



Conseil économique et social

Distr. générale
2 décembre 2021
Français
Original : anglais

Commission de statistique

Cinquante-troisième session

1^{er}-4 mars 2022

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

**Questions soumises pour examen et décision :
données ouvertes**

Rapport du Groupe de travail sur les données ouvertes

Note du Secrétaire général

Le Secrétaire général a l'honneur de transmettre le rapport du Groupe de travail sur les données ouvertes, conformément à la décision 2021/224 du Conseil économique et social et à la pratique établie. Le rapport rend compte des travaux du Groupe de travail visant à fournir aux organismes nationaux de statistique des orientations sur les pratiques relatives aux données ouvertes dans l'établissement de statistiques officielles. Ces orientations portent sur des sujets tels que les données ouvertes par défaut et les meilleures pratiques en matière de concession de licences relatives aux données, l'interopérabilité des données et un modèle de maturité des données commun et les cas d'utilisation des statistiques locales ouvertes. Un document de référence contient des exemples concrets des avantages que suppose une publication plus fréquente de statistiques locales sous la forme de données ouvertes et de la façon dont de telles statistiques sont employées et associées avec d'autres données ouvertes.

La Commission de statistique est invitée à examiner et à approuver les travaux du Groupe de travail. La décision que la Commission est appelée à prendre est énoncée au paragraphe 62 du présent rapport.

* [E/CN.3/2022/1](#).



Rapport du Groupe de travail sur les données ouvertes

I. Contexte

1. À sa cinquantième session, par sa décision 50/105, la Commission de statistique a créé un groupe de travail sur les données ouvertes chargé de poursuivre les travaux sur les données ouvertes, notamment l'élaboration de directives sur l'évaluation et l'application des données ouvertes dans le contexte des statistiques officielles, et l'a prié de présenter les résultats de ses activités à sa cinquante et unième session, puis tous les deux ans.
2. À la cinquante et unième session de la Commission de statistique, en mars 2020, le Groupe de travail sur les données ouvertes a présenté le rapport sur ses travaux (voir le document publié sous la cote [E/CN.3/2020/26](#)), y compris ses orientations concernant les pratiques relatives aux données ouvertes dans l'établissement de statistiques officielles, destinées aux organismes nationaux de statistique. Pour faire suite à cette présentation, dans sa décision 51/124, la Commission a salué les travaux du Groupe de travail et approuvé le mandat proposé, y compris la définition des objectifs et des tâches, la composition et les méthodes de travail, tel qu'il figure dans l'annexe du rapport, et elle a prié le Groupe de travail de lui faire rapport à sa cinquante-troisième session.
3. Le présent rapport décrit les travaux réalisés par le Groupe de travail en 2020 et 2021 comme suite aux demandes de la Commission de statistique.

II. Programme de travail

4. Le Groupe de travail sur les données ouvertes est composé de représentants de pays, d'institutions et d'organisations internationales et d'autres partenaires. Présidé par la Nouvelle-Zélande, il est actuellement constitué des représentants de l'Australie, de la Colombie, de l'Égypte, de la Jamaïque, de la Jordanie, de la Malaisie, de Maurice, du Mexique, des Pays-Bas, de la Pologne, du Royaume-Uni de Grande-Bretagne et d'Irlande du Nord, de la Sierra Leone, de la Suède, du Suriname et de la Tunisie ainsi que de représentants des organisations et institutions internationales suivantes : le Partenariat mondial pour les données du développement durable, l'Institut international de statistique, Open Data Watch, le Partenariat statistique au service du développement au XXI^e siècle et la Banque mondiale.
5. En application de la décision 51/124 de la Commission de statistique, le Groupe de travail a structuré ses travaux en trois volets de sorte à mettre l'accent sur des activités précises :
 - a) Un volet consacré aux données ouvertes par défaut : dans ce cadre, le Groupe de travail donne des orientations concernant l'application du principe selon lequel les données sont ouvertes par défaut s'agissant des statistiques officielles, y compris les licences ouvertes, initiative de première priorité concernant l'ouverture des données, et l'adoption d'une stratégie conçue pour être ouverte¹ afin que la question de l'ouverture des données soit traitée bien en amont de l'étape de la publication de celles-ci. En outre, le Groupe de travail précise et renforce davantage le lien entre le principe d'ouverture par défaut et les Principes fondamentaux de la statistique officielle, en particulier le principe 6 concernant la protection de la

¹ Voir Paul Stone, "« Open by default » is the long game", highlights from the July 2021 Open Data Charter Implementation Working Group discussion. Disponible à l'adresse suivante : <https://medium.com/opendata-charter/spotlight-open-by-default-is-the-long-game-d3744380fd9>.

confidentialité statistique, y compris en ce qui concerne la conformité avec la législation et la réglementation internes. Les travaux menés dans le cadre de ce volet ont été dirigés conjointement par le Suriname et Open Data Watch et des contributions ont été faites par le Mexique et la Sierra Leone, le Partenariat mondial pour les données du développement durable, le Partenariat statistique au service du développement au XXI^e siècle, la Banque mondiale et la Division de statistique du Département des affaires économiques et sociales du Secrétariat ;

b) Un volet consacré à l'interopérabilité : à ce titre, le Groupe de travail poursuit l'élaboration de directives relatives à l'incorporation des données ouvertes et à l'interopérabilité dans l'ensemble des processus de production statistique. Les travaux menés dans ce cadre ont été dirigés par les représentants du Partenariat mondial pour les données du développement durable et de la Division de statistique et la Colombie, le Mexique, la Nouvelle-Zélande, les Pays-Bas, la Sierra Leone et la Banque mondiale y ont apporté des contributions ;

c) Un volet consacré aux données locales ouvertes : à ce sujet, le Groupe de travail met au point un document de référence qui constitue un guide des bonnes pratiques concernant la diffusion de statistiques locales sous la forme de données ouvertes par les organismes nationaux de statistique, l'accent étant mis sur l'équilibre entre l'ouverture et la protection de la vie privée. Les travaux réalisés dans ce cadre ont été dirigés par l'Australie et la Suède et la Malaisie, la Nouvelle-Zélande, les Pays-Bas et la Sierra Leone y ont contribué en donnant des exemples.

6. Le Groupe de travail s'est retrouvé sur Webex à raison d'une réunion toutes les six à huit semaines, environ, en vue d'échanger des renseignements concernant les pratiques nationales en matière de données ouvertes et d'examiner les progrès accomplis dans le cadre des trois volets.

III. Résultats des travaux (synthèse)

A. Le principe d'ouverture par défaut

7. Le principe d'ouverture par défaut constitue le fondement d'un ensemble de règles au titre desquelles les données des pouvoirs publics ou d'une organisation sont disponibles publiquement et conformément aux directives applicables aux données ouvertes², à quelques rares exceptions près (dans un souci de sécurité, par exemple, ou de protection de la vie privée). Au titre de ce principe, il est admis que les données des pouvoirs publics, produites grâce à des ressources publiques, sont précieuses et

² Voir, par exemple, « A review of open data practice in official statistics and their correspondence to the Fundamental Principles of Official Statistics » (Examen des pratiques relatives aux données ouvertes en statistique officielle et leur correspondance avec les Principes fondamentaux de la statistique officielle) (document de référence établi par la Division de statistique à l'intention de la Commission de statistique à sa cinquantième session et disponible à l'adresse suivante : <https://unstats.un.org/unsd/statcom/50th-session/documents/BG-Item3c-Open-Data-guidance-and-mapping-to-FPOS-E.pdf>), où est évoquée l'application pratique, par Open Data Watch, de la définition des données ouvertes pour ce qui est des aspects suivants: a) la lisibilité par des machines ; b) l'emploi de formats non propriétaires ; c) la mise à disposition de plusieurs solutions de téléchargement ; d) la mise à disposition de métadonnées qui explicitent suffisamment le contexte pour permettre de comprendre les données ; et e) les licences ouvertes. Voir également le document de référence intitulé « Guidance on the implementation of open data in national statistical offices » (Directives applicables à l'utilisation des données ouvertes dans les organismes nationaux de statistique), établi par le Groupe de travail sur les données ouvertes à l'intention de la Commission de statistique à sa cinquante et unième session et disponible à l'adresse suivante : https://unstats.un.org/unsd/statcom/51st-session/documents/BG-Item3v-Guidance_OD-E.pdf.

comptent nombre d'utilisateurs et d'utilisations et devraient par conséquent être disponibles publiquement, autrement dit par défaut³.

