

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

CSI/CSSS/24/146

**DÉLIBÉRATION N° 24/074 DU 7 MAI 2024 RELATIVE À LA COMMUNICATION DE DONNÉES À CARACTÈRE PERSONNEL PSEUDONYMISÉES ISSUES DU DATAWAREHOUSE MARCHÉ DU TRAVAIL ET PROTECTION SOCIALE PAR LA BANQUE CARREFOUR DE LA SÉCURITÉ SOCIALE À L'ONDERZOEKSGROEP ECONOMIE DE LA KU LEUVEN, DANS LE CADRE D'UNE ÉTUDE SUR L'AUTOMATISATION ET L'EMPLOI**

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 demande de l'*Onderzoeksgroep Economie* (ECON) - campus Brussel (Faculteit Economie en Bedrijfswetenschappen de la KU Leuven);

Vu le rapport de la section Innovation et Soutien à la décision de la Banque Carrefour de la sécurité sociale;

Vu le rapport du président.

**A. OBJET**

1. L'*Onderzoeksgroep Economie* de la KU Leuven étudie l'impact du développement technologique (automatisation) sur l'emploi en Belgique (notamment l'impact sur le nombre de travailleurs en fonction de leur niveau de formation et de leur domaine de formation). Dans un premier temps, les entreprises dont le processus de production est automatisé seraient identifiées au moyen d'informations contenues dans leurs comptes annuels (à fournir par la Banque nationale de Belgique). Ensuite, il serait analysé, par entreprise concernée, son évolution de l'emploi et celle de ses fournisseurs (aussi à identifier par la Banque nationale de Belgique<sup>1</sup>) (en effet, les chercheurs souhaitent connaître les changements du marché du travail sur toute la chaîne d'approvisionnement). Enfin, il serait fait appel à des données à caractère personnel issues du datawarehouse marché du travail et protection sociale de la Banque Carrefour de la sécurité sociale (informations relatives au niveau de formation, au domaine de formation, à l'âge et à la rémunération des travailleurs).

---

<sup>1</sup> La Banque nationale de Belgique utilise à cet effet les données relatives à la TVA qui sont disponibles auprès du Service public fédéral Finances. L'organisation reçoit ces données à caractère personnel en application de la délibération n° 15/2018 du 8 mars 2018 de l'ancien comité sectoriel pour l'autorité fédérale, mais elle les utiliserait (aussi) pour une autre finalité. Cette problématique spécifique doit être réglée conformément aux dispositions de l'article 35/1 de la loi du 15 août 2012 *relative à l'institution et à l'organisation d'un intégrateur de services fédéral*.

