

Article

Symposium 2008 :
Collecte des données : défis, réalisations et nouvelles orientations

Intégration d'information sur fichiers audio dans la collecte des données

par Charles Loftis et Shane Trahan

2009



Intégration d'information sur fichiers audio dans la collecte des données

Charles Loftis et Shane Trahan¹

Résumé

L'enregistrement audio des interviews peut être un outil efficace et polyvalent de collecte des données. Cependant, il peut produire de gros fichiers dont la gestion est fastidieuse. Les progrès techniques, y compris de meilleurs outils de développement de logiciels audio et l'adoption croissante de connexions Internet à large bande, ont allégé le fardeau de la collecte de données audio. Le présent article porte sur les technologies et méthodes utilisées pour enregistrer et gérer les données d'enquête recueillies par enregistrement audio en se servant d'ordinateurs portatifs, de téléphones et de connexions Internet. Le processus décrit comprend un appareil raccordé directement au récepteur téléphonique qui diffuse les conversations en continu directement vers l'ordinateur portatif pour la sauvegarde et la transmission.

Mots clés : Collecte de données audio, transmission audio.

1. Introduction

1.1. Introduction

La collecte des données peut prendre de multiples formes, allant de la totalisation papier et crayon à l'interview assistée par ordinateur (IAO). Nombre d'entre nous ont été exposés à ces méthodes de collecte des données et d'autres; cependant, l'intégration d'enregistrements audio dans le processus est une forme de collecte des données qui a échappé aux spécialistes des sondages pendant de nombreuses années. Les raisons de ne pas recueillir les données de cette façon sont nombreuses, l'une d'entre elles étant la difficulté de transférer les données de l'endroit où elles sont recueillies à un emplacement centralisé où les fichiers peuvent être transcrits ou écoutés. La prévalence croissante de la connexion Internet à large bande (Horrigan, 2008) a rendu cette tâche plus facile, mais la gestion de gros fichiers audio demeure fastidieuse et leur transfert peut prendre beaucoup de temps.

1.2. Collecte des données

Au fil des ans, la collecte des données a pris de nombreuses formes et les méthodes varient de la simple collecte de données au moyen d'un questionnaire papier-crayon à l'interview Web assistée par ordinateur (IWAO). Chacune a ses avantages et ses inconvénients, si bien que les spécialistes des sondages doivent évaluer chaque type d'instrument de collecte des données afin de choisir celui qui convient à leurs besoins. Les interviews sont effectuées principalement sur place, par téléphone ou à l'aide d'un système d'interview sur place assistée par ordinateur (IPAO). Historiquement, l'interview sur place est celle qui produit les taux de réponse les plus élevés (De Vaus, 2002), mais il s'agit aussi de celle qui prend le plus de temps et qui est la plus coûteuse. Par ailleurs, l'interview téléphonique produit des taux de réponse beaucoup plus élevés que les enquêtes par la poste, mais un biais peut se produire puisque seul le segment de la population ayant un téléphone est interviewé. (Il convient aussi de souligner qu'un grand nombre de personnes sont abonnées uniquement à un service de téléphonie mobile pour lequel l'information pour les contacts peut être difficile à obtenir.) Les méthodes d'interview fondées sur l'IPAO et l'IWAO permettent de saisir les données directement dans un ordinateur, ce

¹ Charles Loftis, RTI International 3040 Cornwallis Rd, Durham NC 27709 USA; Shane Trahan RTI International 3040 Cornwallis Rd, Durham NC 27709 USA

qui raccourcit le temps de traitement et de validation des données, mais leur utilisation est limitée par le bagage informatique des répondants, la connectabilité à Internet, la formation des intervieweurs, l'acquisition du matériel et les coûts de lancement du développement des logiciels. Les questionnaires papier-crayon peuvent être envoyés par la poste à un grand nombre de personnes et les répondants fournissent sans doute des réponses plus honnêtes aux questions délicates (Muijs, 2004), mais nombre de personnes qui reçoivent ce genre de questionnaire n'y répondent pas. Il est indubitable que les méthodes de collecte de données au moyen d'appareils de poche deviendront rapidement plus fréquentes, à mesure qu'augmente le nombre de personnes familiarisées avec ce genre d'appareils, mais les utilisateurs de ces méthodes devront surmonter nombre d'obstacles semblables à ceux mentionnés plus haut.

