



**Conseil économique
et social**

Distr.
GÉNÉRALE

CES/2005/8
23 mars 2005

FRANÇAIS
Original: ANGLAIS

COMMISSION DE STATISTIQUE et
COMMISSION ÉCONOMIQUE POUR L'EUROPE

CONFÉRENCE DES STATISTICIENS EUROPÉENS

Cinquante-troisième réunion plénière
(Genève, 13-15 juin 2005)

**UTILISATION DU WEB DANS LA COLLECTE DE DONNÉES
AUX FINS DE L'ÉTABLISSEMENT DES STATISTIQUES
DES ENTREPRISES EN FINLANDE**

Communication sollicitée émanant de Statistics Finland*

INTRODUCTION

1. Les principales sources des données de base utilisées pour l'établissement des statistiques en Finlande sont les systèmes administratifs. La loi finlandaise sur la statistique stipule qu'il faut utiliser les données existantes pour établir les statistiques plutôt que de procéder à une collecte directe de données. Sans vouloir entrer dans les détails, on a calculé que près de 94 % des données brutes¹ rassemblées par Statistics Finland (SF) proviennent des systèmes administratifs. Dans le reste des données brutes, on compte environ 2,5 % de données recueillies en interrogeant des personnes et à peu près 3,5 % de données obtenues directement auprès d'institutions publiques et d'entreprises elles-mêmes. Sur ces 3,5 %, près de 1,8 % de données sont collectées au moyen de formulaires électroniques (Excel, ASCII et fichiers séquentiels) et 1,2 % en utilisant des formulaires Internet. En 2005, la part des données recueillies au moyen de formulaires sur papier atteindra environ 0,5 %.

* Document établi par Ilkka Hyppönen.

¹ Mesurées en «unités statistiques» multipliées par le «nombre de variables».

2. Il convient de souligner que les données recueillies directement auprès de personnes, d'institutions et d'entreprises ont une valeur bien supérieure à la proportion des données brutes qu'elles représentent. Dans de nombreux cas, les données de source administrative et les données d'enquête sont associées au niveau de l'unité statistique afin d'établir les statistiques requises. La collecte directe est nécessaire pour plusieurs raisons, principalement pour obtenir des données qui ne sont pas disponibles dans les systèmes administratifs, pour disposer plus rapidement de données pour des unités statistiques particulières et pour recueillir directement des informations auprès de grandes unités statistiques de manière à mieux contrôler la qualité des données.

3. Les deux catégories de données sont associées, par exemple dans le cadre des recensements de la population ou aux fins de l'établissement des statistiques sur la répartition des revenus ou des statistiques sur la structure des entreprises. Dans le cas des statistiques des entreprises pour lesquelles certaines données brutes peuvent être tirées de sources administratives, on se contente généralement de procéder à une collecte directe des données uniquement auprès des «entreprises les plus grandes», expression qui désigne les entreprises «employant 20 personnes ou davantage», parce que l'on a besoin de données ventilées en fonction d'une classification détaillée que l'on peut se procurer auprès de ces entreprises au demeurant très petites selon les normes internationales.

4. Nous nous devons également de signaler qu'il n'existe pas de données administratives de remplacement pour la plupart des enquêtes. Cette remarque vaut en particulier dans le cas des administrations locales, pour lesquelles de nombreuses statistiques sur l'éducation et les finances sont établies à Statistics Finland. Il s'agit essentiellement d'informations de nature administrative, et elles ont pour la plupart été transférées des organismes administratifs vers Statistics Finland au cours des 15 dernières années dans le but de rationaliser le processus de collecte de données de l'administration centrale. Malgré ces considérations, ces statistiques sont reconnues comme faisant partie intégrante des «statistiques officielles» et sont prises en considération dans le présent document.

5. On compte deux catégories d'entités économiques auprès desquelles on procède régulièrement à une collecte de données: les entreprises et les institutions administratives locales. Les données recueillies auprès de ces deux groupes représentent à peu près 92 % des données collectées directement par Statistics Finland (abstraction faite des données recueillies auprès des ménages). Ces deux catégories sont examinées plus en détail ci-après.

CYBERADMINISTRATION

6. D'après de nombreuses comparaisons internationales, la société finlandaise est considérée comme assez avancée pour ce qui est de l'utilisation d'Internet. Le raccordement à l'Internet est chose courante en Finlande et plus de 95 % des entreprises et 100 % des administrations locales ont accès à l'Internet (chiffres de 2003). Depuis plusieurs années, le Gouvernement finlandais met en œuvre un programme d'informatisation de l'administration (cyberadministration) et de promotion de la société de l'information, qui a créé un climat favorable à l'utilisation de l'Internet pour les communications avec l'administration.

