



EDITORIAL

Jo Noppe
Président du Conseil d'administration

Poursuivre l'engagement en faveur de la coopération

L'Institut interfédéral de Statistique (IIS) a été mis en place en 2016 dans le but de renforcer la coopération entre les différents services statistiques des niveaux fédéral et régional, à savoir Statbel, la Banque nationale de Belgique, le Bureau fédéral du Plan, l'IWEPS, l'IBSA et Statistiek Vlaanderen.

Les institutions de l'IIS partagent l'ambition de fournir des statistiques publiques de qualité, indépendantes et pertinentes pour la Belgique et les régions, afin que les gouvernements, les entreprises, les organisations et les citoyens puissent prendre des décisions en toute connaissance de cause.

Afin de concrétiser cette ambition, les institutions coopèrent de manière intensive au sein de l'IIS. Cette coopération a lieu à la fois dans un certain nombre de groupes de travail permanents et dans des projets de coopération spécifiques (via des demandes de coopération ou des SLA) visant à produire de nouvelles statistiques ou à améliorer les statistiques existantes.

Un aperçu exhaustif de tous les groupes de travail et projets de coopération en cours figure chaque année dans le rapport d'activité de l'IIS. Cette newsletter de l'IIS se concentre sur un nombre limité d'entre eux, tels que le groupe de travail sur les indicateurs pour le suivi des objectifs de développement durable (ODD) et le projet de collaboration sur la banque de données des contrats de bail. Une mention spéciale mérite d'être accordée à un autre projet également abordé dans cette newsletter: « *gender based violence survey* ». Il s'agit d'une enquête à grande échelle visant à analyser la violence liée au genre dans notre pays et qui constitue la première nouvelle collecte de données initiée par Eurostat et développée entièrement dans le cadre de l'IIS. L'IWEPS, Statbel, l'IBSA et la VSA/réseau Statistiek Vlaanderen, ainsi que d'autres partenaires, ont uni leurs forces à cette fin depuis 2019 et ont dégagé des ressources considérables ainsi que du personnel.

IIS NEWSLETTER

TABLE DES MATIERES

- p. 1-2** [Editorial](#)
- p. 3-4** [Débat sociétal sur les indicateurs pour les SDG](#)
- p. 5-7** [La banque de données des contrats de bail enregistrés](#)
- p. 8** [Cinq ans de collaboration interfédérale pour mesurer et mieux comprendre les violences liées au genre](#)
- p. 9-11** [Microdonnées pour la recherche statistique et scientifique](#)

Ces efforts étaient nécessaires compte tenu de la complexité et de l'ampleur du projet, y compris l'élaboration de la méthodologie d'enquête, la conception du questionnaire, l'élaboration et l'attribution à un prestataire externe du travail de terrain, le suivi du travail sur le terrain, le nettoyage des données, la pondération et l'analyse des résultats. Après presque 5 ans de collaboration, les microdonnées ont été livrées à Eurostat à la mi-2023 et les premiers résultats pourront être publiés.

Mais la coopération au sein de l'IIS ne s'arrête pas aux groupes de travail et aux projets de coopération existants. Des procédures simplifiées ont également été convenues au sein de l'IIS pour le partage des microdonnées entre les différentes institutions, ce qui est d'une importance cruciale pour le développement et la production de statistiques publiques fiables et de haute qualité. Vous trouverez plus d'informations à ce sujet dans cette newsletter.

Enfin, cette newsletter est un excellent exemple de la coopération au sein de l'IIS. Grâce à elle, les institutions concernées peuvent apprendre les unes des autres et échanger des bonnes pratiques. C'est un aspect sur lequel l'IIS souhaite se concentrer encore davantage dans les années à venir, afin de renforcer en permanence l'expertise et les compétences de son personnel. Cette ambition a également été reprise dans le plan d'action élaboré en réponse au peer review d'Eurostat sur le système belge de statistiques publiques.

Il nous appartient à tous de continuer à y travailler ensemble dans les années à venir.



Débat sociétal sur les indicateurs pour les SDG

Alain Henry

Un groupe de travail de l'IIS est chargé de coordonner le suivi des *Sustainable Development Goals* (SDG), définis par l'ONU en 2015, sur la base d'indicateurs (voir article dans la Newsletter de mars 2022). À partir des 231 indicateurs globaux définis par l'ONU pour mesurer le progrès vers les SDG, un ensemble de 116 indicateurs a été défini pour la Belgique, dont 82 sont actuellement publiés sur le site www.indicators.be du Bureau fédéral du Plan (BFP). À partir de cet ensemble, le progrès de la Belgique vers les SDG est évalué par le BFP.

Ce dispositif de mesure gagnerait en pertinence si cet ensemble d'indicateurs faisait l'objet d'un consensus sociétal.

Cet article décrit dans une première section le processus mis en place pour construire un tel consensus et présente dans une deuxième section les résultats de ce processus.

