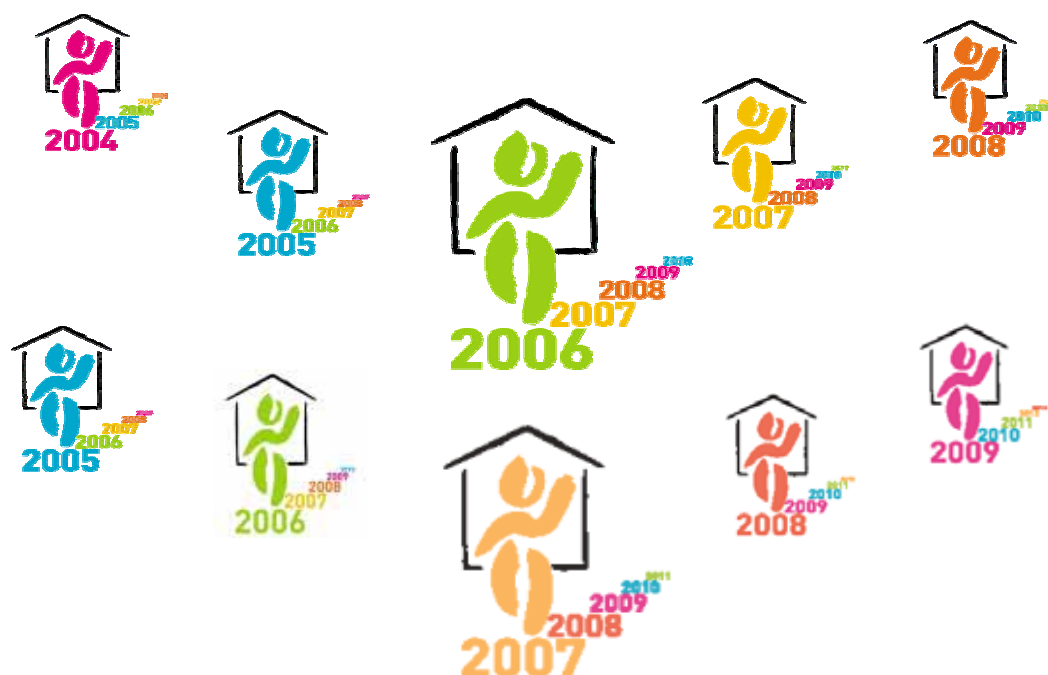




La diffusion des résultats statistiques du recensement de la population

Marcel BELLLOT
Marie-Hélène BOULIDARD
Philippe CLAIRET

- Rapport d'un groupe de travail du Cnis

N° 123

Mars 2011

La diffusion des résultats statistiques du recensement de la population

Rapport du groupe de travail du CNIS - mars 2011

(version 17 mars 2011)

Sommaire

PRÉAMBULE	4
RÉCAPITULATIF DES RECOMMANDATIONS DU GROUPE DE TRAVAIL	5
AMÉLIORER LE CONTENU AU REGARD DES BESOINS EXPRIMÉS PAR LES UTILISATEURS : 5	
FACILITER L'APPROPRIATION DE L'OFFRE PAR TOUS LES UTILISATEURS EN DEVELOPPANT L'INTERACTIVITE DU SITE :	6
Développer l'interactivité dans l'accès aux données	6
Développer l'interactivité entre la donnée et la métadonnée	7
Parfaire le développement de l'interactivité dans l'accès à la documentation, et compléter le contenu de celle-ci.....	8
CONCERNANT LES SERVICES COMPLEMENTAIRES AUX PRODUITS EN LIGNE	9
CONCERNANT D'AUTRES ACTIONS D'ACCOMPAGNEMENT DE LA DIFFUSION DU RECENSEMENT	9
RAPPORT	10
I. INTRODUCTION	10
II. LES CONSTATS	11
II.1. Les axes qui ont guidé la diffusion du recensement	11
II.2. Repères sur la fréquentation du site.....	11
II.3. Les points positifs de la diffusion.....	12
II.4. Les insatisfactions vis-à-vis de la diffusion.....	13
II.5. Les actions d'accompagnement.....	14
II.6. L'écoute et la réactivité de l'Insee sur le recensement	14

III. AVIS ET PROPOSITIONS DU GROUPE	15
III.1. AMÉLIORER LE CONTENU AU REGARD DES BESOINS EXPRIMÉS PAR LES UTILISATEURS ET DES CONTRAINTES.....	15
III.1.1. Activité.....	15
III.1.2. Immigration - nationalités	15
III.1.3. Catégories sociales	16
III.1.4. Logement	16
III.1.5. Mobilités	16
III.1.6. Plus généralement les comparaisons avec 1999	16
III.1.7. Des tableaux plus détaillés pour l'échelon France (et région ?) affichés en ligne.....	17
III.1.8. Enrichir l'analyse statistique des utilisateurs par l'accès à des données de référence ou en permettant d'effectuer facilement des comparaisons	17
III.1.9. Compléter certains fichiers détails de quelques données	17
III.1.10. Pouvoir calculer une densité à l'infra communal	18
III.2. FACILITER L'APPROPRIATION DE L'OFFRE PAR TOUS LES UTILISATEURS EN DÉVELOPPANT L'INTERACTIVITÉ DU SITE OFFERTE PAR LA TECHNOLOGIE INTERNET ..	18
III.2.1. Développer l'interactivité dans l'accès aux données pour accompagner l'internaute dans sa recherche de données.....	18
III.2.2. Apporter quelques améliorations à l'organisation de l'accès aux bases de données ou à leur contenu.	20
III.2.3. Développer l'interactivité entre la donnée et la métadonnée afin de rapprocher les métadonnées des données	22
III.2.4. Développer l'interactivité dans l'accès à la documentation pour améliorer l'efficacité de l'outil d'accompagnement	23
III.2.5. Apporter quelques compléments à la documentation.....	24
III.2.6. Continuer de développer les Outils d'accompagnement en ligne - Aide et foire aux questions	24
III.2.7. Intégrer le recensement dans l'offre d'informations des données locales pour une offre complète et cohérente	25
III.3. Les services complémentaires aux produits en ligne	25
III.3.1. Les produits sur mesure.....	25
III.3.2. Le service Diaf - Diffusion infra communale de zones à façon	26
III.4. Autres actions d'accompagnement de la diffusion du recensement.....	26
III.4.1. Une formation pour les experts ou chargés d'étude et pour l'université à pérenniser.....	26
III.4.2. Des modules pédagogiques en ligne	26
III.4.3. Un système de réponse réactif.....	27
III.4.4. Une nouvelle évaluation dans quelques années.....	27

ANNEXES	28
Mandat du groupe	28
Composition du groupe.....	29
Quelques remarques sur la diffusion des produits du recensement rénové de la population : le point de vue du chercheur. Note de M. Alexandre Kych - réseau QUETELET - le 4 mars 2011.	30
La diffusion des résultats du recensement français depuis 2009 - Insee Intervention à la Conférence des statisticiens européens - Nations unies - 1 ^{er} juillet 2010 -	35
L'enquête en ligne de juillet 2010 auprès des internautes qui ont visité le site du recensement.....	35

Préambule

Les dix huit mois qui viennent de s'écouler ont été une période test pour la diffusion des résultats statistiques du recensement de la population française. Mis en place par le CNIS à l'automne 2009, notre groupe de travail a fourni une riche contribution pour apprécier cette diffusion et son accompagnement.

Le groupe de travail a fait de nombreuses suggestions pour améliorer les conditions de diffusion et d'appropriation du recensement. Réunies dans le présent rapport, nos propositions sont susceptibles d'être mises en œuvre progressivement par l'INSEE, étant entendu que, dans la phase actuelle, la diffusion annuelle reste la priorité.

L'Institut a déjà prouvé sa réactivité en prenant en compte plusieurs des remarques faites au cours des travaux du groupe. Le saut qualitatif à effectuer porte moins sur les contenus offerts que sur la recherche d'une meilleure adéquation entre la logique de recherche de l'utilisateur et le cheminement proposé en ligne pour y accéder.

La diffusion des recensements 2006 et 2007 est une réussite qu'il convient de saluer, mais des progrès peuvent être engagés pour les années futures sur les différents axes mentionnés dans le rapport.

J'invite la Commission "Territoires" du Conseil national de l'information statistique à refaire cet exercice après une phase de cinq années de diffusion du recensement, et de préférence après l'opération coordonnée de diffusion des pays européens, sous une forme à définir.

Je remercie les agents des collectivités locales, les représentants des entreprises et les chercheurs, membres du groupe, les experts invités, qui ont contribué à la richesse du contenu de ce rapport, ainsi que les représentants de l'Insee qui nous ont accompagnés dans nos travaux.

Marcel Belliot
Délégué général de la Fédération nationale des agences d'urbanisme
Président du groupe diffusion RP du Cnis

Récapitulatif des recommandations du groupe de travail

Le groupe a rédigé 48 propositions. Elles sont présentées ici par ordre d'apparition dans le rapport avec mention du degré de priorité et structurées en quatre grands groupes. Vingt propositions sont classées en priorité forte, seize en priorité moyenne et douze en priorité faible. La difficulté de mise en œuvre n'a pas été prise en considération dans le rapport, il appartiendra à l'Insee de l'apprécier pour chaque cas.

AMÉLIORER LE CONTENU AU REGARD DES BESOINS EXPRIMÉS PAR LES UTILISATEURS :

Proposition 1 - priorité forte : modifier pour les 3 prochains millésimes de diffusion le contenu des « Chiffres Clés »- Thématique Activité afin de rendre possible la comparaison avec 99, compte tenu des changements de concepts. Pour les millésimes suivants, la comparaison se fera par rapport à l'année n-5 et ne posera donc plus de difficultés.

Proposition 2 - priorité moyenne : en raison des changements de concepts sur l'activité et l'emploi, développer une offre de données sur les situations mixtes quant au statut d'activité.

Proposition 3 - priorité forte : abaisser, autant que la qualité statistique et le cadre juridique l'autorisent, le seuil de diffusion, actuellement à 50 000 habitants, des données sur la nationalité en 13 modalités.

Proposition 4 - priorité faible : développer l'offre de données selon la nationalité, notamment croiser nationalité et lieu de naissance à l'étranger et nationalité de la personne de référence du ménage et taille des ménages, autant que la qualité statistique et le cadre juridique le permettent.

Proposition 5 - priorité forte : abaisser, autant que la qualité statistique l'autorise, le seuil de diffusion des données par catégories socioprofessionnelles en 42 postes.

Proposition 6 - priorité moyenne : proposer pour les communes de 10 000 habitants ou plus une offre de données prédéfinies (tableaux détaillés) sur la population et les caractéristiques d'habitat pour les logements récents.

Proposition 7 - priorité moyenne : poursuivre la réflexion méthodologique sur le recensement des constructions neuves et clarifier le processus pour l'utilisateur. Dans l'état actuel de la collecte où l'année de construction est déclarative, il convient d'alerter l'utilisateur sur la mobilisation des résultats faisant intervenir cette variable. Elle ne permet pas de dénombrer le volume de la construction neuve.

