

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

Genève, 22-24 juin 2020

Points 5 et 10 de l'ordre du jour provisoire

Nouveaux rôles des organismes nationaux de statistique et d'information géospatiale dans les récents écosystèmes nationaux de données**Gestion des données – prochaines étapes pour les organismes nationaux de statistique****Mise en œuvre du nouveau rôle des organismes nationaux de statistique face à des possibilités accrues¹****Note établie par l'Estonie avec des contributions de l'Albanie, du Canada, de l'Irlande, de l'Italie, de la Nouvelle-Zélande, des Pays-Bas, de la Pologne et du secrétariat de la CEE***Résumé*

Le Bureau de la Conférence des statisticiens européens a demandé à un groupe de pays, dirigé par l'Estonie, d'établir un document sur le rôle des organismes nationaux de statistique dans le nouvel écosystème de données. Il a été présenté à la réunion plénière de 2019 de la Conférence des statisticiens européens en tant que contribution au séminaire organisé sur le thème « Le nouveau rôle des organismes nationaux de statistique en tant que bureaux de statistique et de données ».

À l'issue de ce séminaire, la Conférence a demandé à une équipe spéciale informelle (composée de l'Albanie, du Canada, de l'Estonie (chef de file), de l'Irlande, de l'Italie, de la Nouvelle-Zélande, des Pays-Bas, de la Pologne et du secrétariat de la CEE) d'actualiser le document faisant la synthèse des expériences et des différentes démarches ainsi que des propositions faites lors du séminaire, et de mettre au point un plan d'action portant sur les problématiques à étudier plus avant. Ce document actualisé constitue une contribution essentielle au séminaire de suivi de la Conférence de 2020 sur les « Nouveaux rôles des organismes nationaux de statistique et d'information géospatiale dans les récents écosystèmes nationaux de données ».

La Conférence est invitée à formuler des observations sur les questions et considérations présentées dans le document et à convenir de la marche à suivre.

¹ Il a été convenu que le présent document serait publié après la date normale de publication en raison de circonstances indépendantes de la volonté du soumetteur.



I. Introduction

A. Objectif et propos du présent document

1. Le présent document a pour objectif d'examiner les enjeux auxquels sont confrontés les systèmes statistiques nationaux et les organismes nationaux de statistique en ce qui concerne la mise en place d'écosystèmes nationaux de données. De tels écosystèmes comprennent généralement des données très hétérogènes – statistiques, administratives, géospatiales – et d'autres nouvelles sources, dont les « mégadonnées ». Les écosystèmes offrent de nouvelles possibilités d'établir un lien entre ces différentes sources de données, de manière à créer un patrimoine informationnel utile pour les décideurs et le grand public.

2. Ce document aborde trois aspects : les questions d'ordre stratégique que soulève le nouveau rôle des organismes nationaux de statistique dans les écosystèmes de données, les domaines dans lesquels ces organismes devraient assumer un leadership et comment ils devraient relever ces défis de façon à répondre aux attentes des différents groupes d'utilisateurs, depuis les dirigeants jusqu'aux simples citoyens. Les auteurs ont conscience que les rôles et les capacités des systèmes statistiques nationaux diffèrent selon les pays, de même que les attentes des dirigeants et de la société quant au nouveau rôle que ces organismes devraient jouer. Il existe des différences entre les pays en ce qui concerne le développement économique global et l'utilisation des technologies de l'information. Il existe également des différences culturelles entre les pays, notamment en ce qui concerne la culture de l'information². Tous ces éléments ont un impact substantiel sur la perception de leur rôle par les organismes nationaux de statistique.

3. La version originale de ce document a été soumise lors du séminaire de la Conférence des statisticiens européens de 2019, en réponse à une demande du Bureau de la Conférence réuni les 14 et 15 octobre 2018.

4. Le présent document tient compte des résultats du séminaire de la Conférence de 2019 en présentant une synthèse des expériences et des différentes démarches ainsi que des propositions faites lors de celui-ci. Il inclut un éventuel plan d'action et les prochaines étapes à prévoir afin d'aider les organismes nationaux de statistique à assumer un nouveau rôle et de réaliser des progrès concrets dans ce domaine qui fait l'objet d'un vaste débat.

B. Débat en cours sur le nouveau rôle des organismes nationaux de statistique

5. Le Bureau de la Conférence a commencé à se pencher en février 2018 sur le rôle des systèmes et organismes nationaux de statistique dans les nouveaux écosystèmes de données³. Le document prend également en considération les résultats des réunions pertinentes de la Conférence des statisticiens européens, de l'OCDE et de l'UE.

6. En novembre 2014, le Groupe consultatif d'experts indépendants de l'ONU a présenté le document « A World That Counts. Mobilising the Data Revolution for

² Par culture de l'information, nous entendons à la fois l'habitus, la manière dont l'organisation conserve ses informations, et les informations générales ou l'aptitude à comprendre et utiliser les données dans le pays.

³ Il existe de nombreuses définitions d'un écosystème de données. Dans le document du PNUD intitulé « Data ecosystems for sustainable development » (2017) un écosystème de données est expliqué comme suit : « L'approche écosystémique implique une évaluation systématique de la capacité statistique officielle : les cadres juridiques et politiques en place pour les données ouvertes, les obstacles à la participation des multiples parties prenantes tout au long de la chaîne de valeur des données, pour la mise en œuvre et le suivi des objectifs de développement durable, et des points d'entrée pour des actions futures. Il faut en outre évaluer de manière critique les besoins en infrastructures pour améliorer la collecte, la diffusion et l'utilisation des données, y compris les possibilités de tirer parti des nouvelles technologies pour ce qui est de la participation ». p. 20.