ORGANISATION DU DÉBAT SOCIÉTAL AUTOUR DES INDICATEURS DE SUIVI DES SDG

Dans un rapport de 2020, la Cour des Comptes recommande: «sur la base des listes d'indicateurs existant au niveau international, l'Institut interfédéral de statistique (IIS) doit élaborer des indicateurs visant à couvrir tous les aspects des SDG». En réponse, le Plan fédéral de développement durable (PFDD), adopté le 1er octobre 2021, prévoit un débat sur les indicateurs de suivi des SDG avec les experts des services publics, de l'IIS et des principaux conseils d'avis fédéraux.

Le processus de débat a été structuré en plusieurs étapes. Le BFP et l'IIS ont d'abord rédigé une note préparatoire, présentant l'ensemble de 82 indicateurs actuellement publiés (parmi ceux-ci, 37 sont spécifiques à la Belgique et ne sont donc pas des indicateurs de l'ONU), une liste de 34 autres indicateurs de l'ONU pour lesquels les données existent en Belgique, ainsi qu'une liste de 118 indicateurs de l'ONU pour lesquels un travail de recherche est nécessaire avant de construire ces indicateurs pour la Belgique.

Cette note a été soumise en mars 2023 aux experts des services publics, réunis dans le cadre de la Commission interdépartementale de développement durable (CIDD). Il en est ressorti une nouvelle note, avec un ensemble plus riche de propositions d'indicateurs à inclure ou à exclure du suivi actuel, ainsi qu'à développer en priorité.

Cette nouvelle note a servi d'input à une discussion avec la société civile en avril 2023, organisée par le Conseil fédéral du développement durable (CFDD). Les organisations participant à différents conseils d'avis, tant fédéraux que régionaux, y ont été invitées, de même que d'autres organisations ayant manifesté un intérêt sur les indicateurs lors de la consultation sur le PFDD. À cette discussion était représenté un large échantillon d'associations patronales, syndicales et issues des milieux associatifs (défense de l'environnement et des consommateurs, jeunes, Nord-Sud...).

Sur la base de la note préparatoire BFP/IIS et des discussions avec les experts des services publics et avec la société civile, le BFP a rédigé une note de synthèse¹. Cette note, qui présente les principaux résultats et qui a été validée par le groupe de travail de l'IIS sur les indicateurs pour les SDG, est résumée ci-dessous.

1 BFP (2023). Indicateurs de suivi des Sustainable development goals, Synthèse du débat sociétal.

RÉSULTATS

Lors des débats, un total de 157 indicateurs (soit 82 actuellement publiés et 75 autres indicateurs) ont été discutés. À partir de ce débat, les propositions de modifications de l'ensemble actuel d'indicateurs sont les suivantes.

- Conserver 67 indicateurs sans changement (des ventilations seront ajoutées pour certains d'entre eux).
- Supprimer 10 indicateurs.
- Elargir la définition de 4 indicateurs et remplacer un indicateur lorsqu'une alternative sera développée.
- Ajouter 12 nouveaux indicateurs de suivi des SDG, parmi les 75 nouveaux indicateurs discutés.

Tous ces changements (sauf le remplacement d'un indicateur) pourraient être réalisés d'ici fin 2025 dans le cadre des moyens actuels de la Task force développement durable du BFP. Fin 2025, l'ensemble d'indicateurs de suivi des SDG pourrait donc inclure 84 indicateurs, dont 16 auraient été redéfinis ou ajoutés. En complément, une réflexion sera également menée pour inclure dans cet ensemble les indicateurs composites de bien-être développés par le BFP dans le cadre de la loi sur les indicateurs complémentaires au PIB.



Pour les 63 nouveaux indicateurs discutés mais qui ne seraient pas repris dans l'ensemble d'indicateurs de suivi des SDG d'ici à 2025, les propositions sont les suivantes.

- Développer 23 indicateurs, dont 9 en priorité, sachant que l'un d'entre eux viendrait remplacer un indicateur existant.

- Ne pas intégrer 40 indicateurs, ceci pour diverses raisons (par exemple une redondance avec un indicateur existant).

Le suivi des SDG est réalisé à partir d'indicateurs existants, collectés et suivis par le BFP. Il faut noter que les 23 indicateurs à développer devront l'être par les administrations publiques compétentes (éventuellement dans le cadre d'un *Service Level Agreement* de l'IIS), pour lesquelles des moyens devraient éventuellement être dégagés. Une fois ceux-ci développés, le BFP pourrait alors les intégrer à l'ensemble des indicateurs de suivi, en fonction des ressources disponibles pour ce faire.

Plusieurs considérations transversales ont également été évoquées.

- L'aspect systémique de la liste a été souligné, que ce soit pour le domaine de la santé ou pour d'autres domaines. En effet, un seul indicateur peut souvent être relié à plusieurs thèmes ou SDG.

- Les discussions ont souligné un problème inhérent à la collecte de statistiques officielles. Construire des séries temporelles longues et respectant les critères de qualité statistique demande du temps, ce qui rend plus difficile la prise en compte de problématiques émergentes.

- Compte tenu de l'importance du pilier européen des droits sociaux de l'UE, il est essentiel que l'ensemble d'indicateurs de suivi des SDG soit cohérent avec les indicateurs et les objectifs de cet engagement.