Proposition 8 - priorité faible : pour les champs ayant connu des changements de définition par rapport à 1999, mettre à disposition des indicateurs 1999 avec un contenu recalculé sur la base des concepts en vigueur depuis le millésime 2006 (niveau de détail du type des indicateurs à l'infra communal).

Proposition 9 - priorité forte : mettre en ligne sur le site RP99, les données IRIS Analyse 1999, tout du moins celles issues de l'exploitation principale, à charge pour l'utilisateur, pour toute comparaison, de prendre les précautions qui s'imposent du fait des changements de définition, voire de découpage,

Proposition 10 - priorité forte : enrichir les tableaux détaillés d'un 4^e niveau d'agrégation des données (France entière ou France entière et région si la qualité statistique le permet pour toutes les régions) afin d'offrir un niveau de modalités très détaillées.

Proposition 11 - priorité moyenne : développer l'offre de données mobilisables comme références pour des comparaisons spatiales. Ceci pourrait être réglé par l'enrichissement des niveaux géographiques

supracommunaux offerts dans les fichiers d'indicateurs des Chiffres Clés en proposant, comme pour les régions et départements une diffusion par EPCI (Communautés Urbaines, Communauté d'Agglomération, Communautés de Communes) et par strates démographiques de communes, d'EPCI ou d'unités urbaines. L'affichage par tranches de taille est aujourd'hui totalement absent du plan de diffusion.

Proposition 12 - priorité moyenne : enrichir de variables complémentaires les fichiers détails afin de répondre aux besoins à ce jour non satisfaits de catégories d'utilisateurs du domaine public intervenant sur des territoires supra communaux, notamment la FNAU (Fédération Nationale des Agences d'Urbanisme) et les responsables régionaux en charge de la formation et de l'emploi.

Proposition 13 - priorité moyenne : enrichir l'offre infra communale sur la thématique des mobilités, en codifiant le lieu de travail à l'IRIS, pour obtenir des navettes domicile-travail à cet échelon.

Proposition 14 - priorité faible : enrichir l'offre infra communale, sur la thématique population, en fournissant la superficie de l'IRIS afin de permettre un calcul de densité à l'infra communal,

FACILITER L'APPROPRIATION DE L'OFFRE PAR TOUS LES UTILISATEURS EN DEVELOPPANT L'INTERACTIVITE DU SITE :

Développer l'interactivité dans l'accès aux données

Proposition 15 - priorité forte : elle énonce le principe fondateur de l'ensemble des propositions qui suivent pour répondre à ce grand objectif. Le groupe propose de travailler à développer l'interactivité du site offerte par la technologie Internet, tant pour l'accès aux données qu'à la documentation, afin de guider l'internaute non expert et non familier du site par un questionnement pas à pas, tout en conservant un accès plus direct aux produits et à la documentation pour les autres catégories d'internautes.

Cette ambition amènera à une refonte sensible de la page d'accueil du site.

Proposition 16 - priorité forte : alors que le mode actuel d'accès aux données est unique, à savoir par type de produits, permettre une recherche multicritère, géographique, thématique, datée et, toujours, par type de produits, avec au minimum un seul critère obligatoire et sans ordre de précellence.

Proposition 17 - priorité faible : compléter le processus de recherche géographique par la possibilité d'une sélection pas à pas questionnant d'abord sur le type de zonage recherché puis offrant le choix dans une liste nominative pour le niveau retenu.

Proposition 18 - priorité forte : compléter l'offre de niveaux géographiques en faisant de l'EPCI dans ses diverses déclinaisons un niveau géographique à part entière et en introduisant le niveau par strates démographiques.

Par ailleurs, l'occurrence systématique de l'EPCI englobant la commune recherchée, y compris pour les cas où l'appellation de l'EPCI ne contient pas le nom de la commune centre, viendra compléter la liste de choix proposés lors d'une recherche géographique par saisie directe du nom du territoire recherché.

Proposition 19 - priorité forte : dans le cadre d'une recherche thématique, permettre à l'internaute par un affichage dynamique et interactif, d'obtenir quelques précisions sur le contenu du thème qu'il s'apprête à sélectionner. De même, dans le cadre d'une recherche documentaire, pouvoir préciser synthétiquement le contenu de la documentation que l'utilisateur est en passe d'ouvrir.

Proposition 20 - priorité forte : faciliter l'accès aux visualisations cartographiques des résultats du RP par un accès direct à partir d'un choix dans une liste d'indicateurs précisant également les niveaux géographiques disponibles. Ceci suppose de travailler au rapprochement du site de cartographie et du site RP.

Proposition 21 - priorité faible : développer des solutions permettant à tout moment à l'internaute perdu sur le site RP de se réorienter vers le site Insee ou vers un autre produit du site RP.

Proposition 22 - priorité moyenne : dans le cadre d'une recherche multicritère, introduire la possibilité d'une sélection d'un territoire de référence pour une comparaison spatiale selon une même procédure de recherche géographique et pour un affichage simultané des résultats sur le territoire étudié et sur celui pris pour référence.

Proposition 23 - priorité moyenne : faire de l'IRIS un échelon géographique à part entière et donc accessible dans le cadre d'une recherche multicritère relative aux « chiffres clés », pour un affichage limité à quelques données essentielles, compte tenu des précautions d'usage s'imposant à cet échelon.

Afin de parfaire le développement de l'interactivité dans l'accès aux données, quelques propositions visent en complément à apporter des améliorations à l'organisation de l'accès aux bases de données ou à leur contenu :

Proposition 24 - priorité forte : réinterroger l'appellation des différents types de produits afin de s'assurer d'une compréhension sans ambiguïté, notamment modifier l'intitulé « données téléchargeables » au bénéfice de « bases de données ». Au-delà du site RP, cette proposition pourrait plus largement s'appliquer à l'ensemble du site Insee. Par exemple, dans le cas du recensement, la distinction Chiffres-clés / Tableaux détaillés n'est pas clairement perçue par une partie des internautes.

Proposition 25 - priorité moyenne : afin de faciliter l'approche de nombre d'utilisateurs locaux, proposer en complément à celle existante une autre organisation des fichiers « bases de données », notamment les « bases de tableaux détaillés », à savoir une partition territoriale (régionale, voire départementale) de fichiers multithématiques pour compléter l'actuelle partition thématique de fichiers France entière.

Proposition 26 - priorité forte : introduire davantage de zonages pour les « bases de tableaux détaillés » conformément à ce qui est proposé dans les « bases de chiffres clés ».

Proposition 27 - priorité faible : offrir une diversité de formats pour le téléchargement.

Proposition 28 - priorité moyenne : réfléchir à la possibilité de s'orienter vers un système d'ouverture de « web services », une technologie adaptée aux modes actuels de services de données.

Proposition 29 - priorité forte : fournir chaque année une information complète en ligne sur les modifications du dessin d'enregistrement des fichiers, des dictionnaires de variables.

Proposition 30 - priorité forte : il semble vivement souhaitable pour gagner en lisibilité de revoir l'organisation de l'offre sur les flux de mobilité et d'en réviser le mode d'accès. Il s'agit de faciliter la recherche par la possibilité d'un accès par requête (territoire d'arrivée/ territoire de départ) avec production de résultats par affichage en ligne pour les flux principaux, complété par la livraison de base de données.

Développer l'interactivité entre la donnée et la métadonnée

Proposition 31 - priorité forte : rapprocher les métadonnées des données par un affichage dynamique et interactif d'une définition littérale courte et du lien vers la définition complète correspondante, ainsi que d'une alerte en cas de changement de concept par rapport à 1999.

Proposition 32 - priorité forte : réinvestir la proposition n°13 du rapport « Giblin » sur la qualité des données visant à indiquer la précision des résultats à l'utilisateur de manière aussi simple et intuitive que possible. Selon ce principe, on peut imaginer repérer dans un tableau par un code couleur (rouge et/ou orange) les données les plus fragiles. La voie d'excellence sur cet axe serait de permettre de calculer et d'afficher un intervalle de confiance à 95 % en relation avec une donnée ou un ensemble de données sur la zone affichée.

Proposition 33 - priorité faible : imaginer une solution simple pour traiter la question des arrondis et ainsi rassurer l'internaute inquiet des résultats « inexacts ».

Proposition 34 - priorité moyenne : donner un accès plus facile à la géographie de l'infra communal par un lien direct et visible avec les fichiers « contour IRIS » sur la page d'accueil de l'offre infra communale. Concernant les fichiers des contours numérisés à l'IRIS, il faut viser à réduire les délais de mise à disposition d'une nouvelle version lors des mises à jour, la rendre concomitante à la diffusion des données..

Proposition 35 - priorité moyenne : compléter l'offre de métadonnées pour accompagner l'utilisateur de bases de données en :

- garantissant l'identification précise, y compris sa version, du fichier téléchargé afin d'ôter tout doute lors d'une réutilisation ultérieure
- accompagnant les bases de données d'une métadonnée de description complète du contenu du fichier et de ses caractéristiques techniques ainsi qu'un mode d'emploi.

Parfaire le développement de l'interactivité dans l'accès à la documentation, et compléter le contenu de celle-ci

Proposition 36 - priorité moyenne : mettre à disposition un outil équivalent aux anciens « fascicules gris » qui décrit les traitements et la construction des variables de chaque exploitation.

Proposition 37 - priorité forte : développer un outil d'accompagnement pour repérer rapidement l'ensemble des changements par rapport à 1999.

Proposition 38 - priorité moyenne : proposer un schéma synthétique interactif qui récapitule l'ensemble de l'offre du site RP. La solution la plus aboutie en terme d'interactivité, serait de faire de ce schéma une interface d'accès aux différents produits.

Proposition 39 - priorité forte : valoriser davantage la page « calendrier/produits » du recensement du site RP afin de lui redonner toute son importance et l'enrichir pour donner une vision de la prochaine campagne de diffusion en incluant les nouveautés prévues.

Proposition 40 - priorité faible : enrichir l'Aide et lier ces rubriques à la foire aux questions et à la documentation pour gagner en lisibilité quant à la cohérence d'ensemble des différents outils d'accompagnement.

Proposition 41 - priorité forte : construire une aide pour l'ouverture des fichiers détails avec divers logiciels.

Proposition 42 - priorité faible : pour parfaire l'appropriation de l'offre Recensement par tous les utilisateurs, le groupe recommande que celle-ci se trouve intégrée dans l'offre d'informations de Données locales pour une offre complète et cohérente qui intégrerait les améliorations proposées dans ce rapport.

CONCERNANT LES SERVICES COMPLEMENTAIRES AUX PRODUITS EN LIGNE

Proposition 43 - priorité forte : le groupe constate que les conditions d'application du secret statistique dans le cadre de la réalisation de tableaux sur mesure complexifient les procédures, et donc les délais d'instruction des commandes sur mesure. Il suggère que des solutions soient trouvées sur ce point. Il propose que le cas des chercheurs qui pâtissent de ces nouvelles dispositions soit traité en priorité dans le cadre des diverses dispositions prises pour le monde de la recherche dans l'accès aux données publiques.