7. Comme indiqué précédemment, Statistics Finland a pour principe d'exploiter, lorsque c'est possible, les données existantes, qu'elles soient de source publique ou privée. Parfois, Statistics Finland va même jusqu'à acheter des données au secteur privé. En outre, depuis les années 70,

Statistics Finland s'est attaché à promouvoir la collecte de données par voie électronique. La part des données recueillies directement sous la forme de fichiers atteste du résultat de ces efforts. La première application utilisant des formulaires Web a été lancée en 1997, mais ce n'est qu'en 2001 que les progrès en la matière se sont accélérés. À cette époque, Statistics Finland a fixé une date butoir à laquelle chaque formulaire imprimé de collecte de données devrait avoir un équivalent électronique. Un grand projet baptisé «modèle de production» est en cours d'exécution depuis deux ans. L'un des sous-projets qui le composent concerne le processus de collecte de données, dont l'un des buts essentiels est la mise en œuvre de l'objectif stratégique de Statistics Finland concernant la collecte de données. Cette initiative a sensiblement accéléré les progrès enregistrés durant l'année écoulée.

ENQUÊTES AUPRÈS DES ENTREPRISES ET DES ADMINISTRATIONS LOCALES

8. On compte environ 60 enquêtes différentes réalisées auprès des entreprises. Sur ce total, 15 intéressent les marchés financiers et sont mises au point conjointement avec l'autorité de contrôle financier et en collaboration avec des banques et d'autres institutions financières. Elles sont fondées, actuellement sur des tableurs Excel et leur contenu est très détaillé. Les enquêtes en question ne seront pas examinées plus en détail dans le présent document. Pour de nombreuses autres enquêtes de conjoncture, la possibilité est également donnée depuis longtemps de communiquer les données par voie électronique, en plus de la formule traditionnelle de présentation des données sur papier.

9. La plupart des enquêtes de conjoncture réalisées en Finlande sont de faible ampleur, en termes du nombre d'enquêtés et du nombre d'éléments de données considérés. Plus de la moitié des enquêtes de conjoncture sont des enquêtes annuelles et les autres des enquêtes mensuelles ou trimestrielles. Les enquêtes mensuelles et trimestrielles sont, pour la plupart, limitées à quelques questions/variables par enquêté, tandis que les enquêtes annuelles peuvent comprendre des centaines d'éléments de données.

10. On dénombre une douzaine d'enquêtes réalisées auprès des autorités administratives locales, pour lesquelles on utilise des formulaires sur papier et/ou des formulaires Web. Ces enquêtes ont notamment pour objet d'établir des statistiques sur les finances des administrations locales, sur l'éducation et sur le personnel des administrations en question. Il s'agit de collectes annuelles, sauf dans deux cas où l'enquête est respectivement trimestrielle et semestrielle. En général, ces enquêtes sont d'assez grande ampleur, par le nombre de variables considérées ou le nombre des effectifs ou des étudiants (élèves) pris en compte.

DEUX TYPES DE SOLUTION

11. Deux types de solution sont envisageables aux fins de la communication des données via l'Internet: l'une consiste à faire appel à un prestataire de services extérieur et l'autre à mettre au point des applications «maison». On compte une cinquantaine d'applications qui sont déjà opérationnelles, ou en cours de mise au point, ou encore à l'étude. Vingt et une de ces applications impliquent ou impliqueront le recours aux compétences d'un prestataire de services extérieur et 20 sont ou seront mises au point en interne. Pour les applications restantes, aucune décision n'a encore été prise.

Nombre de collectes de données via le Web, classées selon l'état d'avancement et la nature de la solution choisie, au début de 2005

	Solution interne	Solution extérieure	À déterminer
Application opérationnelle	11	14	
Application en cours de mise au point	3	3	
Application à l'étude	6	4	11
Total	20	21	11

12. Le prestataire de services extérieur fait office d'intermédiaire entre les enquêtés et Statistics Finland. Il met au point les formulaires et s'occupe des questions administratives intéressant les enquêtés, notamment des identificateurs et des mots de passe des utilisateurs.

