



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

Genève, 20-22 juin 2022

Point 5 de l'ordre du jour provisoire

Travaux du Groupe de haut niveau sur la modernisation de la statistique officielle**Rapport sur les travaux menés en 2021 par le Groupe de haut niveau sur la modernisation de la statistique officielle****Établi par le Groupe de haut niveau sur la modernisation de la statistique officielle***Résumé*

Le présent document récapitule les résultats des travaux réalisés en 2021 par le Groupe de haut niveau sur la modernisation de la statistique officielle.

En 2021, deux projets ont été exécutés : l'un était consacré aux techniques de préservation de la confidentialité des données d'entrée et l'autre était un guide des données synthétiques. Le programme de travail comprenait également la poursuite des travaux des groupes de modernisation, ainsi que plusieurs réunions d'experts et ateliers organisés par les comités directeurs.

L'additif 1 présente les travaux prévus dans le cadre du Groupe de haut niveau pour 2022. L'additif 2 est une note décrivant la vue géospatiale du Modèle générique du processus de production statistique, que le Groupe d'appui à la mise en œuvre des normes du Groupe de haut niveau a finalisé en 2021. L'additif 3 est un extrait du document d'orientation intitulé *The Role of Brand and Reputation Management, Marketing, and Crisis Communication for Statistical Organisations* (Le rôle de la gestion de la marque et de la réputation, du marketing et de la communication de crise pour les organismes de statistique), que le Groupe des capacités et de la communication du Groupe de haut niveau a finalisé en 2021.

À la réunion de février 2022, le Bureau de la Conférence des statisticiens européens a passé en revue les résultats des travaux menés en 2021 dans le cadre du Groupe de haut niveau et a formulé des recommandations sur les travaux prévus pour 2022. La Conférence des statisticiens européens est invitée à approuver le rapport et à prendre note des travaux prévus dans le domaine de la modernisation de la statistique.



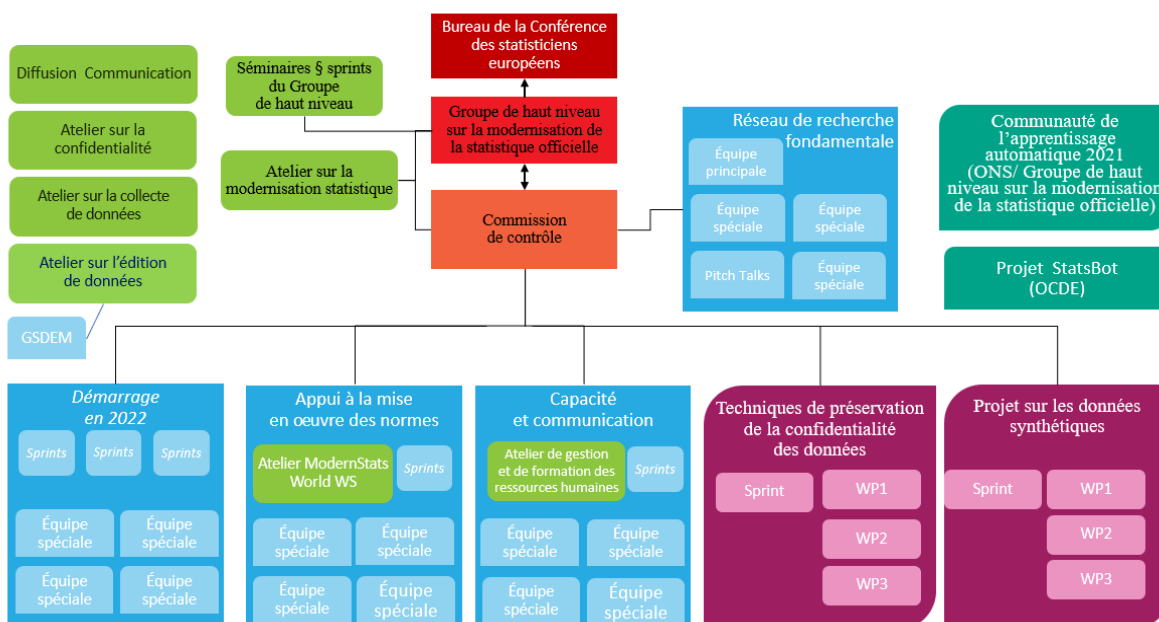
I. Introduction

1. Le Groupe de haut niveau sur la modernisation de la statistique officielle a été créé par la Conférence des statisticiens européens pour promouvoir la modernisation de la statistique officielle. Il est dirigé par les statisticiens en chef de 14 organisations qui en fixent la vision, la mission et les thèmes prioritaires. La Commission de contrôle du Groupe de haut niveau a été mise en place afin de suivre activement l’avancement des travaux, de donner des orientations stratégiques et d’adapter les activités si nécessaire. On trouvera de plus amples informations sur le site Web du Groupe de haut niveau accessible au public (<https://statswiki.unece.org/display/hlgbas/High-Level+Group+for+the+Modernisation+of+Official+Statistics>).