1.3. Projet de la NSDUH

La National Survey on Drug Use and Health (NSDUH) est une enquête nationale réalisée annuellement par Research Triangle Institute (RTI) International pour le compte de la Substance Abuse and Mental Health Services Administration (SAMHSA). Dans le cadre de cette enquête, des données de présélection sont recueillies auprès d'environ 160 000 ménages et des données d'interview, auprès de 67 500 personnes réparties entre les 50 États et le district de Columbia. La NSDUH est, pour l'administration fédérale, la principale source de données nationales sur la consommation d'alcool, de tabac et de substances illicites par la population civile générale de 12 ans et plus non placée en établissement des États-Unis.

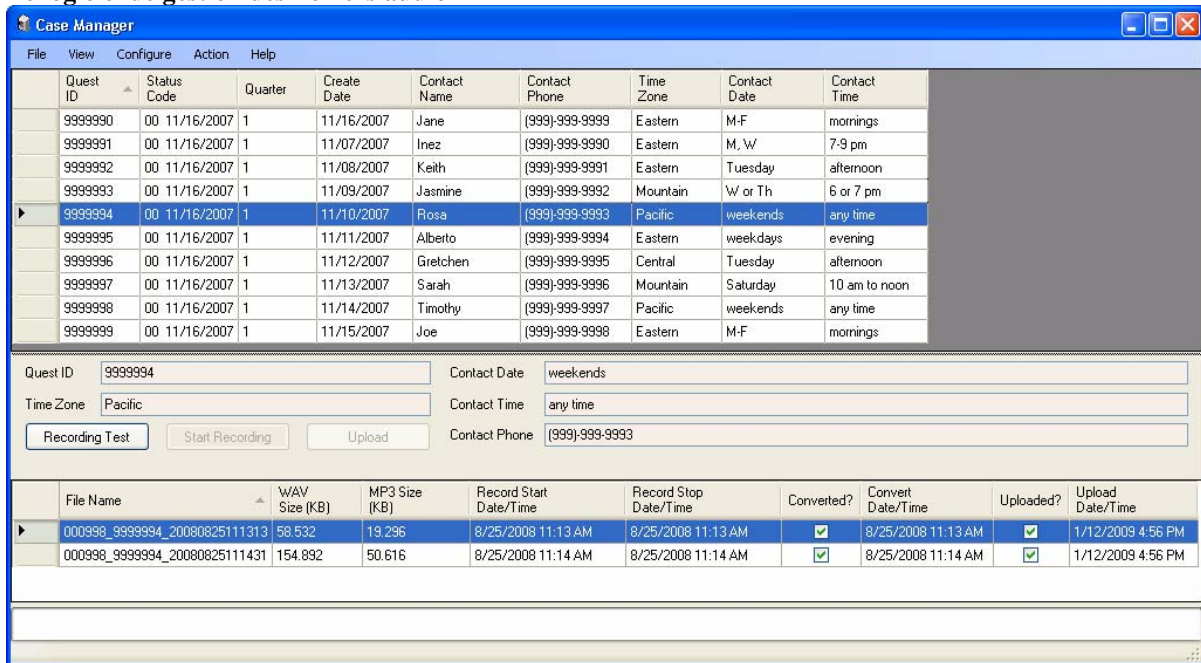
De temps à autre, la SAMHSA demande que l'équipe de la NSDUH procède à des améliorations, à des optimisations ou à des réductions des coûts de l'enquête. Pour l'enquête de 2008, la SAMHSA a demandé que soit exécutée une sous-étude téléphonique en vue d'estimer la prévalence de la maladie grave chez les adultes et d'évaluer la performance de modules de questions d'enquête en vue de les retenir comme instruments d'interview lors de futures vagues de la NSDUH. L'enquête téléphonique a requis des intervieweurs cliniques ayant reçu une formation spéciale pour enregistrer toutes les réponses au moyen d'un questionnaire papier-crayon tout en effectuant, moyennant l'obtention du consentement du répondant, un enregistrement audio de la conversation complète. Les fichiers audio ont été stockés dans un ordinateur portable fourni pour le projet, puis transmis à RTI qui les a utilisés pour des opérations de validation et de contrôle de la qualité. La suite de l'article décrit la solution développée par RTI pour enregistrer, gérer, transmettre et archiver les données audio en vue de leur utilisation dans la sous-étude sur la santé mentale de la NSDUH.