13. La première mise au point d'une application interne a démarré en 2000 et a abouti à l'élaboration d'une application modèle, qui a été perfectionnée et dont l'utilisabilité est devenue plus générale depuis lors. L'application «maison» employée aujourd'hui, s'appelle XCola (application de collecte de données fondée sur XML). Dans XCola, les questionnaires sont définis en tant que document XML, et sont ensuite adaptés à chaque enquêté et transformés en pages Web au moment de l'exécution. On trouvera à l'annexe 1 des renseignements plus détaillés sur XCola.

SÉCURITÉ DES DONNÉES

14. Le prestataire de services extérieur se charge de toutes les questions touchant la sécurité des données. Les données sont envoyées à Statistics Finland sous la forme de fichiers chiffrés.

15. Dans la solution «maison», les composants/modalités techniques ci-après contribuent à assurer la sécurité:

- Toutes les informations circulant sur l'Internet sont chiffrées en SSL (clef de cryptage à 128 octets);
- Le moteur XCola et la base de données de collecte résident sur deux segments distincts du réseau local (LAN), qui sont protégés par des pare-feu à la fois vis-à-vis de l'Internet et vis-à-vis du reste du réseau local de Statistics Finland;
- Il n'est pas possible d'établir une connexion à Internet à partir de ces segments; les serveurs sur ces segments ne peuvent que recevoir/accepter des connexions;
- Le segment où le serveur de base de données est situé ne peut recevoir que des communications émanant du segment où se trouve le serveur XCola;
- Les identificateurs et les mots de passe des utilisateurs sont conservés dans un serveur distinct de services de répertoire;

- L'utilisateur est d'abord authentifié, dans un premier temps, puis autorisé à accéder au système, dans un second temps;
- Les données recueillies sont transférées séparément de la base de données de collecte à la base de données de production selon les besoins, généralement une ou deux fois par jour.

16. Avant la mise en service de la première application de collecte interne élaborée à Statistics Finland, un audit de sécurité des données transmises via Internet a été entrepris par un consultant extérieur. Certaines modifications ont été apportées aux modalités internes et à la solution technique mise au point. Ce niveau de protection est jugé suffisant par Statistics Finland et par les enquêtés qui se servent de la formule de réponse via le Web. Aucun enquêté n'a rencontré des problèmes ou manifesté de la méfiance quant à la sécurité des données assurée dans le cadre des collectes via le Web de Statistics Finland. Il y a lieu de signaler, toutefois, que Statistics Finland ne connaît pas vraiment les raisons pour lesquelles certains enquêtés n'utilisent pas la solution Web. Il se peut que l'incertitude à l'égard de la sécurité des données sur le Web soit l'une des causes de leur réticence.

17. La sécurité des données repose sur des procédures et règles types ainsi que sur des solutions techniques. On peut citer les exemples suivants:

- Des nouveaux identificateurs et mots de passe sont attribués chaque année aux utilisateurs (les anciens sont éliminés);
- Les identificateurs et mots de passe sont initialement expédiés par courrier. Si l'utilisateur le demande, ils peuvent être réexpédiés. Seul l'identificateur de connexion ou le mot de passe peut être expédié par courrier électronique, l'autre élément devant obligatoirement être envoyé par courrier postal ou communiqué par téléphone. Dans ce dernier cas, c'est un employé de Statistics Finland qui appelle l'enquêté et non pas l'inverse;
- Seul un nombre déterminé de nos effectifs a le droit de traiter les informations relatives aux identificateurs/mots de passe des utilisateurs (soit, généralement, deux personnes par enquête pour pallier une éventuelle indisponibilité momentanée de l'une ou de l'autre).

SITUATION AU DÉBUT DE 2005

18. Parmi les sept enquêtes de conjoncture mensuelles et l'enquête de conjoncture trimestrielle proposant une option déjà opérationnelle de réponse via le Web, on en compte quatre pour lesquelles cette option est en service depuis plus d'un an. La proportion de répondants utilisant l'option Internet est de 52 % pour l'une de ces quatre enquêtes et dépasse 70 % pour les trois autres.

19. Sur les six enquêtes de conjoncture annuelles, une seule a proposé l'option de réponse via le Web dès l'année dernière, tandis que pour les autres, cette solution est introduite pour la première fois cette année. Pour l'enquête au titre de laquelle l'option Web existait déjà l'année précédente, 15 % des répondants ont utilisé cette option, bien qu'elle n'ait été proposée qu'après l'envoi initial des formulaires sur papier. Dans l'une des enquêtes pour lesquelles l'option Web vient d'être introduite, on prévoit que la part des réponses via le Web atteindra jusqu'à 30 % et cette estimation devrait aussi se vérifier pour les autres enquêtes.