2. En 2021, deux projets, trois groupes d’experts (appuyés par des équipes spéciales supplémentaires) et plusieurs réunions et ateliers d’experts organisés par les comités directeurs étaient prévus au programme de travail du Groupe de haut niveau. La figure 1 présente un aperçu schématique des groupes, réseaux, projets et ateliers opérationnels en 2021 dans le cadre du Groupe de haut niveau.

3. La pandémie actuelle de COVID-19 a continué de poser des problèmes mais elle a surtout eu des répercussions sur l’organisation des réunions en présentiel. En fin de compte, une seule réunion d’experts a été organisée en présentiel. Faute de pouvoir organiser des sessions « sprint » en présentiel, la concrétisation des résultats du projet a été quelque peu retardée. Cependant, certains effets positifs ont été constatés : davantage de participants ont appris à travailler ensemble à distance, ce qui était déjà la modalité généralement suivie dans le cadre du Groupe de haut niveau avant la pandémie. Au total, les groupes, les équipes spéciales, les comités directeurs et les projets sur la modernisation ont réuni plus de 250 membres issus d’une centaine d’organisations différentes. Plus de 1 000 personnes ont assisté aux réunions et aux ateliers, ce qui était un peu moins qu’en 2020. Même s’il y a eu un groupe en moins et moins de réunions d’experts, d’ateliers¹ et de sessions « sprint », la participation par activité s’est avérée plus élevée.

Figure 1
Structure du Groupe de haut niveau sur la modernisation de la statistique officielle en 2021



¹ En raison de la pandémie et du manque de personnel à la CEE, la Commission de contrôle a décidé que les réunions d’experts et les ateliers suivraient un cycle de vingt-quatre mois.

A. Commission de contrôle

4. La Commission de contrôle a exercé un suivi régulier des travaux et s'est réunie chaque mois, notamment avec les chefs de projet et les présidents de groupe. Le programme de travail proposé pour 2021 comprenait déjà des mesures visant à atténuer les effets de la pandémie de COVID-19 et des pénuries de personnel à la CEE, telles que la réduction de la fréquence des réunions d'experts, de sorte qu'aucun autre ajustement n'était nécessaire.

5. Les réunions virtuelles mensuelles consistent tour à tour à faire le point sur la modernisation et à examiner la stratégie. À tout moment, les présidents et les chefs de projet peuvent demander l'aide ou l'intervention de la Commission de contrôle. Les membres de la Commission se voient également confier la tâche d'être les « champions » d'activités en particulier et de suivre les travaux plus en détail. Cela contribue à s'assurer que les thèmes des réunions d'experts et des ateliers, ainsi que les résultats obtenus par les projets et les groupes, correspondent à la mission, à la vision et aux thèmes prioritaires retenus par le Groupe de haut niveau. La Commission est également chargée d'organiser, en collaboration avec le secrétariat de la CEE, l'atelier de modernisation du Groupe de haut niveau. Des mises à jour mensuelles de l'avancement des travaux des groupes et des projets sont disponibles grâce au lien [Modernisation Updates](#).

II. Projets

6. Pour 2021, deux projets ont été choisis : l'un était consacré aux techniques de préservation de la confidentialité des données d'entrée et l'autre était un guide des données synthétiques.

A. Projet de préservation de la confidentialité des données

7. Le projet de préservation de la confidentialité des données a été reporté de 2020 à 2021 en raison de la pandémie. Statistique Pays-Bas a nommé Dennis Ramondt comme chef de projet à temps partiel. Au total, 24 collègues de neuf organisations ont participé au projet. Les membres du projet ont eu des réunions plénières en ligne une fois par mois, les groupes chargés de lots de travaux et les sous-groupes se réunissant plus régulièrement. Même s'il n'a pas été possible de trouver un spécialiste expérimenté pour codiriger le projet et d'organiser les sessions « sprint » en présentiel, le projet a bien avancé. Les exercices de cadrage déjà entrepris en 2020 y ont contribué.

8. Le projet a été articulé autour de catégories générales. Après avoir affiné le champ d'application et commencé par documenter et étudier divers cas d'utilisation, des catégories générales ont été définies, ce qui a conduit à distinguer trois domaines de travail :

- a) L'intersection d'ensembles privés ;
- b) L'apprentissage automatique privé ;
- c) L'organisation d'une enquête consultative publique.

9. L'intersection d'ensembles privés est une technique cryptographique de calcul multipartite sécurisée qui permet à deux parties au moins possédant des ensembles de comparer des versions cryptées de ces ensembles afin de calculer l'intersection. Dans ce scénario, aucune des parties ne révèle quoi que ce soit à son homologue, à l'exception des éléments figurant dans l'intersection. Cela permet d'effectuer des requêtes (agrégations) et des analyses concernant l'ensemble des données combinées sans révéler les microdonnées ou des informations autres que celles convenues appartenant à l'autre partie.