2. Méthodologie/approche technique

2.1. Modèle

L'objectif, afin d'effectuer la sous-étude, était de rassembler les fichiers audio à RTI pour que les employés affectés au projet puissent les examiner plus en détail. Pour cela, il fallait transmettre ces fichiers à un emplacement centralisé. Comme nous pensions que les fichiers audio risquaient d'être très gros, nous avons décidé qu'une approche similaire à celle employée pour transmettre des paquets de données par Internet serait un modèle idéal pour la conception de notre système de transmission. Ce dernier découperait les fichiers audio en paquets ou fragments plus faciles à gérer et les enverrait à un service Web qui regrouperait ces fragments de fichier afin de recréer un fichier unique. Une fois reçu, chaque fragment serait sauvegardé sur le serveur et une validation de la taille serait effectuée afin de confirmer le succès de la transmission. Tout fragment rejeté au contrôle de validation de la taille serait supprimé, et le service Web demanderait au client de l'envoyer de nouveau. Lorsque tous les fragments seraient reçus et validés, le service Web les fusionnerait pour reconstituer un fichier contenant une chaîne de texte en format base64 unique et convertirait celui-ci en un fichier binaire (audio) qui pourrait être examiné par les utilisateurs du site Web sécurisé.

Figure 2.1-1
Le logiciel de gestion des fichiers audio

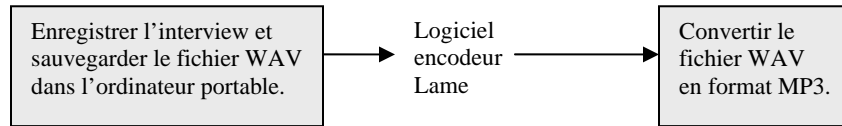


Afin de réussir à transmettre ces données au serveur centralisé, il fallait gérer non seulement les transmissions, mais aussi les interviews proprement dites. Le logiciel de gestion des fichiers audio (figure 2.1-1) que nous avons développé intègre à la fois la saisie et la conversion des fichiers audio, ainsi que les fonctions nécessaires pour communiquer efficacement avec les services Web. L'interface développée permet d'optimiser la gestion de l'acquisition des enregistrements audio tout en laissant à l'arrière-plan les aspects complexes de la transmission et de la conversion des fichiers. L'interface de gestion des cas permet aux intervieweurs d'examiner la situation des enregistrements et comprend une application de planification facile à utiliser pour gérer la charge de travail.

2.2. Acquisition des fichiers audio

Après avoir raccordé correctement le téléphone à l'entrée audio de l'ordinateur portable au moyen d'une clé électronique analogue ou numérique, l'intervieweur clinique commence l'interview et fait démarrer l'enregistrement audio. Chaque enregistrement est stocké sous forme de données brutes (également appelées format « WAV »). Les fichiers WAV fournissent la plus haute qualité possible d'enregistrement, mais ils nécessitent aussi un espace de stockage considérable. Un enregistrement de deux (2) minutes à la qualité de transmission téléphonique typique (mono, 4 à 8 MHz) consomme environ 1 Mo d'espace disque. Des utilitaires de conversion pouvant être obtenus gratuitement permettent la conversion sans perte des données audio WAV au format MP3. En utilisant un « encodeur Lame » téléchargeable gratuitement et en configurant différents taux d'échantillonnage (nombre de fois que les données audio sont stockées par seconde en Hz), canaux (mono, stéréo) et nombres de bits par échantillon (1, 2 octets, etc.), il est possible de contrôler la qualité de la conversion. Nous avons déterminé empiriquement les conditions de conversion au format MP3 comprimé qui reproduisent de la manière la plus approchante l'enregistrement audio original, tout en réduisant significativement la taille du fichier. Le taux de compression était considérable – près de 3 à 1 en taille pour la conversion WAV à MP3. Donc, après la saisie des données audio dans un format sans perte, notre première action consiste à convertir le fichier WAV en un fichier MP3 de taille beaucoup plus petite. La figure 2.2-1 donne un aperçu de ce processus.