8. Les principes d'ouverture par défaut sont souvent appliqués dans le cadre d'une stratégie publique plus large de liberté d'information⁴, mais ils peuvent aussi être appliqués par telle ou telle institution publique, comme un organisme national de statistique. Même si les notions fondamentales qui étayent une stratégie d'ouverture par défaut sont bien assimilées, leur exécution peut être source de préoccupations pratiques et politiques dans un organisme de statistique ou d'autres autorités publiques. En novembre 2020, le Groupe de travail sur les données ouvertes a envoyé un questionnaire à plus de 150 organismes nationaux de statistique en vue d'évaluer leur niveau de compréhension des stratégies d'ouverture par défaut et de leur corrélation avec la publication de statistiques officielles. Quatre-vingt-dix pour cent des répondants estimaient que les organismes nationaux de statistique devraient suivre un modèle d'ouverture par défaut, tandis que 7 % ne savaient pas et 3 % étaient opposés à cette idée. L'étude a servi à confirmer que la plupart des répondants comprenaient généralement le principe d'ouverture par défaut mais ne savaient pas comment l'appliquer aux statistiques officielles. Tous les répondants ont exposé le principe d'ouverture par défaut en faisant référence à l'accès libre aux données, mais seulement 33 % d'entre eux ont su citer des mesures concrètes, par exemple donner des précisions juridiques sur la façon dont les données pouvaient être utilisées grâce au versement au domaine public ou à des licences ouvertes. (Une licence ouverte sert à autoriser un tiers à réutiliser les données dans une certaine mesure tandis que par le versement au domaine public, les propriétaires de données renoncent aux droits d'auteur.)

9. Les licences ouvertes ou le versement au domaine public sont des éléments importants du modèle d'ouverture par défaut, car ils constituent le fondement juridique de l'utilisation des données une fois celles-ci publiées. Par conséquent, la présente section est principalement axée sur les questions relatives au droit d'auteur et à l'octroi de licences. La suivante porte sur le domaine, plus général, de l'ouverture par défaut, qui dépasse la sphère juridique.

Le domaine général de l'ouverture par défaut

10. L'ouverture par défaut des données procure un certain nombre d'avantages aux organismes nationaux de statistique et aux pouvoirs publics en général. Elle contribue à mieux faire connaître aux utilisateurs les activités de l'organisme de statistique, ce qui, à son tour, peut renforcer la confiance et le crédit accordés aussi bien à l'organisme qu'à ses données. En outre, l'application de ce principe maximisera l'utilité que les données sont susceptibles d'avoir pour la société et l'économie. L'ouverture par défaut des données autonomise les citoyens car elle les fait participer à l'action des pouvoirs publics, ce qui rend la prise de décision plus transparente et les services publics plus efficaces et efficients.

11. Il n'y a que peu de ressources qui mettent en lumière des techniques grâce auxquelles les organismes nationaux de statistique peuvent concevoir des stratégies spécialement destinées à appliquer le principe d'ouverture par défaut aux statistiques officielles. La Charte sur les données ouvertes, par exemple, contient des orientations

³ Voir les principes énoncés dans la Charte internationale sur les données ouvertes, disponible à l'adresse suivante : <https://opendatacharter.net/principles/>.

⁴ Voir Directive (UE) 2019/1024 du Parlement européen et du Conseil du 20 juin 2019 concernant les données ouvertes et la réutilisation des informations du secteur public. Disponible à l'adresse suivante : <https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:32019L1024&from=FR>.

générales concernant la publication des données des pouvoirs publics⁵, mais ces recommandations ne s'appliquent pas toutes aux institutions chargées des processus de production statistique officiels. Ci-après figurent les mesures précises qu'il conviendrait de prendre aux fins de la création d'un modèle d'ouverture par défaut destiné à la diffusion de données statistiques.

1. Rallier le soutien de l'ensemble des pouvoirs publics en faveur des stratégies d'ouverture par défaut

12. Les organismes publics concernés devraient tenir des débats destinés à dégager une communauté de vues concernant le principe d'ouverture par défaut et à convenir d'appuyer les procédures y relatives, de sorte que les organismes nationaux de statistique soient dotés des attributions voulues et reçoivent l'appui politique nécessaires pour adopter des stratégies d'ouverture par défaut.

2. Adopter un dispositif de versement au domaine public ou une licence ouverte

13. L'ouverture des données repose fondamentalement sur l'une ou l'autre des conditions suivantes : a) les pouvoirs publics se sont engagés à donner un accès sans restriction aux données versées au domaine public ou b) une licence ouverte octroie à tous les utilisateurs un accès aux données quand celles-ci sont soumises à des droits d'auteur ou font l'objet de réserves de propriété analogues. Dans nombre de pays, les données des pouvoirs publics ne sont concernées par aucun droit d'auteur et ne peuvent par conséquent pas faire l'objet d'une licence. Dans de tels cas, il devrait être précisé sur tous les sites Web ou portails où les données sont publiées que celles-ci sont versées au domaine public. Si les données sont soumises à des droits d'auteur, il convient d'obtenir une licence ouverte. Les questions relatives à l'octroi de licences et aux droits d'auteur sont examinées ultérieurement.

3. Faire preuve de transparence pour ce qui est des ensembles de données qui ne peuvent pas être rendus publics

14. Dans le cadre d'une stratégie d'ouverture par défaut, on publie une « liste négative » des ensembles de données qui ne sont pas en libre accès plutôt qu'une liste des ensembles de données qui sont en accès public. Parmi les motifs recevables pour ne pas donner librement accès à un ensemble de données figurent la nécessité de protéger la sûreté nationale, la nécessité de préserver la confidentialité des renseignements personnels et la nécessité de prévenir tout préjudice découlant de la publication de certains renseignements. Si certaines données ne peuvent pas être publiées ou si elles ne peuvent pas l'être dans un certain délai, il convient de l'énoncer clairement dans la stratégie de diffusion de l'organisme national de statistique et de l'annoncer en ligne sur le site où l'on peut généralement s'attendre à trouver ces données. En prévoyant la publication de la liste des exceptions, les stratégies d'ouverture par défaut témoignent d'un engagement vis-à-vis de la transparence des pouvoirs publics et d'une utilisation efficace des ressources publiques.

4. Instaurer une procédure de demande de données transparente

15. Il est possible que les utilisateurs demandent la publication de données qui ne sont pas disponibles actuellement dans les bases de données publiques. Par conséquent, les organismes nationaux de statistique devraient mettre au point un système qui contient des instructions claires et faciles à comprendre pour ce qui est de demander accès à des ensembles de données particuliers. Il est acceptable de

⁵ Voir la Charte internationale sur les données ouvertes, disponible à l'adresse suivante : <https://opendatacharter.net/principles-fr/>.

facturer le coût marginal d'établissement de demandes spéciales de données. Tout rejet d'une demande doit être motivé et toute éventuelle procédure d'appel connexe doit être expliquée. Il convient de traiter les demandes dans un délai raisonnable.