10. Le projet pilote d'apprentissage automatique privé visait à construire un environnement simulé pour valider le concept d'apprentissage automatique préservant la vie privée à caractère multipartite à la fois pour ce qui est de la formation et de l'interférence. L'objectif était d'étudier les meilleures pratiques et les outils à code source ouvert pour la formation distribuée et collaborative en apprentissage automatique entre plusieurs organisations dans un environnement à faible niveau de confiance, tout en bénéficiant

mutuellement des résultats (le modèle final) ou en permettant un accès sûr à des tiers. Une configuration multiorganisationnelle simulée avec plusieurs organismes nationaux de statistique collectant des données sur les individus (capteurs) pour prédire leurs activités (enquêtes sur l'utilisation du temps et le bien-être) a été testée avec succès.

11. Les résultats ont été partagés lors d'un webinaire qui a été bien accueilli et il a été procédé aux préparatifs d'une enquête consultative publique sur l'informatique privée sécurisée en tant que service. Comme il a été estimé que l'extension du projet procurerait des avantages importants, une proposition visant à poursuivre le projet a été soumise au Groupe de haut niveau, qui l'a acceptée.

B. Projet de guide des données synthétiques

12. Le projet de guide des données synthétiques faisait suite aux travaux sur les données synthétiques entamés dans le cadre du Réseau de recherche fondamentale en 2020. Le projet a été dirigé par Kate Burnett-Isaacs (Statistique Canada)². Y ont participé 50 collègues de 25 institutions, y compris du monde universitaire et du secteur privé. Le groupe de projet global s'est réuni tous les mois, les équipes chargées de lots de travaux et les sous-équipes ayant tenu des réunions fréquentes. Ces réunions ont donné lieu à plusieurs présentations des travaux méthodologiques et pratiques entrepris par les membres.

13. L'objectif principal du projet était de produire un guide pratique sur les données synthétiques pour les organismes nationaux de statistique. Les données synthétiques peuvent être une solution pour fournir des données tout en respectant l'intégrité et la confidentialité. Elles aident ces organismes à assurer une plus grande transparence et ouverture.

14. Le Guide présente des méthodes théoriques permettant de créer des données synthétiques et dégage un consensus international sur les applications concrètes et les meilleures pratiques visant à promouvoir la cohérence, la transparence et la comparabilité au sein des organismes nationaux de statistique et entre eux, ainsi que parmi les utilisateurs du monde universitaire et du secteur privé. Il donne des orientations cohérentes aux décideurs de ces organismes travaillant à tous les niveaux, afin de les aider à savoir si les données synthétiques sont la bonne solution à leur problème de publication de données. Le guide est intentionnellement conçu pour une application pratique ; il n'est pas un manuel exhaustif. Les ressources permettant d'approfondir les concepts techniques sont mises en évidence dans le guide, qui comprend un arbre de décision pratique permettant de choisir la méthode qui convient à chacun en fonction des propriétés des données originales et des propriétés souhaitées des données synthétiques.

15. Le guide a été présenté le 17 novembre, lors d'un webinaire public en ligne qui y était spécialement consacré, dans le cadre d'une manifestation parallèle à l'atelier de modernisation du Groupe de haut niveau de 2021 (15-16 novembre 2021). Un concours de données est prévu au début de l'année 2022 (24-28 janvier 2022). La version finale du guide, qui sera établie en se fondant sur l'évaluation du concours, devrait être publiée en mars 2022. Le guide figure également dans le programme des publications de la CEE pour 2022. Les résultats préliminaires et le projet de guide sont déjà accessibles au public (<https://statswiki.unece.org/x/UQTUE>).

III. Groupes de modernisation

16. En 2021, le Groupe de haut niveau était structuré en trois groupes : le Groupe d'appui à la mise en œuvre des normes, le Groupe des capacités et de la communication et le Réseau de recherche fondamentale. Les groupes coordonnent leurs activités entre eux, mais aussi avec les projets et d'autres activités internationales en cours dans des domaines similaires (par exemple à Eurostat, à l'OCDE et à la Division de statistique de l'ONU). Les présidents des groupes rendent compte chaque mois à la Commission de contrôle des questions

² Elle a été aidée par un comité technique composé de Christine Task (Nexus Research), Gillian Raab (Université d'Edimbourg), Kenza Sallier (Statistique Canada) et Alistair Ramsden (StatsNZ).

examinées. Le lancement d'un nouveau groupe de modernisation remplaçant le Groupe de la mise en commun des outils (qui a fusionné avec le Groupe d'appui à la mise en œuvre des normes) a été reporté.