Sustainable Development⁴ », dans lequel il décrit la situation actuelle comme une révolution des données, un processus continu qu'il faut mettre à profit aux fins du développement durable. Le Groupe consultatif d'experts a recommandé à l'ONU d'élaborer une stratégie détaillée et une feuille de route pour parvenir à un nouveau « Consensus mondial sur les données ». Cette feuille de route devrait inclure la définition et l'adoption de normes sur les plans juridique, technique, géospatial et statistique. Le document met en relief les thématiques suivantes : « Ouverture et échange des données et des métadonnées, y compris l'interopérabilité des données et des systèmes d'information, l'information démographique et géospatiale, notamment la gestion et l'échange des informations géographiques et sémantiques, l'échange mondial d'informations sur les flux financiers illicites et la gestion et l'octroi de licences en matière de données publiques et de droits numériques⁵. »

7. L'OCDE a élaboré un document intitulé « Quelles stratégies pour les organismes nationaux de statistique à l'ère du numérique ? Vers des stratégies de « données intelligentes⁶ ». qui préconise la mise en place d'un cadre pour les données intelligentes – une démarche en rupture avec le battage autour des mégadonnées. Il aborde également la question du nouvel écosystème de données et envisage le rôle des organismes nationaux de statistique comme suit : « Dans le cadre des systèmes statistiques nationaux, l'organisme national de statistique est le mieux placé pour orchestrer la mise en place d'une stratégie nationale en matière de données intelligentes. »

8. Dans le document intitulé « Big data strategies for official statistics », Peter Struijs et Sofie de Boer, de Statistics Netherlands, écrivent que « mégadonnées » et « écosystème de données » sont les mots à la mode⁷.

9. À sa réunion d'octobre 2018, le Bureau de la Conférence a examiné un document du Canada et du Royaume-Uni sur le cadre de partenariat stratégique pour la statistique officielle et un document du secrétariat de la Commission économique pour l'Europe (CEE) sur le rôle des systèmes statistiques nationaux dans le nouvel écosystème des données⁸.

10. Une version de ce document a été présentée lors du séminaire de la Conférence des statisticiens européens sur « Le nouveau rôle des organismes nationaux de statistiques en tant que bureaux de statistique et de données » en juin 2019.

11. À la suite de ce séminaire, la Conférence a demandé à une équipe spéciale informelle (composée de l'Albanie, du Canada, de l'Estonie (chef de file), de l'Irlande, de l'Italie, de la Nouvelle-Zélande, des Pays-Bas, de la Pologne et du secrétariat de la CEE) de réfléchir aux résultats de ce séminaire et d'établir un document faisant la synthèse des expériences et des différentes démarches présentées à cette occasion, ainsi que des propositions faites dans le cadre de la Conférence et de mettre au point un plan d'action portant sur les problématiques à étudier plus avant.

⁴ [A World That Counts. Mobilising the Data Revolution for Sustainable Development](#). Groupe consultatif d'experts indépendants, Organisation des Nations Unies. Novembre 2014.

⁵ [A World That Counts](#), p. 21.

⁶ [Quelles stratégies pour les organismes nationaux de statistique à l'ère du numérique ? Vers des stratégies de « données intelligentes »](#). Quinzième réunion du Comité des statistiques et de la politique statistique de l'OCDE, juin 2018.

⁷ [Peter Struijs, Sofie de Broe. Big Data Strategies for Official Statistics](#). Année 2018

⁸ [Role of National Statistical Systems in the New Data Ecosystem. Réunion du Bureau de la Conférence](#), octobre 2018.
[Modern Partnership Framework for Official Statistics. Réunion du Bureau de la Conférence](#), octobre 2018.

II. Mutations de l'environnement et nouvel écosystème des données

A. La révolution des données et la « datafication » de la société

12. L'expression « révolution des données » a été utilisée dans le rapport « A World That Counts » élaboré en 2014 par le Groupe consultatif d'experts indépendants auprès du Secrétaire général de l'ONU et où il est dit à ce sujet : « Une véritable révolution des données ferait appel aux sources de données existantes et aux nouvelles sources pour intégrer pleinement la statistique dans le processus décisionnel, promouvoir le libre accès et la libre exploitation des données et renforcer les systèmes statistiques » et « On s'accorde généralement à reconnaître que la "révolution des données" désigne les transformations nécessaires pour répondre aux exigences d'un programme de développement complexe, les améliorations dans la façon de produire et d'exploiter les données, la nécessité de combler les lacunes en matière de données afin de prévenir les discriminations, le renforcement des capacités et des connaissances pour l'analytique des données, massives ou non, la modernisation des systèmes de collecte des données, l'ouverture des données pour promouvoir la transparence et la responsabilité ainsi que l'établissement de nouvelles cibles et de nouveaux indicateurs⁹ ».

13. La numérisation des fonctions de la société (ou de l'environnement social) modifie la façon dont la société se comporte. De plus en plus, nous agissons conformément aux supports numériques et nous leur faisons davantage confiance qu'aux documents papier ; nos traces sont, de façon croissante, des traces numériques et constituent donc une source potentielle pour les statistiques.

