



Conseil économique et social

Distr. générale
16 décembre 2016
Français
Original : anglais

Commission de statistique

Quarante-huitième session

7-10 mars 2017

Point 3 d) de l'ordre du jour provisoire*

**Questions soumises à la Commission pour examen et décision :
les mégadonnées en statistique officielle**

Rapport du Groupe de travail mondial sur l'utilisation des mégadonnées en statistique officielle

Note du Secrétaire général

En application de la décision 2016/220 du Conseil économique et social, et conformément à la pratique établie, le Secrétaire général a l'honneur de transmettre le rapport du Groupe de travail mondial sur l'utilisation des mégadonnées en statistique officielle. Au cours des trois dernières années, le Groupe de travail a recensé plusieurs initiatives prometteuses de la part des secteurs public et privé concernant les mégadonnées. D'ampleur nationale ou régionale, ces initiatives (qui sont menées par des organisations publiques ou privées) ont pour objectif de rendre les données, les services et les applications accessibles et de favoriser leur synergie pour les mettre au service de la recherche et du renforcement des capacités. Dans son rapport, le Groupe de travail présente les progrès qu'il a accomplis, ainsi que des idées et propositions d'action à l'avenir. Il y propose également de créer une plateforme mondiale qui permettrait de regrouper toutes les initiatives susmentionnées favorisant l'utilisation des mégadonnées et d'autres sources de données dans un contexte plus large au niveau mondial et qui serait gage d'évolutivité et d'extensibilité des données, services, applications et infrastructures, et permettrait d'exploiter les mégadonnées et de les intégrer dans les sources administratives, l'information géospatiale et les données traditionnelles d'enquête et de recensement. Une telle plateforme serait mise en place sous l'égide de la Commission de statistique, qui est invitée à donner son avis concernant cette proposition, ainsi que sur les progrès accomplis par le Groupe de travail et ses recommandations. Les questions à examiner sont présentées dans la dernière partie du rapport.

* E/CN.3/2017/1.



Rapport du Groupe de travail mondial sur l'utilisation des mégadonnées en statistique officielle

I. Introduction

1. La Commission de statistique a créé le Groupe de travail mondial sur l'utilisation des mégadonnées en statistique officielle à sa quarante-cinquième session en 2014. Conformément à son mandat (voir E/CN.3/2015/4) et à la décision 46/101 de la Commission (voir E/2015/24-E/CN.3/2015/40), le Groupe de travail donne l'orientation stratégique et assure la direction et la coordination d'un programme mondial sur les mégadonnées en statistique officielle, notamment pour les indicateurs du Programme de développement durable à l'horizon 2030. Il encourage aussi le renforcement des capacités, la formation et l'échange d'expériences, et l'utilisation des mégadonnées dans l'élaboration des politiques.

2. Le Groupe de travail a rendu compte de ses travaux aux quarante-sixième et quarante-septième sessions de la Commission de statistique, en 2015 et 2016. Il a présenté les points forts de ses première et deuxième conférences internationales, tenues à Beijing et Abou Dhabi, et de son programme de travail, en abordant des questions de formation, d'amélioration des compétences et de renforcement des capacités; les liens entre les mégadonnées et les objectifs de développement durable; les efforts de promotion et de communication; la question de l'accès aux données et les partenariats; et des questions intersectorielles telles que les classifications et les cadres, ainsi que l'exploration de sources spécifiques de mégadonnées en statistique officielle, à savoir les données tirées des téléphones mobiles et des réseaux sociaux, et l'imagerie satellitaire.

3. En 2016, la Commission a prié le Groupe de travail de prendre un ensemble d'initiatives visant à améliorer l'accès aux données exclusives, telles que les mégadonnées; de mettre au point, en coopération avec des centres de formation régionaux, des cours sur l'utilisation et l'application des mégadonnées, en ciblant plus particulièrement les pays en développement; et d'affiner le cadre de qualité pour les mégadonnées afin de garantir la fiabilité des statistiques officielles établies.

4. La Commission a également demandé que les points ci-après soient ajoutés au programme de travail: la récupération de mégadonnées sur le Web; l'amélioration du site Web du Groupe de travail et du répertoire concernant les mégadonnées; la promotion de l'utilisation des mégadonnées dans les pays en développement; et l'établissement de liens entre les travaux du Groupe de travail mondial et ceux du Groupe de haut niveau pour le partenariat, la coordination et le renforcement des capacités dans le domaine des statistiques relatives au Programme de développement durable à l'horizon 2030.

5. Le Groupe de travail a donné suite aux demandes de la Commission, comme il ressort des progrès qu'il a accomplis en 2016, dont on trouvera un aperçu dans la deuxième partie du présent rapport, et des conclusions de la troisième Conférence internationale sur l'utilisation des mégadonnées en statistique officielle, présentées en troisième partie. La quatrième partie traite plus particulièrement de l'accès aux données exclusives tandis que la cinquième partie est consacrée à la proposition de plateforme mondiale pour la mise en commun des données, services et applications. En conclusion, plusieurs points sont proposés pour examen.

