

---

# Les données personnelles à l'ère numérique

François Pellegrini  
Professeur, Université de Bordeaux  
francois.pellegrini@labri.fr

Ce document est copiable et distribuable librement et gratuitement à la condition expresse que son contenu ne soit modifié en aucune façon, et en particulier que le nom de son auteur et de son institution d'origine continuent à y figurer, de même que le présent texte.

# Identité numérique (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

# Identité numérique (2)

---

- L'identité numérique est le lien créé, au moyen des technologies numériques, entre une personne et ses diverses apparences numériques
- 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
  - Ce que l'on peut connaître
    - Traces que l'on laisse

# Identité numérique (3)

---

- Question des traces et de l'anonymat
  - L'anonymat est (encore) possible dans le monde physique
    - Payer en liquide, etc.
  - Cet anonymat peut-il exister dans le monde numérique ?
    - Les teneurs de compte numériques centralisés sont informés de toutes les transactions
    - D'où développement des crypto-monnaies, telles que le Bitcoin

# Données personnelles et contrôle (1)

---

- La collecte de données personnelles 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)

---

- 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 (3)

---

- 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 (4)

---

- 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
- Retour en force des États qui imposent d'accéder plus ou moins secrètement à ces gisements



# Données personnelles et contrôle (5)

---

- 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 ~ 180 personnels

# Missions de la CNIL

---

- **Autorisation**
  - Étude des projets de traitements
  - Élaboration de doctrines sectorielles
  - Élaboration de normes simplifiées et d'autorisations uniques
- **Contrôle**
  - Contrôles sur place mêlant juristes et informaticiens
  - Contrôles en ligne depuis 2014
- **Sanction**
  - Tribunal administratif spécialisé en matière de données personnelles

# Critères de conformité

---

- Finalité
  - Déterminée, explicite, légitime
- Responsable du traitement
- Destinataire des données traitées
- Durée de conservation
- Mesures de sécurité de conservation
- Conditions d'information, droit d'accès, de rectification et (parfois) d'opposition

# Champ d'application de la loi « 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 d'immatriculation, de téléphone, etc.
- Extension aux « données personnelles »
  - Tout ce qui est, directement ou indirectement, rattaché aux personnes physiques
    - Biométrie, traces comportementales (méta-données), etc.

# Champ d'application de la loi « 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 du NIR au domaine de l'entreprise
  - Usage de la biométrie restreint au contrôle d'accès
    - Mais autorisation pour les cantines scolaires !
    - La biométrie n'est pas révocable
  - Nécessité d'informer les personnes de l'usage des traceurs (« *cookies* ») à vocation publicitaire
    - Distinction entre les traceurs « techniques » et les traceurs « publicitaires »
    - Règlement « e-Privacy » en cours de négociation

# Droit à l'oubli

- Droit récent ayant pour origine l'arrêt « Costeja » de la CJUE du 13 mai 2014
  - Fondé sur les droits de correction et d'opposition reconnus aux personnes concernées par la Directive 95/46 sur la protection des données
- Droit à demander le déréférencement dans les moteurs de recherche
  - Suppression du lien entre le nom de la personne et les ressources
  - Lien « inapproprié » car information devenue inexacte

# Protection intégrée de la vie privée (1)

---

- Prendre en compte la protection de la vie privée dès la conception des dispositifs (« privacy by design »)
- Assurance supplémentaire pour les responsables de traitements quant à leur conformité aux lois « I&L »

# Protection intégrée de la vie privée (2)

---

- 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
  - Alimentation des algorithmes de traitement de « mégadonnées » destinés à détecter des signaux faibles
    - Suivi du comportement des clients : affaire « Target »
    - 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 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
    - Rôle central du consentement dans les règlements RGPD et e-Privacy
  - Aux États-Unis, les données sont un bien dont le contrôle est cessible
    - Pas de loi « Informatique et Libertés »
      - Régulation par la FCC, sur le terrain économique
    - Exemple de la revente des données de navigation par les

FAI

# Gestion de la biométrie (1)

---

- 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

# Gestion de la biométrie (2)

---

- Question de l'architecture des dispositifs
  - Les architectures centralisées permettent les deux
  - Les architectures décentralisées ne permettent que l'authentification
- Question de l'extension des finalités des fichiers
  - « Réquisitions judiciaires » dans TES-2
  - « Recherche en parentèle » dans le FNAEG

# 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

# Bibliographie

---

- *IBM et l'holocauste*, Edwin Black, Robert Laffont, 2001
- *Le profilage des populations*, Armand Mattelard & André Vitalis, La Découverte, 2014
- La législation !
  - C'est facile à lire. Si, si... « *Law is Code* »
- Travaux de Y.-A. de Montjoye et autres sur la réidentification