La protection des données à caractère personnel

François PELLEGRINI
Professeur, Université de Bordeaux

francois.pellegrini@labri.fr

Identité (1)

- Chaque personne a, dans le monde physique, une ou plusieurs apparences :
  - Professionnelle
  - Amicale
  - Associative
  - Etc.
- L'identité est la manière dont une entité est reliée à ses apparences
- L'identité numérique est le lien créé, au moyen des technologies numériques, entre une personne et ses diverses apparences numériques

Identité (2)

- Perçue au travers de nombreuses facettes :
  - Ce que l'on déclare de soi
    - Informations fournies à l'inscription sur un site
  - Ce que l'on montre de soi
    - Ses actions délibérées
  - Ce que l'on peut connaître
    - Traces que l'on laisse

Données personnelles et contrôle (1)

- La collecte de données à caractère personnel est une activité très ancienne
  - Concomitante à l'invention de l'administration
    - Collecte de l'impôt
  - Concomitante à l'invention de l'écriture !
- L'usage de ces données pour le contrôle des populations est également ancien
  - Revenus, hérédité et castes, religion, etc.

Données personnelles et contrôle (2)

- La notion de vie privée est récente
  - Dans les villages, tout le monde sait tout sur tout le monde !
- La création d'« identités de papier » a été rendue nécessaire par l'exode rural
  - Création d'une chaîne de confiance entre l'émetteur et le lecteur du document
  - Rôle assuré par l'administration

Données personnelles et contrôle (3)

- Le danger de la collecte massive des données personnelles est apparu avec l'automatisation de leur traitement
  - Fichage et numérotation des populations « sensibles »
    - Casier judiciaire (et autres « sommiers »)
    - Livret ouvrier
    - Carte d'identité pour les populations nomades et indigènes
  - Utilisation de la mécanographie pour la mise en œuvre de tris a posteriori
    - Cas des Pays-Bas dans les années 1940

Données personnelles et contrôle (4)

- La puissance des outils numériques a encore accru les possibilités de contrôle des populations
  - « Croisements » entre fichiers et non plus seulement tris au sein d'un unique fichier déjà constitué

Données personnelles et contrôle (5)

- Crainte d'une intrusion démesurée des États dans l'intimité des individus
  - À l'époque, seuls les États avaient la capacité de collecter des masses de données
- Glissement ultérieur de la menace vers le secteur privé
  - Disposent de plus d’information que les États
- Retour en force des États qui imposent d'accéder plus ou moins secrètement à ces gisements

Protection des données personnelles (1)

- Création de lois spécifiques
  - En France, loi « Informatique et libertés » de 1978
- Création d'organes de contrôle indépendants de l'exécutif et des administrations
  - Modèle juridique original d'« Autorités administratives indépendantes »
    - Ne peuvent appartenir aux autres pouvoirs en vertu même de la séparation des pouvoirs
  - En France, la CNIL
    - 18 commissaires et ~ 215 personnels

Protection des données personnelles (2)

- Article 1 de la loi « Informatique & Libertés » :

  L'informatique doit être au service de chaque citoyen. Son développement doit s'opérer dans le cadre de la coopération internationale. Elle ne doit porter atteinte ni à l'identité humaine, ni aux droits de l'homme, ni à la vie privée, ni aux libertés individuelles ou publiques.

  Toute personne dispose du droit de décider et de contrôler les usages qui sont faits des données à caractère personnel la concernant, dans les conditions fixées par la présente loi.

- Principe d'*autodétermination informationnelle*
  - Reconnu de rang constitutionnel en Allemagne dès 1983

Missions de la CNIL

- Autorisation
  - Étude des projets de traitements
  - Élaboration de doctrines relatives aux différents types de traitements
  - Élaboration de normes simplifiées, d'autorisations uniques, de référentiels
- Contrôle
  - Contrôles en ligne depuis 2014
- Sanction
  - Tribunal administratif spécialisé en matière de données personnelles

Critères de licéité

- La licéité d'un traitement de données personnelles est jugée selon les critères suivants :
  - Finalité : déterminée, explicite et légitime
  - Responsable du traitement
  - Destinataires des données traitées
  - Durée de conservation
  - Mesures de sécurité de conservation
  - Conditions d'information, droit d'accès et de rectification
- Contrôles effectués :
  - Proportionnalité des dispositions et moyens mis en œuvre pour les respecter

Champ d'application des lois « I&L » (1)

- Concerne exclusivement les personnes physiques
- S'appliquait aux « informations nominatives »
  - Directement associées au nom de l'individu
- Extension de son périmètre aux informations « indirectement » nominatives
  - Numéros de plaque d'immatriculation, de téléphone, etc.
- Extension aux « données à caractère personnel »
  - Tout ce qui est, directement ou indirectement, rattaché aux personnes physiques
    - Biométrie, traces comportementales (méta-données), etc.
  - Il s’agit en fait plutôt de données à « caractère interpersonnel »