- Des ventilations selon le handicap devraient être incluses quand celles-ci sont disponibles.

PROCHAINE ÉTAPES

Suite à ce débat sociétal, l'ensemble des indicateurs de suivi des SDG pour la Belgique va être revu pour incorporer les propositions réalisables sans développement de nouveaux indicateurs. Il s'agit en particulier d'élargir la définition de 4 indicateurs, d'ajouter 12 indicateurs pour lesquels des données sont disponibles et de supprimer 10 indicateurs. Ce travail devrait être achevé pour la fin 2025.

D'autre part, des démarches seront entreprises avec l'IIS et les administrations compétentes pour développer des indicateurs prioritaires parmi les 23 propositions issues du débat. Développer de nouveaux indicateurs peut toutefois être un processus long et laborieux et dans certains cas demander des moyens supplémentaires.

SUSTAINABLE DEVELOPMENT GOALS



<https://www.un.org/sustainabledevelopment/>

La banque de données des contrats de bail enregistrés

David Eyckmans
Eva Debusschere

On utilise actuellement assez peu les bases de données administratives pour analyser le marché locatif privé. Un groupe de travail de l'IIS veut combler (partiellement) cette lacune, via des analyses de la banque de données des contrats de bail enregistrés. Dans ce cadre, une étude de faisabilité a été réalisée, qui pourrait aboutir dans une phase ultérieure, à la production effective de statistiques.

CONTEXTE

Depuis 2007, les bailleurs de biens immobiliers en Belgique sont tenus d'enregistrer leurs contrats de bail. Les données relatives à l'enregistrement sont stockées par le SPF Finances (Administration Générale de la Documentation Patrimoniale, AGDP) dans la banque de données des « contrats de bail enregistrés », mieux connue sous le nom de MyRent.

En 2016, un projet de collaboration a été lancé au sein de l'IIS sur les statistiques relatives au marché locatif privé. Ce projet inclut des représentants des autorités statistiques et des départements thématiques consacrés au logement. Le but principal est de vérifier si et comment les données de cette base de données peuvent être utilisées pour établir des séries statistiques. Les analyses se sont donc concentrées sur une exploration des données, les points d'attention ayant été formulés dans une note méthodologique.

DEMANDE DE DONNÉES

La demande de données tirées de cette base de données a été effectuée auprès de l'AGDP par l'intermédiaire de la Structure de Coordination de l'Information Patrimoniale (SCIP). Le demandeur principal était Statbel. Via une simple procédure de demande, les

autres autorités statistiques ont pu obtenir de Statbel les données pour leur propre région - sur la base de l'art. 15ter de la loi statistique. 'Wonen in Vlaanderen' a aussi demandé les données via la procédure de demande normale auprès de Statbel.

DONNÉES

Le groupe de travail a demandé à Statbel d'enrichir les données de la base de données des baux enregistrés avec :

- Des **informations cadastrales**, afin de compléter les caractéristiques du logement loué, comme l'année de construction, le type de logement, la superficie habitable ;

- Le **Registre national**, pour obtenir des informations plus détaillées sur le profil des bailleurs/locataires ;

- et la **base de données DBRIS**, pour affiner l'image du bailleur (professionnel).

Le couplage avec les informations cadastrales n'a pas encore été effectué. Ces données ont été livrées sous forme pseudonymisée pour l'analyse.

La base de données enrichie et pseudonymisée qui en résulte, appelée ici base de données des contrats de bail, se compose de huit fichiers :

- Les **contrats de bail** de la banque de données des contrats de bail enregistrés comprenant l'identification pseudonymisée du bailleur/locataire et du logement en location, le loyer et les charges locatives, la date de début et la durée prévue du contrat, etc.

- L'**adresse** pseudonymisée du logement en location.

- Le **secteur statistique territorial** du logement en location.

- Des **données démographiques** reprenant la nationalité et la taille du ménage du bailleur/locataire au moment du début du contrat de bail et - pour le locataire - également au moment de la domiciliation dans le logement loué.

- Le **sexe et l'année de naissance** du bailleur/locataire.

- La **typologie LIPRO** - une méthode internationalement reconnue de classification des ménages en différents types - du ménage du bailleur/locataire au 1er janvier suivant le début du contrat de bail.

- Les **déménagements** avec le moment auquel le locataire quitte le logement en location et la commune vers laquelle il déménage.

- Les caractéristiques du bailleur s'il s'agit d'une **entreprise** (code NACE, code NIS établissement, etc.).

La base de données des contrats de bail contient plus de 2 millions de contrats pour la période 2011-2020. Pour les contrats enregistrés avant 2011, le numéro de registre national du locataire n'est pas systématiquement disponible.

POINTS D'ATTENTION

L'exploration de la base de données des contrats de bail a mis en lumière quelques points d'attention concernant les données et l'enregistrement des contrats de bail.