Proposition 44 - priorité faible : dans le cadre du service Diaf (diffusion infra communale à façon), proposer à l'avenir une liste d'indicateurs essentiels (du type « îlots 15 » de 1999) accessibles aux entreprises ou associations à des conditions financières restant à préciser. De même, ce service pourrait être ouvert aux chercheurs. Il convient que cela soit fait en conformité avec les contraintes déterminées par la Cnil.

CONCERNANT D'AUTRES ACTIONS D'ACCOMPAGNEMENT DE LA DIFFUSION DU RECENSEMENT

Proposition 45 - priorité moyenne : pérenniser une formation pour les experts ou chargés d'études, au-delà des partenaires habituels des directions régionales de l'Insee et pour l'université.

Proposition 46 - priorité faible : imaginer à moyen terme la conception de modules pédagogiques en ligne.

Proposition 47 - priorité faible : améliorer l'efficacité de la réponse du service de réponse en ligne « Insee Contact » en imaginant à l'entrée une procédure de repérage rapide du profil et du niveau d'expertise de l'utilisateur posant la question.

Proposition 48 - priorité moyenne : le groupe propose qu'une nouvelle démarche globale d'évaluation de la diffusion des résultats du recensement soit initialisée en 2014 c'est-à-dire après un cycle complet de diffusion. D'ici là, des rencontres à thème avec une dimension méthodologique ou sur différents points de la diffusion jugés sensibles pourraient, selon le groupe, rencontrer l'intérêt des utilisateurs et des experts.

Rapport

I. Introduction

Les travaux du présent groupe de travail du Conseil national de l'information statistique se situent dans la suite de travaux antérieurs menés dans le cadre du Conseil qui ont conduit à orienter la diffusion des résultats du recensement de la population. Il y a eu d'une part une « évaluation de la diffusion des résultats du recensement de la population de 1999 » (rapport Cnis N° 87 - février 2004), d'autre part « Utilisation des données produites par le recensement rénové de la population et leur diffusion » (Rapport du Cnis N° 98 de décembre 2005) et enfin cinq séances d'un sous-groupe de travail dont la dernière a eu lieu en mai 2009 avant la création du présent groupe. Celui-ci rassemble des représentants des collectivités locales, des entreprises et des chercheurs¹ en présence de l'Insee. Les travaux qui font l'objet du présent rapport ont été engagés en septembre 2009 pour se terminer en mars 2011.

La première diffusion générale du recensement qui fait l'objet de la présente analyse a été réalisée en juillet, octobre et décembre 2009 pour les résultats de l'année 2006, cumulant les collectes des années 2004 à 2008, la deuxième en juillet et octobre 2010 pour les résultats de l'année 2007.

Cette diffusion se fait dans le cadre de l'[Arrêté de diffusion du 19 juillet 2007](#).

Le mandat du groupe (cf. annexe) consiste à :

- s'assurer que le dispositif de diffusion est pertinent et correspond bien, de fait, aux besoins des utilisateurs. Il s'agit, d'une part, de vérifier a posteriori la pertinence du dispositif, d'autre part, de proposer les changements qui devront être pris en compte pour la diffusion.
- examiner le dispositif d'appropriation prévu au plan des outils (métadonnées, notes de méthodes, outils de « formation », etc.) et des structures (rôle des directions régionales de l'Insee, des associations de collectivités territoriales, des institutionnels régionaux, des relais nationaux, etc.) et proposer les améliorations souhaitables.

Au cours des travaux, les membres ont fourni leurs remarques d'utilisateurs sur le contenu des produits, la documentation, l'accès aux données et aux métadonnées, le service d'accompagnement mis en place par l'Insee. Le président a souhaité aller au-delà en mobilisant la réflexion des membres sur l'interactivité qui pourrait être encore plus développée dans l'accès aux données et surtout dans l'accompagnement de l'utilisateur compte tenu des ressources d'internet qui est le support privilégié de diffusion du recensement.

Enfin pour enrichir les travaux du groupe, le président du groupe a passé commande d'une enquête de satisfaction en ligne pour avoir le retour de l'avis des internautes. Celle-ci a été réalisée en juillet 2010, donc après la deuxième vague de résultats standards. Les résultats ont été exploités à partir des 1 025 réponses obtenues en un mois. Le groupe a très largement tenu compte des enseignements de cette enquête dans ses propositions.

¹ Cf. contribution en annexe « le point de vue du chercheur »

II. Les constats

II.1. Les axes qui ont guidé la diffusion du recensement

Le Recensement de la population est une pièce maîtresse d'un système d'information locale actualisé chaque année :

une diffusion gratuite et massive des produits sur internet qui couvre les différents besoins, du grand public jusqu'aux experts, avec les outils appropriés ;

la nouvelle méthode de collecte fournit plus rapidement une information actualisée chaque année ; c'est un gain appréciable qui s'accompagne cependant dans certains cas d'une perte de précision, notamment au niveau infra communal ;

des résultats jusqu'au niveau infra communal (IRIS) y compris pour des « zones à façon » (DIAF) définies par l'utilisateur ayant une mission de service public ;

le recensement continue d'apporter des informations analogues et sur les mêmes thèmes que par le passé ;

des concepts améliorés, adaptés aux évolutions de la société et homogénéisés au niveau européen avec donc quelques changements par rapport à 1999 ;

une documentation détaillée sur les variables et des conseils pour l'utilisation des résultats afin de garantir l'autonomie de l'utilisateur.

Ces axes de l'Insee qui ont guidé la diffusion du recensement tiennent compte des débats qui ont eu lieu au Cnis avant la diffusion elle-même et des contraintes fixées par la Cnil (Commission nationale informatique et libertés) mentionnées dans l'arrêté de diffusion.

(voir en annexe un texte général sur la diffusion du recensement français avec un encadré sur la collecte)

II.2. Repères sur la fréquentation du site

Au mois de janvier 2011, 105 500 visites ont eu lieu sur recensement.insee.fr, soit 10 % de plus qu'en janvier 2010. Ces deux mois sont les premiers comparables en terme de contenu de site sur une année, le site RP 2006 n'ayant été complet qu'à la mi-décembre 2009. Ce chiffre ne comprend pas les visiteurs du site « statistiques locales » qui contient les mêmes données « Chiffres clés », qui a connu un saut de plus de dix mille visites par mois fin 2009 lorsque les données du recensement ont été intégrées dans ce site.

Un quart des visiteurs reste entre 30 secondes et 5 minutes sur le site, un autre quart de 5 minutes à une heure, une petite moitié reste moins de 30 secondes.

Parmi les répondants de l'enquête en ligne faite en juillet 2010, 10 % consultaient le site pour la première fois, et 30 % quelquefois par an. Les utilisateurs fréquents, qui consultent plus d'une fois par mois, seraient de l'ordre de 45 %. Dans la réalité, les primo-utilisateurs et les utilisateurs peu

fréquents sont vraisemblablement plus nombreux, comme il avait été constaté sur le site de l'Insee il y a quelques années.

Le thème le plus fréquemment consulté, parmi les six thèmes des Chiffres clés, mesuré par les téléchargements, est celui de la population, viennent ensuite ceux du logement et de l'emploi/activité, deux fois moins, enfin ceux de la formation et de la famille, quatre fois moins.

Les utilisateurs recherchent, pour la moitié d'entre eux, une information précise sur une ou plusieurs zones géographiques, et pour 70% d'entre eux la recherche porte principalement sur une commune, l'échelon géographique, comme le type de données recherchées n'était pas exclusif dans les questions posées. Un peu moins de la moitié (41%) recherchent un ensemble d'informations sur une zone, ils sont aussi nombreux (39%) à souhaiter extraire à partir de bases de données pour travailler ces données dans leur propre environnement.

Les autres échelons constituent aussi des centres d'intérêt : 31% un regroupement de communes, 33% le département ou la région, 14% l'échelon infra communal, 15% l'échelon France (source : répondants à l'enquête en ligne juillet 2010)

Qui consulte le site du recensement ? Ceux qui ont répondu à l'enquête en ligne - un millier - en donnent un aperçu :

12 % d'étudiants ou lycéens,

17 % de particuliers (raisons personnelles ou associatives),

71 % de professionnels, parmi lesquels 24 % de collectivités locales, 22 % de bureau d'études, 18 % d'administration ou assimilé, 13 % d'entreprises, les autres se répartissent entre des organismes locaux (ex agences d'urbanisme ou de développement), des associations ou syndicats, des enseignants, des chercheurs, des élus.

II.3. Les points positifs de la diffusion

Avant l'ouverture du site, un test qualitatif de celui-ci a été réalisé par la SOFRES auprès de personnels d'entreprises, du secteur public ou de l'enseignement. Le site des résultats du recensement a recueilli « une appréciation globale positive en terme d'utilité et de praticité », cette nouvelle interface apparaît « plus simple, plus accessible, plus facile d'utilisation dans sa réalisation par rapport aux autres espaces sur les données locales ou infra communales ». Ceci n'exclut pas quelques critiques précises sur tel ou tel point.

L'enquête faite en ligne après un an de diffusion (1 000 réponses en un mois), témoigne d'une satisfaction générale mais à un niveau moindre qu'attendu compte tenu des éléments très positifs de l'enquête qualitative faite avant l'ouverture. Le groupe en a analysé les raisons et fait des propositions d'axes de progrès pour en tenir compte.

La gratuité a été unanimement appréciée par les utilisateurs. En 1999, les utilisateurs avaient regretté le champ trop limité de la diffusion gratuite même si les prix étaient jugés corrects pour la partie faisant l'objet d'une tarification.

L'Insee a respecté strictement le calendrier annoncé pour l'ensemble de la diffusion du recensement, exception faite de l'actualisation des contours IRIS numérisés. Le non-respect du calendrier avait été un des griefs mentionnés dans le rapport de 2004 du Cnis évaluant la diffusion du recensement de 1999.

Les types de produits proposés en ligne répondent bien aux besoins comme en témoigne la forte satisfaction sur ce point.

Le produit Chiffres clés apporte un très bon niveau de satisfaction pour les utilisateurs pressés qui ont besoins de quelques données de référence sur une commune ou un échelon géographique standard.

La quantité d'information fournie a été saluée comme un progrès malgré la simplification des produits offerts et l'effort à fournir par les utilisateurs professionnels pour construire les résultats correspondant à leurs propres besoins. A ce titre, la mise à disposition de fichiers détails anonymes est un atout majeur.

La fourniture de données 1999 pour les chiffres clés apporte beaucoup de confort pour les comparaisons. A contrario, leur absence dans les autres produits se fait d'autant plus sentir.

L'introduction de labels de qualité pour les données IRIS, dans le sens des recommandations du rapport « Giblin » du Cnis, est un progrès important et remarquable, qui constitue une relative exception dans la mise à disposition des statistiques.

