



Conseil économique et social

Distr. générale
30 mars 2010
Français
Original: anglais

Commission économique pour l'Europe

Conférence des statisticiens européens

Cinquante-huitième réunion plénière

Paris, 8-10 juin 2010

Point 6 de l'ordre du jour provisoire

Statistiques spatiales

Cartographie, Google et statistiques de voisinage*

Note de Statistics Netherlands

Résumé

La note traite du recours aux techniques spatiales modernes pour la diffusion des statistiques. Elle porte sur l'utilisation des cartes par Statistics Netherlands, sur les possibilités qu'offre et les problèmes que pose l'emploi des registres pour établir un lien entre les données et la géographie ainsi que sur les récentes innovations en matière de statistique et de cartographie. Elle expose également un certain nombre de perspectives qui vont s'ouvrir à l'avenir concernant la diffusion des statistiques spatiales.

* Note largement basée sur un article intitulé *Cartography with a capital G: Google and more*, qui a été présenté précédemment à la Conférence sur la refonte des statistiques officielles que l'Association internationale pour les statistiques officielles (AISO) avait organisée à Shanghai (Chine) du 14 au 16 octobre 2008.

I. Introduction

1. Il y a une trentaine d'années, le romancier néerlandais Maarten Biesheuvel a publié l'histoire de deux amis norvégiens, Carl et Nyls, qui rêvaient de créer le globe le plus achevé. Ce globe devait être une reproduction exacte de toute la planète telle qu'elle était le 21 août 1975 à 12 heures. Dans le monde de la statistique, nous qualifierions cela de point de référence. Sur ce globe extrêmement précis, tout serait visible à travers un microscope, que ce soient les gens, les maisons, les trains, les navires, les bureaux, les usines, les hôpitaux, les exploitations agricoles, les écoles, les arbres, voire les vaches dans les champs. Il y aurait en plus des annexes spéciales dans lesquelles seraient ajoutés des registres généalogiques pour tous les individus, animaux, etc. L'intention de ces deux amis était de cartographier et d'enregistrer le monde; pour les statisticiens, il s'agissait de construire une utopie statistique.

2. C'était là un projet démentiel, mais ils ont toute leur vie cherché à réaliser ce rêve. Un dénommé Nwaak ayant eu vent de leur projet leur a envoyé un paquet. À l'instant où le paquet a été posté, une obscurité totale s'est abattue sur le monde. Des catastrophes se sont produites pendant des semaines jusqu'à ce que Carl et Nyls ouvrent le paquet. Celui-ci contenait le globe dont ils rêvaient mais ce dernier s'est révélé plus catastrophique qu'ils n'auraient jamais pu l'imaginer. Lorsqu'il est emballé, l'obscurité s'abat sur la terre, et des tremblements de terre dévastateurs se produisent lorsque les postiers le touchent en écrasant le papier d'emballage. Ce globe préfigure notre planète, c'est son prototype. Avec un microscope, Carl et Nyls peuvent se servir du globe comme d'une fenêtre sur le monde. Ils peuvent alors tout observer, partout et à tout moment sur toute la surface de la terre.

3. Trente ans après ce récit imaginaire, le rêve de Carl et Nyls est en train de devenir réalité. Grâce aux nouvelles techniques novatrices de diffusion, il est possible de voir le monde pratiquement de partout. Il suffit de quelques clics de souris pour que Google Earth et Google Maps, entre autres, ouvrent une fenêtre sur le monde, exactement comme Carl et Nyls l'auraient bien aimé. Dans quelques domaines, par exemple la météo ou la circulation, les informations disponibles apparaissent même en temps réel. Statistics Netherlands participe à ces innovations en menant à bien deux initiatives tout à fait particulières: *Neighbourhood Statistics in Google Earth* et *Statistics Netherlands in Your Neighbourhood*.

4. D'un point de vue cartographique, cela pose cependant un dilemme intéressant. Les cartes thématiques sont une simplification de la réalité. Elles sont conçues pour montrer la répartition géographique d'une variable statistique en laissant de côté tous les détails sans intérêt. Pour utiliser et présenter des photographies aériennes, comme dans Google Earth, il faut réintroduire dans la carte une réalité détaillée. Dans le cas de Carl et Nyls, ce dilemme ne les effleure même pas. Par contre, c'est là un problème sérieux dans la perspective d'une diffusion. La présente note expose comment Statistics Netherlands a résolu ce dilemme et a réussi à utiliser des techniques modernes pour diffuser ses statistiques.

II. Statistiques et cartographie: premières expériences

5. Les cartes ont toujours revêtu une importance déterminante lorsque l'on envisage un voyage, que ce soit sur terre ou sur l'eau. L'orientation est probablement la fonction la plus tangible des cartes. De nos jours, les cartes peuvent également fournir de grandes quantités de données avec une composante spatiale. Elles n'ont plus à être imprimées sur papier. Elles peuvent être également virtuelles ou interactives et peuvent s'afficher sur un écran d'ordinateur. Leur raison d'être est de fournir des informations spatiales. Elles peuvent servir à illustrer des différences ou des similitudes entre zones géographiques. Lorsqu'elles

sont bien conçues, elles montrent clairement les configurations locales ou régionales, lesquelles n'apparaissent pas nécessairement dans des tableaux ou sur des graphiques. Comme de nombreuses statistiques ont une composante géographique, les cartes conviennent particulièrement bien pour une représentation spatiale des résultats statistiques.

A. Diffusion de statistiques régionales au moyen de cartes Web par les services nationaux de statistique

6. De plus en plus, les services nationaux de statistique affichent leurs statistiques régionales sur le Net. Dans l'ensemble, le nombre de leurs sites Web a augmenté, passant de 126 en 2003 à 187 en 2005. De surcroît, un nombre croissant de ces services est en train de découvrir l'utilité potentielle des fonctionnalités cartographiques sur leurs sites Web.

7. Les données peuvent être récupérées sous la forme d'une carte thématique préconçue sur la moitié des sites Web. Un quart des sites ont une interface cartographique pour retrouver et sélectionner des statistiques régionales. Près d'un quart des sites Web offrent un accès en ligne à une base de données pour obtenir des résultats individualisés. Enfin, 19 sites Web au total peuvent produire un résultat individualisé sous la forme d'une carte thématique ou d'une visualisation cartographique interactive.