14. La « datafication » de la société renvoie au fait que les individus, les organisations et les objets inanimés laissent des traces numériques ou échappement de données dont le volume croît de façon exponentielle. Les capteurs sont partout, l'Internet des objets est en plein essor. Aujourd'hui, il semble bien qu'aucun mouvement, aucune action ou opération, aucun changement ne puisse avoir lieu sans générer, de quelque façon et quelque part, des données.¹⁰

B. Infrastructure numérique et écosystème des données

15. Le principal facteur de changement dans l'écosystème des données a été la révolution survenue ces dernières années, marquée par la numérisation croissante de l'information et l'émergence de la notion de « mégadonnées ». Cette révolution soulève inévitablement les questions liées à la réglementation de l'accès aux données, à la qualité des données et aux droits en matière de protection de la vie privée. Elle a notamment suscité la volonté de gérer les données en tant qu'atout stratégique pour une plus grande rentabilité. Des entités autres que les services traditionnels ont donc entrepris de produire des statistiques en mettant à profit les nouvelles sources de données, les nouvelles techniques et les nouveaux outils disponibles. Cette révolution a également contribué à modifier les attentes des parties prenantes – qui réclament désormais un accès plus ouvert et plus rapide aux données, notamment via les portails Web, les interfaces de programmation d'applications (API) et les plateformes de partage. L'exigence d'accès aux données est portée par l'accélération des mutations technologiques, notamment dans les domaines de l'intelligence artificielle (IA) et de l'apprentissage automatique, et la sophistication croissante des compétences parmi les spécialistes de l'analytique et de la science des données¹¹.

16. Pourquoi parler de « données intelligentes » plutôt que de « données massives ou mégadonnées » ? Les projets analysés le montrent : la valeur ne tient pas nécessairement à la « masse » des données. Pour caractériser les mégadonnées, on utilise généralement les

⁹ Voir <http://www.undatarevolution.org/data-revolution/>.

¹⁰ Peter Struijs, Sofie de Broe. Big Data Strategies for Official Statistics, p. 2.

¹¹ Modern Partnership Framework for Official Statistics, par. 5.

trois V – volume, variété et vélocité – auxquels s’ajoutent parfois véracité et valeur. La valeur résulte de l’adoption d’approches créatives des données, parfois en combinant des sources existantes de données (massives ou non) de façon inédite, en exploitant des sources non conventionnelles et en créant des méthodes et des algorithmes novateurs. Elle émane également de l’aptitude à allier les approches statistiques traditionnelles et les techniques les plus récentes en matière de science des données – au lieu de les opposer. Sous cet angle, les « données intelligentes » peuvent être vues comme un rapprochement sémantique à l’appui d’une proposition de valeur qui associe et transcende les deux approches, rassemblant le statisticien et le spécialiste de la science des données¹².

17. La révolution en cours bouleverse les écosystèmes de données depuis plusieurs années, avec la numérisation croissante des informations et l’émergence des « mégadonnées ». Les données sont ainsi devenues un atout stratégique et de nombreuses organisations extérieures au milieu de la statistique produisent désormais des données en exploitant de nouvelles sources et des techniques et outils inédits. Cette évolution a également modifié les attentes des parties prenantes, qui veulent pouvoir accéder aux données plus librement, plus facilement et plus rapidement¹³. Ces sources sont nouvelles dans deux sens : en tant que sources primaires révolutionnant l’utilisation des données, en particulier dans le secteur public établissant tout un écosystème de données administratives, et en tant que sources de données secondaires constituant une opportunité stratégique pour les organismes nationaux de statistique à l’avenir.

III. Questions stratégiques

A. Quel rôle les organismes nationaux de statistique et les statistiques officielles sont-ils appelés à jouer dans la société ?

18. Points soulevés dans les précédentes communications :

a) Pour répondre aux nouvelles demandes, le modèle actuel de production des statistiques officielles doit remédier à des lacunes dans plusieurs domaines :

i) Lacunes en termes de sources de données – capacité de combiner les données provenant de sources traditionnelles avec celles issues des nouvelles sources, de mettre au point de nouvelles méthodes statistiques ou d’adapter les méthodes existantes, afin de traiter efficacement un nombre toujours croissant de sources de données alternatives dans des domaines tels que le couplage des données, l’intégration des données et la validation des données ;

ii) L’accès aux nouvelles données devient une question stratégique qui nécessite de définir une stratégie globale ;

iii) Lacunes en termes de plateformes – de nouvelles plateformes sont nécessaires pour gérer les volumes toujours croissants de données qui sont souvent hébergées à distance (nuage). Il faut pour cela adapter les modèles financiers, les politiques de protection de la vie privée, utiliser les nouvelles techniques d’IA, etc. Nous devons également accélérer le développement et la mise en œuvre de techniques/technologies pour faciliter l’accès aux nouvelles sources de données tout en assurant une confidentialité et une protection de la vie privée encore plus grandes (informatique multipartite sécurisée, techniques de partage des secrets, informatique en périphérie de réseau, virtualisation et abstraction des données, etc.) ;

¹² Quelles stratégies pour les organismes nationaux de statistique à l’ère du numérique ? Vers des stratégies de « données intelligentes », par. 12.

¹³ Rôle des organismes nationaux de statistique dans le nouvel écosystème des données, par. 8.

iv) Lacunes en termes de compétences – une nouvelle expertise est nécessaire (en science des données, modélisation, exploration, analytique, IA, etc.)¹⁴ ;

v) Cadres juridiques – pour garantir l’acceptabilité sociale et un accès suffisant, il faut développer les cadres juridiques ou approfondir et moderniser ceux qui existent déjà afin d’y inclure des considérations relatives à la vie privée et aux avantages pour la société ;

b) Les travaux menés au niveau international sont extrêmement importants. La communauté internationale des statisticiens pourrait élargir son rôle et sa portée, en nouant de nouveaux partenariats pour devenir la communauté internationale des spécialistes des statistiques et des données. Les partenariats avec les organisations mondiales peuvent tirer profit de la coordination à l’échelle mondiale¹⁵.

19. Le rôle des organismes nationaux de statistique peut être abordé dans le cadre de différentes thématiques : législation, stratégie numérique, évolution des fonctions, dimension internationale, besoins des utilisateurs – depuis les dirigeants jusqu’aux simples citoyens. Les organismes nationaux de statistique devraient s’employer activement à combler ces lacunes. Le manque de clarté conceptuelle ou la soi-disant question des définitions est une des lacunes pour laquelle ils jouent un rôle plus actif en abordant différents sujets dans un langage professionnel clair.