2. La méthode de travail suivante serait appliquée. La Banque nationale de Belgique détermine la population des entreprises (il s'agit des entreprises qui ont introduit des comptes annuels et avaient du personnel en service dans la période 2003-2021), définit deux groupes - le groupe d'intervention (les entreprises dont les opérations sont automatisées et leurs fournisseurs) et le groupe témoin (les entreprises dont les opérations ne sont pas automatisées et leurs fournisseurs) -, met plusieurs renseignements au niveau de l'entreprise (en ce compris le numéro d'entreprise et le numéro d'immatriculation) à la disposition de la Banque Carrefour de la sécurité sociale, qui traite ensuite, pour toute entreprise, des données à caractère personnel au niveau du travailleur et les agrège (le numéro d'identification de l'entreprise est, à cet égard, remplacé par un numéro de suite unique, sans signification). Il serait indiqué par employeur, par combinaison de valeurs de critères applicables, le nombre de travailleurs qui y satisfont. On présume que pour de nombreuses combinaisons de ces valeurs de critères applicables, seuls quelques travailleurs y satisferont<sup>2</sup>.
3. Par employeur-personne morale (indiqué à l'aide d'un numéro d'ordre unique, sans signification), seuls des renseignements agrégés au niveau du travailleur seraient traités. Les chercheurs souhaitent en particulier utiliser des informations relatives au niveau de formation et au domaine de formation des travailleurs de l'employeur-personne morale. Il s'agit donc du nombre de travailleurs par combinaison possible du *niveau de formation* (une quinzaine d'options) et du *domaine de formation* (une trentaine d'options). Ces informations permettent de vérifier dans quelle mesure l'automatisation des entreprises a un impact différent sur les travailleurs ayant bénéficié d'une formation différente. Seraient également mis à la disposition, par combinaison de niveau de formation et de domaine de formation, (en classes adéquates), la rémunération annuelle imposable brute moyenne et l'équivalent temps plein moyen. Enfin, l'âge moyen des travailleurs (en classes adaptées) serait également communiqué par employeur (et non par combinaison de niveau de formation et de domaine de formation).
4. Le traitement des données à caractère personnel intervient en deux phases. Au cours de la première phase, les chercheurs reçoivent de la Banque Carrefour de la sécurité sociale un fragment des données (3.000 enregistrements), qui sont certes brouillées, avec lesquelles ils peuvent développer des applications. Cela concerne des données relatives aux employeurs indiqués à l'aide d'un numéro d'ordre spécifique et à leurs travailleurs respectifs (leur statut est, eu égard aux chercheurs, traité dans des tableaux agrégés). Dans une deuxième phase, les chercheurs ont accès aux mêmes types de données de la population complète, et ce sur un ordinateur sécurisé dans les bâtiments de la Banque Carrefour de la sécurité sociale et sous la surveillance permanente d'un collaborateur de cette organisation, pour y appliquer les algorithmes qu'ils ont développés, et ils peuvent emporter les résultats de leurs actions uniquement sous forme de données purement anonyme. À cet effet, la Banque Carrefour de la sécurité sociale réalise, au préalable, une analyse de risque « *small cell* ».

---

<sup>2</sup> Bien que les informations relatives aux travailleurs d'un employeur déterminé (indiqué à l'aide d'un numéro d'ordre sans signification, en remplacement de son numéro d'entreprise et de son numéro d'immatriculation) soient agrégées, il peut tout de même être question de données à caractère personnel dans la mesure où seuls quelques assurés sociaux satisfont à une combinaison déterminée de valeurs de critères. Dans ce cas, il est question de données à caractère personnel pseudonymisées étant donné que seule la Banque Carrefour de la sécurité sociale est en mesure de réidentifier les personnes concernées (le demandeur n'est pas en mesure de le faire).

## B. EXAMEN

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

5. Conformément aux dispositions de l'article 15, § 1<sup>er</sup>, de la loi du 15 janvier 1990 *relative à l'institution et à l'organisation d'une Banque Carrefour de sécurité sociale*, toute communication de données à caractère personnel (il s'agit de données à caractère personnel relatives à une personne physique) par une institution de sécurité sociale ou par la Banque Carrefour de la sécurité sociale à une autre institution de sécurité sociale ou à une organisation autre qu'un service public fédéral, un service public de programmation ou un organisme fédéral d'intérêt public doit faire l'objet d'une délibération préalable de la chambre sécurité sociale de santé du Comité de sécurité de l'information.
6. Les informations ont trait à des employeurs individuels-personnes morales (la Banque Carrefour de la sécurité sociale remplace toutefois, au préalable, leur numéro d'identification réel par un numéro d'ordre fictif). La communication d'informations relatives aux employeurs-personnes morales ne doit, en soi, pas faire l'objet d'une délibération du Comité de sécurité de l'information. Par employeur-personne morale, des informations relatives aux travailleurs salariés sont aussi traitées, et ce sous la forme de tableaux agrégés. Si seuls quelques travailleurs satisfont à une combinaison déterminée de valeurs de critères (et qu'il est donc question d'une *small cell* dans le tableau agrégé), il existe un risque de réidentification des personnes concernées et les renseignements ne peuvent pas être considérés comme des données anonymes. Il s'agit de données à caractère personnel pseudonymisées, étant donné que seule la Banque Carrefour de la sécurité sociale est en mesure de déterminer l'identité des travailleurs déterminés. Vu la procédure proposée, les chercheurs mêmes ne sont pas en mesure de réidentifier les personnes concernées<sup>3</sup>.

