



Conseil économique et social

Distr. générale
19 décembre 2024
Français
Original : anglais

Commission de statistique

Cinquante-sixième session

New York, 4-7 mars 2025

Point 5 h) de l'ordre du jour provisoire*

**Questions soumises pour information : normes
ouvertes communes pour l'échange et le partage
de données et de métadonnées**

Organisations qui parrainent l'Échange de données et de métadonnées statistiques

Note du Secrétaire général

Conformément à la décision 2024/312 du Conseil économique et social et à la pratique établie, le Secrétaire général a l'honneur de transmettre le rapport des organisations qui parrainent l'Échange de données et de métadonnées statistique (SDMX), qui est soumis à la Commission pour information.

* [E/CN.3/2025/1](#).



Rapport des organisations qui parrainent l'Échange de données et de métadonnées statistiques

I. Introduction

1. L'initiative d'Échange de données et de métadonnées statistiques (ou SDMX, pour Statistical Data and Metadata Exchange), lancée en 2001, a pour objet l'élaboration et la tenue à jour de normes et de directives techniques et statistiques, ainsi que d'une architecture et d'outils informatiques à l'usage de la communauté de la statistique officielle.

2. La norme technique SDMX a considérablement évolué depuis sa création, la version SDMX 2.1, lancée en 2011, étant la première à susciter une réelle adhésion. La publication de la version 3.0, en 2021, a apporté une prise en charge déterminante des microdonnées, élargissant la norme, au-delà de son application classique, aux statistiques agrégées et aux séries chronologiques. SDMX 3.0 a de plus introduit le versionnage sémantique des éléments de métadonnées, qui renforce la gestion et l'échange de ces dernières en définissant le contrat de changement. Le versionnage robuste des métadonnées régi par des règles rassure les utilisateurs de métadonnées quant à la stabilité des structures, favorise la réutilisation des métadonnées et soutient les principes « FAIR » (faciles à trouver, accessibles, interopérables et réutilisables) en autorisant les références persistantes aux versions immuables des éléments. Ces dernières années, la norme SDMX a été de plus en plus utilisée au-delà de son application initiale, qui consistait à faciliter la communication et l'échange de données entre les pays et les organismes internationaux : la diffusion, la gestion des données et des métadonnées à l'appui de la gouvernance des données et d'autres cas d'utilisation ont gagné en importance.

3. La Commission de statistique reçoit des rapports d'activité sur l'initiative depuis 2002. En 2008, elle a accepté les normes et directives SDMX et préconisé qu'elles régissent l'échange et la mise en commun des données et métadonnées, demandant ainsi aux organisations qui parrainent l'initiative de poursuivre leurs travaux et d'engager les organismes nationaux et internationaux de statistique à recourir plus largement aux normes et directives.

II. L'Organisation internationale du Travail a rejoint les organisations qui parrainent l'Échange de données et de métadonnées statistiques

4. L'Organisation internationale du Travail (OIT), acceptant une invitation officielle du Comité de parrainage et signant le mémorandum d'accord sur l'échange de données et de métadonnées statistiques, a commencé à parrainer l'initiative SDMX. La gouvernance et le cadre de la coopération entre les organisations parrainantes, notamment les droits de propriété intellectuelle et les modalités de partage des ressources, sont établis dans le mémorandum.

5. L'OIT avait déjà apporté des contributions notables à l'initiative SDMX, avant d'étendre sa collaboration en tant qu'organisation parrainante, en créant des outils qui sont largement utilisés et en participant activement à des groupes de travail et des manifestations, notamment en exerçant un mandat prolongé à la tête du Groupe de travail sur les normes techniques relatives à l'échange de données et de métadonnées statistiques. Dans ce nouveau rôle, l'OIT rejoint la Banque des règlements internationaux (BRI), la Banque centrale européenne (BCE), Eurostat, le Fonds monétaire international (FMI), l'Organisation de coopération et de développement

économiques (OCDE), l'Organisation des Nations Unies (ONU) et la Banque mondiale, qui parrainent l'initiative SDMX depuis 2007.

III. Perspectives ouvertes par l'intelligence artificielle et répercussions

6. Le rythme de l'application de l'intelligence artificielle (IA) à SDMX s'est considérablement accéléré depuis l'émergence de ChatGPT à la fin de 2022, offrant des possibilités de révolutionner chaque étape du cycle de vie des statistiques, de la collecte initiale à la publication finale, en passant par la découverte et l'analyse des données. Malgré les progrès accomplis en matière d'utilisabilité et de contenu des portails de données institutionnels, dans les faits, la découverte de données grâce au langage naturel demeurait difficile jusqu'à récemment, mais la situation a commencé à évoluer. Plusieurs initiatives visant à faciliter l'interrogation de l'interface de programmation d'applications (API) SDMX à l'aide du langage naturel sont en cours. Il sera à terme possible d'utiliser l'IA avec tous types de données publiées par des organisations nationales ou internationales à l'aide d'outils faciles d'accès, libres et clés en main.

7. Le FMI a créé StatGPT, un assistant de génération de requêtes SDMX fondé sur l'IA. La tâche principale de StatGPT consiste à déchiffrer et à extraire comme il se doit tous les paramètres nécessaires d'une invite en langage naturel et à utiliser ces informations pour mettre au point des paramètres de requête SDMX qui permettent de restituer des données statistiques par l'entremise d'une API. Le premier prototype de cette application a été présenté à la neuvième Conférence mondiale sur la norme SDMX, qui s'est tenue à Bahreïn en octobre 2023. Elle a été suivie d'une présentation d'EPAM, le prestataire avec lequel le FMI met au point le produit, à l'atelier SDMX + AI organisé conjointement à Paris, en mars 2024, par l'OCDE et la BRI. À l'issue de cette rencontre, il a été convenu d'exploiter les travaux menés sur StatGPT et de mettre l'application à la disposition de l'ensemble de la communauté SDMX en utilisant des données produites par des organisations partenaires afin de faciliter sa mise à l'essai. Le FMI a publié StatGPT 2.0, qui contient des jeux de données provenant de huit organisations partenaires, en septembre 2024, et plus de 100 participants ont réalisé des essais. Les remontées, les statistiques d'utilisation, les coûts et les résultats d'une enquête seront compilés dans un rapport destiné à la communauté SDMX, qui se réunira pour étudier les possibilités de co-investissement. Le FMI souhaite mettre cet outil à la disposition des institutions publiant des statistiques, mais un modèle de financement durable sera nécessaire pour couvrir les frais de fonctionnement du grand modèle de langage, qui augmentent avec l'utilisation.

8. Autre évolution prometteuse en matière de découverte de données, Sease, une société spécialisée dans la technologie de recherche et l'IA basée au Royaume-Uni de Grande-Bretagne et d'Irlande du Nord, a collaboré avec les organisations qui parrainent SDMX afin d'améliorer la précision des réponses aux questions des consommateurs de données. En associant la technologie classique des moteurs de recherche à de grands modèles de langage, Sease entend désambiguïser les requêtes en langage naturel des utilisateurs, extraire les informations pertinentes et formuler des requêtes structurées. Malgré certaines limites inhérentes aux grands modèles de langage, cette approche constitue une solution relativement simple d'amélioration de la précision des résultats obtenus à partir de questions en langage naturel sur des données statistiques.

9. Les organisations parrainantes ont pour objectif de longue date de rendre SDMX plus accessible. Une initiative de la BRI visant à publier un agent conversationnel

fondé sur SDMX devrait faire considérablement progresser cet effort et améliorer la courbe d'apprentissage pour celles et ceux qui étudient la norme. Cet agent conversationnel propulsé par un grand modèle de langage qui utilise la génération augmentée de récupération permet aux utilisateurs de poser des questions sur les normes techniques et les orientations statistiques de SDMX. Il est prévu d'enrichir les informations de base de supports de renforcement des capacités, de manuels d'utilisation du logiciel SDMX et d'autres contenus utiles. Le développement et les essais qui sont en cours visent à garantir l'exactitude et la pertinence des réponses de l'agent conversationnel.

