

Cadre statistico-géospatial mondial



Principes

Accessible et Utilisable

Interopérabilité
statistique et géospatiale

Géographies communes
pour la diffusion des statistiques

Données d'enregistrement d'unités géocodées
dans un environnement de gestion de données

Utilisation de l'infrastructure géospatiale
fondamentale et du géocodage

Entrée
(Contribution)

Géospatial

- Données fondamentales
- Données complémentaires
- Nouvelles sources de données

- Recensements
- Enquêtes
- Registres de données administratives
- Mégadonnées et autres sources

Statistique

Éléments Clés

Normes et bonnes pratiques

Lois et politiques nationales

Infrastructure technique

Collaboration institutionnelle

Intégration

Informations harmonisées et normalisées

Interopérabilité et Comparabilité

Diffusion

Prise de décision

Analyse

Sortie
(Résultat ou Production)

Département des affaires économiques et sociales
Division de la statistique

Secrétariat de la comité d'experts sur la gestion de l'information géospatiale à l'échelle mondiale

Cadre statistico-géospatial mondial



Nations Unies
New York, 2019



Comité d'experts sur la
gestion de l'information
géospatiale à l'échelle
mondiale



Département des affaires économiques et sociales

Le Département des affaires économiques et sociales (DAES) est une interface essentielle entre les différentes politiques mondiales dans les domaines économique, social et environnemental, et l'action nationale. Le DAES œuvre dans trois grands domaines liés : (i) il génère, analyse et compile un large éventail de données économiques, sociales et environnementales, et des informations sur lesquelles les États membres des Nations Unies s'appuient pour examiner les problèmes communs et évaluer les options stratégiques; (ii) il facilite, dans le cadre de plusieurs organismes intergouvernementaux, les négociations entre États membres afin qu'ils adoptent des mesures communes pour répondre aux préoccupations les plus pressantes du monde; et (iii) il fournit aux gouvernements intéressés des conseils sur la façon de traduire les orientations politiques élaborées lors des conférences et sommets des Nations Unies en programmes dans leur pays et, au moyen d'une assistance technique, il aide à renforcer les capacités nationales.

Division de la statistique

La Division de la statistique des Nations Unies (DS-ONU) s'est engagée à faire progresser les systèmes statistiques et géospatiaux mondiaux. Elle compile et diffuse des données statistiques mondiales, élabore des standards et des normes pour les activités statistiques et géospatiales, et soutient les efforts des pays en vue de renforcer leurs systèmes nationaux d'information statistique et géospatiale. La DS-ONU facilite également la coordination des activités statistiques et géospatiales internationales et soutient le fonctionnement de la Commission de statistique de l'ONU, du Comité d'experts de l'initiative des Nations Unies sur la gestion de l'information géospatiale à l'échelle mondiale et du Groupe d'experts des Nations unies sur les noms géographiques en tant qu'organisme de coordination des systèmes géospatiaux et statistiques mondiaux.

Comité d'experts de l'initiative des Nations Unies sur la gestion de l'information géospatiale à l'échelle mondiale

Le Comité d'experts de l'initiative des Nations Unies sur la gestion de l'information géospatiale à l'échelle mondiale (UN-GGIM) est le principal organe intergouvernemental de discussion, d'amélioration et de coordination des activités de gestion de l'information géospatiale à l'échelle mondiale. À cette fin, il engage les États membres au plus haut niveau et coopère avec les gouvernements pour prendre des décisions communes et établir des orientations concernant l'utilisation de l'information géospatiale dans des cadres nationaux et mondiaux, et pour élaborer des stratégies efficaces de renforcement des capacités géospatiales des pays en développement.

Commission de statistique de l'ONU

La Commission de statistique de l'ONU, qui a vu le jour en 1947, est l'entité qui coiffe le système statistique mondial. Elle réunit les directeurs des statistiques des États membres du monde entier. Il s'agit de la plus haute instance décisionnaire pour ce qui concerne les activités statistiques internationales, en particulier l'établissement des normes statistiques, la mise au point de concepts et de méthodes ainsi que leur mise en œuvre aux niveaux national et international. La Commission de statistique supervise le travail de la Division de statistique de l'Organisation des Nations Unies. C'est une commission technique du Conseil économique et social des Nations Unies.

Remarques

Les appellations employées dans le présent document et la présentation du contenu dans le site Internet n'impliquent de la part du Secrétariat de l'Organisation des Nations Unies aucune prise de position quant au statut juridique des pays, territoires, villes ou zones ou de leurs autorités, ni quant au tracé de leurs frontières ou limites. Le terme « pays » employé dans ce Contenu s'entend également des territoires ou zones. Les termes « régions développées » et « régions en développement » sont utilisés par commodité dans la présentation des statistiques et n'impliquent pas nécessairement de jugement quant au stade de développement atteint par un pays ou une zone donnée.

[Page laissée en blanc intentionnellement]

Sommaire

Le Cadre statistico-géospatial mondial (CSGM) facilite l'intégration de l'information statistique et géospatiale. Cadre établi pour le monde entier, le CSGM permet l'intégration d'un large éventail de données provenant de communautés statistiques et géospatiales et, grâce à l'application de ses cinq principes et éléments moteurs essentiels, il permet la production de données statistiques à référence géospatiale (ou géocodées) harmonisées et normalisées. L'intégration des informations obtenues à partir de données statistiques, géospatiales et autres favorise une prise de décision éclairée et fondée sur des informations factuelles à l'appui des priorités et des programmes de développement aux niveaux local, infranational, national, régional et mondial, tant pour les recensements de population du cycle de 2020 que pour le Programme de développement durable à l'horizon 2030.

« Il est urgent de mettre en place un mécanisme, de type cadre statistico-spatial mondial, favorisant l'élaboration de méthodes cohérentes de production et d'intégration de l'information géostatistique. »

Forum mondial sur l'intégration des données statistiques et géospatiales, New York, 2014

Les renseignements présentés ici aideront les pays et les utilisateurs à comprendre le CSGM, sa valeur, son application, son infrastructure et les exigences de sa mise en œuvre. À cette fin, le document donne une vue d'ensemble du CSGM et une explication de ses cinq principes et d'autres éléments essentiels. D'autres renseignements d'appui à la mise en œuvre sont compilés dans le wiki du Groupe d'experts sur l'intégration de l'information statistique et géospatiale. L'élaboration du CSGM est le fruit d'un processus collaboratif qui a mobilisé des organismes d'information statistique et géospatiale du monde entier.

« ...bâtir le cadre statistico-géospatial mondial comme une méthode commune de production de données statistiques et administratives à référence géospatiale de façon à intégrer les données provenant de diverses sources en fonction de la localisation et de les intégrer à d'autres renseignements géospatiaux. »

Comité d'experts de l'initiative des Nations Unies sur la gestion de l'information géospatiale à l'échelle mondiale, New York, 2015

Fondamentalement, le CSGM permet :

- l'intégration de données permettant de mieux mesurer et suivre les cibles et le cadre mondial d'indicateurs pour les objectifs de développement durable du Programme de développement durable à l'horizon 2030 et les recensements de population du cycle de 2020;
- des comparaisons aux niveaux local, infranational, national, régional et mondial pour les processus décisionnels au sein des pays et entre eux ainsi que dans les domaines thématiques;
- le partage de données entre institutions, grâce à l'interopérabilité de l'information géospatiale et statistique et l'élaboration d'outils et d'applications communs;
- la découverte de nouvelles perspectives et relations entre données qui n'aurait pas été possible si les données socioéconomiques, environnementales ou géospatiales avaient été analysées séparément;
- l'augmentation de la quantité d'informations sur les zones de petite superficie;

- une sensibilisation accrue aux méthodes et aux outils d'évaluation et de gestion des risques de divulgation ainsi que d'amélioration de la protection de la confidentialité des renseignements personnels dans la collecte, le stockage et la diffusion de l'information;
- les conditions nécessaires à l'investissement et au renforcement des capacités en matière d'information statistico-géospatiale;
- l'intégration de nouvelles sources de données qui éclaireront la production d'information géospatiale de grande qualité, par exemple l'observation de la Terre et les autres sources de données complémentaires;
- la consolidation de la collaboration institutionnelle entre les communautés géospatiales et statistiques.

Liste des abréviations

ACPD	Architecture commune de la production statistique
ACPS	Architecture commune de production statistique
API	Interface de programmation d'applications
CDR	Cadre de description des ressources
CEE-ONU	Commission économique pour l'Europe des Nations Unies
CSGM	Cadre statistico-géospatial mondial
CSNU	Commission de statistique des Nations Unies
DGGS	Système de grille discrète mondiale
DSNU	Division de la statistique des Nations Unies
FGDC	Norme sur le contenu des données du Cadre d'information géographique
GAMSO	Modèle générique d'activité des organismes statistiques
GE-IDSG	Groupe d'experts sur l'intégration des données statistiques et géospatiales
GEOSS	Système mondial des systèmes d'observation de la Terre
GFM	Modèle général d'entité
GGRF	Cadre de référence géodésique mondial
GHN-MSO	Groupe de haut niveau sur la modernisation de la statistique officielle
IDD	Initiative sur la documentation des données
INDG	Infrastructure nationale de données géospatiales
INS	Institut national de statistique
ISO	Organisation internationale de normalisation
MAUP	Problème d'unité spatiale modifiable
MGIS	Modèle générique de l'information statistique
MSGPO	Modèle statistique général du processus opérationnel
ODD	Objectifs de développement durable
OEN	Organisation d'élaboration de normes
OGC	Consortium géospatial ouvert (Open Geospatial Consortium)
OHI	Organisation hydrographique internationale
ONG	Organisation non gouvernementale
ONIG	Organisme national d'information géospatiale
PE	Protocoles d'entente
SCEE	Système de comptabilité économique et environnementale
SDMX	Échange de données et métadonnées statistiques
SSE	Système statistique européen
SSN	Système statistique national
TJS	Service de mise en correspondance des données tabulaires
UN-GGIM	Comité d'experts sur la gestion globale de l'information géospatiale
W3C	Consortium World Wide Web
WFS	Service d'entités géographiques Web
WMS	Service de cartographie Web

Table des matières

Sommaire	vii
Introduction	1
Partie 1 : Le cadre statistico-géospatial mondial	6
Principe 1 : utilisation de l'infrastructure géospatiale de base et du géocodage.....	9
Principe 2 : un géocodage unitaire dans un environnement de gestion des données.....	10
Principe 3 : le recours à des informations géographiques communes pour la diffusion des statistiques.....	11
Principe 4 : l'interopérabilité des informations statistiques et géospatiales	12
Principe 5 : l'accessibilité et l'exploitabilité des statistiques géospatiales	13
Initiatives complémentaires	14
Appel à l'action et voie à suivre	15
Partie 2 : Description détaillée du cadre statistico-géospatial mondial	16
Principe 1 : utilisation de l'infrastructure géospatiale de base et du géocodage.....	17
Principe 2 : un géocodage unitaire dans un environnement de gestion des données.....	20
Principe 3 : le recours à des informations géographiques communes pour la diffusion des statistiques.....	24
Principe 4 : l'interopérabilité des informations statistiques et géospatiales	28
Principe 5 : l'accessibilité et l'exploitabilité des statistiques géospatiales	33
Partie 3 : Annexes	36
Annexe A : Définitions des termes.....	37
Annexe B : Normes, qualité et cadres facilitateurs.....	40
Annexe C : Autres lectures.....	50

Liste des figures

Figure 1 Le lieu comme lien entre la société, l'économie et l'environnement	1
Figure 2 Le cadre statistico-géospatial mondial – des entrées aux résultats	5
Figure 3 Les cinq principes du CSGM	7
Figure 4 Zones géographiques administratives et maillées.....	25
Figure 5 Cadre d'interopérabilité européen	29
Figure 6 Les 14 thèmes des données géospatiales fondamentales mondiales	46

Introduction

Le Programme de développement durable à l'horizon 2030, avec ses 17 objectifs de développement durable (ODD), ses 169 cibles et son cadre mondial d'indicateurs, exprime le besoin impérieux de comprendre, soutenir et développer les trois piliers du développement d'un pays : l'économie, la société et l'environnement¹. Le fait de permettre une prise de décision fondée sur des données probantes dans les secteurs public et privé maximalisera les résultats positifs dans ces trois domaines. Il est admis que l'intégration d'informations issues à la fois du domaine statistique (qui comprennent un large éventail de données socioéconomiques et d'entreprises) et du domaine géospatial (qui comprennent des données sur l'environnement et l'observation de la Terre, dont l'importance est croissante) est essentielle pour la compréhension des liens entre ces trois piliers et au sein de chacun. L'intégration de ces données est essentielle à l'amélioration de la qualité de l'information factuelle servant à prendre des décisions aujourd'hui et dans le futur.

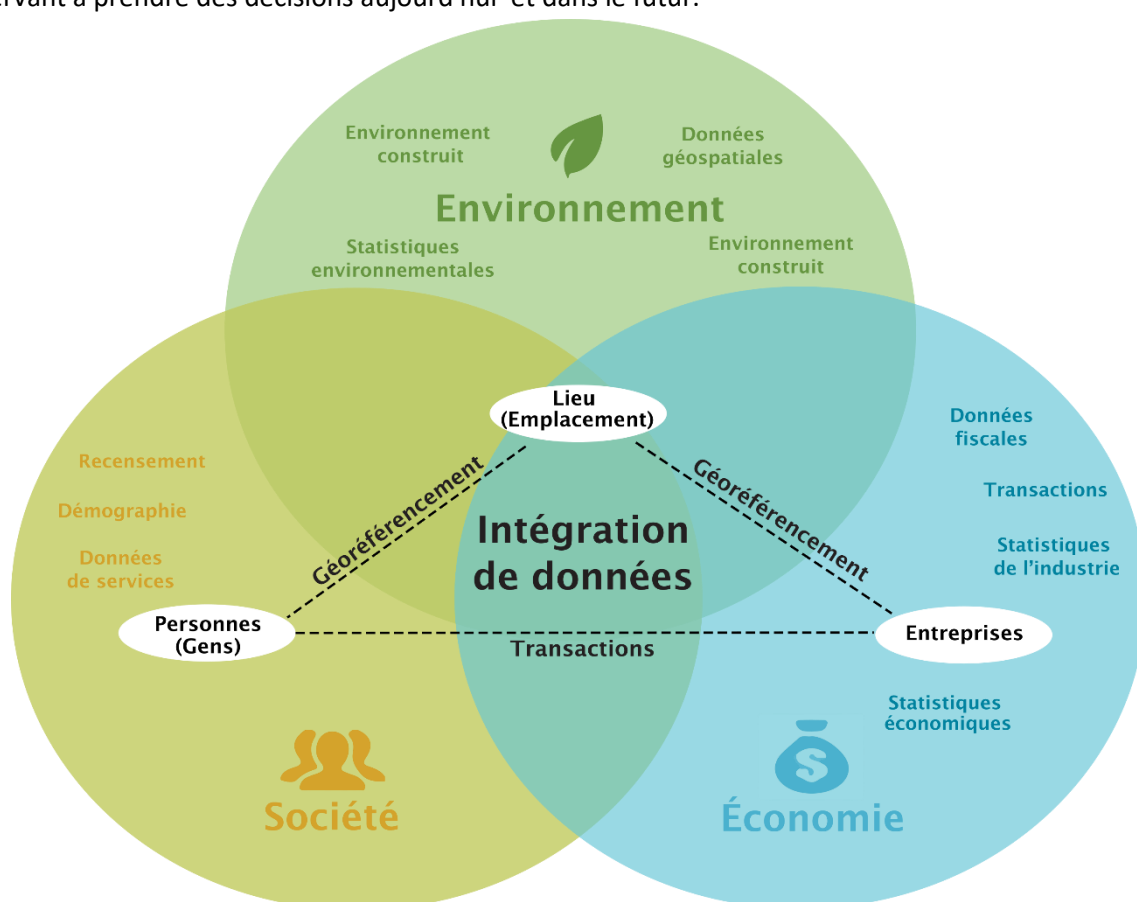


Figure 1 Le lieu comme lien entre la société, l'économie et l'environnement

Plus simplement, **le fait de relier des données statistiques sur les personnes et les entreprises à un lieu ou un emplacement géographique et leur intégration à l'information géospatiale au moyen de la localisation peuvent améliorer notre compréhension des questions sociales, économiques et environnementales, et ce, bien plus que si l'on considère les données statistiques ou géospatiales séparément** (la Figure 1 illustre cette affirmation).

¹ <https://www.un.org/ecosoc/en/sustainable-development>

Le CSGM fournit le mécanisme sous-tendant la réalisation de cette intégration et produit des statistiques géospatiales qui dressent un tableau plus complet de notre monde. Il faut engager des efforts considérables pour tirer pleinement profit des données statistiques et géospatiales intégrées. D'abord et avant tout, en tant que communauté mondiale, nous devons œuvrer ensemble à une sensibilisation accrue aux avantages pour le développement local, infranational, national, régional et global que peut procurer l'intégration de l'information statistique et géospatiale. Lorsque la gestion et l'utilisation de l'information géospatiale deviennent courantes et font partie intégrante de la production de statistiques, l'offre d'informations et de technologies habilitantes se diversifie et augmente. L'action visant à améliorer l'intégration de l'information statistique et géospatiale a lieu dans un contexte où les instituts nationaux de statistique (INS) cherchent également à moderniser leurs systèmes et processus de production statistique, à transformer leurs opérations, à réduire les complexités et à produire de nouveaux indicateurs et mesures pertinents à des fins statistiques. Pour ce faire, il est essentiel d'introduire une infrastructure et des processus interopérables, fondés sur des normes et axés sur les métadonnées.

Les statistiques géospatiales serviront à consolider l'analyse des données de façon à promouvoir la prise de décision éclairée et fondée sur des données probantes. Cette démarche présente plusieurs avantages de taille, notamment la capacité de créer et d'analyser des informations géographiques locales à une échelle plus fine et de tirer parti de nouvelles sources de données. Ces « nouvelles » sources de données peuvent inclure l'utilisation d'observations de la Terre pour produire des données dans des domaines où les méthodes classiques d'information statistique ou géospatiale se révèlent inefficaces, insuffisamment précises temporellement ou inexistantes en raison des capacités nationales actuelles en matière d'information statistique ou géospatiale. La demande faite aux communautés statistiques et géospatiales de fournir de l'information et des renseignements pour les priorités nationales de développement et les 17 ODD renforce la nécessité d'examiner ces autres sources de données.