5. Établir des procédures claires d'accès aux microdonnées et de publication de celles-ci

16. La publication de microdonnées contient des étapes supplémentaires en vue d'empêcher la publication de renseignements confidentiels ou de données personnelles. Les méthodes employées peuvent être complexes et nécessiter des ressources supplémentaires. Par ailleurs, il faut toujours trouver un équilibre entre le niveau d'utilité des microdonnées et le degré de contrôle de la divulgation qu'il est possible d'atteindre. Étant donné qu'il subsiste un risque que les données soient désanonymisées, les organismes nationaux de statistique devraient établir des orientations claires sur les modalités de publication, de consultation et d'utilisation des microdonnées par les utilisateurs. Une fois les risques de divulgation dûment pris en considération, les microdonnées devraient être publiées sous forme de fichiers d'utilisation publique et soumises à des licences (souvent plus restrictives dans ce cadre) et des mécanismes de responsabilisation devraient être en place pour prévenir l'usage impropre et la divulgation de renseignements confidentiels.

6. Faire paraître un calendrier de publication des données

17. Un calendrier de diffusion d'ensembles de données renseigne les utilisateurs quant à la date prévue de mise à disposition en ligne de données. Il s'agit d'un élément important de la transparence. Un tel calendrier est obligatoire quand des ensembles de données sont publiés au titre de la Norme spéciale de diffusion des données du Fonds monétaire international (FMI)⁶.

7. Renforcer les capacités du personnel de produire des données ouvertes

18. Le personnel peut renforcer ses capacités s'il a accès à des programmes de formation, à des ateliers ou à des webinaires axés sur les questions relatives aux données ouvertes. De surcroît, les organismes nationaux de statistique devraient créer un comité interne composé de représentants de toutes les unités qui se réunirait régulièrement pour examiner les stratégies de diffusion des données.

8. Renforcer les capacités des utilisateurs externes d'utiliser et de comprendre les données ouvertes

19. Il importe d'investir des ressources et de collaborer avec d'autres organisations en vue de renforcer l'aptitude des utilisateurs à exploiter et à traiter les données et de mieux faire connaître au public les stratégies d'ouverture de données de l'organisme national de statistique. À cette fin, il est possible d'élaborer des campagnes d'information, de fournir du matériel pédagogique et d'organiser des activités ponctuelles ou périodiques à l'intention de certains groupes d'utilisateurs (par exemple des conférences de presse et des exposés universitaires), en vue d'expliquer des notions et des conclusions statistiques complexes et de donner des explications détaillées sur les ressources dont disposent les utilisateurs pour accéder aux statistiques officielles et les utiliser.

Données soumises à des droits d'auteur et à des licences

20. La naissance du modèle d'ouverture par défaut est révélatrice d'un changement de mentalité des pouvoirs publics vis-à-vis des données, car ce modèle ne fait plus

⁶ Voir <https://dsbb.imf.org>.

porter le débat sur les données qui devraient être ouvertes, mais sur celles qui ne devraient pas l'être. Pour appuyer ce changement, l'organisme national de statistique doit tout d'abord avoir l'autorité juridique voulue et élaborer des normes pour déterminer quelles données ne doivent pas être publiées (cette étape est l'établissement de la liste négative) et ensuite adopter une politique de versement au domaine public ou de licence ouverte.

21. Le nœud de la question est de savoir si les données des pouvoirs publics sont protégées par des droits d'auteur. On entend par droit d'auteur le droit légal de produire des œuvres originales, de les soumettre à des licences de réutilisation ou de donner à un tiers physique ou moral des droits sur elles⁷. Les données des pouvoirs publics, définies comme un ensemble d'éléments factuels, sont généralement considérées comme n'étant pas protégées par des droits d'auteur, mais leur compilation ou représentation stylisée peut l'être⁸ dans certains pays. La législation nationale en matière de droits d'auteur définit les types de produits qui peuvent être protégés par de tels droits. Si un produit (tel que les données produites par un organisme national de statistique) est protégé par des droits d'auteur, son propriétaire peut limiter l'utilisation que des tiers peuvent en faire ; par conséquent, il est nécessaire de disposer d'une licence qui lève les restrictions. Si la législation d'un pays en matière de droits d'auteur prévoit que les données des pouvoirs publics⁹ ne sont concernées par aucun droit d'auteur, celles-ci devraient être versées au domaine public¹⁰ (à l'exception des données personnelles, telles que les microdonnées issues d'enquêtes, dont l'accès est généralement restreint dans le cadre de la législation ou de la réglementation internes).

22. La législation relative aux droits d'auteur varie en fonction des pays, mais il existe en la matière un modèle international dont la plupart des pays sont signataires, à savoir la Convention de Berne pour la protection des œuvres littéraires et artistiques¹¹. Cet accord prévoit de protéger les « œuvres originales » par des droits d'auteur, mais ne traite pas explicitement la question du statut des données des pouvoirs publics ou des renseignements factuels pour ce qui est des droits d'auteur. Dans nombre de pays, la législation relative aux droits d'auteur prévoit un « seuil d'originalité »¹², expression qui a été interprétée comme signifiant que les données des pouvoirs publics ou les ensembles de données factuelles ne sont pas protégés par des droits d'auteur. On considère que la Convention de Berne prévoit un seuil d'originalité¹³, même si elle ne contient pas cette expression.

23. Même s'il est probable que la plupart des données des pouvoirs publics ne sont pas protégées par des droits d'auteur et doivent être publiées au moyen d'un dispositif de versement au domaine public, les organismes nationaux de statistique exploitent des licences ouvertes, telles que la Licence publique Creative Commons Attribution 4.0 International¹⁴ ou des licences qu'ils ont conçues eux-mêmes. D'après l'évaluation Open Data Inventory (ODIN)¹⁵, entre 2018 et 2020, le nombre

⁷ Voir Shaida Badiee *et al.*, « Open data for official statistics: history, principles, and implementation ». *Statistical Journal of the IAOS*, vol. 37, n° 1 (22 mars 2021), p. 139 à 159.

⁸ Voir Catherine Doldirina *et al.*, *Legal Approaches for Open Access to Research Data*, 1^{er} avril 2018, p. 17. Disponible à l'adresse suivante : osf.io/dv6tc.

⁹ Définies comme un « ensemble d'éléments factuels » aux fins du présent rapport.

¹⁰ Voir <https://theodi.org/article/publishers-guide-to-open-data-licensing/>.

¹¹ Voir <https://www.wipo.int/treaties/fr/ip/berne/>.

¹² Voir https://repository.uchastings.edu/cgi/viewcontent.cgi?article=1004&context=hastings_business_law_journal.

¹³ Voir Doldirina *et al.*, « Legal approaches », note de bas de page 54.

¹⁴ Voir <https://creativecommons.org/licenses/by/4.0/deed.fr>.

¹⁵ Voir Open Data Watch, Open Data Inventory annual report 2020/21, disponible à l'adresse suivante : <https://odin.opendatawatch.com/Report/annualReport2020>.

d'organismes nationaux de statistique qui exploitent une licence ouverte (aussi désignée sous le terme « conditions d'utilisation »)¹⁶ est passé de 44 à 58. En outre, 43 pays ont adopté une licence en matière de données qui respecte plusieurs normes relatives aux données ouvertes. Les États-Unis d'Amérique¹⁷, Oman¹⁸, le Luxembourg¹⁹ et la Suède²⁰ annoncent sur le site Web où sont publiées leurs données statistiques que celles-ci sont versées au domaine public. Les autres organismes nationaux de statistique ne disposent d'aucune licence ou utilisent une licence restrictive. La Suède est un exemple unique, car elle publie toutes ses données dans le cadre d'une licence Creative Commons 0 (licence CC0)²¹, mais fait paraître l'ensemble de ses publications et de ses œuvres visuelles au titre d'une licence CC BY 4.0, de sorte à conserver le droit de paternité sur ces produits. Cette association de licences permet de concilier les vues des utilisateurs, qui exigent une plus grande ouverture, et celles des organismes nationaux de statistique, qui ont besoin de disposer d'un droit de paternité pour mieux suivre les réutilisations des données.

