

---

# Données à caractère personnel, données ouvertes et innovation

François Pellegrini  
Professeur, Université de Bordeaux  
francois.pellegrini@u-bordeaux.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.

# Donnée

---

- L'information est un bien immatériel
  - N'a pas de « propriétaire » ni ne se « vole »
- Différents régimes juridiques applicables aux différentes catégories de données
  - Droit d'auteur
    - Pour les œuvres de l'esprit exhibant une originalité
  - Droit des données à caractère personnel
    - Pour les données relatives à une personne physique
  - Secret des affaires
  - Etc.
  - Et leur application contractuelle : NDA, CGU, etc.

# Lois « Informatique & Libertés » (1)

---

- Créées en réaction au mésusage des données à caractère personnel par les États durant la première moitié du XX<sup>e</sup> siècle
  - En France, loi « Informatique et Libertés » du 6 janvier 1978
- Création d'organes de contrôle indépendants de l'exécutif et des administrations
  - Modèle juridique original d'« Autorités administratives indépendantes »
- En France, la CNIL

# Lois « Informatique & Libertés » (2)

---

- Article 1 de la loi « I&L » :
  - « 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 »

# Donnée à caractère personnel

---

- Art. 4 RGPD :

« Toute information se rapportant à une personne physique identifiée ou identifiable [...] ; est réputée être une « personne physique identifiable » une personne physique qui peut être identifiée, directement ou indirectement, notamment par référence à un identifiant, tel qu'un nom, un numéro d'identification, des données de localisation, un identifiant en ligne, ou à un ou plusieurs éléments spécifiques propres à son identité physique, physiologique, génétique, psychique, économique, culturelle ou sociale »

# Donnée sensible

- Art. 9 RGPD :

« Le traitement des données à caractère personnel qui révèle l'origine raciale ou ethnique, les opinions politiques, les convictions religieuses ou philosophiques ou l'appartenance syndicale, ainsi que le traitement des données génétiques, des données biométriques aux fins d'identifier une personne physique de manière unique, des données concernant la santé ou des données concernant la vie sexuelle ou l'orientation sexuelle d'une personne physique sont interdits. [...] »

# Critères de licéité

- Licéité appréciée selon les critères suivants :
  - Finalité : déterminée, explicite et légitime
  - Base légale appropriée
  - Responsable du traitement
  - Destinataires des données traitées
  - Durée de conservation
  - Mesures de sécurité de conservation
  - Exercice des droits des personnes : information, droit d'accès, de rectification, d'opposition, etc.
- Obligation de moyens renforcée
- Proportionnalité des moyens mis en œuvre

# Bases légales d'un traitement

---

- Article 6 du RGPD :

« Le traitement n'est licite que si [...] une des conditions suivantes est remplie:

a) la personne concernée a consenti au traitement de ses données [...];

b) le traitement est nécessaire à l'exécution d'un contrat [...]

[...]

f) le traitement est nécessaire aux fins des intérêts légitimes poursuivis par le responsable du traitement ou par un tiers, à moins que ne prévalent les intérêts ou les libertés et droits fondamentaux de la personne concernée [...] »



# Périmètre du RGPD

- Le RGPD ne s'applique pas aux données :
  - Qui ne concernent pas des personnes physiques
    - P. ex. : données d'entreprises
      - Même si certaines données relatives aux entreprises peuvent être des données personnelles
  - Concernant des personnes mais ne permettant pas de les identifier
    - Données anonymes
      - Et non pas pseudonymes !
- Les données personnelles rendues publiques sont toujours soumises au RGPD
  - En particulier quant au détournement de finalité

# Anonymisation et réidentification (1)

---

- L'anonymisation des données est une tâche délicate
  - Elle ne peut se réduire à la seule pseudonymisation
- Plusieurs types d'attaques permettent de réidentifier les personnes concernées par des données pseudonymisées
  - En particulier les attaques par corrélation

# Anonymisation et réidentification (2)

- La portée des attaques par corrélation dépend de la granularité des données collectées
  - Au menu : spaghetti, macaroni ou semoule ?
- Le niveau de granularité détermine un compromis entre :
  - L'utilisabilité des données pour la finalité recherchée ou des finalités ultérieures
  - La protection des personnes contre la réidentification
- Analyse préalable avant toute « ouverture des données »

# Protection intégrée des données

---

- Nécessité de ne collecter que les données strictement nécessaires aux finalités
  - Principe de minimisation des données
- Prendre en compte la protection des données dès la conception des dispositifs (« *data protection by design* »)
  - Intégrer les mécanismes d'attaque dans l'analyse préalable à la collecte des données
- Gérer l'archivage et le partage ultérieur

# Données personnelles et recherche

---

- Tension entre les obligations des lois « I&L » et les pratiques scientifiques
  - Conservation permanente des données à fin de réfutabilité par des tiers
    - Sans nécessairement les partager
- Prise en compte du cycle de vie de la donnée
  - Depuis la collecte jusqu'à l'archivage
    - P. ex. : registres militaires des « poilus » de 14-18
- Consultation en cours de la CNIL auprès de la communauté scientifique

# Conclusion

---

- Incitations de plus en plus fortes à confier des données à caractère personnel de plus en plus intimes contre la promesse de bénéfices de plus en plus incertains
  - Données de comportement
  - Données médicales et génétiques
- Nécessité :
  - De prise en compte du risque à long terme
  - D'un encadrement fort
  - D'une éducation des personnes à ces enjeux