



---

**Commission économique pour l'Europe****Conférence des statisticiens européens****Groupe d'experts des recensements de la population et des habitations****Douzième réunion**

Genève, 28-30 octobre 2009

Point 6 de l'ordre du jour provisoire

**Collecte de données en ligne****Système électronique de collecte multimodale des données du recensement à partir de registres**

Note présentée par l'Institut national italien de statistique, Italie

**I. Stratégie de collecte des données adoptée pour le quinzième recensement général de la population**

1. La stratégie de collecte de données adoptée pour le quinzième recensement général de la population fait appel aux registres municipaux pour l'identification des unités d'enquête. Le passage d'un *recensement classique* à un *recensement effectué à l'aide de registres* s'effectuera grâce à de nouvelles techniques de collecte, des questionnaires expédiés par la poste et une grande diversité de modalités de renvoi des données (poste, compilation sur le Web et envoi aux centres municipaux de collecte). On utilisera également les adresses contenues dans les archives municipales qui seront géocodées par rapport au secteur de recensement.

2. Une autre caractéristique essentielle de la nouvelle stratégie est son *adaptabilité*, à savoir recourir à différentes méthodes et techniques en fonction de l'effectif démographique des communes. Les classes d'effectif sont énumérées dans le tableau ci-dessous.

Tableau 1  
**Nombre de communes et population en 2001 par effectif démographique**

Type	Effectif démographique	Nombre	Proportion de la population de 2001 (en %)
<b>A</b>	Capitales de provinces et communes d'au moins <b>50 000</b> habitants	165	35,8
<b>B</b>	Communes comptant entre <b>20 000</b> et <b>49 999</b> habitants (à l'exclusion des capitales de provinces)	339	17,0
<b>C1</b>	Communes comprenant entre <b>5 000</b> et <b>19 999</b> habitants (à l'exception des capitales de provinces)	1 859	29,6
<b>C2</b>	Communes de moins de <b>5 000</b> habitants	5 738	17,6

3. Les modules ci-après feront partie de la nouvelle stratégie de recensement pour toutes les communes:

a) **Exploitation des registres municipaux pour orienter la collecte** en parallèle avec les nouveaux modes de diffusion des questionnaires et les techniques multimodales de collecte;

b) **Recherche, par les agents recenseurs, des non-réponses**, c'est-à-dire des questionnaires non renvoyés par les personnes recensées figurant dans le registre;

c) **Recherche, par les agents recenseurs, du sous-dénombrement du registre**, à savoir l'identification des unités d'enquête ne figurant pas dans les registres (personnes supposées vivre dans la commune mais qui n'y sont pas enregistrées);

d) **Comparaison des registres à partir desquels s'effectue le recensement dans le contexte de l'enquête.**

4. D'autres modules (définition du secteur de recensement, dénombrement des numéros d'immeubles) visent à soutenir d'autres innovations concernant les types de communes A et B uniquement. Dans ces dernières, les archives des adresses municipales, géocodées par rapport au secteur de recensement et préparées par *l'Institut national italien de statistique (Istat)*, seront validées par un **dénombrement des numéros d'immeubles**, afin d'intégrer les archives nationales aux éventuelles archives municipales locales. De plus, il sera procédé à un **recensement des logements** afin de lister le nombre d'unités d'habitation possibles pour chaque numéro d'immeuble. Ces informations, ainsi que celles des archives sur les numéros d'immeubles, serviront aux agents recenseurs dans leurs activités de recherche des non-réponses et du sous-dénombrement.

5. La **stratégie de collecte de données à l'aide de questionnaires longs et de questionnaires courts** sera également appliquée pour les seules communes de types A et B. Elle servira à estimer les variables socioéconomiques représentatives au niveau du district de recensement (les agglomérations du secteur de recensement).

6. Dans le recensement pilote qui devrait être réalisé fin octobre 2009, trois types de longueur de questionnaire seront testés: court, long et moyen, ce dernier étant une variante du questionnaire court.

7. En résumé, **dans les communes de types A et B**, la nouvelle stratégie inclura les modules ci-après:

a) Actualisation du secteur de recensement;

- b) Détermination du district de dépouillement;
- c) Élaboration d'un prototype d'archivage des numéros d'immeubles;
- d) Dénombrement des numéros d'immeubles;
- e) Recensement effectué à l'aide de registres;
- f) Stratégie de collecte de données à l'aide de questionnaires longs et de questionnaires courts;
- g) Recherche des non-réponses;
- h) Recherche du sous-dénombrement;
- i) Comparaison des registres à partir desquels s'effectue le recensement.