### Licéité du traitement

7. 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 de données à caractère personnel n'est licite que si, et dans la mesure où, au moins une des conditions explicites mentionnées dans cet article est remplie.
8. La communication de données à caractère personnel pseudonymisées (en ce qui concerne la qualification des informations: voir supra) issues du datawarehouse marché du travail et protection sociale par la Banque Carrefour de la sécurité sociale à l'*Onderzoeksgroep Economie* de la KU Leuven en vue de la réalisation d'une étude sur l'automatisation et l'emploi est par conséquent licite étant donné qu'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.

---

<sup>3</sup> Les données ont trait aux employeurs-personnes morales et à leurs travailleurs respectifs dont les données sont agrégées. Même si les tableaux d'informations agrégées contiennent des *small cells*, les chercheurs ne sont pas en mesure de réidentifier les travailleurs individuels car le numéro d'identification de l'employeur est toujours remplacé par un numéro d'ordre sans signification. Seule la Banque Carrefour de la sécurité sociale est en mesure d'établir un lien entre les données mises à la disposition et les personnes physiques auxquelles elles ont trait. Il s'agit donc de données à caractère personnel pseudonymisées.

### Principes relatifs au traitement des données à caractère personnel

9. 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 (limitation des finalités), elles doivent être adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées (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 (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 (intégrité et confidentialité).

### Limitation de la finalité

10. La communication poursuit une finalité déterminée, explicite et légitime, à savoir la réalisation d'une étude sur l'automatisation et l'emploi. Par leur projet, les chercheurs souhaitent étudier l'impact du développement technologique sur l'emploi en Belgique. À cet effet, ils analyseraient l'impact sur le nombre de travailleurs, répartis en fonction de leur niveau de formation et de leur domaine de formation (pour les entreprises en voie d'automatisation et pour leurs fournisseurs).

### Minimisation des données

11. L'étude a trait à un groupe d'intervention (entreprises ayant un processus de production automatisé et leurs fournisseurs respectifs) et à un groupe témoin (entreprises sans processus de production automatisé et leurs fournisseurs respectifs). Il s'agit in extenso d'employeurs-personnes morales. Leurs numéros d'identification (c'est-à-dire le numéro d'entreprise et le numéro d'immatriculation) ne sont pas communiqués aux chercheurs mais sont remplacés par un numéro d'ordre fictif.
12. Par employeur-personne morale (sélectionné par la Banque nationale de Belgique parce qu'il a introduit des comptes annuels au cours de la période 2003-2021 et avait du personnel en service), des tableaux agrégés relatifs à la situation de leurs travailleurs sont mis à la disposition. Ces tableaux agrégés contiennent le nombre de travailleurs de l'employeur-personne morale, ensuite répartis en fonction de leur niveau de formation, leur domaine de formation, leur âge et leur rémunération.
13. Le nombre de travailleurs est réparti par combinaison possible de niveau de formation et de domaine de formation. Pour toute combinaison semblable (donc par cellule), sont également mis à la disposition (en classes) la rémunération annuelle imposable brute moyenne et l'équivalent temps plein moyen. Par ailleurs, il est communiqué pour tout employeur l'âge

moyen des travailleurs (en classes) (il s'agit de l'âge moyen des travailleurs en tant que tel, et non pour toute combinaison de niveau de formation et de domaine de formation).