La documentation recueille un avis très favorable, car elle constitue un accompagnement de grande qualité, même si son accès peut être amélioré compte tenu de son important volume.

II.4. Les insatisfactions vis-à-vis de la diffusion

L'entrée dans le site Recensement à partir du site Insee est appréciée quant à sa simplicité. Mais au-delà, l'accès à l'information réclame un aménagement pour se rapprocher de la démarche, de la logique intuitive de l'utilisateur, pour lequel l'entrée par type de produit est peu explicite. La géographie et le thème devraient constituer les entrées majeures. L'utilisateur a une difficulté à appréhender la structuration et la présentation du site Recensement par rapport à sa propre logique. Cette difficulté d'appréhension concerne également au-delà des données elles-mêmes, la documentation. Là aussi ce n'est pas tant le contenu, jugé riche, complet, pertinent et compréhensible, qui pose question mais l'accès à la documentation et surtout la compréhension de l'architecture générale de la documentation en ligne. Le rapport d'évaluation de la diffusion de 1999 avait déjà beaucoup insisté sur la nécessité d'accompagner l'utilisateur, bien que le contexte de mise à disposition des produits fût différent.

Il est difficile de constituer des résultats sur sa propre zone si elle ne figure pas parmi les zones standards. La livraison 2010 a déjà amélioré l'accompagnement sur ce point (rubrique « Aide » complétée).

Les membres du groupe regrettent l'absence de résultats statistiques pour les Zones urbaines sensibles, du moins un an après la première diffusion de résultats.

Le volume des fichiers mis à disposition a créé des difficultés à une partie des utilisateurs tant en matière de possibilité de téléchargement que d'intégration dans des tableurs de version parfois ancienne que de manipulation informatique.

La cartographie est peu accessible et « loin » de la logique du site du recensement. La nouvelle configuration des « Données locales » accessible depuis fin 2010 a permis de s'en rapprocher et d'une manière plus générale de se rapprocher de la logique du site Insee.

Les données 1999 sont difficiles à construire pour faire les comparaisons lorsqu'elles ne sont pas fournies et dans certains cas la construction est impossible en raison des changements de concept (activité). Par ailleurs, à l'IRIS, les données 1999, à l'exception de quelques grands indicateurs ne sont pas disponibles en ligne.

Les chercheurs ont l'impression, dans les conditions à l'œuvre au moment de la rédaction du rapport, d'un recul par rapport à 1999 pour ce qui les concerne, en raison des conditions d'application du secret pour les produits sur mesure, d'un moindre détail de modalités de variables qu'ils jugent importantes, de l'absence de constitution de fichiers qui leur sont spécifiquement destinés. Ils

espèrent que les discussions en cours permettront de trouver une issue favorable. La mise en place des centres d'accès sécurisés qui leur sont dédiés prend du temps, à la fois pour des raisons budgétaires et techniques.

Les professionnels qui utilisent un système d'information géographique ont des difficultés à afficher les résultats par IRIS dans leur outil en raison du délai de fourniture d'un produit adéquat des contours numérisés.

II.5. Les actions d'accompagnement

Le bilan des actions d'accompagnement mises en œuvre par l'Insee en région et à l'échelon national est plutôt positif. Pendant l'année 2009 et principalement au second semestre, cinq mille personnes dans les régions, parmi les utilisateurs locaux principaux de la statistique publique ou collaborateurs des acteurs locaux importants, ont suivi une conférence ou une formation sur l'utilisation des résultats du recensement. Des actions ont aussi été menées à l'échelon national et un module, ouvert aux entreprises et organismes publics, a été construit pour pérenniser la formation.

Des actions de communication ciblées ont permis d'informer sur la mise à disposition des résultats et leur contenu, soit directement dans des supports de l'Insee soit dans des supports de médias professionnels.

II.6. L'écoute et la réactivité de l'Insee sur le recensement

Des évolutions ont été déjà mises en œuvre dès les premiers constats d'insatisfaction exprimés au cours des travaux du groupe. L'Insee a aussi suivi précisément tous les messages et toutes les demandes lui parvenant sur les résultats du recensement. Les évolutions constatées au cours des diverses versions témoignent de cette écoute des membres du groupe et du public.

III. Avis et propositions du groupe

III.1. AMÉLIORER LE CONTENU AU REGARD DES BESOINS EXPRIMÉS PAR LES UTILISATEURS ET DES CONTRAINTES

Quelques thèmes sur lesquels le contenu pourrait évoluer :

III.1.1. Activité

Le changement de la définition des actifs et des emplois par rapport au recensement de 1999 crée une des principales difficultés dans l'utilisation des résultats du recensement. Il est en effet impossible de reconstruire des résultats 1999 à partir de la définition actuelle. La catégorie pour laquelle l'Insee conseille d'effectuer les comparaisons concerne les 25-54 ans, et il convient de le faire avec prudence. Or cette catégorie ne fait pas l'objet de diffusion d'accès simple en 1999 ou pour l'actuel recensement. Ainsi dans le produit Chiffres clés, censé être d'accès facile et synthétique, la fiche concernant l'activité est particulièrement complexe car elle mêle des exploitations différentes et comporte des évolutions sur le champ des 15-64 ans alors que les comparaisons sur les parties extrêmes de cette classe sont déconseillées. Il conviendrait de modifier (réduire ou compléter) cette fiche pour éviter des interprétations erronées, du moins jusqu'à ce que 1999 puisse être remplacé lorsqu'on disposera d'un écart de cinq années pour la diffusion des évolutions avec le même concept. Au minimum une alerte devrait être donnée dans la fiche.

Compte tenu de ce changement de concept les professionnels apprécieraient de disposer de données sur les situations mixtes : retraités/actifs, étudiants/actifs ou autres, résultats que devrait permettre le questionnement.

Il est souhaité par ailleurs un tableau complémentaire qui croiserait la condition d'emploi et le secteur d'activité.

III.1.2. Immigration - nationalités

Les nationalités les plus représentées (13 postes) ne sont disponibles que pour des zones égales ou supérieures à 50 000 habitants. Le seuil de diffusion de telles données était 5 000 habitants en 1999 pour les communes, il conviendrait de vérifier si la fiabilité statistique ne permet pas de descendre au moins le seuil à 10 000 habitants dans les tableaux détaillés. De même on peut regretter le manque de tableaux immigrés identiques à ceux diffusés pour les étrangers.

Parmi les nationalités, en particulier dans les zones frontalières, ne figurent pas le pays frontalier dans les principales nationalités.

Deux demandes complémentaires concernent le lieu de naissance à l'étranger par nationalités, et les nationalités de la personne de référence du ménage selon le nombre de personnes du ménage.

III.1.3. Catégories sociales

Les chercheurs font remarquer qu'il est impossible de travailler sur les différenciations sociales au plan local, les travaux dans ce domaine requièrent le code Catégories sociales en 42 postes. Ce niveau de détail est fourni seulement à l'échelon de la région ou pour les actifs au lieu de travail des zones supérieures à 50 000 habitants.

III.1.4. Logement

Les informations sur les moyens de chauffage intégrant les énergies renouvelables sont fréquemment demandées, ces informations ne figurent pas dans la feuille de logement utilisée pour la collecte. Le contexte plaiderait pour une évolution sur cette question.

Le champ de la diffusion concernant les logements neufs pourrait être enrichi pour les grandes communes. Il pourrait être imaginé, sous réserve de qualité statistique suffisante, quelques tableaux sur ce champ, avec des caractéristiques essentielles des occupants de ces logements neufs.

Il était prévu, pour les communes de 10 000 habitants et +, que les logements correspondants à des adresses nouvelles seraient recensés exhaustivement dans les 3 ans suivant leur intégration dans le RIL. L'Insee fait part des écueils méthodologiques de cette mesure qui viendrait déséquilibrer les enquêtes annuelles. Les logements neufs, recensés exhaustivement, sont répartis sur cinq ans.

Le nouveau dispositif de recensement semble sous-estimer localement le volume de la construction neuve (achevée depuis 1999). Il semble que les constructions neuves ne soient pas à ce jour recensées exhaustivement ou couvertes en totalité par le sondage pour celles ne donnant pas lieu à la création d'une adresse nouvelle. Le groupe souhaite que la réflexion méthodologique se poursuive sur cet axe pour améliorer la qualité du recensement des logements neufs.

III.1.5. Mobilités

Le recensement ne permet pas de distinguer les mobilités quotidiennes des mobilités hebdomadaires de travailleurs ou d'étudiants. Ce point suppose une évolution du questionnaire et il traite d'un sujet extrêmement pointu.

De même que sont fournies les mobilités domicile-travail de commune à commune, il serait souhaitable d'avoir ces navettes à l'échelon de l'IRIS. Le lieu de résidence de l'actif est facilement disponible, il conviendrait de codifier aussi le lieu de travail à l'IRIS pour réaliser la diffusion des flux en provenance des communes ou des IRIS vers les IRIS de travail, ce qui serait un grand progrès pour l'analyse des flux vers les pôles d'emploi intra urbains.

III.1.6. Plus généralement les comparaisons avec 1999

Il est regretté que les données 1999 soient absentes des tableaux détaillés ; de même dans les résultats à l'IRIS, au moins lorsque ceux-ci n'ont pas changé de limites. Concernant les données à l'IRIS, on constate que celles de 1999 ne sont pas même disponibles sur le site RP99 (cf. ci-dessous).

La fourniture de bases de données 1999 avec le contenu et les concepts de 2006 aurait été extrêmement appréciée et aurait constitué une économie collective, en évitant éventuellement des erreurs de la part des utilisateurs. Il est certain qu'une telle initiative constituait un coût spécifique supplémentaire important dans un contexte de diffusion annuelle des résultats.

Concernant la diffusion à l'IRIS, l'offre millésimée 2007 de la thématique population, avec la prise en compte supplémentaire de nouvelles bornes pour les données sur l'âge en 8 modalités rendant ainsi possible la comparaison immédiate avec 1999, constitue déjà une avancée.

Au-delà, le groupe propose que soit mises en ligne les données IRIS Analyse 1999, tout du moins celles issues de l'exploitation principale, à charge pour l'utilisateur de prendre les précautions qui s'imposent du fait des changements de définition, voire de découpage, pour toute comparaison.

III.1.7. Des tableaux plus détaillés pour l'échelon France (et région ?) affichés en ligne

Pour certaines recherches relativement simples mais fines, il conviendrait de disposer de données affichées à l'échelon France. La conception d'une déclinaison des tableaux détaillés avec une option de modalités « très détaillées » à l'échelon France, ou France et régions, permettrait de donner accès rapidement à des informations utiles (ex PCS détaillées, nationalités détaillées). Il s'agirait par exemple de créer une quatrième tranche en complément des trois existantes pour les tableaux détaillés (2 000 à moins de 10 000, 10 000 à moins de 50 000, $\geq$ 50 000) ce qui ne modifierait qu'à la marge la logique de diffusion mise en œuvre. Il conviendrait de vérifier que les tableaux détaillés existants, mais déclinés avec des modalités plus détaillées, correspondent bien à ce qui est souhaité par les utilisateurs à l'échelon France.

