



HAL
open science

Données brutes et produits finis. Deux expériences de restitution de données géographiques

Sébastien Oliveau

► **To cite this version:**

Sébastien Oliveau. Données brutes et produits finis. Deux expériences de restitution de données géographiques. Julie Beaujard. La restitution des données dans la recherche en sciences sociales : techniques et enjeux, pp.30-42, 2006. <hal-01140456>

HAL Id: hal-01140456

<https://hal.science/hal-01140456v1>

Submitted on 8 Apr 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

.DONNEES BRUTES ET PRODUITS FINIS :

DEUX EXPERIENCES DE RESTITUTION DE DONNEES GEOGRAPHIQUES.

Sébastien Oliveau, post-doctorant, UMR Géographie-cités.

La présentation faite ici reposera sur deux exemples complémentaires liés à ma (jeune) expérience de recherche. Le premier concernera l’AJEI, qui sera envisagé à diverses reprises comme un exemple de diffusion collective original, le second sera ma propre production (et je dois citer C.Z. Guilmoto avec qui ces travaux ont été développés). Certains points de mon introduction sont volontairement polémiques, afin de susciter une discussion par la suite, mais ils reflètent aussi en partie ma position de chercheur vis-à-vis de la société.

Aborder la restitution des données pose d’abord la question de savoir pourquoi on fait de la recherche (en général) et pourquoi on mène sa recherche (en particulier). Je commencerai par clarifier ce premier point qui constitue, à mon sens, le socle idéologique de ce questionnement. Il me semble que la première chose à faire est de se décomplexer par rapport à notre travail de recherche et son utilité sociale, puisque c’est finalement une des questions sous-jacentes.

Autrement dit, il s’agit de reposer les termes de l’interaction entre recherche dite « fondamentale » et société, en opposition, ou du moins en comparaison avec la recherche-action et le développement. La recherche fondamentale, c’est-à-dire non directement appliquée, est un processus de création de savoir, au sens large. Ce savoir peut, dans un second temps, être appliqué ou servir de matériau à une autre recherche qui pourra être ensuite appliquée.

De ce point de vue, la première action de restitution – et la seule qui soit obligatoire¹ – est la publication (à l’extrême limite on peut inclure la présentation orale, mais celle-ci se rapproche plus de l’enseignement, de la « performance » pour reprendre les termes d’Emmanuel Grimaud). Le chercheur qui ne fait "que" publier remplit déjà son contrat vis-à-vis de la société, et, si je peux l’exprimer ainsi, justifie déjà son utilité sociale.

Les applications possibles issues de son travail sont le rôle d’autres institutions, qu’elles soient gouvernementales ou non.

¹ Elle est obligatoire en ce sens que la recherche n’existe qu’à travers elle : une recherche qui n’est pas exprimée ne peut pas être considéré, du point de vue institutionnel/académique, comme une recherche.

Mais si on en reste là, cette journée n'aurait pas de sens... et peut être que l'AJEI perdrait aussi une partie de sa raison d'être. Donc, on dépassera cette vision juste, mais étroite, de la recherche, et surtout du chercheur. Celui-ci est aussi, ne l'oublions pas, un acteur du système qu'il étudie (c'est peut être encore plus sensible en sciences sociales²).

Deux bonnes raisons de ne pas en rester à la publication peuvent alors amener le chercheur à aller plus loin que la publication académique. La première est sa position de citoyen, particulièrement lorsqu'il travaille dans/sur son propre pays (ce qui n'est pas notre cas ici). La seconde est son rôle perturbateur et/ou prédateur du système qu'il étudie, ce qui est d'autant plus vrai que l'on fait plus de terrain. Je ne commenterai pas ce point, vous savez tous que lorsque l'on arrive quelque part, les gens s'arrêtent, et qu'il est finalement difficile, voire impossible, d'obtenir que les gens fassent « comme si » nous n'étions pas présents ! Ce dernier point justifie à lui seul l'effort de restitution qui nous réunit aujourd'hui. On nous offre du temps, qu'avons-nous à rendre en échange ? Enfin, nos recherches nous amènent souvent à accumuler des matériaux qui peuvent intéresser d'autres personnes, et rien ne justifie de conserver ces matériaux à l'abri de la connaissance générale.