14. Ces informations relatives au niveau de formation et au domaine de formation des travailleurs telles qu'elles sont disponibles dans le datawarehouse marché du travail et protection sociale permettent aux chercheurs de vérifier dans quelle mesure l'impact de l'automatisation varie en fonction de la formation des travailleurs. Ils sont en mesure de contrôler, par employeur-personne morale de la sélection (le groupe d'intervention et le groupe témoin), l'évolution de leur effectif de travailleurs respectif.
15. Par combinaison de niveau de formation et de domaine de formation (pour un employeur individuel-personne morale), la rémunération annuelle imposable brute moyenne est indiquée en classes. Ces informations permettent aux chercheurs de vérifier si les salaires des travailleurs augmentent ou diminuent après l'automatisation. L'équivalent temps plein moyen permet aux chercheurs de corriger, le cas échéant, les informations relatives à l'emploi et au salaire.
16. L'âge moyen des travailleurs est réparti en classes et communiqué au niveau de l'entreprise (et donc pas de manière individuelle, par couplage du niveau de formation et du domaine de formation des travailleurs). Les chercheurs souhaitent analyser le rapport entre l'automatisation de l'entreprise et l'âge des travailleurs (ils partent de l'hypothèse qu'une automatisation donne lieu à une plus grande demande de jeunes travailleurs qui sont davantage familiarisés avec les nouvelles technologies).
17. La Banque Carrefour de la sécurité sociale enrichit les informations des entreprises pour lesquelles elle a reçu le numéro d'entreprise et le numéro d'immatriculation de la part de la Banque nationale de Belgique avec des informations issues du datawarehouse marché du travail et protection sociale qu'elle gère (plus précisément des informations relatives aux travailleurs). Elle remplace ces numéros d'identification par un numéro d'ordre unique, sans autre signification.
18. Dans une première phase, la Banque Carrefour de la sécurité sociale fournit aux chercheurs une fraction des entreprises sélectionnées, toutefois d'une manière brouillée (certaines caractéristiques d'une personne sont attribuées à l'autre personne). Les chercheurs sont ainsi en mesure de développer des applications spécifiques dans leurs propres bureaux. Cependant, ils ne sont pas en mesure de retrouver l'identité des travailleurs concernés.
19. Dans une deuxième phase, les chercheurs traitent les mêmes catégories de données de l'ensemble de la sélection d'entreprises afin d'exécuter les applications qu'ils ont développées antérieurement. Cette exécution a toujours lieu sur un ordinateur sécurisé, dans le bâtiment de la Banque Carrefour de la sécurité sociale, sous la surveillance permanente d'un collaborateur désigné à cet effet. Les chercheurs peuvent emporter les résultats de leurs actions uniquement sous la forme de données anonymes.
20. Les renseignements à communiquer au cours de la première phase semblent adéquats, pertinents et non excessifs par rapport à la finalité précitée. Au cours d'une deuxième phase, les chercheurs exécutent les applications qu'ils ont développées dans le bâtiment de la

Banque Carrefour de la sécurité sociale, dans un environnement sécurisé, et ils quittent le bâtiment de la Banque Carrefour de la sécurité sociale avec les résultats de leurs actions sous la forme de données anonymes (à cet effet, la Banque Carrefour de la sécurité sociale réalise, au préalable, une analyse des risques *small cell*).

#### Limitation de la conservation

21. Les chercheurs détruisent les données à caractère personnel pseudonymisées qu'ils ont reçues, au cours de la première phase, de la Banque Carrefour de la sécurité sociale, dès qu'elles ne sont plus nécessaires pour la réalisation de leur étude relative à l'automatisation et à l'emploi, et ce au plus tard le 31 décembre 2027. Cette durée de conservation peut uniquement être prolongée par une décision explicite de la chambre sécurité sociale et santé du Comité de sécurité de l'information.

#### Intégrité et confidentialité

22. Le demandeur met tout en œuvre pour éviter une identification des personnes concernées et s'abstient, en toute hypothèse, de toute tentative visant à convertir les données à caractère personnel pseudonymisées reçues de la part de la Banque Carrefour de la sécurité sociale au cours de la première phase de l'étude 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, en outre, les résultats du traitement qu'il a réalisé uniquement sous une forme qui ne permet d'identifier d'aucune façon les assurés sociaux concernés.
23. La Banque Carrefour de la sécurité sociale prend les mesures spécifiques nécessaires (en fonction du *cas d'utilisation*) à partir d'une *toolbox* comprenant des actions possibles, afin de garantir que les chercheurs ne puissent à aucun moment réidentifier les intéressés.
24. Pour le surplus, le demandeur 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 disposition réglementaire relative à la protection de la vie privée, en particulier le 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 la communication décrite de données à caractère personnel pseudonymisées provenant du datawarehouse marché du travail et protection sociale, par la Banque Carrefour de la sécurité sociale à l'*Onderzoeksgroep Economie* de la KU Leuven, dans le cadre d'une étude sur l'automatisation et l'emploi, telle que décrite dans la présente délibération, est autorisée moyennant le respect des mesures de protection des données qui ont été définies.

La présente délibération entre en vigueur le 24 mai 2024.

Michel DENEYER  
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).