Taux de couverture

On ne connaît pas la proportion de contrats de bail effectivement enregistrés par rapport à l'ensemble des nouveaux contrats. L'absence d'informations sur le taux de couverture de la base de données des contrats de bail enregistrés a bien entendu des conséquences : ce n'est que si la base de données est (suffisamment) représentative que l'on peut établir des statistiques sur l'ensemble du marché locatif privé (par exemple, le profil du bailleur/locataire, les loyers, etc.). Si ce n'est pas le cas, nous ne pouvons établir des statistiques que sur les contrats enregistrés (par exemple, les nombres).

Une [étude de l'IWEPS sur le marché locatif privé](#) s'est penchée sur le taux de couverture de l'enregistrement. Sur la base du census de 2011, le pourcentage de ménages qui louaient un logement au 1er janvier 2011 et qui avaient déménagé au cours de l'année 2010 a été examiné. Une estimation (grossière) du nombre de nouveaux contrats de bail conclus en 2010 et qui auraient dû être enregistrés a ainsi pu être effectuée. Ce nombre a été comparé aux contrats effectivement enregistrés. On obtient ainsi un taux de couverture de 50% pour la Belgique (avec des différences régionales : 35% à Bruxelles, 50% en Wallonie et 60% en Flandre). Il serait intéressant de répéter cet exercice pour les années les plus récentes, car la part de contrats enregistrés pourrait avoir augmenté.

Caractéristiques du logement

Les fichiers de données actuellement disponibles contiennent peu d'informations sur les caractéristiques du logement. Ces informations nous permettraient de mieux interpréter les données issues de la banque de données des contrats de bail enregistrés. Ce serait une plus-value afin de pouvoir situer le loyer en fonction du type logement (appartement ou logement unifamilial).

Uniquement pour les nouveaux contrats

Actuellement, les données de la base de données des contrats de bail ne se prêtent qu'à des analyses et des statistiques sur les nouveaux contrats de bail. On enregistre bien une durée de contrat mais la date de la fin effective du contrat n'est pas mentionnée dans la base de données.

On utilise des données sur les déménagements pour déterminer approximativement quand un contrat en cours a pris fin. On peut ainsi aussi déterminer le groupe des contrats en cours durant une période donnée.

Comme nous n'avons pas d'identification du locataire avant 2011, nous ne pouvons pas enregistrer les déménagements de locataires ayant un contrat de bail conclu avant 2011. Nous ne pouvons donc pas encore identifier de manière fiable les contrats de bail en cours. Nous ne savons pas non plus si un contrat de bail prend fin quand le (dernier) locataire décède ou émigre.

Etant donné que l'identification du locataire s'améliore avec le temps, nous souhaitons examiner quel serait la meilleure approche méthodologique pour les contrats en cours, mais aussi pour les décès et les émigrations.

Location sociale et logements pour étudiants

La location sociale ainsi que les logements pour étudiants sont (partiellement) inclus dans la banque de données des contrats de bail enregistrés. Il s'agit de segments spécifiques du marché locatif ayant leurs propres durées de contrat, loyers, etc. Pour pouvoir analyser de façon pertinente le marché locatif privé, il est nécessaire de pouvoir identifier ces segments de marché au sein de la banque de données.

Nous pouvons ainsi éliminer ces segments spécifiques du marché locatif des analyses ou réaliser des analyses distinctes afin de suivre leur évolution.

Nous pouvons assez bien identifier les logements sociaux en location sur la base des codes NACE des bailleurs. Pour la Flandre, nous nous attendons à une amélioration quand les sociétés de logement seront établies en tant que bailleur social unique de la commune.

Pour les logements pour étudiants, nous pouvons, sur la base de certaines caractéristiques comme la durée du contrat, mois de début du contrat, la localisation du logement loué, l'âge, etc., effectuer une délimitation fondée, même si la marge d'incertitude reste supérieure à celle des logements sociaux. Ajouter la catégorie 'chambre (d'étudiant)' au type de logement lors de l'enregistrement constituerait une plus-value.



Sous-locations

Dans certains cas, les logements sont sous-loués. C'est par exemple le cas avec une agence immobilière sociale. Elle loue le logement au propriétaire et le sous-loue ensuite au locataire social. Les fichiers de données ne contiennent cependant pas de contrats avec une entreprise en tant que locataire.

SUITE DES TRAVAUX ?

Les objectifs initiaux ont été (pratiquement) atteints. Une exploration en profondeur a été réalisée et les constatations sont consignées dans une note méthodologique. Une poursuite de la collaboration constituera certainement une plus-value. Ci-dessous sont repris quelques aspects auxquels nous souhaitons consacrer les futurs travaux.

Couplage avec certaines informations cadastrales

D'autres informations sur les caractéristiques du logement, comme l'année de construction, le type de logement ou la superficie habitable, permettraient d'améliorer l'interprétation des statistiques (notamment sur le loyer) et de produire de nombreuses analyses et statistiques complémentaires intéressantes. Le couplage des fichiers aux informations cadastrales présenterait une plus-value importante. Nous examinons ce qu'il est possible de faire ici.

Demande récurrente de données

L'enregistrement le plus récent dans la base de données des contrats de bail disponible date du 1er novembre 2020. Il est souhaitable d'avoir davantage d'informations récentes par exemple pour savoir ce qui s'est passé pendant et après la pandémie de coronavirus sur le marché locatif, compte tenu des mesures politiques spécifiques prises durant cette période. Il est envisagé d'introduire une demande récurrente (annuelle par exemple).

