sur des employeurs et non sur des travailleurs identifiés ou identifiables.

Etant donné qu'il y a des personnes physiques parmi les employeurs concernés, il s'agit (en partie) d'une communication de données à caractère personnel qui, en vertu de l'article 15, § 1<sup>er</sup>, alinéa 2, de la loi du 15 janvier 1990 *relative à l'institution et à l'organisation d'une Banque-carrefour de la sécurité sociale*, doit

faire l'objet d'une autorisation de principe de la section sécurité sociale du comité sectoriel de la sécurité sociale et de la santé.

- 2.3.** L'étude vise d'une part, à optimiser le projet OASIS en y intégrant la technique de datamining et, d'autre part, à mieux délimiter le phénomène de fraude sociale. Il s'agit d'une finalité légitime.

Les données à caractère personnel à communiquer concernent uniquement des employeurs. Les données à caractère personnel éventuelles relatives à des travailleurs des employeurs concernés seront agrégées au niveau de l'employeur. En aucun cas la communication ne portera donc sur des travailleurs identifiés ou identifiables.

Lors de la communication des données à caractère personnel, chaque numéro d'identification d'un employeur est remplacé par un numéro d'ordre unique insignifiant. Les données à caractère personnel à communiquer ne peuvent être mises en rapport avec un employeur (personne physique ou morale) qu'au moyen de ce numéro d'ordre unique insignifiant.

Les caractéristiques personnelles proprement dites restent limitées et sont généralement communiquées en classes. Elles ne semblent pas de nature à permettre une (ré)identification de l'employeur concerné.

Les données semblent, par ailleurs, pertinentes et non excessives par rapport à la finalité précitée.

- 2.4.** Il s'agit d'un traitement ultérieur de données à caractère personnel codées à des fins historiques, statistiques ou scientifiques, ce qui implique le respect des dispositions de la section II du chapitre II de 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*.

Les chercheurs ne sauraient pas réaliser la finalité précitée à l'aide de données anonymes. En effet, ils doivent avoir la possibilité de suivre la situation des intéressés.

- 2.5.** La Banque Carrefour de la sécurité sociale ne communiquera les données à caractère personnel codées qu'après avoir reçu de la Commission de la protection de la vie privée, conformément à l'article 13 de 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*, l'accusé de réception de la déclaration par les chercheurs du traitement à des fins historiques, statistiques ou scientifiques.



- 2.6.** Les chercheurs doivent s'engager contractuellement à mettre en œuvre tous les moyens possibles pour éviter une identification des personnes auxquelles les données à caractère personnel communiquées ont trait.

En toute hypothèse, il est interdit, conformément à l'article 6 de 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*, d'entreprendre toute action visant à convertir les données à caractère personnel codées qui ont été communiquées en des données à caractère personnel non codées.

Le non-respect de cette interdiction est assorti d'une amende variant de cent à cent mille euros en vertu de l'article 39, 1°, de la loi du 8 décembre 1992.

- 2.7.** Conformément à l'article 23 de 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*, les résultats d'un traitement à des fins historiques, statistiques ou scientifiques ne peuvent être publiés sous une forme qui permette l'identification de la personne concernée, sauf si la personne concernée a donné son consentement et qu'il ne soit porté atteinte à la vie privée de tiers ou si la publication de données à caractère personnel non codées est limitée à des données manifestement rendues publiques par la personne concernée elle-même ou ayant une relation étroite avec le caractère public de la personne concernée ou des faits dans lesquelles celle-ci est ou a été impliquée. Sous réserve des exceptions précitées, les résultats de la recherche doivent donc être publiés sous forme anonyme.
- 2.8.** Lors du traitement des données à caractère personnel, les chercheurs doivent tenir compte de la loi du 15 janvier 1990 *relative à l'institution et à l'organisation d'une Banque-carrefour de la sécurité sociale*, 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*, de leurs arrêtés d'exécution et de toute autre disposition légale ou réglementaire relative à la protection de la vie privée.
- 2.9.** Les données à caractère personnel peuvent être conservées pour la durée nécessaire à la réalisation de l'étude et au plus tard jusqu'au 31 décembre 2010. Elles devront ensuite être détruites.

Par ces motifs,

**le comité sectoriel de la sécurité sociale et de la santé, section sécurité sociale,**

autorise la Banque Carrefour de la sécurité sociale à communiquer les données à caractère personnel codées précitées, selon les modalités précitées, à l'Université libre de Bruxelles, à la Katholieke Universiteit Leuven et à l'Université catholique de Louvain, en vue d'optimiser le projet OASIS en y intégrant la technique de datamining et de mieux délimiter le phénomène de fraude sociale.

Yves ROGER  
Président

Le siège du Comité sectoriel de la Sécurité sociale et de la Santé est établi dans les bureaux de la Banque-Carrefour de la Sécurité sociale, à l'adresse suivante : Chaussée Saint-Pierre, 375 – 1040 Bruxelles (tél. 32-2-741 83 11)