20. En ce qui concerne les 11 enquêtes auprès des administrations locales comportant une option de réponse via le Web, l'une enregistre un taux de réponse par Internet de 80 %, et les autres un taux de plus de 95 %. Ces chiffres sont assez surprenants étant donné qu'il s'agit d'enquêtes annuelles et qu'un répondant ne peut pas apprendre l'interface utilisateur par cœur mais doit répéter le processus d'apprentissage chaque année.

DONNÉES D'EXPÉRIENCE

Facteurs de réussite

21. Comme on l'a fait observer ci-dessus, les collectes via le Web ont rencontré un grand succès dans le cadre des enquêtes auprès des administrations locales et des enquêtes de conjoncture à court terme. Pour les enquêtes de conjoncture annuelles, on ne dispose encore que d'informations partielles en raison du long intervalle entre deux enquêtes de cette nature et du fait que le moment du démarrage pour ce genre d'enquête est plus tardif que pour les deux autres groupes.

22. Les facteurs ci-après semblent contribuer au succès de la collecte via le Web dans le cas des enquêtes de conjoncture à court terme:

- Les formulaires utilisés pour les enquêtes à court terme sont simples et peuvent être remplis aisément et rapidement, en une seule fois. Dans la plupart des cas, le temps nécessaire pour répondre à une enquête a été sensiblement réduit, ce qui allège la charge de travail des répondants;
- Les répondants appartiennent au personnel de grandes sociétés. Les personnes chargées de remplir les formulaires possèdent les connaissances techniques requises pour se servir de l'Internet. La solution SSL assurant la sécurité des données est elle aussi courante et n'exige pas de compétences techniques supplémentaires;
- Les enquêtes à court terme sont très répétitives et tellement fréquentes qu'il est possible d'apprendre par cœur la procédure connexe;
- La communication de statistiques en retour a suscité des réactions positives de la part des répondants et même des demandes d'informations complémentaires concernant ces statistiques.

23. Pour ce qui est des administrations locales, les facteurs expliquant le succès de l'utilisation du Web dans le cadre des enquêtes qui les concernent sont notamment les suivants:

- Les statistiques fournies en retour: les municipalités obtiennent leurs propres statistiques à partir des données que Statistics Finland recueille auprès des établissements relevant de leur autorité (les différents établissements d'enseignement, par exemple);
- Une partie des collectes de données influe directement sur le niveau des subventions que les administrations municipales reçoivent du gouvernement national. Par conséquent, ces collectes de données sont extrêmement importantes à leurs yeux;

- Les répondants sont enthousiasmés de pouvoir utiliser l'Internet – d'aucuns ont même trouvé cela AMUSANT de remplir les formulaires Web (par opposition au formulaire sur papier);
- S'agissant des statistiques financières des administrations locales, les données des enquêtes antérieures sont préalablement insérées dans le formulaire, ce qui aide le répondant à se souvenir de la manière dont les chiffres ont été calculés l'année précédente. Les contrôles de validité de type logique intégrés dans l'application sont utiles au répondant de même que l'aide contextuelle disponible.

Coûts de mise au point

24. Le coût de la mise au point et de l'exploitation des applications sur le Web a diminué de 60 à 70 % au cours des trois dernières années, tant pour les applications «maison» que pour celles élaborées par un prestataire de services extérieur.

25. De nos jours, les frais d'investissement moyens correspondant aux sommes payées à un prestataire de services par application de collecte sont de l'ordre de 5 000 euros, ce qui équivaut à peu près au coût d'un mois et demi de salaire d'un employé de bureau. Les coûts d'exploitation par an, à partir de la deuxième année, ne représentent que 1/5^e de cette somme. Il ne faut pas oublier que généralement, le nombre de répondants se chiffre en centaines, rarement en milliers, et jamais en dizaines de mille jusqu'ici. Bien entendu, une part du travail incombe au personnel de Statistics Finland, tant au stade de la mise au point que durant la phase d'exploitation, même s'il est essentiellement réalisé par le prestataire de services.