III.1.8. Enrichir l'analyse statistique des utilisateurs par l'accès à des données de référence ou en permettant d'effectuer facilement des comparaisons

Ce point est plus spécifiquement une question d'organisation que de contenu, il sera donc repris plus loin, mais il touche de très près à la donnée. L'Insee a créé, à la demande des utilisateurs, des fichiers d'indicateurs des Chiffres clés pour les départements et régions. Il serait souhaitable de les fournir aussi pour les communautés d'agglomérations et communautés urbaines et éventuellement les communautés de communes. Ceci serait d'autant plus nécessaire en 2013 où l'intercommunalité constituera une partition complète du territoire avec un seuil supérieur à 5 000 habitants hormis quelques exceptions.

De façon plus générale, l'obtention de données de référence pour comparer la zone affichée par l'utilisateur pourrait être développée. En particulier l'affichage de tranches de taille de communes, d'EPCI, d'unités urbaines, est totalement absent des fonctionnalités offertes et réclame un travail important pour créer les indicateurs équivalents aux chiffres clés pour situer dans la tranche correspondante le territoire observé.

III.1.9. Compléter certains fichiers détails de quelques données

La Fnau (Fédération nationale des agences d'urbanisme) a remis à l'Insee des propositions pour compléter certains fichiers détails afin de mener des analyses aujourd'hui impossibles compte tenu de l'organisation des variables fournies les deux premières années. Il convient de le faire dans le cadre des consignes édictées par la Cnil et en tenant compte des contraintes statistiques.

De même, les responsables régionaux en charge de la formation et de l'emploi font part de leur difficulté d'utilisation du recensement pour leurs travaux. En effet les modalités des variables disponibles à l'échelon communal ne permettent pas de mettre en œuvre des diagnostics dans le

domaine de la formation et de l'emploi. Seul l'échelon régional est fourni en ligne avec beaucoup de détail, or les politiques sont mises en œuvre par la région à un échelon infra régional. La commande de tableaux sur mesure, qui constitue un palliatif, n'apporte pas totalement satisfaction en raison des contraintes de diffusion et des coûts, alors qu'il s'agit d'un sujet commun à tous les acteurs régionaux de ce domaine d'action publique. Il conviendrait de mener un diagnostic pour trouver la meilleure manière de progresser dans la satisfaction de ces utilisateurs du domaine public de compétence régionale.

III.1.10. Pouvoir calculer une densité à l'infra communal

Il a été jugé souhaitable par un membre du groupe d'avoir, comme pour la commune, la superficie de l'IRIS, ce qui permet des calculs de densités.

III.2. FACILITER L'APPROPRIATION DE L'OFFRE PAR TOUS LES UTILISATEURS EN DÉVELOPPANT L'INTERACTIVITÉ DU SITE OFFERTE PAR LA TECHNOLOGIE INTERNET

III.2.1. Développer l'interactivité dans l'accès aux données pour accompagner l'internaute dans sa recherche de données

Concernant l'accès aux données ou à la documentation, il conviendrait de développer l'interactivité selon le principe suivant : guider l'internaute non expert ou non familier du site dans sa recherche par un questionnement pas à pas. Pour les autres catégories d'internautes, conserver un accès plus direct aux produits et à la documentation.

III.2.1.1. Permettre une recherche multicritère, géographique, thématique, datée, par type de produits

Il s'agit de s'appuyer sur le pavé déjà disponible en bas à droite de la page d'accueil du site mais de le rendre plus visible, d'en faire l'entrée principale « évidente » d'accès aux données pour tout nouvel utilisateur.

Multiplier les critères d'accès, alors qu'actuellement seul un accès par type de produits est proposé : 3 entrées supplémentaires devraient être proposées pour une recherche multicritères avec au minimum un seul critère obligatoire sans ordre de précellence : la géographie, le thème, et l'année. On garderait néanmoins pour l'année un affichage par défaut du dernier millésime.

III.2.1.2. Compléter la recherche géographique

Conserver l'accès géographique par saisie directe du nom du territoire recherché en complétant par l'occurrence systématique de l'établissement public de coopération intercommunal (EPCI à fiscalité propre) englobant la commune recherchée y compris pour les cas où l'appellation de l'EPCI ne contient pas le nom de la commune centre, ce qui actuellement pose des problèmes d'affichage.

Offrir aussi la possibilité d'une sélection pas à pas questionnant d'abord sur le type de zone recherché (ex. « département », « commune ») puis fournissant une liste nominative pour le niveau retenu (type

module de recherche géographique pour les données du RP99). Le niveau EPCI (Établissement public de coopération intercommunal) sera intégré ici comme un niveau géographique à part entière, éventuellement dans ses diverses déclinaisons (communauté urbaine, communauté d'agglomération, communauté de communes, etc.). Ce type d'entrée pourrait être complété par des modalités de tranches de taille (communes, EPCI, unités urbaines, ...).

III.2.1.3. Mémoriser l'interrogation antérieure

La mémorisation des interrogations antérieures donnerait un confort au moment d'une connexion sur le site. La logique de diffusion retenue ne permet pas comme en 1999, de conserver la construction du territoire composé en ligne. Cette fonctionnalité avait été fort appréciée. Le principe du téléchargement d'un fichier standard reporte la question chez l'utilisateur dans la version actuelle.

III.2.1.4. Préciser le contenu de chaque thème

Il convient de donner un peu de détail sur le contenu du thème par un affichage interactif, pour que l'internaute s'assure que l'information recherchée a de fortes chances de se trouver dans le thème qu'il s'apprête à sélectionner.

III.2.1.5. Faciliter l'accès aux cartes par un accès direct à partir d'un choix dans une liste d'indicateurs

Lister les indicateurs que les cartes permettent de visualiser et les échelons géographiques disponibles pour chaque indicateur cartographié, si possible donner un accès direct correspondant à l'indicateur et l'échelon géographique sélectionné.

L'éloignement du site de cartographie pour afficher les résultats du RP constitue une pénalité pour la mise en valeur de cette offre. Toute solution permettant un rapprochement serait à explorer.

III.2.1.6. Permettre à l'internaute perdu de se réorienter

Offrir des solutions de réorientation vers le site Insee ou vers d'autres produits du recensement. En effet le champ des informations couvertes par le recensement n'est pas forcément connu, par exemple des élèves ou des étudiants, ou même d'autres catégories d'internautes, il convient donc d'en tenir compte en cas de recherche infructueuse.

Tout cet ensemble de propositions contribue à refondre sensiblement la page d'accueil pour se rapprocher de la démarche de recherche de la majorité des internautes en les guidant pour leur favoriser l'accès à l'information, tout en conservant un accès direct par type de produits pour les utilisateurs experts et familiers du site recensement. Le point essentiel est de permettre la recherche d'une information précise sur une zone précise.

III.2.1.7. Introduire la possibilité de sélection d'un territoire de comparaison

Avec une procédure équivalente, les données livrées à l'issue de la recherche seraient produites parallèlement pour le territoire étudié et le territoire choisi comme référence parmi les échelons géographiques supérieurs et des modalités de territoire par strate démographique. Leur affichage simultané reste à imaginer.

III.2.1.8. Rendre plus facile l'accès à l'infra communal en proposant l'affichage de quelques indicateurs

L'échelon IRIS devrait être proposé au même titre que tout échelon géographique. Compte tenu des précautions d'usage liées à la précision des résultats à cet échelon, quelques données essentielles en nombre très limité pourraient être fournies en affichage immédiat.

III.2.1.9. Accentuer la visibilité de la page « calendrier/produits »

Mettre beaucoup plus en valeur et mentionner plus fréquemment la page « calendrier/produits » du recensement présente sur le site insee. En effet l'information sur le calendrier de mise à disposition de chaque produit est très importante.

III.2.2. Apporter quelques améliorations à l'organisation de l'accès aux bases de données ou à leur contenu.

III.2.2.1. Modifier l'intitulé « données téléchargeables »

Réinterroger les différentes appellations pour être sûr de leur caractère explicite et non ambigu, par exemple « données téléchargeables » pourrait être remplacé par « Bases de données », toute l'offre étant constituée de données téléchargeables. De même, l'utilisateur ne perçoit pas d'emblée la différence entre chiffres clés et tableaux détaillés mais cela touche l'ensemble du site de l'Insee.

III.2.2.2. Organiser les fichiers « bases de données » mis à disposition

De nombreux utilisateurs ont réclamé la possibilité de télécharger des bases de données limitées à leur propre territoire, à la fois pour des raisons de volume et aussi pour limiter les traitements à effectuer sur de nombreuses bases. Par exemple un conseil général qui souhaite tous les « tableaux détaillés » de son territoire doit traiter plus de 80 bases sur la France pour en extraire son département à chaque fois. Le choix d'une partition régionale ou départementale de fichiers multi thèmes est à expertiser. Cela ne doit en aucun cas remplacer la livraison actuelle sur tout le territoire

mais pourrait venir en complément. La solution la plus aboutie serait la possibilité de sélectionner un extrait par requête sur la thématique et/ou le territoire.

III.2.2.3. Proposer des d'autres formats pour les fichiers détails.

Le format Beyond a satisfait une partie non négligeable des utilisateurs dans les années antérieures, l'abandon de ce format pour faciliter le traitement des fichiers détails n'est pas compris de cette catégorie qui a fait l'investissement correspondant et en apprécie les avantages. Éventuellement, les fichiers partitionnés ci-dessus pourraient être aussi servis en format beyond si les utilisateurs sont encore potentiellement nombreux. Le format XML est aussi à privilégier à l'avenir (voir aussi ci-dessous III.2.2.5). La cohérence sur les offres multiples de formats concerne l'ensemble du site Insee.

III.2.2.4. Introduire davantage de zonages pour les bases de tableaux détaillés

En effet, les bases de données « Chiffres clés » proposent la plupart des zonages standards à l'intérieur du tableau ; pour les tableaux détaillés seul département/commune figure, ce qui oblige à de nombreuses manipulations pour réintroduire les échelons standards. Il y a évidemment un équilibre à trouver entre volume et contenu en tenant compte aussi des versions les plus utilisées de logiciels permettant d'afficher les données.

III.2.2.5. Une technologie adaptée aux modes actuels de services de données, « les web services »

Pour certains utilisateurs, par exemple les gestionnaires de SIG (systèmes d'information géographique) et beaucoup d'autres professionnels, il convient de s'orienter vers un système d'ouverture de « web services » qui permet de maintenir les données chez le fournisseur de données, lesquelles sont appelées à chaque utilisation avec toutes les métadonnées liées.

III.2.2.6. Fournir chaque année une information complète en ligne sur les modifications du dessin d'enregistrement des fichiers, des dictionnaires de variables...

Il faut pouvoir maîtriser les versions offertes avec leurs différences, un pas a déjà été franchi dans ce sens avec les « actualités » de la page d'accueil. Peut-être faut-il compléter cette offre et la rendre plus interactive ?