Pour les organismes nationaux d'information géospatiale, ce travail se fait simultanément à travers des efforts collectifs qui amélioreront la gestion et l'utilisation de l'information géospatiale. Il s'agit notamment des mesures visant à définir précisément les ensembles de données géospatiales fondamentales ou de base nécessaires au soutien des activités géospatiales dans le cadre des efforts nationaux et internationaux : par exemple, pour appuyer la mise en œuvre des ODD. Ces ensembles de données fondamentales font partie de l'infrastructure géospatiale d'un pays. Ils permettent le géocodage des statistiques et comprennent des données géographiques administratives et statistiques qui facilitent l'intégration de l'information statistique avec l'information géospatiale.

Le CSGM se situe dans un écosystème au développement rapide qui renforce les communautés statistiques et géospatiales mondiales. Les principaux éléments de cet écosystème sont :

- le Cadre intégré d'information géospatiale, à la fois fondement et guide de l'élaboration, de l'intégration et de la consolidation de la gestion de l'information géospatiale;
- le Groupe de haut niveau sur la modernisation de la statistique officielle, qui favorise l'élaboration d'architectures et de modèles de production de statistiques;
- le Groupe de travail sur l'information géospatiale du Groupe interinstitutions et d'experts sur les indicateurs relatifs aux Objectifs de développement durable, qui fournit une expertise et des conseils sur le rôle crucial de l'information géospatiale dans la réalisation des ODD.

L'utilité et le potentiel du CSGM ont été démontrés par son adoption par plusieurs pays, dont l'infrastructure statistique et géospatiale se situe à des stades de développement différents (p. ex. l'Australie, l'Égypte, le Mexique, la Nouvelle-Zélande, l'Afrique du Sud et la Suède).² De plus, sa valeur a été reconnue et son adoption promue au niveau régional (p. ex. en Afrique, en Europe et en Amérique latine). De plus amples informations sur le processus de mise en œuvre se trouvent dans les documents et les renseignements publiés sur le wiki du Groupe d'experts sur l'intégration des données statistiques et géospatiales (GE-IDSG)³, qui présente les bonnes pratiques et les méthodes de mise en œuvre, ainsi que des études de cas par pays.

Mode d'emploi du document

Le présent document décrit la mise en œuvre globale actuellement recommandée pour le CSGM. Il se divise en trois parties.

La **partie 1** présente le CSGM, y compris ses entrées, ses cinq principes, ses éléments essentiels et ses résultats. Il s'agit d'un résumé synoptique qui vise à fournir de l'information distillée sur l'importance de l'information statistique à référence géospatiale et sur la façon de l'intégrer.

La **partie 2** décrit en détail les cinq principes, en fournissant de plus amples renseignements et des éléments de contexte sur chacun d'eux.

La **partie 3** comprend une série d'annexes.

- L'annexe A présente en détail les définitions des termes courants utilisés dans le document.
- L'annexe B fournit des renseignements de base sur les normes et la qualité des données.
- L'annexe C donne des propositions de lecture.

Le GE-IDSG diffusera progressivement des documents qui appuieront la mise en œuvre. Ces documents aideront les pays à mettre en œuvre le CSGM et présenteront en détail les leçons tirées de différentes mises en œuvre au niveau national et régional, ainsi que les bonnes pratiques, l'élaboration de normes et les cadres d'application. Les documents seront accessibles principalement sur le wiki du GE-IDSG, qui offre un espace de collaboration dans lequel le Groupe d'experts peut publier de l'information et mettre à jour le CSGM.

² <http://ggim.un.org/meetings/GGIM-committee/8th-Session/documents/Global-Statistical-Geospatial-Framework-July-2018.pdf>

³ <https://unstats.un.org/wiki/display/ISGI/United+Nations+Expert+Group+on+the+Integration+of+Statistical+and+Geospatial+Information>

[Page laissée en blanc intentionnellement]

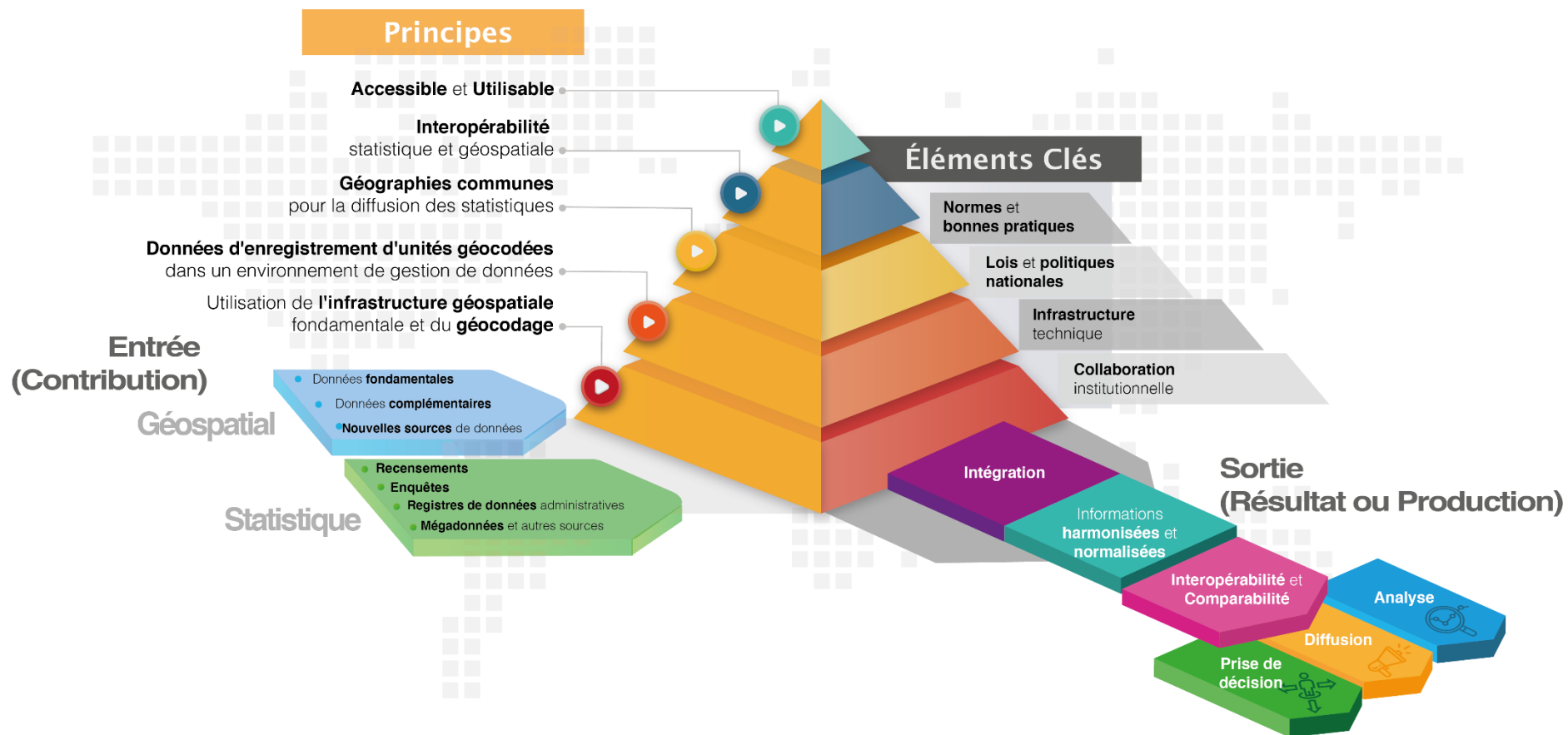


Figure 2 Le cadre statistico-géospatial mondial – des entrées aux résultats

Partie 1 : Le cadre statistico-géospatial mondial

Une vision globale du CSGM est illustrée dans la Figure 2. La figure présente les entrées, les principes, les éléments essentiels et les résultats. Grâce à ces composantes, le CSGM sert de pont entre les professionnels des domaines statistique et géospatial, entre les instituts nationaux de statistique et les organismes nationaux d'information géospatiale (ONIG), ainsi qu'entre les normes, les méthodes, les flux de travail et les outils statistiques et géospatiaux.

Données d'entrée du CSGM

Le CSGM part d'entrées cruciales de données géospatiales fondamentales, qu'il complète par d'autres sources de données géospatiales, au besoin; ces dernières peuvent être fournies par les ONIG et la communauté géospatiale en général⁴. Il sert principalement à ajouter une dimension géospatiale aux données statistiques classiques faisant autorité et, de plus en plus, aux données issues de sources administratives et autres, dont une grande partie provient des INS et des dépositaires de données administratives du système statistique national (SSN) dans son ensemble. Chacune de ces entrées peut être enrichie par des ensembles de données complémentaires, comme celles issues de l'externalisation ouverte ainsi que de technologies ou de sources nouvelles (p. ex. les mégadonnées).

Les cinq principes du cadre statistico-géospatial mondial

Les cinq principes du CSGM guident les processus généraux par lesquels une série d'infrastructures et de processus géospatiaux et statistiques sont appliqués aux entrées de données pour que leur intégration soit possible. Tout d'abord, les données statistiques se voient attribuer une référence géospatiale au niveau le plus fin possible. Ensuite, le recours à des méthodes et des outils géospatiaux, comme des informations géographiques communes et des normes communes de bonnes pratiques, permet de garantir l'interopérabilité, l'accessibilité et l'exploitabilité des données. Les cinq principes sont :

1. l'utilisation de l'infrastructure géospatiale de base et du géocodage;
2. un géocodage unitaire dans un environnement de gestion des données;
3. le recours à des informations géographiques communes pour la diffusion des statistiques;
4. l'interopérabilité des informations statistiques et géospatiales;
5. l'accessibilité et l'exploitabilité des statistiques géospatiales;

⁴Les données géospatiales fondamentales ont été définies par l'initiative des Nations Unies sur la gestion de l'information géospatiale à l'échelle mondiale (UN-GGIM). L'annexe B traite des thèmes des données géospatiales fondamentales mondiales plus en détail.

La Figure 3 illustre la structure hiérarchique de chacun des cinq principes.

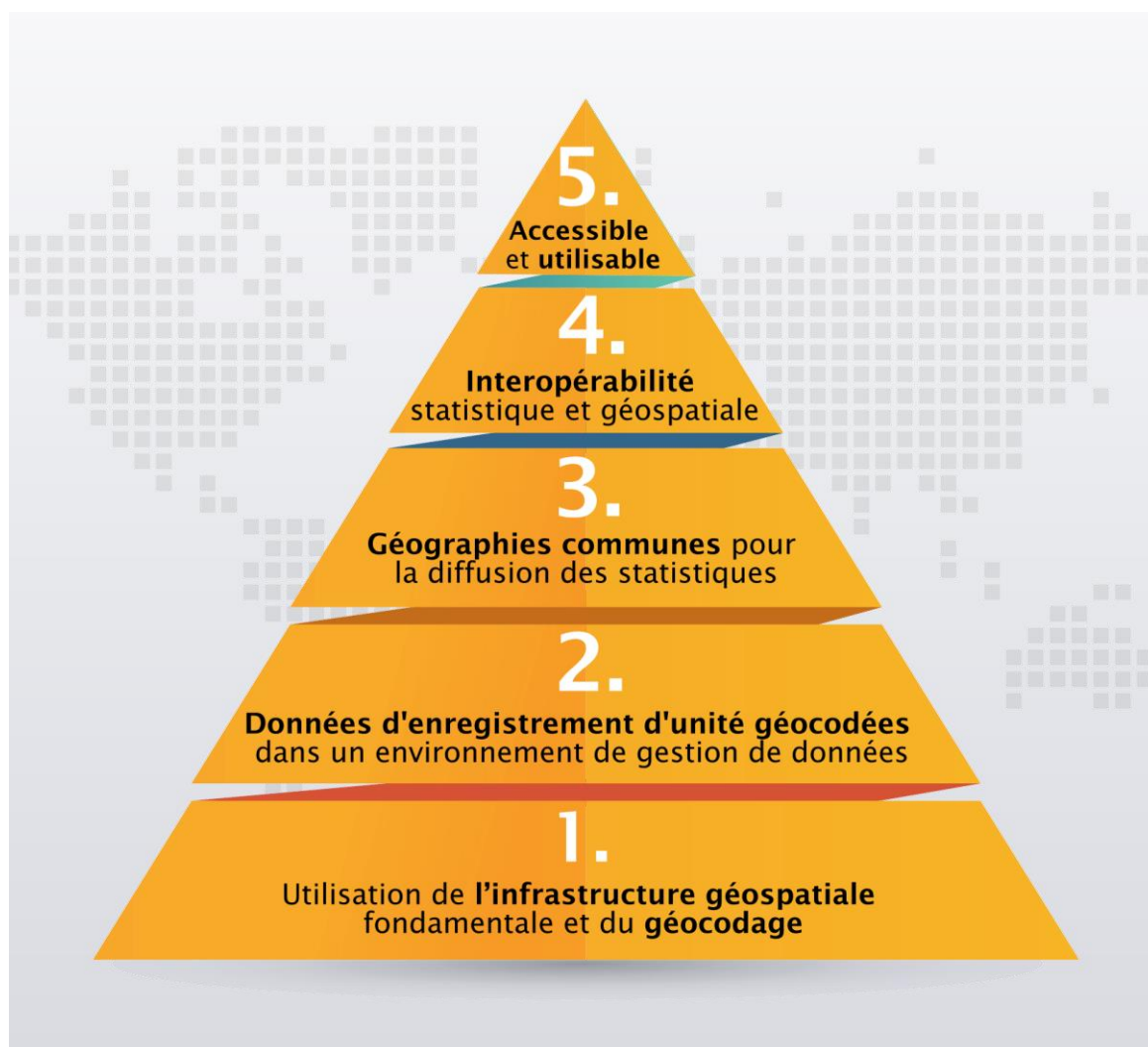


Figure 3 Les cinq principes du CSGM

Ces principes guident d'une part les pays dans l'établissement et la consolidation de leurs processus de collecte de données statistiques et administratives à référence géospatiale, et permettent d'autre part de déterminer les lacunes de leurs infrastructures statistiques et géospatiales nationales respectives, surtout en matière de capacités.

Le CSGM – Quatre éléments essentiels

Quatre éléments essentiels, qui recourent les cinq principes du CSGM, jouent un rôle moteur crucial en permettant l'obtention de données à partir des diverses sources et l'application des principes du cadre statistico-géospatial mondial. Ces éléments sont :

- les normes et bonnes pratiques
 - qui s'appliquent dans l'ensemble des communautés géospatiales et statistiques ainsi qu'aux technologies de l'information et à d'autres domaines;
 - qui comprennent des normes statistiques officielles (notamment des classifications, des concepts et des définitions), d'autres normes (comme les normes ISO), des bonnes pratiques et lignes directrices (comme les meilleures pratiques de l'OGC [Consortium

géospatial ouvert], de l'ISO et du W3C [World Wide Web Consortium]) et d'autres normes *de facto* largement utilisées dans la communauté géospatiale (comme GeoJSON).

- les lois et politiques nationales
 - qui sont des composantes essentielles clés de l'infrastructure législative, professionnelle et sociale permettant et, dans certains cas, restreignant les activités;
 - qui peuvent être la législation visant la protection internationale et nationale des données, la protection de la vie privée et la confidentialité, les exigences en matière d'éthique et d'acceptabilité sociale, les politiques sur les données ouvertes et les accords sur l'accès aux données;
- l'infrastructure technique
 - qui comprend un large éventail de capacités techniques nationales et régionales, englobant les compétences des personnes, les méthodologies et processus établis et convenus, et l'infrastructure des systèmes;
 - qui peut inclure les systèmes de gestion des données, ainsi que l'infrastructure informatique et de communication;
- la collaboration institutionnelle
 - qui nécessite un engagement à l'égard de la collaboration institutionnelle entre les principales parties prenantes, en particulier entre les organismes gouvernementaux responsables des données statistiques, géospatiales et administratives;
 - qui peut être appuyée par des ententes officielles ou des relations de collaboration institutionnelles positives et qui doit inclure des initiatives de sensibilisation et d'éducation.

Résultats du CSGM

Les processus du CSGM produisent des données de sortie ayant un degré plus élevé d'harmonisation structurelle, de normalisation ainsi que de souplesse géospatiale, comparativement aux résultats produits précédemment par les méthodes ad hoc couramment utilisées pour ajouter une référence géospatiale à des données. Les résultats du CSGM permettront un plus grand accès à des données géospatiales et socioéconomiques interopérables, utilisables efficacement dans les différentes disciplines à l'appui des priorités nationales et mondiales.

De par leurs qualités intrinsèques, ces résultats améliorés font plus facilement l'objet d'une intégration fondée sur le lieu et sont plus utilisables dans des processus plus complexes d'intégration de données statistiques. Les résultats obtenus grâce au CSGM ont comme caractéristique importante de permettre de produire des résultats reproductibles de manière plus fiable dans le temps, et ce quel que soit leur auteur. Ils donnent ainsi la possibilité d'accroître l'efficacité et la puissance des analyses et des applications qui éclairent les processus décisionnels.

Il est également important de souligner que les données statistiques et géospatiales servant à la production d'informations statistiques et géospatiales doivent être rendues publiques dans le strict respect de la législation relative à la protection de la vie privée et la confidentialité, ainsi que des attentes usuelles des collectivités concernées. Bien que des renseignements privés et confidentiels puissent être connus et identifiables dans des données recueillies et entreposées en toute sécurité, ils doivent être protégés et rendus confidentiels de manière appropriée en cas de diffusion des données.

Principe 1 : utilisation de l'infrastructure géospatiale de base et du géocodage

Le principe 1 décrit l'adoption d'une méthode commune et uniforme pour placer chaque unité statistique d'un ensemble de données dans le temps et l'espace, au moyen d'une infrastructure géospatiale fondamentale.

