

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

Paris, 9-11 avril 2014

Point 5 de l'ordre du jour provisoire

**Travaux du Groupe de haut niveau sur la modernisation
de la production et des services statistiques****Modèle générique du processus de production statistique****Document établi par le Groupe de haut niveau sur la modernisation
de la production et des services statistiques***Résumé*

Le présent document donne une description du modèle générique du processus de production statistique et de la manière dont il s'articule avec d'autres normes essentielles pour la modernisation des statistiques. Le modèle générique du processus de production statistique décrit et définit l'ensemble des processus d'activité nécessaires à l'établissement de statistiques officielles. Il offre un cadre type et une terminologie harmonisée pour aider les organismes de statistique à moderniser leurs processus de production statistique et à partager leurs méthodes et leurs composantes. Il peut aussi être utilisé pour intégrer les normes relatives aux données et aux métadonnées, servir de modèle pour les documents associés au processus, harmoniser les infrastructures d'informatique statistique et offrir un cadre à l'évaluation et à l'amélioration de la qualité du processus.

Le présent document est présenté à la Conférence des statisticiens européens pour information.



Table des matières

	<i>Page</i>
I. Introduction.....	3
A. Historique	3
B. Principaux changements par rapport à la version 4.0	4
II. Le modèle	4
A. Comprendre le modèle générique du processus de production statistique	4
B. Structure du modèle.....	5
C. Applicabilité	6
D. Utilisation du modèle générique du processus de production statistique	7
III. Articulation avec d'autres modèles et normes	7
A. Architecture commune de la production statistique.....	7
B. Modèle générique d'informations statistiques.....	8
C. Modèle de processus d'activité/générique longitudinal.....	9
IV. Niveaux 1 et 2 du modèle générique du processus de production statistique	10
V. Description des phases et des sous-processus	11
A. Phase de définition des besoins	11
B. Phase de conception	13
C. Phase d'élaboration	15
D. Phase de collecte.....	17
E. Phase de traitement.....	18
F. Phase d'analyse	21
G. Phase de diffusion.....	22
H. Phase d'évaluation.....	24
VI. Processus fondamentaux	25
A. Gestion de la qualité	25
B. Gestion des métadonnées.....	26
VII. Autres utilisations du Modèle générique du processus d'activité statistique	28
Annexe	
1. Liste des sigles et abréviations.....	30

I. Introduction

1. Le modèle générique du processus de production statistique (GSBPM) décrit et définit l'ensemble des processus d'activité nécessaires à l'établissement de statistiques officielles. Il offre un cadre type et une terminologie harmonisée pour aider les organismes de statistique à moderniser leurs processus de production statistique et à partager leurs méthodes et leurs composantes. Il peut aussi être utilisé pour intégrer les normes relatives aux données et aux métadonnées, servir de modèle pour les documents issus du processus, harmoniser les infrastructures d'informatique statistique, et offrir un cadre à l'évaluation et à l'amélioration de la qualité du processus. Ces finalités et d'autres encore pour lesquelles le GSBPM peut être utilisé sont examinées plus en détail dans la section VI. Cette version du GSBPM est entièrement alignée sur la version 1.1 du modèle générique d'informations statistiques (GSIM)¹ et sert de base à la mise en œuvre de l'architecture commune de la production statistique (CSPA)².

A. Historique

2. Les sessions de travail communes CEE/Eurostat/OCDE sur les métadonnées statistiques (METIS) ont établi un Cadre commun de métadonnées (CMF)³. La partie C de ce cadre est intitulée «Les métadonnées et le cycle statistique». Elle se rapporte aux phases du processus de production statistique (connu également sous le nom de chaîne de valeur statistique ou cycle statistique) et donne des termes génériques pour les décrire. Depuis novembre 2013, ces travaux sont pris en charge par le Comité de modernisation des normes sous la tutelle du Groupe de haut niveau sur la modernisation de la production et des services statistiques⁴.

3. Au cours d'un atelier consacré à la mise au point de la partie C du CMF, tenu à Vienne en juillet 2007⁵, les participants sont convenus que le modèle de processus d'activité utilisé par Statistics New Zealand constituerait un bon point de départ pour l'élaboration d'un modèle générique du processus de production statistique. Suite à plusieurs projets et consultations publiques, la version 4.0 du GSBPM a été publiée en avril 2009. Elle a depuis été largement adoptée par les organismes de statistique officiels de par le monde et constitue l'un des piliers de la vision et de la stratégie du Groupe de haut niveau pour une modernisation fondée sur des normes.

4. Un modèle complémentaire, le modèle générique d'informations statistiques (GSIM), est paru en décembre 2012 (pour plus de renseignements, voir la section III). Les travaux entrepris pour élaborer puis mettre en œuvre le GSIM ont permis d'identifier plusieurs améliorations possibles à apporter au GSBPM. En 2013, le Groupe de haut niveau a lancé un projet intitulé «Cadres et normes pour une modernisation de la statistique» qui prévoyait un examen plus étendu du GSBPM et du GSIM destiné à mieux harmoniser les documents associés aux deux modèles et à prendre en compte les retours d'information suscités par leurs applications pratiques. La version actuelle du GSBPM (la version 5.0)

¹ Voir l'adresse: <http://www1.unece.org/stat/platform/display/metis/Generic+Statistical+Information+Model>.

² Voir l'adresse: <http://www1.unece.org/stat/platform/display/CSPA/Common+Statistical+Production+Architecture+Home>.

³ Voir l'adresse: <http://www.unece.org/stats/cmfi/>.

⁴ Voir l'adresse: <http://www1.unece.org/stat/platform/display/hlgbas>.

⁵ Les documents issus de l'atelier sont disponibles à l'adresse: <http://www.unece.org/stats/documents/2007.07.metis.htm>.

résulte directement de ces travaux. Pour autant que cette version soit considérée comme définitive à sa date de publication, des mises à jour pourraient bien être nécessaires dans les années à venir, soit parce que l'on aura une expérience plus approfondie de la mise en pratique du modèle, soit parce que le caractère même de la production statistique aura évolué. Le lecteur est donc invité à consulter le site Web www.unece.org/stats/gsbpm pour s'assurer qu'il dispose de la version la plus récente.

B. Principaux changements par rapport à la version 4.0

5. Le Groupe de haut niveau avait donné des instructions claires quant au fait que l'examen de la version 4.0 du GSBPM devait se borner à introduire des changements pleinement justifiés d'un point de vue économique et bénéficiant d'un large soutien auprès des statisticiens internationaux. La raison en était que de nombreux organismes avaient lourdement investi dans la mise en œuvre du modèle. Tout remaniement important pourrait entraîner des coûts conséquents et, à terme, aller à l'encontre du but recherché, à savoir faire en sorte que le GSBPM soit largement adopté et utilisé. Ainsi, peu de modifications ont été apportées au modèle proprement dit, mais les documents connexes ont bénéficié en revanche de beaucoup d'autres améliorations, ajouts et éclaircissements.

6. Les principaux changements intervenus entre les versions 4.0 et 5.0 du GSBPM sont les suivants:

- La phase 8 (archivage) a été supprimée et incorporée dans le processus global de gestion des données et des métadonnées au motif que l'on peut archiver à tous les stades du processus de production statistique;
- Un nouveau sous-processus «Élaboration ou renforcement des composantes de diffusion» a été ajouté à la phase «élaboration» dans la mesure où l'on a de plus en plus intérêt à disposer de plusieurs options pour la diffusion;
- Plusieurs sous-processus ont été rebaptisés pour plus de clarté;
- Les descriptifs des sous-processus ont été actualisés et étoffés le cas échéant. La terminologie en usage a été modifiée pour être moins axée sur les enquêtes étant donné le recours croissant à d'autres sources (données administratives, données massives, etc.).

II. Le modèle

A. Comprendre le modèle générique du processus de production statistique

7. Le GSBPM devrait être appliqué et interprété avec souplesse. Ce n'est pas un cadre rigide dans lequel toutes les étapes doivent être suivies dans un ordre strict: il définit plutôt les étapes possibles du processus de production statistique ainsi que leurs interdépendances.

8. Bien que la présentation du GSBPM suive l'ordre logique des étapes inscrites dans la plupart des cycles statistiques, les éléments du modèle peuvent intervenir dans des ordres différents suivant les circonstances. Par ailleurs, certains sous-processus peuvent être reproduits plusieurs fois, formant des boucles itératives, en particulier lors des phases de traitement et d'analyse.

9. Le GSBPM devrait donc plutôt être considéré comme une matrice autorisant une grande variété de cheminements. En ce sens, ce modèle se veut suffisamment général pour pouvoir s'appliquer largement et vise à encourager une conception normalisée du processus de production statistique sans pour autant devenir trop restrictif ou trop abstrait et théorique.

B. Structure du modèle

10. Le GSBPM comprend trois niveaux:

- Le niveau 0, qui est celui du processus de production statistique;
- Le niveau 1, qui correspond aux huit phases du processus de production statistique;
- Le niveau 2, qui correspond aux sous-processus de chaque phase.

11. Un diagramme indiquant les phases (niveau 1) et les sous-processus (niveau 2) est reproduit dans la section IV. Les sous-processus sont décrits de façon détaillée dans la section V.

12. Le GSBPM prend également en compte plusieurs processus fondamentaux qui s'appliquent tout au long des huit phases et s'étendent aux divers processus de production statistique. Ils se classent dans deux catégories, ceux qui ont une composante statistique, et ceux qui sont plus généraux et pourraient s'appliquer à tout type d'organisation. Le premier groupe est jugé plus important dans le contexte du GSBPM, mais le deuxième compte aussi dans la mesure où il influe (souvent indirectement) sur plusieurs parties du modèle.

13. Les processus fondamentaux qui ont une composante statistique sont énumérés ci-dessous. Les quatre premiers sont les plus étroitement associés au modèle. La gestion de la qualité et la gestion des métadonnées sont particulièrement mises en évidence dans les diagrammes type et sont décrites de façon plus détaillée dans la section VI.

