Nations Unies ECE/CES/2023/9

Conseil économique et social

Distr. générale
11 avril 2023
Français
Original : anglais

Commission économique pour l'Europe

Conférence des statisticiens européens

Soixante et onzième réunion plénière
Genève, 22-23 juin 2023
Point 2 de l'ordre du jour provisoire
Travaux du Groupe de haut niveau sur la modernisation de la statistique officielle

Rapport sur les travaux du Groupe de haut niveau sur la modernisation de la statistique officielle

Document élaboré par le secrétariat et la Commission de contrôle du Groupe de haut niveau sur la modernisation de la statistique officielle

Résumé

Le présent document récapitule les résultats des travaux menés en 2022 par le Groupe de haut niveau sur la modernisation de la statistique officielle (ci-après « le Groupe de haut niveau »). L'additif 1 au présent document contient le mandat actualisé relatif à la suite des travaux. Dans l'additif 2 sont présentés les travaux prévus pour 2023 dans le cadre du Groupe de haut niveau.

À la réunion de février 2023, le Bureau de la Conférence des statisticiens européens a examiné et approuvé les résultats des travaux menés en 2022 dans le cadre du Groupe de haut niveau, le mandat actualisé et les activités prévues pour 2023. La Conférence des statisticiens européens est invitée à approuver le rapport et le mandat actualisé du Groupe de haut niveau et à prendre note des travaux prévus dans le domaine de la modernisation de la statistique.

GE.23-06649 (F) 280423 010523

Merci de recycler

I. Introduction

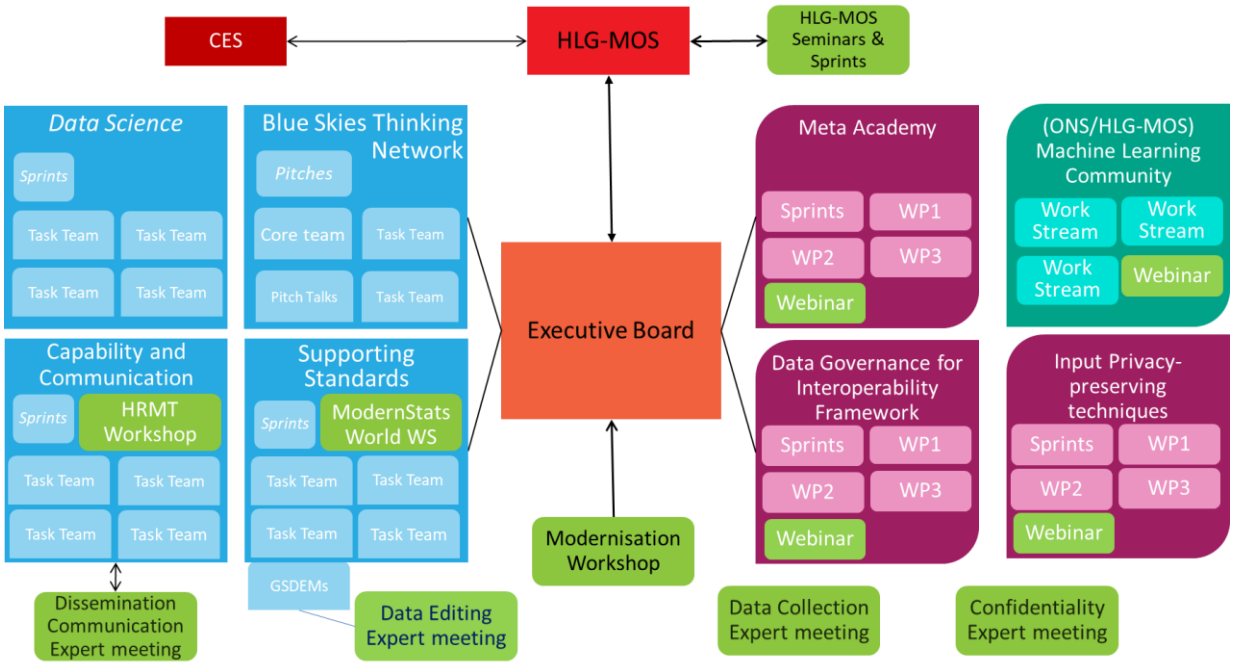
1. Le présent document récapitule les résultats des travaux menés en 2022 dans le cadre du programme de travail du Groupe de haut niveau sur la modernisation de la statistique officielle.

2. Le Groupe de haut niveau 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 13 organisations qui en fixent la stratégie, la mission et les thèmes prioritaires et en approuvent le programme de travail. 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 d'accès public du Groupe de haut niveau.

3. En 2022, trois projets, quatre groupes d'experts (appuyés par des équipes spéciales supplémentaires), un réseau de praticiens et plusieurs réunions et ateliers d'experts organisés par les comités directeurs étaient prévus au programme de travail du Groupe de haut niveau. Un aperçu schématique des groupes, réseaux, projets et ateliers opérationnels en 2022 dans le cadre du Groupe de haut niveau est présenté dans la figure ci-dessous.

4. Au total, les groupes, les équipes spéciales, les comités directeurs et les projets sur la modernisation ont réuni plus de 650 membres issus de plus d'une centaine d'organisations différentes. Ce nombre est nettement supérieur à celui de 2021, principalement en raison de l'inclusion de plus de 350 participants de la communauté de l'apprentissage automatique. Plus de 750 collègues ont participé aux réunions d'experts et aux ateliers, soit moins qu'en 2021, plusieurs réunions s'étant à nouveau tenues en présentiel.