En travaillant en Inde, nous nous plaçons dans une posture quelque peu particulière. Nous ne sommes pas citoyens de cet état, et nous venons travailler ici pour d'autres motifs. Le rendu d'une partie de notre production (la restitution) prend donc une forme particulière, une sorte de dédouanement, plus ou moins conscient, vis-à-vis d'une société qui n'a pas forcément demandé à devenir un terrain. Pour un chercheur français, la première restitution est donc de participer à la vie de son terrain, en s'insérant dans son réseau institutionnel. C'est-à-dire en essayant de participer à la vie scientifique (ce qui est particulièrement important pour les géographes en Inde, vu l'état de la discipline).

Cela passe tout d'abord par la nécessaire publication en langue anglaise et dans des revues et des ouvrages indiens, ainsi bien sûr que par la participation à des conférences en Inde (d'où la présence d'une journée entièrement en anglais au sein des ateliers jeunes chercheurs depuis leur origine).

Passées ces considérations générales, on doit replacer ce discours à l'échelle individuelle des jeunes chercheurs. On ne demandera certes pas à un étudiant de maîtrise (ni forcément à un nouveau doctorant) d'appliquer tous ces éléments évoqués, mais avant tout de les garder en tête (une sorte de fil rouge dans son parcours, une conscience que si une opportunité peut être

² Et on se rappellera ce qu'écrivait Gargantua à son fils Pantagruel : « science sans conscience n'est que ruine de l'âme ».

saisie de s'investir plus, il faut la prendre). Néanmoins, certains éléments sont faciles à mettre en pratique : rendre son mémoire accessible (publication en ligne par exemple - et l'AJEI offre cette possibilité³). Vis-à-vis des personnes rencontrées sur le terrain, une conclusion ou un résumé de son travail, traduit en anglais, est une solution à faible coût qui permet de restituer une partie de ce que l'on a pris.

Mais assez parlé en termes généraux, je vais présenter ici deux actions de restitution de données en Inde, basées sur mon expérience personnelle.

Mon travail de thèse repose sur l'analyse d'une base de données géographiques (pour mémoire, une base de données géographiques est une matrice d'information concernant des objets situés dans l'espace, voir l'intervention de Bertrand Lefebvre). Il s'agit dans ce cas précis de la numérisation des cartes du recensement et de leur croisement avec les tableaux de données effectués de 1998 à 2000 au sein de l'Institut Français de Pondichéry (programme SIFP). Une fois ces données mises en place, l'idée de les restituer est vite devenue une évidence : ces données étaient toutes issues du domaine public, et les restituer, c'était rendre ce que l'on avait emprunté, y ayant ajouté une valeur supplémentaire : leur géoréférencement.

Ces données brutes, pour être utilisées, nécessitaient l'usage d'un logiciel spécifique : un Système d'Information Géographique (SIG). Le poids que représentait cet ensemble de données était une contrainte supplémentaire. 17 000 points (villages) et 250 polygones (villes) représentent un nombre important d'objets à cartographier, analyser et surtout transférer.

Mettre simplement à disposition les données sur Internet nous semblait donc limité, et nous avons dû envisager une autre méthode (la figure 1 résume l'état général du projet).. C'est un point important, car la restitution au public, c'est-à-dire en dehors de la sphère scientifique, nécessite souvent des efforts particuliers, pas uniquement en terme de diffusion, mais aussi en terme de mise en forme.

³ Pour les thèses de doctorat, on doit l'existence d'un serveur de thèse en ligne au CNRS : <http://tel.ccsd.cnrs.fr/>