- Gestion de la qualité – Ce processus comprend les mécanismes d'évaluation et de contrôle de la qualité. Il reconnaît l'importance de l'évaluation et du retour d'information pendant toute la durée du cycle statistique;
- Gestion des métadonnées – Des métadonnées sont produites et traitées durant chaque phase, aussi est-il vraiment nécessaire de prévoir un système de gestion de ces informations pour s'assurer que les métadonnées pertinentes restent liées aux données qu'elles décrivent dans tout le GSBPM. Des considérations indépendantes du processus interviennent aussi ici, notamment, concernant les métadonnées, les questions de garde et de propriété, de qualité, de règles en matière d'archivage, de préservation, de conservation et d'élimination;
- Gestion des données – Des considérations indépendantes du processus interviennent dans ce cadre, notamment, concernant les données générales, les questions de sécurité, de garde et de propriété, de qualité des données, de règles en matière d'archivage, de préservation, de conservation et d'élimination;
- Gestion des données de processus – Ce processus comprend la gestion des données et des métadonnées qui sont produites à tous les stades du cycle statistique et qui fournissent des informations à leur sujet;
- Gestion des connaissances – Ce processus assure la répétabilité des cycles statistiques, principalement par la tenue des documents connexes;
- Gestion du cadre statistique – Ce processus comprend l'élaboration de normes, par exemple de méthodologies, de concepts et de classifications qui s'appliquent à une multiplicité de processus;
- Gestion du programme statistique – Ce processus englobe le suivi et l'examen systématiques des nouveaux besoins en matière d'information et des sources de données nouvelles et évolutives dans tous les domaines statistiques. Il peut aboutir à la définition de nouveaux processus de production statistique ou à la refonte des processus existants;

- Gestion des fournisseurs – Ce processus vise la gestion de la charge de travail d'un processus à l'autre et des questions comme l'établissement de profils et la gestion des coordonnées (il intéresse donc au premier chef les processus de production statistique prévoyant la tenue de registres);
 - Gestion des clients – Ce processus comprend les activités de commercialisation au sens large, les mesures prises pour mieux faire connaître la statistique et la suite donnée aux réactions des clients ne portant pas sur des points spécifiques.
14. Les processus fondamentaux de portée plus générale sont les suivants:
- Gestion des ressources humaines;
 - Gestion financière;
 - Gestion de projets;
 - Gestion du cadre juridique;
 - Gestion du cadre organisationnel;
 - Planification stratégique.

C. Applicabilité

15. Le GSBPM entend s'appliquer à toutes les activités menées par les instances chargées d'établir les statistiques officielles à l'échelle nationale et internationale, lesquelles aboutissent à la production de données. Il est conçu pour être indépendant de la source de données, et peut donc être utilisé pour la description de processus fondés sur des enquêtes, des recensements, des fichiers administratifs et d'autres sources autres que statistiques ou mixtes, et l'évaluation de leur qualité.

16. Si, de façon générale, le processus de production statistique consiste à collecter et traiter des données dans le but d'élaborer des produits statistiques, le GSBPM s'applique aussi aux cas où l'on doit réviser des données existantes et recalculer des séries temporelles, soit parce que les sources de données se sont améliorées, soit parce que l'on a changé de méthode dans l'intervalle. Dans ces cas, les données d'entrée sont des statistiques déjà publiées, qui sont à nouveau traitées et analysées pour donner des produits révisés. Il est probable que plusieurs sous-processus, voire même peut-être certaines phases (en particulier celles qui figurent le plus en amont), seront omis dans ce contexte. Le GSBPM peut aussi s'appliquer à des exercices comme l'établissement des comptabilités nationales et aux activités courantes des organismes de statistique internationaux.

17. En dehors des processus d'élaboration de statistiques, le GSBPM peut aussi s'utiliser pour établir et tenir des registres statistiques pour lesquels les intrants sont semblables à ceux exploités pour la production statistique (bien que généralement les données administratives y tiennent une plus grande place) et les produits sont le plus souvent des bases ou autres formes d'extraction de données, qui vont ensuite alimenter d'autres processus⁶.

18. Le GSBPM devrait être perçu comme suffisamment souple pour s'appliquer à tous les scénarios qui précèdent.

⁶ Voir l'adresse: <http://www1.unece.org/stat/platform/download/attachments/57835551/BR+meeting+paper2.doc>.

D. Utilisation du modèle générique du processus de production statistique

19. Le GSBPM est un modèle de référence. Il est conçu pour être utilisé par les organismes à divers degrés. Un organisme peut choisir soit de l'appliquer directement soit de l'utiliser comme point de départ pour en établir une version adaptée spécifique à ses besoins. Dans certains cas, il ne peut servir que de modèle auquel les organismes se réfèrent pour mieux se comprendre lorsqu'ils communiquent en interne ou avec des homologues. Tous ces scénarios d'utilisation du GSBPM sont valables.

20. Lorsque des organismes ont mis au point des adaptations spécifiques du GSBPM, ils peuvent introduire dans le modèle des spécialisations qui correspondent à leur contexte propre. D'après les constatations qui ont pu être faites à ce jour, ces spécialisations ne sont pas suffisamment génériques pour être incorporées dans le modèle initial.

21. Dans certains cas, il peut être approprié de regrouper certains des éléments du modèle. Par exemple, on pourrait considérer que les phases un à trois relèvent d'une seule phase de planification. Dans d'autres cas, en particulier dans la pratique, il pourrait être nécessaire d'ajouter un ou plusieurs niveaux détaillés à la structure présentée ci-dessous pour bien distinguer les diverses composantes des sous-processus.

22. Il pourrait aussi être nécessaire de clore officiellement une phase avant d'entamer la suivante, à savoir certifier les produits de la première phase comme constituant des intrants valables pour la deuxième. Cet aval officiel est implicite dans le modèle, mais peut s'appliquer de différentes façons selon les besoins de l'organisme concerné.

III. Articulation avec d'autres modèles et normes

23. Le GSBPM a été établi initialement sur la base du modèle générique de processus d'activité de Statistics New Zealand complété par les apports d'autres organismes de statistique possédant une expérience de la modélisation du cycle. Il existe toutefois plusieurs autres modèles et normes connexes qui résultent soit d'initiatives de modernisation soit d'activités analogues n'entrant pas dans le champ des statistiques officielles.

A. Architecture commune de la production statistique

24. Si l'on examine pour commencer les liens entre le GSBPM et les autres cadres et normes nécessaires à la modernisation des statistiques, il peut être utile de les replacer dans le cadre de l'architecture d'entreprise. Cette dernière articule les fonctions et les processus d'activité et les éléments d'information, les applications et les technologies nécessaires à leur déroulement. C'est un moyen de décrire ce que fait un organisme, et de quelle manière, pour tenter de déterminer comment il pourrait améliorer la qualité et l'efficacité de ses prestations.

25. Pour ce qui est de la modernisation des statistiques, son but est d'aligner les architectures d'entreprise de divers organismes en créant une «architecture de secteur» pour l'ensemble du «secteur statistique officiel». Cette approche vise à faciliter la collaboration, le partage et la mise au point conjointe des composantes et des services qui sont nécessaires aux différentes étapes du cycle statistique (définies en rapport avec le GSBPM). Elle a donné naissance à l'architecture commune de la production statistique (CSPA) qui a été publiée pour la première fois à la fin de 2013⁷.

⁷ Voir l'adresse: <http://www1.unece.org/stat/platform/display/CSPA/Common+Statistical+Production+Architecture+Home>.

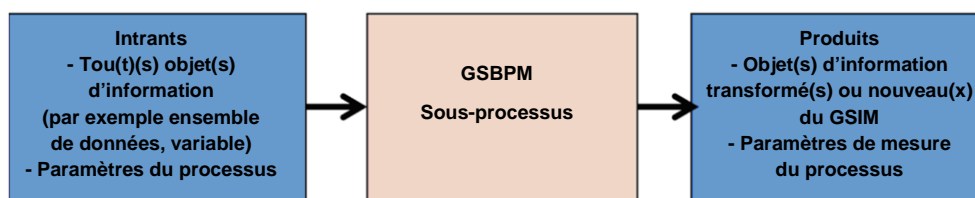
B. Modèle générique d'informations statistiques

26. Le modèle générique d'informations statistiques (GSIM) est un cadre de référence pour l'information statistique conçu de façon à jouer un rôle important dans la modernisation et la simplification des statistiques officielles tant au niveau national qu'international. Il permet de faire une description générale de la définition, de la gestion et de l'utilisation des données et des métadonnées tout au long du processus de production de statistiques. Il offre un ensemble d'objets d'information normalisés, décrits de façon uniforme, qui sont les intrants et les produits dans le processus de la conception et de la production de statistiques. Le GSIM aide à mieux comprendre les liens importants qui existent entre les entités prenant part à la production de statistiques, et peut servir à encadrer l'élaboration et l'utilisation de normes ou de spécifications cohérentes en matière de mise en œuvre.

27. Comme le GSBPM, le GSIM est l'une des pièces maîtresses de la modernisation des statistiques officielles et de l'abandon progressif du cloisonnement par domaine statistique. Le GSIM est conçu pour permettre autant que possible l'adoption d'approches novatrices de la production statistique, par exemple, dans le domaine de la diffusion, où les impératifs de flexibilité et d'innovation s'accroissent. Il appuie aussi les approches suivies actuellement pour la production statistique.

28. Le GSIM recense environ 110 objets d'information, dont par exemple les ensembles de données, les variables, les classifications statistiques, les unités, les populations ainsi que les règles et les paramètres nécessaires au déroulement des processus de production (dont les règles d'édition des données).

29. Le GSIM et le GSBPM sont des modèles complémentaires pour la production et la gestion de l'information statistique. Comme le montre le diagramme ci-dessous, le GSIM sert à décrire les sous-processus du GSBPM en définissant les objets d'information qui circulent entre eux, sont produits par eux et sont utilisés par eux pour produire des statistiques officielles. Les intrants et les produits peuvent se définir comme objets d'information et sont mis en forme selon le GSIM.



30. On peut tirer un meilleur profit du GSIM s'il est appliqué en conjonction avec le GSBPM. Le GSBPM se montre aussi plus utile s'il est appliqué en conjonction avec le GSIM. Il est néanmoins possible (même si ce n'est pas la solution optimale) d'appliquer l'un sans l'autre. De même, les deux modèles appuient la mise en œuvre de la CSPA, mais peuvent s'appliquer que l'on emploie ce cadre architectural ou non.

31. De même que les processus de production statistique pris individuellement ne comportent pas tous les sous-processus prévus par le GSBPM, il est très peu probable que tous les objets d'information du GSIM soient nécessaires dans un processus de production statistique donné.

32. L'application conjointe du GSBPM et du GSIM peut faciliter l'élaboration de systèmes efficaces pilotés par les métadonnées et aider à l'harmonisation d'infrastructures de calcul statistique.

C. Modèle de processus d'activité/générique longitudinal

33. Si l'on s'écarte du domaine des statistiques officielles, les chercheurs qui réalisent des enquêtes sociales ont mis au point le modèle de processus d'activité longitudinal générique (GLBPM)⁸ «pour fournir un modèle générique propre à enrichir le dialogue entre les organismes qui procèdent à des collectes de données longitudinales et d'autres collectes de données répétées dans le temps».