L'objectif du principe 1 consiste à obtenir des références de localisation normalisées de grande qualité (comme des identificateurs d'adresse physique, de bien ou d'immeuble, ou d'autres types de description d'emplacement), afin d'attribuer des coordonnées exactes, une petite région géographique ou une coordonnée de carroyage normalisée à chaque unité statistique au niveau de l'enregistrement de microdonnées ou d'unités. De plus, l'horodatage des lieux permettent de placer l'unité précisément dans le temps et l'espace, ce qui répond à une nécessité statistique forte de l'établissement de données dans une série chronologique. De préférence, les lieux sont enregistrés par la saisie directe ou indirecte de leurs coordonnées exprimées en abscisse et en ordonnée. Quand l'infrastructure géospatiale et statistique actuelle d'un pays ne permet pas ce degré de précision, des adaptations s'appuyant sur des descriptions de lieux plus générales ou des zones géographiques plus étendues seront nécessaires.

Le processus d'obtention de localisations et de géocodes doit employer des données géospatiales fondamentales pertinentes⁵, et prendre en charge les fonctionnalités de système d'une infrastructure nationale de données géospatiales (INDG) ou d'autres sources nationales approuvées. Idéalement, chaque pays devrait avoir inclus dans son infrastructure statistico-géospatiale une adresse géoréférencée, un registre d'immeubles, une parcelle cadastrale ou un nom de localité. S'il n'existe pas de registre de ce type, les pays sont enjoint à rapidement mettre à l'essai la mise en œuvre d'un autre référencement spatial fondé sur des points dans les données d'enregistrement d'unités⁶.

La mise en œuvre du principe 1 permet de réaliser les objectifs suivants :

- les informations relatives à l'adresse, au bien, à l'immeuble et à la localisation sont précises et cohérentes, et elles respectent les normes et les bonnes pratiques approuvées par le pays;
- les résultats du géocodage sont précis et cohérents et, dans la mesure du possible, ont recours à des méthodes et des systèmes communs;
- tous les problèmes de géocodage sont gérés de manière uniforme par la mise en application de procédures normalisées.

1

Le principe 1 est axé sur la création d'infrastructures permettant la mise en œuvre et la mutualisation du CSGM.

Cette infrastructure favorise la création de références de localisation normalisées de grande qualité – comme des identificateurs d'adresse physique, de bien ou d'immeuble ou d'autres types de description d'emplacement – et garantit l'attribution de coordonnées exactes et de coordonnées de carroyage normalisées. L'ajout d'une estampille temporelle lie l'élément temporel à l'unité.

⁵ Comme l'indiquent les Thèmes des données géospatiales fondamentales mondiales présentés en détail dans l'annexe B.

⁶ Voir la définition du géocodage donnée dans l'annexe.

Principe 2 : un géocodage unitaire dans un environnement de gestion des données

Le principe 2 appuie le processus de couplage ou de stockage de références géographiques extrêmement précises (c.-à-d. géocodes, coordonnées, codes de petites régions géographiques ou identificateurs de données couplées) correspondant à chaque enregistrement de microdonnées ou d'unités statistiques. C'est ce qu'on désigne couramment par l'expression « données à référence géospatiale ». Ces dernières doivent être produites dans un environnement de gestion de données sécurisé et fondé sur des normes. Ce processus applique l'infrastructure de codage d'adresses et les données fondamentales du principe 1.

L'objectif du principe 2 est de permettre, dans la mesure du possible, de coupler tous les enregistrements d'unités statistiques à une localisation géographique. Cela permettra ensuite l'intégration de données provenant de sources très variées, notamment d'autres données statistiques socioéconomiques, administratives et géospatiales sur l'environnement bâti et naturel. L'intégration de ces données, par traitement géospatial, peut ensuite produire de nouvelles variables statistiques à référence géospatiale à des fins d'analyse. Le principe 2 permet également d'appliquer de manière souple tout contenu géographique lors de la préparation des données en vue de leur diffusion et de leur analyse. Cela appuiera notamment l'agrégation future de données statistiques dans de nouvelles unités géographiques ou l'adaptation aux changements de régions géographiques qui se produisent au fil du temps.

Le principe 2 comprend le recours à des outils, des techniques, des normes et de bonnes pratiques de gestion des données pour faciliter le couplage et la gestion des géocodes dans les ensembles de données statistiques. Cela permet également d'assurer la gestion adéquate des exigences en matière de protection de la vie privée et de confidentialité pour les données diffusées.

La mise en œuvre du principe 2 permet de réaliser les objectifs suivants :

- toutes les microdonnées statistiques ont une référence géospatiale permettant de les utiliser de façon souple dans les processus d'analyse, de visualisation, de diffusion et d'intégration des données statistiques;
- l'agrégation des données pour les grandes régions géographiques est simplifiée par le stockage d'un identificateur ou d'un code unique par petite région géographique ou par maille normalisée standard pour chaque enregistrement d'unité;
- l'adaptation aux changements de zones géographiques ou la compilation de données pour les nouvelles zones géographiques est possible;
- la gestion des données est efficace, y compris en matière de protection de la vie privée et de confidentialité;
- les rôles relatifs à la tenue à jour et la garde des données sont précisément définis;
- l'information et les métadonnées géocodées sont cohérentes, interprétables et systématiquement mises à jour.

2

Le principe 2 permet de coupler chaque enregistrement d'unité statistique à une référence géographique (une coordonnée ou une petite région géographique par exemple) de façon à pouvoir appliquer les statistiques à tout contexte géographique.

Cela facilitera l'intégration ou la corrélation des données provenant d'autres sources de données et atténuera les difficultés causées par de nouvelles zones géographiques ou des changements de régions.

Principe 3 : le recours à des informations géographiques communes pour la diffusion des statistiques

Le principe 3 emploie la zone géographique comme instrument d'intégration des données. Il utilise un ensemble commun et convenu d'informations géographiques aux fins d'affichage, de stockage, de déclaration et d'analyse de comparaisons sociales, économiques et environnementales entre ensembles de données statistiques provenant de différentes sources. Le principe 3 établit l'importance fondamentale que revêt la mise en place d'un équilibre entre les informations géographiques statistiques et administratives actuelles et d'autres systèmes de référence géographique – par exemple les quadrillages – comme fondement des informations géographiques communes à tous les ensembles de données.

L'objectif du principe 3 consiste à promouvoir un ensemble d'informations géographiques communes qui assure la cohérence de l'agrégation et de la diffusion des données statistiques à référence géospatiale, qu'elles se trouvent dans des découpages géographiques administratifs ou quadrillés. Les données sont allouées uniformément à des segments administratifs ou des unités statistiques plus petits (comme des blocs de grille à mailles), divisés selon les subdivisions politiques, foncières ou topologiques, ou elles sont attribuées de manière cohérente à des unités de carroyage de tailles différentes (c.-à-d. des carrés ou des pixels). De plus, le principe 3 permet la translation et la mise en correspondance d'informations statistiques entre les découpages géographiques administratifs ou quadrillés.

La mise en œuvre du principe 3 permet de réaliser les objectifs suivants :

- les données de différentes sources sont intégrables grâce à des entités géographiques communes;
- la visualisation, l'analyse et l'interprétation des informations statistiques et géographiques sont simplifiées;
- les métadonnées appuient l'agrégation, l'intégration et l'utilisation des données;
- la définition et l'application de méthodes d'agrégation et de désagrégation améliorent la qualité des données ainsi que l'évaluation, la cohérence et l'utilisation accrue des données;
- la conversion des données en informations géographiques est prise en charge par des mécanismes de conversion normalisés (par exemple par des correspondances⁷).

3

Le principe 3 détermine les définitions des régions géographiques ainsi que l'agrégation et la désagrégation de données à des régions; il permet la cohérence et la comparabilité des données statistiques et géospatiales intégrées.

Le fait de disposer d'un ensemble d'informations géographiques communes garantit que les données statistiques ont des références géospatiales uniformes et qu'elles peuvent être intégrées au niveau agrégé. Cela permet aussi aux utilisateurs de découvrir, d'accéder, d'intégrer, d'analyser et de visualiser les informations statistiques de façon transparente dans les régions géographiques d'intérêt.

⁷ Pour en savoir plus sur les méthodes de correspondance, voir : <http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Correspondences>

Principe 4 : l'interopérabilité des informations statistiques et géospatiales

Données, normes, processus et organismes

Le principe 4 définit les conditions préalables au fonctionnement des données statistiques et géospatiales comme un écosystème de données, dans lequel les parties prenantes interagissent pour échanger, produire et consommer des données. Pour dépasser les obstacles structurels, sémantiques et syntaxiques entre les données et les métadonnées issues de communautés et de fournisseurs différents, il est nécessaire que les normes des données statistiques et géospatiales et celles des métadonnées soient interopérables.

De plus, il est nécessaire d'améliorer l'efficacité de la découverte, de l'accessibilité et de l'utilisation des données géospatiales. Souvent, l'interopérabilité complète des données exige d'abord la suppression des obstacles présents dans les lois, politiques et organismes nationaux qui entravent la coopération entre les parties prenantes et forment des barrières entre les producteurs et les utilisateurs finaux.

L'amélioration de l'interopérabilité permet aux communautés statistiques et géospatiales de continuer à exploiter leurs propres modèles généraux de données ainsi que leurs architectures et leurs fonctionnalités de métadonnées tout en intégrant et reliant efficacement les ensembles de données de différents systèmes et applications et en y accédant facilement. Par conséquent, le principe 4 préconise l'utilisation de normes et de bonnes pratiques adoptées internationalement par les deux communautés pour accroître l'interopérabilité des données statistiques et géospatiales, des normes, des processus et des organismes.

La mise en œuvre du principe 4 permet de réaliser les objectifs suivants :

- une plus grande efficacité et simplification dans la création, la découverte, l'intégration et l'utilisation de statistiques à référence géospatiale et de données géospatiales;
- la mise en place de mécanismes d'accès axés sur le service ou exploitables par une machine (par exemple au moyen d'API) pour accroître l'efficacité de l'accès et de l'utilisation et permettre l'adaptation et l'évolution des usages dans le temps;
- l'augmentation des possibilités d'application à un plus grand éventail de données et de technologies.

4

En augmentant la normalisation et l'interopérabilité des données, le principe 4 accroît l'efficacité et la simplification de la création, de la découverte, de l'intégration et de l'utilisation de statistiques et de données géospatiales.

Il devrait ainsi être possible d'employer un éventail plus large de données et de technologies et d'obtenir par conséquent des informations plus variées, qui seront exploitables dans la prise de décision. Cela améliorera également la coopération entre toutes les parties prenantes produisant et utilisant des informations statistiques et géospatiales.

Principe 5 : l'accessibilité et l'exploitabilité des statistiques géospatiales

Le principe 5 souligne la nécessité pour les dépositaires de données de rendre les statistiques à référence géospatiale accessibles et exploitables selon les normes et les bonnes pratiques convenues, afin que les utilisateurs de données puissent découvrir, consulter, intégrer, analyser et visualiser cette information de façon transparente pour les zones géographiques d'intérêt. Il traite de la nécessité de définir ou, au besoin, d'élaborer des politiques, des normes, des bonnes pratiques et des technologies qui appuient ces utilisations.

Les instituts nationaux de statistique (INS) et les organismes nationaux d'information géospatiale (ONIG) doivent tenir compte de questions législatives et opérationnelles très diverses lorsqu'ils publient et analysent de l'information sur des personnes et des entreprises. C'est pourquoi un aspect important de ce principe consiste à veiller à ce que les données soient accessibles au moyen de mécanismes sécurisés, qui protègent la vie privée et la confidentialité tout en permettant une analyse de données favorisant la prise de décision fondée sur des données probantes. Parmi les autres questions intéressantes, citons la qualité des données dans leurs différentes dimensions (en particulier en ce qui concerne leur fiabilité, leur actualité et leur pertinence) et l'accès aux capacités d'analyse, de diffusion et de visualisation.

L'objectif du principe 5 est de promouvoir la diffusion d'informations statistiques à référence géospatiale sous une forme exploitable et accessible. Il favorise expressément l'utilisation de services Web et de méthodes de couplage de données normalisés qui fournissent un accès dynamique et exploitable par machine à ces données, avec les garanties nécessaires quant à l'intégrité des données.

La mise en œuvre du principe 5 permet de réaliser les objectifs suivants :

- les dépositaires de données peuvent publier des données dont la confidentialité et les droits de propriété sont protégés;
- les utilisateurs de données peuvent découvrir des statistiques à référence géospatiale et y accéder;
- les utilisateurs de données peuvent entreprendre des analyses et visualiser les données;
- les services Web et les méthodes de couplage des données permettent un accès machine-machine, ainsi que le couplage dynamique des informations;
- les utilisateurs de données peuvent connaître l'état des modifications ou des changements apportés aux données fournies par les dépositaires de données, au moyen de vérifications de l'intégrité des données.

5

Le principe 5 aide les dépositaires de données à diffuser les données en toute confiance, à améliorer la découverte et l'accessibilité des statistiques à référence géospatiale (en particulier en encourageant les services Web à fournir des couplages dynamiques et exploitables par une machine aux données), et à soutenir l'analyse et l'évaluation des données dans la prise de décision.

Initiatives complémentaires

Le CSGM se situe dans un écosystème informationnel au développement rapide qui renforce les communautés statistiques et géospatiales mondiales. Un grand nombre de cadres et de ressources d'information connexes complètent et interagissent avec le CSGM :

- le Cadre intégré d'information géospatiale⁸, adopté par le Comité d'experts de l'initiative des Nations Unies sur la gestion de l'information géospatiale à l'échelle mondiale (UN-GGIM) au moyen de la résolution 8/113 en 2018, promeut, appuie et offre la direction, la coordination et les normes nécessaires pour fournir des informations géospatiales intégrées exploitables dans la recherche de solutions durables aux problèmes de développement social, économique et environnemental. Composé de 7 principes fondamentaux, 8 objectifs et 9 voies stratégiques, le Cadre intégré d'information géospatiale propose une base, une référence et un mécanisme permettant aux États membres de prendre des mesures avantageuses aux niveaux local, infranational, national, régional et mondial pour traiter les difficultés rencontrées dans la collecte, le traitement, l'analyse et la diffusion de données géospatiales fiables, livrées en temps opportun, accessibles et cohérentes et de leurs informations connexes;
- le Groupe de haut niveau sur la modernisation de la statistique officielle favorise l'élaboration d'architectures et de modèles de production de statistiques modernisés. Ce groupe appuie l'élaboration de l'Architecture commune de production statistique⁹, qui procure un lien pratique entre les modèles conceptuels plus précis que sont le modèle générique de l'information statistique (MGIS) et le modèle statistique général du processus opérationnel (MSGPO) et la production pratique de données statistiques;
- le Groupe de travail sur l'information géospatiale du Groupe interinstitutions et d'experts sur les indicateurs relatifs aux objectifs de développement durable fournit une expertise et des conseils sur le rôle crucial de l'information géospatiale dans la réalisation des ODD, y compris sur des thèmes comme la désagrégation modélisée des données et la possibilité d'utiliser des observations de la Terre.

Ces cadres et ces ressources d'information serviront à consolider et à compléter la mise en œuvre du CSGM. Ainsi, les pays seront en mesure de participer au rapprochement des communautés statistiques et géospatiales, de favoriser et d'accroître la production efficace d'informations statistiques géospatiales, d'améliorer la comparaison des données entre zones géographiques et pays et, finalement, de proposer des informations plus intégrées qui appuieront des analyses et une prise de décision fondée sur des informations factuelles. En outre, le CSGM peut jouer un rôle de catalyseur de ces cadres dans l'écosystème informationnel général d'un pays.

⁸ <http://ggim.un.org/meetings/GGIM-committee/8th-Session/documents/Part%201-IGIF-Overarching-Strategic-Framework-24July2018.pdf>

⁹ <https://statswiki.unece.org/display/CSPA/Common+Statistical+Production+Architecture>

Appel à l'action et voie à suivre

L'élaboration du CSGM vise à répondre aux besoins de données sous-jacents du cycle de recensements de la population et de l'habitat de 2020 et du Programme de développement durable à l'horizon 2030. L'intégration des informations statistiques et géospatiales est essentielle pour la libération du potentiel des données, qui permettra de prendre des décisions éclairées et fondées sur des informations factuelles à tous les niveaux d'un pays, ainsi que d'éclairer l'action régionale et mondiale. Les données intégrées et les statistiques à référence géospatiale qui en résultent contribueront aux plans de développement et fourniront des informations précieuses sur les piliers sous-jacents de la société, de l'économie et de l'environnement d'un pays.

Pour tout pays envisageant de progresser vers la production de statistiques géospatiales, la première étape importante consiste à fournir des données géocodées, qui sont une composante fondamentale et indispensable du CSGM. C'est pourquoi les pays sont invités à réfléchir à une manière de tirer parti des technologies et méthodes nouvelles et émergentes pour assurer la production de données géocodées, et à adopter le CSGM, ainsi que de nouvelles technologies et approches, dans leur architecture de production statistique et géospatiale. Pour cela, ils devront notamment acquérir des compétences en matière de gestion des données et d'analyse géospatiale et statistique, compétences qui, en fin de compte, renforceront l'ensemble de l'écosystème informationnel d'un pays, et non pas seulement les communautés statistiques et géospatiales.

Le CSGM est complété par des cadres généraux, comme le Cadre intégré d'information géospatiale, et les initiatives à venir, notamment le Groupe de travail sur les domaines statistiques de l'OGC (Open Geospatial Consortium)¹⁰. Par conséquent, il est important de retenir que tous les pays auront du chemin à parcourir pour mettre en place le CSGM, quel que soit leur point de départ. Beaucoup auront besoin de l'aide de leurs voisins, de pays de leur région et de ceux ayant déjà progressé dans la mise en œuvre du cadre.

