



Commission économique pour l'Europe**Conférence des statisticiens européens****Soixante-septième réunion plénière**

Paris, 26-28 juin 2019

Point 2 a) de l'ordre du jour provisoire

Nouvelles sources de données – accessibilité et utilisation**Première séance : accéder à de nouvelles sources de données****Accès aux mégadonnées à des fins statistiques****Note de l'Office fédéral de la statistique de l'Allemagne***Résumé*

Le présent document donne un aperçu des activités menées par l'Office fédéral de la statistique de l'Allemagne (Destatis) en vue de garantir l'accès à de nouvelles sources de données numériques. Il reste encore à résoudre de nombreux problèmes pour pouvoir utiliser les mégadonnées, à savoir tout d'abord des questions de qualité et de méthodes telles que la couverture, les biais de sélection ainsi que la disponibilité régulière et en temps voulu. Destatis a déjà entrepris plusieurs études de faisabilité dans ce domaine. Pour pouvoir utiliser plus durablement les nouvelles données numériques aux fins des statistiques officielles, il est indispensable d'assurer un accès permanent aux données détenues par des entreprises privées. Destatis cherche donc à établir une base juridique pour accéder aux données privées, le but étant de renforcer la coopération.

Le document est présenté, pour examen, au séminaire de la Conférence des statisticiens européens de 2019 intitulé « Nouvelles sources de données – accessibilité et utilisation ».



I. Introduction

1. L'Office fédéral de la statistique de l'Allemagne (Destatis) ainsi que le groupe de travail des nouvelles données numériques dans le contexte des statistiques officielles, créé par le conseil consultatif statistique de Destatis, sont convaincus de l'immense potentiel que les mégadonnées présentent en matière de statistiques officielles, lorsqu'elles sont associées à des données tirées d'enquêtes et à des données administratives. On s'attend à ce que les statistiques officielles soient plus d'actualité et plus précises grâce à ces nouvelles sources de données, d'où un allègement de la charge que représente la communication de données, en particulier pour les entreprises privées, dont les données sont à ce jour collectées dans le cadre d'enquêtes. Ainsi, les statistiques officielles devraient répondre aux attentes de la société à l'ère numérique, à savoir que toutes les sources d'information disponibles devraient être utilisées et qu'il faudrait, dans la mesure du possible, éviter le recours à des enquêtes requérant de nombreuses ressources. Reste la question de savoir si cette démarche permettra aux responsables des statistiques officielles d'économiser de l'argent. Les éventuelles réductions du coût des enquêtes peuvent être surcompensées – au moins à court terme – par les investissements dans le matériel informatique et les logiciels, ainsi que dans la formation des statisticiens. En outre, même si l'accès aux données privées est garanti sur le plan juridique, leur traitement préalable à des fins statistiques exigera peut-être (encore) que les instituts nationaux de statistique rémunèrent ce service. Une analyse coûts-avantages spéciale reste à faire avant que les mégadonnées puissent être utilisées pour la production de statistiques et pas seulement pour des études de faisabilité.

2. Pour déterminer si les mégadonnées peuvent être adéquatement utilisées dans le domaine des statistiques officielles, Destatis participe à plusieurs études de faisabilité, par exemple l'étude des données fournies par la téléphonie mobile et par les satellites à diverses fins dans le secteur des statistiques officielles. S'agissant de l'accessibilité et du traitement des mégadonnées, Destatis collabore avec différentes organisations, telles que la compagnie de téléphonie mobile T-Systems International GmbH et la société Motionlogic GmbH (toutes deux filiales de la société Deutsche Telekom AG), le Centre aérospatial allemand et l'Organisme fédéral allemand de cartographie et de géodésie. Destatis est également membre de l'Équipe spéciale du Système statistique européen (SSE), du Groupe directeur du SSE ainsi que du Groupe de travail mondial des Nations Unies sur l'utilisation des mégadonnées en statistique officielle et coopère avec beaucoup d'autres universités et instituts nationaux de statistique dans le cadre de différents projets européens.

3. Pour autant, de nombreux problèmes doivent encore être résolus pour utiliser les mégadonnées, à savoir tout d'abord des questions de qualité et de méthodes telles que la couverture, les biais de sélection ainsi que la disponibilité régulière et en temps voulu. Il faut ensuite intensifier la communication interne et externe concernant l'utilisation des mégadonnées. L'infrastructure informatique et les interfaces indispensables à l'intégration des demi-produits appartenant à des propriétaires de données privées doivent être passés en revue ainsi que les compétences requises. En outre, différentes questions juridiques doivent être résolues avec diverses parties prenantes telles que les ministères, les spécialistes de la protection des données aux niveaux national et régional, l'Agence fédérale des réseaux et, enfin et surtout, les fournisseurs de données.

4. Le présent document traite principalement de l'accès aux mégadonnées, qui est essentiel pour que le secteur des statistiques officielles puisse produire des statistiques plus précises et de meilleure qualité tout en allégeant la charge de travail des répondants. Il s'agit là de l'un des principaux obstacles rencontrés comme en témoigne l'expérience tirée des projets visant les mégadonnées.