26. On en est déjà à la troisième génération de solutions mises au point en interne. La première application autonome a été élaborée en 2000. En 2002-2003, plusieurs applications nouvelles ont été créées sur la base de la précédente. La troisième génération correspond à la mise au point de l'application XCola et à son utilisation dans le cadre des programmes de collecte de données. Au cours des première et deuxième phases de mise au point, l'investissement total en ressources a représenté environ 2,5 années-personnel (créateur d'applications informatiques, statisticien de rang élevé), ce qui englobait la mise en place d'un environnement de communication sûr. L'élaboration de XCola a demandé environ une année-personnel, y compris la mise en place de quatre applications de collecte. Selon les estimations, les opérations de mise en place supplémentaires exigent à peu près 150 heures de travail pour chaque application.

Avantages

27. Des économies peuvent être réalisées sur les frais d'équipement et d'expédition de courrier et sur le plan du traitement manuel des formulaires et des données. En ce qui concerne les quatre séries de statistiques mensuelles/trimestrielles² pour lesquelles la collecte de données via le Web est opérationnelle depuis plus d'une année, la proportion moyenne du travail économisé au stade

² Enquête mensuelle sur la production industrielle, enquête mensuelle sur le chiffre d'affaires, enquête mensuelle sur les coûts de construction, enquête trimestrielle sur les stocks dans l'industrie et le commerce. On dénombre, au total, pour ces enquêtes à peu près 3 300 répondants par mois et 800 par trimestre.

de la collecte des données est de plus de 40 %, soit, pour nous, une économie totale de deux années-personnel. La quantité de courrier expédiée par la poste dans le cadre de ces quatre enquêtes a été réduite de 64 000 unités par an (65 %), ce qui correspond au coût d'environ une demi-année de travail d'un employé de bureau. Au total, les économies réalisées représentent environ 2,5 années-personnel.

28. Le temps de réponse moyen aux enquêtes mensuelles a été réduit dans le meilleur des cas de sept à huit jours, soit de 30 %. Ce gain de temps se répercute également sur la phase d'exploitation suivante. Pour les statistiques à court terme, le temps gagné en termes de distribution du courrier est particulièrement important en ce qui concerne les notes de rappel. Dans le meilleur des cas, le nombre de notes de rappel a diminué de moitié. Pour ce qui est des enquêtes de conjoncture annuelles, rien n'indique clairement jusqu'ici que les enquêtés répondent plus rapidement. L'annexe 2 fournit des renseignements détaillés concernant l'une des enquêtes mensuelles.

29. Les données reçues sont de meilleure qualité. De façon générale, on évalue la diminution des erreurs à 25 %, même si aucune véritable étude n'a été entreprise pour étayer cette estimation. Ce gain de qualité vaut à la fois pour les enquêtes mensuelles et les enquêtes annuelles.

30. S'agissant des statistiques sur les finances publiques locales, on estime avoir pu réduire de 25 % de travail de correction parce que les données reçues comportent moins d'erreurs. Selon l'opinion spécialisée de la personne responsable, la qualité des données s'est également améliorée à d'autres égards, tant d'un point de vue technique que sur le plan de leur validité. On ne relève pas d'amélioration sensible des délais de réponse.

31. Pour les statistiques sur l'éducation, il est très difficile de comparer l'ancien et le nouveau mode de collecte des données, parce que le contenu de l'enquête se modifie lorsque l'on passe à l'option Web. Dans un cas en particulier, la personne responsable a déclaré qu'un élargissement du contenu des statistiques aurait été impossible sans utiliser dans un même temps un formulaire Web.

32. À mesure que le volume des tâches manuelles diminue, celles-ci peuvent être remplacées par l'exécution de fonctions plus valorisantes au sein du service de statistique.

33. Les avantages que retirent les enquêtés du changement sont notamment les suivants:

- Ils sont en mesure de remplir le questionnaire rapidement;
- Dans certains cas, l'insertion préalable de certaines données dans le questionnaire peut les aider à se rappeler de leurs réponses antérieures;
- Des contrôles de validité empêchent l'envoi de données erronées et éliminent la nécessité de demander des informations supplémentaires pour rectifier des erreurs;
- Dans les formulaires complexes où le même élément d'information apparaît à plusieurs reprises, il suffit à présent de le saisir une seule fois;

- Dans un premier temps du moins, les intéressés aiment se servir de l'Internet pour remplir le questionnaire parce que c'est un moyen moderne d'effectuer cette tâche;
- Bon nombre d'enquêtés ont déclaré que leur travail s'était allégé.