A. Réseau de recherche fondamentale

17. Le Réseau de recherche fondamentale est le laboratoire d'idées de la communauté de la modernisation statistique. Il sert de plateforme de recherche et d'innovation dans laquelle les membres peuvent partager des idées et chercher des partenaires pour étudier comment les innovations portant sur les processus de production peuvent bénéficier aux organismes de statistique. L'objectif est d'élaborer et d'évaluer des propositions d'activités à exécuter par le Groupe de haut niveau et, le cas échéant, de réaliser des études de suivi à court terme. Le réseau est une structure qui chapeaute un groupe principal et des équipes spéciales à court terme chargés d'étudier de nouvelles idées et possibilités au moyen de courts projets d'évaluation soutenus par la Commission de contrôle du Groupe de haut niveau. Toute idée qui s'inscrit dans le cadre stratégique du Groupe de haut niveau peut être prise en considération. Pour favoriser l'innovation, un espace est également accordé à la réflexion hors des sentiers battus.

18. Le Réseau est dirigé par Barteld Braaksma, responsable de l'innovation de Statistique Pays-Bas, et appuyé par la CEE. Le groupe principal d'experts compte une quinzaine de membres qui se sont réunis au moins une fois par mois. En outre, plusieurs séances de présentation (« pitch talks ») sont organisées pour cerner de nouveaux sujets. Ces séances sont des « mini sprints » qui permettent aux membres et à des personnes extérieures de présenter brièvement une idée ou un projet dans le domaine de la modernisation, avant que le groupe principal et les experts invités n'en débattent.

19. En 2021, le Réseau a cerné et examiné les sujets suivants :

- a) Le pouvoir des données réseau : études de faisabilité (Daniel Elazar, ABS) ;
- b) La plate-forme du centre commun de biosécurité COVID-19 (Eric Deeben, ONS) ;
- c) La recherche sur les utilisateurs pour les statistiques officielles (Eric Anvar, OCDE) ;
- d) Le système d'enquête rapide (Branko Josipović, SORS) ;
- e) De l'expérimentation à la mise en œuvre dans les statistiques officielles (Kate Burnett, Statistique Canada) ;
- f) Des microdonnées pour comprendre la baisse des taux de réponse (Gary Dunnet, StatsNZ) ;
- g) L'enquête sur la COVID fondée sur Facebook (Barteld Braaksma, Statistique Pays-Bas).

20. L'activité « De l'expérimentation à la mise en œuvre dans les statistiques officielles » (sujet mentionné à l'alinéa e) ci-dessus) a débouché sur l'organisation d'une session consacrée à ce sujet lors du Congrès mondial de la statistique de l'Institut international de statistique (IIS) par les membres du Réseau de recherche fondamentale. Le retour d'information et les discussions ont abouti à une proposition de projet pour le programme de travail 2022 du Groupe de haut niveau : « Meta Academy for the Modernization of Official Statistics » (Méta-académie pour la modernisation des statistiques officielles). Sur la base d'un sujet examiné en 2020, le Réseau a élaboré une autre proposition de projet (en collaboration avec le Groupe d'appui à la mise en œuvre des normes du Groupe de haut niveau) : « Data Governance Framework to Achieve Data Interoperability » (Cadre de gouvernance des données pour réaliser l'interopérabilité des données). À la suite de l'étude de faisabilité menée sur le pouvoir des données réseau (thème mentionné à l'alinéa a)), l'ABS, l'ONS, Statistique Pays-Bas et quelques universités étudient actuellement la possibilité de collaborer à un modèle de perturbation des réseaux de la chaîne d'approvisionnement. La plateforme du centre commun de biosécurité (thème mentionné à

l'alinéa b) du paragraphe 20 et un système d'enquête rapide (thèmes mentionnés à l'alinéa d)) seront examinés plus avant en 2022.

21. La Commission de contrôle du Groupe de haut niveau supervise les travaux et peut demander à ce que des sujets susceptibles de contribuer à la modernisation de la statistique soient explorés. Le Groupe reste prêt à évaluer les propositions d'activité et de projet soumises par la communauté statistique. De plus amples informations figurent sur la page wiki du réseau de réflexion du Réseau de recherche fondamentale ([Blue Skies Thinking Network wiki page](#)).

B. Groupe d'appui à la mise en œuvre des normes

22. Le Groupe d'appui à la mise en œuvre des normes est chargé de mettre à jour et de mettre au point des modèles et des normes tels que le GAMS0, le GSBPM, le GSIM et le CSPA³, qui sont élaborés par le Groupe de haut niveau. L'objectif du Groupe est de concevoir, d'améliorer, d'intégrer, de promouvoir, de soutenir et de faciliter la mise en œuvre de l'éventail de normes nécessaires à la modernisation des statistiques. Le Groupe a été présidé par Zoltán Vereczkei (Hongrie) et a bénéficié de l'appui du secrétariat de la CEE. En 2021, cinq équipes spéciales ont travaillé sous l'égide du Groupe. Le groupe principal est composé de vingt experts issus de douze organismes nationaux de statistique et de quatre organisations internationales. De nombreux autres collègues ont collaboré aux équipes spéciales. Les réunions du groupe plénier ont eu lieu tous les mois et les cinq équipes spéciales se sont réunies fréquemment.