Figure 2.2-1
Convertir le fichier WAV



2.3. Préparation du fichier audio du côté client

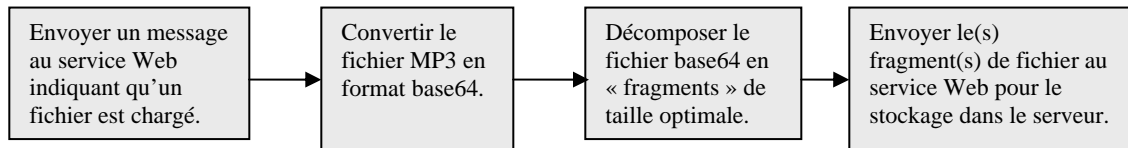
Une fois qu'un enregistrement est achevé et a été converti du format WAV au format MP3, il subit une autre conversion au format base64. Ce dernier est l'ensemble de tous les caractères majuscules et minuscules, de tous les chiffres de 0 à 9 et du trait d'union (« - »). La conversion du fichier MP3 binaire en un fichier texte en base64 facilite la transmission des données audio au moyen d'une connexion utilisant le protocole de transfert hypertexte (HTTP) standard sans interférence des intranet et (ou) des pare-feu.

Une fois qu'il est en format texte, le fichier est prêt à être transmis en vue de son examen et de son stockage. Le logiciel de gestion des fichiers audio avertit alors le service Web qu'un fichier audio est prêt à être téléchargé en amont. Le message est envoyé au service Web dans un format XML et indique la taille du fichier MP3 final, le nom du fichier et d'autres paramètres importants. Une fois que le service Web a reçu cette information, il envoie un message XML d'accusé de réception au logiciel de gestion des fichiers audio pour informer le client qu'il est prêt à recevoir les fragments de fichier. En fonction de la vitesse de connexion Internet, le logiciel de gestion des fichiers audio divise alors le fichier base64 en fragments de fichier de taille optimale et commence à les transmettre à un entrepôt central où ils seront fusionnés. Les fragments de fichier ne sont considérés comme transmis qu'après la réception d'un accusé de réception provenant du service Web de transmission. Le logiciel de gestion des fichiers audio envoie au service Web les paramètres qui suivent avec chaque fragment de fichier de données en base64 :

- ID* – Numéro d'identification unique de l'intervieweur
- Offset* – Décalage du fichier texte
- LengthVal* – Longueur du fragment
- ClientFileName* – Nom de fichier comprenant le fichier complet
- Base64Chunk* – Le fragment de texte en base64

La figure 2.3-1 donne un aperçu de ce processus.

Figure 2.3-1
Transmission du fichier

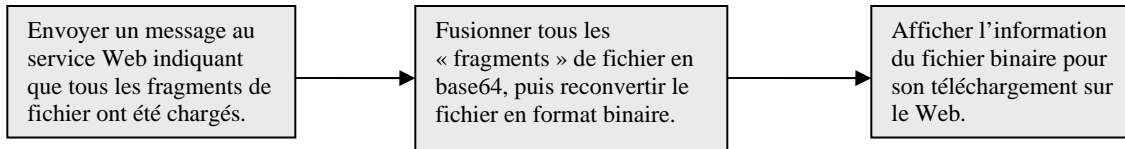


2.4. Manipulation du fichier audio du côté serveur

Tous les fragments vérifiés sont sauvegardés dans un répertoire et l'information correspondante sur les paramètres est mémorisée dans une base de données. Une fois que le dernier fragment de fichier est reçu, le logiciel de gestion des fichiers audio envoie au service Web un message indiquant que tous les fragments ont été reçus et vérifiés, et qu'il peut commencer à fusionner les fichiers texte et à reconverter le fichier résultant en format binaire. Le service Web recherche alors dans la base de données contenant tous les fragments de fichier reçus des fichiers particuliers par ordre décroissant du paramètre de décalage et construit un fichier texte en base64 complet conformément à l'information contenue dans la base de données. Ensuite, le service Web

convertit le fichier texte en format base64 en un fichier binaire et compare la taille de ce dernier à l'information envoyée au départ. Si le résultat de la vérification du fichier est bon, le nouveau fichier est affiché sur le site Web de gestion où il peut être examiné. La figure 2.4-1 donne un aperçu de ce processus.

Figure 2.4-1
Afficher l'information du fichier binaire



2.5. Examen des fichiers audio

Lorsqu'ils ont été téléchargés en amont dans le serveur central, les fichiers audio peuvent être examinés sur un site Web à accès sécurisé. Le système d'examen sur le Web permet aux utilisateurs de trier les fichiers audio en fonction de la date de téléchargement/réception (historique de téléchargement), du numéro d'identification du questionnaire (Questid) ou de l'intervieweur. Les fichiers audio sont en format MP3 et peuvent être écoutés au moyen de divers navigateurs et systèmes d'exploitation.