Aperçu récapitulatif des coûts-avantages

34. En ce qui concerne la mise au point d'applications internes, il semble que l'investissement dans la collecte via le Web, mesuré purement et simplement en euros, se soit avéré rentable au bout d'un an environ en ce qui concerne les statistiques à court terme. Les progrès en termes de qualité des données, la réduction du temps de collecte, l'amélioration de la qualité du travail du personnel de bureau et l'allègement de la charge de travail des enquêtés, telle que les répondants la perçoivent, représentent des avantages supplémentaires.

35. Dans l'ensemble, on constate que l'investissement dans la collecte de données via le Web a été rentabilisé en moins de deux ans dans tous les domaines d'application. La nécessité de conserver le mode de collecte sur papier dans un avenir proche est le seul facteur restreignant les avantages potentiels.

CONCEPTION DU SYSTÈME ET AUTRES CONSIDÉRATIONS

36. L'application XCola mise au point en interne selon le principe «une application, une base de données» offre l'occasion d'utiliser un seul système aux fins de l'exécution de tâches pour lesquelles chaque domaine de la statistique avait auparavant sa propre solution. C'est un pas dans la bonne direction, parce qu'il permet de restreindre les coûts de maintenance et d'exploitation.

37. On ne note aucun changement fondamental dans la répartition des tâches entre les différents services de Statistics Finland. En principe, une équipe de statisticiens et d'employés de bureau est chargée de toutes les phases de l'élaboration d'un ou plusieurs produits statistiques, avec l'appui d'experts en informatique et de spécialistes de la publication.

38. Dans le cadre de son expérience, Statistics Finland n'a rencontré aucun écueil particulier. L'usage de bonnes méthodes de mise au point d'applications et le soin apporté à la conception du dispositif de sécurité se sont avérés être des moyens d'action suffisants. On trouvera ci-après certaines considérations relatives à la conception du système, parce qu'elles mettent l'accent sur les aspects où Statistics Finland a commis des erreurs ou aurait pu mieux faire dans le passé:

- Pour concevoir un système de collecte des données via le Web, il ne suffit pas de reproduire simplement un formulaire sur papier sous forme électronique. En raison des limites imposées par les écrans d'affichage, la présentation du formulaire et le flux des questions doivent être repensés. En outre, il convient de tirer parti des possibilités nouvelles offertes par l'option Web comme la validation en ligne, l'assistance contextuelle, l'insertion préalable de données antérieures dans les questionnaires et le retour d'information aux répondants. L'expérience des employés (de bureau) qui étaient chargés de dépouiller les formulaires devrait toujours être mise à profit;

- La validation en ligne devrait être aussi approfondie que possible. Toutefois, les contrôles devraient être conçus de telle manière qu'ils n'entraient pas inutilement le processus de réponse;
- Pour tester une application de collecte des données, il est utile de l'expérimenter auprès d'un petit groupe de répondants (en plus des essais réalisés au sein du service de statistique même). La durée appropriée d'une période d'essai est de deux à trois mois dans le cas des enquêtes mensuelles, et d'un trimestre dans le cas des enquêtes trimestrielles. Pour les enquêtes moins fréquentes, il n'est pas réalisable de tester la méthode auprès des répondants;
- Les personnes chargées de gérer les identificateurs de connexion et les mots de passe attribués aux utilisateurs doivent recevoir une formation appropriée. Il faut prêter particulièrement attention à l'organisation de la gestion des identificateurs parce que, chaque mois, entre 1 et 3 % des enquêtés égarent leur identificateur et/ou leur mot de passe;
- L'application de collecte doit permettre d'interrompre le processus de réponse et de le reprendre ultérieurement. Lors de la maintenance de l'application de collecte, qui est alors fermée pendant la durée de cette opération, il faut veiller à ce qu'aucun enquêté ne perde les données déjà entrées;
- Lors de la distribution collective de messages aux enquêtés, il faudrait veiller à ce que ces envois n'encombrent pas le serveur du courrier électronique. Les expéditeurs devront s'assurer que tous les messages ont bien été envoyés, par exemple en s'adressant à eux-mêmes le dernier message;
- S'il est fait appel à un agent extérieur, on doit mettre en place toutes les procédures requises en vue de l'instauration d'une bonne relation prestataire-client, notamment des règles régissant la communication et la coopération, ainsi que la réception et l'accusé de réception des tâches de maintenance.