Représentation graphique 1

L'une des interfaces régionales associées à des sites Web de services nationaux de statistique

China Statistical Information Network (Beijing, China)										
National language	English	Other languages	Keyword search	Census data retrieval format	Data download	Census data retrieval cost	Levels of aggregation	Individualized output of Census data	Geographical map interface for retrieval	
•	•	•	•	Textual	•	Free	National	2nd	•	•
				Tabular	•	Registration	Other	Individual	•	•
				Chart/Map	•			Tabular	•	•
					•			Chart/Map	•	•
					•			Interactivity	•	•

Source: Switchboard to NSO websites worldwide, <http://kartoweb.itc.nl/nso>.

B. Utilisation des cartes par Statistics Netherlands

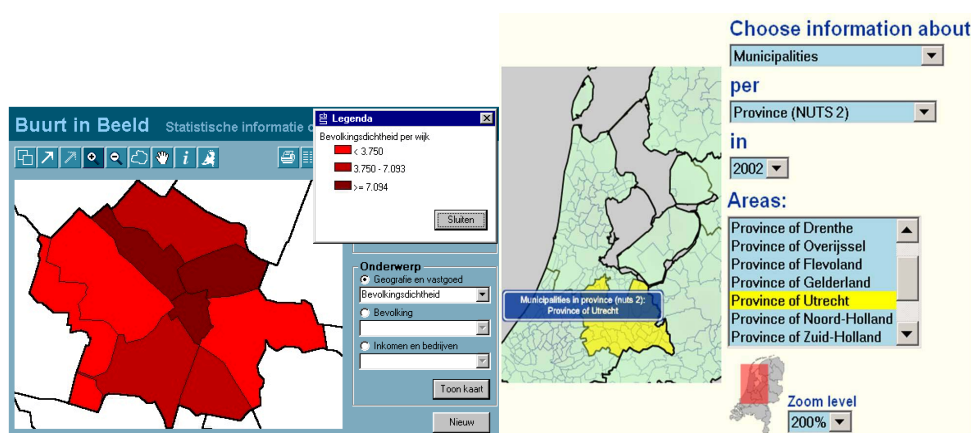
8. Depuis l'apparition sur l'Internet des cartes numériques et interactives, Statistics Netherlands s'est investi dans la diffusion cartographique et la visualisation des statistiques régionales et a produit des applications telles que *Buurt in Beeld* («Neighbourhood in the picture») en 2001 et *Cartographic Entry into StatLine* en 2003. Dans une étude réalisée en 2005 par l'International Institute for Geo-Information Science (ITC), Statistics Netherlands figurait parmi les pionniers dans le monde des services nationaux de statistique. En 2003,

six sites Web seulement dans le monde offraient à leurs utilisateurs une interface géographique pour la recherche et la récupération des données. De surcroît, les deux applications de Statistics Netherlands peuvent être utilisées comme supports de présentation et pour des analyses et recherches en ligne.

9. *Cartographic Entry* a été créée en 2003. Il ne s'agit pas d'une application isolée; elle est au contraire intégrée dans StatLine, la base de données statistiques en ligne, et comporte une interface cartographique utile pour trouver et récupérer des données statistiques régionales depuis le niveau des communes jusqu'à celui des groupes de province. Des deux applications, c'était *Neighbourhood in the picture* qui faisait vraiment figure de pionnier. Instrument autonome placé sur le site Web en 2001, elle a remporté sur-le-champ un énorme succès. Elle apporte aux utilisateurs des informations allant du niveau du district à celui du voisinage, informations que le site Web n'avait pu fournir jusqu'alors.

Représentation graphique 2

Neighbourhood in the picture (2001) et Cartographic Entry (2003)



10. Pour les deux applications, les utilisateurs devaient installer un module d'extension, et c'était là manifestement un inconvénient. Les serveurs n'en ont pas moins été paralysés en raison du grand nombre de visiteurs. Ce nombre et la vitesse à laquelle il augmentait étaient spectaculaires.

C. Facilité d'utilisation et technologie: Quelle est la prochaine étape?

11. Indépendamment de cet immense succès, la question s'est posée à un certain moment de savoir si toute cette fonctionnalité cartographique aidait véritablement les utilisateurs, en particulier ceux qui consultaient la base de données StatLine. Une enquête de satisfaction a été réalisée. Il s'est avéré qu'une interface cartographique était véritablement utile pour la récupération des données régionales. Les réactions des utilisateurs ont également conduit à penser que le potentiel n'en était pas encore totalement exploité.

12. Entre-temps, l'Internet évoluait rapidement. Le site Web de Statistics Netherlands a été modernisé et transformé en un système de gestion du contenu de type professionnel. Il n'a pas été jugé réaliste d'investir dans la technologie de l'application *Neighbourhood* qui avait alors vieilli. Après une demi-décennie, ce pionnier était dépassé et s'est trouvé bruyamment arrêté. Plusieurs questions se sont posées au cours d'une période de réflexion. Une nouvelle génération de supports de diffusion d'informations de voisinage était souhaitable, mais sous quelle forme?

13. En premier lieu s'est posée la question de la communication des données disponibles au niveau du voisinage. La quantité d'informations disponibles est inversement proportionnelle au nombre de régions. Au niveau national et à celui de la province, la communication des données est très au point. Par contre, au niveau du voisinage, il n'est possible d'obtenir des chiffres fiables qu'en utilisant de grands échantillons ou bien des sources dans leur intégralité. Si l'on ajoutait des districts et des voisinages à *Cartographic Entry*, on aboutirait de toute évidence à des incohérences entre les données régionales et les données locales. Par ailleurs, l'enquête de satisfaction faisait aussi apparaître clairement que ce n'était pas une bonne idée d'intégrer une cartographie de voisinage dans *Cartographic Entry* qui existait encore. L'interface utilisateur existante ne convenait pas pour trouver les voisinages et calculer les statistiques correspondantes: trop de manipulations étaient nécessaires avant de parvenir au résultat demandé. Il fallait donc de nouveau créer une application fonctionnant de manière séparée.

14. Sur le plan technique, il a été possible d'utiliser *Cartographic Entry*, qui fait partie de la base de données StatLine, en appliquant au serveur le format SVG (graphique vectoriel adaptable) et en utilisant le format Flash à partir de 2010. Cela dit, on rencontrerait un grand nombre de problèmes techniques si l'on appliquait la fonctionnalité cartographique de SVG ou de Flash au niveau du voisinage. L'affichage de 11 000 voisinages en plus des 500 communes donnerait probablement naissance à un instrument difficile à utiliser. De surcroît, Google qui était alors entré en scène laissait entrevoir de nouvelles perspectives prometteuses dans le domaine de la cartographie interactive.

15. Enfin, l'apparition, en plus de Google, de plusieurs nouvelles sources de données importantes a conduit à penser qu'une nouvelle étape dans la diffusion des statistiques de voisinage était réalisable. C'est ce dont vont traiter les parties 3 et 4.