10. Si les avancées déjà réalisées sont très prometteuses, l'aventure de l'IA appliquée à SDMX ne fait que commencer. Le fait que d'importants investissements soient consentis à la fois par des institutions et par des entreprises privées démontre que l'IA peut apporter une valeur concrète à la chaîne de valeur des statistiques. Les efforts de collaboration et les solutions innovantes qui en cours d'élaboration mettent en lumière l'attachement de la communauté à exploiter l'IA au profit de toutes les parties prenantes.

IV. Examen du dossier de décision sur l'échange de données et de métadonnées statistiques

11. Le premier dossier de décision sur la norme SDMX a été publié en milieu d'année 2020. Cette publication a marqué une étape importante puisque, pour la première fois, les avantages de la norme SDMX et les différents aspects de son utilisation ont été rassemblés dans un document unique et concis. Toutefois, les évolutions rapides survenues depuis ont rapidement rendu ce dossier de décision obsolète. Cela s'explique notamment par le fait que l'utilisation de la norme a largement dépassé son objectif initial d'échange de données entre des pays et des organismes internationaux. Dans de nombreuses régions du monde, le principal cas d'utilisation de la norme SDMX est la diffusion de données facilitée par de puissantes plateformes, telles que la solution .Stat Suite, de la Communauté de collaboration sur les systèmes d'information statistique (SIS-CC), et StatKit, de l'Institut national italien de la statistique (ISTAT), ainsi que la gestion des données et des métadonnées, facilitée par Fusion Metadata Registry (FMR), à l'appui de la gouvernance des données. De plus, les versions les plus récentes de la norme prennent en charge des cas d'utilisation entièrement nouveaux, tels que l'échange de microdonnées et les données géospatiales.

12. Compte tenu des évolutions susmentionnées, les organismes qui parrainent SDMX ont créé, au début de 2024, une équipe spéciale chargée de l'examen du dossier de décision sur SDMX. Cette équipe se rapprochera des parties prenantes des organismes nationaux de statistique, des banques centrales, des universités et des organisations internationales, entre autres, pour répertorier les cas d'utilisation au-delà de l'échange de données et compiler les expériences des utilisateurs quant à la mise en œuvre de la norme SDMX pour les différents cas d'utilisation. Ces informations permettront d'actualiser le dossier de décision, ce qui facilitera la prise de décisions éclairées de la part des organismes nationaux et internationaux de statistique, ainsi que d'autres utilisateurs qui envisagent la mise en œuvre des flux de données SDMX. L'équipe spéciale devrait rendre ses conclusions en 2025.

V. Nouvelles initiatives

A. Écosystème de logiciels libres sdmx.io

13. Nombreux sont les organismes de statistique officielle qui considèrent les données comme un bien collectif ; la BRI est allée plus loin, en s'engageant à faire des logiciels utilisés pour la production ou l'utilisation de données statistiques un bien collectif. Cette idée s'est concrétisée dans une initiative appelée « sdmx.io ».

14. Conformément à la vision de l'initiative sdmx.io, la BRI s'est associée à des organisations internationales (OCDE, FMI, Eurostat), à des organismes nationaux de statistiques (Institut national de la statistique et des études économiques (INSEE) en France), à des banques centrales (Banca d'Italia, Banque nationale de Belgique) et à des organisations du secteur privé (Meaningful Data, HMS, Making Sense) pour améliorer les outils servant à produire et à utiliser des statistiques nationales et internationales. Les résultats et réalisations de ces initiatives ont contribué à la création ou à l'amélioration d'un riche écosystème de logiciels libres, de grande qualité, créés par des entités des secteurs public et privé.

15. Soucieuse de soutenir les concepteurs d'outils, la BRI offre une assistance à la création et la pérennisation d'outils libres et encourage l'adhésion aux bonnes pratiques relatives à l'interopérabilité en limitant le plus possible le chevauchement des fonctionnalités. Pour aider les utilisateurs de ces outils, la BRI s'efforce d'en assurer la pertinence en veillant à leur applicabilité à des cas pratiques d'utilisation statistique tout en favorisant leur adoption et leur utilisation grâce à des approches axées sur la facilité de déploiement et à des ressources d'apprentissage.

16. Ces mesures ont permis de renforcer la collaboration et la coordination entre les concepteurs d'outils, ce qui, dans certains cas, a conduit à la création de nouveaux outils destinés à combler les lacunes de l'écosystème de production de données (gingado, FMR Workbench, SDMX Dashboard Generator) ou, au contraire, à la consolidation de plusieurs outils en un seul outil de pointe (pysdmx, FMR, Matrix Generator) cocréé et tenu à jour grâce à des ressources partagées et à des modèles de bonne gouvernance.

B. UNdata

17. En réponse à la Stratégie du Secrétaire général pour l'exploitation des données par tout le monde, partout, le projet de modernisation d'UNdata a considérablement renforcé la capacité d'intégrer et de diffuser des données statistiques dans le système des Nations Unies. Guidée par le Plan d'action à l'échelle du système pour l'innovation dans le domaine des données et des statistiques des Nations Unies, cette initiative soutient directement le Groupe de travail sur les normes techniques relatives à l'échange de données et de métadonnées statistiques (Groupe de travail technique SDMX) en créant des outils qui améliorent l'interopérabilité des infrastructures de données conformes à la norme SDMX. Le projet de modernisation d'UNdata a notamment donné lieu à la création de cartographies complètes, visant à harmoniser les listes de codes et les concepts utilisés par diverses entités des Nations Unies, qui sont alignées sur les principales classifications statistiques internationales et contribuent ainsi à l'interopérabilité des mises en œuvre de la norme SDMX d'une organisation et d'un domaine statistique à l'autre. Ajoutons que l'alignement d'UNdata sur le cadre d'interopérabilité interdomaines (Cross-Domain Interoperability Framework) contribuera à l'application des principes « FAIR » dans l'écosystème de données SDMX.

18. L'accent mis dans le projet de modernisation de la plateforme UNdata sur le rapprochement entre des données variées grâce à des méthodes fondées sur les graphes de connaissance complète les objectifs du Groupe de travail technique SDMX en facilitant la traduction et la mise en relation des données au format SDMX avec des formats de données ouvertes tels que DDI-CDI (Data Documentation Initiative Cross-Domain Integration), schema.org et SKOS (Simple Knowledge Organization System). Ces travaux devraient ouvrir la voie à une navigation fluide entre les données statistiques structurées et les contextes de métadonnées plus larges, améliorant ainsi l'accessibilité et les moyens d'analyse. L'accent mis dans le projet sur les cartographies normalisées et lisibles par machine soutient l'objectif fixé à travers la norme SDMX d'une mise en commun efficace et évolutive des données tout en favorisant l'utilisation des ressources statistiques qui font autorité et en améliorant l'incidence de ces dernières.

VI. Avancées techniques

A. Langage de validation et de transformation

19. Le langage de validation et de transformation (Validation and Transformation Language, VTL) est un langage standard qui permet de définir des règles de validation et de transformation, notamment un ensemble d'opérateurs, ainsi que la syntaxe et la sémantique de ces derniers, pour tous types de données statistiques. Le langage VTL est fondé sur le modèle « Transformations et Expressions » en reprenant les parties communes des normes du modèle générique d'informations statistiques, de l'initiative SDMX et de la DDI (Data Documentation Initiative) pour représenter les concepts et les données. La formalisation logique des règles de validation et de transformation fournit un langage technologiquement neutre qui peut être utilisé directement par les utilisateurs, sans intervention des technologies de l'information. Ces caractéristiques permettent d'intégrer facilement le langage VTL aux différentes étapes du processus statistique dans des organisations qui adoptent des normes telles que SDMX, DDI ou DPM (Data Point Model).

20. L'équipe chargée du langage de validation et de transformation, créée en 2013 à l'initiative du secrétariat du SDMX et responsable de la création et de la tenue à jour du langage, en a publié la version 2.1 en juillet 2024. La nouvelle version contient des opérateurs temporels standard supplémentaires et apporte quelques correctifs de bogues dans les spécifications du langage. Le manuel d'utilisation et le manuel de référence devraient être disponibles à la fin de 2024 au format Markdown, qui améliore la lisibilité et la maintenabilité, et agrémentés d'un plus grand nombre d'exemples.