8. **Dans les communes de types C1 et C2**, la nouvelle stratégie inclura les modules ci-après:

- a) Actualisation du secteur de recensement;
- b) Recensement effectué à l'aide de registres;
- c) Stratégie de collecte de données à l'aide de questionnaires longs et de questionnaires courts (pour les communes de type C1 uniquement);
- d) Reconnaissance complète du secteur de recensement par l'agent recenseur à partir des informations issues du système de gestion des collectes, afin de rechercher les non-réponses et de réaliser le recensement des logements;
- e) Comparaison des registres à partir desquels s'effectue le recensement.

9. La tâche des agents recenseurs devrait être facilitée et allégée par rapport aux recensements précédents grâce aux **informations** obtenues par le **système de gestion** qui leur permettront, par exemple, de limiter leur parcours quotidien aux domiciles des ménages qui n'ont pas encore renvoyé leur questionnaire.

10. Le recensement pilote, qui sera réalisé à partir d'un échantillon stratifié d'environ 80 000 ménages, permettra de tester toutes les méthodologies décrites, à l'exception du dénombrement des numéros d'immeubles et de la comparaison entre les registres à partir desquels s'effectue le recensement. Il permettra également de voir comment les communes utilisent le système de gestion.

## II. Architecture du système en ligne

11. Le système du recensement pilote du quinzième recensement de la population se décompose en deux sous-systèmes: la collecte de données et la gestion du recensement.

12. Le système de collecte de données est une application en ligne qui offre: un accès aux ménages dûment autorisés; la possibilité de remplir le questionnaire en ligne; la validation automatique des données obéissant à certaines règles; l'exportation du questionnaire rempli vers le serveur; l'affichage et la sauvegarde sur l'ordinateur de la personne recensée d'une version imprimable du questionnaire rempli sous forme d'image, comme preuve de l'exportation de son document; une copie automatique des données définitives dans le système de gestion.

13. Le système de gestion permet aux organismes du réseau constitué pour le recensement de surveiller et de gérer toutes les opérations sur le terrain. Les innovations énoncées doivent être complétées par certaines fonctionnalités qui s'ajouteront au système de contrôle ordinaire: synchronisation et échange de données avec les systèmes externes;

synchronisation avec le système de collecte de données en ligne; complexité accrue des fonctions de contrôle due à la diversification des questionnaires et de la multiplicité des modes de renvoi des questionnaires aux communes; autorisations d'insérer du texte octroyées à un plus grand nombre de catégories d'utilisateurs; gestion claire des informations sur l'état d'avancement de chaque questionnaire; utilisation des informations pour orienter l'activité quotidienne et systématique des agents recenseurs.

14. Le système de contrôle, de par ses fonctions de communication et de diffusion des informations pour la collecte des données et de gestion des opérations, devient un rouage essentiel du dispositif de recensement. Il doit à ce titre offrir tous les services nécessaires au réseau de collecte: accessibilité, fonctions de gestion et de mise à jour utilisables, et rapports récapitulatifs actualisés en temps réel à partir des données saisies.

15. Le dispositif pour le recensement pilote comprendra 31 échantillons de communes, des opérateurs d'Istat et d'autres opérateurs du réseau constitué pour le recensement spécialisés dans les fonctions énumérées ci-dessus.

16. Chaque sous-système utilise sa propre base de données relationnelle et la communication de données entre eux est à sens unique; le sous-système de collecte alimente le sous-système de gestion en copiant les données définitives dès qu'elles sont disponibles. Cette solution, adoptée pour le prototype de recensement pilote, présente les avantages suivants:

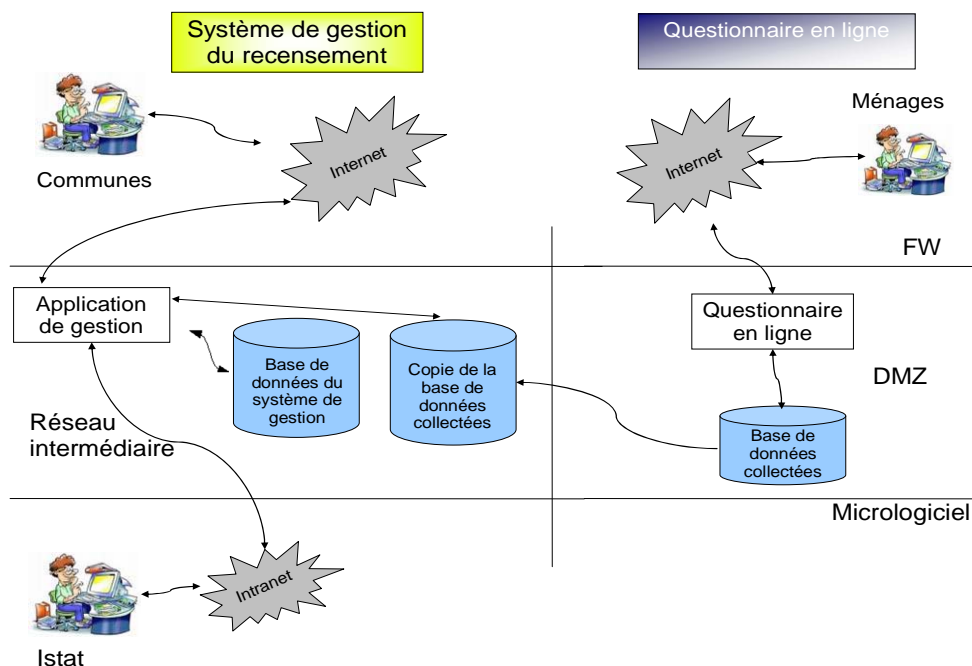
- a) Indépendance des sous-systèmes;
- b) Contrôle des performances séparé et ajustements différenciés;
- c) Meilleure protection de la confidentialité des données.

17. Il convient de noter que cette approche devra être réexaminée en cas de sous-traitance de l'infrastructure pour le questionnaire de recensement en ligne.

18. Le graphique ci-après décrit l'architecture générale et l'interconnexion des sous-systèmes.

Graphique 1

#### Architecture du système en ligne



### III. Système de gestion des enquêtes

19. Le système de gestion des enquêtes permet de gérer l'activité des agents recenseurs, l'élaboration et la fourniture du questionnaire, ainsi que la mise à disposition et la collecte d'informations.

20. L'accès au système sera contrôlé par un module de profils d'utilisateurs; en conséquence, les diverses autorisations devront être définies en fonction des responsabilités de chaque utilisateur, notamment des différents niveaux territoriaux.

21. Dans le prototype, les profils sont les suivants:

- a) ISTAT – Agents des services centraux et régionaux;
- b) Organisations du réseau constitué pour le recensement – Service d'appui du bureau municipal du recensement;
- c) Autres organisations nationales.

22. Dans le prototype du sous-système, les fonctions ci-dessous ont été prévues:

- a) Contrôle de l'état du secteur de recensement;
- b) Surveillance de l'état d'avancement de l'enquête;
- c) Gestion des opérateurs;
- d) Diverses opérations relatives au questionnaire;
- d) Fourniture de documents.

#### A. État du secteur de recensement

23. Le contrôle de l'état du secteur de recensement est la fonction centrale du système de gestion. Il permet d'orienter les opérations sur le terrain grâce à l'agenda du secteur de recensement et de superviser le résultat de la collecte de données pour chaque ménage à partir du registre municipal et à l'aide de formulaires récapitulatifs et de formulaires complémentaires détaillés. Cette fonction peut être utilisée par n'importe quel type d'opérateur autorisé (agent recenseur, coordinateur ou responsable communal).

24. L'agenda du secteur de recensement garde des traces des changements de statut de chaque questionnaire. À l'origine, il est alimenté par les données issues du registre de la population et par la liste des adresses de chaque secteur de recensement. Il est ensuite actualisé en permanence par les données du système de collecte et celles apportées par les agents recenseurs et les opérateurs du service d'appui.

25. Les formulaires récapitulatifs et les formulaires complémentaires détaillés fournissent des informations sur le type de logement et la composition du ménage. Les informations contenues dans les seconds proviennent d'un questionnaire unique rempli soit directement par le biais du système de collecte s'il est transmis par Internet, soit par un opérateur communal lorsqu'il a été reçu en format papier. Les formulaires récapitulatifs fournissent des informations agrégées géographiquement, à partir du secteur de recensement.

