

Que dit la loi Informatique et liberté sur le recueil de données à caractère personnel ?

Le recueil et le traitement statistique de données concernant des personnes est régié par la loi informatique et liberté du 6 janvier 1978, modifiée en 2004, relative à l'informatique, aux fichiers et aux libertés. Cette loi s'intéresse aux données « à caractère personnel » destinées à figurer dans des fichiers informatiques ou non, qui permettent d'identifier directement ou indirectement une personne physique. Elle autorise, à certaines conditions, le « traitement » de ces données personnelles directement ou indirectement nominatives, c'est-à-dire leur recueil, leur enregistrement, leur conservation, leur transmission.

Le traitement de données à caractère personnel est autorisé dans les conditions suivantes :

- 1/ il doit répondre à un « objectif déterminé », « légitime » et clairement « explicite »,
- 2/ le recueil de ces données doit être « loyal » et « licite »,
- 3/ les données ne doivent pas être ultérieurement utilisées à d'autres fins que celles annoncées. (Mais les chercheurs peuvent néanmoins obtenir le droit d'élaborer des statistiques à partir de fichiers administratifs ou d'entreprise, quand bien même la collecte des données n'était pas destinée à l'élaboration de statistique ou à la recherche scientifique).
- 4/ Les données collectées au regard de cet « objectif déterminé » doivent être « adéquates, pertinentes et non excessives ». Autrement dit, la liste des données recueillies doit se limiter à ce qui est strictement nécessaire pour répondre à l'objectif annoncé.
- 5/ Enfin, la conservation de ces données personnelles, sous une forme qui permet l'identification des personnes, ne doit pas excéder la durée nécessaire pour répondre aux objectifs annoncés.

Dans le cas de la réalisation d'enquêtes statistiques, ces dispositions de la loi et notamment le point 5, induisent qu'après la phase de collecte des données auprès des individus (phase qui nécessite généralement de connaître leur nom, leur adresse, leur numéro de téléphone), les données enregistrées dans les fichiers informatiques ne doivent pas permettre l'identification des personnes. Les fichiers doivent donc être anonymes : les noms, les numéros de téléphone, les adresses ne peuvent figurer dans les fichiers. De plus, le recoupement des informations restantes dans le fichier ne doit pas permettre l'identification d'une personne. Seraient considérées comme indirectement nominatives des informations (comme l'âge, la ville de résidence, la profession, le sexe, etc.) si elles leur recoupement permettait d'identifier quelqu'un. Par



Enquête sur la diversité des populations en France

exemple, la mention du département de résidence et de la profession « préfet » permettent immédiatement d'identifier une personne physique car il n'y a qu'un seul préfet par département. Pour éviter une identification aussi facile et protéger les personnes, les informations recueillies sur les professions doivent faire l'objet de regroupements (tels que ouvriers, employés, cadres, cadres dirigeant, etc...) qui empêchent l'identification directe. Cet impératif d'anonymisation implique également que l'effectif des individus contenus dans le fichier soit suffisamment élevé. Enfin, la liste des noms, adresses, numéros de téléphone qui a servi à contacter les enquêtés doit être détruite dès que sa conservation n'est plus nécessaire à la réalisation de l'enquête.