



Conseil économique et social

Distr. générale
14 mars 2012
Français
Original: anglais

Commission économique pour l'Europe

Conférence des statisticiens européens

Groupe d'experts des recensements de la population et des habitations

Quatorzième réunion

Genève, 24 et 25 mai 2012

Point 5 de l'ordre du jour provisoire

Collecte de données par Internet

Recensement de 2011 au Portugal – mise en œuvre du système de recensement électronique

Note de Statistique Portugal

Résumé

Le présent document décrit l'expérience portugaise de l'utilisation d'Internet (recensement électronique) pour remplir les questionnaires du recensement de 2011. Il fait suite au document présenté à la douzième réunion du Groupe d'experts des recensements de la population et des habitations qui s'est tenue du 28 au 30 octobre 2009 (<http://www.unece.org/stats/documents/2009.10.census.html>).

I. Introduction

1. Des recensements de la population et des habitations sont effectués au Portugal depuis 1864, conformément aux recommandations internationales. Jusqu'en 1960, il n'était tenu compte que du recensement de la population qui comportait seulement quelques questions concernant l'habitation. Depuis 1970, le volet consacré à l'habitation est plus étoffé et fait l'objet d'une présentation distincte, raison pour laquelle on en est venu à effectuer simultanément deux recensements (population et habitation). Les recensements organisés en 2011 (recensement de 2011) correspondent donc au quinzième recensement général de la population et au cinquième recensement général des habitations: ils ont été réalisés simultanément, étant donné les liens étroits entre les deux volets.

2. Le modèle de recensement retenu a toujours été le modèle traditionnel, dans la mesure où toutes les unités statistiques (bâtiment, local d'habitation, ménage et individu) sont observées sur une base individuelle et de façon exhaustive, au moyen de questionnaires également traités sous tous les aspects. Aucune méthode d'échantillonnage n'a donc été utilisée dans la collecte et le traitement des questionnaires. La seule exception s'est produite en 1970 en raison de problèmes rencontrés dans le système de traitement des données. Ainsi, la plupart des résultats pour 1970 sont basés sur un échantillon de 20 % des questionnaires recueillis.

3. Dès le lancement, en 2006, de l'opération de recensement de 2011, il a été décidé qu'un système de réponse en ligne serait mis en œuvre, compte tenu des possibilités désormais offertes par un tel système (accès plus aisé pour les citoyens et nombre croissant de personnes déjà initiées à ce type d'infrastructure technologique). En outre, l'expérience connue d'un certain nombre de pays tels que le Canada, la Nouvelle-Zélande et l'Australie s'avérait non seulement encourageante du point de vue de la participation de la population au recensement via Internet, mais montrait aussi clairement que, quel que soit le modèle de recensement adopté, l'utilisation d'Internet dans de telles opérations statistiques satisfaisait manifestement l'exigence de confort et de confidentialité de nombreux citoyens.

4. Les résultats obtenus concernant la participation individuelle au recensement de 2011 par Internet (environ 50 % de la population recensée) ont été très appréciables, voire surprenants, étant donné ceux qui avaient été observés lors des expérimentations. Ces résultats étaient dus à une forte mobilisation de tous les agents sur le terrain qui étaient suffisamment motivés pour promouvoir les réponses en ligne, notamment en raison d'un mode de rémunération ne faisant aucune distinction entre les types de réponses. Comme les réponses en ligne facilitaient le travail des agents recenseurs en réduisant le nombre de visites à effectuer dans les unités d'habitation, ceux-ci ont encouragé le public à répondre par Internet, soit individuellement, soit avec l'aide de membres de leur famille ou d'entités administratives locales.

II. Expérimentations

5. En prévision du recensement de 2011, trois expérimentations ont été menées: deux tests portant sur les questionnaires en 2008 et 2009 et une enquête pilote en 2010. Les expérimentations ont toujours été réalisées à une période de l'année correspondant au jour prévu pour l'opération finale de recensement, de façon à remplir quatre conditions clefs:

- a) Conditions météorologiques et de mobilité de la population analogues à celles de l'opération finale;
- b) Taille croissante et diversification de la population et des zones observées;

c) Répartition des zones à observer entre les principales régions (sept), pour que les diverses structures régionales chargées de l'opération finale puissent y participer et acquérir l'expérience nécessaire;

d) Délais raisonnables accordés pour établir des rapports sur chaque opération expérimentale et intégrer les résultats à l'opération suivante.

6. Les réponses au recensement de 2011 pouvaient être données en ligne (recensement électronique) ou au moyen de questionnaires sur papier, d'où la nécessité de fixer une date limite pour l'achèvement du recensement électronique et les contacts à établir avec tous les logements dont aucune réponse n'avait été reçue. Ainsi, un délai plus court a été accordé pour participer au recensement électronique, délai à l'expiration duquel une réponse ne pouvait être envoyée qu'en utilisant les questionnaires papier.

7. On trouvera dans le tableau ci-après un ensemble d'indicateurs concernant l'organisation des expérimentations et les principaux résultats obtenus.

Tableau 1

Expérimentation et année	Date du recensement	Taille (nombre d'unités d'habitation)	Période de collecte des données		Taux de réponse au recensement électronique	Temps moyen passé à remplir le questionnaire électronique	Rapport final
			Recensement électronique	Recensement électronique + papier			
Test 1 – 2008	7 avril	6 911	7 avril-4 mai	14 avril-4 mai	14 %	46 mn	Juillet 2008
Test 2 – 2009	20 avril	17 161	20 avril-10 mai	27 avril -24 mai	9 %	49 mn	Juillet 2009
Enquête pilote – 2010	12 avril	45 887	12 avril-2 mai	19 avril-16 mai	13,4 %	42 mn	Juillet 2010

8. L'expérience du recensement précédent est toujours un élément clef. Cependant, en raison du temps écoulé – habituellement dix ans – il n'est pas possible de se dispenser d'organiser un important travail d'expérimentation. Dans le cas des réponses en ligne, mis à part le fait qu'il s'agissait de la première expérience de ce type au Portugal, on ne disposait d'aucun registre préexistant des adresses des logements, ce qui représentait un facteur de complexité supplémentaire. Il a fallu opérer un nouveau contrôle des procédures pour identifier chaque réponse envoyée et acceptée par le système.