24. Nombre d'organisations internationales sont peu disposées à adopter les licences ouvertes. Il ressort d'un examen réalisé en septembre 2019 par le Comité de coordination des activités de statistique que seulement 14 des 45 membres de ce comité détenaient une licence ouverte.

Coordination avec les cadres juridiques nationaux

25. Il est devenu de plus en plus fréquent que les organismes nationaux de statistique publient des données dans le cadre d'une licence ouverte ou versent les données au domaine public, mais on sous-estime souvent la nécessité de mettre en œuvre cette pratique en coordination avec d'autres lois et stratégies internes. S'il n'est pas tenu compte de ce besoin, d'autres lois ou stratégies internes qui prévalent peuvent favoriser la poursuite de la stratégie de restriction de l'ouverture des données statistiques ou, à tout le moins, semer la confusion parmi les utilisateurs.

26. Dans l'idéal, les libellés seront harmonisés dans le cadre d'une révision de la loi ou de la stratégie de substitution afin d'avancer sur la voie d'un système national de données intégré²². Si une telle entreprise se révèle impossible, il convient alors d'harmoniser les libellés employés dans la licence ou le dispositif de versement au domaine public. L'un des principaux avantages que présentent les licences ouvertes et les dispositifs de versement au domaine public est qu'ils donnent aux utilisateurs des précisions juridiques concernant la façon dont ils peuvent ou ne peuvent pas utiliser les données, ce qui peut à son tour se traduire par une plus grande utilisation des données. Cependant, si ces licences ou ces dispositifs sont foncièrement incompatibles avec d'autres textes de loi, les utilisateurs n'en tireront aucun avantage et peuvent hésiter à réutiliser des données.

27. Afin de mieux comprendre si les organismes nationaux de statistique avaient examiné leur cadre juridique national général, le Groupe de travail leur a distribué un deuxième questionnaire destiné à les interroger au sujet du cadre juridique et des

¹⁶ Il importe de distinguer les conditions d'utilisation des licences en matière de données. Comme leur nom l'indique, les licences en matière de données concernent précisément les données, alors que les conditions d'utilisation définissent les modalités d'utilisation de contenus de sites Web de façon plus générale.

¹⁷ Voir www.data.gov/privacy-policy#data_policy.

¹⁸ Voir <https://data.gov.om/legal/termsofuse>.

¹⁹ Voir <https://statistiques.public.lu/fr/support/notice/index.html#copyright>.

²⁰ Voir www.scb.se/en/services/open-data-api/.

²¹ Voir <https://creativecommons.org/share-your-work/public-domain/cc0/>.

²² Voir <https://wdr2021.worldbank.org/fr/>.

stratégies du pays concernant les licences en matière de données ou le versement au domaine public. Dans ce cadre, ces organismes ont été questionnés quant aux autres lois et stratégies qu'ils avaient consultées avant l'adoption d'une licence ouverte. Près de 47 % des répondants ont déclaré qu'ils avaient examiné leur loi nationale sur la statistique, tandis que 13 % ont fait savoir qu'ils s'étaient référés à la licence générale en matière de données exploitée par les pouvoirs publics. Seul un organisme national de statistique a fait savoir qu'il avait étudié toutes les lois applicables (c'est-à-dire les lois relatives aux statistiques, à la licence applicable aux données des pouvoirs publics, au droit d'auteur, à la protection de la vie privée, et au droit à l'information). Ces réponses peuvent témoigner à la fois d'une incompréhension des retentissements des autres lois sur les licences en matière de données et de la nécessité d'investir davantage de ressources pour aider les organismes nationaux de statistique et d'autres producteurs nationaux de statistiques à se repérer dans le paysage juridique.

28. Dans certains pays, le libellé des lois et stratégies est compatible avec celui des licences et des dispositifs de versement au domaine public. Néanmoins, en vue de prévenir toute incompatibilité de cet ordre, il importe dès les premières étapes des procédures d'adoption ou de création d'une licence ouverte ou d'un dispositif de versement au domaine public de procéder à un examen juridique de l'ensemble des lois applicables. On retrouvera ci-après une énumération des types de lois communément adoptées dans nombre de pays et dont le libellé peut être incompatible avec celui d'une licence en matière de données.

Législation relative aux droits d'auteur. Comme exposé de manière détaillée ci-dessus, la législation relative aux droits d'auteur devrait être examinée d'entrée de jeu afin de déterminer si les données des pouvoirs publics sont protégées par des droits d'auteur. Si elles ne le sont pas, il convient d'adopter un dispositif de versement au domaine public plutôt que d'avoir recours à une licence ;

Licences applicables aux données des pouvoirs publics. Dans certains pays, tels le Royaume-Uni²³, l'Allemagne²⁴, le Canada²⁵ et Oman²⁶, les pouvoirs publics ont adopté une licence applicable dans tous les organismes publics. Dans certains cas, ces organismes sont tenus d'utiliser cette licence pour leurs données, tandis que dans d'autres, l'utilisation de la licence est recommandée ou la licence peut constituer un ensemble de directives que doivent suivre les organismes publics quand ils adoptent leur propre licence. La Nouvelle-Zélande²⁷, par exemple, prévoit un ensemble de licences et de directives dont l'emploi par les administrations publiques est autorisé. Aux États-Unis²⁸, une liste des données versées au domaine public est publiée, ainsi qu'une liste des licences ouvertes applicables aux données provenant de tierces parties (les données des pouvoirs publics ne sont concernées par aucun droit d'auteur dans ce pays). Si les pouvoirs publics ont déjà mis au point une licence ou un ensemble de directives applicables à l'octroi de licences ou au versement au domaine public, les organismes nationaux de statistique devraient examiner cette documentation afin de vérifier qu'ils ont qualité pour créer ou adopter leur propre licence ou leurs propres directives en matière de dispositif de versement au domaine public.

Lois internes relatives à la statistique. Une loi relative à la statistique devrait définir la portée et l'organisation du système statistique national, instaurer un mandat de recueil de données, imposer la confidentialité statistique, établir les normes de

²³ Voir www.nationalarchives.gov.uk/doc/open-government-licence/version/3/.

²⁴ Voir www.govdata.de/dl-de/by-2-0.

²⁵ Voir <https://ouvert.canada.ca/fr/licence-du-gouvernement-ouvert-canada>.

²⁶ Voir [www.moheri.gov.om/userupload/Policy/Government %20Open %20Data %20Policy.pdf](http://www.moheri.gov.om/userupload/Policy/Government%20Open%20Data%20Policy.pdf).

²⁷ Voir www.data.govt.nz/toolkit/policies/nzgoal/.

²⁸ Voir <https://resources.data.gov/open-licenses/>.

qualité des statistiques officielles et autoriser la diffusion des statistiques. Les renseignements relatifs à l'accès aux données ou aux stratégies de diffusion peuvent comprendre des restrictions concernant les modalités d'utilisation ou de récupération des données, ainsi que des directives relatives à l'obtention des permissions nécessaires. Le libellé contenu dans la loi susmentionnée peut être compatible ou non avec celui d'une licence ouverte ou d'un dispositif de versement au domaine public. S'il existe dans ladite loi des restrictions qui empêchent l'ouverture des données, la licence en matière de données ou le dispositif de versement au domaine public doit remédier à cette contradiction. Si le libellé d'une telle loi appuie la licence ou le dispositif de versement au domaine public, aucune mesure ne sera nécessaire²⁹.