CONCLUSIONS

39. Comme indiqué dans les paragraphes ci-dessus, il est possible, en faisant appel à un prestataire de services ou en élaborant une solution «maison», de mettre en œuvre des applications économiquement avantageuses de collecte de données, fondées sur l'utilisation de l'Internet. Il semblerait que les coûts de mise au point des deux types de solutions (externe et interne) soient à peu près identiques et que l'investissement correspondant puisse être rentabilisé assez rapidement, c'est-à-dire au bout d'un ou deux ans. Pour les enquêtes très fréquentes, c'est-à-dire mensuelles ou trimestrielles, une solution interne est un meilleur choix, parce qu'elle garantit l'autonomie dans un domaine où le facteur temps joue un grand rôle dans les applications. C'est pour ces mêmes enquêtes que l'option de réponse via l'Internet semble se révéler la plus profitable.

40. A Statistics Finland, chacun s'accorde à dire que la qualité des données reçues via l'Internet est meilleure que celle des données communiquées sur papier. Cette constatation vaut pour toutes les catégories d'enquêtes. Il semble aussi que les enquêtés soient favorables aux

nouvelles technologies. On se doit de mettre à profit cet état d'esprit dans le cadre des efforts qu'il faut constamment déployer pour stimuler la bonne volonté des répondants.

41. Il est possible de satisfaire aux exigences en matière de confidentialité et de sécurité des données sans devoir engager des dépenses trop importantes. Le niveau de sécurité offert par nos solutions de réponse via l'Internet est le même que pour n'importe quel service Internet avancé qui implique l'usage de monnaie électronique, et la sécurité interne est assurée par une planification minutieuse, des mesures de bon sens et quelques conseils d'un consultant spécialisé dans la sécurité des télécommunications. Statistics Finland n'a pas eu besoin d'infrastructures spéciales pour son système de réponse via le Web, en plus de celles qui devaient être mises en place de toute manière pour protéger son environnement informatique contre les aspects néfastes de l'Internet. En fait, la mise au point d'applications de collecte des données fondées sur le Web a permis à Statistics Finland de mieux comprendre la technologie de la sécurité de l'Internet.

42. Une stratégie clairement définie s'impose pour obtenir de bons résultats, un objectif ne pouvant être atteint que s'il a été bien défini.

ANNEXE 1

Structure de l'application XCola, ainsi que des bases de données et segments correspondants

1. XCola est un moteur d'application pour les enquêtes fondées sur l'Internet. Dans XCola, les questionnaires sont définis en tant que documents XML, puis adaptés à chaque enquêté et transformés en pages Web au moment de l'exécution. Tout ceci se fait selon un processus très générique applicable à la plupart des enquêtes sans exiger, par exemple, une programmation supplémentaire. L'application est élaborée en utilisant visual.net (Microsoft) et le serveur MS SQL.
2. La base de données de collecte, où sont stockées les données entrées par l'enquêté, se trouve à Statistics Finland. Elle réside sur un serveur qui est raccordé à un segment distinct du réseau local (LAN) de Statistics Finland. Ce segment est séparé de l'Internet par un pare-feu et du reste du réseau local de Statistics Finland par un autre pare-feu. Le logiciel utilisé pour la base de données est le serveur MS SQL.
3. Les informations concernant les enquêtés, par exemple leurs nom et adresse, sont entreposées dans la même base de données du serveur MS SQL que la base de données de collecte.
4. Les identificateurs et mots de passe des utilisateurs sont conservés dans un serveur de services de répertoire résidant sur le même segment que la base de données de collecte.
5. L'entreposage des identificateurs et mots de passe des utilisateurs s'effectue sur un serveur de services de répertoire distinct. Au moment de la connexion, le répondant est tout d'abord authentifié (c'est-à-dire admis après avoir été identifié), puis autorisé à accéder au système d'enquête. Dans le cadre de cette procédure d'autorisation, XCola détermine le rôle qui revient au répondant dans le contexte de l'enquête et configure la session de communication en conséquence.
6. L'application XCola se trouve sur un serveur distinct lui-même placé sur un segment séparé. Les deux segments sont connectés à un pare-feu. Aucun appareil sur ces segments ne peut établir une connexion avec l'Internet, et le segment comprenant la base de données de collecte ne peut accepter que les messages émanant du segment où se trouve l'application XCola qui peut, quant à lui, accepter les messages provenant de l'extérieur.
7. Le format XML utilisé pour l'application XCola est une version améliorée de XML4DR, et reprend aussi quelques éléments de la définition de la norme W3C's XForms.
8. Sur le plan technique, l'interface utilisateur est produite en tant qu'application ASP.NET, qui est envoyée via l'Internet au poste de travail du répondant. Le document XML définissant le questionnaire ne comprend pas de code propre à l'interface utilisateur. Au lieu de cela, tous les aspects de l'interface utilisateur sont pris en charge par le moteur XCola et ASP.NET.
9. Pour créer l'interface utilisateur, XCola se sert de renseignements tirés du formulaire XML, de la base de données de collecte et de la base de données des répondants. À partir des