B. Quelles sont les fonctions traditionnelles que les organismes nationaux de statistique doivent maintenir ? En quoi le nouveau rôle de ces organismes pourrait-il modifier leurs valeurs et leurs traditions ?

20. Points soulevés dans les précédentes communications :

a) Comme l’ont exposé Struijs et De Broe, la valeur ajoutée des statistiques officielles tient aux éléments suivants :

i) Normes de qualité rigoureuses des statistiques officielles ;

ii) Transparence totale des méthodes et des hypothèses ;

iii) Indépendance professionnelle ou soumission au jugement de professionnels indépendants ;

iv) Publication simultanée de l’information pour tous les utilisateurs ;

v) Cohérence et comparabilité¹⁶ ;

b) Jusqu’à présent, les organismes de statistiques se sont toujours appuyés sur une structure de production verticale, qu’ils contrôlent intégralement, exploitant les enquêtes (auprès des ménages, des entreprises) et les données transmises par d’autres institutions (par exemple, les organismes nationaux de statistique collectent des données auprès des ministères d’exécution, des agences et des autorités locales, et les banques centrales collectent des données auprès des établissements financiers). Depuis des années, les organismes nationaux de statistique s’efforcent (a) d’accroître l’efficacité des mécanismes de notification et (b) d’accéder aux sources administratives pour compléter ces données et éventuellement réduire la charge, en particulier pour les entreprises¹⁷.

21. Les statistiques officielles ont certaines traditions opérationnelles, notamment sur des sujets tels que la qualité statistique, la neutralité, la transparence des méthodes et des hypothèses, la cohérence et la comparabilité des statistiques, etc. Lors de l’élaboration de

¹⁴ Quelles stratégies pour les organismes nationaux de statistique à l’ère du numérique ? Vers des stratégies de « données intelligentes », par. 7. Rôle des organismes nationaux de statistique dans le nouvel écosystème des données, par. 16.

¹⁵ Rôle des organismes nationaux de statistique dans le nouvel écosystème des données, par. 6.

¹⁶ Peter Struijs, Sofie de Broe. Big Data Strategies for Official Statistics, p. 4.

¹⁷ Quelles stratégies pour les organismes nationaux de statistique à l’ère du numérique ? Vers des stratégies de « données intelligentes », par. 15.

leur nouveau rôle, quelle est la stratégie des organismes nationaux de statistique pour garantir que ces caractéristiques essentielles resteront importantes ?

22. Quel serait un modèle économique viable et durable pour les organismes nationaux de statistique à l'avenir ? En particulier, comment mobiliser les ressources financières nécessaires à la réalisation de leurs ambitions ? Comment devraient-ils se comporter et se positionner par rapport au marché ?

23. La redéfinition du rôle imparti aux organismes nationaux de statistique modifiera inévitablement leur façon de travailler. Peut-on dès maintenant énoncer de nouveaux principes ou faudrait-il se fonder sur les résultats obtenus par certains pays qui font l'objet d'études de cas ?

C. Comment les organismes nationaux de statistique réagiront-ils à la concurrence des autres organisations et quels partenariats stratégiques mettre en place ?

24. Points soulevés dans les précédentes communications :

a) Un cadre de partenariat stratégique moderne doit être centré avant tout sur le système statistique national et faire fond sur les avantages comparatifs des partenaires qui partagent un objectif commun et une même vision. Les accords de partenariat modernisés doivent être souples, adaptables et fondés sur un programme rentable, efficace et axé sur la qualité¹⁸ ;

b) Dans les dispositifs de partenariat, la force des organismes nationaux de statistique tient à leurs solides capacités en matière de gouvernance et d'intendance des données. Mais, étant donné les mutations accélérées à l'œuvre dans les domaines de la disponibilité des données et des technologies numériques, ces organismes doivent rechercher des partenariats stratégiques avec les fournisseurs de données et de technologies, les scientifiques et les chercheurs ainsi que les médias pour pouvoir continuer de répondre aux besoins des parties prenantes, qui ne cessent d'évoluer¹⁹ ;

c) Pour les organismes nationaux de statistique, les autres institutions qui s'occupent de statistiques et d'analyse des données peuvent être perçues comme des concurrents ou comme des partenaires. Le premier concurrent est le monde universitaire. Bien entendu, les opérations conjointes ou les programmes de recherche conjoints sont légion mais il n'en reste pas moins nécessaire d'examiner le rôle que joueront ces deux modalités à l'avenir. Un autre type d'organisation avec lequel les organismes nationaux de statistique sont en concurrence est celui des sociétés fournissant des services d'intermédiation et d'analyse de données.

25. La question des partenariats revêt une importance croissante. Parmi les partenaires figureront d'autres organismes de service public, des agences géospatiales et cartographiques, des institutions universitaires, des installations de recherche et des groupes de réflexion, des scientifiques et des ingénieurs spécialisés dans les données, les organisations de protection des données, les organisations du secteur privé qui peuvent être en mesure de fournir une formation, une expertise et des solutions logicielles, les communautés du logiciel libre, etc.

26. Lorsqu'ils tentent de résoudre des problèmes par le biais de partenariats, les organismes nationaux de statistique agissent naturellement comme des organes de coordination. Cela soulève la question de l'intendance des données, du partage des données et de leur utilisation protégée et éthique.

¹⁸ Modern Partnership Framework for Official Statistics, par. 7.

¹⁹ Modern Partnership Framework for Official Statistics, par. 14.

D. Comment apporter rapidement des éclairages sur les évolutions de la société tout en maintenant un degré de fiabilité élevé des données ?