II. Progrès du Groupe de travail

6. En 2016, le Groupe de travail a tenu sa troisième Conférence internationale sur l'utilisation des mégadonnées en statistique officielle à Dublin, préparé une proposition de plateforme mondiale visant à encourager la collaboration en matière de données, services et applications, et décidé d'inclure la question de l'accès aux données exclusives dans l'examen général du cadre de qualité des statistiques officielles. Il a également avancé sur d'autres questions, telles que la collaboration sur des projets pilotes par le biais du bac à sable de la Commission économique pour l'Europe (CEE), l'établissement de plusieurs manuels d'utilisation des données satellite, des téléphones mobiles et des réseaux sociaux, et la mise à niveau de son site Web, notamment en rendant disponible un inventaire des projets relatifs à l'utilisation des mégadonnées. On trouvera détaillés dans la présente partie les progrès réalisés par les équipes spéciales dans différents domaines. Un résumé des principales conclusions de la troisième Conférence internationale, la proposition de plateforme mondiale et les recommandations concernant l'accès aux données exclusives seront exposés dans les parties suivantes.

A. Données satellite

7. L'équipe spéciale du Groupe de travail chargée de l'imagerie satellitaire, des données géospatiales et de la télédétection a établi un manuel sur les sources de données d'observation de la Terre et les méthodes de production de statistiques agricoles et autres reposant sur l'utilisation des données d'imagerie satellitaire, dans lequel elle présente aussi dans les grandes lignes ses projets pilotes et ses recommandations aux organismes nationaux de statistique qui n'ont encore jamais utilisé des données d'observation de la Terre. Les projets pilotes portent sur l'application des données issues de l'imagerie satellitaire à la production de statistiques agricoles par l'Institut australien de la statistique, l'utilisation de satellites privés pour évaluer les stocks de matières premières grâce à Terra Bella (une filiale de Google), l'analyse environnementale de scénarios climatiques dans différentes régions du Mexique par l'Institut national de statistique et de géographie mexicain (INEGI), et l'utilisation de l'imagerie satellitaire dans la production de statistiques sur la couverture et l'occupation des sols par le Département administratif national de statistique colombien (DANE)]. Le manuel devrait être finalisé en 2017, après examen lors d'une réunion d'experts de l'agriculture et de la couverture du sol venus de l'Organisation des Nations Unies pour l'alimentation et l'agriculture et de la Commission européenne.

8. L'équipe spéciale met actuellement au point des formations pratiques aux méthodes d'utilisation des données d'observation de la Terre pour produire des statistiques agricoles, et propose des études de cas sur l'application de ces données à la mesure des progrès accomplis sur la voie des objectifs de développement durable. Il s'agira de présenter les possibilités d'utilisation des données issue de l'imagerie satellitaire en statistique officielle et d'offrir des conseils sur le choix des sources et des méthodes de production de statistiques, à la faveur d'exposés, de séances de discussion, de démonstrations de l'utilisation du progiciel R -un exemple de traitement des données satellite libres d'accès à des fins d'analyse avec le moteur Google Earth-, et de l'apport d'une aide concrète à la production de statistiques à partir de données issues de l'imagerie satellitaire. Il est prévu d'adapter par la suite

ces outils pédagogiques pour pouvoir les utiliser pour la comptabilité des écosystèmes dans le domaine agricole. Les supports de formation seront examinés lors d'un premier atelier en janvier 2017, avant la tenue d'autres ateliers en Asie et en Amérique latine.

B. Données issues des téléphones mobiles

9. L'équipe spéciale chargée des données issues des téléphones mobiles a entrepris d'établir un manuel, qui devrait être finalisé en 2017, à partir des leçons tirées de plusieurs projets d'Eurostat et d'un projet réalisé à Oman en collaboration avec l'entreprise Positium. On s'inspirera également du travail accompli dans le cadre de projets des Centres et Réseaux d'excellence du système statistique européen (réseau ESSNet) aux Pays-Bas, en France, en Belgique, en Slovénie et en Espagne. Ce manuel présentera notamment des modèles opérationnels de partenariats entre organismes nationaux de statistique et opérateurs de téléphonie mobile, et le rôle qu'ils peuvent jouer dans la production de statistiques dans les domaines du tourisme (international, national et régional), du recensement de la population diurne (excursions touristiques à la journée), des événements, des transports et de la densité de population. Il offrira également des conseils concrets, des exemples de pratiques nationales et de nombreuses références, qui s'avèreront particulièrement utiles pour les pays en développement.