21. On recense actuellement quatre moteurs VTL libres, mis en œuvre par des sociétés privées et des organisations internationales. Le langage VTL a été adopté par plusieurs banques centrales, organismes nationaux de statistique et organisations internationales. Un certain nombre d'initiatives sont en cours de déploiement. Le VTL devrait être intégré à FMR, et plusieurs développeurs d'outils VTL rejoignent l'initiative sdmx.io.

B. SDMX 3.1

22. La dernière version de la norme technique, prévue pour la fin de 2024, apporte une solution aux problèmes que pose la gestion de jeux de données hautement multidimensionnelles dans les statistiques officielles. Les statisticiennes et statisticiens se sont démenés pour modéliser comme il se doit les jeux de données à

forte dimensionnalité en utilisant le modèle générique d'informations statistiques SDMX ou d'autres environnements, tels qu'une base de données relationnelle. Ces jeux de données nécessitent souvent l'ajout de nouvelles ventilations ou dimensions au fil du temps, ce qui est coûteux du fait de l'effet perturbateur sur les processus d'échange de données déjà en place. SDMX 3.1 atténue ces problèmes en ce qu'elle permet de définir des sous-cubes fixes sur les définitions de structure de données qui caractérisent des cubes analytiques et des schémas de tableaux de données qui demeurent valides même si la définition de structure de données ajoute des dimensions. Cette innovation limite la nécessité de recourir à des dimensions mixtes et à d'autres solutions de contournement, améliore l'interopérabilité et ouvre la voie à une création agile des définitions de structure de données en évitant de perturber les pipelines de collecte de données établis.

C. SDMX 3.2

23. À l'horizon 2025, une nouvelle amélioration incrémentielle du modèle générique d'informations statistiques SDMX devrait prendre en charge les classes de métadonnées structurales personnalisées. À ce jour, les informations fournissent un ensemble de classes normalisées pour les éléments de métadonnées structurales, notamment les concepts, les listes de codes et les définitions de structure. À partir de SDMX 3.2, les utilisateurs pourront créer leurs propres classes. Le premier cas d'utilisation de cette fonction de définition de structure personnalisée est celui des métadonnées de présentation, qui décrivent généralement la manière dont les données doivent être formatées lorsqu'elles sont présentées sur des sites Web et des portails de découverte de données. Les métadonnées de présentation sont souvent propres aux institutions et aux logiciels de visualisation, ce qui rend difficile la définition d'une classe de métadonnées normalisée qui réponde à tous les cas d'utilisation. Les définitions de structure personnalisées devraient trouver une application au-delà des métadonnées de présentation. On recense déjà des exemples de tableaux de publication, de modèles de rapports et de règles de vérification de l'égalité des soldes pour lesquels des classes de métadonnées structurales non normalisées sont utilisées, mais qui suivent par ailleurs les modèles SDMX pour la définition, le stockage et l'échange de métadonnées.

D. Interface de programmation d'application de transfert d'état représentationnel

24. L'interface de programmation d'application (API) de transfert d'état représentationnel (REpresentational State Transfer, REST) est l'un des trois piliers de la norme technique SDMX, avec le modèle générique d'informations statistiques et les formats de transmission. Introduite dans SDMX 2.1, la spécification de la version 1 de l'API REST a simplifié la découverte, l'extraction et la validation des données moyennant diverses requêtes. La version 2 de l'API REST SDMX, publiée avec SDMX 3.0, a introduit des améliorations non rétrocompatibles, notamment une amélioration des requêtes de données, des opérations de maintenance et des requêtes de métadonnées de référence. Elle prend également en charge de nouvelles caractéristiques du modèle générique d'informations statistiques telles que les listes de codes géospatiaux et le versionnage sémantique. D'autres améliorations incrémentielles sont prévues pour le quatrième trimestre de 2024, notamment le « voyage dans le temps », le tri et la pagination, ainsi que les contraintes segmentées. Pour l'avenir, la feuille de route prévoit de nouvelles API pour des fonctions commerciales telles que la validation et le mappage des données, ainsi que des

alternatives à REST, telles que gRPC, pour des cas d'utilisation particuliers, tels que les microservices.

E. Kit de comptabilité technologique

25. Dans le cadre de la contribution qui est la sienne aux travaux du Groupe de travail technique SDMX, Eurostat a mis au point un kit de compatibilité technologique (Technology Compatibility Kit, TCK) qui offre une boîte à outils permettant de tester la compatibilité des implémentations de l'API REST avec la norme SDMX. Le TCK permet aux utilisateurs et aux développeurs de logiciels SDMX de mesurer la conformité et la couverture d'un point de terminaison RESTful SDMX par rapport aux spécifications de l'API REST SDMX disponibles, grâce à une batterie de tests conçue de manière dynamique par le TCK. La dernière version officielle du TCK permet de tester la production de structures, schémas et données au format SDMX-ML (Markup Language) des requêtes de l'API REST par rapport aux spécifications de l'API REST SDMX jusqu'à la version 1.4 de l'API. Des versions d'essai du TCK assorties de fonctionnalités supplémentaires [par exemple, la possibilité de travailler avec les dernières spécifications de l'API ou la couverture de formats de sortie supplémentaires, tels que SDMX-JavaScript Object Notation (SDMX-JSON)] sont régulièrement mises à la disposition de la communauté SDMX.

VII. Services officiels d'échange de données et de métadonnées statistiques

A. Répertoire mondial SDMX

26. Le Répertoire mondial SDMX¹ est un répertoire central de définitions de structure de données convenues au niveau mondial pour plusieurs domaines, notamment la comptabilité nationale et les objectifs de développement durable. Il est également la source ultime de concepts interdomaines normalisés tels que la fréquence, l'unité de mesure et la confidentialité de l'observation, qui sous-tendent la volonté d'améliorer la comparabilité des jeux de données grâce à l'harmonisation mondiale des concepts.

27. Le Répertoire mondial prend en charge la dernière version (3.0) de la norme SDMX. Les évolutions prévues se concentrent sur la prise en charge de SDMX 3.1 et des versions ultérieures ainsi que sur le transfert du service vers l'infrastructure de la BRI.

B. SDMX Global Discovery Service

28. Le Service de découverte mondial SDMX (SDMX Global Discovery Service, SGDS) est un nouveau service officiel de l'initiative SDMX, dont le lancement est prévu au cours du deuxième trimestre de 2025, qui vise à améliorer la découverte et le partage des éléments et des jeux de données SDMX à mesure que les répertoires SDMX et les services de diffusion de données déployés et utilisés en production se multiplient. La nécessité de disposer d'un service de découverte robuste est évidente eu égard au rôle essentiel du Répertoire mondial, qui vise à gérer des éléments de base tels que les schémas d'organismes les plus élevés (top-level Agency Schemes)

¹ Voir <https://registry.sdmx.org>.

et les divers répertoires internationaux et nationaux dans lesquels on tient à jour des métadonnées structurelles largement utilisées.

29. La découverte de la source d'éléments liés à SDMX faisant autorité pour un organisme donné est un cas d'utilisation typique envisagé. Le SGDS, en fonction de l'identifiant de l'organisme SDMX, renverra le point de terminaison d'API de la structure REST SDMX à partir duquel les éléments de l'organisme peuvent être récupérés. Le service d'attribution de l'URN (Uniform Resource Name) est une fonctionnalité supplémentaire essentielle qui permet à SDMX de prendre en charge les principes « FAIR ». Son rôle consiste à traduire l'URN persistant d'un élément en une URL à partir de laquelle l'élément peut être récupéré.