#### B. Rapports récapitulatifs

26. Les rapports récapitulatifs portent sur la section de contrôle de la collecte du système de gestion. Ils permettent une évaluation presque en temps réel de l'état d'avancement des

opérations sur le terrain. Le contrôle de la collecte doit être effectué aux différents niveaux géographiques, à partir du niveau de responsabilité et d'accès aux fichiers accordé aux utilisateurs. En conséquence, les rapports récapitulatifs peuvent être parcourus grâce à la technique de la progression descendante du niveau régional vers le secteur de recensement.

27. Grâce à la fonction d'établissement de rapports, on obtient des tableaux de données agrégées sur le nombre de questionnaires renvoyés ventilés par:

- a) Type de questionnaire (court, moyen, long);
- b) Mode de transmission (Internet, courrier postal, centre municipal, agent recenseur).

28. Des rapports sur l'état d'avancement de la collecte depuis le lancement des opérations sont également disponibles. Un responsable municipal pourra visualiser les rapports concernant tous les secteurs de recensements pilotes tandis que les agents recenseurs n'auront accès qu'à ceux qui concernent les secteurs de recensements dont ils ont la charge et les coordinateurs uniquement à ceux des secteurs assignés aux agents recenseurs qu'ils supervisent.

### **C. Gestion des opérateurs**

29. Cet ensemble de fonctions permet aux communes de gérer en toute autonomie leur personnel sur le terrain et en appui. Les utilisateurs autorisés peuvent entrer dans la base de données personnelle sur les coordinateurs et les agents recenseurs, visualiser ou modifier ces informations. Une fois les coordonnées d'un nouvel opérateur insérées, le système génère des justificatifs d'identité qui lui sont propres et qui peuvent lui être transmis directement par le système lui-même par SMS ou par courriel.

30. Les opérateurs autorisés ont également la possibilité de rattacher des agents recenseurs à des coordinateurs, d'assigner des secteurs de recensements (et des questionnaires) aux agents recenseurs et de réinitialiser les mots de passe des agents recenseurs et des coordinateurs le cas échéant.

### **D. Questionnaires**

31. Cet ensemble de fonctions permet d'imprimer des informations personnelles sur les questionnaires envoyés aux personnes recensées et de visualiser celles figurant sur les questionnaires renvoyés par Internet. On peut également imprimer des questionnaires vierges ainsi que des formulaires individuels. Dans le cadre de la présente section, un opérateur communal peut comparer les données du registre avec les résultats du recensement. Le rapport réalisé donne des informations sur chaque questionnaire individuellement.

## **IV. Système de collecte des questionnaires en ligne**

32. Dans la présente section sont décrites les caractéristiques de l'architecture logicielle et les aspects les plus intéressants du système logiciel qui donne accès au questionnaire par Internet et permet de le remplir facilement en ligne si l'on souhaite.