27. Points soulevés dans les précédentes communications :

a) Les organismes nationaux de statistique fonctionneront-ils demain comme des centres d'échange d'informations proposant des jeux de données exacts, traités et normalisés, mis à disposition dans des catalogues et inventaires de données ouvertes ? Cette fonction sera-t-elle centrée sur les services plutôt que sur les données, notamment sur les services suivants : fourniture de données, y compris données de base et de référence, analyse et simulation²⁰ ?

b) Tandis que le numérique irrigue nos sociétés et que les individus sont en mesure d'obtenir en temps réel des informations (plus ou moins fiables) sur chaque aspect de la vie ou presque, les attentes en matière de statistiques ont-elles aussi évolué. Cette nouvelle demande présente au moins quatre caractéristiques :

i) Des délais de mise à disposition réduits au minimum – les délais d'obtention des statistiques officielles que les utilisateurs sont disposés à accepter ont diminué ;

ii) Granularité : les données doivent être mises à disposition rapidement mais elles doivent aussi être granulaires, ce qui implique, par exemple, qu'elles soient pertinentes au niveau local (« comment va ma communauté locale ? »), qu'elles opèrent une distinction entre les différents groupes socioéconomiques et qu'elles couvrent un large éventail de problématiques sociales, économiques et environnementales ;

iii) La quête d'une « garantie de qualité » – face à une offre de données toujours plus pléthorique (et écrasante), les utilisateurs attendent des organismes nationaux de statistique qu'ils donnent du sens au battage associé aux « mégadonnées », qu'ils tranchent entre des informations parfois contradictoires et valorisent et contextualisent les données orphelines diffusées sur les réseaux sociaux, dans les médias ou d'autres sources nouvelles ;

iv) Les attentes sont également plus élevées en ce qui concerne la gamme des nouveaux services et produits de données, auxquels peuvent contribuer les experts (universitaires, ONG, entreprises), voire les citoyens dans leur ensemble²¹.

28. Pour que les organismes nationaux de statistique passent à une nouvelle étape, ils doivent :

a) Se centrer davantage sur les utilisateurs, atteindre l'ensemble de la société, fournir les données selon des méthodes innovantes et lutter contre les fausses nouvelles ;

b) S'investir davantage dans la gouvernance des données, conseiller les gouvernements sur les questions stratégiques liées aux statistiques et aux données, fournir une expertise pour la mise en place de portails gouvernementaux de données (et probablement pour d'autres que les pouvoirs publics) ?

c) Jouer un rôle accru pour ce qui est de s'assurer de la qualité des informations utilisées dans la prise de décisions.

29. S'agissant des services de données, il sera probablement de plus en plus important que les prestations soient fournies en temps voulu, presque en temps réel. Les services de données devraient également tenir compte de la nécessité d'une granularité accrue.

²⁰ Modern Partnership Framework for Official Statistics, par. 33.

²¹ Quelles stratégies pour les organismes nationaux de statistique à l'ère du numérique ? Vers des stratégies de « données intelligentes », par. 5.

IV. Nouveaux rôles éventuels des organismes nationaux de statistique : considérations générales

30. On trouvera dans la présente section des considérations d'ordre général dont les organismes nationaux de statistique tiendront compte lorsqu'ils réfléchiront au nouveau rôle qu'ils sont appelés à jouer dans les écosystèmes nationaux de données. Elles ont été passées en revue à la réunion plénière de la Conférence tenue en juin 2019. Ces généralités, qui pourraient servir de point de départ pour l'élaboration d'orientations plus détaillées, concernent les organismes nationaux de statistique, les gouvernements et la société dans son ensemble, en particulier ceux qui aspirent à conduire cette « révolution des données ». Une attention spéciale est accordée aux organismes nationaux de statistique et à la communauté internationale des statisticiens, ce groupe cible ayant à la fois la dimension institutionnelle et les compétences nécessaires pour conseiller les autres parties sur les principes régissant la gouvernance des données. Le document est modifié en fonction des questions soulevées par les approches émergentes en matière de systèmes de données et d'intendance des données.

A. Le rôle des organismes nationaux de statistique dans la gouvernance des données publiques

31. Le cadre juridique de la statistique officielle devrait être étoffé pour inclure les nouveaux rôles dévolus aux organismes nationaux de statistique. Le champ des organismes nationaux de statistique s'élargit à mesure que de nouvelles sources de données deviennent disponibles et que les attentes de la société évoluent en ce qui concerne l'utilité éventuelle des données et des statistiques. Les organismes nationaux de statistique ont beaucoup à offrir aux partenaires éventuels : principes rigoureux, utilisation éthique, transparence et qualité. Ils doivent, pour rester dignes de confiance, préserver ces atouts, parmi d'autres, qui découlent des Principes fondamentaux de la statistique officielle ;

Considération 1 : Les organismes nationaux de statistique pourraient entamer des discussions au sein des administrations nationales afin de définir un régime juridique permettant l'utilisation des données. Ils devraient mettre en avant les avantages qu'il y a à suivre les orientations fournies par la Conférence des statisticiens européens dans l'ouvrage « Guidance on Modernizing Statistical Legislation », en particulier concernant :

- a) Les données ouvertes, les données ouvertes corrélées et les statistiques officielles ;
- b) L'échange de données entre producteurs de statistiques officielles ;
- c) L'évolution du recensement de la population et du logement et les aspects juridiques ;
- d) La collaboration avec les banques centrales ;
- e) L'intégration des données géospatiales et des statistiques ;
- f) La collaboration avec différents fournisseurs de données ;
- g) Les bureaux de statistique et la gestion des données gouvernementales.

32. Les organismes nationaux de statistique devraient s'attacher à jouer un rôle de premier plan dans la gouvernance des données à l'échelle de l'ensemble de l'administration. La gouvernance des données comprend les éléments suivants :

- a) Architecture des données ;
- b) Développement de normes de données communes, tant descriptives que techniques ;
- c) Gestion des métadonnées ;

- d) Gestion des classifications (nomenclatures) et des normes au sein du gouvernement ;
- e) Repérabilité et accès aux données ;
- f) Gestion de la qualité des données ;
- g) Sécurité et confidentialité des données ;
- h) Gestion du cycle de vie de l'information.