C. Données issues des réseaux sociaux

10. De même, l'équipe spéciale chargée des données issues des réseaux sociaux travaille à l'établissement d'un manuel qui devrait être achevé en 2017 expliquant la manière dont les techniques de traitement des mégadonnées peuvent être appliquées aux données issues des réseaux sociaux pour en tirer des statistiques. Ce manuel décrira notamment plusieurs sources de données issues des réseaux sociaux (comme Twitter et Facebook), les méthodes et techniques de traitement correspondantes (dont le traitement automatique des langues), et les méthodes d'extraction d'informations à partir d'éléments des réseaux sociaux autres que les messages et du contenu de pages Web. Il renverra aussi à des applications qui permettent d'établir des indices de confiance des consommateurs, des indicateurs de l'opinion publique, et des analyses de la mobilité et des sentiments. En plus du manuel, l'équipe Global Pulse, membre de l'équipe spéciale chargée des réseaux sociaux, a conclu un accord avec Twitter sur l'utilisation d'importants volumes de tweets chaque mois, qui peuvent être partagés avec les membres du Groupe de travail via le bac à sable.

D. Inventaire des projets relatifs à l'utilisation des mégadonnées

11. En parallèle de l'enquête sur les projets relatifs aux mégadonnées menée en 2015 par le Groupe de travail, l'équipe spéciale chargée des mégadonnées et des objectifs de développement durable a étudié l'utilisation des mégadonnées dans le cadre spécifique des objectifs de développement durable dans plusieurs institutions. Les informations recueillies sont venues enrichir l'inventaire des projets relatifs à l'utilisation des mégadonnées, un catalogue de projets sur l'utilisation des mégadonnées aux fins de la production de statistiques officielles et d'indicateurs

d'objectifs de développement durable, et d'autres types de statistiques utiles à la gestion et au suivi des projets et programmes du secteur public. Les procédures de gestion et de mise à jour de l'inventaire ont été standardisées, et l'alignement des projets relatifs aux mégadonnées sur les objectifs de développement durable et leurs cibles devrait permettre une meilleure utilisation de l'inventaire en statistique officielle. En dressant cet inventaire, initiative inédite, le Groupe de travail mondial poursuit ses efforts de promotion de toutes ces pratiques dans le cadre de l'utilisation des mégadonnées en statistique officielle et par les spécialistes des mégadonnées.

E. Formation, compétences et renforcement des capacités

12. Plus les statisticiens recourent aux mégadonnées et aux nouvelles infrastructures informatiques en statistique officielle, plus les besoins en nouvelles compétences et nouvelles méthodes basées sur ces infrastructures sont grands pour que les mégadonnées puissent être appliquées aux processus de production de statistiques issues de plusieurs sources. Cette multiplicité de sources présente plusieurs problèmes, tant dans la manière dont on traite et analyse les données que dans l'utilisation des infrastructures informatiques. Pour résoudre ces problèmes de transfert de connaissances, l'équipe spéciale chargée de la formation, des compétences et du renforcement des capacités a concentré son travail sur quatre grands objectifs :

a) Mettre au point des méthodes et des outils (notamment de coordination et de facilitation) permettant d'évaluer les principaux besoins des systèmes de statistique nationale en matière de compétences s'agissant des mégadonnées et l'aptitude des institutions à utiliser les mégadonnées pour produire des statistiques;

b) Prodiguer des conseils en vue de l'élaboration d'un programme de formation en plusieurs modules adapté aux besoins en compétences et au degré d'aptitude des institutions mis en évidence par l'évaluation ;

c) Faciliter la création d'un réseau mondial d'instituts de formation et de recherche en vue de former du personnel à l'utilisation des mégadonnées en statistique et de renforcer les capacités dans ce domaine;

d) Coordonner des programmes de formation plus avancés et les aligner sur les résultats de l'enquête mondiale sur les mégadonnées en statistique officielle et les cours de formation mis au point par les différentes équipes spéciales du Groupe de travail.

13. Les résultats attendus par l'équipe en 2017 sont les suivants:

a) Évaluation des principaux besoins en compétences et du degré d'aptitude des institutions grâce à la réalisation d'une nouvelle enquête mondiale sur les mégadonnées, y compris auprès des instituts nationaux de statistique;

b) Cours de formation permettant l'acquisition des compétences essentielles nécessaires et préparation des institutions à l'utilisation des mégadonnées dans la production de statistiques, en collaboration avec les organismes nationaux de statistique, les institutions universitaires et les partenaires du secteur privé ;

c) Création d'un réseau mondial de renforcement des capacités rassemblant les institutions nationales et régionales de formation et de recherche, et des centres d'innovation en matière de données.

