

Les initiatives internationales en matière de géographie et de statistique

Loonis Vincent¹

¹ *Insee, Division des Méthodes et des Référentiels Géographiques*

6 juin 2017

- 1 Généralités
- 2 L'ONU et le *Global Statistical Geospatial Framework*
- 3 Eurostat et la géographie, un aperçu
- 4 l'Insee, l'ONU et Eurostat,
- 5 L'Insee et le premier principe du GSGF.

Géoréférencement :

- 1 Insérer les données attributaires (noms, libellés, identifiants) des objets géographiques dans les fichiers statistiques,
- 2 Pour les utiliser directement ou accéder à leurs descriptions géométriques afin de faciliter les travaux de production et de diffusion statistique.

Des problèmes de terminologie (pas spécifiques à la France) :

- 1 Géoréférencement fin (x,y) : Géolocalisation (Insee), Géocodage (IGN)

Un système d'information statistique finement géoréférencé permet :

- d'éclairer le débat public au niveau local, national ou international,
- d'améliorer l'efficacité du processus de production statistique,
- de confronter des données issues de systèmes d'information indépendants,
- la mise en place de collaborations entre les instituts de statistique d'une part et de géographie d'autre part,
- l'exploration de nouvelles sources de données (téléphonie mobile, données satellites,...).

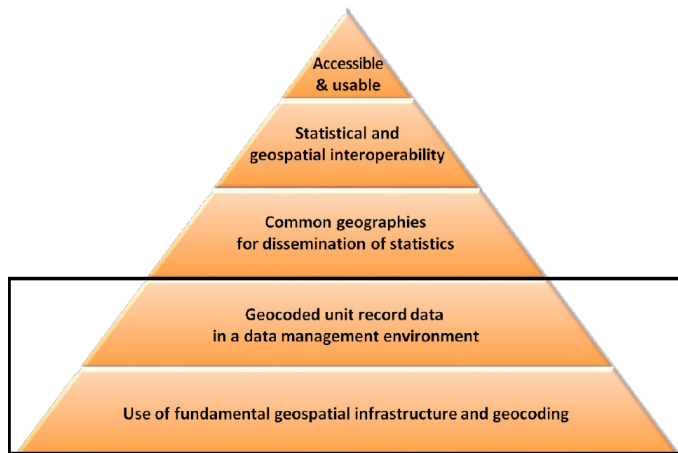
Géoréférencement, des problèmes de méthodes :

- ① Facile quand il existe un identifiant (Cadastré).
- ② Plus compliqué quand il n'y a pas d'identifiant, la qualité et la simplicité des travaux dépendent
 - du *référentiel d'adresses*,
 - du moteur d'appariement,
 - des adresses dans le fichier à géoréférencer,
 - des contours géométriques.

Pour faire face aux défis à venir, l'ONU a éprouvé le besoin de disposer d'une information géographique

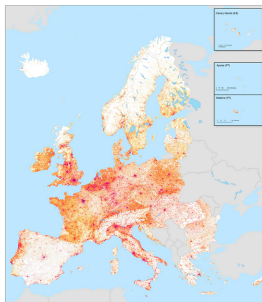
- cohérente et homogène sur l'ensemble de la planète : Global Geospatial Information Management (UN-GGIM),
- intégrée au système d'information statistique, selon des processus répondant à un guide des bonnes pratiques
 - Groupe d'experts sur l'intégration de l'information géographique et statistique, mis en place en novembre 2013
- permettant, par exemple, de calculer les indicateurs associés aux ODD

FIGURE 1 – Les 5 principes du Global Statistical Geospatial Framework



Pour permettre notamment une diffusion de 13 indicateurs issus de l'ensemble des recensements de la population 2021 dans des carreaux de $1km^2$, Eurostat

- 1 a étudié les difficultés techniques de géoréférencement des recensements (GEOSTAT 1), permettant de publier une première carte de la population



- ② a edicté les grands principes de constitution et de maintenance d'un système statistique finement géoréférencé (GEOSTAT 2)
 - a mis en place une *nomenclature* de mode de gestion de la géographie dans les Instituts Nationaux de Statistique (INS)
 - **in-house** : gestion de la géographie dans les INS
 - **hybrid** : production de l'information géographique dans les IGN, intégration dans les INS
 - **data broker** : production et intégration par les IGN pour les INS
 - a insisté sur l'importance de l'existence d'un identifiant des adresses.
- ③ déclinera le guide des bonnes pratiques de l'ONU au niveau européen (GEOSTAT 3).
- ④ a reconnu la *primauté* de l'UN-GGIM pour les questions de géographie et de statistique

Une des actions de Insee 2025 (A1-7) consiste à mettre en place un système d'information statistique géoréférencé de manière plus systématique.

Pour y parvenir, l'Insee :

- participe aux réflexions internationales (UN-GGIM, ODD, Geostat 1 et 2...)
 - l'Insee a organisé en novembre 2016, avec l'IGN, une conférence internationale sur ces thématiques.
- les intègre aux réflexions menées dans le cadre du comité stratégique de gestion de la géographie mis en place à l'Insee
- propose des projets nationaux, financés par Eurostat :
 - un rapport sur la situation de l'adresse en France (towards a French address register 2017)
 - une *French touch* plus orientée méthodologie et étude (manuel de statistique spatiale).

L'Insee et le premier principe du GSGF

Une gestion de la géographie *In-house* jusqu'à très récemment (fin 2016) :

- 1 L'Insee était producteur d'information géographique,
 - pour les données de contours (IRIS, communes, cantons, communes déléguées..),
 - pour les voies,
 - pour les adresses.
 - Les éléments géométriques produits à l'Insee étaient cohérents entre eux mais incohérents avec les systèmes extérieurs à l'Insee (RGE).
- 2 Le géoréférencement était réalisé en interne via deux méthodes à partir de l'adresse :

Une gestion de la géographie désormais semi-*hybride* :

- 1 Les données géographiques :
 - Externalisation vers l'IGN de la gestion de certains contours (IRIS, Communes..) et des voies
 - Maintien en interne de la gestion des cantons, des communes déléguées
 - Maintien en interne de la gestion des adresses **mais cohérentes avec le RGE.**
- 2 Le géoréférencement est toujours interne,
 - deux méthodes **à partir de l'adresse** :
 - **une méthode cadastre pour Fidéli, FiloSofi.**

L'Insee et le premier principe du GSGF

Une gestion *hybride* nécessiterait l'externalisation de la gestion des adresses. L'externalisation est possible sous réserve :

- ① de modifications de la BAN :
 - des modifications globales :
 - gouvernance, pérennisation, environnement juridique
 - des modifications techniques :
 - certification des contributions, historique sur 5 ans, positionnement des adresses cohérent avec les contours...
- ② de la mise en place d'une stratégie par étape
 - ① L'Insee contribue à la BAN,
 - ② L'Insee gère l'adresse de manière mutualisée en interne,
 - ③ Si la BAN répond aux exigences de qualité, elle devient le référentiel d'adresses de l'Insee.

Vers un statut de *data broker* ?

- Une externalisation du géoréférencement n'est pas à ce jour gage d'une hausse significative de la qualité.
- Les améliorations pourraient venir
 - de modifications de l'adressage dans les zones rurales,
 - de la diffusion dans le système d'information de l'État d'un identifiant adresse permettant un géoréférencement par simple appariement,
 - de modifications de la BAN.