III.2.2.7. Réviser l'accès et l'organisation de l'offre sur les flux et la mobilité

Il serait souhaitable de revoir l'organisation des produits offerts sur ce sujet. En effet, la visibilité de l'offre actuelle n'est pas claire, un effort d'organisation et/ou d'explication doit être fait pour bien préciser le contenu des différents produits relatifs à la mobilité. Peut-être faut-il les modifier ? On peut d'ailleurs s'orienter vers une requête avec la commune (ou une intercommunalité) d'arrivée ou de départ aboutissant à un affichage en ligne avec un complément constitué des livraisons de bases de données, au moins pour les flux domicile-travail à l'échelon communal. La solution d'affichage cartographique peut aussi constituer une solution.

III.2.2.8. Rendre concomitante la version actualisée des contours numérisés des IRIS à la diffusion des données à l'IRIS.

Plus d'une année après la diffusion des données, aucun produit n'est disponible fournissant les données vectorielles des contours des IRIS permettant d'intégrer ces contours dans un système d'information géographique. Si l'ensemble de la diffusion du recensement a été une réussite avec des délais annoncés strictement respectés, ce point, qui touche une partie des utilisateurs professionnels, doit être mentionné comme tout à fait défaillant et nuit à la valorisation des résultats. Il est souhaitable que la coopération annoncée avec l'IGN trouve une issue favorable et permanente intégrant rapidement les mises à jour en régime de croisière à des conditions financières raisonnables.

Notons aussi que les plans images des IRIS des communes découpées en IRIS de moins de 10 000 habitants ne sont toujours pas affichés alors que les données statistiques sont disponibles depuis plus d'un an.

III.2.3. Développer l'interactivité entre la donnée et la métadonnée afin de rapprocher les métadonnées des données

III.2.3.1. Rapprocher les métadonnées des données

S'agissant de raccourcir la distance entre la donnée et sa définition, au passage du pointeur sur le titre d'un tableau et/ou sur les entêtes de ligne ou de colonne, imaginer par exemple l'ouverture d'une fenêtre avec présentation

- d'une définition littérale courte et du lien vers la définition complète correspondante,
- d'une alerte en cas de changement de concept et/ou de définition par rapport à 1999 (âge, activité, emploi, nationalité) et du lien avec la « fiche conseil thématique » correspondante.

III.2.3.2. Réinvestir la proposition n° 13 du rapport « Giblin » sur la qualité des données

Le groupe qui a fonctionné avant la diffusion des résultats avait suggéré le principe d'un code couleur selon le niveau de précision des données. Afin de ne pas trop alourdir le visuel, seules les données jugées les plus fragiles pourraient être repérées par exemple en rouge ou orange selon leur degré de fragilité.

Une voie de progrès ou même d'excellence serait de permettre de calculer et d'afficher un intervalle de confiance à 95 % en relation avec une donnée ou un ensemble de données sur la zone affichée.

III.2.3.3. Rechercher une solution simple pour traiter la question des arrondis

Les questions sur les arrondis ont été fréquentes, il serait sain de rechercher une solution pour diminuer ces questions et surtout rassurer l'internaute inquiet des totaux « inexacts ». Au minimum une question dans la foire aux questions ou une rubrique visible de la documentation. L'idéal est peut-être une alerte sur l'intitulé « Total ».

III.2.3.4. Donner un accès plus facile à la géographie de l'infra communal .

Il serait souhaitable de mieux signaler la disponibilité des contours IRIS, de construire un lien plus ergonomique avec les contours IRIS, le fichier au format « pdf » disponible sur la composition par IRIS des communes (plan d'assemblage et plan détaillé de chaque IRIS). De même il sera utile de mentionner la possibilité de commande du fichier des contours numérisés à usage des professionnels.

III.2.3.5. Garantir l'identification du fichier téléchargé

L'export du fichier doit être accompagné de son identification précise y compris de sa version en cas de modification de sorte que la réutilisation ultérieure ne fasse aucun doute. Actuellement le nom du fichier y contribue largement mais il convient d'anticiper des corrections effectuées après une première diffusion de sorte que la source des données soit identifiable sans aucune ambiguïté.

III.2.3.6. Accompagner les bases de données d'une métadonnée de description complète du contenu du fichier et des caractéristiques techniques

Un progrès a été fait dans ce sens en 2010 (données 2007) avec la fourniture des variables et modalités dans un fichier à part ; il faut poursuivre dans ce sens et fournir la même chose pour les tableaux détaillés en format txt.

III.2.4. Développer l'interactivité dans l'accès à la documentation pour améliorer l'efficacité de l'outil d'accompagnement

Proposer une recherche guidée pas à pas dans la documentation

La documentation est très appréciée des internautes ; en revanche, il est nécessaire là aussi d'accompagner l'utilisateur pour que la recherche soit plus aisée. Un questionnaire d'entrée guidée dans la documentation doit être imaginé. Ceci suppose, comme pour l'entrée dans les données par thème, de préciser avec un accès rapide (ex liste de mots) le contenu de la documentation que l'utilisateur s'apprête à ouvrir. Il y aurait aussi avantage à introduire des liens entre les différents univers d'aide à l'utilisateur (ex. définitions et fiches conseils thématiques, géographie et « aide » liée à la géographie).

Une suggestion d'organisation de la documentation a été fournie lors d'une réunion du groupe, l'Insee peut s'en inspirer pour enrichir l'accompagnement.

III.2.5. Apporter quelques compléments à la documentation

III.2.5.1. Mettre à disposition un outil équivalent aux anciens « fascicules gris »

Certains professionnels ont besoin de connaître les traitements effectués et le mode de construction des variables pour interpréter correctement certains résultats. Ces outils sont à construire pour chaque exploitation.

III. 2.5.2. Développer un outil d'accompagnement pour repérer rapidement les changements par rapport à 1999

Un outil d'accompagnement serait utile pour identifier rapidement les changements par rapport à 1999 (ex : liste des variables identiques ou liste des variables ayant subi des modifications, changement de concepts...).

III.2.5.3. Proposer un schéma synthétique interactif qui récapitule l'ensemble de l'offre

En complément du plan du site, ce schéma, qui existe en grande partie, pourrait être affiché en ligne pour avoir une vision complète de l'offre dans un document unique avec des liens.

III.2.6. Continuer de développer les Outils d'accompagnement en ligne - Aide et foire aux questions

III.2.6.1. Enrichir l'Aide

La rubrique « Aide » est constituée d'une série d'informations utiles mais dans des univers différents. Il convient de maintenir la séparation de ces univers et de donner directement accès à chacun pour atteindre plus rapidement l'information recherchée.

Les éléments de cette rubrique doivent être liés à la foire aux questions ou à la documentation de sorte que l'utilisateur perçoive la cohérence de l'accompagnement apporté.

La rubrique Aide devrait permettre de calculer des intervalles de confiance, sauf si cette possibilité est fournie en accompagnement direct des données. Le partage entre Aide et documentation est une question délicate, d'où l'intérêt de donner rapidement accès au sommaire détaillé de ces rubriques.

La foire aux questions (FAQ) est particulièrement succincte et généraliste, son maintien doit faire l'objet d'une expertise, d'autant plus si l'accès à la documentation est beaucoup plus guidé. Si le maintien est décidé, elle doit faire l'objet d'un enrichissement et de modalités spécifiques

d'actualisation. La coexistence avec la foire aux questions de l'ensemble du site insee.fr pose question. Au minimum, il doit y avoir des liens entre les deux, dans les deux sens. Par exemple la FAQ du site insee.fr a une question sur les changements de concepts entre 1999 et 2007, la réponse ne concerne que les aspects de population légale, aucun lien n'est fait sur les changements de concepts relatifs aux résultats statistiques.

III.2.6.2 Construire une aide pour l'ouverture des fichiers détails avec divers logiciels

Les utilisateurs d'un même réseau se communiquent les astuces pour l'utilisation des fichiers détails. Il serait souhaitable que l'Insee les réunisse pour compléter l'aide sur ce point au moins avec les logiciels les plus courants, et signale les points délicats sur lesquels la plupart butent. L'idéal serait de le faire dans un cadre permettant une logique interactive entre les utilisateurs.

III.2.7. Intégrer le recensement dans l'offre d'informations des données locales pour une offre complète et cohérente

Le fait que le recensement constitue une entrée à part est un confort pour les professionnels mais il y aurait de multiples avantages à une intégration future dans l'offre de données locales de l'Insee. Ainsi ceux qui connaissent mal le contenu du recensement seraient moins désespérés par la multiplicité de l'offre et la diversité des produits. En effet, pour eux, c'est l'information recherchée qui prime. Le problème est de ne pas « perdre » les habitués ou les experts. Cette réflexion vaut uniquement pour les résultats statistiques du recensement, tout ce qui concerne la population légale doit rester très identifiable.

Une telle orientation implique qu'une grande partie des propositions faites ci-dessus soient appliquées à l'ensemble des données locales. Cela représente un important chantier de moyen terme.

III.3. Les services complémentaires aux produits en ligne

III.3.1. Les produits sur mesure

Les produits sur mesure, en fait des tableaux, réalisés à la demande, sont des compléments indispensables à l'offre en ligne, car ils permettent de livrer une information non accessible avec les produits offerts sur internet. Leur caractère payant est justifié. Les modalités récentes de traitement du secret pour leur mise en œuvre provoquent des difficultés pour accéder à une information utile. En effet les contraintes sont telles que l'instruction de la demande pour arriver à un résultat utilisable réclame du temps et augmente la complexité. Il serait souhaitable de trouver des solutions sur ce point. Les chercheurs pâtissent dans leurs travaux de ces nouvelles dispositions. Leur cas est à traiter en priorité dans les diverses dispositions prises pour le monde de la recherche dans l'accès aux données publiques.

III.3.2. Le service Diaf - Diffusion infra communale de zones à façon

Ce service concerne les organismes ayant une mission de service public, comme il est précisé dans l'arrêté de diffusion du recensement. Il consiste à fournir des indicateurs équivalents à ceux qui existent pour les IRIS sur des zones de plus de 1 000 logements dont les contours sont définis par l'organisme public. Ce service est limité aux communes de 10 000 habitants et plus, sous réserve de la qualité des données concernant les zones définies. Délivré dans le cadre d'une licence, avec un tarif raisonnable, il constitue un complément appréciable à la diffusion infra communale, en particulier pour les opérations d'aménagement.

Une fois ce service expérimenté et avec un fonctionnement technique éprouvé d'une ou deux années, le groupe suggère que pour une liste d'indicateurs très limités (du type « îlots 15 » de 1999) ce service soit étendu, à des conditions financières à préciser, aux entreprises ou associations, par exemple ayant un identifiant Siren. En effet les besoins des entreprises pour les implantations locales en milieu urbain sont importants et les limites des IRIS sont parfois inadéquates. De même, les chercheurs devraient pouvoir accéder à ce service lorsque leurs travaux le requièrent. La consultation de la Cnil sur l'extension éventuelle du service à ces nouveaux publics est à inscrire compte tenu des règles actuelles qu'elle a fixées.