III. Résultats de la troisième Conférence internationale sur les mégadonnées en statistique officielle

14. La troisième Conférence internationale s'est tenue à Dublin les 30 et 31 août et 1er septembre 2016 en présence d'environ 250 statisticiens venus du monde entier. La Conférence a été ouverte par le Ministre d'État aux services financiers, au cybergouvernement et aux marchés publics de l'Irlande, Eoghan Murphy, qui a souligné tout l'intérêt présenté par les mégadonnées dans le contexte du Programme de développement durable à l'horizon 2030. La Conférence a été l'occasion d'exposer les avancées et les résultats des projets relatifs aux mégadonnées tirées des téléphones mobiles, des réseaux sociaux ainsi que de l'imagerie satellitaire dans diverses applications statistiques. Ont aussi été évoqués les progrès accomplis dans les domaines du renforcement des capacités, de l'accès aux données et des partenariats, de la qualité et de la méthodologie ainsi que dans les stratégies adoptées pour mieux faire connaître l'intérêt des mégadonnées. Tous les exposés sont disponibles sur le site Web de la Conférence.

15. La Conférence a porté sur trois grands thèmes, à savoir l'accès aux mégadonnées et les partenariats en la matière; les stratégies de renforcement des capacités en vue de la transformation et de la modernisation de la statistique officielle; et l'instauration d'un environnement qui permette la production de statistiques officielles à partir de sources variées et de mégadonnées, notamment pour le suivi des objectifs de développement durable.

A. Accès aux données et partenariats

16. Le groupe de haut-niveau sur l'accès aux données et les partenariats a souligné que l'accès aux mégadonnées issues entre autres des téléphones mobiles et des réseaux sociaux, dans le cadre de partenariats public-privé, présenterait d'énormes avantages pour les activités des partenaires privés comme des organismes nationaux de statistique. Ces partenariats devaient mettre à profit l'infrastructure technologique de l'informatique en nuage, l'expertise du secteur privé en matière de données et de leur analyse, et la confiance placée par le public dans les organismes nationaux de statistique pour qu'ils fournissent des données impartiales et de grande qualité en toute indépendance et avec professionnalisme. La question du maintien de la confiance du public a été vue comme essentielle pour permettre l'utilisation de mégadonnées pour la production de statistiques officielles. Dans cette optique, il convient de faire preuve de la plus grande transparence dans la collecte, le stockage et l'utilisation des données issues des téléphones mobiles et des réseaux sociaux, tout en respectant les cadres réglementaires et juridiques applicables aux statistiques officielles et à la protection des données. De tels partenariats permettraient notamment de réduire les coûts et la charge que représente la communication de données et d'améliorer le degré d'actualité et la qualité des statistiques. Le milieu de la statistique officielle devrait toutefois reconnaître qu'il sera nécessaire de réfléchir à des modes de financement novateurs pour tenir compte des intérêts

économiques du secteur privé dans le cadre de ces partenariats. La question de l'accès aux données exclusives est examinée plus en détail dans la quatrième partie du présent rapport.

B. Renforcement des capacités d'innovation en matière de statistique officielle

17. Une évolution des services de statistique proposés aux clients nationaux et internationaux s'impose, puisqu'il faut désormais savoir s'accommoder des facteurs de risque des initiatives externes, de la concurrence des autres fournisseurs de statistiques et des coupes budgétaires, tout en tirant parti des mutations apportées au paysage de l'informatique et de l'information par un développement technologique rapide et la révolution des données.

18. Les groupes sur le renforcement des capacités ont indiqué qu'il fallait modifier différents aspects des systèmes statistiques nationaux pour relever les nouveaux défis et saisir les occasions qui se présentent. Il faudrait élargir l'accès aux programmes existants de renforcement des capacités, et éventuellement faire en sorte qu'ils ciblent la transformation de l'architecture informatique et celle des effectifs, l'exploitation de sources de données supplémentaires, et la réorientation des produits et services. Les groupes ont déclaré que la transformation de l'architecture informatique devrait faciliter le passage d'une utilisation de matériel informatique sur site à celle d'un environnement en ligne basé sur l'informatique en nuage ainsi que l'adoption d'une architecture des application et de services communs pour ce qui est de la collecte de données, des registres, des métadonnées, et de la gestion, de l'analyse et de la diffusion des données.

19. Cet investissement devrait s'accompagner de programmes de renforcement des capacités visant à appuyer la diversification progressive des nouvelles compétences du personnel des systèmes statistiques nationaux, qu'il s'agisse de l'utilisation de nouvelles données multisources et d'une technologie moderne par les spécialistes des données et des systèmes d'information, du renforcement du cadre juridique par les juristes, ou encore de l'action des administrateurs pour faire évoluer la culture d'entreprise en fixant des normes de qualité toujours plus exigeantes. Ces nouvelles capacités devraient permettre l'adoption d'une structure d'entreprise standardisée et flexible qui puisse s'adapter aux besoins émergents, être axées sur les procédures plutôt que sur les produits et s'appuyer de plus en plus sur des sources administratives et sur les mégadonnées (plutôt que sur les sources de données primaires) pour permettre la production de statistiques multiples. Il faudrait également améliorer la stratégie de diffusion et de communication et la rendre suffisamment souple pour cibler différentes catégories d'utilisateurs, en appliquant des techniques variées de dissémination des données, notamment en recourant à des applications sur téléphone mobile et à des graphiques permettant de visualiser les résultats les plus importants.