données de cette dernière, on peut déduire notamment quelles lignes doivent être insérées dans le formulaire (par exemple déterminer sur quels produits le répondant fournit des données à Statistics Finland). Outre qu'elle permet de créer l'interface utilisateur (le formulaire et son format), l'application XCola peut aussi préremplir les questionnaires en y insérant des informations spécifiques au répondant considéré, par exemple des données d'une ou de plusieurs périodes antérieures, à titre de référence et en vue de corrections éventuelles.

10. Le moteur XCola produit aussi une logique d'application permettant de vérifier les données entrées par l'utilisateur, il établit des listes (par exemple des classifications telles que listes/codes de produits ou de pays) et il génère des programmes d'assistance.

11. L'utilisateur a le choix entre trois langues: finnois, suédois et anglais.

12. Une fonctionnalité supplémentaire permet un retour d'information individualisé à chaque répondant. Par exemple, ce dernier peut avoir communication de ses propres chiffres pour la période précédente, mis en parallèle avec ceux du type d'activité dont il est en train de rendre compte.

13. Des notes de rappel sont envoyées par courrier électronique. Il existe une application dédiée au publipostage vers laquelle des demandes d'envoi de notes de rappel peuvent être expédiées à partir d'une application donnée (destinataires, le message). L'application de publipostage se charge ensuite de transmettre la note aux différents enquêtés en espaçant suffisamment les envois pour ne pas surcharger le serveur de courrier électronique.

ANNEXE 2

Étude de cas

Étude de cas: enquête mensuelle sur les ventes (le chiffre d'affaires), 2 000 entreprises, 6 variables:

- Dans les quatre mois qui ont suivi son introduction en mai 2003, le formulaire Web a été utilisé par 70 % des enquêtés (et est actuellement employé par près de 80 % des enquêtés);
- Le nombre d'heures de travail consacrées, par an, à la phase de collecte et d'entrée des données a été ramené de 2 500 heures en 2002 à 1 500 heures en 2004, et il continue de baisser;
- Le pourcentage de réponses reçues à la date d'échéance est passé de 49 à 76 % (au même moment, le délai de réponse a été abaissé de 20 à 17 jours après la fin du mois considéré) et le nombre de notes de rappel est tombé de 1 000 à 500 (autrement dit de 51 à 24 %);
- À la demande de certains enquêtés, l'application Web est désormais accessible aux répondants à partir du premier jour du mois suivant le mois considéré au lieu du cinquième, certains d'entre eux étant pressés de répondre!
- On a enregistré une réduction de sept à huit jours du temps de réponse moyen nécessaire, de même que du délai de préparation des données d'enquête pour la phase suivante de leur exploitation;
- La qualité des données s'est améliorée. Par exemple, il est demandé d'indiquer les «raisons» d'une éventuelle variation très sensible par rapport au mois précédent ce qui facilite la réalisation des ajustements nécessaires;
- Quelques répondants au moins font état d'un allègement de leur travail (le fait qu'autant d'enquêtés utilisent ce mode de réponse montre bien qu'ils le jugent plus avantageux que les formulaires sur papier);
- Le personnel participant à la collecte de données se félicite d'être débarrassé de tâches routinières. L'ampleur du traitement manuel des données a diminué de 50 %;
- L'application considérée est la première des applications de la deuxième génération mises au point par Statistics Finland lui-même. Bon nombre des caractéristiques/solutions techniques ont été modernisées. Cette modernisation a exigé à peu près 850 heures de travail (soit 120 jours), investissement rentabilisé au bout d'un peu plus d'un an. Le travail de maintenance a quant à lui nécessité une centaine d'heures par an;
- Avec XCola, cette application nécessiterait seulement 150 heures de travail environ et l'effort de maintenance serait réduit de moitié.