9. La phase préparatoire de cette opération statistique a constitué un élément essentiel de l'expérimentation des diverses activités, notamment l'évaluation continue des innovations qu'il a été jugé nécessaire de mettre en œuvre afin de garantir des taux élevés de participation au recensement électronique et un contrôle adéquat des réponses. Il est aussi apparu clairement que les taux de réponse au recensement électronique dépendaient dans une large mesure de la mobilisation locale des personnes intervenant sur le terrain. En outre, une bonne campagne de publicité était essentielle pour que les structures administratives locales (bureaux des municipalités et des paroisses civiles) contribuent à apporter leur concours à la population, en l'occurrence par la mise en place de «comptoirs pour le recensement électronique» dotés d'un équipement connecté à Internet et employant des personnes habilitées à aider les recensés. La collecte des données a été systématiquement limitée au recensement électronique pendant la première semaine lors des opérations expérimentales afin de promouvoir les réponses en ligne.

III. Brève description de l'organisation du travail sur le terrain pour la collecte finale des données

10. Depuis 1981, les recensements s'appuient, au Portugal, sur une importante infrastructure cartographique divisant le territoire en secteurs statistiques attribués à chacun des agents recenseurs. Une cartographie numérique de 50 centimètres de résolution a été utilisée en 2011, ce qui a permis de cerner plus facilement les limites de chaque construction sur le territoire concerné. En l'absence de registre préalablement établi des adresses postales des logements, les bâtiments et les logements situés dans chaque secteur ont été identifiés lors de la distribution du questionnaire au moyen d'un code spécial correspondant à la situation administrative et statistique des unités d'habitation.

11. Pour le recensement de 2011, la collecte des données a été organisée conformément aux critères suivants:

- a) Le jour du recensement a été fixé au 21 mars 2011;
- b) La remise des questionnaires a eu lieu pendant les deux semaines précédant la date du recensement, à savoir entre le 7 et le 20 mars. Au cours de cette distribution, l'agent recenseur a consigné l'emplacement des logements dans un dossier correspondant à chaque bâtiment et a laissé sur place une enveloppe contenant l'identifiant et le code confidentiel (ID/PIN) nécessaires pour répondre par Internet. L'identifiant et le code confidentiel ont été fixés de façon à pouvoir identifier la municipalité (308 pour l'ensemble du pays) dans laquelle était situé chaque logement. Tous les logements ont reçu les documents suivants:
 - i) Un questionnaire pour l'unité d'habitation, la partie destinée à l'identification étant dûment remplie par l'agent recenseur;
 - ii) Un questionnaire à l'intention des ménages;
 - iii) Des questionnaires individuels pour les résidents;
 - iv) Un pli fermé avec l'identifiant et le code confidentiel de la municipalité concernée, l'identifiant étant consigné à l'extérieur et le code individuel pour la réponse à l'intérieur;
 - v) Les habitations comptant plus d'un ménage et les logements collectifs n'ont pas reçu de code pour les réponses en ligne en raison de la complexité d'une telle procédure. Cependant, ce type de logements concerne moins de 1 % de la population recensée;
- c) Les questionnaires ont été distribués comme suit:
 - i) Dans la semaine du 21 au 27 mars, les réponses ont été acceptées uniquement en ligne et les agents recenseurs se sont attachés à recueillir les questionnaires dans les logements dont les occupants ne pouvaient envoyer leur réponse via Internet;
 - ii) Dans la période du 28 mars au 12 avril, il était possible de répondre soit en ligne, soit sur papier; cependant, les réponses en ligne ont été privilégiées pour les personnes intéressées; aucune distinction n'a été faite en matière de rémunération selon les différents types de réponses lors du recensement de 2011;
 - iii) Lorsqu'une réponse était envoyée de tel ou tel logement par voie électronique, l'agent recenseur concerné recevait sur son téléphone portable un message contenant l'identifiant correspondant;
 - iv) À partir du 12 avril, seules les réponses sur papier ont été acceptées.