Figure 1
Structure du Groupe de haut niveau sur la modernisation de la statistique officielle en 2022



A. Groupe de haut niveau sur la modernisation de la statistique officielle

5. Le Groupe de haut niveau sur la modernisation de la statistique officielle met régulièrement à jour ses domaines prioritaires dans le droit fil de sa stratégie et de sa mission. La pandémie de COVID-19 a accéléré l'évolution des rôles et des méthodes de travail des organismes nationaux de statistique. Le Groupe de haut niveau a organisé en mars 2022 une « séance éclair » afin d'examiner comment il pouvait axer ses travaux sur la poursuite de l'accélération de l'innovation observée pendant la pandémie de COVID-19. Il en a conclu

que l'accent devrait être mis davantage sur les services statistiques de pointe (par exemple, liés à la gestion des données) que sur la production statistique « classique ». Il était essentiel de bien comprendre ce que les parties prenantes attendaient du Groupe de haut niveau (centrage sur l'utilisateur). Il a également été souligné que l'effet mobilisateur des activités du Groupe de haut niveau en tant que plateforme de collaboration ne devait pas être sous-estimé. Au cours de la pandémie, les activités bilatérales et multilatérales ainsi menées avaient permis de partager des pratiques et des solutions novatrices. Le Groupe de haut niveau devait en outre se pencher sur d'autres points clés soulevés dans le cadre de la séance éclair :

a) Préciser davantage le cadre de modernisation prévu ;

b) Évaluer les résultats du Groupe de haut niveau et les faire connaître ;

c) Appuyer l'amélioration continue des activités de statistique (innovation dans les méthodes de travail).

6. La pandémie de COVID-19 a mis en évidence le besoin de renforcer les capacités et de changer la culture institutionnelle pour mieux répondre aux attentes des utilisateurs. Le Groupe de haut niveau doit donc continuer à favoriser le changement de culture, le passage de l'analyse des données à l'ingénierie des données, ainsi qu'à renforcer sa capacité d'attirer et de fidéliser de nouveaux collaborateurs. Il a l'élan nécessaire pour continuer à définir des orientations concernant les services et les solutions créatives.

7. Le Groupe de haut niveau a bénéficié d'un fort appui et d'une grande reconnaissance pour son travail fondamental d'innovation et de modernisation dans le domaine de la statistique officielle. Il devrait continuer à appuyer la mise en œuvre des résultats obtenus (normes, résultats de projets, etc.). Son rôle clé pour faciliter la transformation a été mis en exergue. La Commission de contrôle avait un rôle central à jouer à cet égard. On ne doutait pas qu'elle alignerait à l'avenir ses activités sur la stratégie et la mission du Groupe de haut niveau, axées sur le soutien à l'innovation. Il a été souligné que pour continuer sur la voie de la réussite, les membres de la Conférence des statisticiens européens devaient contribuer aux travaux. La Commission de contrôle s'est engagée à chercher de nouvelles collaborations au-delà des organismes nationaux de statistique, par exemple avec d'autres organismes internationaux ainsi qu'avec le monde universitaire et le secteur privé, en vue de faire avancer le programme de modernisation du Groupe de haut niveau.

B. Commission de contrôle

8. Dans le droit fil des résultats de la « séance éclair » du Groupe de haut niveau décrite ci-dessus, la Commission de contrôle réorganisera et mettra à jour le cadre de modernisation. Elle veillera à :

a) Redéfinir de manière continue le niveau des objectifs et les domaines sur lesquels concentrer les efforts ;

b) Donner la priorité aux processus statistiques et aux services nouveaux les plus prometteurs ;

c) Passer de la production de statistiques à la fourniture de services ;

d) Donner autant d'importance aux méthodes, aux normes et aux technologies qu'aux personnes (personnel, utilisateurs, parties prenantes, fournisseurs de données) dans les évolutions futures ;

e) Renforcer l'échange et la collaboration pour avancer et obtenir des résultats plus rapidement ;

f) Optimiser encore les réunions d'experts ;

g) Rechercher des experts au-delà des organismes nationaux de statistiques dans les domaines d'intérêt de la modernisation.

9. 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. Après quelques années passées sur fond de pandémie de COVID 19 et de pénurie de personnel à la Commission économique pour l'Europe (CEE), un programme de travail complet a été établi. De plus, un projet supplémentaire a été inclus. Le cycle des réunions d'experts et des ateliers a été porté à vingt-quatre mois, avec la possibilité d'organiser des manifestations en ligne entre ceux-ci. Le Groupe sur l'application des méthodes ModernStats a démarré ses activités en 2022.

10. 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 visant en particulier à 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 dans le cadre des projets et des groupes, correspondent à la mission, à la stratégie et aux thèmes prioritaires retenus par le Groupe de haut niveau. La Commission est chargée d'organiser, en collaboration avec le secrétariat de la CEE, l'atelier sur la modernisation dirigé par le 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

11. Pour 2022, trois projets ont été choisis : le projet relatif aux techniques de préservation de la confidentialité des données d'entrée (prolongé), le projet de méta-académie pour la modernisation de la statistique officielle et le projet de cadre de gouvernance pour l'interopérabilité des données.

A. Projet de préservation de la confidentialité des données d'entrée

12. Le projet relatif à la préservation de la confidentialité des données d'entrée sera poursuivi en 2021. Statistique Pays-Bas a nommé Dennis Ramondt comme chef de projet à temps partiel. Au total, 23 collègues de sept 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 « séances éclair » en présentiel, le projet a bien avancé.

13. 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 en recherche analytique ;

b) L'apprentissage automatique privé ;

c) Une enquête publique consultative.

14. 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.

15. 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 logiciels libres 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.

16. La consultation technique informelle et ouverte visait principalement à déterminer les difficultés de la mise en œuvre pratique d'une infrastructure de services informatiques privés sécurisés multipartites digne de confiance aux fins de la statistique officielle (la consultation était ouverte jusqu'à la fin du mois de novembre 2022). Elle s'adressait aux experts et aux parties prenantes, tels que les spécialistes de la protection de la vie privée et de la sécurité, les utilisateurs potentiels, les cybermilitants et les représentants de la société civile, ainsi que les chercheurs et concepteurs. L'équipe du projet résumera les principaux résultats de la consultation dans un rapport public qui sera mis en ligne sur la page du projet.