Tout pays souhaitant mettre en œuvre le CSGM doit d'abord s'assurer que, d'une part, les domaines d'action prioritaires sont précisément évalués, planifiés et convenus et que, d'autre part, les rôles et les contributions des principales parties prenantes et de plusieurs institutions nationales, en particulier les INS et les ONIG, sont établis d'un commun accord. Les renseignements présentés dans le document apportent une orientation essentielle sur les principaux éléments à prendre en compte dans ce processus et dans le travail de coopération qui devra être accompli par toutes les parties. Les commissions régionales de l'ONU fourniront également des ressources utiles au renforcement des capacités et aux activités de promotion, notamment par l'échange d'informations au sein des régions.

Dans le contexte du cycle de recensements de la population et de l'habitat de 2020 et du Programme de développement durable à l'horizon 2030, l'urgence de la mise en place du CSGM apparaît clairement. Ce processus peut se révéler difficile, mais les efforts des pays seront récompensés par l'amélioration des informations, de l'interopérabilité et de leurs décisions.

¹⁰ <http://www.opengeospatial.org/projects/groups/statisticaldwg>

Partie 2 : Description détaillée du cadre statistico-géospatial mondial



Principe 1 : utilisation de l'infrastructure géospatiale de base et du géocodage

Principe 1 : L'utilisation de l'infrastructure géospatiale de base et du géocodage est axée sur la création d'infrastructures permettant la mise en œuvre et la mutualisation du CSGM. Cette infrastructure favorise la création de références de localisation normalisées de grande qualité – comme des identificateurs d'adresse physique, de bien ou d'immeuble ou d'autres types de description d'emplacement – et garantit l'attribution de coordonnées exactes et de coordonnées de carroyage normalisées. L'ajout d'une empreinte temporelle peut apporter un élément temporel à l'unité.

Pourquoi ce principe est-il nécessaire?

Grâce à la collecte d'informations exactes sur le lieu, on associe aux données d'entrée une localisation de grande qualité, persistante et normalisée, comme un identificateur d'adresse physique, de bien ou d'immeuble ou un autre type de description de lieu. Cela permet d'attribuer des coordonnées exactes, une petite zone géographique ou une coordonnée de carroyage normalisée à chaque unité statistique (c.-à-d. au niveau de l'enregistrement de microdonnées ou d'unités dans les ensembles de données statistiques). L'ajout d'une empreinte temporelle peut apporter un élément temporel à l'enregistrement d'unité statistique. Quand l'infrastructure géospatiale, statistique ou juridique et politique en vigueur dans un pays ne permet pas d'attribuer des géocodes précis de coordonnées aux enregistrements, le recours au géocodage obtenu au moyen de descriptions plus générales ou contextuelles des lieux ou des zones géographiques plus grandes sera nécessaire et constitue un compromis adéquat. De plus, l'utilisation d'identificateurs uniques persistants peut fournir des liens vers des données non spatiales ainsi que des liens vers des renseignements historiques et les mécanismes de contrôle de versions des fonctions.

Que recouvre ce principe?

Le principe 1 renforce les capacités des pays et des responsables de la mise en application en leur fournissant des informations sur les principales données et l'infrastructure nécessaires à une méthode commune et cohérente d'établissement de localisation et de géocode pour chaque unité statistique d'un ensemble de données, comme une personne, un ménage, une entreprise, un immeuble ou encore une parcelle cadastrale ou une unité territoriale, avec un enregistrement d'estampille temporelle correspondant à chaque enregistrement d'unité.

Les données d'entrée du principe 1 doivent provenir de sources de données normalisées afin d'assurer la qualité, l'exactitude, le caractère actualisé et la cohérence des données provenant des dépositaires de données du système statistique national (SSN) et de l'infrastructure nationale de données géospatiales (INDG). Cette exigence permet de respecter les priorités nationales, les programmes internationaux, ainsi que les normes et les bonnes pratiques internationalement reconnues et approuvées. L'un des principaux moteurs du CSGM consiste à appuyer les besoins d'agrégation et de désagrégation statistiques et ceux d'intégration des données, afin d'éclairer la prise de décision fondée sur des données à différents niveaux géographiques. Pour ce faire, la Commission de

statistique¹¹ recommande que toutes les données d'enregistrement d'unité statistique soient recueillies ou associées à une référence de localisation, de préférence à des coordonnées. Quand cela est impossible, il faut associer chaque unité statistique à une zone géographique (c.-à-d. à un polygone) ou une maille. De plus, partout où l'infrastructure¹² existe, les données géospatiales fondamentales¹³ des organismes nationaux devraient servir dans le référencement géospatial et d'autres activités des communautés de données statistiques et administratives.

Objectifs

Le principal objectif du principe 1 est la création d'une infrastructure géospatiale fondamentale et résiliente pour le traitement des données statistiques et géospatiales. En procurant l'interopérabilité nécessaire à la combinaison de données provenant de diverses sources, par exemple grâce à des données de localisation normalisées de grande qualité – comme un identificateur d'adresse physique, de bien ou d'immeuble, ou tout autre élément de localisation précis, actuel, uniforme et normalisé – elle permet d'attribuer des coordonnées précises.

Exigences et avantages

Les avantages principaux suivants sont directement liés à une infrastructure géospatiale fondamentale harmonisée et sont pertinents pour les cinq principes du CSGM :

- il permet le partage de données et d'informations à des fins d'analyse générale;
- il favorise l'utilisation d'informations géocodées;
- il favorise l'utilisation de normes et de bonnes pratiques;
- il rend les données accessibles de façon transparente grâce à un accès facile à l'information;
- il ouvre de nouvelles perspectives et dévoile de nouvelles relations entre les données;
- il favorise la visualisation des données;
- il soutient les processus stratégiques et décisionnels (aux niveaux local, infranational, national, régional et mondial);
- il appuie l'établissement de liens cohérents dans le temps entre les statistiques et les informations de localisation géographique en tenant à jour l'information sur le cycle de vie des caractéristiques géospatiales;
- il apporte des résultats reproductibles permettant de produire des résultats transparents de grande qualité;
- il procure des statistiques influentes et lisibles;
- il propose de nouvelles sources d'information géospatiale pour la mesure et le suivi des cibles et des indicateurs des ODD;
- il permet la mise en œuvre de méthodes, de processus d'évaluation et de techniques de visualisation nouveaux dans les recensements de la population et d'autres actions de collecte statistique, permettant ainsi de comparer dans le temps des données similaires.

¹¹ E/CN.3/2018/33

¹² Quand l'application d'un géoréférencement par coordonnées exprimées en abscisse et en ordonnée n'est pas possible, il est conseillé d'utiliser des descriptions de localisation et des zones géographiques plus générales ou plus grandes comme références géospatiales dans les ensembles de données. Elles se trouvent dans des ensembles de données faisant autorité, comme des registres d'adresses ou d'immeubles.

¹³ Comme les Thèmes des données géospatiales fondamentales mondiales (présentés en détail dans l'annexe B).

Lien avec d'autres principes

Le principe 1 est le fondement sur lequel reposent les principes subséquents, en particulier le principe 2. À son tour, le principe 4 a une incidence importante sur le principe 1, car l'utilisation de normes, de modèles de données et de l'harmonisation du contenu des données est un élément crucial de la construction d'une infrastructure géospatiale fondamentale. Le principe 3 fournit des informations géographiques statistiques pour les données géospatiales fondamentales, qui constituent une composante clé de ce principe.

Entrées

Données statistiques

Unités géographiques territoriales ou administratives ou zones géographiques définies à des fins statistiques (y compris les secteurs de dénombrement et les îlots de recensement). Statistiques sociales, statistiques économiques, statistiques démographiques, données de recensement, statistiques agricoles, statistiques environnementales et autres ensembles de données statistiques et administratives.

Données géospatiales fondamentales

Les thèmes des données géospatiales fondamentales mondiales sont un ensemble de 14 thèmes, approuvés par l'Initiative des Nations Unies sur la gestion de l'information géospatiale à l'échelle mondiale¹⁴, qui visent à soutenir le renforcement des capacités des ONIG et des INS. La réalisation des thèmes nécessitera d'intégrer des données et des informations statistiques provenant des ONIG, des INS et d'autres institutions afin de produire des données géospatiales fondamentales normalisées qui seront employées dans les États membres et dans le soutien à des initiatives comme les ODD. Généralement, la gestion de ces thèmes de données géospatiales fondamentales relève du domaine des INDG. La mise en application de ces thèmes constitue une composante fondamentale de l'infrastructure géospatiale des pays.

Principales parties prenantes

Les parties prenantes susceptibles de contribuer au principe 1, de l'adopter et de le mettre en œuvre comprennent les INS (englobant le SSN), les ONIG (y compris le système d'information géospatiale du pays), les ONG, la société civile, les entreprises privées, les fournisseurs de données et les citoyens. Étant donné le nombre important de parties prenantes, il est recommandé de recourir à un environnement de données ouvertes en raison des efficacités et des possibilités qu'offre un tel contexte. Ces avantages sont contrebalancés par la nécessité d'étudier la meilleure façon d'assurer la protection des renseignements personnels des citoyens conformément à l'évolution des lois et des règlements sur la protection de la vie privée en vigueur à l'échelle nationale ou régionale, tout en tenant compte des normes et des recommandations internationales.

Afin d'améliorer la prise de décision éclairée, les responsables de la mise en œuvre (principalement les INS et les ONIG) sont invités à établir des liens solides de communication et de collaboration institutionnelle. Ils peuvent prendre la forme, entre autres, de lois et de politiques nationales, de protocoles d'entente ou de communautés de pratique.

¹⁴ Comme il est indiqué en annexe du rapport : http://ggim.un.org/meetings/GGIM-committee/8th-Session/documents/E-C20-2018-7-Add_1-Global-fundamental-geospatial-data-themes.pdf

Principe 2 : un géocodage unitaire dans un environnement de gestion des données

Principe 2 : un géocodage unitaire dans un environnement de gestion des données. À partir de l'infrastructure du principe 2, ce principe permet de coupler chaque enregistrement d'unité statistique à une référence géographique (une coordonnée ou une petite région géographique par exemple) de façon à pouvoir appliquer les statistiques à tout contexte géographique. Cela facilitera l'intégration ou la corrélation des données provenant d'autres sources de données et atténuera les difficultés causées par de nouvelles zones géographiques ou des changements de régions.

Pourquoi ce principe est-il nécessaire?

Tous les enregistrements d'unités statistiques doivent avoir une référence géospatiale – être géocodés selon une localisation exacte – dans un environnement de gestion des données sécurisé et fondé sur des normes. Cela permet ensuite d'intégrer ou de coupler des données provenant de sources de données très diverses et d'appliquer les produits statistiques à tout contexte géographique, y compris à des zones géographiques nouvelles ou modifiées.

Que recouvre ce principe?

Le principe 2 porte sur le processus de couplage et de stockage d'une référence géographique très précise (c.-à-d. des géocodes – des coordonnées ou de petites régions géographiques) pour chaque enregistrement d'unité statistique dans un environnement de gestion des données sécurisé et fondé sur des normes. Cela permet de produire à partir de ces ensembles de données des statistiques pour un éventail large de contextes géographiques – par exemple, diverses zones géographiques administratives et statistiques, systèmes de quadrillage (carroyage) – et de contribuer à l'agrégation future de données statistiques dans de nouvelles unités géographiques ou à l'adaptation à des modifications de zones géographiques.

Il est ainsi possible d'intégrer ou de coupler des données provenant de sources très diverses, notamment des statistiques administratives ou des données géospatiales sur l'environnement bâti et naturel, comme des données d'observation de la Terre. Ces données peuvent ensuite faire partie d'analyses géospatiales, par exemple la mesure de la distance entre les ménages et un espace libre public, des services de santé et d'autres services ou activités. De plus, ce principe permet d'une part des processus de couplage de données qui dépendent d'informations précises sur le lieu; et d'autre part une amélioration de ces mêmes processus, comme par exemple, dans l'appariement des dossiers administratifs et du recensement en fonction de données de localisation et d'autres caractéristiques démographiques. Enfin, il garantit une gestion sécurisée des données détaillées de sorte que le droit d'accès aux renseignements susceptibles d'être confidentiels ou sensibles soit approprié et sans risque.

Objectifs

Dans la mise en œuvre du principe 2, les objectifs suivants doivent être visés :

- permettre la production de statistiques à référence géospatiale;
- favoriser l'association de tous les enregistrements d'unités statistiques à une référence géographique précise (idéalement des coordonnées exprimées en abscisse et en ordonnée), ou sinon à la plus petite zone géographique possible. Cela favorise au maximum la réutilisation et l'agrégation de données statistiques à référence géospatiale. Lorsque l'infrastructure géospatiale de base ne suffit pas pour cette activité, il faut employer d'autres méthodes, comme la capture directe de coordonnées et l'utilisation de zones statistiques ou de dénombrement de recensement;
- assurer l'efficacité de la mise en œuvre de l'infrastructure géospatiale fondamentale et de géocodage de façon à montrer sa valeur générale. Cela implique de collaborer avec les INS, les ONIG et d'autres fournisseurs de données fondamentales pour que les exigences soient comprises et que les données et l'infrastructure sont utilisées de façon appropriée;
- mettre en œuvre une gestion efficace des données statistiques et géospatiales. Cela exige de bonnes pratiques de gestion des données techniques et des métadonnées, conformément aux normes et aux bonnes pratiques nationales et internationales;
- protéger adéquatement la vie privée et la confidentialité des ensembles de données au niveau des enregistrements unitaires ou des microdonnées. Pour cela, il faut protéger les données sensibles et confidentielles tout en assurant un accès approprié qui réponde aux besoins des utilisateurs en matière d'analyse et de prise de décision;
- stocker des références géographiques cohérentes et interprétables, de préférence couplées à partir d'un « point fiable » (p. ex. celles liées à un registre d'adresses géré de façon centralisée). Cela nécessite d'établir et de mettre en œuvre des normes sur les données et les métadonnées pour veiller à ce que les références géographiques soient correctement documentées et cohérentes entre ensembles de données, ce qui permet de les utiliser efficacement dans le temps et dans différentes applications;
- veiller à ce que les données soient stockées de manière à faciliter l'application souple de tout contenu géographique lors de la préparation des données à des fins de diffusion et d'analyse. Cela implique de soutenir l'agrégation future des données statistiques dans de nouvelles unités géographiques ou l'adaptation aux changements de régions géographiques qui se produisent au fil du temps ainsi que de veiller à ce que les données soient protégées pour prévenir toute utilisation abusive.

Lien avec d'autres principes

Il est étroitement lié au principe 1, car les activités du principe 2 dépendent des données fondamentales fournies par l'infrastructure de données géospatiales et les capacités de géocodage d'un pays et doivent utiliser ces données.

De plus, le principe 2 est un catalyseur essentiel du principe 3, car ces zones géographiques dépendent de la définition des régions géographiques communes aux fins de diffusion des données prévue dans le principe 3, et elles utilisent cette définition. Elles peuvent également être déterminées dans les données géospatiales fondamentales ou à l'échelle du pays décrites dans le principe 1. Les métadonnées et les données connexes pour les régions géographiques communes définies dans le principe 3 devraient également être appliquées aux activités du principe 2. De plus, les zones géographiques, les données connexes, les normes et les bonnes pratiques du principe 3 peuvent être

tirées des données, des outils, des normes et des bonnes pratiques de la communauté géospatiale ou statistique.

Entrées

Données statistiques

Du point de vue statistique, le MSGPO et le MGIS¹⁵ gérés par le Groupe de haut niveau sur la modernisation de la statistique officielle (GHN-MSO) de la Commission économique pour l'Europe des Nations Unies (CEE-ONU) sont des modèles conçus pour améliorer, intégrer, promouvoir, soutenir et faciliter la modernisation de la statistique¹⁶. Le projet GEOSTAT¹⁷ est une initiative du réseau SSE (système statistique européen) qui promeut l'intégration de l'information statistique et géospatiale pour la communauté statistique dans le but de proposer des descriptions et des analyses de la société et de l'environnement de plus grande qualité. GEOSTAT3 a formulé des recommandations visant à harmoniser la mise en œuvre du CSGM au sein de toute la communauté statistique européenne, en tenant compte des conditions actuelles, des initiatives et des cadres européens et nationaux.

Les ressources supplémentaires sont :

- l'architecture commune de la production statistique (ACPD) du GHN-MSO de la CEE-ONU¹⁸;
- les principes de gestion des données¹⁹;
- les lois, politiques ou protocoles relatifs à la protection de la vie privée adoptés par un pays et à l'échelle internationale. Voici d'autres ressources actuelles :
 - Principes fondamentaux de la statistique officielle²⁰;
 - Principes et lignes directrices concernant la gestion de la confidentialité et de l'accès aux microdonnées²¹;
 - résultats de l'étude de la Division de statistique des Nations Unies (DSNU) et de la CEE-ONU sur le contexte organisationnel et les projets individuels de mégadonnées²²;
 - les lois et politiques en vigueur au niveau national et régional.

Données géospatiales

Du point de vue géospatial, les normes de l'OGC, ISO et de l'OHI (Organisation hydrographique internationale), y compris celles énoncées dans le « Guide sur le rôle des normes dans la gestion de l'information géospatiale », orientent sur les meilleures pratiques relatives aux normes géospatiales internationales dans l'infrastructure de données géospatiales. Ces normes évolueront et changeront dans le cadre des processus de l'OGC, de l'Organisation internationale de normalisation (ISO) et de l'OHI et comprendront les remarques du Groupe de travail sur les domaines statistiques récemment établi²³ sous les auspices de l'OGC.

¹⁵ Les versions actuelles de ces modèles comportent des outils, des méthodes et des processus géospatiaux dans leurs définitions et ressources. Au début de l'année 2019, le nombre d'exemples d'applications nationales intégrant des processus géospatiaux n'était pas élevé, mais il devrait augmenter.