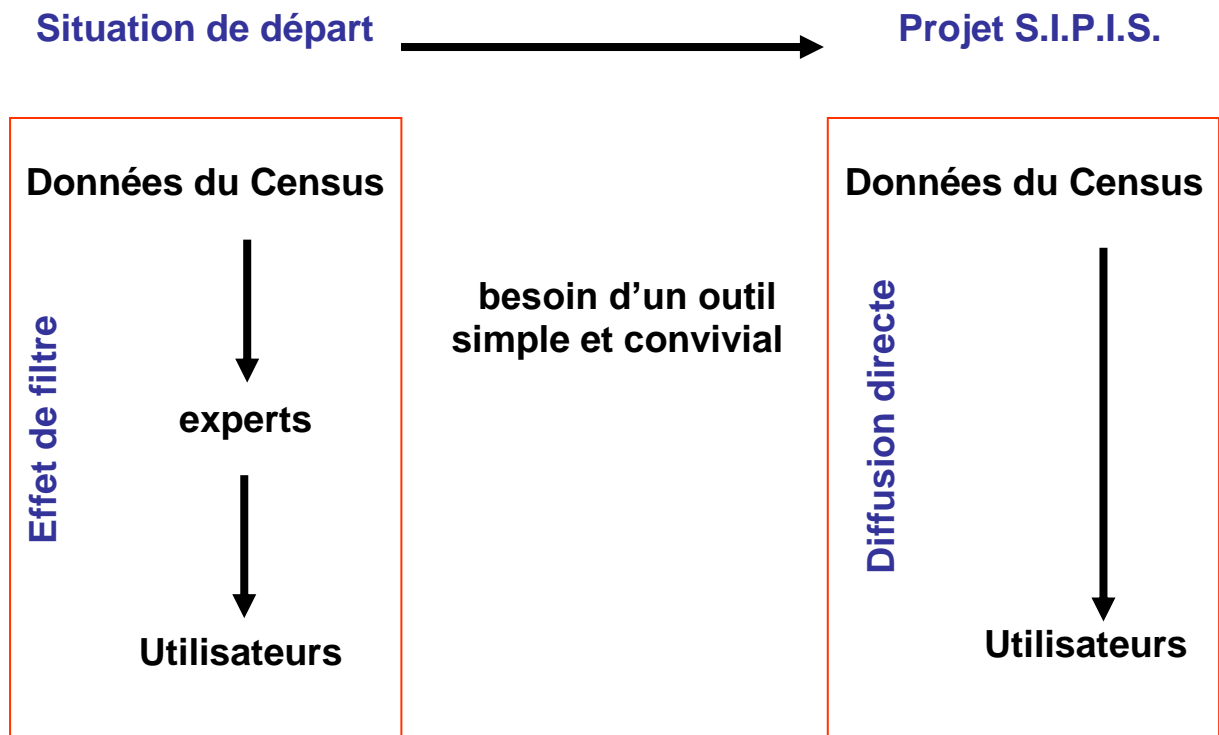


Figure 1 : rendre accessibles les données censitaires.

Nous avons donc cherché un financement qui nous permettrait de développer un SIG spécifique, dédié à l'utilisation de ces données. Les SIG, rappelons-le, sont des logiciels assez complexes, qui nécessitent des formations particulières. Le Fonds des Nations Unis pour la Population (FNUAP) a accepté de subventionner ce projet, qui a nécessité ensuite un an de développement [figure 2]. La figure 3 présente les apports et les limites spécifiques à cette solution.