20. Il a été proposé d'élargir le champ des programmes régionaux de formation à la statistique en ce qui concerne le renforcement des capacités, dans les domaines de la transformation et de la modernisation des instituts pour les cadres supérieurs, et de la création de compétences relatives au génie des systèmes d'information, à l'analyse et à la science des données pour le personnel technique et les membres des services informatiques. Afin de permettre le renforcement des capacités dans le

cadre de ces nouveaux programmes, les instituts régionaux de formation à la statistique pourraient se munir de plateformes régionales et mondiales d'informatique en nuage puisant dans des sources multiples et dotées de capacités de traitement des mégadonnées. Corrélativement, il a été proposé d'instituer un réseau international des instituts régionaux de formation et de recherche, articulé autour de plateformes régionales et mondiales, dans le but de créer l'infrastructure institutionnelle, technologique et informationnelle nécessaire à la mise en place d'une formation continue aux nouvelles méthodes d'établissement de statistiques, compétences d'administration et techniques informatiques. Les réseaux régionaux connectés par une plateforme mondiale seraient complémentaires des instituts nationaux de formation existants et des nouveaux centres nationaux de données et de recherche.

21. Le groupe chargé de l'investissement intelligent pour l'innovation dans le domaine de la statistique officielle a indiqué qu'en raison de l'évolution rapide des paysages de la technologie et de l'information, il fallait repenser la façon dont la communauté des donateurs soutenait les systèmes nationaux, régionaux et internationaux de statistique. Face à cet environnement en mutation rapide, il importe en premier lieu de centrer les investissements intelligents sur des projets à petite échelle visant à lancer et tester des méthodes et des technologies nouvelles avant de les appliquer à plus large échelle. Ces projets et programmes innovants passent par la conclusion de nouveaux partenariats public-privé collaboratifs et stratégiques entre le milieu de la statistique officielle, les universitaires, les entreprises du secteur des technologies et les propriétaires de données exclusives. Il paraît en outre nécessaire de changer de modèle pour pouvoir financer les dépenses renouvelables liées à l'accès continu aux sources privées de mégadonnées et à l'édification d'une infrastructure technologique moderne reposant sur l'informatique en nuage. Pour ce faire, les donateurs doivent passer d'un mode de financement en bloc aux fins d'investissements structurels dans les systèmes statistiques nationaux à un mode de financement en série, sur le long terme, à l'appui de plateformes régionales et mondiales communes et adaptables permettant l'accès à des données, des services et des applications par l'intermédiaire de partenariats stratégiques.

C. Mégadonnées et indicateurs des objectifs de développement durable

22. Les experts des groupes sur les mégadonnées et les objectifs du développement durable ont également souligné qu'il fallait adopter une approche multimodale et multisource pour répondre aux exigences du cadre d'indicateurs pour les objectifs de développement durable. La difficulté ne tient pas au seul nombre des indicateurs, mais aussi à la nécessité de produire des mesures extrêmement précises pour "ne pas faire de laissés-pour-compte". C'est en particulier pour atteindre ce degré de précision qu'il sera nécessaire de recourir à des sources administratives et aux mégadonnées. Sur fond d'un accroissement de la demande en matière d'accès à la technologie, aux données et à des applications à l'appui des principaux programmes de statistique, on a observé des évolutions prometteuses au Sénégal, au Cap-Vert et en Afrique du Sud en ce qui concerne la création de centres nationaux et régionaux de technologie et de données. Dans l'ensemble, les spécialistes réunis à la troisième Conférence internationale ont

accueilli l'idée d'une plateforme mondiale comme une évolution positive pour la coopération internationale et régionale aux fins de la mise en commun des données, services et applications, qui devrait contribuer à résoudre les difficultés inhérentes au suivi des indicateurs relatifs aux objectifs de développement durable.

23. On trouvera davantage de détails sur les résultats de la troisième Conférence internationale dans un document d'information relatif au présent rapport.

IV. Recommandations sur l'accès aux mégadonnées en tant qu'élément constitutif d'un cadre d'assurance qualité des statistiques officielles

24. L'accès aux données est visé par le cinquième Principe fondamental de la statistique officielle. On y lit que les données utilisées à des fins statistiques peuvent être tirées de toutes sortes de sources, qu'il s'agisse d'enquêtes statistiques ou de fichiers administratifs. Les organismes responsables de la statistique doivent choisir leur source en tenant compte de la qualité des données qu'elle peut fournir, de leur actualité, des coûts et de la charge qui pèse sur les personnes sondées. On peut considérer que ce principe s'applique également à l'accès aux mégadonnées.