III.4. Autres actions d'accompagnement de la diffusion du recensement

III.4.1. Une formation pour les experts ou chargés d'étude et pour l'université à pérenniser

Les partenaires des directions régionales de l'Insee ont été servis ainsi que quelques partenaires nationaux comme il a été indiqué plus haut au cours de l'année 2009 en particulier. Cependant il faut imaginer un service stable qui aille au-delà de ce public. Des entreprises et organismes divers ont des besoins. Le module proposé par le CEPE, organisme de formation lié à l'Insee, est susceptible de bien répondre à cette demande. Il faut que cet organisme ait les moyens d'en faire la promotion pour que tous les utilisateurs potentiels en aient connaissance, et que son maintien soit ainsi assuré.

Il est aussi souhaité que les enseignants universitaires qui constituent un relais important pour la formation aient pour assurer celle-ci de façon optimale tous les éléments concernant l'utilisation des données du recensement. Pour ce faire, il conviendrait de renforcer sur ce sujet les rapprochements entre le milieu universitaire et l'Insee.

Le CNFPT (Centre national de formation du personnel territorial) a défini une offre pour l'Île-de-France (1^{ère} couronne) sur ce sujet de l'utilisation des résultats statistiques. Les candidats ont été peu nombreux au moment où les actions de l'Insee étaient intensives pour accompagner le recensement. Il convient de s'assurer qu'en régime de croisière la formation des techniciens et ingénieurs des collectivités qui ont à travailler pour la première fois sur le recensement trouvent bien l'accompagnement nécessaire.

III.4.2. Des modules pédagogiques en ligne

Le site de l'Insee n'a pas de développement important dans ce domaine, le recensement pourrait être un des points d'application au cas où une orientation dans ce domaine serait mise en œuvre à moyen terme.

III.4.3. Un système de réponse réactif

Ceci concerne principalement une interrogation directe de l'Insee par un expert via le service de réponse en ligne « Insee Contact ». Comment éviter l'envoi de questions et de réponses qui prennent du temps pour avoir la réponse souhaitée ? Comment arriver à faire formuler à l'utilisateur expert dès le départ une version détaillée de la question, témoignant de son niveau d'expertise, pour éviter les itérations qui créent de l'insatisfaction car ils accroissent les délais ? L'Insee devrait rechercher des pistes de progrès sur ce point reconnu comme particulièrement délicat.

III.4.4. Une nouvelle évaluation dans quelques années

Le groupe a fonctionné un peu comme un club d'utilisateurs.. Peut-être conviendrait-il réaliser une nouvelle enquête de satisfaction postérieurement aux modifications. Indépendamment de cette opération limitée, un nouvel examen global de la diffusion peut être mené par un groupe équivalent à celui-ci au-delà d'une période minimale de cinq ans pour la prise en compte de données d'évolutions dans les nouvelles modalités de recensement, c'est-à-dire après la diffusion des résultats 2011 qui se fera en 2014. D'ici là, des rencontres thématiques avec une dimension méthodologique pourraient vraisemblablement rencontrer de l'intérêt.

* * *

•

Annexes

Mandat du groupe

Mandat pour le groupe de travail du CNIS

« Diffusion des données du recensement de la population »

Le premier cycle de collecte du recensement (2004-2008) étant achevé, la diffusion complète des données de recensement débutera en juillet 2009 et portera sur des données 2006. Un sous-groupe de travail du Cnis a été constitué en 2005 pour préparer cette diffusion. Ce sous-groupe, qui a tenu 5 réunions, a permis l'élaboration d'un plan de diffusion tenant compte des demandes et des besoins des utilisateurs. Les grandes lignes de ce plan de diffusion ont été présentées lors de la rencontre moyen terme du Cnis consacrée, le 19 septembre dernier, aux statistiques régionales et locales à l'horizon 2013. Le contenu précis et détaillé de ce plan a été rendu public en début d'année 2009, la diffusion proprement dite commençant le 1^{er} juillet 2009, avec l'ouverture d'un site qui lui sera consacré.

Au cours de cette rencontre moyen terme, les utilisateurs ont manifesté un fort besoin d'accompagnement de la part des producteurs de l'information. Si les données du recensement n'ont pas changé de nature, elles s'appuient désormais sur une nouvelle méthode de collecte et de traitement qui amène à un dispositif de diffusion nouveau, qui va affronter l'épreuve des faits et que les utilisateurs doivent s'approprier.

Afin d'assurer au mieux l'efficacité de cette diffusion, l'assemblée plénière du Cnis dans son avis moyen terme 2009-2013, a souhaité la création d'un groupe de travail, dans la continuité du sous-groupe constitué en 2005. Ce groupe de travail a deux objectifs principaux :

- s'assurer que le dispositif de diffusion est pertinent et correspond bien, de fait, aux besoins des utilisateurs. Il s'agit, d'une part, de vérifier a posteriori la pertinence du dispositif, d'autre part, de proposer les changements qui devront être pris en compte pour la diffusion de 2010, portant sur les données 2007.
- examiner le dispositif d'appropriation prévu au plan des outils (métadonnées, notes de méthodes, outils de « formation », etc.) et des structures (rôle des directions régionales de l'Insee, des associations de collectivités territoriales, des institutionnels régionaux, des relais nationaux, etc.) et proposer les améliorations souhaitables.

Composition du groupe

LISTE DES MEMBRES PERMANENTS

Pierre AUDIBERT	Insee – Direction de la diffusion et de l'action régionale
Marcel BELLLOT - <i>Président</i>	Fédération nationale des agences d'urbanisme
Marie-Hélène BOULIDARD- <i>Rapporteur</i>	Démographe territorial La Roche-sur-Yon - Consultante
Camille BOURON	Communauté d'agglomération de La Rochelle
Gilles BOUTONNET	Bnp Paribas
Sophie BURDIN	Mairie de Saint Etienne du Rouvray
Thierry CHALUMEAU	Opinionway
Philippe CLAIRET - <i>Rapporteur</i>	Insee - Info service
François CLANCHÉ	Insee – Direction des statistiques démographiques et sociales
Marie-Hélène COUSIN	Agence d'urbanisme de la région havraise
Jean-Philippe DAMAIS	Mairie de Fontenay aux roses
Frédéric DELBOS	Aucame, agence d'urbanisme de Caen métropole
Françoise DUPONT	Insee - Direction des statistiques démographiques et sociales
Thérèse FERRE	Insee - Direction de la diffusion et de l'action régionale
Anne FOUSSAT	Mairie d'Aubervilliers
Laurent FRAPPART	Leroy merlin France
Pascal FUCHS	Conseil général des Hauts de seine
Sylvie GILLOUX	Mairie de Nantes
Jacques JACOB	Insee – Direction des statistiques démographiques et sociales
Marie-Hélène KEROUANTON	Datar
Etienne KOSIANSKI	Mairie de Gennevilliers
Alexandre KYCH	CNRS
Jean-Luc LIPATZ	Insee – Direction de la diffusion et de l'action régionale
Hervé MARTIN-QUINQUIS	Mairie de Sevrans
Lisa MERLOT	Mairie de Nanterre
Brigitte MICHEL	Insee - Direction régionale des Pays de la Loire
Franck OSSORO	Kiabi - Développement & immobilier
Edmond PRETECEILLE	Observatoire sociologique du changement
Stéphane RASSE	Géoconcept SA
Phouné THIAM	Mairie de Montreuil

Dates des réunions et comptes rendus sur le site du Cnis :

[Accéder à la rubrique concernée sur le site du Cnis](#)

Quelques remarques sur la diffusion des produits du recensement rénové de la population : le point de vue du chercheur. Note de M. Alexandre Kych - réseau QUETELET
- le 4 mars 2011.

Quelques remarques sur la diffusion des produits du recensement rénové de la population :

le point de vue du chercheur

L'offre de données peut se résumer ainsi :

Des produits standards

- 2 bases de données interrogeables en ligne
- 4 bases de données téléchargeables (des tableaux pour des zones géographiques)
- 8 fichiers détail (des fichiers d'individus)

Des produits à façon

- Les tableaux à façon (les PSM, Produits Sur Mesure)
- Les zonages à façon (ZAF)
- L'accès aux CASD (Centres d'Accès Sécurisés aux Données)

Dans ses grandes lignes, l'offre actuelle ne diffère pas sensiblement de l'offre au RP99, aux ZAF et CASD près qui n'existaient pas encore. On verra qu'il n'en est plus de même dans le détail.

Les 2 bases de données interrogeables en ligne

On les laissera de côté. D'abord parce qu'on retrouve les mêmes données dans 2 des 4 bases de données téléchargeables. Ensuite parce que ce sont des produits grand public et qu'ils n'ont pas vocation à répondre à des demandes complexes. Ce qui n'empêche les chercheurs d'y recourir et d'en apprécier la convivialité.

Les 4 bases de données téléchargeables

Ce sont des collections de tableaux à la commune ou à l'IRIS

- Chiffres-clés
- Tableaux détaillés
- Données infracommunales
- Flux de mobilités

Le nombre de collections a diminué par rapport au RP99 où il y en avait 8. Ont disparu les tableaux :

- Vie politique : Circonscriptions législatives
- Cantons

- Politique de la ville : Zus
- Localisation très fine : Îlots
- Rétrospectifs : Profils

Références

On soulignera que les cantons et les circonscriptions législatives ne peuvent être reconstruits à partir des communes et des IRIS et qu'il y a là un véritable problème pour les politistes, dont les séries historiques s'arrêteront en 1999. C'est d'autant plus regrettable que l'INSEE, à la demande des

chercheurs, avait reconstitué des tableaux « Cantons » pour les recensements anciens qui n'en n'avaient pas.

Pour les zones de la politique de la ville, on regrettera la très grande richesse des tableaux « ZUS » du RP99 pour lesquels l'INSEE avait réalisé un effort remarquable dans les thématiques et la construction des données. On ne comprend pas pourquoi ces tableaux n'ont pas été prolongés. En revanche, on trouve d'autres données dans les pages des « données urbaines infra-communales par quartier », des données nouvelles et souvent issues d'autres sources que le recensement.

On laissera de côté les « chiffres clés » et les « flux de mobilités ». Les premiers étant destinés au grand public et les seconds étant avantageusement remplacés par les fichiers détail Individus Mobilités. En revanche, on analysera ensemble les tableaux « détaillés » et les « données infracommunales », tant les 2 collections apparaissent comme un tout aux yeux des chercheurs.

Les tableaux « détaillés » et les « données infracommunales »

Ces tableaux sont souvent moins détaillés : moins de variables croisées par tableau, et moins de modalités par variable. Quant aux IRIS, ils sont décrits par 16 fois moins de tableaux.

