

COMMENT AMÉLIORER L'EFFICACITÉ DU CODAGE ?

François BULOT

Insee, Unité Méthodes Statistiques

-Résumé de la communication-

Sicore est le logiciel de codification automatique utilisé par l'INSEE notamment pour coder la profession dans les enquêtes ménages.

Actuellement, pendant la collecte, l'enquêteur a toujours la même série de questions à poser quel que soit le libellé de profession rencontré. Après la phase de collecte, l'ensemble des libellés de profession est soumis à Sicore qui code automatiquement une grande partie des libellés. Les libellés non reconnus sont alors envoyés aux équipes de reprise qui peuvent modifier le libellé puis relancer Sicore ou donner manuellement le code de la profession.

Le travail de reprise représente une lourde charge pour une enquête. L'idée est donc venue d'insérer Sicore directement à la collecte.

L'avantage serait double :

- D'une part, il y aurait la possibilité d'avertir l'enquêteur que Sicore n'a pas reconnu le libellé collecté (et qu'il y a une faute de frappe par exemple), ce qui lui permettrait de modifier le libellé et de limiter ainsi la charge de reprise.
- D'autre part, il y aurait la possibilité de restreindre la passation du questionnaire aux questions qui permettent réellement de lever les ambiguïtés sur la profession, ce qui limiterait le caractère redondant ou inutile de nombre de questions. Les réponses aux questions non posées peuvent alors être directement imputées.

On aurait ainsi un gain de temps non négligeable à la fois au niveau de la collecte, mais aussi et surtout au niveau de la reprise.

Le risque serait une perte de qualité : un enquêteur pourrait remplacer le libellé que lui a donné l'enquêté par un autre assez proche (mais pas totalement identique) dont il serait sûr qu'il serait reconnu par Sicore. Le risque est notamment une simplification trop systématique des libellés pour être sûr que Sicore les reconnaîtra.

Il pourrait y avoir également une perte de temps à la collecte due au fait que l'enquêteur pourrait faire plusieurs tentatives pour trouver un libellé reconnu par Sicore. Mais seuls des tests permettront de vérifier que cette perte de temps sera inférieure au gain en temps de collecte.