23. Les principales activités et les principaux résultats du Groupe et de ses équipes spéciales en 2021 ont été les suivants :

a) Intégration du GSBPM et du GSIM : une vue intégrée du GSBPM et du GSIM a été créée. Elle permet de mieux comprendre la relation entre le GSBPM et le GSIM en décrivant pour les sous-processus du GSBPM les spécifications d'entrée et de sortie en termes d'objets du GSIM. C'est un moyen important de progresser dans la mise en œuvre des modèles ModernStats (notamment du GSIM, puis du GSBPM ou des deux modèles en même temps) ;

b) GeoGSBPM : La vue géospatiale du Modèle générique du processus de production statistique rend opérationnels les principes du Cadre statistico-géospatial mondial tout au long du processus de production. Le Cadre définit des principes qui étayent la production de données statistiques géospatiales harmonisées et normalisées. Le GeoGSBPM présente les activités et les considérations liées au géospatial en utilisant le cadre du GSBPM ;

c) Mise à jour du GSIM : il était prévu à l'origine de faire une « mise à jour douce » pour clarifier les points ambigus et corriger les erreurs mineures recensées au cours d'autres activités du Groupe telles que la relation entre la GSBPM et le GSIM et les travaux sur l'ontologie de base. Cependant, il en est rapidement ressorti qu'un examen plus approfondi était nécessaire. L'activité se poursuivra sous la forme d'un examen complet du GSIM en 2022 ;

d) Ontologie de base pour les statistiques officielles (COOS) : L'objectif est de mettre au point un cadre formel qui permet l'interopérabilité, l'activation et l'identification unique au niveau mondial des différents modèles ModernStats (GSBPM, GSIM, GAMS0, CSDA). L'ontologie de base propose des représentations formelles pour les concepts fondamentaux utilisés dans notre domaine. L'équipe spéciale COOS a jusqu'à présent publié la spécification de l'ontologie, un document de gouvernance, une politique URI et une ontologie OWL (actuellement en cours d'examen par des experts). Des précisions sont disponibles sur GitHub : <https://linked-statistics.github.io/COOS/coos.html>. Les activités se poursuivront en 2022 ;

³ Modèle générique d'activité des organismes statistiques (GAMS0), Modèle générique du processus de production statistique (GSBPM), Modèle générique d'informations statistiques (GSIM), Architecture commune de la production statistique (CSPA).

e) **Tâche GSBPM** : l'équipe spéciale de la tâche GSBPM a commencé ses travaux à l'automne 2021. Les tâches sont les activités de niveau inférieur ajoutées à l'échelon national pour adopter les descriptions d'activités génériques de niveau supérieur pour les sous-processus du GSBPM. De nombreux exemples nationaux ont maintenant été recueillis et peuvent aider d'autres pays à appliquer le GSBPM. Les travaux se poursuivront en 2022 avec pour objectif principal de définir les tâches communes qu'il est envisagé d'inscrire dans la prochaine version du GSBPM ;

f) **Gouvernance de ModernStats** : le Groupe a élaboré des orientations sur les politiques de gouvernance pour les différents modèles ModernStats et d'autres produits du Groupe de haut niveau portant sur l'appui à l'utilisation de ces modèles. Il clarifie les fréquences de révision, les procédures de révision et les principes de révision, ainsi que les principes et la propriété des versions ;

g) **Enquête sur l'utilisation de ModernStats** : des enquêtes régulières sur l'utilisation des modèles ModernStats sont menées. Quarante-cinq pays ont répondu à l'enquête de 2021. Les résultats montrent que le GSBPM est largement mis en œuvre, tandis que le GAMSO, le GSIM et la CSPA sont moins utilisés. D'autres résultats sont disponibles à l'adresse : <https://statswiki.unece.org/x/rQixEw>.

C. Groupe des capacités et de la communication

24. Le Groupe des capacités et de la communication est chargé des aspects liés à la gestion et à la formation des ressources humaines, ainsi qu'à la communication des organismes de statistique. Les services de ressources humaines et de communication ont été fortement mobilisés dans la lutte contre la COVID-19. Après s'être concentrés sur les différents problèmes posés par la pandémie, ils ont mis l'accent en 2021 sur les conséquences de la pandémie.