17. Les résultats ont été communiqués lors d'un séminaire qui a été bien accueilli et l'élaboration d'une enquête consultative publique sur l'informatique privée multipartite en tant que service a été lancée. Les exposés sont accessibles sur le site Web de la réunion : https://statswiki.unece.org/x/FIECFg. Le projet s'est achevé à la fin de l'année 2022. Un rapport final sera publié.

B. Méta-académie pour la modernisation de la statistique officielle

18. Le projet de méta-académie pour la modernisation de la statistique officielle était dirigé par Kate Burnett-Isaacs, de Statistique Canada, et Eric Anvar, de l'OCDE. Vingt-neuf collaborateurs de treize institutions y ont participé. Ils ont régulièrement tenu des réunions vidéo. Ces réunions ont consisté notamment à présenter plusieurs travaux théoriques et pratiques réalisés par les membres. Un séminaire lié à la méta-académie et au modèle Carpentries s'est tenu parallèlement à l'atelier organisé par le Groupe de haut niveau, au cours duquel un exposé sur les éléments de base de Git et de la gestion des versions a été fait à l'intention des organismes nationaux de statistique et où des méthodes d'intégration d'une formation générique ont été définies et les obstacles à la formation de leur personnel mis en évidence.

19. Les responsables du projet ont étudié les possibilités pour les instituts de statistique d'élaborer un cadre de conception en commun de formations de haute qualité et d'en diffuser le contenu au niveau international. Il a été conclu que des obstacles importants subsistaient :

a) Les besoins en formation ne sont pas internationalement définis et il n'y a pas de méthode commune de création de contenus d'apprentissage ;

b) Il n'existe aucun forum ni aucune communauté qui permette aux « gestionnaires d'académie » ou aux « formateurs » de mettre en commun leurs pratiques et d'apprendre les uns des autres, de partager des priorités ou encore de s'investir collectivement dans des activités de formation sur un sujet particulier, en s'adaptant en permanence face aux évolutions du secteur et aux nouveaux besoins.

20. L'étude des formations existantes aux niveaux national et international a permis de mieux définir les contours d'un cadre commun permettant aux instituts de statistique de collaborer à la modernisation des organismes nationaux de statistique grâce à la mise à niveau des compétences de leurs employés. Ce projet exploratoire a mis en évidence plusieurs possibilités, les plus importantes étant les suivantes :

a) Pédagogie : Les méthodes d'auto-apprentissage et l'apprentissage simultané avec un formateur (en présentiel ou en ligne) sont complémentaires. L'auto-apprentissage (y compris les certifications) peut être une base intéressante pour tirer le meilleur parti de la formation simultanée. Cela étant, en particulier dans le cas de compétences plus élevées, la formation simultanée, qui permet de passer du temps avec son formateur et de créer des liens, est essentielle ;

b) Contenu : Certains projets suivent une approche large couvrant toutes les activités régulières de la statistique officielle (y compris les normes statistiques) ; d'autres se concentrent sur la science des données ou les activités scientifiques à forte intensité de données, ces derniers pouvant être des projets issus du marché ou extérieurs à la statistique officielle ;

c) Public : Certaines initiatives ont un objectif purement national (système statistique national ou administration publique), tandis que d'autres visent à créer un vaste réseau international (par exemple, dans le cadre d'un programme d'échange de données à des fins de développement). Cependant, le multilinguisme reste une difficulté dans nombre d'initiatives ;

d) Plateforme : Les initiatives s'articulent toutes autour d'une plateforme numérique, qui permet de créer et d'échanger des contenus et des services de formation et de promouvoir une marque afin de fédérer une communauté de formateurs et d'apprenants. Une condition essentielle à la réussite est de ne pas se contenter d'exposés, mais d'inclure plusieurs supports (notamment des échantillons de données, des algorithmes et des carnets de notes, des maquettes et des scénarios de formation, des questionnaires de certification, etc.)

21. Le modèle le plus prometteur recensé, intitulé « The Carpentries », met l'accent sur l'apprentissage simultané dans le domaine des compétences liées à la science des données. Un tel modèle appliqué aux méthodes ModernStats pourrait ainsi permettre aux organismes nationaux de statistique de mettre en place un écosystème national de formateurs/d'apprenants en utilisant une plateforme existante (afin de réduire au minimum les coûts informatiques). Dans le modèle « Carpentries », tous les éléments mentionnés ci-dessus sont mis en commun en tant qu'actifs ouverts sous licence Creative Commons dans GitHub, réutilisables à volonté dans d'autres contextes. Ils pourraient également être gérés sous la marque et la gouvernance de ModernStats. Une proposition de projet a donc été élaborée en vue de lancer et de développer le projet ModernStats Carpentry.

C. Projet de cadre de gouvernance pour l'interopérabilité des données

22. Le projet de cadre de gouvernance pour l'interopérabilité des données est dirigé par Juan Muñoz, de l'INEGI (Mexique). Des réunions régulières ont eu lieu entre les vingt participants issus de treize organisations. L'équipe se compose d'experts venus de milieux différents et ayant des points de vue variés. Un travail important a été réalisé pour déterminer la portée du projet et les principaux éléments à prendre en compte dans son cadre. Divers concepts de base relatifs à l'interopérabilité des données dans le contexte des instituts de statistique ont été définis. La structure de la gouvernance pour l'interopérabilité des données statistiques a été fixée au terme d'un nouvel examen de la portée du projet et de l'objectif à atteindre et après avoir déterminé quels produits présenteraient la plus grande valeur ajoutée.

