

Réflexions sur l'externalisation des opérations de saisie de données d'enquêtes

Papa Thiécouta NDIAYE
Informaticien ANSD

I – CONTEXTE ET OBJECTIFS

L'Agence Nationale de la Statistique et de la Démographie (ANSD) est une structure administrative chargée d'assurer la coordination technique des activités du système statistique national et de réaliser elle-même les activités de production et de diffusion des données statistiques pour les besoins du gouvernement, des administrations publiques, du secteur privé, des partenaires au développement et du public.

La Direction du Management de l'Information Statistique, créée au sein de l'Agence, regroupe toutes les fonctions allant du traitement informatique des données, de leur archivage dans des bases de données, la gestion de centres de documentation, jusqu'à leur diffusion, sous toutes les formes (électronique, papier, CD Rom) aux utilisateurs.

C'est ainsi que toutes les tâches allant de la saisie à l'apurement des données des enquêtes réalisées par l'Agence lui reviennent, ceci en rapport avec la qualité et les délais requis pour le respect des engagements pris envers les partenaires.

L'objectif principal de cette réflexion est de se donner les moyens de disposer des données de qualité dans les délais requis.

II - DIAGNOSTIC DE LA SITUATION

L'idée d'externaliser la saisie de données d'enquêtes a été posée depuis l'ex DPS à cause des retards constatés dans les délais d'exécution de certaines opérations d'enquêtes et de recensement.

Il faut dire aussi que l'appréciation négative faite au traitement des données est très souvent basée sur l'expérience de l'ex DPS et d'enquêtes menées pendant la période de transition vers l'Agence.

Les phases de collecte rencontrent souvent des difficultés en effet les questionnaires qui reviennent du terrain ne sont pas toujours de la meilleure qualité possible et on a noté souvent des problèmes de suivi et de remplissage des questionnaires (ratures, inversion de l'ordre de remplissage, données manquantes, calculs inexacts etc.).

L'étude fera donc un diagnostic de tout le processus de traitement des données d'enquête, notamment celui du Recensement de la Population et de l'Habitat (RGPH) de 2002 afin de proposer des solutions.

II.1 Le traitement du RGPH2002

Il sera intéressant de rappeler quelques points qui pourraient expliquer les difficultés rencontrées à ce niveau :

- 1) L'opération n'a pu démarrer que deux années après la collecte faute de financement ;
- 2) Le local qui a servi au démarrage de l'opération n'était pas adéquat étant donné qu'il était logé dans un centre social, et qu'il était situé à côté d'un terrain d'entraînement de basket et d'une salle de répétition d'un groupe de ballet folklorique ce qui occasionnait beaucoup de bruit ;
- 3) De plus, des coupures intempestives de courant ont été la raison principale du retour de l'opération au Point E. Ces difficultés ont généré un important retard. En effet, on a constaté que le volume de documents saisis durant les quatre mois passés au centre social équivalait à la quantité de travail fournie pendant les deux premiers mois passés au Point E ;
- 4) Cet éloignement des locaux de la Direction a aussi occasionné beaucoup de problèmes de supervision;
- 5) Après un mois de saisie, 55 agents (le tiers du personnel) ont été envoyés en renfort à la codification qui n'arrivait pas à alimenter convenablement la saisie ; situation qui s'est prolongée jusqu'à la fin de l'opération ;
- 6) A cause des lenteurs constatées pour alimenter la saisie, il a été demandé à la codification de ne plus mettre l'accent sur les contrôles, mais que la résolution des problèmes soit différée à la fin de la saisie par des imputations et redressements sur lesquels on n'avait malheureusement pas beaucoup réfléchi, ce qui occasionna un relâchement des contrôles à ce niveau ;