III. Utilisation de registres: perspectives et enjeux

16. Les statistiques régionales consistent essentiellement à rassembler et associer des chiffres concernant différents sujets et différentes zones et à créer ainsi une valeur ajoutée. Les *chiffres* et les *zones* sont des éléments qui ont énormément évolué.

A. Utilisation de registres: les données (chiffres)

17. En 2003 et 2004, la loi régissant Statistics Netherlands a fait l'objet d'un réexamen et d'une révision. L'un des principaux changements a consisté à autoriser Statistics Netherlands à utiliser, explicitement et gratuitement, des informations tirées de registres d'autres sources publiques. Dans le même temps, la demande d'informations au niveau local était en augmentation. Il en est résulté un déplacement de ressources qui s'est traduit par la mise en route d'un projet distinct d'élaboration de statistiques de voisinage dans un cadre innovant plus large intitulé «*Statistiques spatiales, statistiques du logement et statistiques de la mobilité*».

18. Rétrospectivement, les résultats sont impressionnants. Les chiffres correspondant au niveau du voisinage avaient plutôt concerné des sujets classiques: population, logement, revenu et surface habitable. En 2001, l'ensemble de la palette comprenait seulement 40 chiffres. L'utilisation de nouveaux registres a permis de recueillir des informations sur de nouveaux sujets: éducation, emploi, sécurité sociale, possession d'un véhicule et utilisation des terres. Le nombre de chiffres disponibles au niveau du voisinage a augmenté au point de dépasser largement la centaine. Les travaux ne sont pas terminés. Actuellement, un thème, la *proximité* des équipements collectifs, est à l'étude et les premiers résultats ont été publiés en 2009.

19. Par ailleurs, le cycle biennal de publication des statistiques de voisinage a été abandonné. Après 1995, 1997, 1999 et 2001, une nouvelle publication a été élaborée et son premier numéro a paru en 2003. Depuis lors, des chiffres ont été publiés chaque année. Enfin, la préparation des publications a été accélérée: les chiffres pour 2001 ont été connus au bout de vingt-cinq mois. Par contre, les premiers chiffres pour 2005 ont été publiés au bout de onze mois, soit deux fois plus vite. L'objectif est de faire paraître les publications dans les neuf mois, soit près de trois fois plus vite qu'auparavant.

20. L'incorporation des sujets mentionnés plus haut a entraîné un grand nombre d'incohérences et de nombreux problèmes au niveau de la publication des données, mais ces dernières années ont permis d'améliorer la possibilité d'obtenir des données à trois égards: quantité, fréquence et rapidité.

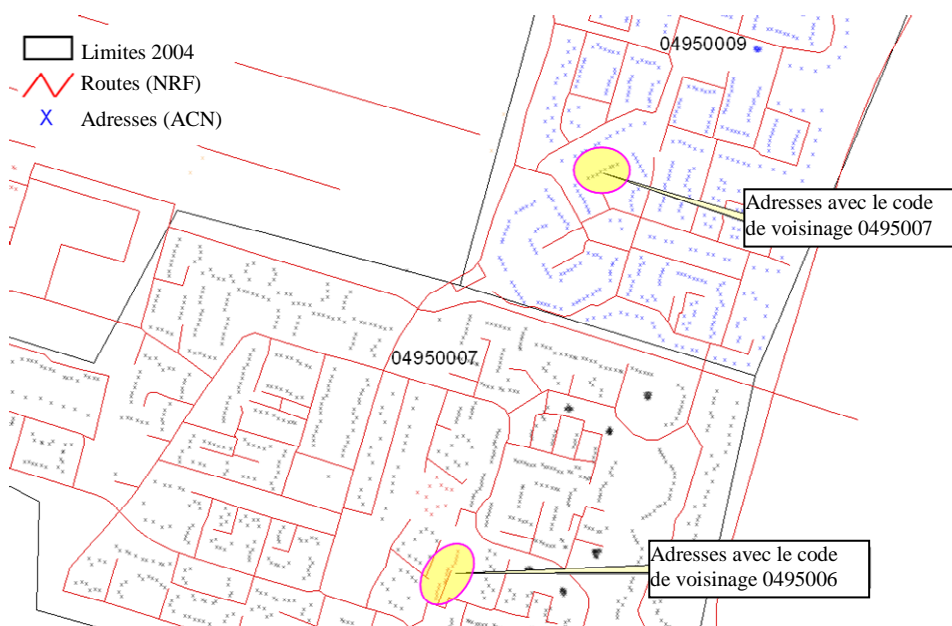
B. Utilisation de registres: la géographie (zone)

21. Grâce à la nouvelle loi régissant Statistics Netherlands, il est devenu possible d'utiliser deux registres géographiques importants pour l'établissement des statistiques. Ces registres contiennent des informations sur les voies de circulation routière, des autoroutes aux simples rues, et des données sur l'emplacement des adresses, c'est-à-dire les coordonnées x et y. Vers le milieu de 2004, ces nouvelles informations tirées des registres ont été confrontées pour la première fois à nos propres informations géographiques.