## A. Architecture logicielle du questionnaire en ligne

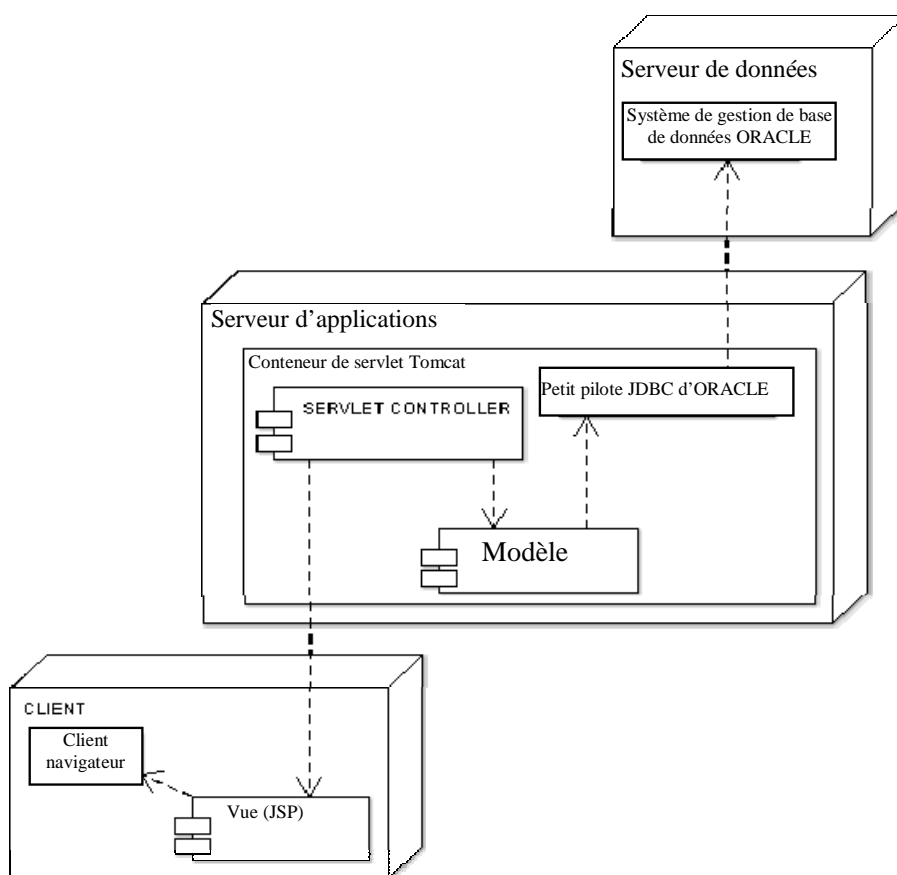
33. Grâce à cette application, les ménages concernés par le recensement pilote ont accès au questionnaire en ligne comme alternative au questionnaire classique en format papier.

## B. Architecture de l'application

34. Le graphique ci-dessous présente de façon synthétique l'architecture de l'application.

Graphique 2

**Architecture logicielle du système**



35. Le *serveur d'applications* est le module central de l'architecture en ce qu'il fournit l'infrastructure et les fonctions d'appui nécessaires à l'exécution du programme.

36. Le *serveur de données* fournit l'infrastructure et les fonctions d'appui nécessaires pour l'accès à la base de données par les protocoles de réseau. Dans ce système, le serveur de données est une base de données Oracle.

37. Le *client* se connecte au serveur d'applications grâce à n'importe quel navigateur pour accéder aux services logiciels disponibles.

38. Le modèle d'architecture utilisé pour concevoir l'application en ligne est le modèle MVC (Model-View-Controller). Il divise l'application en trois niveaux logiques distincts,

qui gèrent le statut de demande, la logique de la présentation des données et la transformation des requêtes des utilisateurs en actions exécutées par le modèle.

39. Il s'agit là du seul choix possible lorsqu'on veut concevoir un système complexe, modulaire et facile à exploiter, car il offre les garanties ci-après:

a) Souplesse: l'application est subdivisée en plusieurs composants selon leur rôle; il est alors possible d'ajouter de nouvelles actions ou de nouvelles vues sans devoir réécrire tout le programme;

b) Réutilisation: les composants sont facilement intégrables dans d'autres applications;

c) Extensibilité: la séparation des composants permet d'y ajouter de nouveaux éléments le cas échéant;

d) Sécurité: la gestion centralisée des actions facilite le contrôle d'accès aux actions et aux données.

### C. Mise en œuvre

40. L'application Web a été développée à l'aide du langage Java Enterprise Edition (cinquième version) qui donne accès aux technologies *servlet* et aux *pages de serveur Java* (JSP).

41. Les servlets sont des applications logicielles qui amplifient les fonctions du serveur Web pour lui permettre de gérer les requêtes http qui nécessitent un traitement particulier. Grâce à la technologie JSP, on crée des pages Web dynamiques en HTML adaptées à la requête des clients ou au statut de demande.

42. Le cadre de l'application a été conçu de façon à ce que cette dernière soit facilement réutilisable dans des enquêtes similaires. Il s'agit d'une architecture logicielle modulaire constituée de plusieurs classes et interfaces qui offrent les fonctions de base pour la collecte et la validation des données et l'interaction avec les bases de données.

### D. Caractéristiques du questionnaire en ligne

43. Le système de collecte est accessible grâce à une adresse Internet qui figure sur la lettre qui accompagne le questionnaire papier envoyé aux ménages.

44. Les justificatifs d'identité autorisant l'accès se composent d'un identifiant et d'un mot de passe. Un des justificatifs est communiqué dans la lettre d'accompagnement et l'autre est un identifiant connu à la fois d'Istat et du ménage concerné (par exemple, le numéro d'immatriculation fiscale du destinataire du questionnaire). Le système guide les utilisateurs à travers les différentes étapes du remplissage du questionnaire grâce à des messages indiquant l'action à réaliser sur chaque page.