- 7) On a aussi noté un problème de gestion du personnel puisque l'équipe de supervision a été souvent confrontée à des actes d'indiscipline (retard, absence non justifiée, bavardages dans la salle, utilisation du téléphone portable, repas dans la salle etc.). Citons le cas de l'agent renvoyé qui a été repris après qu'il se soit excusé. Une certaine complaisance des superviseurs vis-à-vis des agents de saisie a aussi été remarquée. Il a fallu recourir aux services d'anciens retraités de la Direction de la Statistique pour asseoir une plus grande autorité au niveau de la supervision ;
- 8) La gestion des sacs de DR a aussi été source de problème notamment en ce qui concerne leur déplacement entre les salles de stockage du Point E, de la codification de Liberté VI et leur retour à la saisie. Cela occasionna souvent des ruptures d'approvisionnement au niveau de la saisie faute de moyens de transport et l'égaré de DR qui n'ont été finalement ni codifiés ni saisis ;
- 9) A un mois et demi de la fin de l'opération, la saisie a marqué une pause pendant un mois par manque de financement. Des moyens financiers ont été par la suite débloqués par l'ANSD et il a fallu recourir aux services d'un privé et un plan d'exécution (barème de paiement, nombre d'agents de saisie et de superviseurs) a été proposé. L'exécution du marché a été confiée à un GIE. Le même schéma de supervision a été reconduit avec comme innovation majeure que les agents ont été payés à la tâche selon le nombre de questionnaires saisis. Il faut reconnaître qu'avec cette méthode, la saisie a pu être finalisée dans les délais prévus. Cependant, un bon dispositif de contrôle de la qualité des données n'a pas suivi, comme par exemple, un système efficace de double saisie pour valider les données (moyen efficace de contrôle de qualité des données). Par ailleurs, en suivant les agents qui ont confectionné le répertoire des villages, il a été noté que la plupart des erreurs corrigées provenait des fichiers saisis durant cette période ;
- 10) Le Bureau Américain du Recensement, lors de ses différentes missions de préparation, avait recommandé et schématisé un certain nombre d'activités (le plan de rayonnage et de stockage des DR, la saisie en réseau, le mode de transmission et de gestion des sacs etc.), mais malheureusement, aucune de ces recommandations n'a été appliquée.

II.2 La saisie des enquêtes

En général, on constate trois principales difficultés dans l'exploitation des enquêtes :

- Le dépassement des délais (saisie et apurement trop longs) ;
- La mauvaise qualité du remplissage des questionnaires ;
- Le défaut d'accompagnement et de communication entre informaticiens et statisticiens; problème qui a été constaté au démarrage de la saisie de l'EDMC.

Elles sont souvent dues :

- A une mauvaise estimation du temps de saisie et d'apurement liée à un manque de coordination ou à des erreurs d'appréciation. En effet, une bonne estimation devrait se baser sur les éléments suivants : vitesse moyenne par heure d'un agent de saisie, taille de l'échantillon, nombre de variables du questionnaire, nombre de machines etc., ce qui n'est souvent pas le cas. L'avis des informaticiens chargés du traitement des données d'enquête n'est pas sollicité lors de la réalisation du chronogramme de l'enquête ; ceux-ci se trouvent dès lors confrontés à un délai trop court imparti à la saisie.
- Au mauvais remplissage des questionnaires occasionnant beaucoup de données manquantes, de codes erronés, etc. L'expérience de la dernière exploitation de l'enquête SASDE a montré que des étapes importantes ont été ignorées dans les opérations telles que les phases de relecture et de correction des questionnaires avant la saisie ;
- A la non participation des statisticiens aux ateliers de saisie. Les statisticiens chargés de l'enquête devraient assurer une présence effective auprès des informaticiens pour mieux les aider à gérer certains problèmes (erreurs de cohérence, données manquantes, etc.). Ils doivent être disponibles durant toute la période de saisie et d'apurement. De plus, ces derniers doivent être impliqués dès le départ au montage du projet afin de bien maîtriser les outils (questionnaire, manuel). Enfin, les spécifications de contrôle doivent leur parvenir à temps pour qu'ils les traduisent en programme ;
- Au fait que l'enquête pilote ne soit pas suffisamment exploitée pour tester les programmes de saisie et d'apurement afin de déceler tous les problèmes et ainsi permettre une bonne estimation de la durée de la saisie.

Cependant, Il faut noter que des réflexions ont été engagées sur de nouvelles méthodes de traitement de données d'enquêtes. Celles ci ont été appliquées sur les dernières enquêtes pour essayer de pallier ces difficultés notamment le verrouillage du programme de saisie, la relecture des questionnaires avant la saisie et enfin la double saisie. Leur efficacité s'est avérée avec un bouclage dans les délais avec une meilleure qualité des données.