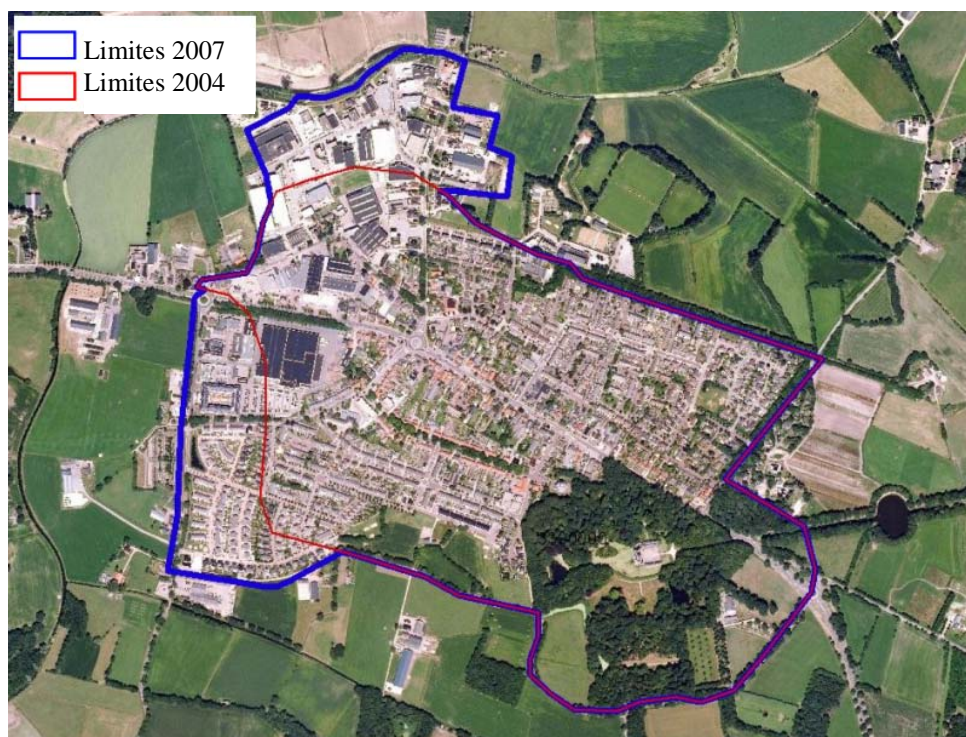
22. Jusqu'en 2004, les statistiques de voisinage ont été calculées à partir de deux sources géographiques importantes qui ont pour nom *Geographical Base File (GBF)* et *CBS/TOP boundaries*. Le GBF est un registre d'adresses qui contient des informations sur le lieu de résidence, le nom de la rue et le numéro du logement pour les sept millions d'adresses aux Pays-Bas. Il a été tenu à jour pendant des décennies, mais il est de notoriété publique qu'il contenait des erreurs datant de l'époque où la numérisation n'existait pas. Le *CBS/TOP boundaries* indique les limites de tous les voisinages aux Pays-Bas. Comme le font tous les polygones (qui expriment les limites), ce registre comprend les noms et codes des 11 000 voisinages. L'association de ces deux registres avec un troisième, le registre de la population par exemple, permet d'obtenir les chiffres de la population correspondant aux divers voisinages.

23. L'adjonction du *National Road File (NRF)* et de l'*Address Coordinates Netherlands (ACN)*, permet d'ajouter le lien manquant: l'emplacement. L'association de tous les registres a produit deux résultats: les adresses qui, dans le GBF, ne correspondaient pas au voisinage (c'est-à-dire qui avaient un code de voisinage erroné) et les polygones qui, dans le *CBS/TOP boundaries*, avaient un tracé inexact. Grâce à l'association des routes et de l'emplacement des adresses, complétée par des photos aériennes qui étaient également disponibles, la qualité de nos propres sources et, par voie de conséquence, celle des statistiques de voisinage, s'est améliorée. Deux exemples illustrent ces deux situations.

Représentation graphique 3
Amélioration qualitative à Bleiswijk



Représentation graphique 4
Développement du parc immobilier et modification des limites à Renswoude



24. Finalement le voisinage a été modifié pour plus de 50 000 adresses, soit environ 1 % de l'ensemble des adresses. Au total, une centaine de limites ont aussi été modifiées, soit environ 1 % également. Dans l'ensemble, les pourcentages ne sont pas particulièrement

impressionnants mais, comme on pouvait s'y attendre, la plupart des adresses étaient regroupées.

25. À partir de 2004, les deux registres géographiques ont été utilisés pour les mises à jour. La mise à jour chaque année des adresses correspondant en particulier aux nouveaux logements était une opération généralement longue et coûteuse. L'utilisation du NRF et de l'ACN a marqué la fin des enquêtes dans les communes et de l'envoi d'agents sur le terrain afin d'établir les codes de voisinage pour les nouvelles adresses, ce qui a permis aux pouvoirs publics en général et à Statistics Netherlands en particulier de réaliser des économies importantes, de sorte que l'amélioration de la qualité n'a pas été le seul point positif de cette évolution.

C. Améliorations et nouvelles idées: la prochaine étape!

26. Jusqu'à présent, nous avons considéré les progrès concernant la communication et la qualité des données ainsi que la réduction des coûts. Les données ont été communiquées en plus grand nombre, plus fréquemment et plus rapidement, et il s'agit maintenant, avec l'amélioration de la qualité sur le plan géographique, d'en faire profiter l'utilisateur. Dans le passé, la composante géographique était pratiquement passée sous silence. L'amélioration de la qualité sur le plan géographique a cependant suscité de nouvelles idées pour la diffusion des statistiques de voisinage. De surcroît, avec l'entrée en scène de Google, le moment est venu de passer à une prochaine étape, celle de la géovisualisation.

IV. Statistiques et cartographie: travaux récents

27. Après l'inventaire qui a été dressé des contributions des utilisateurs et après les grandes améliorations qui ont été apportées sur le plan de la qualité, le moment était venu de passer à une nouvelle étape de la diffusion des statistiques de voisinage. Mais comment la concevoir?

28. Ces dernières années, l'utilisation d'applications cartographiques pour mener à bien des tâches aussi courantes que chercher à éviter des embouteillages ou prendre connaissance des prévisions météorologiques locales a connu un essor spectaculaire. L'achat de maisons en est un autre exemple, de même que les agences de voyage qui ont très rapidement adopté ces applications. La population est de plus en plus habituée à utiliser des moteurs cartographiques dans ses activités de tous les jours. Sur un plan pratique, qu'y a-t-il de mieux que de présenter des statistiques qui font appel à des mécanismes déjà bien connus.

29. De surcroît, le recours à des technologies cartographiques utilisant la photo en mode continu a lui aussi connu un essor spectaculaire aux Pays-Bas. Il s'agit là d'une technologie de pointe pour afficher des cartes sans difficulté sur l'écran des utilisateurs qui sera expliquée par la suite. L'une des causes de cet engouement tient probablement aux images des Pays-Bas qui sont très détaillées étant donné que les photos aériennes sont de plus en plus utilisées. Naturellement la qualité du service rendu s'en trouve améliorée du tout au tout, étant donné que les gens reconnaissent des détails de leur propre voisinage et de leur propre maison, voire de leur propre voiture. C'est exactement la raison pour laquelle ces technologies sont tellement appréciées pour diffuser des statistiques de voisinage.

A. Échecs et réussites: une «course à trois chevaux»

30. L'année 2006 a marqué la première étape, mais des doutes ont surgi au bout de quelques mois. Les objectifs étaient très ambitieux mais il y avait aussi beaucoup trop d'incertitudes. De surcroît, les attentes des parties prenantes n'étaient plus les mêmes et le projet a pâti de multiples restrictions techniques et limitations en rapport avec les logiciels: il a malheureusement échoué et nous n'avons pu qu'en tirer un grand nombre d'enseignements.