Considération 2 : Les organismes nationaux de statistique devraient engager des discussions avec les administrations nationales pour examiner selon quelles modalités ils pourraient être associés à l'élaboration des stratégies nationales relatives aux données, soit à titre consultatif, soit en jouant un rôle directeur.

Considération 3 : Les organismes nationaux de statistique devraient travailler en étroite collaboration avec la communauté internationale pour partager les résultats obtenus dans le domaine de la gouvernance des données.

33. L'intendance des données englobe toutes les activités de gouvernance. Certaines d'entre elles sont d'ordre managérial, d'autres ont une dimension plus technique et d'autres encore pourraient inclure des compétences analytiques.

34. Les responsabilités et les compétences en matière d'intendance des données incluent les activités suivantes, sans toutefois s'y limiter :

- a) Promouvoir l'élaboration de données d'une haute qualité et les moyens d'optimiser leur utilisation ;
- b) Faciliter l'accès aux données, ce qui pourrait couvrir l'accès aux sources de données primaires, les fichiers de micro-données de recherche, la publication de données ouvertes, la maintenance des formats de données, etc., selon les fonctions de l'organisation ;
- c) Promouvoir l'expertise, l'éthique, les compétences et la maîtrise des techniques numériques ;
- d) Promouvoir des cadres, normes et politiques communs en matière de données ;
- e) Élaborer des stratégies, notamment pour le partage des données, la collaboration et d'autres aspects.

Considération 4 : Les organismes nationaux de statistique devraient encourager les gouvernements à nommer un responsable en chef ou autre haut fonctionnaire chargé des données, et inciter les institutions à désigner leur personnel chargé de l'intendance des données. On pourrait aussi mettre en place un conseil de gouvernance des données.

Considération 5 : Les organismes nationaux de statistique devraient contribuer à promouvoir la connaissance et la maîtrise des technologies numériques, spécialement en ce qui concerne les statistiques.

Considération 6 : Les organismes nationaux de statistique devraient travailler avec d'autres organismes gouvernementaux pour définir un ensemble minimum de responsabilités essentielles pour un responsable de la gestion des données. Il faudrait fournir aux organismes nationaux de statistique une panoplie d'instruments comprenant des bonnes pratiques, des conseils et des outils permettant d'exercer ce nouveau rôle.

B. Écosystèmes nationaux de données et services de données

35. À l'appui des écosystèmes nationaux de données, on pourrait élargir la portée des principes relatifs à l'architecture des données élaborés par la communauté des statisticiens, tels que l'Architecture commune des données statistiques et le Modèle générique

d'informations statistiques, de façon que ces principes puissent être appliqués en tant que normes régissant la mise en place d'une architecture pour les écosystèmes nationaux. Il est essentiel de procéder à un changement de culture en créant des conditions propices aux partenariats, au leadership et à une judicieuse prise de risques pour tirer parti des possibilités offertes par les nouveaux systèmes de données ;

Considération 7 : Les organismes nationaux de statistique pourraient travailler de concert avec les responsables des administrations numériques et d'autres organismes gouvernementaux pour mettre progressivement en place une gestion et une gouvernance des données qui soient conformes aux principes de l'architecture commune des données. Un argumentaire commun sur ce sujet s'avérerait utile.

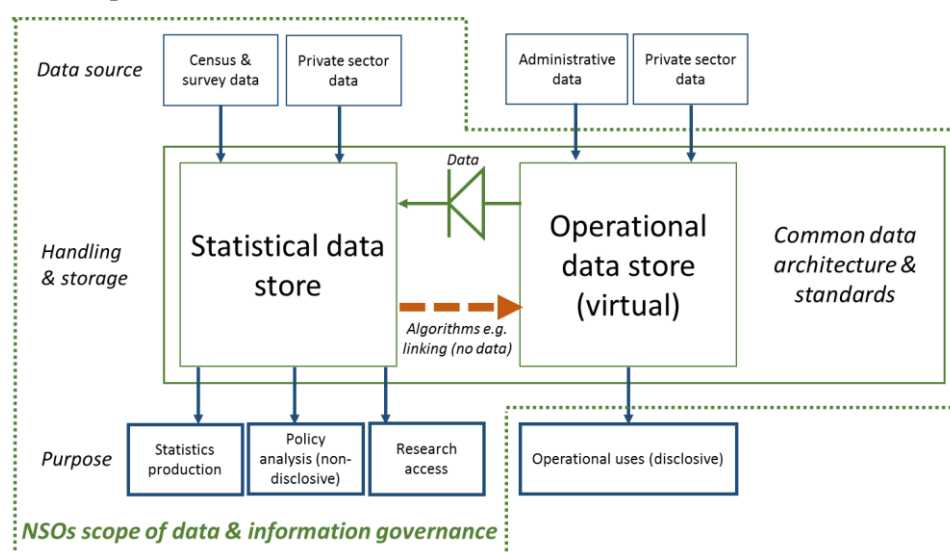
36. Pour appuyer la gouvernance des données et les services de données, les organismes nationaux de statistique doivent convenir d'une architecture qui permette de gérer des sources de données hétérogènes. La figure 1 présente une architecture de haut niveau pour la gouvernance des données. Le champ de la gouvernance est défini comme un lac de données comportant deux zones de stockage ou bases de données. Dans la base de données opérationnelles, sont stockées des données particulières et des données agrégées provenant de sources autres que statistiques. La base de données statistiques contient les données qui sont soumises aux impératifs de confidentialité. On peut transférer des données de la base de données opérationnelles vers la base de données statistiques mais l'inverse n'est pas possible.