Figure 2.5-1
Examen des fichiers audio

Date Posted	Title/Download	Contact	File Size	Download History
30-Mar-08	Audio File: 676916_5788819_20080330151121.mp3	Interviewer 2	7,375 KB	
25-Mar-08	Audio File: 676916_2382889_20080325200127.mp3	Interviewer 2	6,421 KB	
24-Mar-08	Audio File: 676916_8162016_20080320175940.mp3	Interviewer 2	7,302 KB	Clinical034
24-Mar-08	Audio File: 676916_6661230_20080324193319.mp3	Interviewer 2	5,259 KB	Clinical034
15-Mar-08	Audio File: 676916_8239756_20080315101130.mp3	Interviewer 2	13,143 KB	
07-Mar-08	Audio File: 676916_2931634_20080307194204.mp3	Interviewer 2	7,894 KB	
04-Mar-08	Audio File: 676916_7491582_20080304163942.mp3	Interviewer 2	9,947 KB	Clinical033
03-Mar-08	Audio File: 676916_5516596_20080303145240.mp3	Interviewer 2	10,582 KB	Clinical033
27-Feb-08	Audio File: 676916_7971402_20080227194224.mp3	Interviewer 2	11,489 KB	
26-Feb-08	Audio File: 676916_2174863_20080226185206.mp3	Interviewer 2	12,847 KB	
24-Feb-08	Audio File: 676916_3481878_20080224081503.mp3	Interviewer 2	10,202 KB	Clinical029
19-Feb-08	Audio File: 676916_5161763_20080219191335.mp3	Interviewer 2	14,065 KB	Clinical034

Le site Web est en outre un site à diffusion restreinte qui n'autorise que l'accès des superviseurs ou des cliniciens, et ce, uniquement aux fichiers de l'entrepôt correspondant à leur région particulière. Par exemple, les superviseurs sur le terrain de la région un n'ont accès qu'aux fichiers audio de leurs intervieweurs, mais les cliniciens du projet peuvent avoir accès à tous les fichiers audio téléchargés en amont. En plus de la diffusion restreinte, l'examen d'un fichier audio par un clinicien ou par tout autre utilisateur du site est journalisé dans une base de données qui est utilisée pour dépister les enregistrements de nature éventuellement délicate. Le regroupement des fichiers dans un entrepôt central accessible au moyen d'une simple interface Web permet à toute personne ayant les autorisations adéquates d'examiner facilement les enregistrements audio.

3. Résultats

3.1. Résultats du système de saisie des enregistrements audio

Le logiciel et le système développés pour l'étude décrite ici ont été intégrés dans la sous-étude de la santé mentale de la NSDUH de 2008 et se sont avérés très précieux. Au cours de 2008, nous avons reçu et géré environ 6 604 fichiers (audio) d'interview répartis entre 451 répertoires contenant globalement plus de 12 Go de données. La taille moyenne des fichiers était de l'ordre de 7,3 Mo (ce qui équivaut à environ 90 minutes de données audio) avec un temps moyen de chargement d'environ 3,5 minutes à 34,06 Ko/s. Nous avons exigé que chaque ordinateur portable soit raccordé au moyen d'une connexion à large bande, mais notre méthode de transmission des fichiers audio convient également pour les utilisateurs de l'accès par réseau commuté. Néanmoins, ces derniers doivent s'attendre à être fortement désavantagés pour la transmission de gros fichiers.

Les cliniciens ainsi que les gestionnaires de l'enquête ont pu examiner et suivre les réponses à l'enquête sur fichier audio à partir de l'interface Web, et notre méthode de transmission des données et de gestion de ces fichiers s'est avérée plus qu'adéquate et souple pour la gestion d'une enquête fondée sur une réponse audio. La sous-étude sur la santé mentale de la NSDUH a donné de tellement bons résultats qu'elle a été poursuivie en 2009.

3.2. Éléments à prendre en considération

Certaines difficultés se sont posées durant la collecte des données audio, dont l'enregistrement d'interférences et certains problèmes de transmission. Notre équipe a réussi à surmonter les obstacles liés à l'enregistrement audio, mais elle s'est heurtée à des problèmes difficiles à résoudre à cause de contraintes sur lesquelles elle n'avait aucun contrôle.