Définitions univoques et analyses transparentes

Lors des analyses – et en particulier lors de la comparaison avec d'autres études –, nous avons constaté qu'il y avait des écarts (importants) dans des résultats qui, à première vue, devraient être identiques. Ces différences peuvent probablement s'expliquer par un couplage de bases de données administratives qui est légèrement différent, par des concepts dont le contenu est différent, par un nettoyage qui n'est pas réalisé exactement de la même façon... Lorsque l'on travaillera davantage sur la banque de données des contrats de bail enregistrés et que des publications suivront, il conviendra d'harmoniser les définitions et les analyses et de bien les documenter.

Statistiques à suivre

Le groupe de travail a examiné s'il est possible d'établir des séries statistiques sur le marché locatif privé sur la base de la banque de données des contrats de bail enregistrés. De nombreuses analyses statistiques descriptives ont été réalisées lors de l'exploration de la base de données. Il est apparu que toutes les analyses n'étaient pas d'une qualité suffisante pour être publiées sous forme de série statistique. Toutefois, certaines présentaient une qualité suffisante (en fonction du taux de couverture). Il s'agit notamment:

- Du nombre de nouveaux contrats de bail enregistrés par an ;
- Du loyer médian (surtout intéressant selon le type de logement) ;
- Du 'Repeated rent index', qui reflète l'évolution de prix d'un nouveau contrat pour le même logement.

Certaines de ces statistiques peuvent être affinées géographiquement, par exemple par type de zone ou jusqu'au niveau de la commune.

Le groupe de travail déterminera dans une phase ultérieure quelles statistiques concrètes peuvent être proposées et produites.

Cinq ans de collaboration interfédérale pour mesurer et mieux comprendre les violences liées au genre

Chloé Janssen
Frédéric Vesentini

ÉTUDIER LES VIOLENCES DE GENRE : UNE DEMANDE INSTITUTIONNELLE ET SOCIALE FORTE

Ces dernières années, la problématique des violences basées sur le genre est devenue un enjeu majeur des politiques publiques. La [Convention d'Istanbul](#), signée et ratifiée par la plupart des pays occidentaux, entrée en vigueur en 2014 et ratifiée par la Belgique en 2016, définit un socle commun d'intervention pour ces politiques publiques. Parmi les engagements de la Convention d'Istanbul, l'article 11 prévoit la production de statistiques et d'études permettant aux décideurs de mieux connaître les situations visées et de mettre en œuvre une *evidence-informed policy*. Cette Convention d'Istanbul s'inscrit par ailleurs comme un point d'aboutissement d'une suite de déclarations et d'engagements internationaux enjoignant de (mieux) quantifier la prévalence des violences faites aux femmes et de (mieux) comprendre leurs causes et conséquences, afin de mettre en œuvre de meilleures réponses pour lutter contre ces formes de violence.

UN BESOIN DE DONNÉES NOUVELLES

En Belgique, au tournant du siècle passé, deux études de grande échelle portant sur la prévalence de la violence exercée à l'encontre des femmes (et des hommes) avaient été menées, en 1988 et en 1998. En 2008, l'Institut pour l'égalité des femmes et des hommes (IEFH) a fait réaliser une [nouvelle étude](#) d'ampleur sur la survenance, l'expérience et le recours à la violence physique, sexuelle et psychologique. En 2012, c'était au tour de l'*European Union Agency for Fundamental Rights* (FRA) de mettre en œuvre, au niveau européen, une [enquête](#) portant sur les violences contre les femmes. Depuis, des événements tels que le mouvement [#metoo](#) ou les confinements stricts imposés

dans le cadre de la lutte contre la pandémie de COVID-19 ont mis la problématique des violences de genre davantage encore au centre du débat public. Le besoin de données nouvelles pour éclairer les plans d'action devenait donc particulièrement criant.

UNE ENQUÊTE EUROPÉENNE DANS LE CADRE D'UNE COLLABORATION INTERFÉDÉRALE

En 2016, un groupe de travail a été constitué au sein du Système statistique européen pour construire une enquête sur les violences de genre. Le périmètre est large puisqu'il envisage les violences subies au travail, entre (ex-)partenaires ou encore dans l'enfance. Une fois finalisée, cette [enquête « GBV Survey »](#) a ensuite été proposée aux États (membres de l'UE ou partenaires), avec pour horizon des collectes de données échelonnées entre 2020 et 2023. En Belgique, c'est bien la collaboration interfédérale qui a été le vecteur permettant de mettre en œuvre cette enquête, à travers un groupe de travail piloté par l'IWEPS et associant Statbel, l'IBSA et Statistiek Vlaanderen¹. Il faut dire que la lutte contre les violences de genre a été affichée comme priorité de tous les gouvernements du pays et qu'il y avait de réels avantages à la collaboration : données cohérentes et comparables, à l'intérieur de la Belgique mais aussi à l'international, réduction des coûts, partage d'expertise, etc.