III - PROPOSITIONS DE SOLUTIONS

Plutôt que de parler de choix de scénario, nous allons dans un premier temps revisiter tous les problèmes posés et proposer des solutions dont la mise en œuvre dépendrait en grande partie de la volonté de la Direction Générale. Aussi une réflexion globale et profonde devrait être menée pour se donner toutes les chances de réussir la mise en œuvre et le traitement des enquêtes. Il s'agira surtout de ressortir tous les points à satisfaire ou à améliorer :

1) Sur l'organisation de la collecte :

- Améliorer le processus de recrutement des enquêteurs en ne choisissant pas systématiquement les anciens ;
- Améliorer l'encadrement sur le terrain en mettant d'importants moyens en œuvre tels qu'une meilleure implication du personnel permanent, des moyens de déplacement importants, une meilleure sensibilisation des populations cibles, etc. Cette phase devrait davantage relever du personnel permanent ;
- Evaluer les agents enquêteurs et mettre en place un système de motivation lié à la validation des questionnaires reçus ;
- Etablir une liste rouge des agents qui n'ont pas donné satisfaction lors de précédentes opérations ;
- Insister surtout sur les bonnes pratiques pour mener une enquête.

2) Sur l'exploitation des données :

- Adopter complètement ou en partie le modèle de traitement des données des enquêtes MICS et EDS où le programme est complètement verrouillé avec un système de codification complet (créer des codes pour les données manquantes, inclure la modalité « ne sait pas » etc.) pour ne pas bloquer la saisie ;
- Prévoir la relecture des questionnaires avant la saisie (vérification, comptage, corrections des questionnaires) ;
- Prévoir la double saisie ;
- Etablir aussi une liste rouge des agents qui n'ont pas donné satisfaction lors de précédentes opérations.

Ces méthodes de traitement qui ont démontré leur efficacité, vont nécessiter des ressources financières importantes entraînant le doublement de la charge salariale lorsqu'il s'agit d'une double saisie intégrale. Elles ont été élaborées et conçues par la société MACRO. Tous les experts travaillant avec ces méthodes ont été formés par cette société et travaillent comme consultants dans la plupart des pays africains sur l'exploitation des MICS et EDS.

Une analyse de tous les points évoqués plus haut, amène à dire qu'il y a une forte interaction de toutes les étapes dans une enquête. On ne peut pas extirper le traitement informatique de tout le processus, Il faut plutôt chercher à régler les difficultés globalement. En effet beaucoup de problèmes proviennent du terrain et il faut les solutionner avant, pendant et après la saisie. Souvent, des retours de terrain doivent être envisagés. Aussi, la solution d'externaliser uniquement la saisie ne parait pas tout régler, car il doit y avoir une forte implication de tous les acteurs dans le processus pour assurer une bonne réussite de l'opération. Avant de passer au traitement informatique, il est important d'essayer de régler les problèmes de la collecte pour garantir une meilleure qualité des données. La question est de voir s'il est possible de d'assurer une bonne qualité des questionnaires avant de les transmettre à un cabinet pour ne pas recourir constamment aux compétences des équipes de statisticiens et d'enquêteurs de l'Agence. Ceci nous amène à proposer la

forte recommandation de ne jamais s'orienter vers une semi externalisation (terrain exécuté par l'agence et la saisie par un cabinet) mais plutôt que toutes enquêtes soient entièrement réalisées et exploitées soit par l'Agence ou soit par un cabinet privé avec une collaboration du personnel de l'Agence.

Il y a aussi d'autres éléments d'appréciation qui peuvent influencer les orientations que l'Agence pourrait décider. Il s'agit :

1. du type d'enquête ;
2. du financement ;
3. du mode d'exécution ;
4. d'asseoir une meilleure compétitivité de l'Agence.

Les types d'enquêtes les plus connus sont :

- les enquêtes auprès des ménages ;
- les enquêtes auprès des unités de production ;
- les enquêtes MICS, EDS, SASDE ;
- l'urbanisation et la migration ;
- le recensement de la population ;
- l'emploi, le travail des enfants ;
- les monographies ;
- les enquêtes de conjoncture ;
- les enquêtes particulières (TDCI, etc..).