34. Comme le GSBPM, le GLBPM est un modèle de référence pour la configuration de processus d'activité effectifs. Il sert de référence aux organismes de recherche pour l'exercice de collecte de données longitudinales et transversales répétées. Il décrit les activités entreprises en corrélation avec leurs intrants et produits courants.

35. Le GLBPM présente beaucoup de similitudes avec le GSBPM bien qu'il diffère pour certaines activités particulières correspondant aux différents besoins et pratiques des chercheurs en sciences statistiques et sociales. Il opte pour l'approche de la trajectoire non linéaire au travers d'une matrice d'options possibles résultant directement du GSBPM.

⁸ Voir l'adresse: <http://dx.doi.org/10.3886/DDILongitudinal05>.

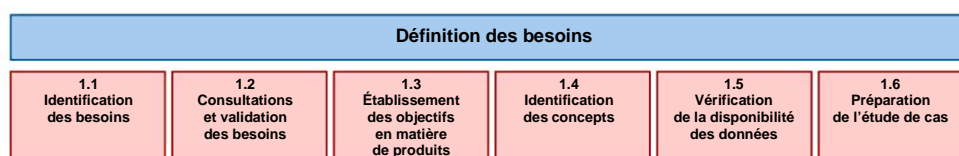
IV. Niveaux 1 et 2 du modèle générique du processus de production statistique

Gestion de la qualité/Gestion des métadonnées							
Définition des besoins	Conception	Élaboration	Collecte	Traitement	Analyse	Diffusion	Évaluation
1.1 Identification des besoins	2.1 Conception des produits	3.1 Élaboration de l'instrument de collecte	4.1 Création d'un cadre et sélection d'un échantillon	5.1 Intégration des données	6.1 Élaboration du projet de produits	7.1 Actualisation des systèmes de produits	8.1 Recueil des produits d'évaluation
1.2 Consultations et validation des besoins	2.2 Conception des descriptions de variables	3.2 Élaboration ou renforcement des composantes du processus	4.2 Organisation de la collecte	5.2 Classification et codage	6.2 Validation des produits	7.2 Élaboration des produits de diffusion	8.2 Conduite de l'évaluation
1.3 Établissement des objectifs en matière de produits	2.3 Conception de la collecte	3.3 Élaboration ou renforcement des composantes de diffusion	4.3 Réalisation de la collecte	5.3 Examen et validation	6.3 Interprétation et explication des produits	7.3 Gestion de la publication des produits de diffusion	8.3 Adoption d'un plan d'action
1.4 Identification des concepts	2.4 Conception du cadre et de l'échantillon	3.4 Configuration du déroulement des travaux	4.4 Finalisation de la collecte	5.4 Édition et imputation des données	6.4 Mise en place du contrôle de la divulgation	7.4 Promotion des produits de diffusion	
1.5 Vérification de la disponibilité des données	2.5 Conception du traitement et de l'analyse	3.5 Mise à l'essai du système de production		5.5 Calcul de nouvelles variables et unités	6.5 Finalisation des produits	7.5 Gestion de l'assistance aux utilisateurs	
1.6 Préparation de l'étude de cas	2.6 Conception de systèmes de production et du déroulement des travaux	3.6 Mise à l'essai du processus de production statistique		5.6 Calcul des coefficients de pondération			
		3.7 Finalisation du système de production		5.7 Calcul des agrégats			
				5.8 Finalisation des fichiers de données			

V. Description des phases et des sous-processus

36. La présente section examine chaque phase en succession en énumérant les divers sous-processus qu'elle comporte et en décrivant le contenu.

A. Phase de définition des besoins



37. Cette phase se déclenche lorsque le besoin de nouvelles statistiques se fait sentir, ou lorsque les réactions suscitées par les statistiques existantes poussent à entreprendre une étude. Elle comprend toutes les activités qui consistent à amener les clients à cerner leurs besoins statistiques dans le détail, à proposer des solutions possibles de haut niveau et à préparer des études de cas pour répondre à ces besoins.

38. Pendant cette phase, l'organisme concerné:

- Identifie le besoin en matière de statistiques;
- Confirme, de façon plus détaillée, les besoins des parties prenantes en matière de statistiques;
- Établit les objectifs de haut niveau des produits statistiques;
- Identifie les concepts et variables pertinents pour lesquels on a besoin de données;
- Vérifie la mesure dans laquelle les sources de données existantes peuvent répondre à ces besoins;
- Prépare l'étude de cas pour obtenir l'autorisation de produire les statistiques.

39. Cette phase se divise en six sous-processus. Ces derniers sont généralement séquentiels, de gauche à droite, mais peuvent aussi se dérouler en parallèle et être itératifs. Ces sous-processus sont les suivants:

1. Identification des besoins

40. Ce sous-processus comprend l'enquête initiale, l'identification des statistiques nécessaires et la définition de ce que l'on en attend. Il peut être engagé suite à une nouvelle demande d'information ou un changement contextuel tel qu'une réduction budgétaire. Des plans d'action établis à partir d'évaluations des itérations antérieures du processus en question ou d'autres peuvent servir d'intrants. Ce sous-processus prévoit aussi l'examen des pratiques en vigueur dans d'autres organismes de statistique (nationaux et internationaux) qui produisent des données similaires et, en particulier, des méthodes utilisées par ces organismes. Il peut prévoir l'étude des besoins particuliers de différentes communautés d'utilisateurs comme les handicapés, ou différents groupes ethniques.

2. Consultations et validation des besoins

41. Ce sous-processus met l'accent sur les consultations avec les parties prenantes et la validation détaillée des besoins en matière de statistiques. Une bonne compréhension des besoins des utilisateurs est indispensable pour que l'organisme de statistique sache non

seulement ce qui est attendu de lui, mais quand, comment, et, plus important encore peut-être, pourquoi il lui est demandé d'intervenir. Pour la deuxième itération de cette phase et les suivantes, on s'attachera à déterminer si les besoins précédemment identifiés ont évolué. Cette compréhension étroite des besoins des utilisateurs est un élément essentiel de ce sous-processus.

3. Établissement des objectifs en matière de produits

42. Ce sous-processus identifie les produits statistiques requis pour répondre aux besoins des utilisateurs définis lors du sous-processus 1.2 (Consultations et validation des besoins). C'est à ce stade que l'on convient avec les utilisateurs de la pertinence des produits proposés et du niveau de leur qualité. Les cadres juridiques (par exemple, en rapport avec la confidentialité) et les ressources disponibles ont toutes les chances de poser des difficultés au moment de fixer les objectifs en matière de produits.

4. Identification des concepts

43. Ce sous-processus permet de clarifier les concepts requis que le processus d'activité évaluera du point de vue de l'utilisateur. À ce stade, les concepts identifiés peuvent ne pas s'aligner sur les normes statistiques existantes. Cet alignement, ainsi que le choix ou la définition des concepts et variables statistiques à utiliser, s'opère dans le cadre du sous-processus 2.2.

5. Vérification de la disponibilité des données

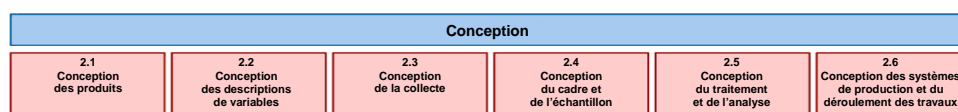
44. Ce sous-processus permet de vérifier si les sources de données existantes peuvent répondre aux besoins des utilisateurs, et d'établir les conditions dans lesquelles elles seraient disponibles, en tenant compte des restrictions qui pourraient peser sur leur utilisation. L'évaluation des autres options possibles consiste ordinairement à rechercher des sources de données administratives ou autres sources de données non statistiques potentielles pour déterminer si ces dernières peuvent servir à des fins statistiques. Lorsque les sources existantes ont été évaluées, une stratégie est élaborée pour répondre aux besoins en matière de données qui restent éventuellement à pourvoir. Ce sous-processus prévoit aussi une évaluation plus générale du cadre juridique dans lequel les données seraient collectées et utilisées, et peut donc permettre de dégager des propositions de modifications de la législation existante ou d'introduction d'un nouveau cadre juridique.

6. Préparation de l'étude de cas

45. Ce sous-processus rassemble les conclusions des autres sous-processus de la phase considérée sous la forme d'une étude de cas qui sert à demander l'autorisation de mise en marche du processus de production statistique nouveau ou modifié. Cette étude de cas devra se conformer aux prescriptions de l'organisme chargé d'approuver le projet mais comprend généralement des éléments tels que les suivants:

- Un descriptif du processus de production statistique «actuel» (s'il en existe déjà un), contenant des renseignements sur la manière dont les statistiques existantes sont produites et soulignant les éventuelles insuffisances et questions à régler;
- La solution «future» proposée, assortie d'une présentation détaillée de la manière dont le processus de production statistique sera mis au point pour produire les statistiques nouvelles ou révisées requises;
- Une analyse coûts-avantages, ainsi qu'une évaluation des éventuelles contraintes extérieures.

B. Phase de conception



46. Cette phase décrit les activités de développement et de conception et tous les travaux de recherche pratiques connexes nécessaires pour définir les produits, concepts, méthodes, instruments de collecte⁹ et processus opérationnels statistiques. Elle englobe tous les éléments de conception nécessaires pour définir ou affiner les produits ou services statistiques mis à jour dans l'étude de cas. Elle précise toutes les métadonnées pertinentes prêtes à être utilisées ultérieurement dans le processus de production statistique ainsi que les procédures d'assurance de la qualité. Pour les produits statistiques élaborés à intervalles réguliers, cette phase intervient lors de la première itération, et chaque fois que des mesures propres à améliorer le processus sont identifiées dans la phase d'évaluation de l'itération précédente.

47. Les activités de conception s'appuient pour une large part sur les normes internationales et nationales afin de réduire la longueur et le coût du processus de conception et de renforcer la comparabilité et la possibilité d'utilisation des produits. Les organismes sont aussi incités à réutiliser ou adapter des éléments de conception des processus existants. En outre, les produits des processus de conception peuvent servir de base aux normes futures à l'échelle des organismes et à l'échelle nationale ou internationale.

48. Cette phase se divise en six sous-processus, qui sont généralement séquentiels, de gauche à droite, mais peuvent aussi se dérouler en parallèle et peuvent être itératifs. Ces sous-processus sont les suivants:

1. Conception des produits

49. Ce sous-processus englobe la conception détaillée des résultats statistiques, produits et services à obtenir, y compris les travaux de développement connexes et la préparation des systèmes et des outils employés pendant la phase de «diffusion». Les méthodes de contrôle de la divulgation, ainsi que les processus régissant l'accès aux éventuels produits confidentiels sont aussi mis au point dans ce cadre. Les produits devraient chaque fois que possible être conçus de façon à respecter les normes existantes, de sorte que les intrants fournis à ce processus peuvent être des métadonnées provenant de collectes analogues ou précédentes, des normes internationales et des informations sur les pratiques en vigueur dans d'autres organismes de statistique résultant du sous-processus 1.1 (Identification des besoins).