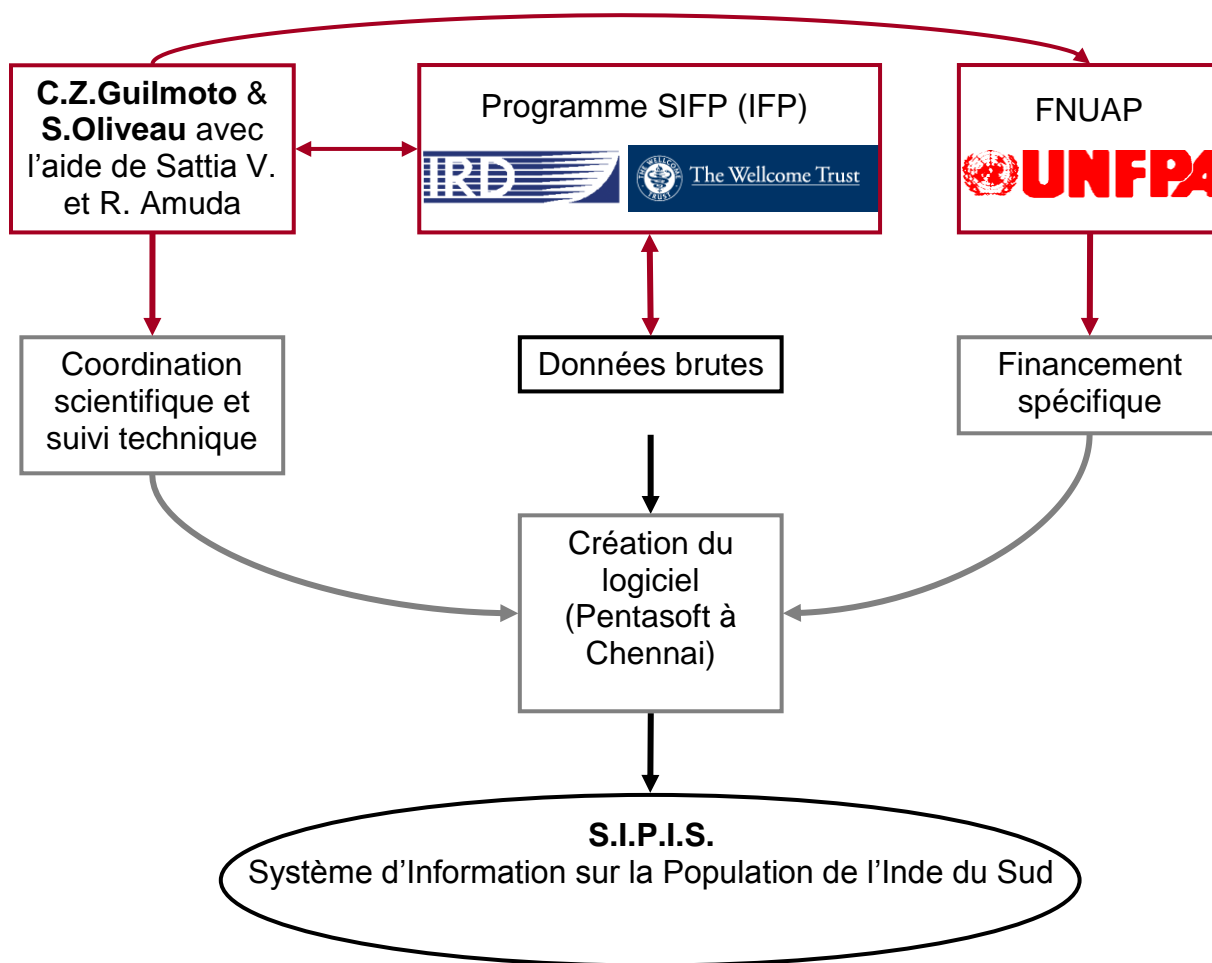


Figure 2 : schéma de l'organisation du projet SIPIS.

Apports	Limites
Corpus complet de données brutes disponible	Pas d'analyse fournie
Données supplémentaires ajoutées	Utilisateur seul face aux données (cartographie)
Coût nul pour l'utilisateur (grâce au financement)	Coût élevé de production
Variété du public atteint : autres chercheurs et étudiants, ONG, administrations, entreprises privées	Difficulté d'accès au produit (CDROM => support physique)

Figure 3 : Apports et limites du projet de CD-ROM SIPIS pour la diffusion de l'information géographique du recensement.

Suite à la diffusion de ce produit, nous avons eu quelques retours, insistant notamment sur les limites du produit, qui concernaient la « solitude » de l'utilisateur face à cette masse de données ainsi que la difficulté d'accès au CDROM (distribué uniquement par l'Institut Français de Pondichéry). De plus, des échanges, notamment avec les membres du census⁴, ont souligné l'intérêt de mettre en ligne ces données (d'ailleurs, le recensement l'a fait à l'échelle des districts sur son site www.censusindiamap.net). Cela a déterminé le choix de développer un second outil, conçu comme complétant le précédent, à savoir un atlas en ligne.

L'atlas guide le lecteur, puisqu'il fournit une carte finie et commentée. De plus, l'utilisation de cartes dans le discours des sciences sociales indiennes étant limitée, nous avons fait des choix spécifiques qui n'auraient certainement pas été les mêmes en France. Les choix présidant à la réalisation de cet atlas ont été de deux ordres : scientifique et technique.

- Scientifique en ce qui concerne la cartographie choisie (lissage spatial des données) et les commentaires (courts pour faciliter la lecture, volontairement généraux pour ne pas rebuter les non scientifiques).
- Technique dans sa mise en forme en ligne : interactivité limitée, page "légère" permettant un affichage rapide même avec une connexion lente.

La figure 4 montre un schéma similaire à celui de la figure 2 pour souligner les différences d'organisation dans notre démarche de mise en place de l'atlas par rapport à l'organisation du SIPIS.