Considération 8 : Les organismes nationaux de statistique devraient revoir leur infrastructure en ce qui concerne les données, les applications et les aspects techniques, pour s'assurer que les données sont hautement sécurisées et protégées, tout en facilitant l'intégration des données lorsque cela est nécessaire pour les besoins des statistiques et les services de données.

Considération 9 : Les organismes nationaux de statistique devraient faire savoir que le caractère confidentiel des données statistiques sera préservé dans la base de données correspondante. Les organismes nationaux de statistique doivent mener des discussions pour clarifier la différence entre les concepts de vie privée, de sécurité et de confidentialité (statistique).

Figure 1

Architecture de haut niveau pour la gouvernance des données faisant appel à la métaphore du lac de données



37. La qualité des données porte sur la capture et la documentation des métadonnées (y compris les définitions, les règles, les modèles logiques, les jeux de codes et les classifications), ainsi que l'identification des dépositaires des données et l'application d'un cadre de qualité des données.

Considération 10 : Les organismes nationaux de statistique devraient encourager les autres organismes gouvernementaux à revoir la qualité de leurs données, en particulier dans le cas des sources administratives, et à prendre des mesures pour les améliorer continuellement. Les organismes nationaux de statistique devraient contribuer à l'élaboration et à la mise en œuvre de normes pour la qualité des données dans les différents domaines d'application.

Considération 11 : Les organismes nationaux de statistique devraient envisager de constituer une équipe chargée de conseiller et d'aider les autres organismes gouvernementaux dans la mise en œuvre de la gestion de la qualité des données.

38. Les utilisateurs attendent un certain niveau de qualité de la part des organismes nationaux de statistique et il est donc nécessaire de prévoir des modalités de certification pour les statistiques publiées, les jeux de données et les services de données fournis par ces organismes et d'autres prestataires.

Considération 12 : Les organismes nationaux de statistique pourraient faire office d'organe certificateur pour les données détenues par les autres entités qui font partie du système statistique national. Cette certification devrait être fondée sur les mêmes normes que celles appliquées aux données détenues par lesdits organismes.

Considération 13 : Les organismes nationaux de statistique devraient participer aux initiatives en matière de données ouvertes et faire profiter celles-ci de leur expertise dans les domaines de la gouvernance et des services de données.

39. Une communication efficace visant à faire connaître ce que les organismes nationaux de statistique ont à offrir est cruciale pour créer des partenariats gagnant-gagnant avec les décideurs, d'autres organismes gouvernementaux et le secteur privé. La stratégie médiatique doit être considérée en lien étroit avec l'écosystème des données. Il convient de prendre tout particulièrement en considération l'engagement des citoyens et l'explication aux citoyens de l'utilisation des données par les organismes nationaux de statistique car les citoyens ont le droit de vivre dans une société informée et les organismes nationaux de statistique doivent le garantir.

Considération 14 : Les organismes nationaux de statistique doivent considérer la communication interne et externe comme une question stratégique pour sensibiliser les décideurs politiques, les milieux professionnels et le personnel du bureau. Les organismes nationaux de statistique devraient accorder une attention particulière aux besoins d'information des citoyens et à leurs préoccupations concernant l'utilisation des données.

C. La prochaine génération de statistiques et d'analyse statistique

40. La révolution des données permet une plus grande rapidité de l'analyse et une granularité plus poussée. Cette évolution facilitera la production d'informations plus détaillées sur les indicateurs socioéconomiques et les indicateurs du développement durable.

Considération 15 : Les organismes nationaux de statistique devraient travailler en étroite collaboration avec les clients pour établir en temps réel – ou presque – les indicateurs nécessaires et abandonner les indicateurs superflus.

Considération 16 : Les organismes nationaux de statistique devraient développer les capacités analytiques, plus précisément dans les domaines de la maîtrise des technologies numériques ainsi que de la science et de l'ingénierie des données, pour pouvoir appuyer efficacement la gestion des données.

Considération 17 : Les organismes nationaux de statistique devraient examiner attentivement les rôles qu'ils peuvent jouer en ce qui concerne les métadonnées. Ils pourraient par exemple gérer des métadonnées et non pas des données (en expliquant où se trouvent les données, ce qu'elles signifient, quelle en est la qualité) et contribuer au développement de systèmes de métadonnées et à leur interopérabilité.

41. La science des données et les statistiques fondées sur l'intelligence artificielle joueront un rôle de plus en plus important.

Considération 18 : Les organismes nationaux de statistique devraient appliquer la méthodologie et les outils de la science des données pour la production des statistiques et la prestation des services de données.

Considération 19 : Les organismes nationaux de statistique auraient intérêt à renforcer les partenariats avec les spécialistes de l'intelligence artificielle, tant dans les milieux universitaires que dans les différents secteurs d'activité.

42. Il existe une possibilité de soutenir le gouvernement et l'administration au sens large pour promouvoir une culture des données dans l'ensemble de l'écosystème des données administratives en tenant compte des questions d'éthique à cet égard.

Considération 20 : Les organismes nationaux de statistique devraient envisager de consulter d'autres administrations publiques.

Considération 21 : Les organismes nationaux de statistique devraient jouer un rôle actif dans les débats sur l'éthique des données.

V. La voie à suivre

A. Études de cas sur la mise en œuvre

43. Pour aider les organismes nationaux de statistique à jouer un rôle dynamique dans la conception et l'établissement d'écosystèmes nationaux de données conformes aux considérations ci-dessus, il serait utile de partager les résultats des initiatives concrètes mises en œuvre dans ce domaine. Il conviendrait également de créer à cette fin une plateforme et d'élaborer le ou les formats nécessaires pour conserver les informations de façon structurée.