23. Le cadre consiste en un système de règles et de théories contribuant à harmoniser les travaux, à unifier les concepts et à organiser les éléments de base nécessaires à la création et à la gestion d'une plateforme de données, de métadonnées et de systèmes interopérables. Il servira de structure à un document décrivant le cadre de référence qui rassemble les principaux éléments nécessaires à l'interopérabilité des données. Ce cadre guidera le lecteur au long d'un processus simple de mise en place d'un système de gouvernance permettant d'assurer l'interopérabilité des données au sein de l'organisation considérée. Le document vise à exposer les raisons pour lesquelles un cadre d'interopérabilité est nécessaire, les objectifs visés et la manière de les atteindre.

24. Les membres de l'équipe ont commencé à créer du contenu en vue d'alimenter les différentes sections convenues. Toutefois, en raison de la complexité du projet et des ressources limitées dont elle dispose, l'équipe a élaboré une proposition visant à prolonger le projet jusqu'en 2023.

D. Projet sur les données synthétiques (activités de suivi)

25. Le projet sur les données synthétiques s'est achevé en 2021, mais d'importantes activités de suivi ont été menées en 2022 sous la direction de Statistique Canada. Ces activités comprenaient un hackathon organisé en janvier 2022 afin de mettre à l'essai les lignes directrices élaborées dans le cadre du projet sur les données synthétiques. Ces lignes directrices ont ensuite été publiées sous la forme d'un rapport finalisé en décembre 2022. La publication, intitulée *Synthetic Data for Official Statistics: A Starter Guide (*Données synthétiques pour les statistiques officielles : Guide de démarrage), sera publiée en janvier 2023.

III. Groupes de modernisation

26. En 2022, le Groupe de haut niveau était structuré en quatre groupes : le Groupe d'appui à la mise en œuvre des normes, le Groupe des capacités et de la communication, le Réseau de recherche fondamentale et, depuis 2022, le Groupe de l'application de la science des données et des méthodes modernes. Dans le cadre de chaque groupe, différentes équipes spéciales ont été mises en place pour mener les activités définies dans le cadre du programme de travail pour 2022. Chaque groupe et les équipes spéciales se réunissent régulièrement, en général au moins une fois par mois. Les groupes coordonnent leurs activités entre eux, mais aussi avec les projets du Groupe de haut niveau 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 et se concertent avec elle.

A. Groupe de l'application de la science des données et des méthodes modernes

27. Le Groupe de l'application de la science des données et des méthodes modernes a été mis en place au début de l'année 2022. En raison de l'importance croissante des nouvelles sources de données et des méthodes d'établissement des statistiques officielles, il vise à dépasser les cadres conceptuels de la science des données et des méthodes modernes et à trouver des moyens concrets de moderniser les processus opérationnels des organismes nationaux de statistique.

28. Le Groupe, qui se compose de 21 collaborateurs issus de 10 organisations, a analysé le marché pour dresser un état des lieux des travaux existants dans le domaine de la science des données et des méthodes modernes et définir le champ de ses activités. Au total, huit thèmes potentiels ont été retenus, notamment la collecte et l'intégration des données, l'édition des données, la confidentialité des données et la diffusion des données.

29. En fonction de l'expertise et des centres d'intérêt des membres, trois thèmes ont été classés par ordre de priorité et des équipes spéciales ont été créées. Le Groupe s'est réuni une fois par mois et les équipes spéciales se sont réunies mensuellement ou, si nécessaire, plus fréquemment.

a) **Équipe spéciale 1 − Compréhension et choix des modèles :** Il existe un large éventail de méthodes modernes et de modèles d'apprentissage automatique qui évoluent rapidement, mais les informations sont fragmentées et les orientations générales sont rares. Cette équipe spéciale a donc pour objet de regrouper les connaissances des organismes nationaux de statistique sous forme d'orientations visant à donner aux utilisateurs une compréhension globale de la manière dont les différentes méthodes et les différents modèles s'adaptent à différents types de problèmes, de données et d'objectifs et à faciliter ainsi le choix des solutions les plus adéquates ;

b) **Équipe spéciale 2 − Accélération de la mise en œuvre de solutions fondées sur l'apprentissage automatique pour l'édition de données** : L'apprentissage automatique est très prometteur pour ce qui est d'améliorer l'efficacité, en complément ou en remplacement des méthodes traditionnelles, ainsi que pour ce qui est d'améliorer la qualité d'une manière difficilement réalisable avec les moyens traditionnels. Cette équipe spéciale a

pour tâche d'élaborer des orientations génériques sur les questions essentielles d'organisation liées à la mise en œuvre de solutions fondées sur l'apprentissage automatique dans le domaine de l'édition de données, afin d'accélérer cette mise en œuvre ;

c) **Équipe spéciale 3 − Cadre international sur l'intelligence artificielle (IA) responsable au service de la statistique officielle :** Avec l'utilisation et l'adoption à grande échelle de l'IA et de l'apprentissage automatique, les principes fondamentaux que sont l'éthique, la protection de la vie privée, l'équité et la légalité sont aujourd'hui plus importants que jamais. L'objectif de cette équipe spéciale est de mettre en place une norme commune de conception, de mise au point et de déploiement responsables de solutions fondées sur l'IA dans les organismes nationaux de statistique, en veillant à ce que ces nouvelles méthodes soient acceptables du point de vue éthique et humain.

30. Ces trois thèmes ont été présentés au cours de la manifestation internationale de collaboration du Data Science Campus de l'Office national de statistique (ONS) britannique, qui s'est tenue à Newport (Royaume-Uni) du 12 au 14 juillet 2022. Ils ont suscité un fort appui et un grand intérêt de la part des participants.

B. Réseau de recherche fondamentale

31. Le Réseau de recherche fondamentale est le laboratoire d'idées de la communauté de modernisation de la 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.