Ces différentes opérations ont une portée aussi bien nationale que régionale avec un échantillon très variable et elles peuvent être initiées de plusieurs façons :

- Par l'Agence Nationale de la Statistique et de la Démographie ;
- Par un partenaire ou bailleur de fonds (UNICEF, Banque mondiale, membres du SSN) ;
- Dans le cadre d'une coopération régionale (AFRISTAT), la même enquête étant réalisée dans tous les pays membres.

Les coûts supplémentaires (marges bénéficiaires du cabinet, etc.) qui pourraient en découler, seront-ils bénéfiques pour l'Agence ?

Le faire faire va aussi entraîner le dépérissement technique du personnel de l'agence intervenant dans les enquêtes et finalement ces structures extérieures commanditaires d'études ne finiront elles pas par confier ces opérations directement à ces bureaux d'études ?

Les cabinets spécialisés dans le traitement de données d'enquêtes ne sont pas très nombreux dans le pays, les plus connus sont SYSCOM, CRDH et SONAGROSOL alors que l'Agence devrait se positionner au moins à un même niveau d'exécution que ces structures.

En fonction de ces éléments, les orientations de l'Agence pourraient tournées autour des points suivants :

1. Une autonomie totale de l'Agence qui peut décider de confier toute l'exploitation à un cabinet, son rôle se limitant à valider les données. Dans ce lot, on pourrait retrouver les enquêtes monographiques, les enquêtes auprès des unités de production et des entreprises, les enquêtes auprès des ménages. Il s'agit dans ce cas de ne pas dissocier les opérations de terrain et le traitement informatique, le tout étant confié à un cabinet. Cependant l'avis du bailleur de Fonds ou du partenaire pourrait être aussi sollicité.
2. Une dépendance de l'Agence vis-à-vis du commanditaire qui peut proposer l'utilisation d'experts qui vont venir avec leurs méthodes et programmes, le rôle de l'ANSD se limitant à recruter un personnel et à assurer le suivi de l'exploitation. Comme exemple, citons les enquêtes EDS où c'est le cabinet MACRO qui réalise les programmes de traitement et le SASDE. Dans ce cas, il est préférable que l'Agence fasse entièrement le terrain et le traitement.

3. Une semi autonomie avec les opérations à dimension régionale comme les enquêtes MICS avec l'UNICEF, les enquêtes 123 ou encore EDMC pilotées par AFRISTAT où des programmes informatiques sont réalisés par l'organisme partenaire et mis à la disposition de l'Agence à l'occasion de séminaires d'appropriation ou par d'autres moyens pour être adaptés au contexte du pays. Les décisions prises par consensus sont mises en œuvre au niveau de chaque pays (le processus de création de variables d'analyse, le calcul de certains indicateurs clé), surtout que ces partenaires apportent des appuis techniques du début à la fin du projet. Il est proposé que l'Agence fasse les enquêtes et le traitement de bout en bout.
4. Concernant le Recensement Général de la Population et de l'Habitat, il mérite beaucoup d'attentions dans les choix futurs puisque s'agissant d'une opération très délicate et coûteuse où un système d'organisation efficient doit être mis en place à partir de la confection du budget (phase très importante), la cartographie, la collecte, la codification et la saisie qui sont toutes très liées. On devra plutôt se limiter à confier le recrutement et la supervision administrative des agents à un cabinet.
5. Une surcharge d'enquêtes menées par l'Agence peut amener à confier des activités à un cabinet.

En conclusion, concernant les points 2, 3 et 4 mentionnés ci avant, il est proposé que l'Agence prenne en charge l'exploitation technique (écriture de programme, hébergement et supervision technique) et lorsque le personnel à recruter dépasse un nombre qui est fixé (15 par exemple) comme pour l'exploitation d'un recensement, que la gestion des ressources humaines soit confiée à un cabinet qui devra fournir et superviser administrativement le personnel qu'il mettra à la disposition de l'Agence. L'Agence aura toujours la possibilité de s'orienter vers un choix après une analyse de la situation avec les points 1 et 5.

Deux autres éléments doivent être pris en compte dans la réflexion sur l'externalisation du traitement des enquêtes :

1. La décentralisation de la saisie au niveau des régions.

Lorsque la saisie doit se faire à un niveau régional, avec la connexion future des services régionaux au réseau de l'ANSD, il sera plus judicieux que l'Agence prenne en charge toute l'exploitation.