1. Première réussite: initiative n° 1

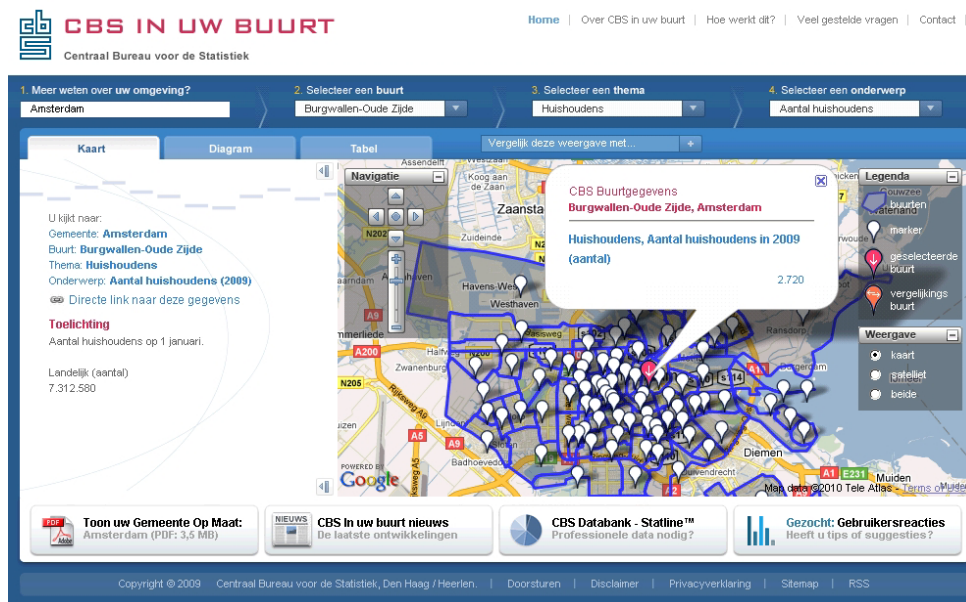
31. Dans les derniers mois de 2006, Statistics Netherlands a été invité à participer à une initiative ouverte au public dans le cadre de l'infrastructure nationale de géodonnées, qui avait pour nom *OnzeGeo*. Comme c'était un projet de grande ampleur, l'inconvénient était que nous n'avions pas beaucoup d'influence sur sa mise en œuvre et son orientation. Sur le plan interne, nous en étions encore à étudier ce qui nous avait arrêtés précédemment. Cependant, le projet gouvernemental était facilement accessible et il a été décidé d'y participer et d'en retirer des enseignements. Ce projet nous a fourni l'occasion de réfléchir sur les positions initiales, alors même qu'«un premier cheval était en fait dans les starting blocks». On attendait beaucoup de ce premier cheval et on espérait bien voir apparaître rapidement des résultats. Au bout de trois mois à peine, les premiers résultats ont été présentés dans une version beta au cours d'une conférence nationale. Cependant, l'objectif, les tâches à effectuer et l'encadrement n'étaient toujours pas suffisamment clairs. En particulier, l'absence d'appropriation a conduit à y mettre provisoirement un terme. Toutefois, Statistics Netherlands avait alors une meilleure idée des intervenants mondiaux tels que Google Earth et Microsoft Virtual Earth qui étaient entrés en scène. On en était venu à mieux voir comment utiliser ces interfaces cartographiques perfectionnées qui rencontraient un grand succès et ouvraient des perspectives prometteuses. Par contre, le cheval n° 1 n'avait pas terminé la course.

2. Une véritable réussite: les initiatives n°s 2 et 3

32. Le mois de janvier 2007 a marqué un tournant. Statistics Netherlands a décidé d'utiliser les outils modernes qu'offre la cartographie basée sur la photographie en mode continu pour la présentation de ses statistiques de voisinage. Il n'existait que peu d'options quant à la technologie à utiliser. Microsoft proposait *Microsoft Virtual Earth*, et Google la technologie employée dans *Google Map* et la version plus perfectionnée *Google Earth*. Il existait également des logiciels libres: *Open Streetmap* et *NASA World Wind*. D'après les expériences réalisées, ce que Google proposait convenait techniquement mieux pour nos projets. Mais il y avait plus important encore: considérant nos futurs utilisateurs, la population néerlandaise, nous avons remarqué que les deux technologies de Google étaient bien plus prisées que les autres.

33. L'initiative n° 2 prenait racine dans le projet lancé en 2006, qui avait été arrêté. Sa réalisation a été confiée à des sous-traitants et, le moment venu, un site dédié qui avait pour nom *Statistics Netherlands in Your Neighbourhood* a été ouvert au public en février 2008. Il remporte depuis lors un grand succès. Dans les premiers jours, on a compté plus de 20 000 visiteurs uniques. À présent, le site est consulté environ 200 fois par jour. Il utilise Google Maps pour projeter des variables statistiques sur une carte Google.

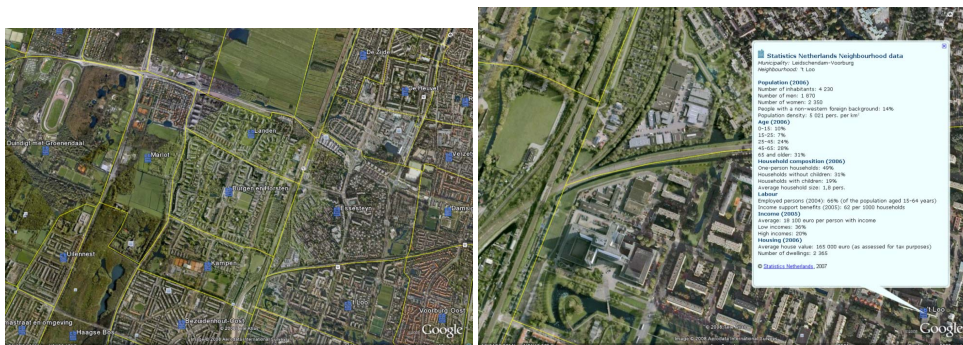
Représentation graphique 5 Site Web de Statistics Netherlands in Your Neighbourhood



34. Le site Web ne contient pas toutes les informations disponibles sur les voisinages, mais seulement les principales informations, et les plus récentes, sur la population, les ménages, l'emploi, le revenu et le logement. Des liens pour prendre connaissance des renseignements associés sont ajoutés en bas de l'écran, par exemple les publications de la commune en format PDF, les articles les plus récents sur les voisinages publiés par Statistics Netherlands, un lien direct à la base de données StatLine de Statistics Netherlands et la possibilité pour les utilisateurs d'apporter des informations en retour.