30. Le SGDS fournit une API REST qui facilite l'accès par programmation, mais aussi une interface utilisateur simple qui permet de parcourir et de rechercher des éléments liés à SDMX. Le service est actuellement en phase d'essai et d'évaluation sur le site <https://gds.sdmx.io>. Les remontées des parties prenantes ont déjà permis d'améliorer la conception, notamment en rationalisant les messages de manière à les aligner plus étroitement sur les pratiques SDMX et en améliorant les requêtes axées sur les clients afin d'optimiser la découvrabilité.

31. La finalité du SGDS est d'exposer la nouvelle API à des clients potentiels, principalement d'autres services ayant besoin de découvrir les points de terminaison SDMX. L'interface utilisateur sera en outre mise à la disposition de la communauté SDMX, sur sdmx.io, afin de recueillir leurs observations et leurs réactions. La création de procédures d'intégration de nouveaux organismes et de leurs services est également une priorité. Ajoutons qu'un Test Configuration Kit (Technology Compatibility Kit) sera ajouté pour mettre à l'essai et configurer les accès aux services SDMX, notamment les ressources et les formats pris en charge.

VIII. Interopérabilité

A. Travaux généraux sur l'interopérabilité

32. Le Groupe d'appui à la mise en œuvre des normes du Groupe de haut niveau sur la modernisation de la statistique officielle de la Commission économique pour l'Europe a publié un rapport dans lequel il répertorie les éléments des normes SDMX et DDI qui sont pertinents pour chaque phase et sous-processus du Modèle générique du processus de production statistique, de la spécification des besoins statistiques à la diffusion des statistiques². Les normes SDMX et DDI sont toutes deux largement adoptées, chacune ayant des avantages et une communauté d'utilisateurs qui leur sont propres. L'utilisation du Modèle générique du processus de production statistique pour contextualiser les éléments liés à SDMX et DDI de cette manière vise à aider les spécialistes de l'une de ces normes à repérer facilement quels éléments de l'autre pourraient être utiles à un stade donné du processus de production de statistiques. Cette condition préalable est importante pour celles et ceux qui utilisent les deux normes et souhaitent les faire interagir, par exemple pour construire un pipeline de données. Les éléments des normes SDMX et DDI présentent de nombreuses similitudes et ces normes ont évolué, ces dernières années, jusqu'à commencer à faire parfois doublon dans les rôles qu'elles peuvent jouer. Le rapport place également les normes SDMX et DDI dans le contexte du Modèle générique du processus de production statistique et comporte des conseils visant à les rendre interopérables entre elles et avec le VTL. Composante utile de la boîte à outils des développeurs, ce

² Voir <https://unece.org/sites/default/files/2024-10/Implementation%20Standards%20in%20the%20context%20of%20GSBPM.pdf>.

rapport devrait en outre inspirer de nouveaux travaux sur la prise en compte d'autres normes de mise en œuvre ouvertes.

33. Les organisations qui parrainent SDMX ont créé en 2024 une équipe spéciale chargée des classifications SDMX, dont la mission consiste à faciliter la création, la publication et la gestion des classifications statistiques officielles sous forme de listes de codes dans le Répertoire mondial SDMX. Bien que de nombreuses classifications soient déjà publiées sous forme de listes de codes SDMX et utilisées dans les structures mondiales d'échange et de diffusion de données, la couverture est incomplète. La disponibilité des classifications officielles devrait à la fois simplifier la création de structures de diffusion de données par les organismes nationaux et internationaux de statistique et améliorer considérablement l'interopérabilité de ces structures. La mise au point de procédures normalisées de tenue à jour des classifications statistiques est un aspect essentiel des travaux de l'équipe spéciale, l'objectif étant de réduire au minimum la charge de travail et d'instaurer des procédures durables de tenue à jour des classifications. Complétée par les progrès de la modélisation des données dans les versions récentes de SDMX, ainsi que par la mise en œuvre prochaine des principes « FAIR » et la passerelle établie entre SDMX et le Web sémantique, cette initiative devrait ouvrir la voie à une amélioration non négligeable de l'utilisabilité des données statistiques, à la réduction du coût de la préparation préalable des données (*data wrangling*) issues de sources multiples et à la possibilité pour les applications d'IA d'utiliser facilement les données.

34. Le Groupe de travail statistique SDMX, en collaboration avec des partenaires, a travaillé sur les directives axées sur le contenu, dont certaines ont été publiées et d'autres sont en cours d'élaboration, suivantes :

- Directives sur l'unité de mesure. L'unité de mesure employée dans la modélisation des données statistiques est importante, car elle permet de garantir la cohérence et l'exactitude des données. Les unités de mesure sont utilisées pour quantifier les données et permettent d'évaluer de manière immédiate la comparabilité et le champ d'application du calcul des données. Des directives axées sur le contenu applicables aux unités de mesure s'adressent aux modélisateurs et modélisatrices de données statistiques qui travaillent avec des données multidimensionnelles, principalement dans le cadre de SDMX, mais avec un champ d'application qui peut aller au-delà (par exemple l'entreposage de données en général). Les modèles proposés sont ancrés dans des normes internationales existantes et bien établies dans les domaines de la science et de la technique.
- Directives relatives aux hiérarchies dans SDMX. L'objectif des directives est d'illustrer les cas d'utilisation applicables aux hiérarchies dans SDMX, de fournir des exemples et de recommander des meilleures pratiques de mise en œuvre. Les hiérarchies sont couramment utilisées pour représenter diverses relations et classifications. Leur rôle est essentiel, dans les systèmes de gestion des données, pour des tâches telles que la modélisation et la diffusion des données.
- Révision des directives relatives à la création et à la tenue à jour des listes de codes SDMX. La révision ajoute de nouvelles fonctionnalités de la version 3.0 de SDMX et répond à d'autres questions et recommandations pratiques, telles que celles qui sont liées aux ventilations à variantes multiples.
- Directives sur les annotations dans SDMX et révision du vocabulaire contrôlé. La révision ajoute des fonctionnalités de SDMX 3.0 et de nouvelles annotations au vocabulaire contrôlé, améliorant ainsi l'interopérabilité.

- Directives sur l'utilisation des millésimes d'acquisition dans SDMX. Dans les directives, le millésime est défini comme un jeu de données publiées à un moment donné. Les millésimes d'acquisition permettent aux universitaires de reproduire les travaux d'autres chercheurs et chercheuses, de mettre au point des modèles de prévision plus précis et d'analyser les décisions de politique économique à l'aide des données disponibles à l'époque où ces décisions ont été prises.
- Mise en œuvre de SDMX au niveau institutionnel. Nouvelles directives sur la mise en œuvre de SDMX au niveau institutionnel (multidomaines et tout au long du cycle de vie statistique), qui pourraient être utilisées par les organismes pour lancer leur mise en œuvre institutionnelle, économiser des ressources et du temps et atteindre les objectifs des utilisateurs et utilisatrices de SDMX.
- Étude des avantages de l'interopérabilité entre les ontologies SDMX et LOD (Linked Open Data, données ouvertes liées). À mesure que les normes applicables aux données gagnent en maturité, le besoin d'interopérabilité, qui peut être vu comme un simple moyen d'exprimer des informations d'une norme dans l'autre, se fait de plus en plus sentir. Si l'interopérabilité peut servir à élargir le public visé pour des informations qui sont les mêmes, la question plus complexe qui est posée est celle de savoir si les deux normes sont complémentaires en tirant parti du potentiel de l'une et de l'autre. Pour ce faire, il faut établir l'ontologie de SDMX et apporter d'éventuelles modifications à SDMX, par exemple fournir des identificateurs de ressources uniques pour les objets SDMX.
- Microdonnées pour SDMX 3.0. Les nouvelles directives décriront les meilleures pratiques applicables à la modélisation, au partage, à l'interopérabilité et la confidentialité des microdonnées.