2. Conception des descriptions de variables

50. Ce sous-processus définit les variables statistiques à recueillir au moyen de l'instrument de collecte, ainsi que toutes les autres variables qui en seront tirées dans le cadre du sous-processus 5.5 (Calcul de nouvelles variables et unités) et toutes les classifications statistiques qui seront utilisées. On s'attache en principe à respecter les normes nationales et internationales existantes chaque fois que possible. Ce sous-processus devrait peut-être se dérouler en parallèle au sous-processus 2.3 (Conception de la collecte) car la définition des variables à collecter et le choix des instruments de collecte peuvent être interdépendants dans une certaine mesure. La préparation des descriptions de variables et classifications recueillies et obtenues sous forme de métadonnées est un préalable nécessaire aux phases suivantes.

⁹ Aux fins du GSBPM, on entend généralement par instruments de collecte tout outil ou programme destiné à recueillir ou extraire des données et des métadonnées, du questionnaire sur support papier à l'outil de «web-scraping» (extraction de données de sites Web). Dans la version 1.1 du GSIM, les instruments de collecte sont les «circuits d'échanges» utilisés pour l'information entrante.

3. Conception de la collecte

51. Ce sous-processus détermine la ou les méthode(s) et instrument(s) les plus appropriés. Les activités menées effectivement au cours de ce sous-processus dépendront des types d'instrument de collecte requis qui peuvent être des entretiens assistés par ordinateur, des questionnaires sur support papier, des interfaces de données administratives et des techniques d'intégration de données. Ce sous-processus englobe la conception d'instruments de collecte, de modèles de questions-réponses (en conjonction avec les variables et les classifications statistiques établies dans le cadre du sous-processus 2.2 (Conception des descriptions de variables)). Il comprend aussi la conception de tous les accords formels en vue de la fourniture de données, comme les mémorandums d'accord et la validation des bases juridiques de la collecte des données. Ce sous-processus est facilité par des outils tels que les collections de questions (qui permettent de réutiliser des questions et leurs caractéristiques), les outils de création de questionnaire (qui permettent de compiler rapidement et facilement des questions dans des formats adaptés aux essais cognitifs) et les modèles d'accord (pour faciliter la normalisation des conditions et modalités). Ce sous-processus comprend aussi la conception de systèmes de gestion de fournisseurs adaptés au processus.

4. Conception du cadre et de l'échantillon

52. Ce sous-processus s'applique seulement aux processus prévoyant une collecte de données par échantillonnage, notamment au moyen d'études statistiques. Il permet d'identifier et de désigner la population visée, de définir une base de sondage (et, le cas échéant, le registre dont elle provient) et de déterminer les critères et les méthodes d'échantillonnage les plus appropriés (dont éventuellement le recensement complet). Les sources communes utilisées pour établir la base de sondage sont les registres administratifs et statistiques, les recensements et les informations tirées d'autres enquêtes par sondage. Ce sous-processus décrit la manière de conjuguer ces sources si nécessaire. Il faut examiner à ce stade la mesure dans laquelle le cadre recouvre la population visée. Un plan de sondage devrait être établi, l'échantillon à étudier étant désigné dans le cadre du sous-processus 4.1 (Création d'un cadre et sélection d'un échantillon) en utilisant la méthode énoncée dans ce sous-processus.

5. Conception du traitement et de l'analyse

53. Ce sous-processus englobe la conception de la méthode de traitement statistique qui doit être appliquée au cours des phases de «traitement» et d'«analyse». Il peut comprendre la définition des routines de codage, d'édition, d'imputation, d'estimation, d'intégration, de validation et de finalisation des ensembles de données.

6. Conception des systèmes de production et du déroulement des travaux

54. Ce sous-processus détermine le déroulement des travaux du stade de la collecte au stade de la diffusion des données, en faisant le tour de tous les processus requis par l'ensemble du cycle statistique, et en veillant à ce que ces derniers puissent s'enchaîner efficacement sans lacunes ni redondances. Divers systèmes et bases de données sont nécessaires tout au long du cycle. En général, on réutilise les mêmes processus et techniques pour de nombreux processus de production statistique, de sorte qu'il faut commencer par examiner les produits existants (par exemple, des services, des systèmes et des bases de données) pour voir s'ils conviennent en l'espèce, et dans les cas où ils ne donnent pas entièrement satisfaction, chercher de nouvelles solutions. Ce sous-processus devrait aussi permettre d'étudier les modalités d'interaction du personnel avec ces systèmes et de définir les responsabilités de chaque agent et le stade auquel il intervient.

C. Phase d'élaboration

Élaboration						
3.1 Élaboration de l'instrument de collecte	3.2 Élaboration ou renforcement des composantes du processus	3.3 Élaboration ou renforcement des composantes de diffusion	3.4 Configuration du déroulement des travaux	3.5 Mise à l'essai du système de production	3.6 Mise à l'essai du processus de production statistique	3.7 Finalisation du système de production

55. Cette phase permet d'élaborer et de mettre à l'essai le produit retenu jusqu'à ce qu'il soit prêt à être utilisé dans des conditions réelles. Les produits de la phase de «conception» orientent les choix de processus, d'instruments, d'informations et de services réutilisables qui sont regroupés et configurés à ce stade pour créer l'environnement opérationnel complet permettant de lancer le processus. De nouveaux services sont élaborés par exception, pour combler les lacunes constatées dans le catalogue de services existant provenant de l'organisme lui-même ou de l'extérieur. Ces nouveaux services sont conçus de façon à être en grande partie réutilisables dans l'architecture de production statistique.

56. Pour les produits statistiques élaborés à intervalles réguliers, cette phase intervient habituellement à la première itération ou suite à un réexamen ou un changement de méthode ou de technique plutôt qu'à chaque itération.

57. Cette phase se divise en sept sous-processus, qui sont généralement séquentiels, de gauche à droite, mais peuvent aussi se dérouler en parallèle et être itératifs. Ces sous-processus sont les suivants:

1. Élaboration de l'instrument de collecte

58. Ce sous-processus décrit les activités visant à concevoir les instruments à utiliser au cours de la phase de «collecte». L'instrument de collecte est produit ou élaboré d'après le cahier des charges établi au cours de la phase de «conception». Une collecte peut faire appel à un ou plusieurs modes de réception des données, par exemple des interrogatoires directs ou des entretiens téléphoniques; des questionnaires sur support papier, sur support électronique ou sur Internet; des plates-formes SDMX. On peut aussi utiliser comme instruments de collecte les routines d'extraction de données utilisées pour réunir des informations issues d'ensembles de données statistiques ou administratifs existants. Ce sous-processus prévoit aussi la préparation et la mise à l'essai du contenu et du fonctionnement de l'instrument choisi (par exemple, mise à l'essai des questions figurant dans un questionnaire). Il est recommandé d'étudier le lien direct entre les instruments de collecte et le système de métadonnées statistiques, de façon à faciliter le recueil de métadonnées pendant la phase de collecte. La corrélation des métadonnées et des données au point de saisie permet de réduire la charge de travail lors des phases suivantes. La saisie des paramètres de mesure de la collecte de données (paradonnées) revêt aussi de l'importance pour ce sous-processus.

2. Élaboration ou renforcement des composantes du processus

59. Ce sous-processus décrit les activités visant à élaborer de nouveaux services et composantes ou à renforcer les composantes et services existants pour les phases de «traitement» et d'«analyse», telles qu'elles ont été conçues lors de la phase de «conception». Parmi ces services on peut citer les usages et fonctionnalités des tableaux, les services d'information, les fonctions de transformation, les organigrammes des tâches, les services de gestion des fournisseurs et des métadonnées.

3. Élaboration ou renforcement des composantes de diffusion

60. Ce sous-processus décrit les activités visant à construire de nouveaux services et composantes ou à renforcer les composantes et services existants pour la diffusion de produits statistiques, telles qu'elles ont été conçues lors du sous-processus 2.1 (Conception des produits). Tous les types de composantes et de services de diffusion sont inclus, de ceux utilisés pour éditer des publications traditionnelles sur support papier à ceux qui fournissent des services Web, des données en libre accès ou qui ouvrent l'accès à des microdonnées.

4. Configuration du déroulement des travaux

61. Ce sous-processus permet de configurer le déroulement des travaux, les systèmes et les transformations utilisés dans le cadre des processus de production statistique, de la collecte des données à leur diffusion. Il fait en sorte que le déroulement des travaux défini lors du sous-processus 2.6 (Conception de systèmes de production et du déroulement des travaux) fonctionne dans la pratique.

5. Mise à l'essai du système de production

62. Ce sous-processus se rapporte à la mise à l'essai des services regroupés et configurés et du déroulement des travaux qui en résulte. Il comprend les essais techniques et l'agrément de nouveaux programmes et routines et permet de valider l'utilisation pour le processus en cours des routines existantes issues d'autres processus de production statistique. Bien qu'une partie de cette activité de mise à l'essai de composantes et de services individuels puisse logiquement être reliée au sous-processus 3.2 (Élaboration ou renforcement des composantes du processus), ce sous-processus permet aussi de mettre à l'essai les interactions entre les services regroupés et configurés et de veiller à ce que la solution retenue fasse jouer un ensemble cohérent de processus, d'informations et de services.

6. Mise à l'essai du processus de production statistique

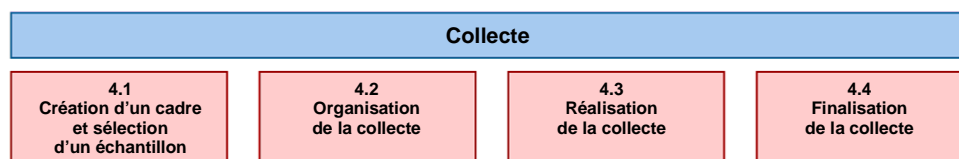
63. Ce sous-processus décrit les activités de gestion d'un essai sur le terrain ou d'un essai pilote du processus de production statistique. Il prévoit en général une collecte de données à petite échelle, qui permet de tester les instruments de collecte, puis le traitement et l'analyse des données recueillies, et ce, afin de s'assurer que le cycle statistique fonctionne comme prévu. Après la phase expérimentale, il peut être nécessaire de revenir un cran en arrière pour ajuster les instruments, systèmes ou composantes. Pour un processus de production statistique de grande envergure, un recensement de population par exemple, plusieurs itérations peuvent être nécessaires avant que le processus ne fonctionne de manière satisfaisante.