Lois relatives au droit à l'information. Aussi désignée sous le terme de « loi relative à la liberté d'information », la loi relative au droit à l'information instaure les procédures qui permettent au public de demander des renseignements détenus par le secteur public et peut établir une présomption d'accès par défaut. Dans la plupart des cas, une telle loi devrait appuyer et non pas limiter les licences ouvertes et les dispositifs de versement au domaine public. Néanmoins, en vue de démontrer leur respect des pratiques en matière de transparence et de mettre en avant leurs engagements dans ce domaine, les organismes nationaux de statistique peuvent mentionner ces lois dans leur licence en matière de données ou ailleurs.

Lois relatives à la protection et à la confidentialité des données. Ces lois régissent le recueil, l'actualisation, l'utilisation et la diffusion des renseignements relatifs aux personnes collectés par des organismes publics. Elles peuvent aussi s'appliquer aux données regroupées par le secteur privé. Les organismes nationaux de statistique ne devraient jamais publier au titre d'une licence ouverte les renseignements qui sont susceptibles d'être sensibles ou verser ces informations au domaine public. Dans un souci de protection des renseignements sensibles, même les microdonnées anonymisées issues d'études peuvent être diffusées dans des conditions plus restrictives que celles établies dans une licence ouverte ou dans un dispositif de versement au domaine public.

Remédier aux préoccupations des producteurs de statistiques

29. À plus d'un titre, les producteurs de statistiques hésitent à publier des données dans le cadre d'un dispositif de versement au domaine public ou d'une licence ouverte s'inscrivant dans un modèle d'ouverture par défaut. Cette réticence provient des nombreuses difficultés auxquelles se heurtent généralement les organismes nationaux de statistique quand ils accordent un accès ouvert aux données qui ont été examinées en détail par d'autres³⁰. Certaines des préoccupations à ce sujet ainsi que les solutions qu'il est possible d'y apporter sans nuire à l'ouverture des données sont analysées ci-dessous.

30. La part de 10 % de répondants qui ne souhaitaient pas adopter une stratégie d'ouverture par défaut dans le cadre de la publication de statistiques officielles étaient principalement préoccupés par la protection de la confidentialité des données et la prévention du risque d'usages abusifs des données, entre autres choses.

²⁹ Voir, par exemple, Commission économique pour l'Europe, *Guidance on Modernizing Statistical Legislation* (Geneva, 2018), chap. 8, sect. A. Disponible à l'adresse suivante : <https://unece.org/fileadmin/DAM/stats/publications/2018/ECECESSTAT20183.pdf>.

³⁰ Voir Banque mondiale, « Open data challenges and opportunities for national statistical offices » (Washington, DC, 1^{er} juillet 2014). Disponible à l'adresse suivante : <https://documents.banquemonde.org/fr/publication/documents-reports/documentdetail/740381468128389452/open-data-challenges-and-opportunities-for-national-statistical-offices>.

Prévention du risque d'usages abusifs ou d'erreur d'interprétation des données

31. Bon nombre d'organismes nationaux de statistique craignent que leurs données soient mal interprétées ou employées pour faire des déclarations dans le domaine de la publicité ou affirmer une contrevérité dans des débats politiques. Certains se sont attelés à prévenir le risque d'usages abusifs de leurs données dans le cadre de leurs licences en la matière. En 2020, Open Data Watch a conclu dans son rapport ODIN que 21 pays avaient adopté des licences en matière de données ou instauré des conditions d'utilisation qui interdisaient l'usage abusif. Il est problématique que, dans tous les cas, la définition de l'usage abusif ne soit pas établie et par conséquent, qu'elle soit ouverte à interprétation, car une telle décision pourrait être considérée comme favorable à la censure des opinions.

32. Au titre du principe 4 des Principes fondamentaux de la statistique officielle, il est établi que les organismes nationaux responsables de la statistique « ont le droit de faire des observations sur les interprétations erronées et les usages abusifs des statistiques »³¹. Néanmoins, les directives régissant l'application des Principes fondamentaux³² ne laissent pas entendre que ce point devrait être traité dans le cadre de licences en matière de données ou d'autres instruments employés pour restreindre l'utilisation de données. Au lieu de cela, elles mettent l'accent sur l'élaboration d'une stratégie de réponse aux cas d'usages abusifs et les activités de collaboration avec les utilisateurs destinées à renforcer leurs capacités d'appréhender et d'utiliser les données, qui peuvent comprendre des conférences de presse visant à expliquer à des journalistes les principales conclusions présentées dans de nouveaux rapports et la publication de supports non techniques qui permettent d'expliquer des résultats statistiques à des utilisateurs moins chevronnés. Dans les cas où l'usage abusif des données est nuisible, le droit interne devrait prévoir de réparer le préjudice subi en faisant intervenir des mécanismes de responsabilisation adaptés et non en employant d'autres statistiques.

33. Les lois internes relatives à la statistique peuvent prévoir expressément d'autres formes d'usage abusif, telles que la désanonymisation des données visant à remonter aux personnes ou aux organismes dont elles émanent. Par ailleurs, les lois applicables aux organismes de statistique peuvent interdire la publication prématurée ou « la fuite » des informations commerciales sensibles. Ces lois et d'autres qui visent à protéger les fournisseurs de données ou à garantir un accès équitable aux informations statistiques ne sont pas incompatibles avec les stratégies d'ouverture de données.

Protection de la réputation ou de l'intégrité de l'organisme national de statistique (droits moraux)

34. Les organismes nationaux de statistiques se préoccupent de la prévention de l'usage abusif et, dans cette même perspective, craignent fréquemment qu'une déclaration controversée ou même volontairement trompeuse dans laquelle leurs données sont citées leur soit, à tort, attribuée, ou appréhendent d'être associés à une telle déclaration, qui pourrait nuire à leur crédibilité ou rejaillir négativement sur la confiance du public en leurs statistiques. Certains organismes nationaux de statistique exploitent des licences en matière de données ou emploient d'autres moyens d'intervention destinés à interdire ce type d'usage en application de leurs droits moraux. Cette méthode présente les mêmes inconvénients que celles qui servent à prévenir l'usage abusif : elle peut être employée pour sanctionner les opinions peu populaires ou contraindre les utilisateurs de données à se censurer. Si les données des pouvoirs publics ne sont pas protégées par des droits d'auteur au titre du droit interne,

³¹ Voir <https://unstats.un.org/unsd/dnss/gp/FP-New-F.pdf>.

³² Voir https://unstats.un.org/unsd/dnss/gp/Implementation_Guidelines_FINAL_without_edit.pdf.

les dispositifs de versement au domaine public, tels que la licence CC0 et la licence d'Open Data Commons appelée Public Domain Dedication and License (PDDL)³³, peuvent servir à contourner les droits moraux. Si les organismes nationaux de statistique adoptent une licence ouverte en remplacement, ils peuvent y inclure des clauses de non-approbation pour remédier à ces préoccupations sans porter préjudice à l'ouverture des données.

Protection de la vie privée

35. Les organismes nationaux de statistique se soucient, à juste titre, de protéger la confidentialité des données que les citoyens leur communiquent. Même si la question de l'atteinte à la vie privée ne se pose pas pour ce qui est de la plupart des données agrégées fournies par les organismes nationaux de statistique, elle se pose dans le cas d'autres produits statistiques, tels que les microdonnées. Dans le cadre du modèle d'ouverture par défaut de la diffusion des statistiques, la confidentialité des informations personnelles est reconnue et soutenue. Toute stratégie relative à la publication des données doit respecter le sixième principe fondamental de la statistique officielle. Les microdonnées brutes ne devraient pas être publiées dans le cadre de la même licence ouverte ou du même dispositif de versement au domaine public que les produits statistiques dont la confidentialité pose peu de problèmes. Dans le cadre de leur stratégie d'ouverture par défaut, les organismes nationaux de statistique devraient définir expressément les types de données (ou de produits) qui sont disponibles au titre d'une licence ouverte. La plupart de ces organismes publient des microdonnées dans un catalogue distinct qui se trouve sur un autre site Web que le leur, où les conditions et la licence associées à la publication de ces données apparaissent clairement. Il est fréquent que la publication de microdonnées soit assortie des conditions d'utilisation qui défendent d'essayer de désanonymiser les données.