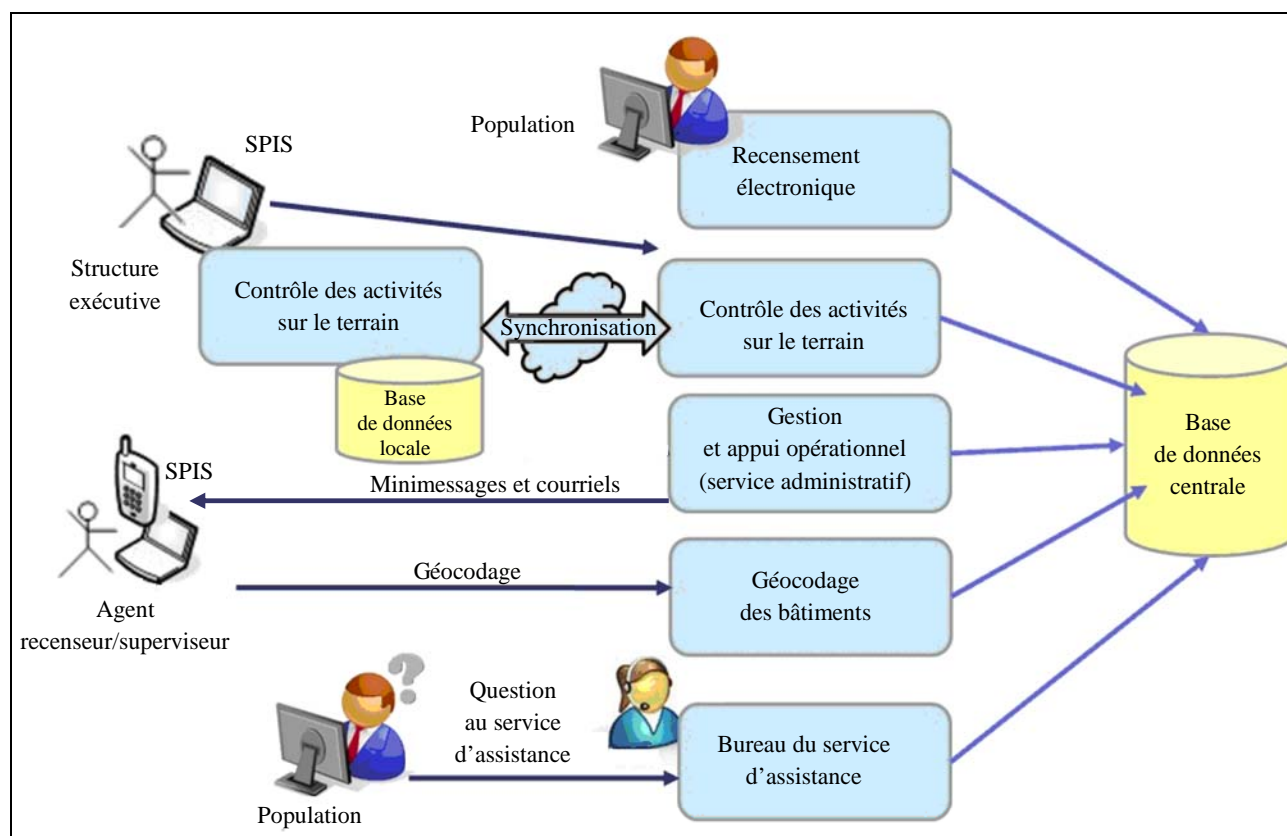
12. À la fin avril, 97 % environ des logements ayant reçu le questionnaire avaient été recensés. Comme il est fréquent pour ce type d'opération statistique, il a fallu, pour un petit nombre de logements, effectuer un travail considérable de vérification des données, tâche qui s'est prolongée jusqu'à la fin mai. Les premiers résultats du recensement de 2011 (résultats préliminaires) ont été communiqués à la fin juin 2011 et les résultats provisoires ont été publiés début décembre 2011.

IV. Organisation et fonctionnement de l'infrastructure technologique finale

A. Structure technologique de base

13. L'architecture d'appui au recensement de 2011 était composée des applications ou modules fonctionnels suivants.

Figure 1



Recensement électronique ou système de collecte par Internet: application permettant de communiquer le questionnaire électronique aux unités d'habitation, aux ménages et aux personnes concernées.

Contrôle des activités sur le terrain: fonction permettant aux équipes de terrain de saisir des données sur les activités réalisées, d'effectuer des tâches de soutien et de suivre le déroulement des travaux;

Gestion et appui opérationnel (service administratif): communication d'informations à la structure interne, définition des paramètres et configuration de l'ensemble de l'opération;

Permanence téléphonique (service d'assistance): service permettant de répondre aux questions et aux demandes d'assistance des recensés. Un site Web d'assistance technique a également été établi pour les équipes de terrain;