7. Finalisation du système de production

64. Ce sous-processus comprend les activités visant à mettre en phase de production, prêts à être utilisés par domaine d'activité, les processus et services regroupés et configurés, y compris les services qui ont été modifiés et les nouveaux services. Il faut à cette fin:

- Établir la documentation sur les composantes du processus, y compris la documentation technique et les manuels d'utilisation;
- Former les utilisateurs professionnels aux modalités de fonctionnement du processus;
- Mettre les composantes du processus dans l'environnement de production statistique, et veiller à ce qu'elles s'y comportent comme prévu (cette activité peut aussi s'inscrire dans le cadre du sous-processus 3.5 (Mise à l'essai du système de production)).

D. Phase de collecte



65. Cette phase permet de collecter ou de recueillir toutes les informations nécessaires (données et métadonnées), en utilisant des modes de collecte différents (y compris des extractions de registres et bases de données statistiques, administratifs et d'autres registres et bases de données non statistiques) et de les transférer dans l'environnement approprié pour la suite du traitement. Si la validation de formats d'ensembles de données peut entrer dans ce cadre, il n'en est pas de même de la transformation des données elles-mêmes, car cette tâche relève entièrement de la phase de «traitement». Pour les produits statistiques élaborés à intervalles réguliers, cette phase intervient à chaque itération.

66. La phase de «collecte» se divise en quatre sous-processus, qui sont généralement séquentiels, de gauche à droite, mais peuvent aussi se dérouler en parallèle et être itératifs. Ces sous-processus sont les suivants:

1. Création d'un cadre et sélection d'un échantillon

67. C'est au titre de ce sous-processus que l'on établit le cadre et sélectionne l'échantillon pour cette itération de la collecte comme indiqué dans le sous-processus 2.4 (Conception du cadre et de l'échantillon). Il en est de même pour la coordination d'échantillons entre reprises du même processus de production statistique (qui permet par exemple de gérer les chevauchements ou les rotations) et entre différents processus mettant en œuvre un cadre ou un registre commun (le but étant par exemple de gérer les chevauchements ou de répartir la charge de travail des répondants). Les questions d'assurance et d'agrément qualité du cadre et de l'échantillon retenu relèvent aussi de ce sous-processus bien que la tenue des registres sur lesquels reposent les cadres de plusieurs processus de production statistique soit considérée à part. La partie échantillonnage de ce sous-processus n'est généralement pas utile lorsque l'on n'utilise que des sources préexistantes (par exemple, des sources administratives) car les processus de ce type créent généralement des cadres à partir des données disponibles puis suivent la méthode du recensement.

2. Organisation de la collecte

68. Ce sous-processus fait en sorte que les personnes, les processus et les technologies soient prêts à collecter des données et des métadonnées sous tous les modes désignés. Il prend un certain temps, car il comprend les activités liées à la stratégie, la planification et la formation menées pour préparer le processus de production statistique en question. Dans les cas où le processus se répète à intervalles réguliers, certaines de ces activités (voire leur totalité) peuvent ne pas être explicitement requises pour chaque itération. En cas de processus ponctuel ou de nouveau processus, ces activités peuvent durer longtemps. Les tâches ci-après s'effectuent dans ce cadre:

- Établir une stratégie de collecte;
- Former le personnel chargé de la collecte;
- Veiller à disposer de moyens de collecte, par exemple des ordinateurs portables;
- Convenir de modalités avec des organes de collecte intermédiaires, par exemple des sous-traitants pour la conduite d'entretiens téléphoniques assistés par ordinateur;
- Configurer des systèmes de collecte pour demander et recevoir les données;
- Veiller à la sécurité des données à recueillir;

- Préparer des instruments de collecte (par exemple, impression de questionnaires, préremplissage de ces questionnaires avec les données existantes, téléchargement des questionnaires et des données vers les ordinateurs des enquêteurs, etc.).

69. Pour les sources autres que les enquêtes, ce sous-processus consistera notamment à faire en sorte que les processus, systèmes et procédures de confidentialité nécessaires soient en place pour obtenir ou extraire de la source les informations dont on a besoin.

3. Réalisation de la collecte

70. Ce sous-processus est celui dans le cadre duquel la collecte est mise en œuvre, les différents instruments étant utilisés pour collecter ou recueillir l'information, qui peut prendre la forme de microdonnées brutes ou d'agrégats produits à la source, et de toutes les métadonnées associées. Il comprend le contact initial avec les fournisseurs et les mesures de suivi ou de rappel ultérieures éventuelles. Il peut inclure l'entrée manuelle des données au point de contact, ou la gestion du travail d'enquête, suivant la source et le mode de collecte. C'est à ce stade que l'on enregistre la date à laquelle les fournisseurs ont été contactés, la manière dont cela a été fait, et si les intéressés ont répondu. Ce sous-processus englobe aussi la gestion des fournisseurs prenant part à la collecte en cours, pour faire en sorte que la relation entre l'organisme de statistique et les fournisseurs de données reste positive, ainsi que l'enregistrement des observations, demandes de renseignements et plaintes, et des réponses qui leur sont apportées. Pour les sources administratives et autres sources non statistiques, ce processus est rapide: soit le fournisseur est contacté pour qu'il envoie l'information, soit il l'envoie comme prévu. Une fois que la collecte a atteint ses objectifs, elle se termine et fait l'objet d'un rapport. Il est possible de procéder à une première validation de la structure et de l'intégrité de l'information reçue dans le cadre de ce sous-processus, par exemple en vérifiant que les fichiers sont dans le bon format et contiennent les champs prévus. Tout le travail de validation des contenus s'effectue lors de la phase de traitement.

4. Finalisation de la collecte

71. Ce sous-processus englobe le téléchargement des données et métadonnées recueillies dans un environnement électronique adapté pour la suite du traitement. Il peut comprendre la saisie manuelle ou automatique des données, ce qui suppose par exemple la collaboration d'employés de bureau ou le recours à des outils de reconnaissance des caractères optiques pour extraire l'information des questionnaires sur support papier, ou la conversion des formats des fichiers reçus d'autres organismes. Il peut aussi prévoir l'analyse des métadonnées de processus (paradonnées) associées à la collecte si l'on veut s'assurer que les activités de collecte sont conformes aux prescriptions. Si on utilise un instrument de collecte physique, comme un questionnaire sur support papier, qui ne sert plus ensuite lors du traitement, c'est au cours de ce sous-processus que le document en question est archivé.

E. Phase de traitement

Traitement							
5.1 Intégration des données	5.2 Classification et codage	5.3 Examen et validation	5.4 Édition et imputation des données	5.5 Calcul de nouvelles variables et unités	5.6 Calcul des coefficients de pondération	5.7 Calcul des agrégats	5.8 Finalisation des fichiers de données

72. Cette phase décrit le nettoyage des données et leur préparation aux fins d'analyse. Elle regroupe les sous-processus permettant de vérifier, de nettoyer et de transformer les données d'entrée afin de pouvoir les analyser et les diffuser comme produits statistiques. Elle peut être répétée plusieurs fois si nécessaire. Pour les produits statistiques élaborés

régulièrement, cette phase intervient à chaque itération. Les sous-processus de cette phase peuvent s'appliquer aux données provenant de sources statistiques comme de sources non statistiques (exception faite éventuellement du sous-processus 5.6 (Calcul des coefficients de pondération) qui est généralement spécifique aux données d'enquête).

73. Les phases «traitement» et «analyse» peuvent être itératives et parallèles. L'analyse peut révéler une acception plus large des données et rendre nécessaire un traitement supplémentaire. Les activités qui relèvent des phases de «traitement» et d'«analyse» peuvent commencer avant la fin de la phase de «collecte». Cela autorise la compilation de résultats provisoires lorsque les utilisateurs veulent une réponse rapide, et accroît le temps disponible pour l'analyse.

74. Cette phase se divise en huit sous-processus qui peuvent être séquentiels, de gauche à droite, mais peuvent aussi se dérouler en parallèle et être itératifs. Ces sous-processus sont les suivants:

1. Intégration des données

75. Ce sous-processus permet d'intégrer les données provenant d'une ou plusieurs sources. C'est dans ce cadre que sont regroupés les résultats des sous-processus de la phase de «traitement». Les données d'entrée peuvent provenir de diverses sources externes ou internes, et faire intervenir divers modes de collecte, y compris des extraits de fichiers administratifs. Il en résulte un ensemble de données corrélatives. L'intégration des données consiste notamment à:

- Combiner des données émanant de sources multiples pour créer des statistiques intégrées comme les comptes nationaux;
- Effectuer des routines de mise en correspondance/couplage de fichiers afin de corrélérer des micro et des macrodonnées issues de diverses sources;
- Établir un ordre de priorité lorsque deux ou plusieurs sources contiennent des données pour la même variable, assorties de valeurs potentiellement différentes.

76. L'opération d'intégration des données peut s'effectuer à tout moment pendant cette phase, avant ou après n'importe lequel des autres sous-processus. Il est aussi possible de renouveler l'opération à plusieurs reprises au cours d'un cycle statistique. Une fois cette opération terminée, suivant le degré de protection des données exigé, les données peuvent être rendues anonymes, c'est-à-dire débarrassées de leurs marqueurs comme le nom et l'adresse, pour que la confidentialité soit davantage protégée.

2. Classification et codage

77. Ce sous-processus comprend la classification et le codage des données d'entrée. Par exemple, des routines de codage automatiques (ou manuelles) peuvent assigner des codes numériques aux réponses données de façon textuelle suivant un système de classification prédéterminé.

3. Examen et validation

78. C'est dans le cadre de ce sous-processus que les données sont examinées dans le but d'identifier les problèmes, erreurs et disparités potentiels comme les valeurs aberrantes, les cas de non-réponse et les erreurs de codage. On parle aussi d'étape de validation des données d'entrée. Ce sous-processus peut se dérouler de façon itérative, la validation s'opérant d'après des règles d'édition prédéterminées, habituellement dans l'ordre qui a été fixé. Il peut désigner des données en vue d'une inspection ou d'une édition automatiques ou manuelles. L'examen et la validation peuvent porter sur des données provenant de sources de toutes catégories, avant et après l'opération d'intégration. Même si la validation relève

en principe de la phase de «traitement», dans la pratique, certains éléments de validation peuvent se tenir en même temps que les activités de collecte, en particulier pour des modes comme la collecte par Internet. Si ce sous-processus permet de détecter les erreurs effectives ou potentielles, toutes les activités de correction qui modifient effectivement les données se rattachent au sous-processus 5.4.