¹⁶ <https://statswiki.unece.org/display/hlgbas>

¹⁷ Les projets GEOSTAT donnent aux pays et aux régions un exemple de mise en œuvre du CSGM, ainsi que des ressources plus détaillées et des conseils pratiques concernant son application <https://www.efgs.info/geostat/>

¹⁸ <https://statswiki.unece.org/display/hlgbas/Modernisation+Groups>

¹⁹ <https://statswiki.unece.org/display/DA/VI.+Key+principles>

²⁰ <https://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx>

²¹ <https://unstats.un.org/unsd/statcom/doc07/BG-Microdata-E.pdf>

²² <https://unstats.un.org/unsd/statcom/doc15/BG-BigData.pdf> (page 10)

²³ <http://www.opengeospatial.org/projects/groups/statisticaldwg>

Les ressources supplémentaires sont :

- les principes de gestion des données du Système mondial des systèmes d'observation de la Terre (GEOSS)²⁴;
- les normes et l'infrastructure d'adressage ou de détermination des positions (voir le principe 1 du CSGM);
- les outils d'infrastructure de géocodage, les normes de métadonnées et les bonnes pratiques, y compris le géocodage et la validation des adresses par point de contact et par lots (voir le principe 1 du CSGM);
- la promotion du géocodage et de la validation des adresses aux points d'entrée (voir le principe 1 du CSGM);
- les lois, politiques ou protocoles relatifs à la protection de la vie privée adoptés par un pays et à l'échelle internationale. Voici d'autres ressources actuelles :
 - régions géographiques convenues, et autres, et l'infrastructure connexe de données et de métadonnées (voir le principe 3);
 - cadres de référence géodésiques mondiaux, nationaux ou régionaux, comme le repère de référence géodésique mondial²⁵.

Principales parties prenantes

Communauté géospatiale

Fourniture :

- de données et d'infrastructures géospatiales fondamentales, ainsi que de capacités de géocodage (selon le principe 1 du CSGM);
- de cadres de référence géodésiques nationaux, régionaux et mondiaux et leur mise en œuvre;
- de cadres de gestion des données géospatiales;
- de normes et de bonnes pratiques en matière de données géospatiales, en particulier de spécifications de métadonnées de géocodage;
- d'un appui aux concepts et aux zones géographiques communs.

Communauté statistique

- Protocoles de protection de la vie privée nationaux et internationaux (p. ex. Principes fondamentaux de la statistique officielle des Nations Unies);
- cadres de gestion des données statistiques, y compris les registres de base (p. ex. registres d'adresses, d'immeubles et d'entreprises);
- appui aux concepts et aux zones géographiques et statistiques communs.
- données non issues d'enquêtes, y compris les données provenant de scanners et de capteurs et les données obtenues par externalisation ouverte;
- mise en œuvre des principes relatifs aux données d'enregistrement d'unités statistiques et administrative, à leur stockage et à leur gestion.

²⁴ https://www.earthobservations.org/documents/dswg/201504_data_management_principles_long_final.pdf

²⁵ <http://www.unggrf.org>

Mise en œuvre des principes relatifs aux données d'enregistrement d'unités administratives, à leur stockage et à leur gestion²⁶.

Principe 3 : le recours à des informations géographiques communes pour la diffusion des statistiques

Principe 3 : le recours à des informations géographiques communes pour la diffusion des statistiques. Ce principe détermine les définitions des régions géographiques ainsi que l'agrégation et la désagrégation de données à des régions; il permet la cohérence et la comparabilité des données statistiques et géospatiales intégrées. Le fait de disposer d'un ensemble d'informations géographiques communes garantit que les données statistiques ont des références géospatiales uniformes et qu'elles peuvent être intégrées au niveau agrégé. Cela permet aussi aux utilisateurs de découvrir, de consulter, d'intégrer, d'analyser et de visualiser les informations statistiques de façon transparente dans les régions géographiques d'intérêt.

Pourquoi ce principe est-il nécessaire?

La prise de décision éclairée et fondée sur des informations factuelles nécessite la comparaison de données statistiques entre régions géographiques. Cela est possible grâce à des zones géographiques communes. Les zones géographiques communes peuvent comprendre des îlots de recensement, des quadrillages, ou des découpages administratifs et cadastraux. Les zones géographiques communes permettent de produire des déclarations statistiques de base, des analyses géostatistiques et des visualisations à différentes échelles (par exemple locale, infranationale, nationale, régionale²⁷ et mondiale, ainsi qu'à l'échelle institutionnelle²⁸), puis de comparer ou d'évaluer les résultats de façon cohérente. Ces zones géographiques fournissent également un mécanisme de gestion de la protection de la vie privée et des produits de données statistiques et géospatiales confidentiels.

Que recouvre ce principe?

L'instauration d'informations géographiques communes permet de produire et diffuser des statistiques et des informations géospatiales intégrées afin d'éclairer la prise de décision. Il est possible d'y parvenir en encourageant la participation des parties prenantes à la création de classifications géographiques (une classification constituée d'une ou de plusieurs régions géographiques de diffusion communes) et de quadrillages géographiques. Cela entraîne l'augmentation de la cohérence et de l'efficacité des données chez les INS, les ONIG, les organisations internationales et régionales et les autres institutions concernées qui y participent, grâce à la création et la tenue à jour de la couche géographique de référence.

Par conséquent, l'objectif du principe 3 est d'appuyer la fourniture d'un ensemble commun de zones géographiques qui garantissent que les références géospatiales des données statistiques sont compatibles, qu'elles soient sous forme de maillage ou qu'elles utilisent des découpages géographiques administratifs ou statistiques. Il s'agit soit de découpages géographiques attribués uniformément par taille quand il s'agit d'un quadrillage soit de découpages administratifs ou d'unités

²⁶La distribution de ces rôles de parties prenantes entre communautés peut différer selon les pays.

²⁷ Comme l'Union européenne, l'Amérique du Nord et l'Afrique.

²⁸ Comme la Francophonie, le G20?

statistiques (comme des blocs de grille à mailles ou des zones fonctionnelles) divisés selon des subdivisions politiques, immobilières ou topologiques.

Un des principaux éléments de ce principe consiste à permettre l'agrégation précise des données depuis la plus petite zone géographique commune, jusqu'aux régions géographiques de diffusion de niveau supérieur afin de répondre aux besoins du plus vaste éventail possible de décideurs, quels que soient les découpages géographiques ou l'échelle. Cela comprend les agrégations à la fois pour les niveaux géographiques administratifs et statistiques supérieurs (p. ex. circonscriptions électorales, secteurs de recensement et autres zones fonctionnelles)²⁹ et pour la plus petite taille de cellule possible quand on utilise un maillage. Il est également important de noter qu'il existe deux possibilités de diffusion des statistiques géospatiales, à savoir les zones géographiques administratives ou statistiques ou les données maillées (voir Figure 4), qui sont examinées plus en détail dans l'annexe B.

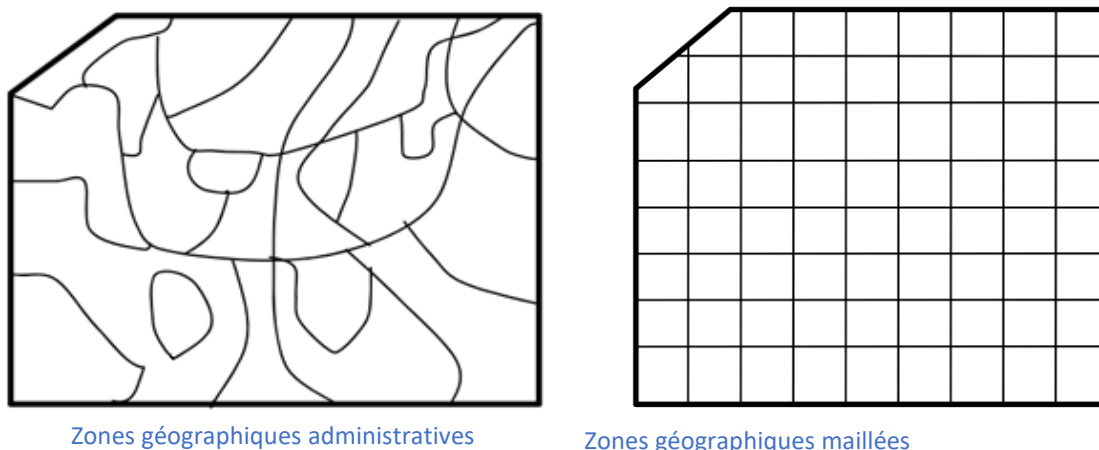


Figure 4 Zones géographiques administratives et maillées

Les INS sont invités à tenir compte des avantages des données maillées en plus des zones géographiques administratives établies par les ONIG et d'autres institutions géographiques. Les données obtenues par quadrillage peuvent être à la fois une source d'informations riche et une zone géographique uniforme aux fins de diffusion et d'intégration des informations³⁰.

Objectifs

La mise en œuvre du principe 3 doit viser les objectifs suivants :

²⁹ Rien n'empêche intrinsèquement une méthode d'enregistrement d'unité par petite région géographique. Toutefois, cette démarche perdrait de son efficacité au fil du temps, à mesure que les parties prenantes demanderont des données statistiques concernant les nouvelles régions géographiques communes et d'autres régions géographiques communes de niveau supérieur aux institutions responsables de l'intégration des données statistiques et de l'infrastructure géospatiale.

³⁰ Les efforts récents de la communauté internationale ont abouti à l'élaboration d'une norme de système de grille discrète mondiale (*Discrete Global Grid System, DGGs*), élaborée sous les auspices de l'OGC. Cette norme offre de nouvelles perspectives en matière d'utilisation de quadrillages dans le contexte du principe des zones géographiques communes et des statistiques géospatiales. Pour en savoir plus sur le DGGs : <http://www.opengeospatial.org/projects/groups/dggsswg>

- veiller à ce que le plus vaste éventail possible de zones géographiques de diffusion communes et de types de zones géographiques juridico-administratives³¹, statistiques³² et plus récemment intégrées³³ soient pris en compte et représenté;
- s'assurer que les régions géographiques de diffusion communes sont évaluées et reconnues conjointement par les parties prenantes intéressées avant leur adoption. Elles doivent être intégrables aux infrastructures géospatiales statistiques existantes et nouvelles. L'information concernant les changements doit être documentée au moyen d'un estampillage temporel et des renseignements sur le cycle de vie, à savoir sur les naissances, les divisions, les fusions et les déplacements de frontière dans ces régions géographiques (comme la création de nouveaux découpages géographiques cadastraux ou politiques). Elles devraient également pouvoir inclure la production de données sociales, économiques et environnementales, d'indicateurs et d'autres renseignements qui soient comparables et intégrés. Cela améliorera la capacité de production de données et d'indicateurs à des fins nationales et répondra aux besoins de suivi et d'établissement de rapports relatifs aux initiatives mondiales (comme le cycle de recensements de la population et des logements de 2020 et les ODD);
- se rappeler de l'importance cruciale à considérer les zones géographiques de diffusion communes proposées et adoptées comme des compléments des zones géographiques administratives et statistiques existantes et à permettre leur intégration;
- fournir des métadonnées et des documents connexes pour toutes les zones géographiques, par exemple, de l'information sur les méthodologies de délimitation pour toute nouvelle zone géographique de diffusion commune ainsi que le cycle de vie de ces zones géographiques et les changements qui leur sont apportés. Cela est essentiel pour faire en sorte que les parties prenantes comprennent et utilisent les concepts potentiellement complexes étayant les méthodes de délimitation employées dans la tenue à jour des zones géographiques de diffusion communes. C'est pourquoi l'adoption continue des normes internationales en vigueur en matière de métadonnées statistiques et géospatiales est un moyen d'atteindre cet objectif (voir le principe 2);
- établir des outils et des méthodes simplifiant l'agrégation géographique des données. Cela implique la mise en œuvre de listes de codes normalisés, ou de tables d'allocation, qui permettent aux outils et applications statistiques d'agréger et d'afficher ou de cartographier de façon uniforme les agrégations géographiques des données;
- veiller à ce que les principes, cadres et pratiques en matière de protection de la vie privée et de qualité des données à l'échelle nationale et internationale, ainsi que leurs modifications, soient pris en compte et respectés dans la conception des zones géographiques de diffusion communes, leur adoption et leur mise en œuvre (par exemple pour les questions de déclaration, de distribution et de visualisation des données et des indicateurs);
- effectuer des examens réguliers et coopératifs de toutes les zones géographiques de diffusion communes adoptées afin de vérifier qu'elles restent pertinentes pour les principaux programmes nationaux, régionaux et mondiaux.

³¹ Les régions géographiques définies dans les lois, les règlements ou la constitution (p. ex. la définition juridique des principales régions géographiques infranationales, comme les États et les provinces).

³² Les zones géographiques définies par un ensemble de règles ou une méthodologie visant à représenter un concept géographique (p. ex. les régions métropolitaines ou les zones fonctionnelles de base, les zones de marché du travail à l'extérieur des régions métropolitaines).

³³ Les zones géographiques conçues pour intégrer efficacement les données sociales, économiques et environnementales (p. ex. la méthode fondée sur le quadrillage). En Europe, l'ensemble de données de carroyage de population GEOSTAT est présenté comme un premier exemple de carroyage de la population de l'Union européenne (UE).

Lien avec d'autres principes

La mise en œuvre des principes 1 et 2 est une condition préalable à la réalisation complète du principe 3 dans un système de production statistique, car leur mise en œuvre permet l'agrégation souple des données dans toute zone statistique. Le principe 3, quant à lui, est une condition importante du principe 5, car les zones géographiques communes constituent le fondement de la diffusion de statistiques géospatiales. Si le principe 2 est entièrement mis en œuvre, l'agrégation de statistiques géocodées est facilement réalisable pour toute zone statistique.

Entrées

Zones géographiques statistiques

- Unités statistiques ou mailles (cellules de grille);
- régions statistiques;
- secteurs de dénombrement de recensement.

Données géospatiales

- découpages géographiques administratifs;
- zones fonctionnelles;
- localisations d'adresses;
- topographie;
- immeubles/logements ou parcelles cadastrales.

Principales parties prenantes

Les principales parties prenantes de ce principe sont celles du principe 2.

Principe 4 : l'interopérabilité des informations statistiques et géospatiales

Données, normes, processus et organismes

Principe 4 : L'interopérabilité statistique et géospatiale (données, normes, processus et organismes) permet de plus normaliser et utiliser les données, et par conséquent d'accroître l'efficacité et la simplification de la création, de la découverte, de l'intégration et de l'utilisation des statistiques géospatiales. Elle augmente aussi la possibilité de mettre en application un plus large éventail de données et de technologies, ce qui permet, d'une part, d'accroître la quantité d'informations disponibles et accessibles aux fins de prise de décision et, d'autre part, de prendre en charge certains aspects de l'amélioration de la coopération entre toutes les parties prenantes produisant et utilisant des informations statistiques et géospatiales.

Pourquoi ce principe est-il nécessaire?

Pour dépasser les obstacles structurels, sémantiques et syntaxiques entre les données et les métadonnées issues de communautés et de fournisseurs différents, il est nécessaire d'augmenter l'interopérabilité entre les normes des données statistiques et géospatiales et celles des métadonnées. Cela améliore également la découverte, l'accès et l'utilisation des données statistiques à référence géospatiale. La plus grande interopérabilité améliore l'adéquation entre données géospatiales et statistiques et leur utilisation harmonisée dans une série d'applications et de systèmes de gestion des données, y compris la modélisation des données et la planification de la production. Il est donc essentiel de convenir précisément des normes et de s'engager à les mettre en œuvre pour tirer les avantages escomptés de l'interopérabilité.

Que recouvre ce principe?

Le principe 4 porte sur l'interopérabilité de toutes les données, métadonnées, normes et bonnes pratiques qui facilitent l'intégration et la production de données statistiques à référence géospatiale. Il comprend les outils et méthodes employés à toutes les étapes du processus de production de la donnée statistique et géospatiale. Il traite également des processus connexes, comme la reproductibilité, la gestion de la qualité et les mécanismes d'interaction entre parties prenantes et utilisateurs. Le principe 4 admet que les communautés statistiques et géospatiales exploitent leurs propres modèles généraux de données, fonctionnalités de métadonnées, architectures et infrastructures de données. Par exemple, la communauté statistique utilise les mécanismes du MGIS (GSIM en anglais), de SDMX et de DDI (Initiative sur la documentation des données ou IDD en français). Parallèlement, la communauté géospatiale utilise couramment GFM (*general feature model* ou modèle général d'entité) et elle a élaboré la norme ISO:19115 sur les métadonnées, ainsi que plusieurs normes et bonnes pratiques propres à certaines applications³⁴ en vue d'appuyer l'interopérabilité des données.

³⁴ Pour lire des analyses de ces modèles statistiques et géospatiaux et des normes de métadonnées, consulter : http://ggim.un.org/meetings/2015-2nd_Mtg_EG-ISGI-Portugal/documents/Connecting%20Geographic%20and%20Statistical%20Information%20Standards%20EG-ISGI%202015.pdf and http://ggim.un.org/meetings/2015-2nd_Mtg_EG-ISGI-Portugal/documents/Metadata%20interoperability%20cover%20paper%20EG-ISGI%202015.pdf

Au sein de la communauté statistique, l'intégration uniforme des processus géospatiaux, des normes et des bonnes pratiques aux processus opérationnels statistiques et aux systèmes de gestion des données est nécessaire, en tant qu'élément fondamental, et ne sert pas seulement à la diffusion des données statistiques. Pour ce faire, les pays sont invités à réfléchir aux moyens d'intégrer plus explicitement les cadres géospatiaux, les normes, les bonnes pratiques et les processus existants à l'architecture commune de production statistique (ACPS ou CSPA en anglais) et ses composantes. Cela permettrait d'accroître l'efficacité et la simplification de la création, de la diffusion, de la découverte et de l'utilisation de statistiques à référence géospatiale, tout en appuyant l'intégration de données géospatiales, statistiques et administratives.