25. En 2021, 46 collègues de 18 organisations ont participé au Groupe et à ses équipes spéciales et sous-groupes. Le Groupe est dirigé par Anna Borowska (Statistique Pologne) et Maria Hurley (CSO Ireland). Le Groupe et ses équipes spéciales et sous-groupes se sont réunis une fois par mois, voire plus souvent lorsque cela était nécessaire. Les activités du Groupe sont regroupées dans trois domaines :

a) **Gestion de l'éthique** : l'activité comprend le leadership éthique et couvre aussi bien l'éthique des affaires que l'éthique des données. Au début de l'année 2021, une courte enquête a été menée auprès des organismes nationaux de statistique afin de déterminer si ceux-ci avaient déjà mis en place des politiques ayant trait aux principes clefs de la gestion de l'éthique et du leadership éthique. Elle a été suivie d'une deuxième enquête visant à recueillir des informations sur l'existence de pratiques de gestion de l'éthique et de leadership éthique tant au niveau des processus organisationnels (éthique des affaires) que de la production et de la recherche (éthique des données). Les résultats sont disponibles dans le rapport de l'enquête sur la gestion de l'éthique ([report on survey ethics management](#)).

b) **Études de marché, marketing numérique et stratégies de communication** : une marque et une réputation bien développées constituent un atout stratégique essentiel pour promouvoir la valeur des statistiques officielles et combattre les effets de la désinformation sur nos sociétés. L'équipe spéciale a élaboré un document d'orientation sur le rôle de la gestion de la marque et de la réputation, du marketing et de la communication de crise dans la résolution des problèmes auxquels sont actuellement confrontés les organismes nationaux de statistique et les organisations statistiques internationales. Le guide est accompagné de cas d'utilisation qui seront partagés sur les wikis du Groupe de haut niveau. Tout le matériel sera disponible sur le site <https://statswiki.unece.org/x/cRGxEw> (disponible fin février 2022).

c) **Le travail de demain, le lieu de travail de demain et les compétences de demain** : la pandémie a fortement accéléré l'évolution de la façon dont nous travaillons et des compétences dont nous avons besoin. Les équipes spéciales ont conçu, à l'intention de tout organisme national de statistique en transition vers un mode de travail mixte, trois boîtes à outils destinées aux employés, aux administrateurs et aux employeurs qui serviront de cadre pour structurer les politiques et les solutions locales et individuelles et pour s'adapter aux

besoins locaux et individuels. Ces boîtes sont censées examiner les considérations liées à la transition vers un modèle de travail mixte ou hybride et étudier les obstacles, les besoins et les perspectives propres à l'employeur, à l'employé et à l'administrateur. Elles permettent de soulever des questions qui appellent des solutions, de cerner les domaines communs de préoccupation ou ceux requérant une attention particulière et de proposer des actions contributives de haut niveau. Enfin, elles servent à lancer des conversations internes.

D. Groupe de l'apprentissage automatique en 2021

26. Le Groupe de l'apprentissage automatique est le prolongement du projet d'apprentissage automatique 2019-2020 du Groupe de haut niveau. La communauté est dirigée par le Data Science Campus de l'ONS avec le soutien de la CEE. Les objectifs sont notamment de fournir une plateforme pour faciliter la recherche afin de moderniser les statistiques officielles, de renforcer les capacités en matière d'apprentissage automatique et de partager les connaissances. Le Groupe est impulsé par la communauté et toutes les activités sont dirigées par les membres. Il est ouvert à tous les organismes officiels de statistique et accessible à différents niveaux de compétence. Des ressources sont disponibles pour la communauté au sens large, car l'objectif est d'offrir des ressources ouvertes, partageables et facilement accessibles à la communauté et de faciliter le renforcement des capacités d'apprentissage automatique dans le domaine des statistiques officielles.

27. Près de 250 membres de plus de 30 pays ont participé à la communauté. Il existe une forte demande de partage des connaissances et de renforcement des capacités et une forte participation aux réunions. Les travaux s'articulent autour de cinq domaines de travail censés aborder les différentes questions qui se posent lorsque l'apprentissage automatique est utilisé pour les statistiques officielles :

a) **Domaine de travail 1 – De l'idée à la solution valide** : Les études pilotes sont menées pour évaluer la valeur ajoutée de l'apprentissage automatique dans divers domaines thématiques : codage et classification, édition et imputation, utilisation de données d'imagerie, modélisation et optimisation des voies. Une étude menée sur l'expérience de la réplication a mis en évidence les avantages d'un partage de ces projets d'apprentissage automatique ;

b) **Domaine de travail 2 – De la solution valide à la production** : Les études ont porté sur les questions entourant l'application des solutions d'apprentissage automatique et trois activités sont menées par le FMI et l'INEGI (Mexique). Les rapports sont disponibles sur l'espace wiki de l'apprentissage automatique : <https://statswiki.unece.org/x/eAB-EQ> ;