4. Édition et imputation

79. Dans les cas où les données sont jugées incorrectes, manquent ou ne sont pas fiables, de nouvelles valeurs peuvent être introduites dans le cadre de ce sous-processus. Les termes édition et imputation recouvrent diverses méthodes permettant de réaliser ces opérations, l'approche suivie étant souvent fondée sur des règles. Les étapes spécifiques de ce processus sont les suivantes:

- Déterminer s'il convient d'ajouter ou de modifier des données;
- Choisir la méthode à utiliser;
- Ajouter/modifier des valeurs de données;
- Inscrire les nouvelles valeurs dans l'ensemble des données et marquer ces dernières comme étant modifiées;
- Produire des métadonnées sur le processus d'édition et d'imputation.

5. Calcul de nouvelles variables et unités

80. Ce sous-processus permet de tirer des données de variables et d'unités qui ne proviennent pas expressément de la collecte, mais dont on a besoin pour livrer les produits voulus. Il permet de calculer de nouvelles variables en appliquant des formules arithmétiques à une ou plusieurs des variables qui figurent déjà dans l'ensemble de données, ou en appliquant différentes hypothèses types. Il pourra être nécessaire de répéter l'opération car certaines variables calculées peuvent être elles-mêmes fondées sur d'autres. Il importe donc de faire en sorte que les variables soient calculées dans le bon ordre. De nouvelles unités peuvent être calculées par agrégation ou dissociation de données pour les unités de collecte, ou par diverses autres méthodes d'estimation. On peut citer en exemple les calculs concernant les ménages, pour lesquels les unités de collecte sont des personnes ou ceux concernant les entreprises pour lesquels les unités de collecte sont des services juridiques.

6. Calcul des coefficients de pondération

81. Ce sous-processus comprend le calcul des coefficients de pondération pour les enregistrements de données unitaires suivant la méthode établie au titre du sous-processus 2.5 (Conception du traitement et de l'analyse). Dans le cas des enquêtes par sondage, les coefficients peuvent être utilisés pour «arrondir» les résultats et les rendre représentatifs de la population cible, ou pour procéder aux ajustements exigés par les non-réponses dans les énumérations totales. Dans d'autres situations, il peut être nécessaire de pondérer les variables à des fins de normalisation.

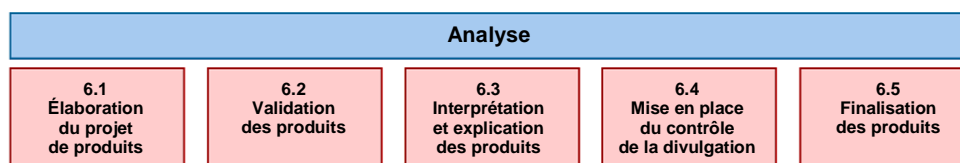
7. Calcul des agrégats

82. Ce sous-processus permet d'obtenir des données agrégées et des totaux de population à partir de microdonnées ou de plus petits agrégats. Il inclut la mise en commun des données dans les cas où les enregistrements partagent certaines caractéristiques, la détermination des mesures de moyenne et de dispersion, et l'application des coefficients de pondération issus du sous-processus 5.6 pour le calcul des totaux appropriés. Dans le cas des enquêtes par sondage, les erreurs d'échantillonnage peuvent aussi être calculées dans le cadre de ce sous-processus, et associées aux agrégats pertinents.

8. Finalisation des fichiers de données

83. Ce sous-processus regroupe les résultats des autres sous-processus de la même phase pour produire un fichier de données (le plus souvent de macrodonnées) qui sert d'intrant pour la phase d'«analyse». Parfois, cela peut être un fichier intermédiaire plutôt qu'un fichier définitif, en particulier pour les processus d'activité où les délais sont très serrés et où il est demandé de produire à la fois des estimations préliminaires et des estimations définitives.

F. Phase d'analyse



84. Cette phase est celle où les produits statistiques sont obtenus, examinés dans le détail et préparés en vue de leur diffusion. C'est à ce stade que le contenu statistique est établi (y compris les observations, les notes techniques, etc.) et que l'on vérifie que les produits sont «conformes aux objectifs» avant leur diffusion auprès du public. Cette phase comprend aussi les sous-processus et activités qui permettent aux analystes statistiques de comprendre les statistiques produites. Pour les produits statistiques élaborés à des intervalles réguliers, cette phase est reproduite à chaque itération. La phase d'«analyse» et les sous-processus sont génériques pour tous les produits statistiques, quelle que soit la source des données.

85. La phase d'«analyse» se divise en cinq sous-processus qui sont généralement séquentiels, de gauche à droite, mais peuvent aussi se dérouler en parallèle et être itératifs. Ces sous-processus sont les suivants:

1. Élaboration du projet de produits

86. C'est au cours de ce sous-processus que les données sont transformées en produits statistiques. C'est là que l'on procède aux mesures additionnelles telles qu'indices, tendances ou séries corrigées des variations saisonnières et que l'on enregistre les caractéristiques de qualité.

2. Validation des produits

87. C'est au cours de ce sous-processus que les statisticiens valident la qualité des produits obtenus, conformément à un cadre général d'évaluation de la qualité et aux attentes exprimées. Ce sous-processus comprend aussi les activités en rapport avec la constitution d'un corpus d'informations par accumulation de connaissances sur un domaine statistique spécifique. Ces connaissances sont alors appliquées à la collecte en cours, dans l'environnement qui est le sien, pour déceler tout écart par rapport à ce qui était prévu et pouvoir analyser les produits en connaissance de cause. Les activités de validation ci-après peuvent être effectuées:

- Vérifier que la population visée et les taux de réponse sont conformes aux prescriptions;
- Comparer les statistiques aux cycles précédents (le cas échéant);
- Vérifier la présence des métadonnées et des paradonnées (métadonnées de processus) associées et leur conformité avec les prévisions;
- Confronter les statistiques à d'autres données pertinentes (internes et externes);

- Rechercher les incohérences qui pourraient demeurer dans les statistiques;
- Procéder à la macroédition;
- Valider les statistiques en fonction des attentes et des connaissances que l'on a du domaine.

3. Interprétation et explication des produits

88. C'est au cours de ce sous-processus que les statisticiens approfondissent leur connaissance des produits. Ils s'appuient sur cette connaissance pour interpréter et expliquer les statistiques produites pour le cycle en évaluant la mesure dans laquelle ces statistiques répondent aux attentes initiales, en examinant les statistiques sous tous les angles au moyen de divers outils et médias et en procédant à des analyses statistiques approfondies.

4. Mise en place du contrôle de la divulgation

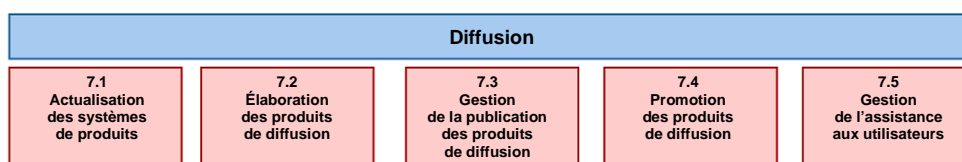
89. C'est au titre de ce sous-processus que l'on veille à ce que les données (et métadonnées) à diffuser n'enfreignent pas les règles en vigueur en matière de confidentialité. On peut à cette fin effectuer des contrôles de la divulgation primaire et secondaire et appliquer des techniques de suppression ou de perturbation des données. Le degré et la méthode de contrôle de la divulgation peuvent varier pour différents types de produits, par exemple on n'utilisera pas la même approche pour les ensembles de microdonnées destinés à la recherche et pour les tableaux ou cartes destinés à être publiés.

5. Finalisation des produits

90. Ce sous-processus permet de faire en sorte que les statistiques et les informations connexes répondent à leurs objectifs et atteignent le niveau de qualité requis, et soient donc immédiatement utilisables. Il comprend les activités suivantes:

- Procéder à des contrôles de cohérence;
- Déterminer le niveau de diffusion, et publier des mises en garde;
- Collationner des informations complémentaires telles qu'interprétations, commentaires, notes techniques, notes d'orientation, mesures des incertitudes et toutes autres métadonnées nécessaires;
- Produire les documents internes complémentaires;
- Avant la diffusion, procéder à des échanges avec les experts internes compétents pour la question à l'examen;
- Approuver le contenu statistique destiné à être publié.

G. Phase de diffusion



91. Cette phase est celle de la remise des produits statistiques aux clients. Elle comprend toutes les activités associées au regroupement et à la diffusion d'une gamme de produits statiques et dynamiques par diverses voies. Ces activités aident les clients à accéder aux produits publiés par l'organisme de statistique et à les utiliser.

92. Pour les produits statistiques élaborés à intervalles réguliers, cette phase intervient à chaque itération. Elle se divise en cinq sous-processus qui sont généralement séquentiels, de gauche à droite, mais peuvent aussi se dérouler en parallèle et être itératifs. Ces sous-processus sont les suivants:

1. Actualisation des systèmes de produits

93. Ce sous-processus gère l'actualisation des systèmes par lesquels les données et métadonnées sont stockées, prêtes à être diffusées, et ce, au moyen des activités suivantes:

- Formater les données et les métadonnées prêtes à être saisies dans les bases de données produites;
- Télécharger les données et métadonnées dans les bases de données produites;
- Faire en sorte que les données soient corrélées aux métadonnées pertinentes.

94. Le formatage, le téléchargement et la corrélation des données devraient s'opérer plus en amont de préférence, mais c'est à ce stade que l'on procède au contrôle final tendant à vérifier que toutes les métadonnées nécessaires sont en place prêtes à être diffusées.

2. Élaboration des produits de diffusion

95. Ce sous-processus est celui de l'élaboration des produits conçus au préalable (dans le sous-processus 2.1) pour répondre aux besoins des utilisateurs. Il peut s'agir d'imprimés, de communiqués de presse et de sites Web. Les produits peuvent prendre de nombreuses formes y compris celle de graphiques interactifs, de tableaux, d'ensembles de microdonnées d'usage public et de fichiers téléchargeables. Le processus comprend ordinairement les étapes suivantes:

- Établir les composantes du produit (texte explicatif, tableaux, graphiques, déclaration de qualité, etc.);
- Assembler les composantes sous forme de produits;
- Éditer les produits et vérifier qu'ils respectent les normes de publication.

3. Gestion de la publication des produits de diffusion

96. Ce sous-processus permet de s'assurer que tous les éléments sont en place pour la diffusion et notamment de gérer le calendrier de publication. Il comprend les notes d'orientation à l'intention de groupes spécifiques comme les journalistes ou les responsables ministériels, ainsi que les dispositions relatives aux embargos avant publication. Il comprend aussi la fourniture des produits aux abonnés et la gestion de l'accès à des données confidentielles par les groupes d'utilisateurs agréés comme les chercheurs. Parfois, un organisme peut se voir contraint de retirer un produit, par exemple si une erreur a été décelée. C'est aussi à ce stade qu'il le fera.