Finalement, depuis 2019, chaque année apparaît dans le programme statistique intégré de l'IIS un SLA² lié à cette enquête. Le premier visait à réaliser une étude de faisabilité financière d'un tel projet. Se sont ensuite succédées jusqu'au programme statistique 2023 des collaborations annuelles visant notamment la formalisation d'un accord entre les gouvernements finançant l'enquête, la mise au point de la

méthodologie d'échantillonnage, la traduction du questionnaire, la rédaction d'un cahier de charges pour la réalisation de l'enquête, l'attribution du marché, le suivi du prestataire, la formation des enquêteur-trices ou encore la vérification et le nettoyage des données. Cette année 2023 verra l'aboutissement d'un processus de cinq ans de travail, avec la livraison des données à Eurostat³, la mise à disposition par Statbel des micro-données (plus de 5750 réponses⁴) aux acteurs de la recherche ainsi que la rédaction d'un rapport statistique piloté par l'IWEPS.

BILAN ET PERSPECTIVES

Cette collaboration a assurément été une expérience très enrichissante pour les parties impliquées, tant sur les plans méthodologique et scientifique qu'organisationnel. En effet, vu la forme interfédérale de la collaboration et la structure du financement, il s'est agi à chaque étape de veiller à assurer une représentativité statistique minimale pour chaque région, et ce malgré des taux de réponse très variables, mais aussi d'assurer une comparabilité complète des données entre régions.

Cette expérience partenariale apparaît aujourd'hui d'autant plus importante que l'enquête pourrait devenir à terme une enquête régulière et obligatoire au sein du Système statistique européen. Si telles étaient les perspectives, l'architecture statistique belge serait sans aucun doute bien préparée pour relever le défi.

1 Les instituts statistiques ont en outre associé à la démarche et au groupe de travail une série d'institutions partenaires, en prise avec la problématique : l'Institut pour l'égalité des femmes et des hommes, Equal Brussels, Bruxelles Prévention & Sécurité, l'Agentschap Binnenlands Bestuur, le Département Welzijn, Volksgezondheid en Gezin.

2 Les collaborations entre les partenaires de l'IIS sont réglées par voie de Service Level Agreements (SLA).

3 La collecte proprement dite s'est étalée entre l'été 2021 et l'été 2022.

4 Dont plus de 80% de femmes.

Microdonnées pour la recherche statistique et scientifique

Erik Meersseman

INTRODUCTION

Les autorités statistiques ont pour mission de produire des statistiques de qualité de manière indépendante et de les diffuser.

Afin de produire ces statistiques, les autorités statistiques ont accès à des données existantes provenant d'institutions publiques ou privées ou collectent des données auprès des citoyens ou des entreprises au moyen d'enquêtes.

La diffusion de statistiques se fait principalement via des statistiques globales et anonymes. Il s'agit de résultats où un chiffre se rapporte toujours à plusieurs citoyens ou entreprises. Aucune information sur des entreprises ou citoyens individuels identifiables n'est donc divulguée.

Pour certains chercheurs, ces statistiques globales et anonymes ne suffisent pas pour réaliser leurs travaux. La législation prévoit que les autorités statistiques peuvent, sous certaines conditions, communiquer des microdonnées (données individuelles), où chaque ligne du fichier contient les données d'une personne, d'un ménage ou d'une entreprise.

Cet article décrit les possibilités et la procédure permettant aux chercheurs d'obtenir des microdonnées dont dispose Statbel, l'office belge de statistique, à des fins statistiques et scientifiques.

CADRE LÉGAL

Le traitement des données par Statbel ainsi que la communication de ces données à des tiers sont régis légalement par la *Loi du 4 juillet 1962 relative à la statistique publique* (ci-après la 'loi statistique'). La loi statistique part du principe que toutes les données obtenues et traitées par Statbel pour la production de statistiques sont protégées par le secret statistique.

De même, le règlement (CE) n° 223/2009 du Parlement européen et du Conseil du 11 mars 2009 relatif aux statistiques européennes (ci-après le 'règlement sur la statistique') protège les données collectées par Statbel contre la réutilisation à des fins autres que statistiques et scientifiques.

Par ailleurs, en ce qui concerne les données à caractère personnel, le Règlement général sur la protection des données (Règlement (UE) n° 2016/679 du Parlement européen et du Conseil du 27 avril 2016 - ci-après le 'RGPD') et la *Loi du 30 juillet 2018 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel* (ci-après la 'loi-cadre RGPD') sont également applicables.

Il convient de noter que la loi statistique et le règlement sur la statistique protègent également les données portant sur des entreprises ou des personnes décédées, ce qui n'entre pas dans le champ d'application du RGPD.