⁴ Ceux-ci avaient choisi de se réunir à l'IFP pour lancer officiellement la campagne de préparation des Districts Census Handbooks.

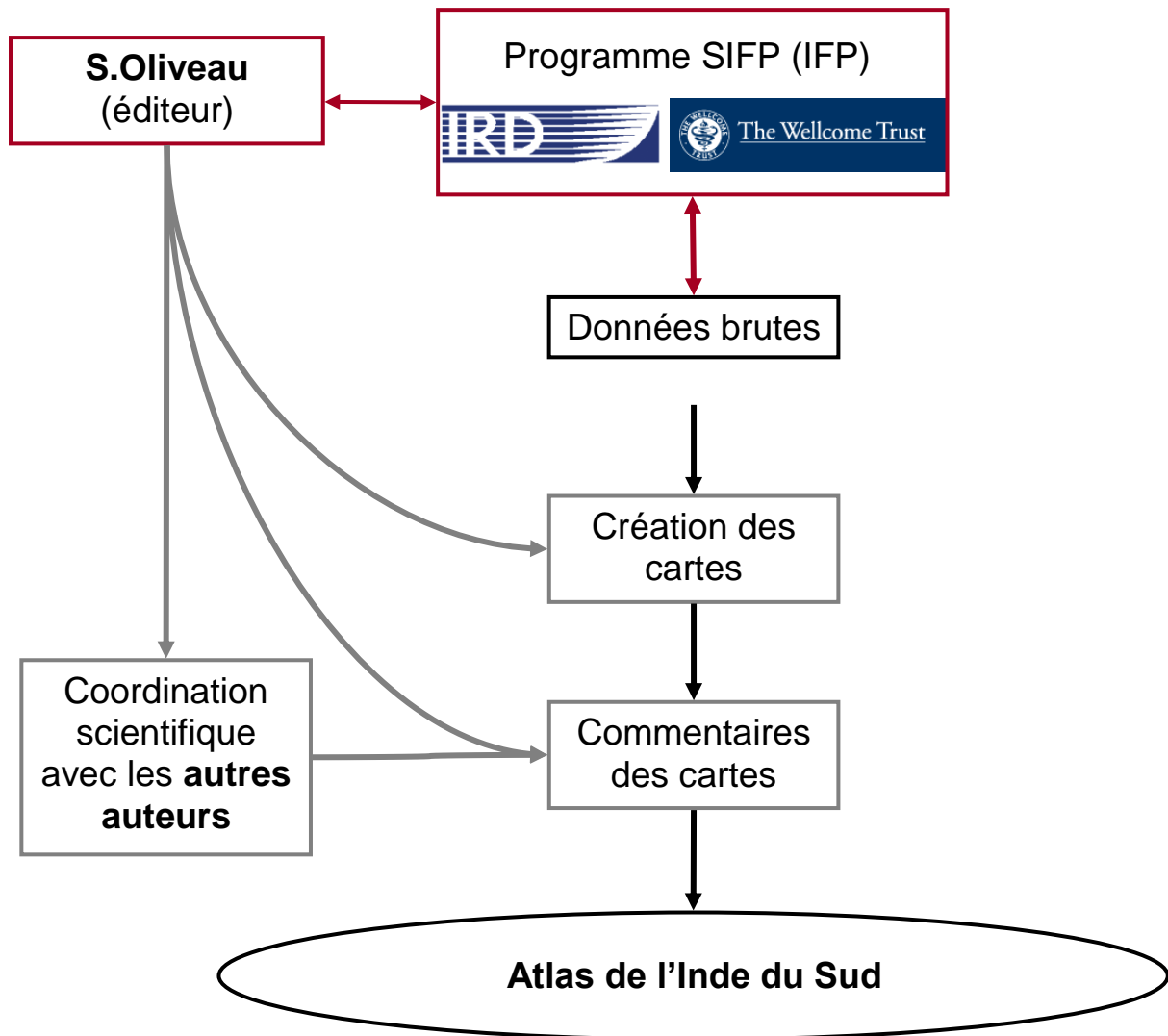


Figure 4 : schéma de l'organisation du projet d'atlas en ligne.