Le principe 4 porte non seulement sur les aspects techniques, mais aussi sur tous les aspects concernant la manière dont les organisations produisent et partagent des informations entre elles et avec les utilisateurs de données. L'intégration des statistiques et de l'information géospatiale est une opération véritablement transversale, qui concerne en particulier différents services au sein des INS et des ONIG, mais qui nécessite également la contribution et la collaboration institutionnelle d'autres organisations et organismes gouvernementaux, ainsi que d'autres producteurs et utilisateurs de ces informations. C'est pourquoi, outre les questions techniques, les modèles et accords de coopération doivent être modifiés et harmonisés de manière à ne pas entraver l'intégration et l'utilisation des données. Le cadre d'interopérabilité européen, illustré à la figure 5 ci-dessous³⁵, a servi à fournir le mécanisme de description des dimensions requises.



Figure 5 Cadre d'interopérabilité européen

1. L'**interopérabilité légale** permet la collaboration d'organismes opérant dans le contexte de différents cadres législatifs, de politiques et de stratégies nationales. Les lois et les politiques nationales ne doivent pas entraver la coopération et des ententes précises sur la façon de traiter

³⁵ https://ec.europa.eu/isa2/eif_en

les différences de législation entre pays doivent être conclues. Par exemple, les lois et les politiques nationales en matière de statistiques devraient comprendre le droit des INS d'avoir accès à des informations géospatiales essentielles de qualité définie et, idéalement, sans frais.

2. L'**interopérabilité organisationnelle** désigne la façon dont les administrations publiques (c.-à-d. les organisations et les organismes gouvernementaux) harmonisent leurs processus opérationnels, leurs responsabilités et leurs attentes pour atteindre les objectifs convenus. En pratique, l'interopérabilité organisationnelle signifie la documentation et l'intégration ou l'harmonisation des processus opérationnels et des informations pertinentes échangées. Il s'agit également de répondre aux exigences de la communauté d'utilisateurs et du SSN.
3. L'**interopérabilité sémantique** garantit que le format et la signification des données et de l'information échangées sont précisément préservés et compris : « ce qui est envoyé est compris ». Cette interopérabilité comprend les aspects syntaxiques, comme la terminologie servant à décrire les concepts, ainsi que la description du format exact de l'information.
4. L'**interopérabilité technique** recouvre les systèmes de couplage et les services d'applications et d'infrastructures. Il s'agit notamment des spécifications des interfaces et des services, ainsi que des normes et formats des données et des métadonnées.

Chacun de ces éléments est crucial pour l'intégration et la production de données statistiques à référence géospatiale et tous sont étroitement interdépendants.

Objectifs du principe

La mise en œuvre du principe 4 doit viser les éléments clés suivants :

Normes et bonnes pratiques

- Permettre aux spécialistes de la communauté statistique et géospatiale de comprendre pleinement la nature, le potentiel et les limites d'autres domaines de données;
- favoriser une communication harmonieuse entre spécialistes au moyen d'une terminologie uniforme et intelligible;
- mettre en place des dispositions permettant de documenter les données, outils, processus et méthodologies dans les langues officielles du pays;
- garantir la préservation et la permanence des données et des outils;
- veiller à ce que seules des normes et des bonnes pratiques ouvertes et internationales soient mises en œuvre, idéalement par l'utilisation ou la modification de normes existantes ou l'établissement de liens entre les normes existantes, en ne créant de nouvelles normes et pratiques qu'en cas de besoin et collectivement;
- assurer la transparence et la visibilité des données et des métadonnées;
- protéger les principes communs de qualité.

Lois et politiques nationales

- Soutenir la coopération entre parties prenantes au moyen d'accords et de lois.

Infrastructure

- Veiller à ce que les données à référence géospatiale ou intégrées puissent circuler librement entre les producteurs de données statistiques et géospatiales, et entre les producteurs de données et les utilisateurs de données, sans que leur circulation soit entravée par des lois et politiques nationales et techniques, des obstacles organisationnels, économiques, linguistiques et conceptuels ou des frontières nationales;
- mettre en place des mécanismes d'accès axés sur le service ou exploitables par une machine (par exemple au moyen d'API) pour accroître l'efficacité de l'accès et de l'utilisation et permettre l'adaptation et l'évolution des usages dans le temps;
- élaborer des solutions communes, afin que les données et les outils puissent être réutilisés et ainsi éviter le dédoublement des efforts, grâce à une infrastructure géospatiale fondamentale nationale unique;
- veiller à ce que tous les utilisateurs, qu'ils soient issus du domaine géospatial ou statistique, puissent découvrir l'information requise et y accéder au moyen d'interfaces techniques et d'interfaces utilisateur définies qui ne nécessitent pas de connaissances intersectorielles;
- s'assurer que les données et les outils soient ouverts et gratuits, dans la mesure du possible, afin que les utilisateurs aient accès à l'ensemble des renseignements sans perte d'information due à des problèmes techniques ou d'interopérabilité.

Lien avec d'autres principes

L'interopérabilité concerne la façon dont les données circulent de la source à l'utilisateur final, par exemple dans l'ensemble du processus de production statistique, y compris la diffusion de données aux utilisateurs intermédiaires et finaux. Ainsi, dans la plupart des cas, les questions d'interopérabilité recourent les autres principes du CSGM et ne sont pas propres à un seul principe. L'interopérabilité est essentielle à la réussite de la mise en œuvre du CSGM.

La mise en œuvre complète de l'interopérabilité décrite dans ce principe est particulièrement importante pour le principe 5, car l'absence d'interopérabilité dans l'un des autres principes se traduit souvent par des renseignements incomplets ou moins utiles pour l'utilisateur final.

Principales parties prenantes

Souvent, aux INS et aux ONIG s'ajoutent des dépositaires de données administratives, qui agissent également comme fournisseurs de données statistiques, mais dont les données ne sont généralement pas interopérables avec les statistiques et les informations géospatiales (les découpages géographiques administratifs; par exemple – voir le principe 3). Les autres parties prenantes sont les principaux organismes de normalisation mondiaux tels que l'ISO, l'OGC et l'OHI ainsi que les organisations dirigeant la modernisation de la statistique officielle, comme la CEE-ONU³⁶.

³⁶ La Commission ne se limite pas aux pays situés dans les limites géographiques de l'Europe et comprend plusieurs États membres non européens : https://www.unece.org/oes/nutshell/member_states_representatives.html

La **Commission européenne** est responsable d'INSPIRE, principal cadre d'établissement de normes visant les informations géospatiales en Europe, **Eurostat** assurant la maintenance du système statistique européen et contribuant aux activités d'établissement de normes, telles que SDMX et ModernStats. L'aperçu régional est appuyé par la Commission économique des Nations Unies pour l'Europe, la **CEE-ONU**, notamment par son soutien à l'initiative « ModernStats » de modernisation de la statistique officielle.

Le Forum européen sur la géographie et la statistique (EFGS) se concentre sur l'élaboration de pratiques exemplaires en matière de production de statistiques à référence géospatiale en Europe. Sous la forme d'un réseau professionnel, il organise la conférence annuelle européenne sur l'intégration des statistiques et des informations géospatiales, renforçant ainsi le partage de connaissances et la communication.

À l'échelle régionale, l'Europe a mené des initiatives et des actions soutenant l'interopérabilité. Parallèlement la Commission statistique de l'ONU, en tant que dépositaire mondial des statistiques et des informations géospatiales et de leur intégration, soutient le renforcement des capacités.

[Actions européennes en faveur de l'interopérabilité](#)

Principe 5 : l'accessibilité et l'exploitabilité des statistiques géospatiales

Principe 5 : l'accessibilité et l'exploitabilité des statistiques géospatiales. Ce principe aide les dépositaires de données à diffuser les données en toute confiance, à améliorer la découverte et l'accessibilité des statistiques à référence géospatiale (en particulier en encourageant les services Web à fournir des couplages dynamiques et exploitables par une machine aux données), et à soutenir l'analyse et l'évaluation des données dans la prise de décision.

Pourquoi ce principe est-il nécessaire?

Les responsables politiques doivent pouvoir accéder aux données obtenues à partir de statistiques à référence géospatiale et les exploiter pour prendre des décisions éclairées et fondées sur des données probantes. C'est pourquoi les dépositaires de données doivent être en mesure de communiquer leurs données en toute confiance, conformément aux lois et aux politiques nationales et régionales, en s'assurant du respect de la protection de la vie privée et des bonnes pratiques pertinentes.

Que recouvre ce principe?

Le principe 5 soutient l'accessibilité, l'analyse, la visualisation et la diffusion des informations statistiques géospatiales nécessaires à la prise de décision éclairée. Pour ce faire, il faut employer, déterminer et élaborer des politiques, des normes et des bonnes pratiques applicables qui augmentent l'accessibilité et l'utilisation de statistiques à référence géospatiale.

Le principe 5 a notamment comme aspect important de veiller à ce que les données soient tenues à jour à intervalles réguliers et qu'elles puissent être consultées en temps opportun au moyen de mécanismes sûrs qui protègent la vie privée et la confidentialité et permettent l'analyse des données aux fins de prise de décision éclairée et fondée sur des données. Cela comprend la fourniture de métadonnées indiquant le moment de la mise à jour, de la publication et de la future mise à jour prévue de chaque information. Parmi les autres questions intéressantes, citons les différents facteurs de la qualité des données (en particulier leur fiabilité, leur actualité et leur pertinence), ainsi que l'analyse, la diffusion et la visualisation des données.

L'accessibilité consiste à fournir les politiques, les normes, les bonnes pratiques et les technologies nécessaires pour que les données géospatiales soient facilement accessibles. Ces politiques, normes, bonnes pratiques et lignes directrices soulignent les nombreuses questions législatives et opérationnelles que les organismes doivent prendre en compte quand ils publient et analysent des informations sur les personnes et les entreprises. La facilité d'utilisation consiste à s'assurer que les résultats sont facilement compris par un large éventail d'utilisateurs aux compétences techniques et non techniques et qu'ils sont exploitables dans les politiques et la prise de décision. En protégeant la vie privée et la confidentialité des données, les dépositaires des données peuvent diffuser des données en toute confiance, et les utilisateurs des données peuvent facilement découvrir et consulter des statistiques à référence géospatiale réutilisables qui peuvent éclairer la visualisation, l'analyse et la prise de décision.

Il est possible d'obtenir des données à la fois accessibles et facilement exploitables en les marquant et les cataloguant de manière adéquate au moyen de métadonnées normalisées appropriées, puis en les diffusant par des services Web ou des formats de fichiers normalisés et ouverts. Cela permet d'accéder à des données lisibles par machine et de prendre en charge des données couplées, ce qui crée ensuite des possibilités de fourniture et d'élaboration de systèmes de connaissances intégrées dans tous les secteurs et dans le cadre de mises en œuvre locales, régionales et mondiales.

Objectifs

La mise en œuvre du principe 5 doit viser les objectifs suivants :

- fournir des données statistiques à référence géospatiale qui peuvent être consultées rapidement et facilement et qui peuvent être utilisées efficacement dans l'analyse de politiques et la prise rapide de décisions fondées sur des données probantes;
- encourager les pays à protéger adéquatement et conformément aux attentes la vie privée et la confidentialité des données d'entrée, ainsi que les personnes et les entreprises qui fournissent ces informations, tout en produisant des statistiques spatio-temporelles informatives de grande qualité;
- organiser la disponibilité de plateformes de diffusion qui prennent en charge la découverte automatisée de données pour faciliter l'échange, l'agrégation et la présentation des données;
- s'assurer que les « modifications » font l'objet d'un suivi et d'une déclaration afin que soient différenciées les modifications purement statistiques (augmentation ou diminution de la population) des modifications géospatiales (vaste annexion ou cession de territoire) et des changements représentationnels, y compris les changements de système de référence géodésique et les modifications au cadastre;
- veiller à ce que les données géospatiales, les statistiques, les données administratives et les autres sources de données soient intégrées, disponibles et accessibles. De préférence, une des méthodes d'accès devrait inclure des services Web, qui permettent un accès lisible par machine, permettant de consulter, de découvrir et de fournir des données dans un format normalisé, digeste et indépendant (multiplateforme);
- assurer la redondance qui doit être intégrée au système, sous la forme de plusieurs plateformes, points d'accès et méthodes de distribution. Des sites Web miroirs, des formats de données de substitution et d'autres méthodes doivent être mis en œuvre;
- rendre possible l'accès à un public large grâce à des plateformes Web de SIG ayant la capacité de générer des visualisations personnalisées, des agrégations et finalement des analyses de données.

Lien avec d'autres principes

Le principe 5 influe sur la mise en œuvre de tous les autres principes, car il met en avant le besoin de données à référence géospatiale intégrées de grande qualité pour la prise de décision fondée sur des données probantes et il est structuré de manière à permettre leur accessibilité et leur utilisation. Les données de sortie à référence géospatiale devraient être conçues pour répondre à des exigences précises en matière de rapports telles que les ODD, le cadre du Système de comptabilité économique et environnementale (SCEE) ou les priorités de développement nationales. Par conséquent, les données d'entrée des principes précédents doivent être conçues selon ces exigences.

Les principes 1 et 2 sont importants en raison de leur effet sur le degré d'actualité et la qualité (pertinence, exhaustivité, définitions, classifications, topologie, etc.) des données mises à la disposition des utilisateurs. En retour, les décisions relatives à la collecte et à la tenue à jour des données au moment de l'application du principe 1 devraient tenir compte de la facilité d'utilisation et des applications définies dans le principe 5.

Le principe 3 lui est lié en raison du besoin évident de zones statistiques communes, qui constituent la base de l'analyse et des comparaisons, et le principe 4 est étroitement lié au principe 5 par les normes techniques qui favorisent l'interopérabilité des données et de l'infrastructure technique. L'importance de l'interopérabilité dans l'infrastructure technique s'étend aux interfaces utilisateur frontales et aux bases de données dorsales, aux logiciels et aux composants matériels.

Il est important de souligner que les objectifs généraux du principe 5 peuvent et doivent être visés dans toute la mesure du possible, y compris quand il est impossible de mettre en œuvre l'ensemble du CSGM. Ainsi, les données à référence géospatiale seront fournies de la façon la plus accessible possible.

Entrées

Les entrées de ce principe sont en fait les données, les éléments et les processus définis dans les autres principes. Il s'agit en effet :

- de données à référence géospatiale précises et actualisées;
- de métadonnées normalisées et de données adéquatement marquées;
- de la participation, des connaissances et des compétences des secteurs géospatial et statistique, qui sont importantes;
- d'une infrastructure technique en mesure de donner accès aux ensembles de données publiés, y compris dans des formulaires lisibles par machine;
- de normes techniques à jour pour les services Web et l'accessibilité.

Principales parties prenantes

Les communautés statistiques et géospatiales joueront un rôle important dans la réussite de l'application du principe 5. La nature exacte de ce rôle, accompagnée de cas d'utilisation détaillés aux fins de sa mise en œuvre, devrait être déterminée par des ententes de collaboration institutionnelle, convenues mutuellement. Le principe 5 représente l'interface entre l'utilisation publique et les données recueillies et organisées dans le cadre des processus du CSGM. Les parties prenantes extérieures et les INS devraient jouer un rôle important dans l'élaboration des meilleures pratiques empêchant la divulgation de données confidentielles et une utilisation de données susceptible de permettre d'identifier personnellement les répondants ou d'accéder à des renseignements sensibles. Ils devraient aussi assumer un rôle de conseil en la matière.

Partie 3 : Annexes



Annexe A : Définitions des termes

Informations géographiques communes

Les informations géographiques communes se composent d'un ensemble convenu de zones géographiques employées dans l'affichage, le stockage, la déclaration et l'analyse de comparaisons sociales, économiques et environnementales entre des ensembles de données statistiques provenant de différentes sources. Les informations géographiques communes permettent de produire et diffuser des statistiques intégrées et des informations géospatiales dans un pays ou une région afin d'appuyer la prise de décision éclairée.

Environnement de gestion des données

Un environnement de gestion des données englobe la totalité des outils et de l'infrastructure d'acquisition, de validation, de stockage, de protection et de traitement des données nécessaires pour garantir aux utilisateurs l'accessibilité, la fiabilité et l'actualité des données.

Système de grille discrète mondiale (DGGS)

Les DGGS représentent la Terre comme une hiérarchie de cellules de surface égales avec une résolution géospatiale de plus en plus fine. Des observations individuelles peuvent être attribuées à une cellule correspondant à la fois à la position et à la taille (ou à l'incertitude) du phénomène observé. Les DGGS sont très intéressants dans l'intégration et l'analyse de données géospatiales.

Infrastructure géospatiale fondamentale

Une infrastructure géospatiale fondamentale englobe l'infrastructure nationale de données géospatiales (INDG) (voir ci-dessous), les normes, les technologies, les politiques et les bonnes pratiques, ainsi que d'autres éléments clés permettant la fourniture d'informations géospatiales dans un pays. Le Cadre intégré d'information géospatiale³⁷ approuvé par l'initiative des Nations Unies sur la gestion de l'information géospatiale à l'échelle mondiale (UN-GGIM) peut fournir cette infrastructure, faute d'infrastructure nationale.

Géocodage

Aux fins du CSGM, le géocodage est généralement défini comme le processus d'ajout d'une référence géospatiale aux enregistrements d'unités statistiques afin qu'ils soient exploitables dans des analyses géospatiales. Plus précisément, le géocodage est le processus de couplage d'informations de localisation non référencées (p. ex. une adresse), qui est associé à une unité statistique, à un géocode (c.-à-d. un objet à référence géospatiale). Le géocode peut aussi être directement intégré à l'enregistrement d'unité statistique. Les géocodes sont, de préférence, des objets à référence géospatiale à petite échelle stockés en tant que type de donnée géométrique, comme des coordonnées de localisation (c.-à-d. coordonnées x, y, z), ou de petites zones géographiques (p. ex. blocs de grille à mailles, côtés d'îlots ou zones géographiques similaires de petite unité de base). Les unités géographiques plus grandes, comme les régions géographiques de dénombrement, peuvent servir de géocodes à défaut d'unités géospatiales à plus petite échelle. Le couplage d'un géocode à un

³⁷ <http://ggim.un.org/meetings/GGIM-committee/8th-Session/documents/Part%201-IGIF-Overarching-Strategic-Framework-24July2018.pdf>

enregistrement d'unité statistique se fait au moyen de systèmes de codage géographique normalisés, d'un identifiant URI ou d'autres mécanismes de couplage informatisés.