B. Évolutions propres à un domaine

35. Dans le domaine des statistiques du travail, l'Organisation internationale du Travail (OIT) a continué d'aider les pays à mettre en œuvre des systèmes d'information sur le marché du travail (SIMT) en élaborant un ensemble d'activités de renforcement des capacités SDMX et en fournissant une boîte à outils constituée de logiciels libres basés sur SDMX, à savoir la solution .Stat Suite, SDMX Constructor, Statistical Metadata-driven Analysis & Reporting Tool (SMART) et LMIS Excel Add-in. Ces deux dernières années, quatre pays (Afrique du Sud, Botswana, Chili et Uruguay) ont lancé leur système, mais quatre autres en sont aux dernières étapes du téléchargement initial des données, et 15 projets (13 systèmes nationaux et 2 systèmes régionaux) sont en cours de mise en œuvre. Une définition de structure de données mondiales du travail élaborée par l'OCDE, l'OIT, Eurostat, la BCE et la Banque mondiale a été publiée et peut être consultée dans le Répertoire mondial pour l'échange de données et de métadonnées sur le travail.

36. Le FMI a continué de faciliter la diffusion des données au moyen de la norme SDMX dans le cadre de la mise en œuvre de ses initiatives relatives aux normes de diffusion des données. Ces deux dernières années, il a aidé sept pays à mettre en place des pages de données nationales récapitulatives compatibles avec SDMX, à savoir trois adhérents au Système général de diffusion des données amélioré, deux abonnés à la Norme spéciale de diffusion des données (NSDD) et deux abonnés à la norme NSDD Plus. À la fin de 2024, 112 pays diffuseront des données au format SDMX via leurs pages de données nationales récapitulatives.

37. Une note d'orientation sur la norme SDMX applicable au commerce de marchandises et de services a été élaborée dans le cadre de la révision des *Statistiques*

du commerce international de marchandises 2010 et du Manuel des statistiques du commerce international des services 2010 par l'équipe spéciale chargée des statistiques du commerce international du Comité d'experts chargé des statistiques relatives aux entreprises et au commerce³ et a fait l'objet d'une consultation mondiale en février 2024. La note d'orientation comporte une liste de variables minimales et une mise à jour des définitions de structure de données tenant compte des versions à venir des *Statistiques* et du *Manuel*, et fait état des stratégies relatives à la mise en œuvre de la stratégie SDMX dans les *Statistiques*. Le résultat positif de la consultation mondiale allait dans le sens d'un appui au cas d'utilisation de la norme SDMX pour les statistiques commerciales, en soulignant les avantages potentiels que présente ce cas d'utilisation pour le partage des données, la normalisation et l'efficacité. L'équipe spéciale continuera d'exécuter le plan décrit dans les notes d'orientation en 2025 et au-delà.

38. Les définitions de structure de données SDMX mondiales dans le domaine des statistiques macroéconomiques continuent d'être tenues à jour et d'être utilisées pour l'échange de données. En particulier, la Nomenclature des fonctions de la consommation individuelle de 2018, récemment finalisée, a été mise en œuvre en 2024 dans plusieurs de ces définitions de structure de données pour les statistiques macroéconomiques mondiales. Les domaines actuellement couverts comprennent la comptabilité nationale, la balance des paiements, les investissements étrangers directs, les indices des prix à la consommation et la comptabilité économique et environnementale. Une définition de structure de données mondiales pour les prix de l'immobilier résidentiel est également en cours d'élaboration, sous la coordination de la BRI.

39. Dans le cadre des travaux menés par le Comité d'experts de la comptabilité environnementale et économique, la Division de statistique, en collaboration avec l'OCDE, a lancé en 2023 une collecte mondiale conjointe de données pour les comptes des émissions atmosphériques et des flux d'énergie conformément au Système de comptabilité environnementale et économique. La collecte mondiale de données complète la collecte de données menée par Eurostat pour les pays du système statistique européen. La collecte de données se fonde sur les questionnaires Excel compatibles avec SDMX qui correspondent aux définitions de structure de données du Système de comptabilité environnementale et économique. Des questionnaires analogues ont été mis à disposition des pays du système statistique européen par Eurostat. La Division de statistique diffuse les données par l'entremise de son API SDMX. Elle a en outre commencé en 2024 à échanger avec Eurostat et l'OCDE des données sur les comptes des émissions atmosphériques et des flux d'énergie au moyen d'API SDMX. La Division de statistique diffuse donc désormais les données du Système de comptabilité environnementale et économique pour 42 États Membres.

40. Le Groupe de travail sur l'échange de données et de métadonnées statistiques concernant les indicateurs relatifs aux objectifs de développement durable a continué de tenir à jour les structures mondiales de données et de métadonnées et les flux de données relatifs aux indicateurs des objectifs. L'échange de données SDMX avec les organismes responsables des indicateurs relatifs aux objectifs de développement durable a considérablement augmenté ; plus de 40 % de la Base de données mondiale relative aux indicateurs de suivi des objectifs de développement durable sont désormais transmis sous la forme de jeux de données SDMX. La fourniture volontaire de jeux de données sur les objectifs par les États Membres s'est également poursuivie, l'échange de données ayant été mis en place avec une quarantaine de pays. Le jeu de données mondiales sur les objectifs, ainsi que les métadonnées relatives aux

³ Disponible (en anglais) à l'adresse suivante : https://unstats.un.org/wiki/download/attachments/311230515/GNV.15_SDMX-IMTS-MSITS.pdf?version=1&modificationDate=1735849291423&api=v2.

indicateurs, est diffusé grâce à des API SDMX de la Division de statistique. La disponibilité des métadonnées dans un format lisible par machine a permis d'associer les indicateurs de suivi des objectifs à leurs métadonnées sur le site Web de la Base de données mondiale et d'afficher les données assorties des métadonnées. Le Groupe de travail a également contribué à la création de la plateforme Open SDG⁴, plébiscitée, qui est gérée par l'organisme national de statistique du Royaume-Uni, qui prend en charge les structures de données et métadonnées relatives aux objectifs et qui est utilisée par quelque 18 États Membres pour diffuser les indicateurs de suivi des objectifs qui sont les leurs. L'expérience acquise dans le cadre de l'échange de données sur les objectifs a permis au Groupe de travail d'apporter d'importantes contributions à la création de SDMX 3.0 et SDMX 3.1. Le Groupe de travail a également mis au point une formation en ligne sur l'emploi de la norme SDMX aux fins du suivi des indicateurs relatifs aux objectifs de développement durable et a coordonné les efforts de renforcement des capacités dans ce domaine.

IX. Logiciels et applications

41. L'OIT a continué de promouvoir sa boîte à outils SDMX, composée de SDMX Constructor pour l'édition des métadonnées structurales et de SMART pour la compilation, le recodage et le reformatage des données. Constructor a été téléchargé plus de 1 700 fois et SMART, près de 1 100 fois. Un nouveau complément Excel, LMIS Excel Add-in, a été publié et permet de télécharger dans Excel n'importe quel flux de données SDMX à partir de n'importe quelle API SDMX. En cas de connexion à une plateforme .Stat Suite, ces données Excel peuvent être éditées et renvoyées dans l'entrepôt de données.

42. L'OIT a mis au point une méthode simplifiée de modélisation des données pour la mise en place des SIMT, en utilisant un classeur partagé dont une feuille permet de saisir la liste des concepts, une autre feuille est réservée aux listes de codes associées à chaque concept par son identifiant et une troisième feuille permet de définir les tableaux dimensionnels à publier, en fonction de l'indicateur et des ventilations. Une fonction de SDMX Constructor permet à l'utilisateur d'importer ces feuilles de calcul et de créer tous les éléments SDMX (à savoir le schéma conceptuel, les listes de codes, les définitions de structure de données, les flux de données et les contraintes de contenu) nécessaires au téléchargement initial des données en une fois.

43. L'outil SDMX Central du FMI a continué d'aider les pays membres à convertir, à valider et à enregistrer les données SDMX sur leur page de données nationales récapitulatives. Depuis 2016, le FMI met une plateforme en nuage et des services Web gratuitement à la disposition des utilisateurs, promouvant ainsi SDMX comme un bon moyen d'échange pour les pays qui appliquent le Système général de diffusion des données amélioré et sont abonnés aux normes NSDD et NSDD Plus. SDMX Central est un rouage essentiel de la collecte automatisée de données entre le FMI et les pays membres. Le FMI a opéré la transition de SDMX Central à FMR, en 2023, ce qui a eu pour avantage non négligeable d'ouvrir la voie à une collaboration stratégique avec d'autres organismes qui parrainent SDMX et à l'adoption d'une approche commune de la priorité donnée à la création d'une plateforme unique et unifiée. FMR, qui est une plateforme libre et gratuite, est gérée par la BRI.