35. Au cours de l'été 2007, une étape plus ou moins fortuite a été franchie. Après avoir consulté Google Netherlands pour la deuxième initiative, Statistics Netherlands a constaté qu'il était possible de placer une couche d'information statistique sur Google Earth en utilisant le bon savoir-faire. De surcroît, ce qui a aussi son importance, cela pouvait se faire plus facilement et beaucoup plus rapidement qu'avec le nouveau site Web en construction. Un troisième cheval était entré dans la course. À la fin de 2007, le résultat pouvait être présenté au public et a été publié sur le site Web principal de Statistics Netherlands. Il s'agit d'une démarche complémentaire qui découle en fait de *Statistics Netherlands in Your Neighbourhood*.

Représentation graphique 6 Photos aériennes de Google Earth, avec indication des limites des voisinages, noms et statistiques de Statistics Netherlands



36. Les noms et limites des voisinages deviennent visibles si l'utilisateur zoome sur une région bien déterminée. L'icône (CBS) vieille d'une centaine d'années de Statistics Netherlands apparaît dans chaque voisinage lorsque l'image est suffisamment agrandie. Avec un clic, on fait apparaître le signet classique de Google Earth sous la forme d'une bulle sur la photo sous-jacente.

37. Actuellement, *Statistics Netherlands in Your Neighbourhood* et *Neighbourhood Statistics in Google Earth* sont des portails de StatLine qui contiennent tout l'ensemble de données sur les voisinages. Les deux initiatives ne comportent donc qu'un sous-ensemble du total. Priorité est donnée à la présentation de statistiques de voisinage au moyen d'outils et de supports utilisés dans la vie de tous les jours. Au terme de la course des trois chevaux, deux ont passé la ligne d'arrivée.

B. Comment cela fonctionne: survol, pavage, mode continu et autres caractéristiques

38. *Statistics Netherlands in Your Neighbourhood* et *Neighbourhood Statistics in Google Earth* utilisent l'un et l'autre la technologie de Google. Les photos satellites ou photos aériennes servent à afficher des superficies avec plus ou moins de détails. Ces photos sont *pavées* dans le bon ordre de sorte que l'utilisateur a une impression de *survol* lorsqu'il consulte les données sur la carte. En raison de l'énorme quantité de photos nécessaires pour afficher les cartes possibles pratiquement sans fin, ces logiciels utilisent des techniques de diffusion en *mode continu*: l'utilisateur ne reçoit que les photos pertinentes et ces photos sont assemblées dans son ordinateur. Ce type de création de cartes découle d'une technique de cartographie à base de photos sous forme de pavés ou en mode continu.

39. Dans *Neighbourhood Statistics in Google Earth*, nous exploitons les possibilités en mode continu offertes par Google Earth. La couche correspondant à Google Earth contient 20 variables pour 11 000 voisinages et leurs bordures détaillées. Une simple application aboutirait à une couche de Google Earth de 60 MB. Une couche de cette taille est trop importante pour que le grand public puisse la télécharger facilement. De surcroît, elle entraînerait une dégradation de la qualité. Nous divisons donc les 11 000 voisinages en 450 segments, à savoir les communes. Un segment (c'est-à-dire les voisinages d'une commune) est uniquement téléchargé lorsque l'utilisateur agrandit suffisamment l'image d'une commune. D'après nos expériences, si une commune occupe plus de 450 pixels, ses voisinages peuvent être téléchargés dans la couche de Google Earth. Le téléchargement des données ne signifie pas que les voisinages apparaissent immédiatement. Pour de nombreuses régions, cela entraînerait un affichage fouillis dans lequel les limites et les noms seraient superposés. Nous ne montrons donc un voisinage que lorsqu'il nécessite plus de 150 pixels sur l'écran. De la sorte, les utilisateurs ont le sentiment d'une interaction immédiate.

40. Les données et limites à afficher sur la carte sont en KML (Keyhole Markup Language), langage conçu par l'ex-entreprise Keyhole qui a été rachetée par Google. Ce format, qui a été activement perfectionné par Google, est basé sur XML. Depuis 2007, KML 2.2 est une norme ouverte adoptée par l'OGC (Open Geospatial Consortium). D'autres géonavigateurs, Virtual Earth de Microsoft par exemple, utilisent de plus en plus le format KML. Il est donc possible d'afficher le même fichier KML sur différents géonavigateurs. Le format KML contient de nombreuses fonctionnalités pour l'affichage des marqueurs, des fenêtres d'information, des lignes, des rectangles, voire des modèles entiers en 3D (COLLADA).

41. Google Maps contient des images satellites et des images aériennes. Google Earth va un peu plus loin en offrant un éventail de possibilités interactives. Cette application n'est pas exploitée par navigateur mais doit être téléchargée. La surface de la terre est représentée à l'aide d'images satellites et d'images aériennes reportées sur un globe en trois dimensions. Une fonction «pan and zoom» (panoramique et zoom) efficace permet aux utilisateurs d'explorer n'importe quelle région du globe. Des couches peuvent indiquer la hauteur (renseignements en trois dimensions), ce qui fournit une dimension supplémentaire. Les fonctions d'analyse de Google Earth sont limitées. Elles sont plus développées dans *Statistics Netherlands in Your Neighbourhood* qui comporte par exemple des graphiques et des tableaux.

42. Google Earth a été lancé en 2005. Une version néerlandaise est disponible depuis mai 2007. Son introduction a été le point de départ d'une recherche à l'échelle mondiale d'images spectaculaires ou prêtant à controverse. On trouvera ci-après trois exemples choisis arbitrairement parmi une offre infinie.

Représentation graphique 7

Trois images détaillées de Google Earth aux Pays-Bas



43. Depuis le lancement de Google Earth, un certain nombre d'images ont fait l'objet d'une large attention de la part des médias. Les parties qui étaient classées et apparaissaient floues ont elles aussi suscité une grande attention. Dans plusieurs pays, l'affichage de zones sensibles telles que les centrales nucléaires, les installations militaires, les agences spatiales et les résidences des personnalités royales a donné lieu à un débat public. On continue de trouver de nouvelles utilisations à Google Earth. De nouvelles applications apparaissent dans les médias pratiquement chaque semaine. De plus, Google, la société mère, s'emploie activement à établir de nouveaux partenariats.

C. Objectifs et opportunités

44. L'un des principaux objectifs en passant à une nouvelle étape dans la diffusion de statistiques de voisinage était de les présenter de manière plus naturelle et dans un contexte déjà familier. Google est bien connu, et lorsque sa technologie donne en outre la possibilité de reconnaître son propre voisinage elle devient un moyen efficace et excellent de diffusion des statistiques de voisinage de Statistics Netherlands. De surcroît, la localisation des informations sur le voisinage est dorénavant plus élaborée qu'avec les moteurs de recherche classiques.