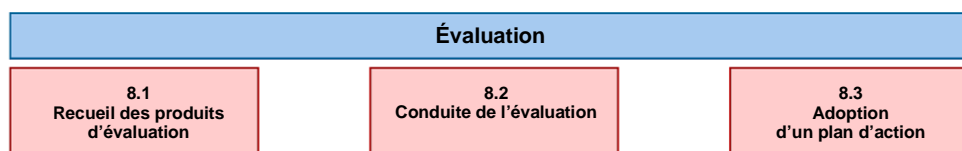
4. Promotion des produits de diffusion

97. Bien que de façon générale la commercialisation puisse se classer parmi les processus fondamentaux, ce sous-processus concerne la promotion active des produits statistiques obtenus dans le cadre d'un cycle statistique donné dans le but de leur faire atteindre le public le plus vaste possible. Il prévoit l'utilisation d'outils de gestion des relations avec la clientèle, pour mieux cibler les utilisateurs potentiels des produits, ainsi que le recours à des outils tels que les sites Web, les wikis et les blogs pour faciliter le processus de communication d'informations statistiques aux utilisateurs.

5. Gestion de l'assistance aux utilisateurs

98. Ce sous-processus veille à l'enregistrement des demandes de renseignements et de services des clients, comme l'accès aux microdonnées, et à la fourniture de réponses dans les délais convenus. Ces demandes devraient être examinées à intervalles réguliers pour contribuer au processus fondamental de gestion de la qualité car elles peuvent signaler de nouveaux besoins chez les utilisateurs, ou des changements dans les besoins exprimés.

H. Phase d'évaluation



99. Cette phase fait référence à l'évaluation d'un cas particulier de processus de production statistique, à ne pas confondre avec le processus fondamental de gestion de la qualité statistique décrit à la section VI qui aborde la question de façon plus générale. Elle se tient logiquement au dernier stade du déroulement du processus mais repose sur des intrants réunis tout au long des diverses phases. Elle peut consister à évaluer le succès du cycle statistique particulier en s'appuyant sur divers intrants quantitatifs et qualitatifs et en identifiant et en hiérarchisant les améliorations possibles qui pourraient lui être apportées.

100. Pour les produits statistiques élaborés à intervalles réguliers, on devrait au moins en théorie faire une évaluation à chaque itération, pour déterminer s'il convient de répéter le processus et, dans l'affirmative, s'il y a lieu de lui apporter des améliorations. Toutefois, dans certains cas, en particulier pour les cycles statistiques réguliers et bien établis, il peut ne pas être utile de procéder à une évaluation officielle à chaque fois. Dans ces cas, cette phase pourra éventuellement servir à déterminer si l'itération suivante doit démarrer dès la phase de l'identification des besoins, ou à partir d'une phase ultérieure (qui est souvent la phase de collecte).

101. Cette phase se divise en trois sous-processus qui sont généralement séquentiels, de gauche à droite, mais peuvent aussi se chevaucher jusqu'à un certain point dans la pratique. Ces sous-processus sont les suivants:

1. Recueil des produits d'évaluation

102. Des supports d'évaluation peuvent être produits au cours de n'importe quel autre phase ou sous-processus. Ils peuvent prendre diverses formes, réactions des utilisateurs, métadonnées de processus (paradonnées), paramètres de mesure des systèmes ou suggestions du personnel par exemple. Le compte rendu des progrès réalisés au regard d'un plan d'action convenu lors d'une précédente itération peut aussi constituer un intrant pour les évaluations des itérations suivantes. Ce sous-processus regroupe ces différents apports et les met à la disposition de la personne ou de l'équipe chargée de l'évaluation.

2. Conduite de l'évaluation

103. Ce sous-processus permet d'analyser les intrants de l'exercice d'évaluation et d'en faire la synthèse dans un rapport d'évaluation. Ce rapport devrait consigner toutes les questions relatives à la qualité qui concernent particulièrement une itération du processus de production statistique, et devrait faire des recommandations de modifications le cas échéant. Ces recommandations peuvent concerner des modifications à apporter à une phase ou un sous-processus quelconque lors des itérations futures du processus, ou aller à l'encontre de la répétition du processus.

3. Adoption d'un plan d'action

104. Ce sous-processus réunit tous les décideurs voulus en vue de l'établissement et de l'adoption d'un plan d'action fondé sur le rapport d'évaluation. Il devrait aussi comporter l'examen d'un mécanisme de suivi de l'impact de ces mesures, lequel pourra, à son tour, contribuer aux évaluations des itérations futures du processus.

VI. Processus fondamentaux

105. Le GSBPM reconnaît aussi plusieurs processus fondamentaux qui s'appliquent tout au long des phases de production et au travers des cycles statistiques. Certains de ces processus fondamentaux sont répertoriés aux paragraphes 13 et 14. Les processus de gestion de la qualité et de la gestion de métadonnées sont détaillés dans la présente section.

A. Gestion de la qualité

106. La qualité intéresse les organismes, les processus et les produits. Dans le cas présent, le processus fondamental de gestion de la qualité se rapporte principalement à la qualité des produits et des processus.

107. Le principal objectif de la gestion de la qualité dans le cycle statistique est de mieux appréhender et gérer la qualité des produits statistiques. Les organismes de statistique sont généralement d'accord pour définir la qualité selon la norme ISO 9000-2005: «aptitude d'un ensemble de caractéristiques intrinsèques à satisfaire des exigences»¹⁰. Ainsi, la qualité d'un produit est un concept complexe et pluridimensionnel qui se définit habituellement en tenant compte de différents aspects de la qualité. Les aspects de la qualité jugés les plus importants dépendent des perspectives, des besoins et des priorités de l'utilisateur, lesquels varient d'un processus et d'un groupe d'utilisateurs à l'autre.

108. Pour améliorer la qualité des produits, la gestion de la qualité devrait intervenir à tous les stades du modèle du processus de production statistique. Ce processus est en rapport direct avec la phase 8 (évaluation) qui a plus particulièrement pour fonction de postévaluer des cas individuels de cycle statistique. La gestion de la qualité a toutefois un champ plus vaste et plus profond. Outre l'évaluation des itérations d'un processus, elle sert aussi à évaluer des phases et des sous-processus pris individuellement, dans le meilleur des cas chaque fois qu'ils s'appliquent, mais au moins suivant le calendrier établi. Les métadonnées inhérentes aux divers processus présentent aussi de l'intérêt comme intrants pour la gestion de la qualité du processus. Ces évaluations peuvent s'insérer dans un processus donné, ou s'étendre à plusieurs processus qui utilisent des composantes communes.

109. Il faut noter en outre la place fondamentale que tiennent dans la gestion de la qualité les diverses mesures de contrôle de qualité qui doivent normalement être prévues au sein des sous-processus pour prévenir les erreurs et en assurer le suivi. Cette stratégie pourrait s'illustrer dans un plan d'assurance-qualité.

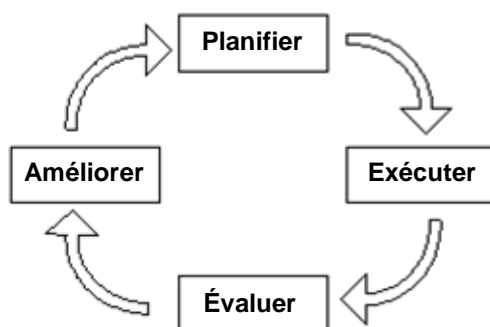
110. Dans un organisme, la gestion de la qualité porte en général sur un cadre d'évaluation de la qualité bien particulier, et peut donc prendre diverses formes et donner différents résultats d'un organisme à l'autre. La multiplicité actuelle des cadres d'évaluation de la qualité donne encore plus d'importance à l'évaluation comparative et à l'évaluation par les pairs et même si ces deux approches ont peu de chances d'être

¹⁰ ISO 9000:2005, Systèmes de management de la qualité – Principes essentiels et vocabulaire. Organisation internationale de normalisation.

réalisables pour chaque itération ou chaque volet du cycle statistique, elles devraient être utilisées d'une manière systématique suivant un calendrier prédéterminé qui autorise l'examen de toutes les grandes composantes du processus dans un certain laps de temps¹¹.

111. Pour étendre la portée du processus fondamental de gestion de la qualité, on peut aussi envisager d'évaluer des groupes de processus de production statistique pour voir s'ils présentent des chevauchements d'activité ou des lacunes.

112. Toutes les évaluations suscitent des réactions qui devraient à leur tour être exploitées pour améliorer le processus, la phase ou le sous-processus pertinent en formant une boucle de la qualité.



113. On peut citer les activités de gestion de la qualité ci-après en exemple:

- Établir et maintenir le cadre d'évaluation de la qualité;
- Établir les critères de qualité généraux;
- Établir les objectifs de qualité du processus et en contrôler la réalisation;
- Demander aux utilisateurs leurs appréciations et les analyser;
- Examiner le fonctionnement du processus et rendre compte des enseignements tirés;
- Examiner les métadonnées du processus et les indicateurs de qualité;
- Procéder à un audit interne ou externe des processus statistiques.

114. La gestion de la qualité fait aussi intervenir des facteurs institutionnels et organisationnels. Ces facteurs relèvent d'autres processus fondamentaux du GSBPM (par exemple, la gestion des ressources humaines, la gestion des programmes statistiques) bien qu'ils puissent avoir un impact sur la qualité.

B. Gestion des métadonnées

115. Une gestion avisée des métadonnées est indispensable à la bonne marche des processus d'activité statistique. Des métadonnées sont associées à chaque phase, soit qu'elles soient produites pendant cette phase, soit reportées d'une phase précédente. Dans le cadre de ce modèle, le processus fondamental de la gestion des métadonnées met l'accent sur la création, l'utilisation et l'archivage de métadonnées statistiques, bien que les métadonnées associées aux différents sous-processus soient aussi intéressantes, y compris comme intrant pour la gestion de la qualité. Ce qui compte, surtout, c'est de veiller à ce que les métadonnées

¹¹ Parmi les cadres généraux qui pourraient convenir, on citera le cadre national d'assurance de la qualité élaboré par un groupe d'experts mondial relevant de la Commission de statistique des Nations Unies. Voir l'adresse suivante: <http://unstats.un.org/unsd/dnss/QualityNQAF/nqaf.aspx>.

soient saisies aussitôt que possible et stockées et transférées de phase en phase ainsi que les données qu'elles décrivent. La stratégie et les systèmes de gestion des métadonnées sont donc indispensables au fonctionnement de ce modèle et peuvent être facilités par le GSIM.

