

ACTIONS

par N. Anciaux et B. Nguyen

limiter la collecte des données personnelles

Un problème juridique NP-difficile

Oui, un problème mathématique tout à fait théorique à l'origine peut avoir une application très pratique ! La preuve avec le problème de n -exposition, qui concerne la collecte de données personnelles, et la problématique de l'attribution d'aides sociales. Un défi avec données réelles est proposé aux lecteurs.

Le concept de protection des données personnelles existe depuis longtemps. En France, par exemple, la loi relative à l'informatique, aux fichiers et aux libertés, dite *loi informatique et libertés*, date de 1978 et pose l'existence d'un certain nombre de principes fondamentaux liés au traitement des données personnelles, protégeant ainsi la vie privée des citoyens. Parmi ses nombreux principes, l'un d'entre eux régit le *traitement* des données personnelles : la collecte limitée de données personnelles en vue d'un traitement informatique (voir en encadré).

En effet, la loi de 1978 énonce dans ses articles 6 (sur la collecte et conservation des données) et 7 (sur le consentement) les conditions de licéité d'un traitement (informatique). On voit que la *finalité* du traitement doit être « *déterminée, explicite et légitime* », que les données traitées doivent être « *adéquates, pertinentes et non excessives* », mais également « *exactes et complètes* ». En d'autres termes, il n'est pas légal de collecter des données sans savoir si elles seront utiles ou pas pour le processus.

Traitement des données personnelles

Un *traitement de données personnelles* est un terme général, correspondant à un traitement automatique (informatique) ou non portant sur des données personnelles, quel que soit le procédé utilisé, et notamment la collecte, l'enregistrement, l'organisation, la conservation, l'adaptation ou la modification, l'extraction, la consultation, l'utilisation, la communication par transmission, diffusion ou toute autre forme de mise à disposition, le rapprochement ou l'interconnexion, ainsi que le verrouillage, l'effacement ou la destruction.

Or, nous allons voir que ce principe (juridique) de *minimalité* des données traitées est très difficile à mettre en place de manière pratique puisqu'il pose des problèmes mathématiques (informatiques) complexes : il est en effet NP-difficile de décider, compte tenu d'un traitement modélisé sous la forme de règles logiques permettant une prise de décision multicritère, si un ensemble de données collecté est minimal ou non. À un problème de décision NP-complet (voir l'article consacré à ce thème dans ce hors-série), on peut lier un problème d'optimisation qui cherche à calculer la