Peur de perdre des recettes

36. Étant donné que les données ouvertes doivent être gratuites, nombre d'organismes nationaux de statistique craignent que l'ouverture de leurs données entraîne une perte des recettes qu'ils tiraient des ventes de leurs publications statistiques ou d'autres données. Cependant, l'utilisation accrue des données et le développement de l'innovation qu'entraîne l'ouverture des données peuvent compenser la perte de recettes éventuellement essayée dans un premier temps, en particulier si l'on tient compte des coûts de transaction qui accompagnent les régimes de licence restrictifs³⁴. Tant que les données issues de publications sont librement disponibles en ligne, les organismes nationaux de statistique peuvent continuer de facturer des frais minimes pour rembourser les coûts des publications papier. En outre, ils peuvent facturer les coûts marginaux de prestation de services spéciaux en dehors du cadre de la fourniture classique de données.

37. Les licences et les dispositifs de versement au domaine public les plus utiles sont concis, clairs, adoptés par toutes les administrations publiques qui produisent des données, et s'inscrivent dans le cadre d'un régime juridique qui vient les renforcer.

³³ Voir <https://opendatacommons.org/licenses/pddl/>.

³⁴ Voir <https://theodi.org/wp-content/uploads/2016/04/The-economic-value-of-data-assets-under-alternative-policy-regimes.pdf>.

B. L'interopérabilité : la clef de l'ouverture des statistiques

Contexte

38. À sa cinquantième session, en mars 2019, la Commission de statistique a été invitée à exprimer ses vues concernant le guide intitulé « Data interoperability: a practitioner's guide to joining up data in the development sector » (Interopérabilité des données : guide pratique aux fins de l'unification des données relatives au développement)³⁵. Ces travaux ont été présentés par le Groupe des Amis de la présidence sur les Principes fondamentaux de la statistique officielle et les données ouvertes dans le cadre de son document de référence, intitulé « A review of open data practices in official statistics and their correspondence to the Fundamental Principles of Official Statistics »³⁶.

39. Dans sa décision 50/105, la Commission a approuvé les directives sur l'interopérabilité des données et, consciente de l'importance pour les pays de pouvoir accéder aux outils permettant l'interopérabilité des données, a prié le Groupe de poursuivre l'élaboration des directives.

40. En application de ce mandat, le Partenariat mondial pour les données du développement durable, en collaboration avec le Secrétariat, a établi le document « Connecting data communities: introducing the joined-up data maturity assessment » (connecter les communautés de données : présentation de l'évaluation conjointe de la maturité des données)³⁷ (le modèle de maturité) avec l'appui du groupe Collaborative on SDG Data Interoperability³⁸. Ledit groupe ayant atteint ses objectifs principaux, il a été dissous en novembre 2020 ; cependant, plusieurs membres d'origine continuent d'intervenir dans le cadre des travaux du Groupe de travail sur les données ouvertes et ont diffusé pour avis le document contenant le modèle de maturité dans leurs organisations respectives.

41. Les avis reçus des organismes nationaux de statistique étaient dans l'ensemble positifs, ce qui souligne qu'il importe que lesdits organismes disposent d'un outil stratégique conçu pour contribuer à favoriser une meilleure interopérabilité des données et la publication de données statistiques ouvertes.

L'importance de l'interopérabilité dans l'ouverture des statistiques

42. Comme énoncé dans le guide, l'interopérabilité englobe la capacité de regrouper des ensembles de données hétérogènes sans perdre de sens et de les intégrer dans des plateformes, systèmes et outils différents. Elle est essentielle pour favoriser une publication à plus grande échelle des données statistiques ouvertes et permettre aux organismes nationaux de statistique de compiler et d'intégrer les données nécessaires pour suivre et mesurer les progrès accomplis sur la voie de la réalisation des objectifs de développement durable conformément au cadre mondial d'indicateurs relatifs aux objectifs de développement durable.

43. Les capacités d'intégrer les données statistiques à d'autres ensembles de données géospatiales, d'utiliser des données anonymisées de géolocalisation afin de suivre les mouvements de la population pendant la pandémie de maladie à coronavirus 2019 (COVID-19) et de mieux incorporer plusieurs ensembles de données

³⁵ http://www.data4sdgs.org/sites/default/files/services_files/Interoperability - A practitioner's guide to joining-up data in the development sector.pdf.

³⁶ Disponible à l'adresse suivante : <https://unstats.un.org/unsd/statcom/50th-session/documents/BG-Item3c-Open-Data-guidance-and-mapping-to-FPOS-E.pdf>.

³⁷ Disponible à l'adresse suivante : http://www.data4sdgs.org/sites/default/files/file_uploads/Joined_Up_Data_Maturity_Assessment_draft5.pdf.

³⁸ www.data4sdgs.org/initiatives/data-interoperability-collaborative.

administratives dans les procédures statistiques reposent toutes sur l'interopérabilité. En outre, l'interopérabilité est un catalyseur essentiel des stratégies et pratiques en matière de données ouvertes associées à la publication de données statistiques. Afin que les ensembles de données s'inscrivent dans des architectures interopérables et emploient des classifications communes et que les champs de métadonnées soient alignés et normalisés et, si possible, contiennent des libellés communs, il convient de tenir compte de l'interopérabilité à chaque étape de la production de statistiques.

44. Faisant fond sur le cadre conceptuel établi par le groupe Collaborative on SDG Data Interoperability et publié dans le guide pratique, le modèle de maturité est l'aboutissement de six ans de travail destinés à conjuguer les différentes façons dont les acteurs qui s'emploient à obtenir des résultats en matière de développement appréhendent l'interopérabilité dans la pratique. À ce sujet, le modèle de maturité³⁹ constitue un outil très efficace que les organismes nationaux de statistique peuvent utiliser dans les cas suivants, entre autres :

- a) La production d'un nouveau cadre ou d'une nouvelle stratégie de gestion des données ;
- b) L'examen stratégique ou l'actualisation d'un cadre de gestion des données ou des modalités de fonctionnement en place, depuis l'élaboration d'une stratégie nationale de production de statistiques à l'adoption du Modèle générique du processus de production statistique⁴⁰ ;
- c) Une évaluation plus large de la maturité de la gouvernance des données ;
- d) L'augmentation de la portée des systèmes statistiques nationaux, de sorte qu'ils comptent davantage d'acteurs et publient plus largement leurs données ; dans ce cadre, le modèle a une valeur stratégique.

45. Il convient de noter que le modèle de maturité est axé sur le volet stratégique de la gestion de l'organisation et ne vise pas à être une source de conseils concernant des fonctions statistiques spécifiques.

Mesurer les progrès accomplis sur la voie de l'interopérabilité des données statistiques

46. Le modèle de maturité compte trois composantes : les couches d'interopérabilité, les dimensions de maturité et les niveaux de maturité. La première composante est constituée des quatre couches d'interopérabilité établies dans le document d'orientation et le guide, comme conçues à l'origine par Palfrey et Gasser (2012)⁴¹, à savoir : a) l'interopérabilité organisationnelle, b) l'interopérabilité humaine, c) l'interopérabilité des données et d) l'interopérabilité technologique.