44. SDMX Matrix Generator est un outil basé sur Excel qui permet de modéliser et de concevoir visuellement des éléments de données SDMX et de générer le balisage SDMX-ML pour la mise en œuvre. SDMX Matrix Generator est une amélioration de Generic SDMX Design Matrix, qui a fait ses preuves en tant qu'outil de conception

⁴ Voir <https://open-sdg.org/>.

collaboratif pour les non spécialistes de SDMX. L'objectif principal de l'outil est d'intégrer les directives SDMX en matière de modélisation et de pouvoir créer les éléments sans disposer d'une grande connaissance technique de SDMX tout en mettant l'accent sur les aspects statistiques du modèle de données. Ces aspects lui ont valu d'être largement utilisé dans la formation à la modélisation. L'outil permet de trouver des solutions de modélisation souples, comme utiliser une définition de structure de données pour plusieurs flux de données. Il est largement utilisé par les organismes parrainants et la communauté SDMX dans son ensemble, et s'est avéré particulièrement utile dans la phase initiale de conceptualisation des modèles de données. Les projets mondiaux de définitions de structure de données se fondent sur SDMX Matrix Generator pour concevoir de manière collaborative leurs métadonnées structurelles. Les nouvelles fonctionnalités qui ont été demandées sont la création de schémas de catégories et de schémas d'organismes, la définition facile de contraintes, l'extension aux enquêtes de modélisation et aux grandes listes de codes, la sérialisation de la décomposition, les dimensions non codées, les références à plusieurs schémas de concepts, les rôles de concepts personnalisés et de nombreuses améliorations de la performance et de la stabilité.

45. Microsoft Power BI étant devenu l'un des outils d'informatique décisionnelle les plus utilisés sur le marché aujourd'hui, la SIS-CC (voir par. 48) a jugé nécessaire de faciliter l'obtention et la visualisation des données statistiques en mettant au point un connecteur Power-BI pour SDMX. Le connecteur étant certifié par Microsoft, il est intégré à chaque installation de Power BI dans le monde. C'est la première fois que SDMX est intégré à la suite d'outils d'une entreprise du secteur privé, et cette initiative a considérablement développé la base d'installation de SDMX.

46. FMR est un registre de métadonnées structurelles SDMX perfectionné et libre d'utilisation, mis au point et tenu à jour par la BRI. Il est largement utilisé dans des organisations du monde entier en tant que plateforme de modélisation des données et répertoire contrôlé pour les métadonnées statistiques, contribuant à améliorer la gouvernance, la maintenabilité, la normalisation et l'harmonisation des métadonnées. Outre son déploiement dans des institutions, FMR propulse plusieurs services publics de métadonnées structurelles SDMX, notamment le Répertoire mondial SDMX et IMF Central. Depuis sa première version en 2007, FMR fournit une interface utilisateur qui permet aux statisticiens et aux responsables des données de créer et de tenir à jour de manière interactive des éléments de métadonnées SDMX. Cette capacité a été étendue à FMR Workbench, une variante de FMR qui fonctionne avec n'importe quel répertoire de métadonnées SDMX qui présente l'API REST standard. L'une des principales motivations de la création de FMR Workbench était la nécessité pour la Banque nationale de Belgique de faciliter la gestion des métadonnées structurelles dans la solution .Stat Suite de la SIS-CC, qui est très utilisée pour diffuser des données statistiques. L'équipe FMR de la BRI, en collaboration avec des collègues de l'OCDE et de la Banque nationale de Belgique, a depuis conçu une interface utilisateur FMR supplémentaire aux fins particulières du cas d'utilisation de la modélisation dans le domaine statistique. Cet outil de modélisation des données FMR met en œuvre la méthode matricielle bien établie, qui permet de décrire des domaines statistiques complets sous forme de modèles abstraits. Ces derniers peuvent ensuite être matérialisés sous forme d'éléments SDMX afin d'être utilisés dans la pratique. Les responsables du projet recherchent actuellement des co-investisseurs pour financer la phase de mise sur pied, un projet viable *a minima* étant prévu pour le début de 2026, sous réserve de son financement. La version 11 de FMR, publiée en 2021, prend en charge le modèle générique d'informations statistiques et les fonctionnalités de base de SDMX 3.0., et a depuis fait l'objet de 19 itérations mineures. FMR 12, qui sera la prochaine version majeure, est prévue pour le début de 2025 et ajoute la prise en charge du « voyage dans le temps des métadonnées », qui permet de récupérer les métadonnées structurelles telles qu'elles étaient à

n'importe quel moment dans le passé. La validation des données SDMX à l'aide des métadonnées structurales telles qu'elles étaient au moment où les données ont été créées est un cas d'utilisation essentiel du voyage dans le temps des métadonnées, qui est primordial pour les responsables de la collecte et de la communication de données dans les domaines dans lesquels les métadonnées évoluent rapidement. FMR 13, qui est prévu pour le milieu de l'année 2025, prendra en charge le versionnage sémantique des éléments et les améliorations prévues dans SDMX 3.1 relatives à la modélisation de données comportant un grand nombre de dimensions.

47. Eurostat a continué de perfectionner et de mettre à jour ses outils SDMX. Il a mis à niveau son application SDMX Converter et les différentes composantes de son infrastructure de référence SDMX de manière à prendre en charge les structures et les formats de données de SDMX 3.0. La prise en charge des principales fonctionnalités de SDMX 3.0, telles que les valeurs de tableau et les mesures multiples, a fait l'objet d'un soin particulier. Les outils d'Eurostat sont de type logiciel libre.

48. La solution .Stat Suite est une plateforme SDMX native, perfectionnée, de type logiciel libre et libre d'utilisation sur laquelle repose le cycle de vie des données pour les statistiques officielles dans plus de 50 organisations à travers le monde. Elle est développée et tenue à jour par la SIS-CC, une communauté mondiale de plus de 20 membres et partenaires qui œuvrent sous la direction de l'OCDE. La plateforme .Stat Suite permet de gérer le cycle de vie des données pour les statistiques officielles (conception, collecte, traitement, diffusion), d'explorer les données et de mettre au point diverses expériences de communication de l'information et de diffusion, et fait fond sur les meilleures pratiques en matière de modélisation des données statistiques. La solution .Stat Suite se fonde sur l'outil SDMX Reference Infrastructure (SDMX-RI) d'Eurostat et FMR Workbench. Ces deux dernières années, la communauté a commencé à travailler, en collaboration avec la BRI et Eurostat, sur une mise en œuvre SDMX de référence. Par ailleurs, un certain nombre de projets ont été mis en ligne, et la solution .Stat Suite a été lancée dans plusieurs organismes de statistique et organisations internationales, notamment en El Salvador, en Grèce, au Luxembourg, à Madagascar, aux Maldives, à Malte, en Thaïlande, en Uruguay et à l'Organisation des Nations Unies pour l'alimentation et l'agriculture. La solution .Stat Suite est également un pilier de l'initiative de modélisation des données SDMX, menée par l'OCDE. Dans le cadre des efforts constants qu'elle déploie en faveur de l'adoption de logiciels libres, la SIS-CC a mis en place un support, pour la base de données libre MariaDB, qui permettra aux responsables de la mise en œuvre de dégager des économies substantielles.

49. Le secrétariat de l'initiative SDMX a récemment lancé une nouvelle page Web sur les outils SDMX destinés aux développeurs et aux responsables de la mise en œuvre où l'on peut trouver une liste complète des outils SDMX mis à disposition par les organismes parrainants et les membres de la communauté SDMX au sens large.