116. Le GSIM constitue un cadre de référence d'objets d'information qui permet de faire une description générale de la définition, de la gestion et de l'utilisation des données et des métadonnées tout au long du processus de production de statistiques. Il appuie une approche cohérente des métadonnées, en facilitant le rôle premier que prévoit pour ces dernières la partie A du Cadre commun des métadonnées¹², «Métadonnées statistiques dans un contexte professionnel», à savoir que les métadonnées devraient seulement et formellement définir les contenus et les liens entre les objets et les processus dans le système d'information statistique.

117. La partie A du Cadre commun des métadonnées identifie aussi les 16 principes fondamentaux ci-après de la gestion des métadonnées, qui devraient tous être suivis dans le processus global de gestion des métadonnées et pris en compte au moment de concevoir et de mettre en œuvre un système de métadonnées statistiques. Ces principes se classent dans quatre groupes:

1. Traitement des métadonnées

a) *Modèle de processus de production statistique*: Gérer les métadonnées dans l'optique du modèle générique de processus de production statistique;

b) *Métadonnées actives et non passives*: Faire en sorte que les métadonnées soient aussi actives que possible. Les métadonnées actives sont celles qui commandent d'autres processus. En traitant les métadonnées de cette façon, on fera en sorte qu'elles soient précises et adaptées à la situation du moment;

c) *Recyclage*: Recycler, autant que possible, les métadonnées aux fins de l'harmonisation des statistiques ainsi que par souci d'efficacité;

d) *Versions*: Sauvegarder l'historique (c'est-à-dire les versions anciennes) des métadonnées.

2. Autorité chargée des métadonnées

a) *Enregistrement*: Veiller à ce que le processus d'enregistrement (déroulement du travail) associé à chaque élément de métadonnées soit étayé par les documents voulus, afin de pouvoir déterminer clairement les droits de propriété, le statut d'approbation, la date de mise en circulation, etc.;

b) *Source unique*: Veiller à ce qu'il n'existe qu'une seule source faisant autorité («autorité d'enregistrement») pour chaque élément de métadonnées;

c) *Saisie/actualisation unique*: Réduire au minimum les risques d'erreur en n'entrant les métadonnées qu'une seule fois et en ne les actualisant qu'en un seul endroit;

d) *Écarts par rapport aux normes*: Faire en sorte que les écarts par rapport aux normes soient surveillés de près, approuvés, documentés et mis en évidence.

3. Relation au cycle/processus statistique

a) *Intégrité*: Veiller à ce que les travaux relatifs aux métadonnées fassent partie intégrante des processus d'activité dans l'ensemble de l'organisme;

b) *Métadonnées correspondantes*: Faire en sorte que soient communiquées aux utilisateurs finals des métadonnées qui correspondent à celles qui ont guidé le processus d'activité ou ont été élaborées au cours de ce processus;

¹² Voir l'adresse: <http://www.unece.org/stats/cmf/PartA.html>.

c) *Définition du flux*: Définir le flux des métadonnées en l'associant au processus de production de statistiques et au processus d'activité (en même temps que le flux de données et la logique métier);

d) *Saisie à la source*: Saisir les métadonnées à leur source, si possible de manière automatique, en tant que sous-produit d'autres processus;

e) *Échange et utilisation*: Échanger des métadonnées et les utiliser comme base d'information tant pour le traitement et l'analyse informatiques que pour l'interprétation humaine. Il faudrait baser l'infrastructure d'échange de données et de métadonnées connexes sur des composantes associées de manière souple, permettant d'utiliser le langage d'échange standard de son choix, comme le langage XML.

VII. Autres utilisations du Modèle générique du processus d'activité statistique

118. Initialement, le GSBPM avait pour but de fournir aux organismes de statistique une base leur permettant d'arrêter une terminologie normalisée propre à faciliter leurs débats sur l'élaboration de systèmes et de processus de métadonnées statistiques. Toutefois, au fur et à mesure de l'élaboration du modèle, il est devenu de plus en plus manifeste qu'il pouvait être utilisé à de nombreuses autres fins, en particulier pour la modernisation des statistiques officielles. Plusieurs documents décrivant les utilisations effectives et potentielles du GSBPM sont disponibles sur la plate-forme wikipedia de la CEE¹³. On trouvera ci-après une liste de tâches pour lesquelles le GSBPM est actuellement employé dont le but est de susciter de nouvelles idées d'applications pratiques du modèle:

- Harmoniser les architectures de production statistique – le GSBPM peut être considéré comme un modèle de configuration opérationnelle d'architecture de calcul statistique. Il identifie les principales composantes du cycle statistique, promeut une terminologie normalisée et des modalités normalisées de mise en œuvre des processus de production statistique. C'est un élément déterminant pour l'architecture commune de production statistique¹⁴;
- Faciliter le partage de logiciels statistiques – en lien avec ce qui précède, le GSBPM définit les composantes des processus statistiques de façon à non seulement encourager le partage d'outils logiciels entre processus de production statistique, mais à faciliter aussi le partage entre différents organismes de statistique qui appliquent le modèle. Il a été utilisé pour «classer» les logiciels qu'il est possible de partager dans l'inventaire établi par le Conseil consultatif pour le partage des systèmes d'information statistique¹⁵;
- Décrire lesquelles des normes sont ou pourraient être utilisées pour différentes phases du processus de production statistique. Par exemple, l'annexe 2 du Manuel de l'utilisateur du SMDX 2.1¹⁶ étudie les modalités d'application du SMDX aux travaux statistiques dans le cadre d'un modèle de processus d'activité;
- Fournir un cadre à l'évaluation et à l'amélioration de la qualité des processus – pour qu'une approche comparative de l'évaluation de la qualité des processus réussisse, il faut normaliser ces derniers dans toute la mesure possible. Le GSBPM offre un mécanisme le permettant;

¹³ Voir l'adresse: <http://www1.unece.org/stat/platform/display/metis/Papers+about+the+GSBPM>.

¹⁴ Voir l'adresse: <http://www1.unece.org/stat/platform/display/CSPA/Common+Statistical+Production+Architecture+Home>.

¹⁵ Voir l'adresse: <http://www1.unece.org/stat/platform/display/msis/Software+Inventory>.

¹⁶ Voir l'adresse: http://sdmx.org/index.php?page_id=38.

- Mieux associer les travaux sur les métadonnées et la qualité statistiques – en lien avec ce qui précède, le cadre commun fourni par le GSBPM peut aider à associer les travaux internationaux sur les métadonnées statistiques avec ceux afférents à la qualité des données en fournissant un cadre commun et une terminologie commune pour décrire le processus d'activité statistique;
- Fournir un modèle fondamental pour la normalisation méthodologique – des normes méthodologiques peuvent être associées aux phases ou sous-processus auxquels elles se rapportent et être ensuite classifiées et stockées dans une structure fondée sur le GSBPM;
- Fournir une structure pour la documentation des processus statistiques – le GSBPM peut permettre de structurer l'organisation et le stockage de la documentation dans un organisme en favorisant la normalisation et l'identification des bonnes pratiques;
- Fournir un cadre au renforcement des capacités des organismes – le GSBPM peut être utilisé comme cadre d'évaluation des connaissances et des capacités qui existent déjà au sein d'un organisme et peut permettre d'identifier les lacunes à combler pour améliorer son efficacité opérationnelle;
- Contribuer à la planification de haut niveau des travaux des institutions;
- Mettre au point un référentiel de modèles de processus d'activité – Statistics New Zealand a mis au point une base de données pour stocker des modèles de processus et pouvoir les mettre en correspondance avec son modèle de processus d'activité statistique. Cet organisme prévoit aussi de constituer un réseau d'échanges sur la modélisation de processus d'activité, à savoir un forum régulier qui aura pour mission d'approfondir les connaissances sur la modélisation des processus, de promouvoir son modèle et de mieux le faire comprendre et de débattre sur la modélisation des processus et le rôle des modèles comme facteurs d'amélioration des processus;
- Mesurer les coûts opérationnels – le GSBPM peut être utilisé comme repère pour mesurer les coûts des différents volets du processus de production statistique. Cela permet de cibler les activités de modernisation qui peuvent améliorer l'efficacité des volets du processus les plus coûteux;
- Mesurer la performance du système – en lien avec ce qui précède concernant les coûts, le GSBPM peut aussi être utilisé pour identifier les composantes qui ne fonctionnent pas efficacement, qui se chevauchent inutilement, ou qui doivent être remplacées. Il peut aussi permettre d'identifier les lacunes à combler qui exigent l'élaboration de nouvelles composantes.

Annexe 1

Liste des sigles et abréviations

Note: La liste ci-après contient les principales abréviations utilisées dans le présent document. Pour disposer d'un glossaire plus complet des termes en rapport avec le processus de production statistique, voir la documentation relative au GSIM.

CEE	Commission économique pour l'Europe
CMF	Cadre commun de métadonnées: ensemble de ressources liées à l'utilisation des métadonnées par les organismes de statistique, y compris des informations sur les normes et les meilleures pratiques. Voir l'adresse: http://www.unece.org/stats/cm/
CSPA	Architecture commune de la production statistique: une architecture des professionnels du secteur qui regroupe le GSBPM et le GSIM, outre les nouveaux cadres sur les services statistiques, pour établir un descriptif agréé de haut niveau du «système» de production statistique conformément à l'initiative de modernisation
DDI	Data Documentation Initiative: norme internationale pour la description des données issues des sciences sociales, comportementales et économiques
GLBPM	Modèle de processus d'activité longitudinal générique: modèle fondé sur le GSBPM mis au point par les chercheurs qui réalisent des enquêtes sociales
GSBPM	Modèle générique du processus de production statistique: outil flexible permettant de décrire et de définir l'ensemble des processus d'activité nécessaires à la production de statistiques officielles.
GSIM	Modèle générique d'informations statistiques: cadre de référence des objets d'information, qui permet de faire une description générale de la définition, de la gestion et de l'utilisation des données et des métadonnées tout au long du processus de production de statistiques
METIS	Dénomination des travaux menés sur les métadonnées statistiques sous l'égide de la Conférence des statisticiens européens
OCDE	Organisation de coopération et de développement économiques
SDMX	Échange de données et de métadonnées statistiques: ensemble de normes techniques et de directives concernant le contenu, utilisées avec les réseaux et outils informatiques, qui facilitent l'échange et le partage de données et de métadonnées statistiques
XML	eXtensible Mark-up Language (Langage de balisage extensible): langage qui définit une série de règles pour le codage de documents sous un format à la fois lisible par l'homme et par l'ordinateur