Cet atlas est aujourd'hui publié dans la revue *Cybergéo* (www.cybergeo.presse.fr). Ce qui tend à montrer (en passant) qu'une œuvre de restitution peut aussi être bénéfique, puisque cela constitue une publication (même si les atlas sont considérés comme des formes de valorisation de la recherche et non pas des produits scientifiques réels). Si nous avons proposé la publication à cette revue, c'est qu'elle participe depuis 9 années déjà à ce travail de diffusion libre des travaux géographiques sur Internet⁵, en ayant acquis (grâce à une vraie politique scientifique) la reconnaissance de ses pairs. Je ne veux pas développer ici un discours déjà tenu ailleurs (voir l'édito du moi de mai 2005 de *Cybergéo*, ou l'argumentaire développé dans Oliveau 2004), mais cette approche de la libre diffusion en ligne est encore loin d'être acquise

⁵ Je ne peux m'empêcher ici de faire un parallèle avec le travail de l'AJEI, qui met en ligne depuis 1999 tous ses travaux et ceux de ses membres.

dans le milieu scientifique et fait toujours l'objet de débat⁶. La figure 5 rappelle les apports et les limites de ce projet d'atlas en ligne.

Apports	Limites
Données déjà traitées (analyse préparée par des experts)	Données traitées selon un angle défini en amont (top-down)
Coût nul de production	Pas d'interactivité
Accès décentralisé autonome et gratuit (Internet en accès libre) www.cybergeogeo.presse.fr	

Figure 5 : apport et limites de l'atlas en ligne.

L'avenir est certainement à la mise en ligne libre et au mélange des genres. Ainsi, les atlas interactifs en ligne qui permettront au lecteur de participer à la création de la carte, d'en interroger les données directement tout en ayant à disposition des cartes déjà finies et commentées. Bertrand Lefebvre et Pierre Chapelet y travaillent sur leur projet, et nul doute que ce genre de méthode va se répandre.

En conclusion, je rappellerai que la restitution, quelque soit sa forme, renvoie souvent le chercheur vers de nouvelles questions et intervient donc dans son évolution personnelle. Ce n'est donc jamais un exercice à pure perte (entendue comme altruiste au sens où l'on n'en retirerait aucune forme de bénéfice), même si ce n'est pas le plus "rentable"⁷ en terme de carrière quand elle ne vise pas la communauté scientifique (le monde académique ne juge pas sur la restitution non académique), c'est-à-dire qu'il n'est pas une fin en soi mais un détour sur un parcours scientifique.

⁶ Mais on doit souligner l'engagement fort du CNRS, à travers son serveur de thèse, évoqué plus haut, et son projet intitulé, avec humour, HAL (Hyper Article en Ligne). Sur le site <http://ccsd.cnrs.fr/>

⁷ Soyons très clair cependant, cela ne constitue absolument pas un handicap, et est souvent valorisé de façon indirecte, comme une capacité à s'investir en dehors d'une vision étroite de la science.

La bibliographie qui illustre ce texte est la suivante (une bibliographie plus complète sur ce thème est disponible dans Oliveau (2004) :

SIPIS :

Guilmoto, C. Z., Oliveau, S., 2000, *South Indian Population Information System*, Volume 1 : Tamil Nadu & Pondicherry, Institut Français de Pondichéry, CDROM.
(<http://www.demographie.net/sipis/>)

Atlas :

Oliveau, S., 2003, « Atlas of South India », *Cybergéo*. (<http://www.cybergegeo.presse.fr>).

L'atlas en ligne du recensement: <http://www.censusindiamap.net/>

A propos de l'expérience du SIG sur le Sud de l'Inde :

Guilmoto, C. Z., Oliveau, S., Chasles, V., Delage, R., Vella, S., 2005, *Mapping out social change in South India*, Pondy Papers in Social Sciences n°31, Institut Français de Pondichéry, Pondichéry, 117 p.

Guilmoto, C. Z., Oliveau, S., Vingadassamy, Sattianarayanan, 2002, « Un système d'information géographique en Inde du Sud : Théorie, mise en œuvre et applications thématiques », *Espace, Populations et sociétés*, Lille, p. 147-163.

Oliveau, S., 2004, « Geographic Information in the ICT's era: what has changed, and how ? », in Prasad K., (Ed.), *Information and Communication Technology: Recasting Development*, B.R. Publishing, New Delhi, p. 225-252.