44. On pourrait aussi aider les organismes nationaux de statistique en élaborant une feuille de route ou stratégie commune en matière de données. L'Irlande s'est ainsi dotée d'une stratégie pour les données de la fonction publique, le Canada a élaboré une feuille de route pour la stratégie du secteur public en matière de données, la Nouvelle-Zélande a mis au point une feuille de route et une stratégie en matière de données et l'Estonie a défini une stratégie pour la gouvernance des données. Globalement, ces stratégies offrent de grandes similitudes et reflètent les travaux en cours qui définissent les orientations et les cibles à atteindre. Elles suivent étroitement la plupart des thématiques évoquées, à savoir :

- a) Vision, comment traiter les données ;
- b) Confidentialité et transparence, comment appréhender les données et les utiliser ;
- c) Gouvernance et normes, définir les rôles et les responsabilités ;
- d) Analytique et science des données ;
- e) Infrastructures ;
- f) Partenariats ;
- g) Gestion des collections numériques et prestation de services de données.

45. Par suite, les auteurs du présent document proposent de passer à la prochaine étape. Sur la base des réponses à l'enquête réalisée pour le séminaire de juin 2019 et qui portait sur 48 pays, certains d'entre eux pourraient être invités à partager des informations plus détaillées sur leurs approches et les enseignements tirés dans la mise en œuvre de leur

stratégie de gouvernance des données. Ces informations pourraient être publiées sous la forme d'études de cas. Le Bureau de la Conférence pourra ainsi mieux identifier les questions nécessitant un débat plus approfondi.

B. Plan d'action

46. Il faudrait prévoir un plan d'action concret pour soutenir les organismes nationaux de statistique dans le cadre d'un écosystème de données plus vaste, ainsi que le partage des bonnes pratiques en ce qui concerne l'utilisation des données.

47. Les auteurs du présent document proposent d'envisager les étapes ci-après pour élaborer des orientations supplémentaires sur le rôle des organismes nationaux de statistique dans le nouvel écosystème de données :

a) Les points de la section IV A « Le rôle des organismes nationaux de statistique dans la gouvernance des données publiques » pourraient être examinés avec différentes parties prenantes, tant au niveau national qu'international. Les auteurs proposent d'organiser un séminaire sur le sujet afin de partager les nouvelles pratiques qui ont évolué au sein de la communauté statistique ;

b) Les points de la section IV B « Écosystèmes nationaux de données et services de données » pourraient être développés davantage avec le soutien et les conseils du Groupe de haut niveau sur la modernisation de la statistique officielle en vue de formuler des recommandations plus spécifiques. Dans ce domaine, il est également important de suivre les meilleures pratiques de la communauté statistique et d'en tirer des enseignements ;

c) Les points de la section IV C « La prochaine génération de statistiques et d'analyse statistique » pourraient être intégrés à d'autres groupes de travail (par exemple, la communication). L'évolution de la science des données aura une incidence considérable sur la production et la diffusion des statistiques. Ces sujets peuvent être abordés dans le cadre d'un débat plus large sur la manière dont les partenariats en matière de données sont constitués et sur les moyens de gérer ces réseaux.

48. Les auteurs se proposent d'examiner les résultats de l'enquête présentés au paragraphe 45 ainsi que les rapports établis par certains pays contenant des informations plus détaillées sur leurs approches et les enseignements tirés lors du séminaire de la Conférence, en juin 2020.

VI. Sources

- Role of National Statistical Systems in the New Data Ecosystem. Document établi pour la réunion du Bureau de la Conférence, octobre 2018.
https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/bur/2018/October/04-Redefining_national_statistical_systems.pdf.
- Modern Partnership Framework for Official Statistics. Document établi pour la réunion du Bureau de la Conférence, octobre 2018.
https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/bur/2018/October/08-Framework_for_establishing_partnerships_approved.pdf.
- Quelles stratégies pour les organismes nationaux de statistique à l'ère du numérique ? Vers des stratégies de « données intelligentes ». Document établi pour la quinzième réunion du Comité des statistiques et de la politique statistique de l'OCDE, juin 2018.
[http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/CSSP\(2018\)7&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/CSSP(2018)7&docLanguage=En).
- Report to the Clerk of the Privy Council: A Data Strategy Roadmap for the Federal Public Service (Gouvernement du Canada, septembre 2018).
https://www.canada.ca/content/dam/pco-bcp/documents/clk/Data_Strategy_Roadmap_ENG.pdf.
- Stratégie en matière de données et feuille de route pour la Nouvelle-Zélande. Présentation par le responsable gouvernemental de l'intendance des données, décembre 2018.
<https://data.govt.nz/assets/Uploads/data-strategy-and-roadmap-dec-18.pdf>.
- Peter Struijs, Sofie de Broe. Big Data Strategies for Official Statistics. 2018
<http://www.dgins2018.ro/wp-content/uploads/2018/10/20-NL-DGINS-Session-2-paper-Struijs-Big-Data-Strategies.pdf>.
- Coopération pour le développement, Rapport 2017. Données au service du développement (OCDE, octobre 2017)
https://www.oecd-ilibrary.org/development/development-co-operation-report-2017_dcr-2017-en.
- A World That Counts. Mobilising the Data Revolution for Sustainable Development. Groupe consultatif d'experts indépendants, Organisation des Nations Unies, Novembre 2014.
<http://www.undatarevolution.org/wp-content/uploads/2014/12/A-World-That-Counts2.pdf>.
- Data ecosystems for sustainable development. An assessment of six pilot countries. PNUD, septembre 2017
<http://www.undp.org/content/dam/undp/library/Sustainable%20Development/Data%20Ecosystems%20for%20Sustainable%20Development.pdf>.
- Conférence des statisticiens européens de la CEE « Guidance on modernising statistical legislation ».
<http://www.unece.org/index.php?id=51141>.