25. Le Groupe de travail s'inquiète de ce que cet accès aux mégadonnées suppose, dans de nombreux cas, d'accéder à des données détenues par des propriétaires de données exclusives issus du secteur privé et ayant des intérêts économiques. C'est en cela que l'accès aux mégadonnées se distingue de l'accès aux sources traditionnelles et les inquiétudes soulevées à cet égard tiennent à la capacité des instituts de statistiques à passer des accords avec les propriétaires de données tout en conservant leur indépendance, en préservant la vie privée des personnes sondées et la confidentialité des données, et en produisant des statistiques de grande qualité. Les questions qui doivent être examinées ont notamment trait à la répartition égale de la charge entre les propriétaires de données, au coût de la fourniture de données et aux efforts déployés à cette fin, à la place des données dans les propositions de valeur des entreprises, au juste équilibre entre les intérêts publics et privés et aux accords opérationnels entre instituts de statistique et fournisseurs de données. Ces accords devront peut-être prévoir des dispositions relatives à la transparence de la fourniture des données et être assortis d'une description précise des métadonnées et de mentions légales publiques portant sur l'utilisation des données afin d'assurer à la fois la qualité des statistiques officielles et le maintien de la confiance du public.

26. En 2015, le Groupe de travail a élaboré des principes pour l'accès aux sources de métadonnées aux fins de statistiques officielles en vue d'aider les organismes de statistique à accéder à de nouvelles sources de données tout en instaurant un équilibre entre les intérêts légitimes des organismes, souvent privés, qui détiennent les métadonnées, et l'intérêt public qu'il y a à disposer de meilleures statistiques officielles. Ces principes d'accès sont liés aux Principes fondamentaux de la statistique officielle, et en appellent à la responsabilité sociale des fournisseurs de données. Le Groupe de travail en a fixé les termes après avoir consulté les parties concernées à plusieurs occasions en 2015, et le texte a reçu un accueil globalement positif.

27. À la réunion qu'il a tenue en août 2016 à Dublin, le Groupe de travail a discuté des prochaines mesures à prendre pour que la communauté mondiale de la

statistique connaisse et reconnaisse ces principes. Rappelant qu'il avait fallu 20 ans pour faire adopter les Principes fondamentaux, les participants sont tout d'abord convenus que le terme "principe" était peut-être trop fort pour que le texte soit approuvé par les différents organes internationaux, et qu'il serait préférable d'utiliser plutôt celui de "recommandation". Le processus de consultation avec les parties prenantes se poursuivrait en outre par l'intermédiaire d'une enquête ciblée. Plus important encore, il a été convenu à la réunion que les "recommandations sur la question de l'accès" devraient être examinées par la Commission dans un cadre exhaustif comprenant les Principes fondamentaux de la statistique officielle, les codes d'usage de la statistique officielle et les cadres nationaux d'assurance de la qualité.

28. Dans un rapport sur les Principes fondamentaux de la statistique officielle (E/CN.3/2017/9) soumis à la Commission à sa quarante-huitième session, la Commission a été priée d'envisager la création d'un groupe des « amis de la présidence », dont le mandat consisterait entre autres à réexaminer la façon dont les Principes fondamentaux devraient être interprétés ou élargis pour comprendre le secteur privé, en particulier en ce qui concerne l'utilisation des données exclusives. Le groupe des "amis de la présidence" pourrait souhaiter prendre en compte les inquiétudes du Groupe de travail au sujet de l'accès à ces données. On trouvera davantage de détails sur la question de l'accès aux données exclusives dans un document d'information relatif au présent rapport.

V. Plateforme mondiale pour la mise en commun des données, services et applications

29. En novembre 2014, le Secrétaire général a transmis le rapport du Groupe consultatif d'experts indépendants sur la révolution des données pour le développement durable intitulé « A world that counts: mobilizing the data revolution for sustainable development » (Un monde qui compte : mettre la révolution des données au service du développement durable). Y figuraient des recommandations sur les questions suivantes : a) la technologie, les innovations et les analyses nécessaires pour établir un réseau de centres d'innovation en matière de données permettant d'exploiter et de mettre en commun les données et les recherches sur les données; b) le renforcement des capacités et le transfert de technologie ainsi que les ressources connexes, l'aptitude à exploiter les données et la mobilisation de ressources au moyen de modes de financement novateurs développés en partenariat avec le secteur privé; et c) la gouvernance et l'encadrement des partenariats et des activités de coordination entre les gouvernements, le secteur privé, les organisations non gouvernementales, les médias et le milieu universitaire visant à promouvoir les pratiques et principes recommandables en ce qui concerne la mise en commun des données, les données ouvertes et le droit des données.