II. Approche des organismes de statistique

5. Du point de vue de Destatis, les données des entreprises privées, qui constituent en général un sous-produit de la principale activité de ces entités, devraient être accessibles à des fins publiques, c'est-à-dire aux fins des statistiques officielles. Par conséquent, il est

nécessaire d'avoir un accès permanent aux données des entreprises privées, accès qui doit être indépendant de la volonté de ces dernières ou de la situation du marché. À cet égard, Destatis cherche à établir une base juridique pour accéder aux données privées, ce qui renforcera la coopération.

6. En ce qui concerne la réglementation juridique, Destatis et l'Office national de l'information et des technologies de la Rhénanie du Nord-Westphalie ont fait établir un rapport sur le droit comparé concernant l'accès aux mégadonnées à des fins publiques en Europe. Ainsi, on peut assurer que les organismes de statistique officiels allemands produisent des statistiques dans les mêmes conditions que les partenaires du Système statistique européen. De plus, le rapport est censé passer en revue la manière dont il faut modifier le cadre juridique des organismes de statistique fédéraux allemands pour qu'ils puissent utiliser des mégadonnées.

7. Au niveau européen¹, Eurostat a attribué en 2016 un contrat dont l'objet était d'analyser la situation juridique en matière d'accès, d'utilisation et de diffusion des statistiques extraites des mégadonnées pour tous les États membres de l'Union européenne. À ce jour, deux références sont applicables dans l'Union européenne s'agissant de l'accès aux données privées. La première est la « loi numérique » française², qui fait obligation aux sociétés – dans certaines conditions – d'autoriser l'accès aux données définies comme étant des « données d'intérêt général », par exemple au profit de l'Institut national de la statistique et des études économiques (INSEE). La deuxième référence est la loi de 2017 du Royaume-Uni relative à l'économie numérique (Digital Economy Act)³ (voir partie 5, chapitre 7, article 80, nouvel article 45D), qui permettra à l'Office for National Statistics (ONS) d'accéder aux données détenues par les grandes et moyennes entreprises et les associations caritatives pour qu'il puisse s'acquitter d'une ou de plusieurs de ses fonctions statutaires, à savoir la production de statistiques officielles et la recherche statistique.

8. En outre, le Groupe de travail mondial sur l'utilisation des mégadonnées en statistique officielle a publié en 2016 le document « Recommendations for Access to Data from Private Organisations for Official Statistics » (recommandations sur l'accès aux données d'organisations privées aux fins des statistiques officielles)⁴, qui peut servir de lignes directrices pour les négociations menées avec les entreprises privées possédant des sources de mégadonnées très utiles pour les statistiques officielles.

9. En outre, en avril 2018, la Commission européenne a publié une communication intitulée « Vers un espace européen commun des données »⁵ ainsi que des orientations concernant le partage des données du secteur privé dans l'économie européenne des données⁶, documents qui font partie intégrante de la stratégie de la Commission pour un marché unique numérique et concernent la mise en commun de données entre les entreprises et le secteur public. La Commission européenne soutient l'idée qu'il faut améliorer l'accès des organismes publics comme les instituts nationaux de statistique aux données privées, sans cependant oublier l'intérêt des entreprises privées, notamment en

¹ Voir la version non publiée du document d'Eurostat intitulé « Legal initiatives for data access », établi dans le cadre de l'équipe spéciale des mégadonnées (avril 2017).

² Voir sur le site Legifrance : <https://www.republique-numerique.fr/pages/in-english> et <https://www.republique-numerique.fr/pages/digital-republic-bill-rationale> (consultation des deux pages le 18 mai 2018).

³ Voir sur le site « National Archives » du Royaume-Uni : <http://www.legislation.gov.uk/ukpga/2017/30/contents/enacted> et http://www.legislation.gov.uk/ukpga/2017/30/pdfs/ukpgaen_20170030_en.pdf (consultation des deux pages le 18 mai 2018).

⁴ Voir sur le site de la Division de statistique de l'ONU : <https://unstats.un.org/unsd/bigdata/conferences/2016/gwg/Item%20%2028i%29%20a%20%20Recommendations%20for%20access%20to%20data%20from%20private%20organizations%20for%20official%20statistics%20Draft%2014%20July%202016.pdf> (consultation le 18 mai 2018).

⁵ Voir sur le site de la Commission européenne : <https://ec.europa.eu/digital-single-market/en/news/communication-towards-common-european-data-space> (consultation le 18 mai 2018).

⁶ Voir sur le site de la Commission européenne : <https://ec.europa.eu/digital-single-market/en/news/staff-working-document-guidance-sharing-private-sector-data-european-data-economy> (consultation le 18 mai 2018).

accordant une indemnisation pour les investissements effectués par les entreprises pour adapter les données.

10. En octobre 2018, le Comité du Système statistique européen a adopté le Mémoire de Bucarest et a reconnu qu'il fallait étoffer le cadre juridique aux niveaux européen et national pour réduire les obstacles à l'accès, à l'utilisation et à l'intégration des données hétérogènes afin de produire durablement des statistiques intelligentes fiables.