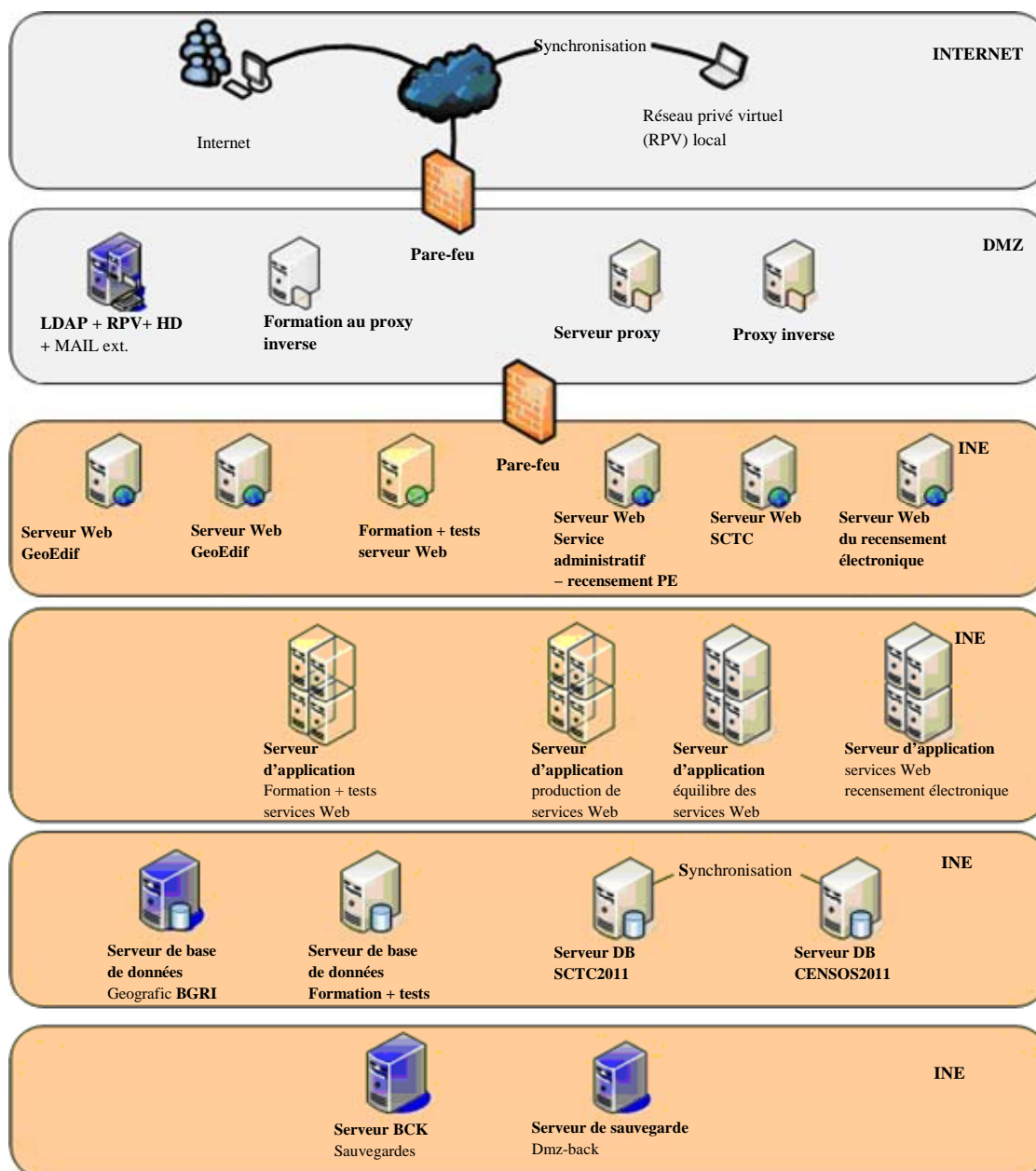
Géocodage des bâtiments: enregistrement des coordonnées des habitations recensées et des numéros respectifs des logements.

14. Plusieurs personnes ont été employées pour communiquer aux collaborateurs sur le terrain, notamment par des minimessages, des renseignements sur chaque ménage ayant rempli le questionnaire par voie électronique.

15. La conception de l'infrastructure technologique utilisée a graduellement été renforcée au cours des opérations d'expérimentation, ce qui a permis d'améliorer un certain nombre de caractéristiques fonctionnelles et technologiques.

16. Les différents éléments de l'infrastructure technologique ont été soumis à des audits indépendants visant à en évaluer le degré de sécurité et la fiabilité. Lorsque l'équipement de production a été installé, l'ensemble de la structure a subi des tests de sécurité et de résistance. En raison de son importance et de son rôle crucial, l'application prévue pour le recensement électronique, destinée à l'utilisation la plus intensive, est celle qui a fait l'objet du plus grand nombre d'essais.

Figure 2

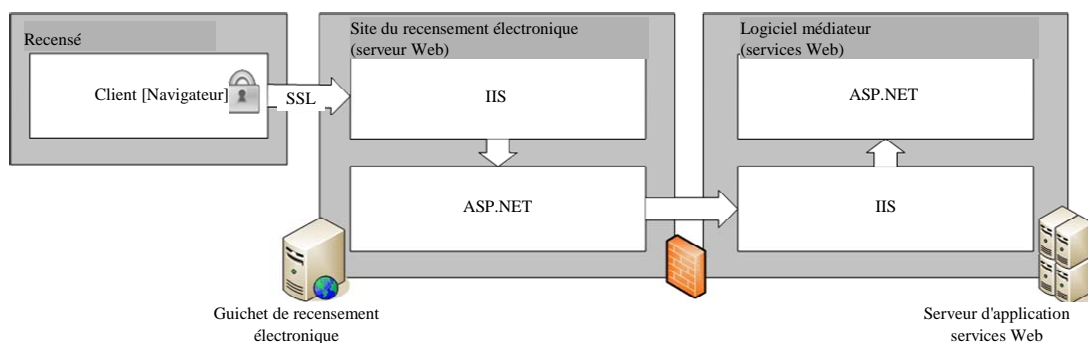


17. Pendant la période de fonctionnement de l'infrastructure technologique, un appui informatique a été fourni aux différentes composantes du système. Qu'il s'agisse de matériel, de logiciel ou de communication, l'assistance a été renforcée et assurée suivant des horaires élargis pendant la période d'envoi des réponses par voie électronique.

18. Le schéma ci-après indique de manière sommaire les modalités d'accès au site Web de recensement électronique, d'authentification dans le système et de circulation des informations. Les recensés ont pu accéder au site Web du recensement électronique au moyen d'un logiciel de navigation, sur une adresse sécurisée de type HTTPS/SSL. Le serveur de ce site avait lui-même uniquement accès (à travers un pare-feu) au serveur d'application, lorsque les règles prévues étaient appliquées au pare-feu. Seuls les serveurs

autorisés avaient un accès direct au serveur d'application. Les services Web étaient assurés au moyen de la technologie WCF (Windows Communication Foundation).

Figure 3



B. Sécurité

19. Avant de remplir le questionnaire, les recensés se sont conformés à une procédure d'authentification en saisissant un identifiant d'accès (ID), assorti d'un chiffre de vérification et d'un code confidentiel (PIN), communiqué au préalable par l'agent recenseur lors de la phase de remise du questionnaire.

20. À cette fin, l'enveloppe fermée et opaque utilisée faisait apparaître, à l'extérieur, le code d'identification, et contenait, à l'intérieur, l'identifiant d'accès et le code confidentiel.