SECRET STATISTIQUE

Tant la loi statistique que le règlement sur la statistique prévoient que Statbel, en tant qu'autorité statistique, peut collecter des données à l'aide d'enquêtes et accéder aux données administratives existantes pour produire des statistiques. Afin de protéger les citoyens et entreprises sur lesquels portent les informations, le législateur a disposé que les données collectées par Statbel pour la production de statistiques sont protégées par le secret statistique. Concrètement, cela signifie que les données qui portent sur des citoyens ou des entreprises peuvent uniquement être utilisées pour produire des statistiques globales et anonymes. Les données permettant l'identification directe ou indirecte d'un citoyen ou d'une entreprise ne peuvent donc pas être utilisées par les administrations pour compléter ou améliorer des fichiers de données ou pour prendre une décision ayant des conséquences pour un citoyen ou une entreprise au niveau individuel. Les informations ne peuvent pas non plus faire partie d'une enquête pénale ou d'un témoignage devant un tribunal.

Outre Statbel, les autorités statistiques régionales sont également soumises au secret statistique sur la base de l'accord de coopération du 15 juillet 2014¹.

1 Accord de coopération du 15 juillet 2014 entre l'Etat fédéral, la Région flamande, la Région wallonne, la Région de Bruxelles-Capitale, la Communauté flamande, la Communauté française, la Communauté germanophone, la Commission communautaire commune de Bruxelles-Capitale et la Commission communautaire française concernant les modalités de fonctionnement de l'Institut interfédéral de statistique, du conseil d'administration et des Comités scientifiques de l'Institut des comptes nationaux.

POUR QUELLES RAISONS LES MICRODONNÉES PEUVENT-ELLES ÊTRE DEMANDÉES ?

Les données peuvent être demandées uniquement dans un but de recherche statistique et scientifique. Cet objectif peut être interprété au sens large. Les recherches destinées à soutenir la prise de décisions politiques entrent également en ligne de compte.

Le résultat du traitement doit aboutir à des statistiques globales et anonymes ou à des rapports d'enquête qui ne peuvent entraîner aucune conséquence individuelle pour les citoyens ou les entreprises. Naturellement, ces résultats globaux et anonymes publiés peuvent ensuite être librement utilisés à des fins non statistiques ou non scientifiques.

QUI PEUT DEMANDER DES DONNÉES ?

La loi statistique énumère les destinataires autorisés à recevoir des microdonnées de Statbel. Il s'agit tout d'abord des administrations fédérales, régionales, provinciales et locales et des organismes publics (p.ex. Sciensano, SPF ETCS, SPF Sécurité sociale). Appartiennent également à cette catégorie les autorités statistiques régionales, le Bureau fédéral du Plan et le service statistique de la Banque nationale de Belgique, qui ont un statut privilégié (voir plus loin). Les organisations multilatérales (p.ex. OCDE) dont l'État belge est membre, peuvent également recevoir des données, dans la mesure où celles-ci sont traitées exclusivement à des fins statistiques ou de recherche scientifique.

Les personnes physiques et morales qui poursuivent un objectif scientifique sont également prises en compte, à condition d'être reconnues par Statbel comme une entité de recherche. Les critères utilisés par Statbel à cet égard sont établis à l'article 4 du règlement (UE) n° 557/2013 de la Commission du 17 juin 2013 mettant en œuvre le règlement (CE) n° 223/2009 du Parlement européen et du Conseil du 11 mars 2009 en ce qui concerne l'accès aux données confidentielles à des fins scientifiques et abrogeant le règlement (CE) n° 831/2002. Il s'agit concrètement d'universités nationales et étrangères et de services d'études. Pour les personnes physiques, il est pratiquement impossible de satisfaire aux exigences en matière de sécurité imposées par Statbel pour l'obtention de microdonnées. Par conséquent, les étudiants (doctorants) ne sont éligibles que si la demande émane de l'université elle-même et que celle-ci se porte garante de la protection des données.

QUELLES SONT LES DONNÉES CONCERNÉES ?

Tant les données (administratives) existantes que Statbel a demandées que les données collectées par Statbel auprès des citoyens ou entreprises au moyen d'enquêtes sont prises en compte.

Statbel ne met à la disposition des chercheurs que des données *pseudonymisées*. Il s'agit de données au niveau individuel (personne, ménage ou entreprise) qui ne permettent pas au chercheur de déduire l'identité de la personne, du ménage ou de l'entreprise sans utiliser des données supplémentaires auxquelles il n'a pas accès. Cela signifie tout d'abord que l'ensemble de données ne peut contenir aucune information permettant une *identification directe*, par exemple le numéro de Registre national, le numéro d'entreprise, le nom, l'adresse, le numéro de téléphone, etc. Les données doivent également être protégées de *l'identification indirecte*. Cela signifie qu'une combinaison des données ne peut pas donner lieu à l'identification d'une personne, d'un ménage ou d'une entreprise.

CHERCHEURS PRIVILÉGIÉS

Les services statistiques de Flandre (autorité statistique flamande), de Wallonie (Institut wallon de l'Évaluation, de la Prospective et de la Statistique) et de Bruxelles (Institut Bruxellois de Statistique et d'Analyse), le département statistique de la Banque nationale de Belgique, le Bureau fédéral du plan et l'Observatoire des prix du SPF Économie peuvent légalement recevoir des microdonnées *non pseudonymisées* dans le cadre de leur mission au sein de l'Institut des Comptes Nationaux (ICN). Ils constituent une exception, Il existe toutefois une condition : les informations doivent être nécessaires pour remplir leur mandat légal, à savoir la production de statistiques. Il convient de noter que ces institutions et leur personnel sont soumis, dans le cadre de leur mission statistique, au secret statistique, tout comme Statbel.