32. Le Réseau est dirigé par Barteld Braaksma, responsable de l'innovation à Statistique Pays-Bas, et est appuyé par la CEE. Le groupe central d'experts compte une quinzaine de membres qui se réunissent au moins une fois par mois.

33. Le Réseau de recherche fondamentale a examiné plusieurs brèves présentations au cours de l'année. Ces présentations, sous forme de « séances éclair », permettent aux membres du réseau et à des personnes extérieures d'exposer brièvement une idée ou un projet de modernisation. Le groupe central et les experts invités ont fait des propositions sur la manière dont ces présentations pourraient être affinées ou étoffées. Il en est résulté une proposition d'activité sur le partage des connaissances pour le passage aux logiciels libres dans les organismes nationaux de statistique, ainsi que quatre propositions de projets au titre du programme de travail du Groupe de haut niveau pour 2023 :

a) Prolongation du projet de cadre de gouvernance pour l'interopérabilité des données ;

b) ModernStats Carpentries (phase 2 du projet de méta-académie) (OCDE/Canada) ;

c) Automatisation robotisée des processus dans les statistiques officielles (Hongrie) ;

d) Informatique en nuage au service de la statistique officielle (Irlande).

34. D'autres activités ont été menées à bien par le réseau en 2022 :

a) **Informatique en nuage au service de la statistique officielle** : Le 29 juin, le Réseau de recherche fondamentale a organisé à Belgrade (Serbie) une séance sur l'utilisation de l'informatique en nuage dans la statistique officielle, au cours de laquelle a été mise en lumière la demande potentielle de collaboration internationale dans ce domaine. Cette séance

a permis de s'accorder sur les aspects de l'adoption et de l'utilisation de l'informatique en nuage qui bénéficieraient le plus d'un projet. Elle a été suivie d'une autre séance sur le sujet qui s'est tenue le 12 juillet à Newport (Royaume-Uni) lors de la manifestation internationale de collaboration du Data Science Campus de l'ONS. Celle-ci a contribué à mieux cerner les domaines que les projets devraient couvrir et l'Irlande s'est portée volontaire pour diriger l'élaboration de la proposition de projet qui a été présentée à l'atelier du Groupe de haut niveau ;

b) **Enquêtes rapides** : Une séance a été organisée au cours de la manifestation de juillet deux Newport afin de mieux déterminer les possibilités de réaliser des enquêtes rapides. Différentes expériences menées au Canada, en Irlande, en Serbie et aux Pays-Bas, comme des enquêtes de confiance auprès des entreprises, des enquêtes sur l'impact de la COVID-19 et des enquêtes sur la perception par le public, ont été présentées et discutées. Les participants ont proposé de mettre en commun les données des enquêtes plurinationales dans un système en nuage afin de faciliter la comparaison des résultats. Un projet d'activité du Réseau de recherche fondamentale a été élaboré sur la base des exemples présentés d'enquêtes non probabilistes, menées sur une plateforme de médias sociaux ;

c) **Jumeaux numériques** : Les travaux dans ce domaine se sont intensifiés dans le cadre d'une autre « séance éclair » organisée en marge de la manifestation de Newport. Plusieurs applications des technologies de télédétection fixées à l'infrastructure ont été présentées. Les participants ont examiné la possibilité d'intégrer des jumeaux numériques distincts liés aux infrastructures dans un écosystème de partage de données afin d'améliorer l'accès aux informations et la prise de décisions sur des sujets tels que la surveillance de l'environnement, l'utilisation des ressources, ou encore l'interaction et les effets d'entraînement entre les différents types d'infrastructures en cas de perturbation des systèmes. Ces travaux ont abouti à une proposition d'activité du Réseau de recherche fondamentale pour 2023.

35. Le groupe se tient prêt à évaluer les propositions d'activité et de projet soumises par la communauté statistique. De plus amples informations figurent sur la plateforme wiki du Réseau de recherche fondamentale.

C. Groupe des capacités et de la communication

36. 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 2022 sur les conséquences de la pandémie.

37. En 2022, 52 collègues de 22 organisations ont participé au Groupe et à ses équipes spéciales et sous-groupes. Le Groupe se réunit une fois par mois et les différents groupes de travail et sous-groupes se réunissent mensuellement ou, si nécessaire, plus fréquemment. Il est dirigé par Anna Borowska (Statistique Pologne) et Maria Hurley (CSO Irlande). Chaque équipe spéciale a un président ou une présidente. Les activités portent sur trois domaines :

a) **Gestion de l'éthique** : En plus des deux enquêtes menées en 2021, une autre enquête complémentaire a été envoyée aux personnes ayant répondu aux enquêtes précédentes. Celle-ci portait sur l'éthique des affaires, c'est-à-dire l'éthique en tant qu'élément transversal dans l'ensemble de l'organisation et en tant que moyen essentiel d'améliorer la qualité des résultats. Les résultats préliminaires des enquêtes ont été présentés lors de l'atelier et peuvent être consultés à l'adresse suivante : https://unece.org/statistics/events/HRMT2022 ;

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 des lignes directrices sur la gestion de la marque,

le marketing et la communication à l'intention des organismes de statistique, qui peuvent être consultées sur la page wiki consacrée à la gestion de la marque et de la réputation : https://statswiki.unece.org/x/cRGxEw ;

c) **Travaux à venir, lieux de travail de demain et compétences futures :** L'équipe de l'avenir de l'emploi s'est concentrée sur trois domaines : les panoplies d'outils pour assurer la mixité de l'environnement de travail, à l'intention des employeurs, des employés et des cadres ; les jeunes (en tant que futurs employés, utilisateurs des statistiques officielles et source de données pour la statistique officielle) ; le futur de l'emploi. L'équipe a envoyé un questionnaire aux pays, portant sur des questions telles que le bien-être et l'environnement de travail mixte, la définition du travail mixte dans les différents pays, l'évolution de l'apprentissage et de la formation, etc. Les résultats du questionnaire et des travaux des équipes peuvent être consultés sur la page wiki « Human resources and more... » : https://statswiki.unece.org/x/UgAMDQ.