c) **Domaine de travail 3 – Prise en compte de l'éthique dans l'utilisation de l'apprentissage automatique aux fins de la recherche et de la production de statistiques** : Ces orientations de haut niveau ont étudié les aspects éthiques de l'utilisation des techniques d'apprentissage automatique aux fins de la recherche et de la production de statistiques. Elles ne sont pas exhaustives, mais visent à aider et à soutenir les analystes, les chercheurs, les spécialistes des données et les statisticiens qui traitent les aspects éthiques des projets basés sur l'apprentissage automatique ;

d) **Domaine de travail 4 – Reconfiguration du modèle** : les circonstances dans lesquelles un modèle d'apprentissage automatique doit être reconfiguré pour préserver le pouvoir prédictif et la qualité du modèle ont été définies ;

e) **Domaine de travail 5 – Cadre de qualité pour l'algorithme statistique** : les études portent sur les aspects du cadre de qualité pour l'algorithme statistique (QF4SA) qui a été mis au point dans un projet consolidé relevant du projet d'apprentissage automatique qui vise à analyser un résultat basé sur un ensemble de paramètres et de procédures standard.

28. Les résultats de la collaboration en matière de recherche dans les cinq domaines de travail et les 18 projets de recherche, ainsi que les ressources d'apprentissage et de formation et d'autres matériels, sont (ou seront) disponibles sur les pages des communautés : <https://statswiki.unece.org/x/eAB-EQ>. Les travaux ont également été partagés lors de sessions « Coffee and Coding » qui ont été organisées par le Groupe et qui ont servi à échanger les connaissances ainsi qu'à obtenir un retour d'information.

IV. Réunions et ateliers

29. En réponse à la pandémie de COVID-19 et au manque de personnel à la CEE, la Commission de contrôle et le secrétariat de la CEE ont décidé de modifier le cycle des réunions d'experts et des ateliers pour le porter à vingt-quatre mois. En 2021, les réunions et ateliers d'experts suivants ont eu lieu :

a) **Réunion du Groupe de haut niveau** (29 janvier, en ligne) : Comme la réunion habituelle du Groupe de haut niveau le week-end précédant la session de la Commission de statistique de l'ONU à New York n'a pas pu avoir lieu, les statisticiens en chef qui sont membres du Groupe de haut niveau se sont réunis en ligne. Le programme de travail a été approuvé. Les statisticiens en chef ont beaucoup apprécié les résultats obtenus en 2020. Tous les participants ont été remerciés pour l'excellent travail réalisé, malgré des circonstances très difficiles. La Commission de contrôle a été chargée de continuer à superviser les progrès accomplis et de veiller à ce que les activités et les projets soient conformes à la déclaration de mission et de vision et aux thèmes prioritaires du Groupe de haut niveau. Elle a également été priée de faire des propositions de création d'un nouveau groupe de modernisation et des propositions de projets et d'activités pour 2022 ;

b) **Réunion d'experts de la collecte de données statistiques** (27-30 septembre, en ligne) : Cette réunion en ligne a été suivie par 151 participants de 30 pays. Elle comprenait 32 présentations qui ont été réparties en quatre sessions de fond : i) Adapter la collecte de données dans un contexte de crise ; ii) Gérer et moderniser la collecte des données ; iii) Multimode et intégration des sources de données ; iv) Attention accordée aux répondants. Les discussions en petits groupes ont permis de définir plusieurs futurs domaines de travail, par exemple les nouvelles techniques et technologies de l'information pour la collecte automatisée (de machine à machine, system2system, API, portails Web), les aspects multidimensionnels de la qualité et la conception d'enquêtes adaptatives. Tous les résumés, documents, présentations et autres résultats de l'atelier sont disponibles sur le site Web de la CEE (<https://unece.org/statistics/events/DC2021>) ;

c) **Réunion d'experts de la diffusion et de la communication des statistiques** (11-14 octobre, en ligne) : 151 participants de 33 pays y ont assisté. Vingt-sept contributions ont été soumises. La réunion comprenait des sessions consacrées aux thèmes suivants : i) L'ère numérique ; ii) L'élargissement des publics ; iii) L'explication des statistiques ; iv) L'adaptation de la communication au public visé. Au cours des discussions en petits groupes, de futurs sujets de travail ont été définis. Les sujets les plus appréciés étaient les suivants : « Comment associer et recruter les jeunes générations qui sont nées avec le numérique ? », « Comment mesurer le succès et l'impact de notre communication ? » « Étudier les plateformes/outils de communication non conventionnels tels que les bandes dessinées, TikTok et Spotify », « Les capacités de communication, les compétences, les profils, les experts dont nous avons besoin » et « Les stratégies permettant de combattre et d'anticiper la désinformation ». Les autres sujets évoqués et tous les documents de la réunion sont disponibles sur le site Web de la CEE : <https://unece.org/statistics/events/DissComm2021> ;