47. Les quatre couches comptent en tout 19 dimensions, ou fonctions de gestion des données, rattachées à chacune d'entre elles. Le modèle de maturité recense les caractéristiques qui sont propres à chaque dimension et qui témoignent du niveau de maturité d'une organisation, puis les classe par ordre croissant de maturité dans les catégories « *non défini* », « *nouveau* », « *en cours d'apprentissage* », « *en cours de renforcement* » ou « *en cours de consolidation* ». Ces niveaux de maturité ne s'excluent pas mutuellement : il est tout à fait possible que les progrès d'un utilisateur

³⁹ Le modèle d'évaluation de maturité des données commun est disponible à l'adresse suivante : https://www.data4sdgs.org/sites/default/files/file_uploads/Interoperability_Maturity_Model_Module.pdf.

⁴⁰ Voir <https://statswiki.unece.org/display/GSBPM/>.

⁴¹ John Palfrey et Urs Gasser, *Interop: The Promise and Perils of Highly Interconnected Systems* (New York, Basic Books, 2012).

relèvent de la catégorie « *en cours de consolidation* » dans une dimension et uniquement de la catégorie « *nouveau* » dans d'autres.

Représentation schématique des différentes dimensions de l'interopérabilité contenues dans le modèle de maturité

<i>Couches d'interopérabilité</i>	<i>Dimensions</i>
1. Interopérabilité organisationnelle	Objectifs stratégiques Direction et gestion Contrôle et responsabilisation Conformité juridique Éthique des données Achats Liens vers des écosystèmes de données plus larges
2. Interopérabilité humaine	Intendance des données Respect de la vie privée et de la confidentialité Connaissances et compétences du personnel Communication interne et externe Adaptabilité
3. Interopérabilité des données	Capacités de modélisation des données et des métadonnées Capacités d'organisation et de classement des données Accès, ouverture et diffusion des données Analyse et automatisation des données Protection des données
4. Interopérabilité technologique	Infrastructures numériques Cybersécurité et intervention en cas d'infraction

48. Le modèle de maturité est accompagné de directives dans lesquelles est expliqué l'importance de chaque dimension. Les organismes nationaux de statistique peuvent les utiliser dans leur intégralité ou avoir recours à certaines parties seulement, selon ce qui est le plus utile compte tenu du contexte qui leur est propre. À ce jour, aucune directive n'a été établie pour aider les organismes nationaux de statistique à évaluer leur « résultat » s'agissant des 19 dimensions. Ce choix est délibéré, car le modèle de maturité est conçu pour mieux mesurer les progrès et appuyer la prise de décision en fonction d'un contexte, pas pour classer les organisations dont les dispositifs d'interopérabilité sont plus avancés par rapport à celles qui ont moins progressé sur ce plan.

49. Depuis la cinquantième session de la Commission de statistique, plusieurs membres du Groupe de travail sur les données ouvertes ont formulé des observations concernant différentes composantes du modèle maturité et ont fait des suggestions en vue de l'améliorer à terme. Les avis suivants ont été donnés en retour du document :

- Il a été admis que le modèle de maturité et les directives connexes peuvent être d'un grand secours pour appuyer la planification stratégique des organismes nationaux de statistique et que l'inclusion des dimensions relatives à l'interopérabilité qui ont trait aux aspects organisationnels et humains peuvent contribuer à distinguer ce modèle d'autres qui sont principalement axés sur l'interopérabilité technologique et l'interopérabilité des données.
- Il a été suggéré d'améliorer le modèle de maturité en y incorporant des questions d'orientation qui non seulement précisent davantage la possibilité que l'outil soit employé différemment par différents organismes nationaux de statistique mais aussi déterminent la façon dont ces organismes peuvent attribuer une note à leurs progrès et à leur maturité s'ils le souhaitaient.
- Il a été suggéré d'incorporer à la note d'orientation un encadré dans lequel est illustrée la façon dont les diverses dimensions de l'interopérabilité influent sur les différents éléments du modèle générique du processus de production statistique et sont reliées à eux.
- Il a été suggéré d'ajouter des « mécanismes de retour d'informations entre les clients et les utilisateurs » à titre de dimension organisationnelle, étant donné qu'il faut faire en sorte que les nécessités des utilisateurs fassent l'objet d'échanges et trouvent une réponse.
- Il a été suggéré que des « services Web » soient ajoutés à la couche technologique à titre de dimension particulière, compte tenu de la dépendance croissante vis-à-vis des services de stockage en nuage, de traitement et de nettoyage des données et qu'il fallait déduire de quelle façon la délégation de ces fonctions influe sur l'interopérabilité.

50. À ce sujet, les organisations statistiques sont encouragées à adapter, peaufiner et adopter le modèle de maturité afin de l'employer dans leur propre contexte et, à l'issue de sa mise en service, d'échanger des idées avec des pairs, des partenaires et le grand public dans le cadre de la procédure d'amélioration continue.

C. Statistiques locales sous la forme de données ouvertes : une approche centrée sur l'utilisateur

51. À la cinquante et unième session de la Commission de statistique, en mars 2020, le Groupe de travail sur les données ouvertes a distribué un document de référence sur les statistiques locales sous la forme de données ouvertes⁴². Celui-ci contenait une description des avantages d'une publication accrue des statistiques locales sous la forme de données ouvertes et a défini les exigences élémentaires concernant les zones géographiques, la teneur, la confidentialité et la visualisation. Il a été souligné que les organismes nationaux de statistique pouvaient jouer un rôle important en tant que fournisseurs solides de statistiques locales s'ils étaient incorporés aux ressources dont disposait un pays pour produire des données ouvertes. Néanmoins, le document de référence ne contenait aucun exemple concret d'utilisation des statistiques locales et d'association de celles-ci avec d'autres types de données ouvertes.

⁴² Disponible à l'adresse suivante : https://unstats.un.org/unsd/statcom/51st-session/documents/BG-Item3v-Local-level_OD-E.pdf.

52. À sa cinquante-troisième session, la Commission de statistique a pris connaissance d'un deuxième document de référence, intitulé « Local-level statistics as open data: a user-centric approach » (statistiques locales sous la forme de données ouvertes : une approche axée sur l'utilisateur). Le Groupe de travail a mis au point le document en y incorporant des exemples concrets recueillis auprès de ses membres et au moyen d'activités de communication organisées à l'issue d'une manifestation parallèle, tenue le 26 février 2021 pendant la cinquante-deuxième session de la Commission, intitulée « The open data revolution: the power of open data for engaging the public with statistics that matter to them » (la révolution des données ouvertes : les données ouvertes peuvent faire participer le grand public à l'élaboration de statistiques qui lui importent)⁴³.

53. Les exemples produits par un pays sont classés selon cinq catégories de cas d'utilisation, détaillées dans le document de référence. Ci-dessous figure une liste de ces catégories ainsi que de brèves descriptions des exemples fournis :

a) Portails de données nationaux où sont recueillis les témoignages d'utilisateurs. Les exemples fournis par la Malaisie, les Pays-Bas et la Nouvelle-Zélande témoignent de l'ampleur des possibilités qu'offrent les données ouvertes pour ce qui est du caractère abordable du logement, des marchés immobiliers, du secteur de l'agriculture et des préoccupations en matière de santé et d'environnement à différentes échelles locales, y compris les banlieues, les municipalités, les districts et les États ;

b) Réutilisation intelligente au moyen des tableaux de bord et des applications. L'Australie, la Malaisie, les Pays-Bas et la Sierra Leone ont donné des exemples de portails de données ouvertes, de répertoires de témoignages et d'applications pour smartphone, y compris pour les ménages à faibles revenus et les peuples autochtones ;

c) Création de services par l'échange avec les utilisateurs finaux. La Nouvelle-Zélande et la Sierra Leone donnent des exemples de mesures de riposte face à la pandémie de COVID-19 et d'un GovHack ;

d) Hackatons et acteurs aux service des développeurs. La Malaisie, les Pays-Bas et la Nouvelle-Zélande donnent des exemples de collaboration avec des développeurs, y compris des journalistes et des établissements d'enseignement supérieur ;

e) Collaboration avec les médias. La Suède donne un exemple de journalisme de données.