11. D'une manière générale, le règlement général sur la protection des données (RGPD) européen est censé renforcer la confiance dans les services numériques et leur sécurité en définissant des normes communes pour la protection des données et en garantissant le droit fondamental à l'autodétermination en matière d'information. Depuis le 25 mai 2018, le RGPD est en vigueur dans tous les États membres, l'objectif étant d'harmoniser les lois relatives à la confidentialité des données dans l'ensemble de l'Europe.

12. Un traitement partiel des données par les entreprises privées pourrait aussi simplifier la mise en conformité avec la législation ayant trait à la protection des données. Des produits statistiques intermédiaires agrégés qui ne comportent pas d'informations personnelles suffisent largement aux statistiques officielles et ainsi les responsables de la protection des données acceptent plus facilement de recourir aux mégadonnées en vue de la production de statistiques officielles. Pour utiliser des produits statistiques ayant fait l'objet d'un traitement préalable par les entreprises privées, il est indispensable d'élaborer des normes claires et contraignantes. Au niveau européen, ce processus a déjà été engagé sous le nom de « statistiques intelligentes fiables ». À cet égard, les produits statistiques intermédiaires provenant de tierces parties doivent satisfaire à des normes qualitatives, éthiques et méthodologiques pour pouvoir être associés à des données issues d'enquêtes et à des données administratives avant d'être utilisés aux fins de la production de statistiques officielles.

III. Solution envisageable en Allemagne

13. En l'état actuel des choses, les mégadonnées elles-mêmes devraient continuer à être gratuites pour le secteur des statistiques officielles, quel que soit le type de données. Autrement dit, elles devraient être fournies par les répondants sans remboursement des coûts.

14. La législation dans ce domaine pourrait être similaire à celle qui régit actuellement l'utilisation des données administratives. Pour évaluer si les données sont adéquates, les organes administratifs doivent communiquer aux organismes de statistique qui le demandent des informations sur l'origine, la structure et le contenu des données ainsi que sur d'autres métadonnées. Si un ministère fédéral décide alors que de nouvelles études devraient être menées, les organes administratifs doivent également transmettre à Destatis des données individuelles anonymisées (sans nom ni adresse). Si les données sont déclarées adéquates, un règlement spécifique doit spécifier les statistiques à établir.

15. De plus, il faut encore décider si les organismes de statistique allemands recourent aux services des entreprises privées pour le traitement des données et en remboursent le coût à ces entreprises. Le fait que le traitement des données effectué par les entreprises privées sera en partie d'un meilleur rapport coût-efficacité montre que la meilleure solution consistera probablement en une participation aux coûts de la part des organismes de statistique ou en une coopération entre les entreprises privées et ces mêmes organismes. En cas de coopération entre les entreprises privées et les organismes de statistique, le principe de la neutralité des statistiques officielles doit être assuré et garanti par des procédures transparentes et juridiquement contraignantes.

IV. Autres questions (d'ordre juridique)

16. Outre l'accès aux mégadonnées, d'autres questions d'ordre juridique doivent encore être examinées, par exemple celle de savoir si une législation spéciale sur la protection des données est nécessaire pour les données que les entreprises privées fourniront au secteur

des statistiques officielles. En ce sens, il faut encore étudier la question de savoir si les résultats ou les produits qui découlent de la coopération avec les entreprises privées pourraient donner lieu à une concurrence déloyale et la question de savoir comment protéger les informations commerciales sensibles.

17. D'autres questions ont trait à la propriété des mégadonnées, questions auxquelles il est particulièrement difficile de répondre lorsqu'il s'agit de données générées par des machines, par exemple dans le secteur agricole. Il faudrait aussi aborder les questions éthiques, notamment celle de savoir si l'utilisation des mégadonnées influe sur les normes professionnelles du secteur des statistiques officielles.

18. Il est probable que les mégadonnées ne révéleront tout leur potentiel que lorsqu'elles seront associées à des données tirées d'enquêtes et à des données administratives. Par conséquent, les législateurs devraient aussi jeter les bases des prescriptions juridiques nécessaires en ce qui concerne la combinaison de différentes sources de données, notamment pour ce qui est des données personnelles.

19. Les statistiques officielles sont actuellement accessibles à la communauté scientifique par le truchement des centres de données de recherche de l'Office fédéral de la statistique et des organismes de statistique des Länder, compte tenu des dispositions du paragraphe 2 de l'article 16 de la loi sur les statistiques fédérales. À l'avenir, la disponibilité des données à des fins scientifiques devra être possible pour les mégadonnées. Les résultats des statistiques officielles doivent être librement disponibles en vue d'une évaluation scientifique.

V. Conclusion

20. Il est primordial d'assurer un accès permanent aux mégadonnées pour produire régulièrement des statistiques officielles. Cet accès ne doit pas dépendre de la situation du marché ni de la volonté des entreprises privées de partager leurs données.

21. Destatis est convaincu que, pour maintenir la viabilité des statistiques officielles, le cadre juridique de l'Office fédéral allemand de la statistique doit être révisé pour ce qui est de l'accès aux données privées, en particulier à l'ère de la numérisation générale.