d) **Atelier et webinaires sur la modernisation organisés par le Groupe de haut niveau** (15-19 novembre, en ligne). La manifestation principale a été suivie par 148 représentants de 41 pays. Tout comme le présent document, l'objectif principal de cet atelier était de rendre compte des activités et des résultats des différents groupes, projets et ateliers inscrits dans le programme de travail du Groupe de haut niveau pour 2021. Un autre objectif majeur était de présenter et d'examiner le programme de travail prévu pour 2022. Tous les documents de la réunion sont disponibles sur le site Web de la CEE (<https://unece.org/statistics/events/HLG2021>). En marge de la manifestation, trois webinaires ont été organisés pour mettre en commun les résultats des deux projets du Groupe de haut niveau et de la communauté de l'apprentissage automatique menés en 2021 :

- **Webinaire sur les données synthétiques** (17 novembre, en ligne). La présentation du guide des données synthétiques a été suivie par 73 participants originaires de 24 pays ;

- **Webinaire IPP** (18 novembre, en ligne). Les résultats du projet ont été présentés ainsi que les plans de suivi, notamment une consultation sur l'informatique privée sécurisée en tant que service. Le webinaire a été suivi par 70 participants représentant 21 pays ;
- **Webinaire du Groupe de l'apprentissage automatique** (19 novembre, en ligne). Lors de cette manifestation conjointe ONS-CEE, les travaux réalisés dans le cadre des différents domaines de travail et les travaux prévus ont été présentés à 279 participants originaires de 48 pays.

e) **Réunion d'experts sur la confidentialité des données statistiques** (1^{er}-3 décembre, Poznan, Pologne) : il était initialement prévu que cette réunion se déroule en présentiel. Toutefois, en raison de l'aggravation de la pandémie, la plupart des participants ont dû s'inscrire en ligne. La réunion a été accueillie par Statistique Pologne et l'Université d'économie et de commerce de Poznań et coorganisée avec Eurostat. Au total, elle a été suivie par 169 participants provenant de 32 pays, de 10 organisations internationales et de 16 universités. Trente participants ont pu y assister en personne. Trente-cinq contributions innovantes et les travaux menés dans le cadre des deux projets du Groupe de haut niveau pour 2021 ont été présentés lors de 6 sessions. Les idées d'avenir qui ont reçu le plus de soutien sont les suivantes : « Contrôle automatisé des sorties : rêve ou réalité ? », « Comment rattacher les paramètres formels de mesure du risque à la perception du risque », « Réexaminer les règles pour les tableaux de fréquence » et « Suppression ou perturbation ? ». Tous les résumés, documents, présentations et autres résultats de l'atelier sont disponibles sur le site Web de la CEE (<https://unece.org/statistics/events/DC2021>) ;

V. Autres activités et résultats

30. **Publications** : la plupart des résultats des travaux réalisés dans le cadre du Groupe de haut niveau sont partagés publiquement sur les wikis du Groupe de haut niveau et sur le site Web de la CEE. En 2021, deux publications officielles de la CEE relatives aux travaux de modernisation ont été élaborées. Les résultats du projet de cadre de communication stratégique ont été regroupés dans une publication imprimée⁴, tandis que les résultats du projet d'apprentissage automatique ont été compilés dans une publication électronique⁵.

31. Les travaux du Groupe de haut niveau ont été facilités par un grand nombre de sites wiki ([wiki sites](#)) et de pages Web ([web pages](#)). Au total, le secrétariat administre et actualise plus de 50 sites wiki à accès public ou restreint, qui ont été créés à des fins collaboratives ou pour diffuser les résultats des activités du Groupe de haut niveau. Divers articles de presse et rapports destinés à des organismes intergouvernementaux ont été rédigés (par exemple, pour la Conférence des statisticiens européens ou le Secrétariat de l'ONU). Le secrétariat du Groupe de haut niveau a également répondu à diverses demandes de publications, d'informations, de présentation de travaux ou de demandes d'assistance émanant de collègues de la région de la CEE et d'ailleurs. Les travaux du Groupe de haut niveau ont été présentés à plusieurs reprises lors de manifestations nationales ou internationales.

32. Les travaux ont également été coordonnés avec d'autres organisations internationales s'occupant de la modernisation des statistiques officielles (par exemple Eurostat, l'OCDE et la Division de statistique du Département des affaires économiques et sociales de l'ONU). Des liens ont été établis avec des activités internationales menées dans des domaines similaires grâce à la participation croisée de membres des différents groupes.

33. On trouvera de plus amples informations sur le portail en ligne principal du Groupe de haut niveau ([main HLG-MOS online portal](#)). Le secrétariat de la CEE ([UNECE Secretariat](#)) peut également être contacté à cette fin.

⁴ <https://unece.org/info/Statistics/pub/357507>

⁵ <https://unece.org/statistics/publications/machine-learning-official-statistics>