30. Au cours des trois dernières années, le Groupe de travail a recensé de nombreuses et prometteuses initiatives publiques et privées lancées au niveau national et régional relatives aux mégadonnées d'origine publique ou privée visant à rendre les données, les services et les applications accessibles et à favoriser leur synergie pour les mettre au service de la recherche et du renforcement des capacités. Le rôle des équipes spéciales du Groupe de travail spécialisées dans les données

satellites et celles qui sont issues des téléphones mobiles et des réseaux sociaux consiste à rendre ces applications accessibles grâce à des manuels accompagnés de supports de formation et de cours. Des équipes spéciales supplémentaires sont en cours de création, et il est proposé d'en établir encore davantage, qui travailleront dans des domaines prometteurs dont l'utilité en statistique officielle est avérée, par exemple la compilation d'indices de prix à l'aide de données provenant de scanners de caisses enregistreuses et de la capture de pages Web. Cependant, aucun effort concerté n'a encore été fait pour créer une plateforme que l'on puisse qualifier de « réseau de centres d'innovation en matière de données ».

31. À la suite d'une décision prise lors de sa réunion du 29 août 2016 à Dublin, le Groupe de travail a proposé la mise en place d'une plateforme mondiale pour la mise en commun des données, services et applications. Fondée sur les meilleures pratiques recensées dans le cadre des initiatives publiques et privées relatives aux mégadonnées et offrant aux statisticiens officiels une infrastructure et un réseau consacrés à l'innovation en matière de données, une telle plateforme répondrait aux besoins en pôles mondiaux dans les domaines suivants :

a) Échange, entre statisticiens officiels, scientifiques spécialistes des données et autres experts venus des secteurs public et privé, d'idées et de méthodes pour traiter, analyser et représenter visuellement les mégadonnées;

b) Stockage des mégadonnées et méthodes pour les traiter, les analyser et les représenter visuellement, et services et applications pour la production de données et leur utilisation à de nouvelles fins;

c) Démonstration que les mégadonnées peuvent améliorer la prise de décisions, grâce à l'exploitation de statistiques officielles dans le cadre d'études pilotes et d'études de cas;

d) Regroupement des ressources éducatives et organisation d'ateliers sur les mégadonnées pour participer au renforcement des capacités.

32. Le Groupe de travail a également proposé que la plateforme soit développée sous l'égide de la Commission, qui déciderait donc de son modèle d'activité et de la marche à suivre pour sa conception et sa mise en place. Il s'agirait d'une plateforme placée sous la supervision de la Commission, créée par et pour les producteurs de statistiques officielles, destinée avant tout à être utilisée par les membres des services statistiques nationaux des pays développés et en développement. Comme indiqué ci-dessus, les divers groupes de la troisième Conférence internationale sur l'utilisation des mégadonnées en statistique officielle et des conférences régionales sur l'élaboration d'un programme de transformation de la statistique officielle, organisées en 2016, ont examiné et appuyé la proposition de création d'une plateforme mondiale. Cette proposition est conforme aux objectifs de l'axe innovation et modernisation des systèmes statistiques nationaux du projet de plan d'action mondial pour les données du développement durable du Groupe de haut niveau pour le partenariat, la coordination et le renforcement des capacités dans le domaine des statistiques relatives au Programme de développement durable à l'horizon 2030 (voir E/CN.3/2017/3) et aux orientations données dans les documents finaux des conférences régionales (voir E/CN.3/2017/5).

Modèle d'activité proposé pour la plateforme mondiale

33. Afin d'aider la Commission dans sa réflexion sur le contenu de la plateforme mondiale et la forme qu'elle prendra, on trouvera proposés dans les paragraphes suivants une ébauche de modèle d'activité, ainsi qu'un scénario envisageable pour organiser sa mise en place. Le Groupe de travail a créé un comité chargé d'explorer d'autres options en vue de la mise en place de la plateforme.

34. La plateforme mondiale pourrait représenter une solution évolutive pour assurer l'accès aux mégadonnées et leur bonne utilisation, les intégrer aux sources administratives, aux données géospatiales et aux données traditionnelles provenant d'enquêtes et de recensements, ainsi que pour permettre l'utilisation des services, applications et infrastructures connexes. En ce qui concerne les produits et les services, la plateforme pourrait : a) faciliter l'accès à l'infrastructure et aux sources de données exclusives, ainsi que leur utilisation; b) accueillir des projets spécifiques organisés à l'échelle mondiale, portant par exemple sur les registres mondiaux d'entreprises ou sur les asymétries de données bilatérales ou multilatérales; c) héberger et rendre accessibles de nouvelles applications libres permettant d'exploiter des ensembles de données provenant de sources multiples, pour rendre plus rapide et flexible la production de statistiques officielles; et d) offrir des services de formation et de renforcement des capacités dans le domaine de l'utilisation de nouvelles sources de données pour la production de statistiques. Pour faire de la plateforme un outil véritablement global de production de statistiques officielles, il est envisagé de faire en sorte qu'elle donne aussi accès à des micro-ensembles de données statistiques fiables provenant de sources officielles ou encore à des données vérifiées et conservées par des organisations partenaires des secteurs privé et public. La plateforme pourrait offrir, entre autres, des services de stockage dans le nuage et d'informatique en nuage, des services statistiques techniques (notamment des services autonomes ou qui puissent servir de composantes lors du développement d'applications) ou encore des services connexes de renforcement des capacités.