Entité géographique

Représentation géométrique d'une entité géographique. Il peut s'agir d'une caractéristique physique comme un enregistrement d'unité, un logement, un bien ou une zone fonctionnelle comme un découpage géographique administratif ou une zone économique.

Géoréférencement

Le géoréférencement est un ensemble de processus généraux qui comprennent le géocodage. Le processus consiste à référencer des données en fonction d'un système de coordonnées géographiques connu, par une correspondance à des points de référence connus dans le système de coordonnées (par ex. rectification d'image par rapport à des points de relevé ou couplage d'adresses à des centroïdes de parcelle), de sorte que les données puissent être affichées, traitées, interrogées et analysées avec d'autres données géographiques.

Statistiques à référence géospatiale

La localisation ou l'étendue géospatiale sont les principales caractéristiques des statistiques à référence géospatiale. De plus, il est recommandé que toutes les données d'enregistrement d'unité statistique soient collectées ou associées à une référence de localisation et qu'idéalement, cela permette de produire des coordonnées géospatiales avec des valeurs d'abscisse et d'ordonnée pour chaque enregistrement.

Interopérabilité

L'interopérabilité est la capacité d'un système d'échanger et d'utiliser de l'information par l'application de normes ouvertes.

Informations de localisation

Les informations de localisation peuvent comprendre des adresses, des identificateurs de bien ou d'immeuble, ainsi que d'autres descriptions d'emplacement, comme les régions géographiques de dénombrement et d'autres descriptions textuelles de lieu, normalisées et non normalisées (p. ex. noms de village).

Infrastructure nationale de données géospatiales

L'infrastructure nationale de données géospatiales fournit la technologie, les politiques, les normes, les bonnes pratiques et les ressources humaines nécessaires pour acquérir, traiter, stocker, distribuer et améliorer les données géospatiales. Toute mise en œuvre réussie d'INDG tient compte des facteurs suivants :

- la tenue à jour des données et la maintenance des systèmes;
- l'intégration des redondances à la solution de diffusion afin d'éviter les points de panne uniques;
- l'examen final et le prétraitement avant diffusion (divulcation des données et confidentialité) pour prévenir les problèmes de divulgation;
- la nécessité de mettre en œuvre la généralisation et l'élimination sélective des données géospatiales afin qu'elles correspondent au niveau minimal de qualité et qu'elles soient

exploitables à des échelles définies répondant à la fois aux besoins à grande et à petite échelle. Cela peut avoir des répercussions sur les questions de cartographie et de stockage des données.

Enregistrements d'unité statistique

Les enregistrements d'unités statistiques peuvent comprendre des personnes, des ménages et des logements, des entreprises, des bâtiments ou des parcelles et unités territoriales.

Annexe B : Normes, qualité et cadres facilitateurs

Tout comme la coopération est fondamentale pour la mise en œuvre du CSGM, la nécessité d'adopter et de mettre en œuvre les normes et les bonnes pratiques statistiques et géospatiales et celle d'évaluer la qualité des données d'entrée et des résultats des données utilisés dans l'ensemble du CSGM sont cruciales.

Il est important de reconnaître que les normes des communautés statistiques et géospatiales sont nombreuses et que, bien que les renseignements ci-dessous appuient et orientent la mise en œuvre du CSGM, il ne s'agit aucunement d'une liste exhaustive. Les INS et les ONIG sont vivement invités à connaître les principes de qualité et les mesures de normalisation qui ont cours dans les communautés statistiques et géospatiales, ainsi que d'autres communautés connexes, compte tenu notamment des travaux d'élaboration de normes de grande qualité en cours. Les pays sont encouragés à contribuer à ces efforts et à considérer le CSGM comme faisant partie d'un large ensemble d'actions qui vise la modernisation de la production de statistique et d'informations géospatiales, et à tenir compte des progrès généraux de la technologie et de l'information.

Les sections suivantes présentent brièvement les principaux efforts de normalisation pertinents liés à l'information statistique et géospatiale, ainsi que les principaux principes de qualité des deux communautés. La section ne cherche pas à présenter la liste exhaustive des normes applicables, mais à sensibiliser au fait que l'intégration des données nécessite de comprendre l'utilisation des normes et le contexte général, d'autant plus que l'élaboration de normes peut être très dynamique.

Normes

Les normes sont essentielles pour faciliter l'élaboration, le partage et l'utilisation de données et de services statistiques et géospatiaux, mais il est important de savoir que les données et les normes statistiques et géospatiales ont des critères et des modalités d'élaboration différents. Le CSGM préconise l'utilisation de normes adoptées internationalement par les deux communautés pour accroître l'interopérabilité des données statistiques et géospatiales, des normes et des processus.

Dans le domaine statistique :

- le GHN-MSO a élaboré et soutient le Modèle générique d'activité des organismes statistiques (GAMSO)³⁸, le MSGPO³⁹, le MGIS⁴⁰ et l'architecture commune de production statistique (CSPA)⁴¹;
- la ligne directrice SDMX ISO/TS 17369⁴², SDMX 2.1⁴³;
- les « Principes fondamentaux de la statistique officielle nationale » approuvés en 1994 par l'Assemblée générale des Nations Unies et accompagnés d'une résolution (renouvelée en 2014) résumant les Principes fondamentaux pour permettre la production d'informations

³⁸ <https://statswiki.unece.org/display/GAMSO>

³⁹ <https://statswiki.unece.org/display/GSBPM>

⁴⁰ <https://statswiki.unece.org/display/gsim>

⁴¹ <https://statswiki.unece.org/display/CSPA>

⁴² <https://www.iso.org/standard/52500.html>

⁴³ https://sdmx.org/?page_id=5008

statistiques officielles de haute qualité à des fins d'analyse et de prise de décision éclairée en matière de politiques;

- le guide sur l'interopérabilité des données⁴⁴ élaboré par la DS-ONU, qui appuie les efforts d'interopérabilité dans le secteur du développement.

Dans le domaine géospatial :

- des organismes internationaux de normalisation comme l'OGC, l'OHI et l'ISO travaillent à la définition de normes géospatiales pertinentes et à la transformation de normes; ils ont publié le « Guide sur le rôle des normes dans la gestion de l'information géospatiale »⁴⁵;
- la série ISO 19000 (ISO/TC 211) sert à décrire les services et les informations géographiques. Les normes ISO 19111 « Système de références par coordonnées » ou ISO 19112 « Système de références spatiales par identificateurs géographiques » et ISO 19115 « Information géographique – Métadonnées » sont généralement utiles. La norme ISO 19115-2 est une extension de cette dernière pour l'imagerie et les matrices. La norme ISO 19157, quant à elle, traite des principes de qualité des données géographiques. De plus, la norme ISO 19160 définit un modèle conceptuel pour les informations d'adresse;
- l'OGC propose des normes sur les services normalisés aux fins de publication et d'exécution des processus informatiques géospatiaux sur le Web, comme WFS (service d'entités géographiques Web) et WMS (service de cartographie Web).
- Quand on s'intéresse aux données géospatiales, il faut aussi tenir compte du repère de référence géodésique mondial (p. ex. ETRS89). Le DGGS fournit également une norme de définition de maillage aux niveaux local et mondial, ce qui est particulièrement utile dans les comparaisons de pays.

Bien qu'il existe un riche corpus de normes statistiques et géospatiales, peu de normes traitent de l'interopérabilité des informations statistiques et géospatiales. C'est pourquoi il est plus que recommandé de réfléchir à une méthode qui surmonterait les difficultés d'interopérabilité découlant des différences de modèles de données, d'interfaces ou même de formats de fichiers. Par exemple, le standard Table Joining Service (TJS, Service de mise en correspondance des données tabulaires) défini par l'OGC peut combler les lacunes en matière d'échange de données tabulaires codées au format SDMX pour plusieurs zones géographiques.

Les domaines prioritaires suivants, tirés des principes du CSGM, nécessitent des mesures de normalisation supplémentaires :

- une classification géographique commune, y compris l'utilisation de zones géographiques administratives et statistiques complétées par des zones géographiques en quadrillage (principe 3);
- des normes ou des lignes directrices sur l'utilisation des zones géographiques aux fins de diffusion des données (principe 3);
- des normes internationales en matière de métadonnées statistiques et géospatiales (principe 4);

⁴⁴ <https://unstats.un.org/wiki/display/InteropGuide/Introduction>

⁴⁵ http://ggim.un.org/meetings/GGIM-committee/8th-Session/documents/Standards_Guide_2018.pdf

- des systèmes et des méthodes permettant de faire correspondre les données entre zones géographiques⁴⁶; (principes 2 et 3);
- des lois, des politiques ou des principes de protection de la vie privée au niveau du pays, conformes aux Principes fondamentaux de la statistique officielle des Nations Unies (principes 3 et 5).

Toutefois, le CSGM demande explicitement de n'élaborer de nouvelles normes qu'en cas de nécessité absolue. Il est préférable que les communautés œuvrent à l'établissement de ponts entre les normes ou à l'extension des normes existantes. Ainsi, Eurostat a étendu SDMX à la diffusion conforme à INSPIRE des données du recensement de 2021.

Zones géographiques administratives et maillées

Historiquement, les découpages des zones administratives ont été perçus par certains comme un système plus fiable pour les enquêtes sur les populations, les logements et les zones géographiques parce qu'elles sont conformes aux caractéristiques physiques constatables sur le terrain. Ces zones correspondent aux limites physiques (c.-à-d. routes, chemins de fer et autres entités visibles), aux entités du paysage (c.-à-d. rivières, lacs et autres bassins hydrographiques), ainsi qu'aux modèles de développement urbain. La vérification géospatiale au sol est peut-être l'argument le plus solide pour favoriser les zones administratives plutôt que les zones maillées. Cependant, s'y oppose un contre-argument fort valide selon lequel le fait de se fier à des limites physiques susceptibles de changer est précisément la raison pour laquelle on ne doit pas préférer les zones administratives aux zones maillées. Le tableau 1 présente un aperçu et une comparaison des aspects positifs et négatifs des zones géographiques administratives et maillées⁴⁷.

Qu'on emploie des zones géographiques administratives ou maillées, il est important de tenir compte du biais potentiel introduit par des unités statistiques ayant des formes et des tailles différentes. Ce problème est connu sous le nom de problème d'unité spatiale modifiable (Modifiable Areal Unit Problem – MAUP)⁴⁸, c'est-à-dire que l'agrégation des unités peut être affectée par la façon dont les limites sont tracées en segment et ensuite agrégées.

Zones géographiques administratives

L'avantage des découpages géographiques administratifs ou statistiques établis employés comme informations géographiques communes est qu'ils se fondent habituellement sur les tendances historiques des peuplements humains. Dans la plupart des cas, les êtres humains comprennent bien les découpages géographiques administratifs. L'exactitude géospatiale des zones délimitées et les données elles-mêmes peuvent être vérifiées sur le terrain grâce à une vérification sur le terrain, à l'imagerie par satellite, au géocodage et à la prospection d'adresses. Souvent, les sources faisant autorité pour ces données proviennent des INS, des ONIG et d'autres organismes nationaux. Parce que les zones géographies statistiques sont conçues pour la diffusion de statistiques, elles peuvent

⁴⁶ <http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Correspondences>

⁴⁷ Il est tiré de l'exposé de Tim Trainor à l'atelier international sur l'intégration de l'information géospatiale et statistique, qui a eu lieu à Beijing (Chine) du 9 au 12 juin 2014 http://ggim.un.org/meetings/2014-IGSI-Beijing/documents/04_USA_UN_Grid_Admin_Trainor_6_5_14.pdf

⁴⁸ Gehlke, Charles E. et Katherine Biehl. « Certain effects of grouping upon the size of the correlation coefficient in census tract material » (Certains effets du regroupement sur la taille du coefficient de corrélation dans le matériau des secteurs de recensement). *Journal of the American Statistical Association*, 29.185A (1934): 169-170.

surmonter bon nombre des problèmes associés aux découpages géographiques administratifs. Elles servent généralement à délimiter des zones dont la population est de taille identique ou similaire. Cela est toutefois impossible dans de nombreux cas, c'est pourquoi il faut normaliser les données avant toute comparaison entre zones. Pour obtenir des zones à la taille de population uniforme, il faut parfois délimiter des zones statistiques en milieu rural très étendues et représentant un mélange disproportionné d'utilisations des terres et de zones de population.

Les inconvénients de l'utilisation de zones administratives comme zones géographiques communes découlent des changements et de la variabilité causés par la croissance démographique, l'expansion des profils de peuplement urbain et des incohérences entre les autorités juridiques et les gouvernements, qui peuvent avoir une incidence sur la comparabilité au sein d'un même pays et entre pays. La difficulté se fait vivement sentir dans les pays en développement, caractérisés par une forte croissance démographique, de rapides changements au niveau urbain et rural, entre autres nombreux facteurs démographiques. Il est par conséquent difficile de comparer les zones géographiques de recensement qui sont fondées sur les densités de population et les profils de peuplement susceptibles de changer entre les années de recensement. Les découpages géographiques invisibles sont subjectifs et peuvent aussi changer, en fonction des systèmes locaux de gestion des terres et de la législation. Par exemple, en cas de construction de route ou de dérivation de rivière, une limite géographique invisible peut être modifiée et correspondre par conséquent à une entité visible.

En outre, la difficulté de comparaison est accrue par les différences encore plus grandes entre pays quant à la méthode de détermination des zones administratives ainsi qu'au nombre et au type de zones géographiques différentes utilisées pour la diffusion de données. Les îlots de recensement peuvent être la plus petite région d'un pays, mais être des districts ou des provinces dans un autre. La généralisation cartographique diffère également d'un pays à l'autre. Les différences de systèmes de classification, d'unités de mesure, de tailles et de formes des découpages géographiques administratifs rendent difficile la création d'un modèle statistique ou d'une mesure uniforme de la densité de la population et de l'utilisation des terres.

Les définitions régionales fondées sur l'économie, la culture, la couverture terrestre ou l'utilisation des terres, ainsi que la topographie physique des zones administratives ajoutent un degré de variabilité au sein des pays, entre pays et entre domaines thématiques. À l'intérieur même des pays, les taux de participation ou de réponse peuvent varier en raison de biais régionaux, ou de l'isolement physique des répondants dans des zones administratives données. À l'échelle mondiale, la principale préoccupation pourrait être l'accessibilité des zones administratives dans les régions éloignées, surtout compte tenu de la mission globale du Programme 2030 qui consiste à « ne laisser personne de côté ».

Zones géographiques maillées

Le principal avantage présenté par l'utilisation de maillages comme zones géographiques communes est leur comparabilité. En effet, ils sont stables et non affectés par les changements fréquents des découpages administratifs (par exemple, lors des redécoupages de zones administratives destinés à

mieux refléter la population d'une zone géographique)⁴⁹. Étant donné que la plupart des unités statistiques à l'échelle des pays utilisent des cadres de localisation par points – c.-à-d. des registres d'adresses et d'immeubles – et que l'utilisation de coordonnées exprimées en abscisse et en ordonnée est recommandée pour chaque enregistrement d'unité statistique, des algorithmes d'agrégation et de désagrégation de données peuvent être appliqués à tout type de zone, y compris aux mailles. On peut ensuite ajouter les mailles pour former des zones dans un but précis ou pour une zone d'étude, comme les régions urbaines, rurales, montagneuses ou côtières, etc. La taille est un facteur très important dans un maillage et la taille de chaque cellule doit être identique. Il est aussi possible d'établir un système hiérarchique qui soit proportionnel à la taille de la zone d'étude, de globale à locale. La plus petite unité de base est une microdonnée géoréférencée par points. La dimension des mailles peut être réalisée de façon à éviter l'« effet de dilution » causé par une unité de base trop grande.

Les zones géographiques maillées sont également en train de devenir une norme mondiale qui permettra des comparaisons mondiales de données à l'échelle des pays. Les quadrillages, en particulier les maillages globaux comme le système DGGs, peuvent servir d'élément d'infrastructure fournissant à la fois un système de référencement de localisation et un conteneur de données.

Les inconvénients de l'utilisation de zones statistiques maillées sont principalement liés à la taille des mailles, à la projection et aux systèmes de codage. La nature arbitraire de la définition des limites de maillage et leur capacité à couper des entités homogènes qui ne peuvent être subdivisées davantage (comme les grands immeubles) peuvent poser des problèmes pratiques de répartition et d'utilisation des données. Étant donné que les données maillées peuvent représenter une zone statistique de petite superficie, les responsables de la mise en œuvre doivent tenir rigoureusement compte de la protection de la vie privée afin de prévenir la divulgation de renseignements identificatoires. Plus la population visée est petite, plus ce risque est élevé. La suppression de données de ces zones peut éliminer des données essentielles nécessaires aux allocations et aux aides gouvernementales. La qualité des données des quadrillages dépend beaucoup de l'exactitude des données d'entrée géocodées, dans lesquelles les erreurs sont difficiles à trouver en cas de statistiques fondées sur un maillage.

⁴⁹ Plus de 20 pays et institutions ont amorcé une transition vers un système statistique basé sur des quadrillages aux fins de production de rapports statistiques, notamment l'Autriche, la Belgique, Chypre, le Danemark, l'Estonie, la Finlande, la France, l'Irlande, les Pays-Bas, la Norvège, la Pologne, la République du Kosovo, la Serbie, la République slovaque, la Slovénie, la Suède et d'autres pays.