Interférences dans les enregistrements audio

Durant l'enregistrement audio des interviews, nous avons découvert que l'utilisation de la clé électronique de liaison téléphonique pour enregistrer les conversations téléphoniques quand l'ordinateur portable était branché au réseau (en utilisant l'identificateur de circuit CA mis à la terre) causait un bourdonnement audible dans l'enregistrement. Dans de nombreux cas, l'interférence était suffisamment bruyante pour que l'enregistrement soit jugé inintelligible. Nous avons appris que le bruit statique était causé par un brouillage au sol arrivant par la fiche de mise à la terre de l'unité d'alimentation en courant alternatif. Par conséquent, pour éliminer l'interférence, nous n'avons utilisé que la pile de l'ordinateur portable comme source d'alimentation. L'inconvénient était que les interviews pouvaient être longues et que la durée de vie de la pile posait alors un problème durant l'interview. Par conséquent, on a appris aux intervieweurs à garder leur pile complètement chargée en tout temps.

Problèmes de transmission

En cas de problèmes durant la transmission ou la conversion du fichier Web en format MP3, il fallait trouver une autre méthode d'obtention des enregistrements. Nous avons donc mis en place un processus permettant aux employés du bureau central de marquer un fichier particulier sauvegardé dans l'ordinateur portable afin qu'il soit transmis au serveur central même s'il n'avait pas été converti au format MP3 (comme dans le cas où un fichier audio était corrompu à cause d'une coupure soudaine d'électricité). Le fichier était encore converti en format base64 et envoyé en fragments à RTI comme n'importe quel autre enregistrement audio. Tout fichier marqué de cette façon par le personnel de RTI afin qu'il soit poussé par le client vers le dépôt centralisé de fichiers ou qu'il soit retransmis a fait l'objet d'une tentative de chargement à chaque connexion, jusqu'à ce que sa réception ait été confirmée.

Notre approche technique a permis d'éliminer nombre de problèmes que posent habituellement les enquêtes par enregistrement audio. Notre méthodologie réduit considérablement la taille du fichier nécessaire pour

mémoriser l'interview et notre système de transmission s'est avéré à la fois souple et fiable pour ce qui est d'envoyer l'information au dépôt central de fichiers.

4. Conclusion

4.1. Conclusion

Dans l'ensemble, l'approche décrite dans le présent article en vue d'enregistrer, de gérer et de transmettre des fichiers audio pour la recherche par sondage a permis de surmonter les obstacles initiaux rencontrés, dont l'entreposage et la gestion de fichiers pouvant être très gros. Selon nous, l'utilisation de la technologie pourrait évoluer dans l'avenir afin de mieux saisir les enregistrements audio. Cela pourrait inclure la production automatique de transcriptions et l'assignation de descripteurs aux mots-clés pour chaque interview. L'élimination de la tâche, à haute intensité de main-d'œuvre, de transcription des interviews permettrait aussi aux cliniciens et aux méthodologistes d'enquête de mieux cibler certaines interviews pour une étude plus approfondie. L'avantage le plus important de l'interview avec enregistrement audio est la capacité de poser des questions d'approfondissement, mais celles-ci requièrent que l'intervieweur établisse une relation avec le répondant, ce qui peut inclure un bavardage sans importance durant l'interview; la transcription automatique peut aider à filtrer ces bavardages. La transcription automatique des fichiers audio en texte donnerait aussi la possibilité de traduire les données textuelles en d'autres langues. Il existe à l'heure actuelle de nombreux services qui traduisent des textes d'une langue à une autre. Ces services sont relativement coûteux, mais leur prix baisse, si bien qu'ils deviendront probablement beaucoup plus fréquents dans les années à venir. La traduction des résultats des interviews audio en plusieurs langues permettra non seulement d'élargir l'auditoire des travaux de recherche, mais contribuera aussi à atténuer les obstacles linguistiques rencontrés durant l'interview de diverses populations. Nous nous attendons également à ce qu'il soit possible d'apporter des modifications mineures au logiciel de gestion des cas en vue d'accroître son utilité sur le terrain. L'un des changements qui pourraient être envisagés est la « transmission en continu » automatique des fichiers de données au serveur central pendant les temps morts d'utilisation des ordinateurs et du réseau.

Bibliographie

- Horrigan, J.B. (2008). Home Broadband Adoption 2008. Pew Internet & American Life Project. <http://www.pewinternet.org/pdfs/PIP_Broadband_2008.pdf>.
- De Vaus, D. (2002). *Surveys in Social Research*. 5^e édition. New York: Routledge, 127
- Muijs, D. (2004). *Doing Quantitative Research in Education: with SPSS*. 1^{ère} édition. London: Sage Publications, 41.