45. Le questionnaire est complété lorsqu'on a répondu aux questions relevant les rubriques ci-après:

a) Liste des personnes vivant normalement dans le logement (liste «A»);

b) Descriptif codifié de la structure du ménage;

c) Liste des personnes ne vivant qu'occasionnellement dans le logement;

d) Informations générales concernant le ménage et le logement;



e) Chaque personne vivant normalement dans le logement remplit un formulaire individuel.

46. Une page d'index permet à l'utilisateur de savoir où il en est du remplissage de chacune des rubriques et établit des liens entre les rubriques et les formulaires individuels.

47. La façon de remplir le formulaire individuel est exactement la même que pour le questionnaire papier dans tous les cas. Le chemin emprunté est généré automatiquement par la séquence de questions/réponses afin de filtrer les questions. Le questionnaire en ligne a été subdivisé en pages qui s'affichent sur l'écran mais dont le nombre a été limité de façon à ne pas nuire à la facilité d'utilisation du système. L'utilisateur a à tout moment la possibilité de sauvegarder partiellement ses données sur le serveur, s'il souhaite suspendre temporairement le processus. Au cours du remplissage, les données sont automatiquement sauvegardées lorsqu'on accède à une nouvelle page.

48. L'interface utilisateur recourt aux éléments graphiques Internet ordinaires: zones de texte simple, boutons radio, cases à cocher, etc., rendant ainsi le système conforme aux procédés les plus utilisés avec lesquels les utilisateurs devraient déjà être familiarisés.

49. Le chemin emprunté pour remplir le questionnaire répond à des règles logiques qui donnent ou non l'accès à d'autres questions en fonction des réponses apportées. Par exemple, les questions concernant la fréquentation d'une école maternelle sont désactivées dès lors que le questionnaire concerne une personne âgée de plus de 6 ans (l'âge est calculé à partir de la date de naissance renseignée).

50. Des règles prédéfinies permettent également au système de détecter des informations manquantes et de faire apparaître des messages d'erreur lorsque les utilisateurs tentent d'enregistrer leurs données; il leur est alors impossible de remplir le reste du questionnaire tant qu'ils n'ont pas fourni les réponses manquantes. De la même façon, le système de gestion détecte les anomalies entre les réponses des utilisateurs et les valeurs attendues pour les variables correspondantes et affiche alors des messages d'erreur. Par exemple, des contrôles d'étendue sont effectués sur les valeurs jour, mois et année pour les dates de naissance ou sur les heures et les minutes pour les données relatives au temps.

51. Le téléchargement des données définitives est possible une fois toutes les rubriques et les questionnaires individuels remplis. Ce faisant, l'utilisateur indique qu'il a achevé de remplir le questionnaire. Avant le téléchargement définitif, l'utilisateur peut enregistrer un résumé de ses réponses en format PDF sur son ordinateur et l'imprimer.

52. Une fois les données dûment transférées, l'utilisateur reçoit un message le remerciant qui confirme la réception des informations; ce document est généré en format PDF afin qu'il puisse être enregistré et imprimé.

## V. Conclusions

53. Le présent document décrit les principales caractéristiques d'un système intégré en ligne de collecte multimodale de données et de gestion du recensement. Ce système sera testé dans le cadre d'un recensement pilote qui sera réalisé à la fin d'octobre 2009. Les spécifications du système trouvent leur origine presque directement dans les innovations méthodologiques du recensement de la population italienne de 2011. Le degré élevé de complexité du processus de collecte vise à réduire le nombre d'agents recenseurs, ce qui est possible en perfectionnant la façon de travailler des services d'appui communaux. Cela suppose également un système de contrôle plus complexe qui tend à devenir un système de gestion des opérations sur le terrain: les actions des agents recenseurs sont plus ciblées par rapport au secteur de recensement et le contrôle du processus de collecte multimodale des questionnaires gagne en efficacité.

54. Le recensement pilote devrait fournir des indications utiles à la fois sur la facilité de remplissage du questionnaire en ligne par les ménages qui choisissent ce mode de transmission et sur l'efficacité supposée des opérateurs communaux qui utilisent le système de gestion du recensement

---