38. Le Groupe a également été chargé d'organiser l'**atelier sur la gestion des ressources humaines et la formation** (11-13 octobre, Bruxelles (Belgique)). Cet atelier, organisé conjointement avec l'Association européenne de libre-échange (AELE) et la CEE, a réuni 60 participants de 28 pays et organisations. L'atelier comprenait quatre sessions : i) POURQUOI devons-nous repenser le travail, le personnel et les lieux de travail des organismes nationaux de statistique ? ii) QU'avons-nous essayé, qu'avons-nous appris des récentes initiatives de réponse aux crises ? iii) COMMENT créer les organismes nationaux de statistique de demain ? a) Organisation, culture et éthique ; b) Personnel ; c) Formation et compétences ; iv) Et ensuite ? L'atelier comportait également une session spéciale à l'intention des pays partenaires de l'AELE dans l'après-midi du 13 octobre. Les exposés et tous les documents de la réunion sont disponibles sur le site Web de la CEE : https://unece.org/statistics/events/HRMT2022.

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

39. Le Groupe d'appui à la mise en œuvre des normes est chargé de mettre à jour et de mettre au point les modèles ModernStats. Il s'agit de modèles de modernisation et de normes élaborés par le Groupe de haut niveau, tels que le GAMSO, le GSBPM, le GSIM et la CSPA[1]. 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.

40. Le Groupe, présidé par Zoltán Vereczkei (Hongrie), bénéficie de l'appui du secrétariat de la CEE. Outre les réunions plénières mensuelles du groupe d'experts, les équipes spéciales se sont réunies fréquemment. Chaque équipe spéciale est présidée par un responsable. 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.

41. Les principales activités et les principaux résultats du Groupe et de ses équipes spéciales en 2022 ont été les suivants :

a) **Révision du GSIM :** La mise à jour progressive initialement prévue a évolué vers un examen complet en 2022. Les commentaires de la communauté ont été recueillis concernant les améliorations possibles du GSIM. Une relation étroite lie le GSIM au travail de l'Ontologie de base de la statistique officielle et les questions d'intérêt commun ont donné lieu à des échanges d'idées. Pour plus de détails, consulter le groupe GitHub à l'adresse suivante : https://github.com/UNECE/GSIMRevision/. L'activité se terminera au premier trimestre 2023, avant un examen public et la parution d'une publication sur le sujet. La publication des résultats finaux est prévue pour juin 2023 ;

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

b) **Ontologie de base pour la statistique officielle** : Les travaux visent à élaborer un modèle d'intégration de l'ensemble des normes de ModernStats en s'appuyant sur des éléments du vocabulaire normalisé bien connu. Cet outil définit un cadre d'intégration conceptuel dont le but est d'assurer la cohérence sémantique des différents modèles en proposant une nomenclature commune, des définitions et un ensemble de représentations des relations inter-modèles et intra-modèle, établies au moyen de l'outil Resource Description Framework (RDF), du langage Web Ontology Language (OWL) et de divers vocabulaires de normes numériques (SKOS, PROV, DCAT, DC, ORG, etc.). La nouvelle version de l'Ontologie de base pour la statistique officielle était prête et un examen public s'est achevé le 25 novembre. De plus amples informations sont disponibles sur GitHub : https://linked-statistics.github.io/COOS/coos.html. Conformément au plan d'action, les activités se poursuivront en 2023 ;

c) **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. Sur la base d'exemples provenant de huit pays, un niveau plus fin d'activités du GSBPM a été mis au point. Ce niveau pourrait faciliter la mise en œuvre du modèle dans d'autres pays. Les résultats et d'autres informations sont disponibles sur la page wiki du GSBPM : https://statswiki.unece.org/x/c4OuFQ ;

d) **SDMX/DDI/GSBPM** : les travaux visent à fournir une description de chaque sous-processus GSBPM du point de vue du SDMX et de la DDI. Les utilisateurs pourront ainsi comprendre comment le SDMX et la DDI contribuent à la mise en œuvre des éléments logiques décrits dans le GSBPM et le GSIM (à partir de résultats de la précédente équipe spéciale sur l'établissement d'un lien entre le GSBPM et le GSIM). Les activités se poursuivront tout au long de l'année 2023.

42. **Atelier mondial ModernStats 2022** : cet atelier a été accueilli par l'organisme national serbe de statistique du 27 au 29 juin à Belgrade. Les utilisateurs des modèles ModernStats ont pu y faire part de leurs expériences et des enseignements qu'ils en ont tirés et ont pu avancer dans la mise au point et l'actualisation de ces modèles. Les débats et les activités de groupe ont été très animés et de nombreuses contributions ont été reçues de la communauté sous différentes formes (commentaires directs, discussions de groupe, séances de remue-méninges, clinique ModernStats, etc.). Les conclusions de l'atelier (principalement des suggestions concernant les travaux futurs, les difficultés et les besoins) font l'objet d'un suivi par le Groupe d'appui à la mise en œuvre des normes. Les documents de l'atelier, y compris le rapport de la réunion, sont disponibles à l'adresse suivante : https://unece.org/statistics/events/MWW2022.

43. En raison du manque de ressources et de la nécessité pour le Groupe de se concentrer d'abord sur d'autres activités, les travaux portant sur les processus généraux du GSBPM et le renforcement des capacités de l'Architecture commune de la production statistique n'ont pas commencé en 2022. Ces deux activités ont été reportées à 2023.