21. Un dispositif de blocage de l'identifiant d'accès a été introduit dans la procédure de validation, afin de protéger la base de données contre d'éventuelles attaques et de contrôler un nombre maximum de sessions actives tout en assurant des services d'une qualité suffisante aux personnes complétant le formulaire électronique.

22. Avant de remplir le questionnaire, il fallait indiquer le code géographique du logement (validé jusqu'au niveau de l'unité territoriale de la zone géographique considérée), ainsi que son adresse. Après s'être correctement identifiés, les recensés pouvaient modifier leur code confidentiel.

23. Une fois le questionnaire dûment rempli et envoyé, les recensés étaient informés que la paire identifiant/code confidentiel n'était plus acceptée.

24. Un relevé des erreurs et un système d'exploitation centralisé ont été mis en place pour les différentes applications, de façon à surveiller et à caractériser les erreurs constatées.

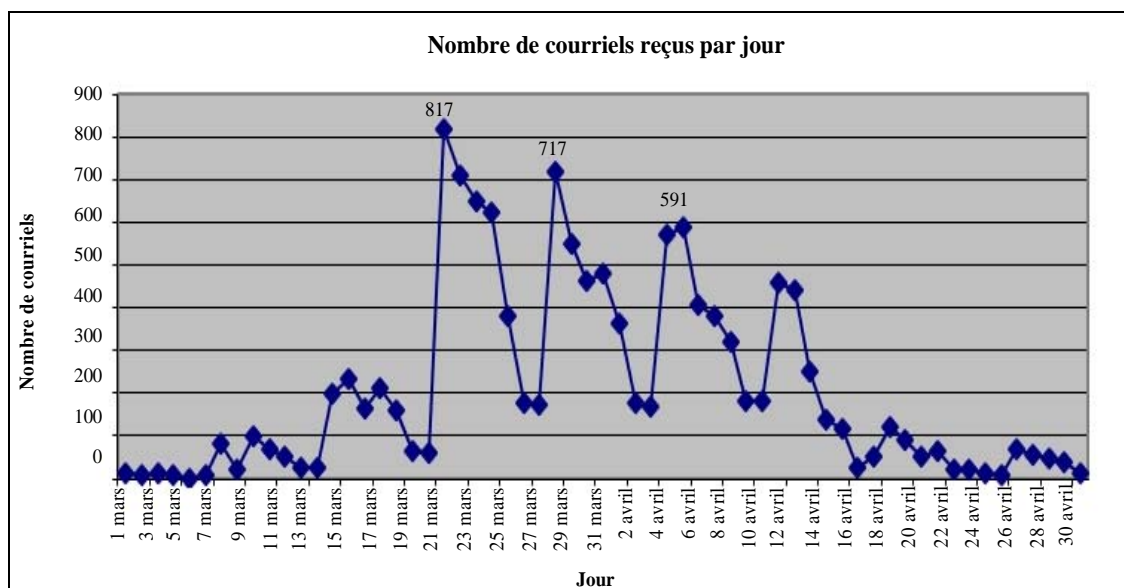
25. La sécurité des différents services (services Web et WCF) a été assurée au moyen d'un mot de passe pour l'utilisateur et d'un système d'identification pour les machines. Un niveau de sécurité a été défini pour valider l'identité de l'émetteur d'informations, c'est-à-dire pour sécuriser les informations échangées par le biais de tel ou tel service. Cet impératif de sécurité a prévalu tant pour la procédure d'accès que pour la procédure de modification des informations.

C. Services administratifs et indicateurs opérationnels

26. Une adresse électronique a été communiquée à la population pour tous contacts et demandes d'assistance. Entre le 1^{er} mars et le 30 avril, 12 713 courriels ont été reçus, dont 81 % (10 236) durant la période d'ouverture du site Web (du 21 mars au 12 avril). En

moyenne, 504 courriels ont été reçus chaque jour pendant la première semaine, 417 pendant la deuxième, puis 376 au cours de la troisième.

Graphique 1



27. La plupart des courriels reçus (70 %) concernaient les activités menées sur le terrain, plutôt que des difficultés liées à l'envoi des réponses via Internet:

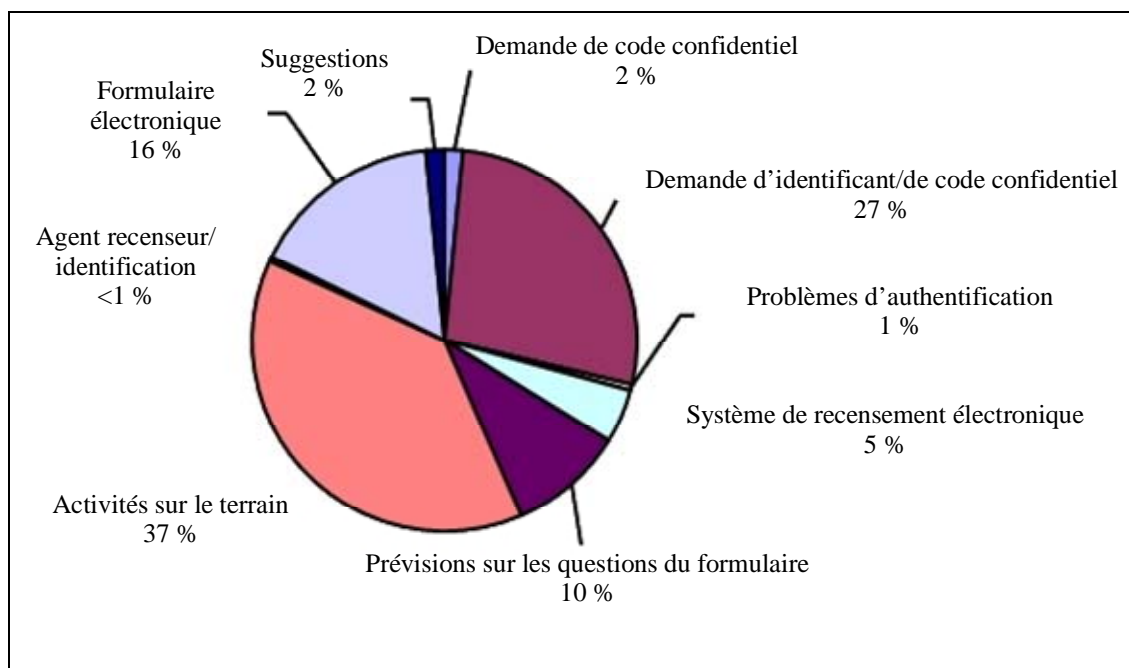
28. Il a été répondu à tous les courriels dans des délais variables suivant leur objet:

- Demande d'identifiant/code confidentiel – jusqu'à vingt-quatre heures pour les jours les plus chargés;
- Demande d'assistance concernant les informations nécessaires pour répondre par Internet – quarante-huit heures au maximum;
- Autres types de demande – délai ne dépassant jamais soixante-douze heures.

D. Service d'assistance téléphonique

29. Un service d'assistance téléphonique gratuite a fonctionné chaque jour entre 9 heures et 22 heures (y compris les week-ends). Du 11 mars au 27 avril, 127 446 appels ont été reçus. La durée moyenne de chaque communication était de trois minutes.

Graphique 2

Principaux motifs des appels adressés au service d'assistance téléphonique

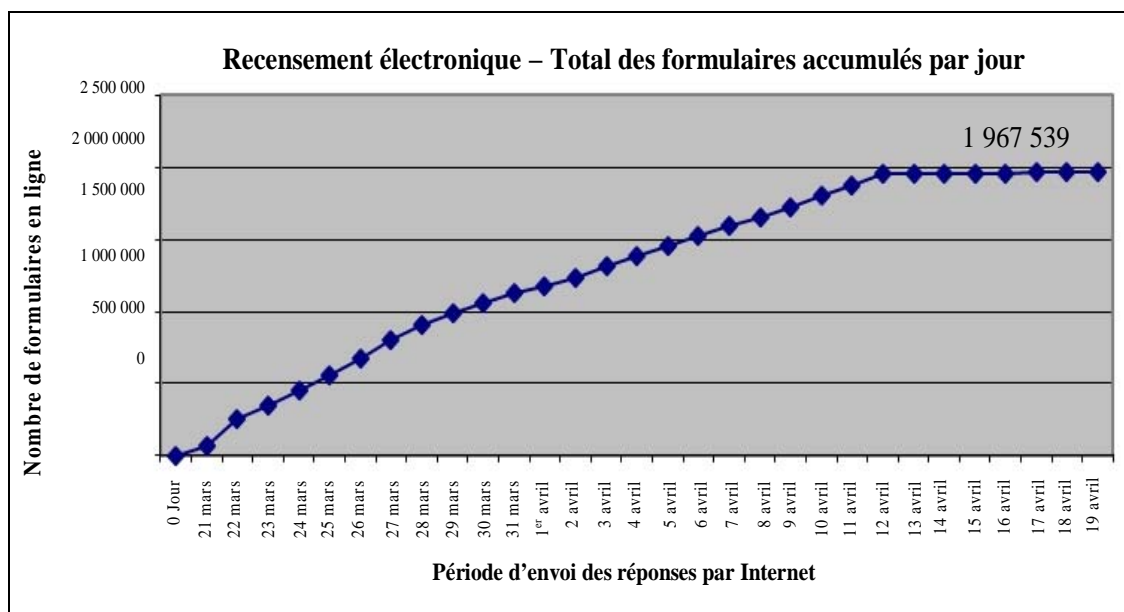
30. Dans 40 % environ des cas (distribution et collecte, confirmation par un membre du personnel, suggestions), l'assistance fournie en ligne n'était pas liée à des difficultés rencontrées dans l'utilisation du système de recensement électronique.

V. Résultats

A. Taux de réponse aux niveaux national et régional

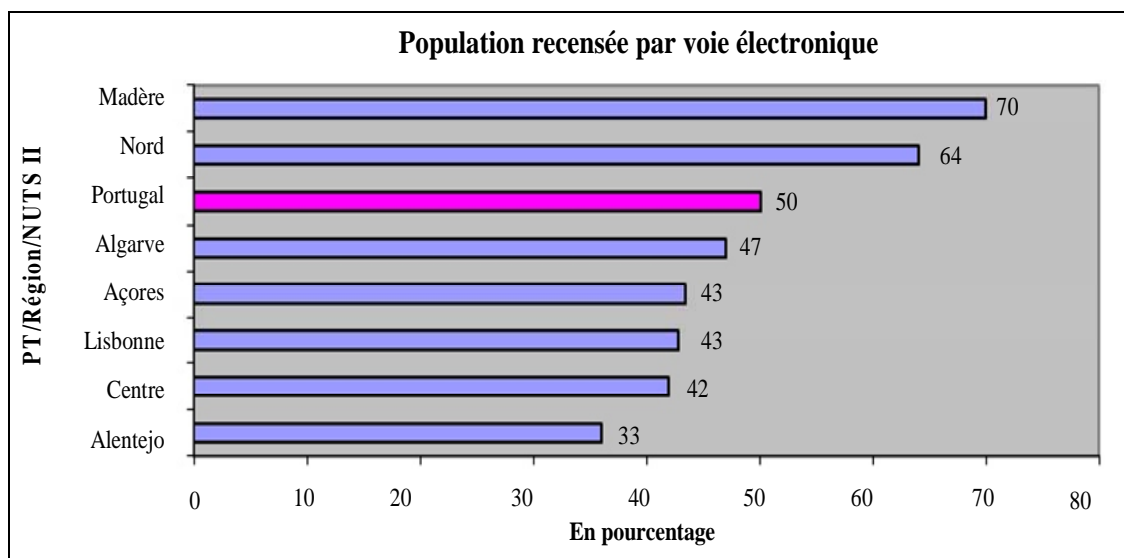
31. Les réponses en ligne ont représenté environ 50 % du total, ce qui correspond à 1 967 539 unités d'habitation et 5 328 044 résidents. Pour le premier jour, 69 094 réponses ont été reçues, après quoi le nombre de réponses s'est établi en moyenne à 85 500 par jour.