COMMENT DEMANDER DES MICRODONNÉES AUPRÈS DE STATBEL ?

Seules les données dont Statbel dispose et qui présentent une valeur ajoutée pour la recherche peuvent être demandées. Afin de vérifier si cette condition est remplie, le chercheur doit, avant d'introduire une demande, se **concerter avec un statisticien de Statbel** familiarisé avec le thème de la recherche et les données concernées. Pendant cet entretien, il sera déterminé sur la base des besoins du chercheur quelles microdonnées entrent en ligne de compte pour une requête. Parmi les aspects importants figurent la qualité des données, l'univers, les périodes de référence et le moment où les données sont disponibles. Au cours de cette phase, on vérifiera également si la demande n'implique pas un risque (accru) en matière de sécurité des données, car si tel est le cas, la consultation préalable du *Data protection officer* (DPO) de Statbel est requise.

Si Statbel peut apporter une valeur ajoutée à la recherche en rendant les microdonnées disponibles et qu'il n'y a pas de preuve à première vue d'un risque (accru) en matière de sécurité des données ou d'une violation du secret statistique, le chercheur peut **soumettre une demande formelle** en utilisant un formulaire standardisé. Le formulaire se trouve sur le [site web](#) de Statbel. Il est composé de deux parties. La première partie est le *formulaire de demande de microdonnées* et le deuxième volet comporte une *déclaration relative aux mesures techniques et organisationnelles* visant à prévenir les violations du secret statistique ou les fuites de données. Un formulaire de demande spécifique est disponible pour les autorités statistiques régionales et pour les demandes de données pour la production de statistiques dans le cadre de l'Institut des Comptes Nationaux (ICN).



DÉLIBÉRATION ET AUTORISATION DE STATBEL

Chaque demande est évaluée par un comité multidisciplinaire au sein de Statbel. Cette évaluation permet de vérifier non seulement si les conditions légales sont satisfaites, mais également si les données demandées sont proportionnelles par rapport à l'objectif de la recherche et si les mesures techniques et organisationnelles sont proportionnelles aux risques en matière de sécurité des données. Le cas échéant, des mesures supplémentaires sont imposées pour lutter contre le risque d'identification indirecte, une réutilisation des données à des fins non statistiques ou non scientifiques ou des fuites de données.

Sur la base de l'avis du comité multidisciplinaire, le DPO de Statbel rédigera un avis formel. Enfin, c'est le directeur général qui, en tant que représentant légal de Statbel, donne l'autorisation de communiquer des (micro) données en signant l'avis du DPO. Un contrat de confidentialité est alors rédigé et soumis à la signature du responsable du traitement du chercheur. Les avis du DPO et les contrats de confidentialité signés sont publiés sur le [site web](#) de Statbel.

PROCÉDURES SIMPLIFIÉES

La procédure formelle développée par Statbel pour la communication des microdonnées se limite aux informations nécessaires pour une évaluation correcte de la demande. Si le chercheur souhaite adapter une autorisation existante, Statbel prévoit un *formulaire de demande simplifié*. Cela ne peut concerner qu'une prolongation du délai de conservation, l'ajout de variables supplémentaires, l'ajout de périodes de référence supplémentaires et l'extension de la finalité de la recherche. La simplification par rapport à la procédure normale se situe dans le formulaire de demande qui est moins détaillé.

Pour certaines recherches, des *données agrégées avec un faible risque d'identification indirecte* suffisent pour effectuer la recherche. En raison de ce faible risque, les données ne peuvent être considérées comme globales et anonymes et une délibération de Statbel doit avoir lieu et, le cas échéant, un contrat de confidentialité doit être conclu. Ici aussi, un formulaire de demande simplifié suffit.

CONCLUSION

Statbel dispose de données de grande qualité, du mandat légal pour mettre ces données à la disposition de chercheurs et de l'expertise nécessaire pour créer des produits de données personnalisés de haute qualité. Statbel est donc devenu une organisation qui va à la rencontre des chercheurs et les aide en mettant à leur disposition des données qui répondent à leurs besoins afin qu'ils puissent mener à bien leur mission dans des conditions optimales.

COLOPHON

Ont participé à ce numéro :

Jo Noppe - Hoofdstatisticus Statistiek Vlaanderen – Président du Conseil d'administration de l'IIS

Alain Henry - Expert – BFP

David Eyckmans – Chargé de mission – Wonen in Vlaanderen

Eva Debusschere – Chargé de mission – Wonen in Vlaanderen

Chloé Janssen - Chargée de recherche - IWEPS

Frédéric Vesentini - Directeur scientifique - IWEPS

Erik Meersseman – DPO – Statbel

L'IIS est le résultat d'un accord de coopération de 2014.

L'IIS doit permettre d'améliorer en continu la qualité des statistiques publiques, venant en soutien des politiques publiques.



www.iis-statistics.be