E. Communauté de l'apprentissage automatique

44. Le Groupe de l'apprentissage automatique ONS/CEE est une plateforme internationale de collaboration dans les domaines de la recherche, de l'échange de connaissances et du renforcement des capacités en matière d'apprentissage automatique au service de la statistique officielle. Il rassemble des statisticiens, des scientifiques des données et des académiciens de 45 pays qui étudient comment l'apprentissage automatique peut améliorer la production statistique et comment intégrer efficacement cet apprentissage dans la production. Le Groupe a été créé en tant que communauté de praticiens assurant le suivi du projet 2019-2020 sur l'apprentissage automatique du Groupe de haut niveau, qui avait été une réussite. La communauté est dirigée par le Data Science Campus de l'ONS avec l'appui de la CEE. Elle vise notamment à créer une plateforme qui facilitera la création, la mise au point et la mise en œuvre de projets de recherche et d'activités de renforcement des compétences répondant aux besoins de la communauté statistique mondiale. En outre, elle a pour mission de constituer et de mobiliser une forte communauté de l'apprentissage

automatique en mettant en commun les ressources et les bonnes pratiques, en échangeant des idées et des données d'expérience et en se tenant au courant des évolutions dans ce domaine. Elle met à la disposition de la communauté des ressources en libre accès, partageables et facilement accessibles, et facilite le renforcement des capacités dans le domaine de l'apprentissage automatique au service de la statistique officielle.

45. Plus de 400 collègues de 45 pays ont participé à la communauté. Il y a une forte demande de partage des connaissances et de renforcement des capacités et une large participation aux réunions. Les travaux s'articulent autour de cinq domaines de travail devant permettre d'aborder les différentes questions qui se posent en matière d'apprentissage automatique au service de la statistique officielle :

46. En 2022, les travaux se sont concentrés sur le passage de la validation de principe à la production. Au terme de 34 propositions d'activités et de discussions approfondies, des groupes thématiques ont été constitués autour des points suivants :

a) Classification des textes ;

b) Analyse de l'imagerie ;

c) Données du système d'identification automatique (AIS) ;

d) Modélisation ;

e) Moissonnage de données sur le Web ;

f) Qualité des données de formation ;

g) Reconfiguration des modèles ;

h) Infrastructure informatique.

47. Les autres domaines clés portaient sur l'éthique, la qualité des données de formation, la collaboration dans le domaine de la recherche et le renforcement des capacités. Des séances éclair ont contribué à faire avancer les travaux, notamment celles que le Data Science Campus a organisées au Royaume-Uni dans le cadre de la manifestation internationale tenue à Newport du 12 au 14 juillet. Le renforcement des capacités a fait l'objet d'activités de type « café et codage » et donné lieu à la création de divers groupes thématiques. Des échanges d'idées sur les travaux ont eu lieu dans le cadre de diverses activités de communication et de sensibilisation, telles que des sites web et wiki, des forums de discussion, des guides, des articles, une chaîne YouTube, divers exposés et un webinaire (le 30 novembre).

48. Pour une vue d'ensemble de toutes les réalisations de la communauté de l'apprentissage automatique, on trouvera le rapport final de 2022 du Groupe de l'apprentissage automatique ONS/CEE, ainsi que tous les rapports des groupes thématiques, les documents présentés lors des réunions mensuelles, les réunions « café et codage » et les « séances éclair » organisées à Newport, sur la page wiki consacrée à l'apprentissage automatique : https://statswiki.unece.org/x/34AqF.

IV. Réunions et ateliers

49. En réponse aux appels à l'action pour lutter contre les changements climatiques et faciliter l'accès à un public plus large tous les deux ans, la Commission de contrôle et le secrétariat de la CEE ont décidé de modifier le cycle des réunions d'experts et des ateliers en présentiel pour le porter à vingt-quatre mois. Entre ces cycles, des réunions d'experts et des ateliers en ligne peuvent être organisés. En 2022, les réunions d'experts et les ateliers suivants ont été organisés (par ordre chronologique) :

a) **Réunion du Groupe de haut niveau sur la modernisation de la statistique officielle** (10 mars 2022, en ligne) : La réunion habituelle du Groupe de haut niveau le week-end précédant la session de la Commission de statistique à New York n'ayant pas pu se tenir en raison des restrictions liées à la COVID-19, les membres du Groupe se sont réunis en ligne. Le programme de travail a été approuvé et les statisticiens en chef se sont félicités des résultats obtenus en 2021. La « séance éclair » a permis de remanier le cadre de la modernisation et d'actualiser les thèmes prioritaires. 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 stratégie et aux nouvelles priorités du Groupe de haut niveau ;

b) **Réunion d'experts sur la diffusion et la communication des statistiques** (13-15 septembre, en ligne) : La réunion a rassemblé 156 participants issus de 36 organismes nationaux de statistique, de neuf organisations internationales et du milieu universitaire. Elle a donné lieu à 22 exposés répartis en trois séances sur les thèmes suivants : i Marque, réputation, confiance − Qui sommes-nous, que représentons-nous, quelle est notre place dans la société ? ; ii) Études de marché, stratégies de commercialisation et de communication en ligne ; iii) Vulgarisation de la statistique. Des discussions en séance plénière ont eu lieu à la fin de la réunion pour faire le point sur les enseignements tirés et les thèmes à aborder. Les sujets évoqués et tous les documents de la réunion sont disponibles sur le site Web de la CEE : https://unece.org/statistics/events/DissComm2022 ;

c) **Réunion d'experts sur l'édition de données statistiques** (3-6 octobre, en ligne) : La réunion d'experts a porté sur le partage de données d'expérience sur la modernisation de l'édition de données et de la production statistique, y compris l'utilisation de données administratives et le recours à l'apprentissage automatique et à l'intelligence artificielle, ainsi que sur la qualité. En outre, des conférences éclair et plusieurs activités collaboratives visant à définir les thèmes de travaux futurs ont été organisées. Les exposés et les résultats de la réunion sont disponibles à l'adresse suivante : https://unece.org/statistics/events/SDE2022 ;