C'est un véritable problème dans la mesure où les IRIS ont été créés pour être les briques élémentaires de tous les zonages. Prendre des précautions pour éviter des utilisations erronées de petites zones géographiques ne doit pas empêcher de constituer des zones au gré des chercheurs en regroupant communes et IRIS. Or les données disponibles à l'IRIS sont à la limite de la compatibilité avec les données disponibles à la commune.

Dès à présent, on rencontre le même problème qu'avec les tableaux « Analyses » du RP99. La CS en 8 postes est tout à fait insuffisante et tous réclament la CS en 42 postes. Sans doute, la majorité des utilisateurs se satisfait pleinement de la CS en 8 postes et l'INSEE a raison de la proposer des ses tableaux. Mais les tableaux « Analyses » du RP99 étaient souvent déclinés en 2 versions, simple et détaillée, sans compter les tabulations sur les fichiers détail principal et complémentaire. Cette démultiplication des versions d'un même tableau en a rebuté plus d'un. Mais, sans aller jusqu'à là, ne peut-on envisager des versions superdétaillées pour la dizaine de tableaux qui mettent en jeu la CS, versions qui seraient considérées comme des tableaux standard et non comme des tableaux à façon, versions qui pourraient être diffusées via Quetelet pour éviter de gêner le grand public ? On peut tenir un raisonnement semblable pour les secteurs d'activité en 38 ou 88 postes.

Reste le problème des tabulations sur les variables sensibles. Au RP99, l'INSEE a doublé le nombre de tableaux sur variables sensibles, à la demande de Quetelet. C'est dire combien ces tableaux sont demandés par les chercheurs. Mais avec les « tableaux détaillés » le nombre de tableaux sur variables sensibles a diminué, leur contenu s'est considérablement restreint et les nomenclatures ont été réduites au minimum : français/étranger ou immigré/non immigré. À vrai dire, ce ne sont même plus des tableaux sur variables sensibles. Certes, cela résout le problème des zones géographiques compatibles avec les variables sensible : toutes les communes sont décrites. Mais pour quel résultat ? Des tableaux de type « sexe*indicateur de nationalité en 2 postes*1 autre variable regroupée en 4 à 6 postes ». Il faut absolument proposer une série de tableaux sur de véritables variables sensibles et pour des zonages géographiques adaptés (TRIRIS, communes, arrondissements, départements, UU, AU et ZE). C'est la solution qui avait été retenue au RP99. Là encore, il serait judicieux que ces tableaux soient considérés comme des tableaux standards.

Les 8 fichiers détail

- Logements
- Individus par régions
- Individus par pseudocantons (ou villes et cantons)
- Individus mobilités domicile-travail
- Individus mobilités domicile-étude
- Individus migrations domicile-domicile antérieur à la commune
- Individus migrations domicile-domicile antérieur au département

Individus migrations domicile-domicile antérieur pour les grandes communes

Contrairement aux bases de tableaux, le nombre s'est accru, passant de 3 à 8. Le fichier détail Logement a été grandement amélioré avec l'introduction de variables sur le ménage. Les fichiers détail Individus Mobilités ou Migrations sont sans équivalents au RP99. Tout ceci est très positif et présente un véritable progrès.

Les 2 fichiers Individus par régions et par pseudocantons

Ils sont complémentaires, en proposant soit un détail des variables sensibles (localisation et origine), soit un détail des autres variables. C'est une solution bien plus satisfaisante que celle retenue au RP99.

Le fichier Individus par régions tient ses promesses avec des nomenclatures détaillées : profession en 786 postes, branche d'activité en 732 postes et origine en 49 postes.

Le fichier Individus par pseudocantons pourrait être amélioré. Peut-être les nomenclatures réduites sont-elles trop réduites (CS en seulement 8 postes et branche d'activité en 17 postes) et l'on peut regretter que l'origine ait quasiment disparu.

Le zonage est déroutant : pseudocantons et IRIS de plus de 200 habitants. Au final, on aboutit à un zonage dont la maille est très déséquilibrée selon que l'on est en ville ou à la campagne. Et surtout, il est quasiment impossible de regrouper les zones pour retrouver les zonages d'études usuels (UU, AU et ZE).

Point qui peut être important pour des travaux avancés : ces 2 fichiers Individus ont perdu la structuration physique des individus en familles et ménages. Faute d'identifiants, il est impossible de réexplorer les individus d'un même ménage pour créer de nouvelles variables de famille ou de ménage.

Enfin, il faut signaler que le zonage communal intégral est absent des 2 fichiers (cf plus haut). Il en résulte que ces fichiers ne peuvent servir pour compléter les bases de tableaux par des tabulations que le chercheur réaliserait lui-même. Dans ces conditions, la seule solution possible est-elle de recourir aux tableaux à façon.

La possibilité de tableaux à façon (les PSM, Produits Sur Mesure)

Il s'agit là d'un point important et qui est devenu crucial avec le nouveau recensement, dans la mesure où, dans beaucoup de cas, ni les tableaux standards, ni les fichiers détail standards, ne peuvent répondre aux interrogations des chercheurs (cf plus haut).

Jusqu'en 2009, pour tous les RP de 99 à 62, on pouvait demander n'importe quel tableau, quelque soit le nombre de variables et de cases, à condition de respecter les règles de diffusion des variables d'origine. Cela a permis de fabriquer des versions très détaillées des tableaux standards du RP99. La procédure fonctionnait à la satisfaction de tous.

Depuis 2010, pour les RRP comme pour les RP antérieurs, de nouvelles règles ont été mises en place.

- 1) On ne peut tabuler que les variables retenues dans une liste établie par l'INSEE.
- 2) Les cases correspondant à 1 à 3 unités (individus, familles, ménages, logements) sont mises au secret (codées selon un code sans réponse).

Ce n'est sans doute pas la première contrainte qui est la plus gênante, même si on peut se demander pourquoi le pays de naissance est codé en 12 postes, alors que la nationalité l'est en plus de 200 postes. En revanche, quand le code IRIS est absent de la liste (c'est le cas pour le RRP), on voit bien que cela revient à interdire toute tabulation à l'IRIS. Ce qui est un handicap majeur tant le zonage en IRIS est devenu une norme.

La deuxième contrainte est devenue infernale. La mise au point de l'architecture d'un tableau est devenue une épreuve qui peut prendre des mois, le temps de trouver la combinaison de variables et de modalités de variables qui satisfait à la règle tout en répondant quand même à la demande du chercheur.

Les conséquences sont lourdes

En effet, si l'on avait appliqué les nouvelles règles aux demandes de 2009, 9 sur 10 d'entre elles auraient été refusées. On est actuellement dans une situation de blocage où très peu de tableaux ont été demandés et qui est largement sous-estimée car les gens le savent et ne demandent pas.

Demander des tabulations sur l'ensemble des 50000 IRIS et petites communes de France est devenu tâche impossible

Même une variable unique tabulée pour la France entière, sans aucune distinction spatiale, peut être censurée, par exemple la nationalité.

Les démarches entre le demandeur, l'ADISP et l'INSEE pour mettre au point une demande prennent un temps considérable ... pour souvent arriver à un échec. Il faut bien voir que le contenu de certaines des tabulations auxquelles on aboutit in fine est quasiment trivial.

C'est une procédure totalement inadaptée à la façon de travailler des chercheurs qui procèdent par essais et erreurs et ne peuvent prévoir à l'avance la façon la plus opportune de construire leurs données.

Dans l'état actuel, les PSM apparaissent davantage comme un produit grand public que comme une procédure utile aux chercheurs. Ce n'était pas le cas précédemment, où elle a joué un grand rôle pour certains.

Une solution ?

On a évoqué la possibilité de PSM sur les recensements pour la recherche (PSM-R). Ce ne serait pas un retour à la situation antérieure. Le chercheur devrait déposer un projet devant le Comité du Secret qui statuerait et donnerait la possibilité de réaliser les tabulations demandées que le chercheur rapatrierait sur sa machine. La procédure est lourde et contraignante pour tous : le Comité du Secret, le chercheur et son institution.

Des suggestions

Ne pourrait-on pas distinguer les PSM qui ne font que détailler des tableaux standards et qui pourraient être tabulés comme auparavant, et les PSM qui constituent une demande vraiment particulière et qui relèveraient de la nouvelles procédure ?

Ne pourrait-on pas utiliser le critère de l'effectif des cases en conjonction avec la taille de la population concernée par la tabulation ou la grossièreté du zonage géographique employé, en distinguant zonage absolu (commune par exemple) et zonage relatif (tranche de taille de commune). Pour résumer, ne pourrait-on pas concevoir une règle plus subtile ?

L'accès aux CASD

Dans l'état actuel, les CASD ne pourraient sans doute pas répondre à l'afflux des demandes.

Quand un chercheur recourt aux PSM, c'est pour obtenir une base de données qu'il pourra utiliser chez lui et combiner avec toutes les données qu'il a déjà. C'est un objectif assez incompatible avec l'objet initial des CASD.

Les CASD et les PSM répondent à des besoins différents.

La possibilité de zonages à façon (ZAF)

C'est un service attrayant : mêmes données que pour les IRIS, avec ses contraintes (seuil de nombre de logements) et ses limites (seulement pour les communes de plus de 10000 hab., RIL oblige). Mais il est réservé aux institutions ayant une mission de service public. Dans ces conditions quelles chances un chercheur aura-t-il de pouvoir en bénéficier ? Quant à travailler dans le cadre des collectivités locales ou des ministères, c'est précisément ce que l'on voulait éviter il y a une dizaine d'années en créant le Centre Quetelet et en ouvrant un égal accès aux données pour tous les chercheurs.

Et les FPR (fichiers de production et de recherche) pour le recensement ?

L'INSEE les a annoncés il y a déjà plusieurs années. Pour l'instant, on en ignore tout. Cela pourrait être aussi un moyen de débloquer le problème des PSM.

Divers

La collection des fichiers annexes sur les zonages usuels et les IRIS (zonages et modifications) n'est pas évidente à trouver quand on parcourt les pages du site www.insee.fr, alors qu'elle se trouve sur les pages dédiées au RP www.recensement-2006.insee.fr. Il faudrait mieux annoncer cette adresse.

De même ne pourrait-on pas signaler les pages sur les données urbaines infra-communales par quartier (même si elles ne proviennent pas que du RP) à partir des pages sur les bases de données infra-communales ?

Alexandre Kych
Réseau Quetelet
Le 4 mars 2011

\$

La diffusion des résultats du recensement français depuis 2009 - Insee
Intervention à la Conférence des statisticiens européens - Nations unies - 1^{er}
juillet 2010 -



18.f.pdf

référence de ce texte en anglais :

<http://www.unece.org/stats/documents/2010.07.census.htm>).

L'enquête en ligne de juillet 2010 auprès des internautes qui ont visité le site
du recensement.

Les résultats de cette enquête peuvent être fournis par le CNIS, le document de synthèse comporte une vingtaine de pages et les résultats complets (questionnaire, tableaux, graphiques, réponses aux questions ouvertes) constituent un document de 130 pages. Un résumé figure dans le compte rendu de la réunion du groupe du 25 novembre 2010 (voir le site cnis.fr).

* * *