Champ d'application des lois « I&L » (2)

- Exemples de doctrines :
  - Existence d'identifiants distincts selon les secteurs
    - Séparation du NIR (le « numéro de sécu ») et du NUMEN
    - Mais extension passée du NIR au domaine de l'entreprise
  - Restriction de l'usage de la biométrie au contrôle d'accès
    - La biométrie n'est pas révocable
  - Nécessité d'informer les personnes et de leur permettre le contrôle de l'usage de traceurs (« *cookies* », « *device fingerprinting* ») selon les catégories de finalités :
    - Technique
    - Mesure d'audience
    - Traçage publicitaire

Protection intégrée des données

- Prendre en compte la protection des données à caractère personnel dès la conception des dispositifs (« *privacy by design* »)
- Sept principes :
  - Mesures proactives et non réactives
  - Protection implicite de la vie privée
  - Protection dès la conception des systèmes et pratiques
  - Fonctionnalité intégrale à somme positive, pas nulle
  - Sécurité de bout en bout tout au long de la conservation
  - Assurances de visibilité et de transparence
  - Respect des utilisateurs

Anonymisation et réidentification (1)

- L'anonymisation des données est un sujet critique à l'ère du numérique
  - Conflit avec le mouvement d'ouverture des données
    - Anonymisation des données de justice
    - « Droit à l'oubli »
  - Alimentation des algorithmes de traitement de mégadonnées (« *big data* ») destinés à détecter des « signaux faibles »
    - Suivi du comportement des clients
    - Gestion des populations : « prédiction » de crimes
  - Respect de la vie privée des citoyens
    - Nécessité de ne pas pouvoir réidentifier

Anonymisation et réidentification (2)

- La réidentification est un problème ancien
  - Loi du 7 juin 1951 « sur l'obligation, la coordination et le secret en matière de statistiques »
  - Création du « Comité du secret statistique »
    - L'INSEE effectue une « dilution » (« *binning* ») de ses données avant de les transmettre à ses usagers
- Expériences récentes sur la réidentification des masses de données
  - Seulement quatre points de mesure nécessaires !
  - Une dilution significative n'augmente pas significativement le nombre de points de mesure nécessaires

La guerre des données personnelles

- Deux conceptions juridiques des données personnelles s'affrontent au niveau mondial
  - En Europe, les données personnelles sont attachées à la personne et leur contrôle est inaliénable
    - Pas de « propriété des données »
    - Rôle central du consentement dans les règlements RGPD et ePrivacy
  - Aux États-Unis, les données sont un bien dont le contrôle est cessible
    - Pas de loi « Informatique et Libertés »
      - Régulation par la FTC, sur le terrain économique
    - Exemple de la revente des données de navigation par les FAI

Gestion de la biométrie

- Les données biométriques sont extrêmement sensibles
  - Non révocables car intimement associées aux personnes
- Deux principaux usages :
  - L'authentification : pouvoir assurer qu'une personne est bien la bonne
  - L'identification : retrouver une personne dans une base de données à partir de ses traces
- Question de l'architecture des dispositifs
  - Les architectures centralisées permettent les deux
  - Les architectures décentralisées ne permettent que l'authentification

La guerre de la biométrie

- Actions offensives des États-Unis pour amasser des stocks de données biométriques relatives à l'ensemble de la population mondiale
  - « Piratage » des outils de collecte biométrique des services de renseignement alliés (CIA/ExpressLane)
  - Collecte massive de documents sur l'Internet (NSA/PRISM et NSA/MUSCULAR)
    - A permis l'identification de « Satoshi Nakamoto » par stylométrie (biométrie du style d'écriture)
  - « Dons » à certains États d'équipements biométriques de contrôle de passage aux frontières
  - Etc...

Éléments de plan stratégique (1)

- Protéger fortement les données personnelles
  - Faire respecter la notion de consentement
- Rejeter la surveillance généralisée des échanges numériques
  - Pas de portes dérobées dans les protocoles et logiciels de communication
- Rendre difficile la surveillance de masse
  - Garantie démocratique

Éléments de plan stratégique (2)

- Appuyer les développements de l'État sur des architectures décentralisées plutôt que centralisées
  - Risque moindre en cas de mésusage
  - Refuser l'« identifiant unique des usagers »
    - Mettre en place une fédération d'identités permettant l'échange de données entre administrations (à la « FranceConnect »)
  - Créer un système d’identité « faillible par défaut »
    - Pas de base biométrique centralisée !
- Rejeter la surveillance généralisée des échanges numériques

Bibliographie

- *IBM et l'holocauste*, Edwin Black, Robert Laffont, 2001
- *Le profilage des populations*, Armand Mattelard & André Vitalis, La Découverte, 2014
- Travaux de Y.-A. de Montjoye et autres sur la réidentification