35. La fourniture de services et d'applications grâce à la plateforme pourrait fonctionner de la manière suivante : on pourrait organiser un projet dont l'objectif serait de donner accès à des données tirées d'images satellite et aux infrastructures connexes, et de transférer la méthodologie nécessaire pour exploiter ces données aux fins de l'établissement de statistiques agricoles pour un pays développé ou en développement. Pour mener à bien un tel projet, il faudrait avoir accès à des données issues de satellites pour pouvoir organiser des essais et des formations, disposer d'un environnement informatique adéquat et sûr où préparer les données, tester les méthodes employées, compiler les statistiques agricoles et former le personnel afin qu'il soit capable de mener régulièrement des opérations liées au projet. La plateforme mondiale serait à même d'offrir de tels services, y compris dans les domaines de la formation et du renforcement des capacités, puisque les membres de ce réseau de partenaires viendraient d'organismes de statistique, d'instituts de recherche et d'entreprises privées et qu'il serait équipé des infrastructures nécessaires.

36. Il faut également prendre en compte le cadre institutionnel dans lequel s'inscrira la plateforme. Le Groupe de travail a créé un comité composé de représentants des organismes nationaux de statistique, notamment ceux ayant récemment mis en place un centre national de données, chargé de régler les détails

de l'étude de viabilité et d'établir une feuille de route en vue de la mise en place de la plateforme. Le comité fera fond sur l'expertise de ses membres et les meilleures pratiques pour concevoir et mettre en place la plateforme. De plus, il organisera des dialogues séparés avec des entreprises des secteurs de la technologie et des données privées, ainsi qu'avec des organismes régionaux et internationaux. Pour procéder à l'étude de viabilité de la plateforme et organiser sa mise en place, le comité prendra en compte des questions telles que le déploiement de l'infrastructure technologique nécessaire, l'accès aux données exclusives appartenant à des entités privées, la mise en commun des applications existantes pour l'utilisation des mégadonnées et le développement de nouvelles applications, la fourniture de services de renforcement des capacités fondés sur des cours proposés dans les secteurs public et privé, la gouvernance et les mécanismes de financement de la plateforme, notamment en ce qui concerne les fonctions et rôles respectifs de la Commission et du Groupe de travail, ou encore les liens avec d'autres initiatives lancées au niveau national, régional et mondial comme les centres nationaux de données, le « bac à sable » de la Commission économique pour l'Europe, les centres régionaux africains et les laboratoires de l'Initiative Global Pulse. On trouvera dans un document d'information concernant le présent rapport un plan initial concernant la mise en place de la plateforme, assorti d'objectifs d'étape et d'un calendrier, notamment pour les essais de la version bêta.

37. En conclusion, le Groupe de travail propose de créer, sous l'égide de la Commission, une plateforme mondiale pour la mise en commun des données, services et applications, qui prendrait la forme d'un réseau de partenaires, notamment des organismes de statistique, des instituts de recherche et des entreprises privées, qui aurait à sa disposition une infrastructure technologique adaptée pour pouvoir offrir aux producteurs de statistiques officielles des données, des applications et des services, notamment de formation et de renforcement des capacités. Cette plateforme serait mise en place progressivement, selon le plan décrit dans le document d'information.

VI. Questions à examiner

38. **La Commission de statistique est invitée à donner son avis sur les points suivants :**

a) Les progrès faits par les équipes spéciales du Groupe de travail mondial des Nations Unies sur l'utilisation des mégadonnées en statistique officielle;

b) Les résultats de la troisième Conférence internationale sur l'utilisation des mégadonnées en statistique officielle;

c) La recommandation du Groupe de travail tendant à ce qu'il soit tenu compte de la question de l'accès aux données exclusives lors des débats concernant la proposition de création d'un groupe des amis de la présidence chargé des Principes fondamentaux de la statistique officielle;

d) La proposition du Groupe de travail de créer à l'intention des spécialistes de la statistique officielle une plateforme mondiale pour la mise en commun des données, services et applications, qui serait mise en place progressivement sous l'égide de la Commission.