d) **Réunion d'experts sur la collecte de données statistiques** (26-28 octobre, Rome (Italie)) : La réunion a suscité un vif intérêt et un fort appui, avec plus de 210 inscriptions. En raison des mesures de prévention prises en Italie au lendemain de la crise liée à la COVID-19, le lieu de la réunion n'a pu accueillir qu'une cinquantaine de personnes et une réunion hybride a été organisée. Au total, 53 participants de 20 pays ont assisté en personne à la réunion et environ 60 participants y ont participé en ligne. L'accent a été mis sur la gestion de la collecte rapide et multimodale de données, l'amélioration de la pratique des répondants depuis le premier contact jusqu'à la fin de l'enquête et la conception réactive permettant de tirer parti des données collectées dans les premières phases d'une enquête pour optimiser la conception de la collecte dans les phases ultérieures. Soixante contributions ont été soumises et trente-quatre d'entre elles ont été retenues par le Comité d'organisation. Les exposés, les documents et les résultats sont disponibles sur le site web de la réunion : https://unece.org/statistics/events/DC2022 ;

e) **Atelier et séminaires du Groupe de haut niveau sur la modernisation de la statistique officielle** (22-24 novembre, Genève (Suisse), et 30 novembre, en ligne). La manifestation principale a été suivie en présentiel par 55 représentants de 27 organisations et en ligne par environ 70 personnes. L'objectif principal de l'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 2022. Un autre objectif majeur était de présenter le programme de travail prévu pour 2023. Les projets ont été classés par ordre de priorité et des suggestions ont été faites quant à leur portée. Tous les documents de la réunion sont disponibles sur le site web de la CEE : https://unece.org/statistics/events/HLG2022. En marge de la réunion, trois séminaires ont été organisés afin d'échanger des informations sur les travaux réalisés dans le cadre de deux projets menés en 2022 par le Groupe de haut niveau et la communauté de l'apprentissage automatique :

- **Séminaire de la méta-académie** (24 novembre, en présentiel). Le séminaire a consisté en deux séances interactives. La première a porté sur les résultats du module de travail 2 du projet de méta-académie sur Git et la gestion des versions, la deuxième a consisté en un exercice de remue-méninges visant à déterminer les insuffisances et les difficultés en matière de formation du personnel ;
- **Webinaire sur le projet de préservation de la confidentialité des données d'entrée** (24 novembre, hybride). Les résultats du projet portant sur l'intersection d'ensembles privés, l'apprentissage automatique privé et le calcul privé multipartite en tant que service ont été présentés. Le séminaire s'est conclu par une présentation des

applications possibles et des travaux restant à mener pour utiliser ces applications dans les processus de production ;

- **Webinaire du Groupe de l'apprentissage automatique** (30 novembre, en ligne). Les travaux des différents groupes de travail ont été présentés au cours de cette manifestation ONS/CEE. L'accent a été mis sur diverses applications (moissonnage de données sur le Web, classification des textes, données d'imagerie, données maritimes AIS, qualité des données de formation, reconfiguration des modèles et infrastructure.

V. Collaboration, communication et coordination

50. **Collaboration** : Les travaux du Groupe de haut niveau sont facilités par le grand nombre d'espaces wiki, de pages Web, d'espaces GitHub et d'autres plateformes favorisant la collaboration et l'échange d'informations. Ainsi, le secrétariat de la CEE gère et tient à jour une cinquantaine de sites wiki, publics ou à accès restreint, qui sont disponibles à des fins de collaboration ou de communication des résultats des activités du Groupe de haut niveau.

51. **Communication** : Pour mieux faire connaître les travaux du Groupe de haut niveau, divers articles de presse et rapports destinés à des organismes intergouvernementaux ont été rédigés. En outre, le secrétariat du Groupe de haut niveau a répondu à plusieurs demandes, provenant de l'intérieur et de l'extérieur de la région de la CEE, en vue de recevoir nos publications, d'échanger des informations, de présenter des travaux ou d'obtenir de l'aide. De plus amples informations sont disponibles sur la page wiki principale du Groupe de haut niveau. Le secrétariat de la CEE peut également être contacté à cette fin.

52. **Publications** : La plupart des résultats des travaux réalisés dans le cadre du Groupe de haut niveau sont accessibles au public sur les pages wiki du Groupe de haut niveau et sur le site Web de la CEE susmentionnés. En 2020, il a été décidé d'inclure une publication relative à la modernisation de la statistique dans le programme officiel de publication de la CEE. Le principal résultat du projet sur les données synthétiques a été publié en 2022 sous le titre « Synthetic Data for Official Statistics: A Starter Guide » (Données synthétiques au service de la statistique officielle : Guide de démarrage). Puisque la possibilité nous a été offerte de figurer dans une publication numérique officielle, les résultats du projet d'apprentissage automatique ont également été intégrés dans une publication officielle de la CEE (au format pdf)[2]. Ces possibilités nous permettent de faire connaître plus largement les travaux réalisés dans le cadre du Groupe de haut niveau.

53. **Coordination internationale :** Les activités ont été coordonnées avec d'autres organisations internationales concernées par la modernisation de la statistique officielle (comme Eurostat, l'OCDE et la Division de statistique de l'ONU). La coordination a été assurée et des liens ont été établis avec des initiatives menées internationalement dans des domaines similaires. Pour cela, une participation croisée des membres aux différents groupes a été la meilleure solution. En outre, le secrétariat de la CEE et la Commission de contrôle rencontrent régulièrement des représentants d'autres initiatives internationales de modernisation afin de mieux coordonner et harmoniser les différentes activités.

[2] https://unece.org/statistics/publications/machine-learning-official-statistics.