X. Communication, information et collaboration des parties prenantes

A. Manifestations mondiales

50. La neuvième Conférence mondiale sur la norme SDMX, qui s'est tenue à Bahreïn du 29 octobre au 2 novembre 2023, a suscité un vif intérêt dans le monde entier, attirant plus de 550 participants. Quelque 230 personnes y ont assisté sur place, les autres ayant suivi les réunions plénières à distance en direct et en différé. Cette manifestation, la première à se tenir en présentiel depuis la pandémie de maladie à coronavirus (COVID-19), a été organisée conjointement par les organismes qui

parrainent SDMX et l'autorité chargée de l'information et de l'administration en ligne de Bahreïn, avec le soutien de la Commission économique et sociale pour l'Asie occidentale (CESAO) et de la Banque d'Italie, et a réuni avec succès un large éventail de parties prenantes représentant des organismes nationaux, régionaux et internationaux, des universités et le secteur privé. Le thème, « Donner aux communautés de données les moyens d'agir » (*Empowering data communities*), a été l'occasion d'explorer la manière dont les outils et les technologies SDMX pouvaient être exploités de façon à améliorer la collaboration et le partage des connaissances. Les neuf sessions plénières organisées sur trois jours couvraient un large éventail de sujets allant de l'IA aux avancées en matière de gouvernance des données grâce à la norme SDMX. La manifestation s'est terminée par deux sessions de formation parallèles de renforcement des capacités SDMX qui s'adressaient aux débutants et aux praticiens plus avancés. Les participants à la table ronde de clôture ont insisté sur le poids de SDMX dans le monde de la gestion et de la gouvernance des données et souligné qu'il fallait constamment s'adapter et collaborer pour répondre à l'évolution des besoins et des difficultés en matière de données. Ils sont convenus qu'une approche centrée sur l'utilisateur et l'accent mis sur l'innovation et la gouvernance des données constituaient la voie à suivre pour le SDMX de demain.

51. Le douzième atelier des spécialistes de SDMX s'est tenu à Amsterdam du 7 au 11 octobre 2024. La manifestation a attiré 87 spécialistes en présentiel, les autres l'ayant suivie à distance grâce à une retransmission en direct. La manifestation était organisée par Eurostat, en collaboration avec les organismes qui parrainent SDMX. Les 11 sessions plénières tenues sur quatre jours étaient complétées par une session avancée de renforcement des capacités le cinquième jour, ainsi que plusieurs réunions parallèles de différents groupes chargés de la gouvernance de l'échange de données et de métadonnées statistiques et équipes spécialisées dans la norme SDMX. La manifestation, qui a mis en évidence l'évolution des cas d'utilisation de la norme SDMX au fil du temps, bien au-delà de l'objectif initial de la norme, a été l'occasion de montrer le potentiel de cette dernière dans la découverte, l'extraction et l'analyse de données assistées par l'IA.

B. Forum des utilisateurs de SDMX

52. Le forum des utilisateurs de SDMX est une plateforme de partage des connaissances, de règlement des problèmes et de collaboration. Lancé en novembre 2022 et fruit d'une collaboration entre des organismes nationaux de statistique, les organisations qui parrainent SDMX et les commissions régionales de l'ONU, il est un maillon du Réseau mondial des responsables des données et des statisticiens, qui est dirigé par la Division de statistique. Le forum est depuis devenu l'une des communautés les plus actives du Réseau mondial, avec plus de 600 messages publiés et 42 000 vues enregistrées depuis son lancement. Le nombre d'utilisateurs et utilisatrices se monte actuellement à plus de 430 et continue de croître. Le forum est devenu un lieu de rencontre prisé et incontournable pour les praticiens et praticiennes de l'échange de données et de métadonnées statistiques, où une communauté d'experts apporte son soutien, où des manifestations et des initiatives sont annoncées et où des ressources de formation sont proposées. La participation au forum est ouverte à tous les utilisateurs, quel que soit leur niveau d'expérience et de connaissance de la norme SDMX.

C. Modernisation du site Web consacré à SDMX

53. Le site Web officiel de l'initiative SDMX, hébergé à l'adresse <https://sdmx.org>, fournit un bon point d'entrée pour les personnes qui découvrent la norme et contient

des ressources, des informations et des conseils qui s'adressent à l'ensemble de la communauté. Il est en cours de modernisation sous la direction du secrétariat et, en particulier, de la BCE, qui finance les interventions menées sur le site. La modernisation est progressive, en commençant par l'amélioration de la présentation des logiciels et des catalogues de renforcement des capacités et la rationalisation de leur tenue à jour. Notons que les créateurs d'outils et de supports d'apprentissage en ligne peuvent désormais soumettre leurs propres ajouts et mises à jour sous le contrôle de l'équipe éditoriale du site, chapeautée par Eurostat, ce qui devrait permettre d'assurer la fraîcheur et l'exhaustivité du contenu. La prochaine étape du programme de modernisation consiste à rafraîchir la page d'accueil en actualisant son style, sa palette de couleurs et son contenu afin de susciter l'intérêt du public et de promouvoir la norme. Les études de conception ont été approuvées et les maquettes ont été converties en une prévisualisation qui est prête pour une mise à l'essai au début de 2025.

XI. Renforcement des capacités

54. La vision de sdmx.io s'étend au-delà de la création et de la durabilité d'un écosystème d'outils de premier plan utilisés pour produire et exploiter des statistiques officielles : elle prévoit aussi la mise à disposition de connaissances visant à aider les organismes de statistique officielle, quel que soit leur stade de développement, à tirer parti de ces technologies et méthodes modernes afin d'améliorer leur capacité de production statistique. En bref, sdmx.io met également l'accent sur les moyens qui permettent de réduire les risques, d'améliorer la qualité et d'accélérer l'arrivée à maturité des méthodes de production de données à tous les stades du cycle de vie des données. Ces objectifs sont atteints en proposant des cours d'apprentissage en ligne de courte durée et des webinaires pratiques réguliers dans lesquels sont abordés des cas d'utilisation pratiques des statistiques officielles. Les webinaires pratiques de sdmx.io comptent régulièrement 100 à 150 participants, et les enregistrements des webinaires sont mis à la disposition, pour consultation, de celles et ceux qui n'ont pas pu assister aux manifestations en direct.

55. L'OIT organise chaque année, dans son Centre international de formation de Turin (Italie), un cours de formation en résidence d'une semaine, sur la mise en œuvre des SIMT, auquel participent une trentaine de personnes venues de plusieurs pays. Le cours comprend une introduction à SDMX, à la modélisation des données SDMX et à la préparation des données en vue de l'ETL (extraire, transformer et charger) sur une plateforme .Stat Suite. Il est par ailleurs possible de suivre chaque année une formation en ligne intitulée « Introduction to SDMX ».

56. Dans la région Afrique, la collaboration entre le FMI et la Banque africaine de développement (BAD) sur la mise en œuvre de SDMX bénéficie du lancement du portail de données ouvertes (ODP) 2.0 de la BAD. Cette plateforme permet aux pays de modéliser leurs jeux de données avant la diffusion, ce qui a pour objectif d'améliorer l'utilisabilité et l'ouverture de ces données. La mise à disposition de la plateforme sera complétée par la fourniture conjointe d'une assistance au développement des capacités de modélisation des données. Les jeux de données de la plateforme deviennent partie intégrante d'un ensemble global de statistiques officielles accessibles au public. Ce travail conjoint sur la modélisation permettra d'aligner cet ensemble sur les normes internationales et de le préparer pour l'IA.

57. Le Groupe de travail sur l'échange de données et de métadonnées statistiques concernant les indicateurs relatifs aux objectifs de développement durable a élaboré et publié un cours de formation en ligne sur l'emploi de la norme SDMX au service des objectifs grâce à un financement fourni par le Royaume-Uni. La formation, qui

s'adresse aux personnes qui connaissent les rudiments de SDMX et souhaitent se familiariser avec les structures et les flux de données relatifs aux objectifs de développement durable, a amélioré la courbe d'apprentissage, facilité la mise en œuvre de l'échange et de la diffusion de données sur les objectifs de développement durable dans les États Membres, ainsi que dans les organismes responsables des indicateurs relatifs aux objectifs de développement durable, et complété la formation en ligne et en présentiel qui est dispensée par la Division de statistique et d'autres organismes internationaux.