45. Les deux initiatives devaient en principe produire un va-et-vient sur le site Web de Statistics Netherlands. Et c'est ce qui s'est passé! Il en a été fait état à diverses reprises dans les médias nationaux et locaux ainsi que dans les magazines spécialisés de certaines branches d'activité. La possibilité de reconnaître le voisinage incite également de nouveaux groupes cibles à s'intéresser à Statistics Netherlands dont les publications touchent de ce fait une plus large audience. Plus important encore, les personnes interviewées ou qui pourraient l'être à l'avenir lui accordent davantage de crédit. Dans leur forme actuelle, les statistiques de voisinage sont un investissement pour l'avenir. Elles rapprochent l'Institut

national de la statistique du grand public. Tout un monde s'est ouvert, aux utilisateurs non professionnels en particulier. Statistics Netherlands, qui était à l'origine un institut chargé de produire des tableaux et publications statistiques traditionnels, se trouve maintenant à l'avant-garde de la diffusion des données par des méthodes novatrices.

V. Un regard tourné vers l'avenir: quelles sont les perspectives?

46. L'utilisation de même que la production des cartes se sont démocratisées avec l'Internet. Les résultats semblent souvent bons, en tout cas d'un point de vue technique. Ils peuvent cependant présenter une image totalement faussée de la réalité. Il est loin d'être évident de produire une carte (thématique) exacte. Les effets des choix concernant les symboles, les couleurs, la généralisation des limites et l'échelle utilisée sont souvent négligés. Certains utilisateurs ont véritablement l'impression que les images affichées avec Google Earth sont littéralement en temps réel. Nous en revenons alors à Carl et Nyls. Nous pouvons en conclure qu'en tant que service national de statistique nous portons une lourde responsabilité dans la diffusion des mesures statistiques.

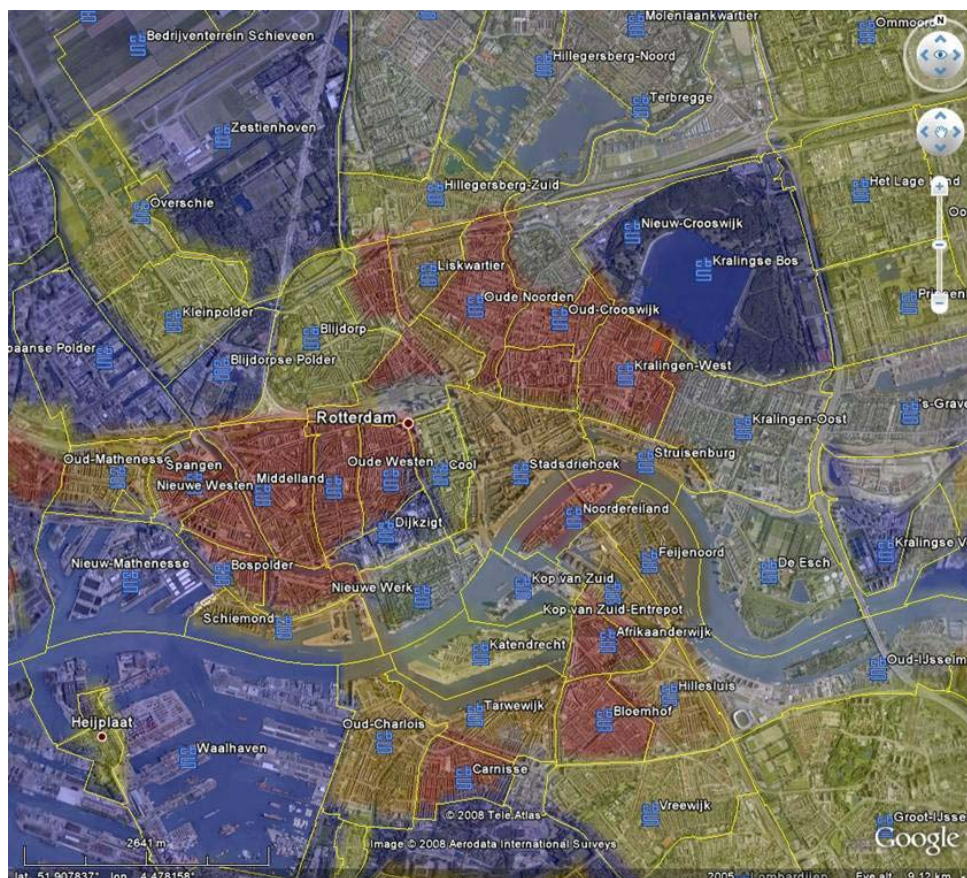
A. Comment résoudre ce dilemme

47. Les cartes sont une simplification de la réalité. La règle en matière d'élaboration de cartes thématiques est de ne montrer que la répartition géographique des données. Tous les détails superflus distrayant l'utilisateur doivent être laissés de côté. En utilisant des photos aériennes, on réintroduit la réalité dans les cartes. S'agissant des statistiques de voisinage, les photos aériennes présentent un avantage: les utilisateurs reconnaissent leur maison, leur rue et leur voisinage en utilisant le zoom de Google Earth. Pour créer une couche thématique correspondant au voisinage, il nous faut tenir compte de l'effet produit en associant des cartes abstraites aux images réalistes produites par Google Maps et Google Earth.

48. Le dilemme à résoudre consiste à trouver le moyen d'associer une carte thématique dans laquelle les caractéristiques géographiques sont superflues et des images aériennes très détaillées offertes par Google Earth. Nous avons trouvé une solution à ce problème.

49. Cette solution consiste à «mixer» une carte thématique et des photos aériennes en faisant appel à la transparence. De surcroît, nous faisons dépendre la transparence du niveau de zoom. Lorsqu'un utilisateur commence par avoir un aperçu du pays, la couche thématique (chloroplèthe) est opaque: ce n'est qu'une carte thématique juxtaposée sur Google Earth. Dès qu'il commence à agrandir l'image, nous augmentons la transparence de la couche et mélangeons concrètement la carte thématique et les photos aériennes. Avec cette solution, les utilisateurs passent sans encombre d'une couche purement thématique à une version composite dans laquelle ils peuvent reconnaître leur voisinage. Dans les expériences que nous avons réalisées, nous avons produit des cartes chloroplèthes très détaillées (en utilisant les polygones de voisinage). Au niveau du pays, ces cartes montrent des configurations de voisinage détaillées mais pas de photos aériennes. Lorsqu'un utilisateur agrandit l'image de son voisinage, la toile de fond aérienne commence à apparaître au travers. La densité de population dans les voisinages offre un exemple de ce qui peut être représenté de la sorte.