Graphique 3



32. Toutes les régions du pays ont connu un taux de réponse élevé, compris entre 36 % et 70 %.

Graphique 4



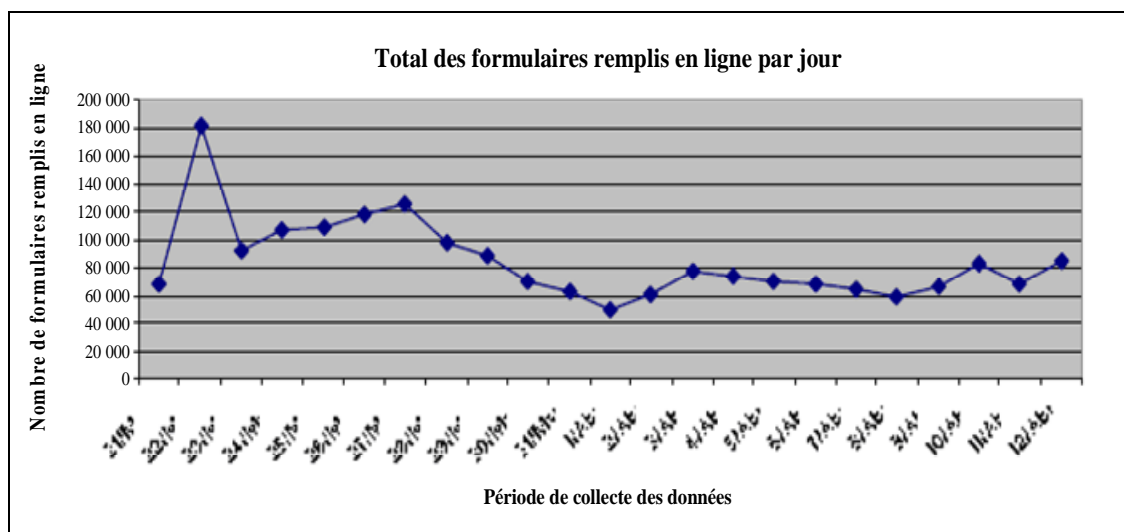
B. Tendances constatées dans l'envoi des réponses en ligne

33. Pendant les vingt-trois jours durant lesquels le site Web a été accessible, l'évolution observée dans l'envoi des réponses en ligne a été comparable à celle des essais préparatoires. Une bonne partie des réponses est parvenue dès les premiers jours, le site Web étant consulté essentiellement après le dîner, entre 21 heures et 23 heures.

34. Dans la première semaine, le nombre de réponses envoyées a été relativement plus important, de l'ordre de 115 000 par jour. Un pic a été enregistré le deuxième jour avec

182 046 réponses envoyées, soit plus du double du nombre enregistré le premier jour (70 000).

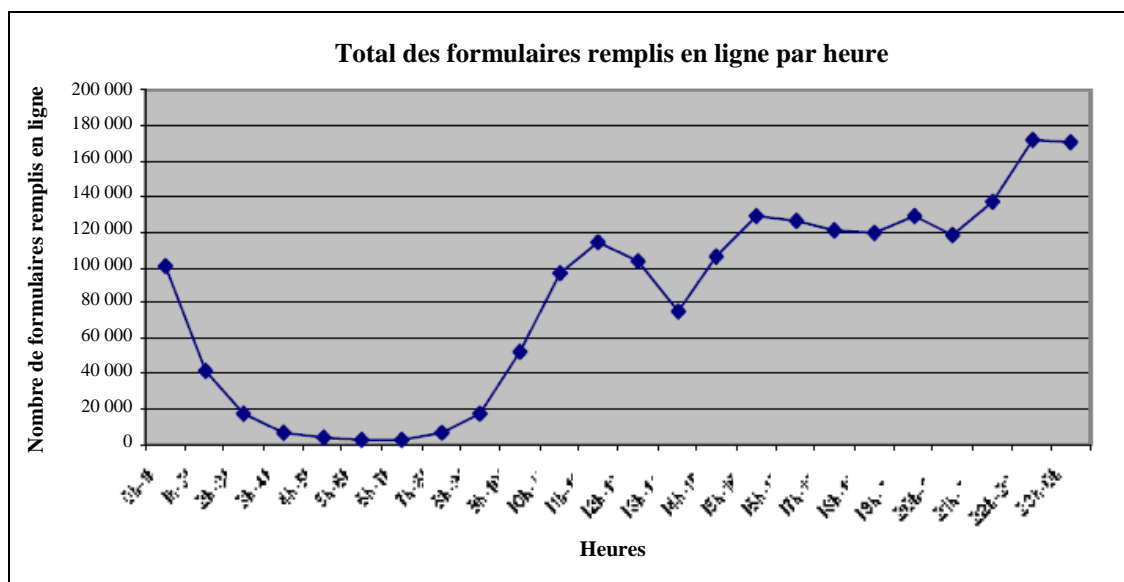
Graphique 5



35. Dans les deux semaines suivantes, le rythme s'est ralenti, passant à 72 000 réponses en moyenne par jour (soit 33 000 de moins que dans la première semaine). À la fin de la période de collecte et malgré la prolongation du délai, il n'y a pas eu une nouvelle pointe le dernier jour.