58. La Division de statistique a continué de mener des activités de renforcement des capacités, en ligne et en présentiel, en partenariat avec des organisations internationales, notamment les commissions régionales de l'ONU, d'autres organismes qui parrainent SDMX et des banques régionales de développement. Une formation en ligne sur l'emploi de la norme SDMX aux fins du suivi des indicateurs relatifs aux objectifs de développement durable⁵ a été dispensée aux organismes responsables, ce qui a entraîné une nette augmentation de la quantité de données transmises au format SDMX et réduit, de ce fait, le fardeau des entités chargées de communiquer et de collecter ces données. Une formation de base en ligne sur SDMX a été dispensée à l'organisme national de statistique de l'Ukraine et a attiré plus de 50 participantes et participants, ce qui a facilité la mise en œuvre de flux de données basés sur SDMX au sein de l'organisme ; d'autres formations sont prévues. En partenariat avec l'OCDE et la Banque asiatique de développement (BAsD), une formation de base sur SDMX a été dispensée à trois pays d'Asie du Sud-Est en septembre 2024 ; elle devrait ouvrir la voie à une série d'autres formations de plus en plus poussées. D'autres activités de formation ont été organisées en partenariat avec les commissions régionales de l'ONU, notamment la Commission économique pour l'Afrique (CEA) et la Commission économique et sociale pour l'Asie occidentale (CESAO).

59. La BAsD, en coopération avec la Division de statistique, la Commission économique et sociale pour l'Asie et le Pacifique (CESAP) et l'Institut de statistique pour l'Asie et le Pacifique (ISAP), a lancé SDMX Foundation Course, un cours en ligne qui s'est déroulé d'avril à mai 2022 et a attiré plus de 500 participantes et participants du monde entier. Le cours a été relancé en février 2023 et les inscriptions sont ouvertes⁶. En collaboration avec la Division, la CESAP et l'ISAP, la BAsD a également lancé SDMX Tools Course, un cours d'apprentissage en ligne qui a été dispensé pour la première fois en novembre et décembre 2023. Le cours, relancé en juillet 2024, est désormais disponible sur demande⁷.

60. La CEA, en partenariat avec la Division de statistique et l'organisme national de statistique du Royaume-Uni, a dispensé une formation sur la modélisation et la diffusion des indicateurs relatifs aux objectifs de développement durable à sept États Membres d'Afrique dans le cadre d'un projet financé par l'Union européenne sur la diffusion de données relatives aux indicateurs. Les sept pays ont à présent créé des prototypes de tableaux de bord sur les objectifs de développement durable qui sont fondés sur la plateforme de diffusion Open SDG, et plusieurs d'entre eux devraient être officiellement lancés à courte échéance.

61. La CESAO a lancé plusieurs initiatives importantes de renforcement des capacités et d'échange de données sur les indicateurs relatifs aux objectifs de développement durable. En collaboration avec la Division de statistique et d'autres institutions, notamment l'Institut arabe de formation et de recherches statistiques et la Ligue des États arabes, la CESAO a organisé, depuis 2020, sept ateliers régionaux

⁵ Voir <https://learning.officialstatistics.org/course/view.php?id=96>.

⁶ Voir <https://elearn.adb.org/course/view.php?id=486>.

⁷ Voir <https://elearn.adb.org/course/view.php?id=520>.

sur l'emploi de la norme SDMX aux fins de la communication des données sur les objectifs de développement durable, la finalité étant de permettre aux États membres de mettre en œuvre les dernières avancées en matière d'automatisation et d'échange de données conformément à la norme. L'outil SDMX Converter for SDGs de la CESAO, mis au point pour faciliter l'échange de données sur les indicateurs des objectifs, a considérablement rationalisé la communication de données nationales sur les objectifs et est actuellement utilisé par 17 pays arabes, rendant ces informations facilement accessibles grâce à la plateforme Arab SDG Monitor. En conséquence, la diffusion des données a augmenté d'au moins 40 %, la qualité des données a progressé et le fardeau des États membres en matière de communication des données a été considérablement réduit. De plus, le guide complet sur SDMX Converter qui a été élaboré par la CESAO et qui contient des conseils techniques sur l'échange et la validation des données aide les États membres à améliorer leur infrastructure de données. La Commission a également enrichi les ressources de la région en fournissant un glossaire SDMX détaillé, en arabe, dans lequel figurent plus de 250 nouveaux termes statistiques visant à promouvoir les normes internationales au sein de la communauté de la statistique.

62. La CESAP continue d'investir dans la mise en œuvre de la norme SDMX dans la région qui est la sienne par l'entremise de projets de renforcement des capacités visant à établir des portails de données nationaux. Les Maldives ont lancé leur première plateforme de diffusion, en octobre 2023, dans le sillage de deux ateliers pratiques organisés par la CESAP et auxquels ont participé des membres du personnel de l'organisme national de statistique et d'autres entités fournissant des données. La modernisation du processus métier a permis de rationaliser la diffusion des données de celles et ceux qui les produisent vers le public, en tirant parti d'outils libres et gratuits tels que .Stat Suite et FMR. Faisant fond sur cette expérience, la CESAP s'est associée à la BASD et à la Communauté du Pacifique afin de lancer une initiative conjointe pour Samoa et Fidji, à la demande de leurs organismes nationaux de statistiques respectifs, visant à appliquer SDMX aux processus de diffusion de données et à créer des portails de données. Le projet favorise en outre le renforcement de la coopération technique entre les deux pays, ce qui permet d'assurer la viabilité à long terme de l'initiative dans la région.

63. La .Stat Academy, créée par la SIS-CC sous les auspices de l'OCDE, offre une formation gratuite en ligne, à suivre à son propre rythme, qui permet d'accompagner le renforcement des capacités pour la solution .Stat Suite et la modélisation des données dans SDMX pour les informaticiens et les producteurs de données. Les parcours d'apprentissage de la .Stat Academy ont été conçus pour développer les connaissances et les compétences au fil du temps, en fonction des objectifs de chaque apprenant ou apprenante. En octobre 2024, plus de 1 000 apprenants avaient été certifiés, et les utilisateurs enregistrés provenaient de 152 pays.

64. L'OCDE devrait organiser en décembre 2024 une manifestation régionale de renforcement des capacités SDMX, dans son centre régional nouvellement établi à Istanbul (Türkiye), à laquelle doivent participer plus de 45 participants représentant 22 organisations. La formation est conçue pour renforcer et améliorer les capacités SDMX des organisations qui sont déjà familiarisées avec l'utilisation de la norme et la pratique de la modélisation des données dans le cadre de leurs activités quotidiennes, et s'adresse aux praticiens des données de niveau intermédiaire. La manifestation est soutenue par un certain nombre d'organisations parrainantes et de partenaires, dont la Division de statistique, la CESAO, l'OIT, l'UNESCO, la Banque centrale de la République de Türkiye, la BRI et le Centre de recherches statistiques, économiques et sociales et de formation pour les pays islamiques.

65. En 2023 et 2024, le Programme des Nations Unies pour l'environnement (PNUE) a dispensé une formation sur SDMX au Bénin et au Botswana dans le cadre

du programme de développement transversal des capacités financé par le Fonds pour l'environnement mondial. La formation s'est déroulée en deux phases, la première portant sur le Cadre pour le développement des statistiques de l'environnement et la seconde sur les systèmes d'information sur l'environnement, au cours de laquelle ont été présentés le modèle générique d'informations statistiques SDMX et une définition de structure de données pour les statistiques de l'environnement. Le PNUE devrait dispenser en décembre 2024 une formation SDMX, à l'intention de BPS-Statistics Indonesia, dans laquelle il mettra l'accent sur les indicateurs relatifs à l'environnement qui ont été recensés dans le pays à l'occasion du premier atelier sur le Cadre, organisé en août 2024.

XII. Décisions que la Commission de statistique est invitée à prendre

66. La Commission de statistique est invitée à prendre note du présent rapport.
-