Représentation graphique 8 Densité de population dans plusieurs voisinages à Rotterdam



B. Travaux futurs

50. La présente note a décrit les démarches entreprises par Statistics Netherlands pour passer à la prochaine étape en matière de diffusion des statistiques de voisinage. Les techniques traditionnelles de cartographie ont été remplacées par des techniques de cartographie basées sur des photos en mode continu telles que celles utilisées par Google Maps et Google Earth. Ces initiatives sont d'un grand secours pour les utilisateurs désireux de voir rapidement les caractéristiques d'un voisinage particulier. Par contre, elles ne les aident pas beaucoup lorsqu'ils veulent avoir une vue d'ensemble de l'état d'une certaine caractéristique au niveau d'un district, d'une commune ou de tout le territoire des Pays-Bas. Il faudrait pour cela un mécanisme plus perfectionné qui ferait éventuellement appel à des «cartes à points de chaleur» (*heat maps*) pour afficher d'un seul coup l'état de la caractéristique aux Pays-Bas. Google Earth offre des possibilités très intéressantes d'utiliser à la fois la couleur et des modèles en 3D pour présenter les variables statistiques. Les premières expériences qui ont été faites sont très prometteuses.

51. Un autre domaine prometteur correspond à l'adjonction à Google Earth de *statistiques animées*. Ce logiciel comporte déjà une fonctionnalité pour tenir compte du temps, ce qui permet de visualiser des tendances statistiques dans les photos fournies en mode continu. On peut l'utiliser pour montrer l'évolution des limites des voisinages ou les changements intervenus dans la population urbaine au fil du temps. Les premières expériences qui ont été faites semblent elles aussi prometteuses.

52. La conception actuelle de ces deux initiatives présente un inconvénient, à savoir qu'elles sont tributaires de Google. Il est de notoriété publique que Google enregistre ce que les utilisateurs de ses services font sur Internet. Il y a également d'autres inconvénients. D'autres solutions que celle proposée par Google sont actuellement à l'étude dans l'espace public. Dès qu'une solution gratuite sera disponible, Statistics Netherlands redéterminera sa position en matière de cartographie à base de photos.

53. Sous l'impulsion de l'initiative européenne INSPIRE, Statistics Netherlands utilise également des normes convenues pour la publication de ses informations géographiques. D'ici peu, une technologie encore plus avancée, les WMS (Web Map Services) et les WFS (Web Feature Services) par exemple, sera utilisée et disponible pour d'autres portails.

54. Enfin, l'innovation la plus prometteuse consiste à exploiter plus avant les informations provenant de registres. L'association des SIG et des sources administratives offre un nouvel éventail d'options. En 2009, l'achèvement et la mise en application d'un «réseau dorsal d'adresses» ont permis d'élaborer de nouvelles statistiques spatiales, ce qui fournira à la longue des résultats à l'échelle spatiale concernant les murs du son, les crues ou la concentration des particules fines dans l'air. Ces résultats ne vont plus dépendre nécessairement des limites administratives, ce qui va ouvrir la voie à encore plus d'interactions, en géographie également. En conclusion, nous n'en sommes qu'au début.

VI. Bibliographie

Beeckman, D.J.D. (2002), Ruimte voor Statistiek: Regionaal statistische informatie in StatLine. Internal report Statistics Netherlands, Voorburg.

Beeckman, D.J.D. (2003), Kaarten en kaartgebruik bij het CBS. CBS website.

Beeckman, D.J.D. (2006), Regionale statistieken anno 2006: horizontaal, verticaal en digitaal. In: Geo-Info 2006-3.

Beeckman, D.J.D., de Jonge, E. (2008), CartoGraphy with a capital G: Google and more. CBS website

Beeckman, D.J.D., van Houwelingen, C.E., de Jonge, E. (2009), Regionale statistieken anno 2009: Het CBS in uw buurt. In Geo-Info 2009-12.

Biesheuvel, J.M.A. (1979), De kaartenmakers. In: De verpletterende werkelijkheid.

Bosch, ten O. and de Jonge, E. (2008), Visualisation of Neighbourhood Statistics Using Google Earth. Invited paper in Meeting on the Management of Statistical Information Systems (MSIS).

Bunschoten, L. (2007), Framework project 'Spatial, housing and mobility statistics'. Internal Project Initiation Document, Voorburg.

Bunschoten, L. (2008), Hoe ver woon ik van In: CBS Bevolkingstrends 2008-2.

CBS Law of 20 November 2003, Staatsblad 2003, 516, lastly amended by law of 15 December 2004. In: Staatsblad 2004, 695.

Elzaker, van C.P.J.M., de Graaf, P.J. and Beeckman, D.J.D. (2007), Testing the use of web maps in the retrieval and dissemination of regional statistical data. In: Proceedings of the ICA commission on Maps and the internet, Warsaw, Poland, 31 July - 2 August, 2007. Warsaw : University of Warsaw, 2007. pp. 25-31.

Graaf, de P.J. (2006), Users and functionalities of the cartographic entry. Dissemination of statistical data through interactive maps at Statistics Netherlands. Unpublished report of an

internship in the framework of the GIMA (Geographical Information Management and Applications) MSc programme.

Hijink, M. (2005), De Globe volgens Google. In: Nrc newspaper.

Houwelingen, van C.E. (2008), Statistics Netherlands' neighbourhood figures in Google Earth. In: CBS webmagazine.

Kraak, M.J. and Brown A. (2001), Webcartography.

Meer, van der A. (2008), Juist vertekenen om onvertekend over te komen: afscheidscollege prof.dr. F.J. Ormeling. In: Geo-Info 2008-5.

Melser, C. and Duimelaar, P.M. (2008), Buurtcijfers van het CBS op nieuwe website. In: CBS webmagazine.

Spee, B. (2001), Buurt in Beeld: Cartografische ontsluiting van de ruimtelijke statistieken bij het CBS. Thesis university of Utrecht.

Verweij, P. (2003), Geografische Informatie Systemen in de journalistieke praktijk.

Web-visualisations (some interesting examples):

- FlowingData, Examples of web-visualisations: <http://flowingdata.com/>
- GENeSIS (a research node of the National Centre for e-Social Science in the United Kingdom), Examples of web-visualisations: <http://www.casa.ucl.ac.uk/genesisblog/?cat=18>.
- Statistics Estonia, Exploring population statistics with the Internet-based mapping application Google Earth: <http://www.stat.ee/32707>.

United Kingdom Office for National Statistics, Boundary Viewer; discover and view boundaries: <http://www.neighbourhood.statistics.gov.uk/dissemination/LeadBoundaryViewer.do?xW=1680&xH=1050>