	Éléments positifs	Éléments négatifs
Zones administratives	<ul style="list-style-type: none"> • Exactitude spatiale des données • Vérification sur le terrain • Vérification de l'imagerie • Géocodage/vérification d'adresses • Sources faisant autorité • Participation des autorités locales • Connaissances locales • Relations d'imbrication avec d'autres zones géographiques • Découpages cadastraux • Seuils des données • Distinction entre zones terrestres et aquatiques • Taux de réponse • Options de réponse • Qualité des réponses • Base de sondage • Contrôles de la divulgation • Exploitable dans une analyse utilisant la technologie SIG • Définition des lieux et des collectivités • Normalisable par la population 	<ul style="list-style-type: none"> • Comparabilité • Modifications de découpage géographique • Diminution de la fréquence de la collecte classique de données du recensement dans les pays • Utilisation de découpages géographiques invisibles • Nombre de zones géographiques différentes • Considérations cartographiques/généralisation des découpages géographiques • Participation variable • Tailles irrégulières • Formes irrégulières • Mesures de la densité variables • Coûts élevés de la tenue à jour des données • Différences légales • Différences régionales • Différences topographiques • Compréhension insuffisante des microcaractéristiques à l'intérieur des unités à grande échelle • Intégration des données difficile • Possibilité de différences dans les zones géographiques administratives d'un pays • Souvent mélange d'utilisations des terres et d'unités géographiques • Inadaptées aux fins d'analyse de séries chronologiques • Inadéquates pour de nombreuses techniques d'analyse géospatiale • Non esthétiques pour la cartographie
Maillages	<ul style="list-style-type: none"> • Portée mondiale et locale – entièrement évolutifs • Échelle uniforme propice aux études transfrontalières • Comparabilité; plus adaptés aux INDG • Plus grande attention portée à la science orientée sur les problèmes • Possibilité de localisation plus exacte des populations • Cadre territorial propice aux échantillonnages • Possibilité d'agrégation à différents types d'unités territoriales • Exploitable dans une analyse utilisant la technologie SIG • Facilement générés à partir de données de points géoréférencés • Possibilité de voir des grappes • Facilement et rentablement produits • Analyse à micro-échelle au moyen de maillage à taille flexible • Possibilité d'intégration des données à de nouvelles sources de données (c.-à-d. au sol, imagerie, Internet) • Stables dans le temps; séries chronologiques non touchées par les changements d'unité administrative • Indépendants des procédures classiques de collecte de données • Très utilisés en science et en pratique • Peuvent comprendre des résolutions plus élevées dans les zones rurales • Peuvent devenir une norme mondiale 	<ul style="list-style-type: none"> • Arbitraires • Peuvent diviser des entités significatives, notamment des entités et des immeubles non fractionnables • Contrôle de la divulgation/taille des cellules/mailles • Taille des cellules/mailles dans les régions rurales • Nécessite que le géocodage des unités soit très précis • Lors de la fusion d'ensembles de données, nécessité de passer d'un système de coordonnées à un autre avant la compilation de données dans des quadrillages • Le système européen de référence terrestre (ETRS80) est basé sur le système de référence de coordonnées de la projection azimutale équivalente de Lambert ayant un centre de projection fixe; or, différentes projections pourraient être nécessaires dans d'autres régions du monde. • Systèmes de codage [intervalles d'échelle contre solutions à arbre de quadrants] • En raison du volume élevé de données, difficulté à trouver et corriger les erreurs • Des maillages différents peuvent être adoptés au sein de régions ou de pays • Les régions présentant des fluctuations dynamiques ou transitoires de la population compliquent les analyses régionales de plusieurs manières. • Nécessité d'employer des modèles de validation croisée spatiale et temporelle utilisant plusieurs sources de données géographiques, physiographiques et socioéconomiques en conjonction avec l'analyse d'images au besoin. • Taille de cellule différente. Nécessité de tenir compte des conditions de taille de cellule au moment de la collecte de données et de la taille de cellule au moment de la diffusion des données • Possibilité de résolution plus faible en milieu urbain

Tableau 1 Comparaison des avantages et des inconvénients des zones géographiques administratives et maillées

Thèmes des données géospatiales fondamentales mondiales

Thème	Aperçu
 <p>Cadre de référence géodésique Global</p>	<p>Le cadre de référence géodésique global est le référentiel permettant de déterminer et d'exprimer avec exactitude des localisations sur la Terre, ainsi que de quantifier les changements de la Terre dans l'espace et le temps. Il s'agit d'une condition préalable essentielle à la collecte, l'intégration et l'utilisation exactes de toutes les autres données géospatiales.</p>
 <p>Adresses</p>	<p>Les adresses sont une étiquette structurée, contenant généralement un numéro de bien, un nom de rue et un nom de localité. Les adresses servent souvent de substituts à d'autres thèmes de données, comme les parcelles cadastrales, et peuvent habituellement être liées à des coordonnées géographiques.</p>
 <p>Immeubles et habitats</p>	<p>Un immeuble s'entend de toute structure couverte construite ou érigée en permanence sur son site, à des fins de protection d'êtres humains, d'animaux, d'objets ou de production de biens économiques. Les habitats sont des ensembles d'immeubles et d'entités connexes où une collectivité mène des activités socioéconomiques.</p>
 <p>Élévation et profondeur</p>	<p>Le thème de l'élévation et de la profondeur décrit la croûte terrestre à la surface de la terre et sous un plan d'eau par rapport à un point de référence vertical.</p>
 <p>Zones fonctionnelles</p>	<p>Les zones fonctionnelles sont l'étendue géographique des zones administratives, législatives, réglementaires, électorales, statistiques, de gouvernance, de prestation de services et de gestion des activités.</p>
 <p>Noms géographiques</p>	<p>Les noms géographiques désignent l'orientation et l'identité des lieux. Ils servent à identifier les lieux correspondant aux entités culturelles et physiques du monde réel, comme les régions, les peuplements ou les autres entités d'intérêt public ou historique.</p>
 <p>Géologie et sols</p>	<p>La géologie désigne la composition et les propriétés des matériaux géologiques (roches et sédiments) souterrains et des affleurements à la surface de la Terre. Le sol est la partie supérieure de la croûte terrestre, formée par des particules minérales, de la matière organique, de l'eau, de l'air et des organismes vivants.</p>
 <p>Occupation des sols et utilisation des terres</p>	<p>L'occupation des sols représente la couverture physique et biologique de la surface de la Terre. L'utilisation des terres désigne la gestion planifiée actuelle et future et la modification de l'environnement naturel à différentes fins humaines ou économiques.</p>
 <p>Parcelles de terre</p>	<p>Les parcelles de terre sont des terres, des parcelles cadastrales ou des zones de la surface de la Terre (terrestres ou aquatiques) sous un régime commun de droits (comme la propriété ou les servitudes), de revendications (comme les minéraux ou les terres indigènes) ou d'utilisation.</p>
 <p>Ortho-imagerie</p>	<p>L'ortho-imagerie présente des données d'images rectifiées géoréférencées de la surface de la Terre, prises à partir de capteurs satellitaires ou de détecteurs aériens.</p>
 <p>Infrastructure physique</p>	<p>Le thème de l'infrastructure physique comprend les installations industrielles et de services publics, ainsi que les installations de prestation de services associées aux services administratifs et sociaux gouvernementaux comme les administrations publiques, les services publics, les écoles, les hôpitaux, etc.</p>
 <p>Répartition de la population</p>	<p>Le thème de la répartition de la population recouvre les données relatives à la répartition géospatiale de la population et ses caractéristiques, ainsi que l'incidence de la population sur l'urbanisation, le développement régional ou la durabilité.</p>
 <p>Eau</p>	<p>Le thème de l'eau recouvre l'étendue géographique et les conditions de toutes les entités aquatiques, y compris les rivières, les lacs et les entités marines.</p>
 <p>Réseaux de transport</p>	<p>Le thème des réseaux de transport recouvre l'ensemble des voies de communication routières, ferroviaires, aériennes et maritimes et les liens entre elles.</p>

Figure 6 Les 14 thèmes des données géospatiales fondamentales mondiales

La Figure 6 énumère et décrit les 14 thèmes des données géospatiales fondamentales mondiales⁵⁰. La mise en œuvre des thèmes nécessite l'intégration des informations provenant des ONIG, des INS et d'autres dépositaires de données.

Cadres statistiques et géospatiaux de qualité

Selon les INS, la grande qualité des statistiques officielles représente un « ... *avantage concurrentiel dans un monde qui connaît une tendance croissante à l'information instantanée, dont la qualité est rarement prouvée comme il le faut* ». ⁵¹

C'est pourquoi les statistiques officielles ont élaboré un cadre complet de gestion de la qualité⁵².

L'Europe a par exemple mis au point un cadre de qualité global internationalement accepté, au sommet duquel se trouve le Code de bonnes pratiques de la statistique européenne⁵³. En plus de l'environnement institutionnel des statistiques officielles et de la qualité des processus statistiques, il traite de la qualité des produits statistiques.

La qualité d'un produit statistique peut être définie comme l'adaptation du produit aux besoins et elle se compose des qualités suivantes :

- pertinence;
- exactitude et fiabilité;
- actualité et ponctualité;
- accessibilité et clarté;
- cohérence et comparabilité.

L'intégration des informations statistiques et géospatiales peut contribuer à toutes ces dimensions, mais en particulier à la pertinence des produits statistiques pour l'utilisateur. Elle doit aussi tenir compte de toutes les étapes du processus statistique, comme la détection et l'analyse des besoins, la conception et la mise à l'essai, l'exécution, l'analyse et la diffusion⁵⁴. Les systèmes de qualité dans les statistiques officielles sont habituellement structurés selon deux niveaux, l'un comportant les principes et les objectifs et l'autre les indicateurs mesurant la réalisation des objectifs.

À titre d'exemple, la mise en œuvre du principe d'**exactitude et de fiabilité** visant l'objectif de « statistiques décrivant avec précision et fiabilité la réalité » peut être mesurée à l'aide des indicateurs suivants : *les données de base, les résultats intermédiaires et les produits statistiques qui font l'objet d'évaluations et de validations à intervalles réguliers*. Par exemple, les mesures de la qualité de données de registre peuvent comprendre les éléments suivants :

⁵⁰ <https://undesamaps.arcgis.com/apps/Cascade/index.html?appid=4741ad51ff7a463d833d18cbcec29fff>

⁵¹ Déclaration de qualité du système statistique européen <https://ec.europa.eu/eurostat/documents/4031688/8188985/KS-02-17-428-EN-N.pdf/116f7c85-cd3e-4bff-b695-4a8e71385fd4>

⁵² L'ONU a élaboré les principes fondamentaux de la statistique officielle <https://unstats.un.org/unsd/dnss/gp/FP-New-F.pdf> et les lignes directrices de l'ONU pour la qualité des statistiques. <https://unstats.un.org/unsd/dnss/qualitynqaf/nqaf.aspx>

⁵³ <https://ec.europa.eu/eurostat/web/quality/overview>

⁵⁴ Mise en œuvre en Colombie (en espagnol) : https://www.dane.gov.co/files/sen/normatividad/NTC_Proceso_Estadistico.pdf

Critère	Indicateur	
	Nom	Valeur
Exactitude	Surdénombrement – unités hors de la population	Pourcentage d'unités n'appartenant pas à la population
	Sous-dénombrement – unités de population manquantes	Pourcentage d'unités de population manquantes
	Répétition injustifiée d'enregistrements concernant une même unité de population	Pourcentage d'enregistrements répétés
	Données manquantes pour des variables	Pourcentage de caractéristiques d'information pour lesquelles il manque des valeurs
		Pourcentage d'unités pour lesquelles il manque des valeurs de caractéristiques d'information particulières
	Ajustement, imputation	Pourcentage d'unités ajustées
		Pourcentage de valeurs ajustées
		Pourcentage de valeurs complétées
	Intégration de données de diverses sources	Appariement exact – pourcentage d'unités appariées
		Erreurs d'intégration – pourcentage d'unités incorrectement appariées
		Absence d'appariement – pourcentage d'unités non appariées

Tableau 2 Exactitude du principe de la qualité appliqué aux données du registre

À cet égard, leur structure est semblable à celle du cadre statistique géospatial, qui devrait faciliter la communication entre les statisticiens et les décideurs de façon générale. Par ailleurs, les statisticiens connaissent généralement moins les concepts de qualité relatifs aux informations géospatiales relatives à la production de statistiques à référence géospatiale, qui sont également abordés dans ce cadre, en particulier les principes 1 et 3.

C'est pourquoi la communication et les échanges concernant la qualité doivent être bidirectionnels :

- pendant la mise en œuvre du cadre, les INS devraient également prêter attention au concept de qualité qui y est énoncé et comprendre l'incidence de la qualité géospatiale sur leurs produits et leurs processus;
- les ONIG concernés et auxquels on demande de contribuer à la mise en œuvre du CSGM, devraient apprendre à connaître la façon dont les INS utilisent différentes sources de données, produisent des statistiques officielles, évaluent la qualité des sources de données, des processus de production et des produits finaux, et documentent la qualité pour les utilisateurs finaux.

Néanmoins, la définition générale de la qualité de tout produit comme étant son adaptation aux besoins s'applique également à l'information géospatiale et aux résultats de son intégration aux statistiques. De même, plusieurs des principes de qualité des statistiques officielles s'appliquent également aux ONIG et à leurs produits. Pour les statisticiens et les utilisateurs, il est donc essentiel

de comprendre la qualité de l'information géospatiale telle qu'elle est définie dans le CSGM et dans l'ensemble du SSN. Cela est valable à la fois dans les utilisations temporaires et l'intégration à la production statistique qui n'est pas nécessairement visible pour l'utilisateur final et dans les produits finals de statistiques géospatiales qui sont conçus pour différents utilisateurs. Pour ce qui est des statistiques, des données géospatiales de bonne qualité sous-tendent les notions de crédibilité et d'autorité, qui sont également importantes pour les produits intégrés en résultant.

En plus de la gestion classique de la qualité des données statistiques et géospatiales, cinq domaines sont particulièrement importants en ce qui concerne la qualité des données géospatiales, et non pas seulement pour les statistiques issues de la numérisation et de la révolution des données :

1. la disponibilité, l'échange et l'utilisation de données géospatiales augmentent;
2. un nombre croissant d'utilisateurs sont moins conscients de la qualité des données géospatiales;
3. les systèmes d'information géospatiale permettent d'utiliser des données géospatiales dans des applications très diverses, sans tenir compte de la pertinence de la qualité des données;
4. les systèmes d'information géospatiale actuels offrent peu d'outils de gestion de la qualité des données géospatiales;
5. la distance entre les utilisateurs de données géospatiales (utilisateurs finaux intermédiaires) et les instances les mieux informées de la qualité des données géospatiales (producteurs) ne cesse de croître.

Les cibles et les indicateurs devraient être intégrés aux cadres de qualité statistique actuels afin qu'ils soient adoptés et utilisés par les statisticiens et mieux communiqués. Enfin, ces cibles et indicateurs devraient constituer une partie importante du futur guide de mise en œuvre du CSGM.

Annexe C : Autres lectures

Cadres nationaux et internationaux

- INSPIRE⁵⁵
- Thèmes des données géospatiales fondamentales mondiales⁵⁶
- Cadre intégré d'information géospatiale⁵⁷
- Australian Statistical Spatial Framework (Cadre spatial statistique de l'Australie)⁵⁸
- GEOSTAT3⁵⁹
- Geographic Information Framework Data Content Standard – FGDC (Norme sur le contenu des données du Cadre d'information géographique)⁶⁰

Accessibilité

- World Wide Web Consortium Web Accessibility Initiative (Initiative pour l'accessibilité du Web du W3C)⁶¹
- OGC Spatial Data on the Web Working Group (groupe de travail sur les données spatiales sur le Web de l'OGC)⁶²
- Article 508 du programme d'accessibilité du gouvernement fédéral des États-Unis⁶³
- Guide pour l'intégration de l'accessibilité dans les normes ISO/CEI⁶⁴

Normes relatives au format des données et des métadonnées

- JSON ou JavaScript Object notation⁶⁵
- ISO/TC 211 Information géographique/Géomatique (catalogue de normes)⁶⁶
- SDMX⁶⁷
- Web Feature Service (WFS) de l'OGC⁶⁸
- Geopackage ⁶⁹
- Norme sur les métadonnées ISO 19115⁷⁰
- Standard Table Joining Service (TJS, service de mise en correspondance des données tabulaires) de l'OGC⁷¹

Protection du respect de la vie privée et prévention de la divulgation de données

- U.S. Census Bureau – Contrôle de la divulgation de données statistiques⁷²
- Série sur la confidentialité de l'Australian Bureau of Statistics⁷³

⁵⁵ <https://inspire.ec.europa.eu/about-inspire/563>

⁵⁶ <https://undesa.maps.arcgis.com/apps/Cascade/index.html?appid=4741ad51ff7a463d833d18cbcec29fff>

⁵⁷ <http://ggim.un.org/meetings/GGIM-committee/8th-Session/documents/Part%201-IGIF-Overarching-Strategic-Framework-24July2018.pdf>

⁵⁸ <https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Statistical+Spatial+Framework>

⁵⁹ <https://www.efgs.info/geostat/geostat-3/>

⁶⁰ http://www.fgdc.gov/standards/projects/framework-data-standard/GI_FrameworkDataStandard_Part0_Base.pdf

⁶¹ <https://www.w3.org/WAI/fundamentals/accessibility-principles/>

⁶² <http://www.opengeospatial.org/projects/groups/sdwwg>

⁶³ <https://www.section508.gov/>

⁶⁴ <https://www.iso.org/standard/57385.html>

⁶⁵ <http://www.ecma-international.org/publications/files/ECMA-ST/ECMA-404.pdf>

⁶⁶ <https://www.iso.org/committee/54904/x/catalogue/p/1/u/0/w/0/d/0>

⁶⁷ <https://sdmx.org/>

⁶⁸ <http://www.opengeospatial.org/standards/wfs>

⁶⁹ <https://www.geopackage.org/>

⁷⁰ <https://www.iso.org/standard/53798.html>

⁷¹ <http://www.opengeospatial.org/standards/tjs>

⁷² <https://www.census.gov/srd/sdc/>

⁷³ <http://www.abs.gov.au/ausstats/abs@.nsf/mf/1160.0>

Ressources du GE-IDSG – disponibles sur le wiki du GE-IDSG

- Définitions statistiques et géospatiales communes⁷⁴
- Processus et histoire du CSGM⁷⁵

⁷⁴ <https://unstats.un.org/wiki/display/ISGI/Common+Statistical+and+Geospatial+Definitions>

⁷⁵ <https://unstats.un.org/wiki/display/ISGI/The+Process+of+Developing+the+GSGF>