36. La courbe horaire des réponses fait apparaître un nombre plus élevé de consultations dans la tranche 21 heures-23 heures, qui concentre environ un quart des réponses de la journée.

Graphique 6



37. La période de collecte en ligne, initialement fixée à trois semaines (du 21 mars 2011 au 10 avril 2011), a été prolongée de deux jours supplémentaires, jusqu'au 12 avril. Cette prolongation visait principalement à atténuer tout encombrement du site le dernier jour et à

permettre à un plus grand nombre de personnes de participer au recensement de 2011 par Internet.

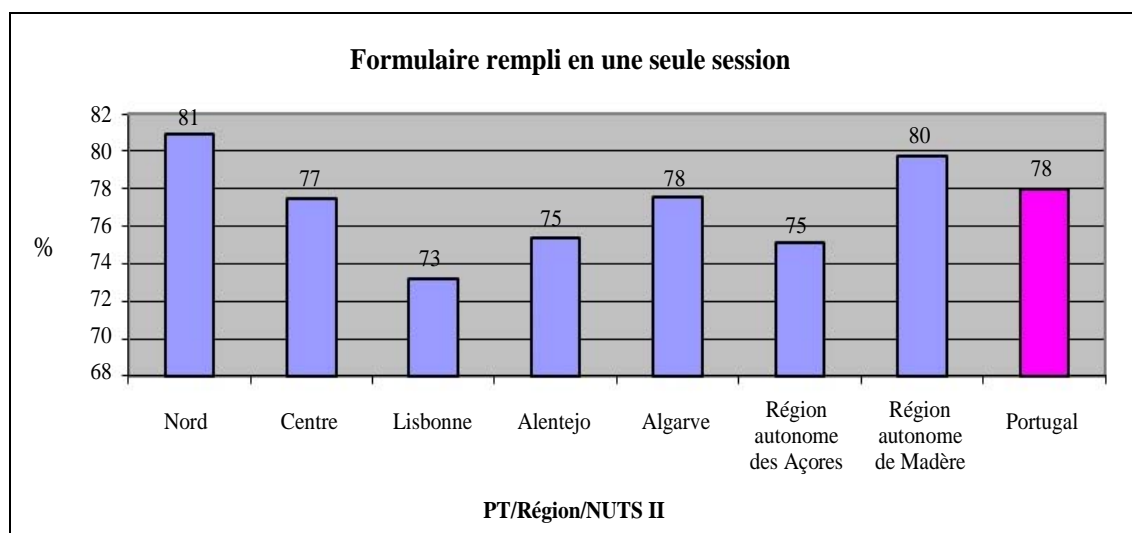
C. Temps nécessaire, en moyenne, pour remplir le questionnaire

38. Il a fallu en moyenne quarante-cinq minutes pour répondre au questionnaire en ligne. Ce délai moyen est légèrement plus élevé que dans les essais antérieurs, notamment l'enquête pilote (quarante-deux minutes), ce qui est sans doute imputable à la lenteur du site Web en fonction des jours et à certaines heures. La moyenne régionale s'échelonne entre quarante minutes dans la région du Nord et cinquante-trois dans celle de Lisbonne.

D. Réponses envoyées en une seule session

39. Environ 78 % des réponses – à savoir 1 526 924 – ont été complétées en une seule fois. Cela signifie que la plupart des répondants ont commencé à remplir le questionnaire du recensement de 2011 et l'ont achevé dans la foulée. La fonction «Sauvegarder et reprendre ultérieurement» a été utilisée dans 22 % des cas.

Graphique 7



E. Réponses au sondage d'opinion

40. Un sondage, comportant deux questions seulement, a été effectué auprès des personnes qui participaient par Internet au recensement de 2011. La réponse à ce questionnaire étant facultative, les résultats font apparaître uniquement l'opinion des personnes de l'échantillon concerné.

Tableau 2
Résultats du questionnaire d'évaluation en ligne

	<i>Oui</i>	<i>%</i>	<i>Non</i>	<i>%</i>	<i>Sans objet</i>	<i>%</i>	<i>Total</i>
Êtes-vous satisfait(e) d'avoir répondu en ligne au Recensement 2011?	192 067	94,5	9 567	4,7	967	0,5	202 601
Avez-vous l'intention d'utiliser Internet pour répondre à une autre enquête statistique?	196 673	97,1	4 915	2,4	1.013	0,5	202 601

41. Au total, 202 601 foyers ont participé au sondage, soit un taux de participation de 10 % par rapport au nombre de réponses en ligne dans le cadre du recensement de 2011 (1 967 539).

42. Les résultats montrent que 95 % des personnes ayant répondu à ce questionnaire facultatif se sont déclarées satisfaites d'avoir participé au recensement de 2011 via Internet et que 97 % d'entre elles entendent répondre à d'autres enquêtes de Statistique Portugal de cette manière.

VI. Principales conclusions

- Le taux de réponse – de l'ordre de 50 % – dénote une forte participation de la population à la collecte des données du recensement en ligne.
- Le système mis en place a montré qu'il pouvait répondre aux besoins réels.
- Les opérations menées à titre expérimental en 2008, 2009 et 2010 ont permis de mettre au point un serveur fiable et de matérialiser la configuration finale du recensement.