2. L'introduction des NTICs.

L'Agence a la volonté de s'orienter vers la collecte en ligne, notamment par le WEB, les PDA (ordinateurs de poche) ou le traitement des questionnaires par lecture optique. Actuellement, ces technologies ne sont pas maîtrisées par le personnel de l'Agence et demande une introduction progressive. Il sera difficile à un expert de l'Agence de recommander que le lecteur optique soit utilisé pour le prochain recensement de la population et de l'habitat. De par les échos de quelques expériences en Afrique, hormis le Maroc où on a annoncé la réussite de l'opération, on parle aussi de fiasco dans d'autres pays. Cette option doit demander l'avis et des études sérieuses d'experts qui les ont bien expérimentées ailleurs avec succès et qui accepteraient d'accompagner l'Agence dans tout le processus de traitement et ensuite de procéder à un transfert de technologie et de compétences pour les opérations futures. Ces nouvelles technologies notamment le lecteur optique sont très sensibles et la mauvaise manipulation des questionnaires du fait de la chaleur ou de la poussière peut amener beaucoup de difficultés au niveau de la lecture et entraîner des pertes de données considérables. Il faut aussi veiller à la bonne qualité du papier, à un écritoire adapté, à un remplissage très lisible des questionnaires en évitant les ratures. D'où la nécessité de bien étudier la faisabilité et de prendre toutes les précautions avant de s'engager dans de telles opérations surtout dans le cas d'un recensement ou d'une enquête de très grande envergure. Par contre, l'utilisation progressive des PDA ou de l'Internet pose moins de problème. Il faut en tout cas expérimenter ces nouvelles méthodes d'abord sur des enquêtes de petite envergure.

Par cet exercice, on a voulu réunir tous les éléments pour permettre à l'Agence de prendre à chaque fois la décision qui lui convient. Il s'agira alors d'étudier au cas par cas les situations se présentant et de s'orienter vers un choix en fonction du contexte. Mais au départ, la possibilité de s'orienter vers une externalisation, doit être prise en compte dès la confection du budget.

IV. ELEMENTS DE CHOIX D'UN CABINET

L'ANSD doit d'être rigoureuse dans le processus de sélection des cabinets. Il sera attendu d'eux, d'avoir des compétences avérées dans la conduite des grandes enquêtes ou des enquêtes nécessitant des techniques spécifiques. L'Agence doit s'entourer de toutes les précautions pour garantir son image et sa crédibilité. Elle doit agréer d'abord un ensemble de cabinets prestataires et établir un cahier de charges spécifique au type d'enquête. De plus, un processus de validation doit être clairement établi.

V. CONCLUSIONS

L'externalisation de la saisie des données d'enquête doit être abordée avec un maximum de précautions puisqu'on parle de confidentialité des données individuelles. Aussi les questions posées sont celles-ci lorsqu'une enquête est confiée à une structure externe, qu'en sera-t-il de la gestion des données ? Sera t-il possible à l'agence de mettre suffisamment de garde fous pour les protéger ? On a aussi tendance à bruler des étapes dans le processus de traitement par exemple pour le recensement, les méthodes d'imputation et de redressement auraient du être abordées sérieusement au début de l'opération et le programme informatique qui en découle prêt au moment de l'enquête pilote. Un pays comme le Bénin a constitué une équipe de 3 démographes et 2 informaticiens pour réfléchir pendant 3 mois uniquement sur les critères d'imputation et de redressement.

Le défaut de personnel qualifié peut militer aussi pour l'externalisation, mais si l'Agence a la vocation de vouloir générer des ressources dans l'exécution de sa mission, n'est-il pas plus judicieux de renforcer le personnel et de mettre les moyens adéquats afin de pouvoir la remplir correctement ? Il faut alors promouvoir le développement de compétences internes à tous les niveaux : personnel d'enquêtes, statisticiens et informaticiens. Les agents contractuels (agents de saisie, de traitement, enquêteurs et contrôleurs) ne sont pas très souvent en phase avec les objectifs de l'Agence et peuvent avoir des comportements irresponsables. Il est impératif à ce niveau de bien les conscientiser à défaut de pouvoir en faire un personnel permanent. Cet effort à mener, pour assurer une bonne qualité des enquêtes, doit être un début pour permettre à l'agence de bien réussir ses opérations en interne.