54. Les organismes nationaux de statistique peuvent accroître l'utilisation, la réutilisation et la valeur des données ouvertes par le renforcement de la gouvernance et de l'intendance des données et l'échange avec les utilisateurs. Les exemples fournis par différents pays démontrent qu'une approche axée sur l'utilisateur aide les organismes nationaux de statistique à publier des données utiles et à faire des statistiques locales une partie importante de l'écosystème de données ouvertes d'un pays, conformément à l'optique adoptée dans le cadre de la troisième vague des données ouverte⁴⁴, pour reprendre la notion mis au point par le Open Data Policy Lab⁴⁵.

⁴³ Voir www.unescap.org/events/2021/open-data-revolution-power-open-data-engaging-public-statistics-matter-them.

⁴⁴ <https://opendatapolicylab.org/third-wave-of-open-data/>.

⁴⁵ Le Open Data Policy Lab a été créé à l'initiative du Governance Lab (GovLab), de la Tandon School of Engineering de la New York University, et vise à améliorer les conditions de vie des populations en modifiant les modalités de gouvernement. Pour de plus amples renseignements, voir <https://opendatapolicylab.org/team/>.

55. Le Groupe de travail a mis au jour plusieurs difficultés relatives aux travaux menés au moyen de statistiques locales sous la forme de données ouvertes. Ces problèmes ont été exposés à la manifestation parallèle susmentionnée, tenue à la cinquante-deuxième session de la Commission de statistique. Ils ont trait aux environnements opérationnels, aux méthodes de mesure des conséquences des données locales ouvertes, à la diversité des publics auxquels des informations sont communiquées, au respect de la vie privée et de la confidentialité, à la technologie, aux normes et à la qualité des métadonnées.

56. La publication d'un nombre plus important d'ensembles de données contenant des données locales ouvertes aide les organismes nationaux de statistique à conserver leur pertinence, car ces statistiques peuvent souvent être utilisées et réutilisées avec d'autres types de données ouvertes, telles que les données géospatiales. L'accroissement de la publication de données locales ouvertes est une solution dont disposent les organismes nationaux de statistique qui cherchent à renforcer leur rôle dans l'écosystème national de données, où différents acteurs peuvent échanger, produire et utiliser des données⁴⁶. Quand les données locales sont utilisées de manière sûre, c'est-à-dire dans le respect de la vie privée et des conditions de sécurité et de confidentialité, elles peuvent constituer une source abondante d'éclaircissements concernant les populations et leurs différents groupes et peuvent contribuer à favoriser les mesures de riposte face aux difficultés prioritaires telles que les catastrophes, les changements climatiques et la réduction de la pauvreté.

57. Chaque organisme national de statistique choisit sa solution en fonction de son contexte national. Toutefois, tous ces organismes ont de fait toujours la même difficulté : tirer parti du potentiel de tous les types de données créés dans la société et protéger et cultiver ce potentiel afin de redistribuer à la société les fruits de cette entreprise, à savoir des statistiques officielles et des services de statistiques et de données de qualité. Une solution en vue de comprendre comment accroître les connaissances des pouvoirs publics et de la société consiste à échanger avec les utilisateurs et à travailler avec des intermédiaires tels que les développeurs et les médias. Le rôle des statisticiens et des statistiques ne constitue qu'un pan de l'écosystème des données : des mesures collectives peuvent contribuer à réaliser la transformation envisagée dans le Programme de développement durable à l'horizon 2030⁴⁷, le Cadre de Sendai pour la réduction des risques de catastrophe (2015-2030)⁴⁸ et l'Accord de Paris adopté dans le cadre de la Convention-cadre des Nations Unies sur les changements climatiques⁴⁹.

IV. Poursuite des travaux relatifs aux données ouvertes

58. Le Groupe de travail sur les données ouvertes a avancé l'idée que les organismes nationaux de statistique sont bien placés pour contribuer à créer et à encourager une culture des données ouvertes s'ils renforcent leur activité principale, qui consiste à fournir des statistiques officielles et, dans certains cas, s'ils s'acquittent de leur rôle d'intendant des données.

59. En tout temps, il convient de trouver un équilibre entre une utilisation sûre des données et la liberté d'explorer et de faire des expériences innovantes en vue de tirer de nouveaux avantages des données. Forts de la confiance et de la transparence qui

⁴⁶ Voir « Approaches to data stewardship » (approches en matière d'intendance des données), document de référence établi dans le cadre de la cinquante-troisième session de la Commission de statistique (2021).

⁴⁷ Résolution 70/1 de l'Assemblée générale.

⁴⁸ Résolution 69/283 de l'Assemblée générale, annexe II.

⁴⁹ Voir FCCC/CP/2015/10/Add.1, décision 1/CP.21, annexe.

sont au fondement de leurs travaux, les organismes nationaux de statistique peuvent accélérer la publication de données ouvertes fondées sur la facilité d'accès, la réutilisation et l'intégration.

60. Le Groupe de travail sur les données ouvertes est convenu que le Groupe de travail sur l'intendance des données, établi par la Commission de statistique à sa cinquante-deuxième session, pourrait poursuivre les travaux qu'il mène en vue de faire progresser la culture des données ouvertes, ainsi que les travaux qu'il réalise avec les acteurs du domaine des données ouvertes, en particulier en ce qui concerne les données locales ouvertes. En outre, il a recommandé que la réalisation des travaux relatifs aux directives en matière de données locales ouvertes et de culture des données ouvertes fasse partie des tâches qui incombent au Groupe de travail sur l'intendance des données.

61. Dans sa décision 51/124, la Commission de statistique a chargé le Groupe de travail sur les données ouvertes de fournir de nouvelles orientations relatives aux données ouvertes, notamment dans trois domaines : le principe d'ouverture par défaut, l'interopérabilité et les données locales ouvertes. Le présent rapport et le document de référence qui l'accompagne, ainsi que les rapports et documents de référence publiés depuis 2019, contiennent une gamme complète d'orientations qui peuvent servir de ressource pour les offices statistiques nationaux qui visent à faire éclore une culture des données ouvertes et à établir des procédures dans ce domaine. Par la création de ces directives – qui peuvent servir de modèle à d'autres groupes de travail, y compris le Groupe de travail sur l'intendance des données – le Groupe de travail sur les données ouvertes s'acquitte du mandat que lui a confié la Commission. Par conséquent, il est recommandé que l'achèvement et l'adoption de ce matériel de référence marquent la conclusion des travaux du Groupe de travail sur les données ouvertes et que les ressources qu'il a fournies soient utilisées par d'autres groupes et intégrées dans leurs propres travaux.

V. Mesures que la Commission de statistique est appelée à prendre

62. La Commission est invitée à prendre les mesures suivantes :

- a) Prendre note du présent rapport ;
- b) Examiner et approuver les projets d'orientations concernant les modalités de mise en place d'un modèle d'ouverture par défaut pour la diffusion des statistiques ;
- c) Prendre note de l'examen des meilleures pratiques en matière d'octroi de licences applicables aux données et de leur harmonisation avec la législation interne ;
- d) Exprimer son point de vue sur le modèle de maturité des données commun et les orientations connexes ;
- e) Prendre note de l'examen des cas d'utilisation des statistiques locales ouvertes ;
- f) Exprimer son point de vue sur la proposition d'intégrer les questions relatives aux données ouvertes dans les discussions du Groupe de travail sur l'intendance des données ;
- g) Examiner et approuver la proposition faite par le Groupe de travail sur les données ouvertes de conclure ses travaux.