

INFORMATION SCIENTIFIQUE ET TECHNIQUE
LICENCE 1 SHA UPMF 2015-2016
SÉANCE 4

MES DONNÉES ?
QUELQUES REFLEXIONS À PROPOS DE LA
PROBLÉMATIQUE DES DONNÉES PERSONNELLES

ROMAIN VANEL

Note : *document mis à jour le plus souvent possible, notamment la bibliographie. Dernière MAJ : 17 mars 2016.*

Les catalogues de bibliothèques contiennent des données : titres, auteurs, etc. Par ex. `idref.fr` : la base d'autorité de l'Abes qui contient des données personnelles.

Qu'est ce qu'une donnée ?

- Personnelle Loi IL Art 2 « Constitue une donnée à caractère personnel toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement, par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres. Pour déterminer si une personne est identifiable, il convient de considérer l'ensemble des moyens en vue de permettre son identification dont dispose ou auxquels peut avoir accès le responsable du traitement ou toute autre personne. »

Soit : « un nom, un prénom, une photo, une image, une adresse IP ».

La loi régit toutes les informations qui permettent d'identifier directement ou indirectement une personne physique.

1. HISTOIRE ET CONTEXTE

1.1. De la réglementation.

1.1.1. *Internationale*. Conférence Internationale sur la protection des données personnelles en 1983. Création d'un groupe de travail.

Divulgarion du réseau dit "ECHELON" : réseau international d'interception de données privées et publiques¹.

Art. 3. Directive 2006/24/CE sur la conservation des données.

Date: 17 mars 2016.

Key words and phrases. IST, information, données personnelles, accès, data.

1. <https://fr.wikipedia.org/wiki/Echelon>.

« Les articles 5, 6 et 9 de la directive 2002/58/CE définissent les règles applicables au traitement, par les fournisseurs de réseaux et de services, de données relatives au trafic et de données de localisation générées par l'utilisation de services de communications électroniques. Ces données doivent être effacées ou rendues anonymes lorsqu'elles ne sont plus nécessaires à la transmission d'une communication, sauf les données requises pour établir les factures et les paiements pour interconnexion ; moyennant l'accord de l'intéressé, certaines données peuvent également être traitées à des fins commerciales ou de fourniture de services à valeur ajoutée. »

Un « Contrôleur européen de la protection des données ».
Journée de la protection des données : 28 janvier.

1.2. **Du Web...** Né en 1990 au Cern. Beaucoup d'évolution depuis le premier serveur `nxoc01.cern.ch`. Présenté en 1992 par Tim Berners-Lee.

La première page <http://nxoc01.cern.ch/hypertext/WWW/TheProject.html>²

Grandes évolutions du Web³.

- Web 1.0. Statique. Aucune interaction avec l'utilisateur, qui peut seulement lire le contenu de la page. Ex : <https://www-fourier.ujf-grenoble.fr/~vanelro/>. Souvent écrit seulement en HTML.
- Web 2.0. Web dynamique. Ex : tous les sites capables de renvoyer des infos en fonction d'autres informations de diverses sources : IP, formulaire, Ex : Wikipedia. La météo de votre ville. Programme TV de ce soir, réseaux sociaux... Basés sur diverses technologies : Php/MySQL ; Ajax etc.
- Web 3.0 Web de données. Données « cachées », réseau de ces données. Ex : les bases qui permettent l'échange de données comme data.bnf.fr, wikidata.org...

Toute cette évolution est génératrice d'un volume de données de plus en plus importants. De toutes natures. L'utilisateur est au coeur du système.

2. La première conservée : <http://www.w3.org/History/19921103-hypertext/hypertext/WWW/Link.html>

3. Voir notamment dans : http://www.univ-bpclermont.fr/Ressources_Num/Les_reseaux_sociaux_web_web/co/module_Les_reseaux_sociaux_2.html

2. DE LA COLLECTE DES DONNÉES...

2.1. **Collecte « volontaire »**. Une partie des données collectées le sont avec le consentement de celui qui est visé. L'utilisation de certains services et outils génère des données. Mais ces utilisations sont facultatives⁴.

Par quels canaux les données sont-elles collectées ?

- boîte email (par exemple, Google, qui affiche de la publicité contextuelle à côté des messages)
- réseaux sociaux : données ajoutées par l'utilisateur (nom, prénom, goûts etc.), puis mises à jour en permanence en fonction de l'utilisation du réseau : pages consultées, aimées, commentées, localisation des publications, heure d'utilisation, etc.
- achats en ligne, par correspondance (déjà La Redoute ! Mailing papier etc). Mouse Tracking : analyse des déplacements sur un site ou page web. Par exemple, laisser la souris positionner sur un produit, puis un autre...
- cloud (espaces de stockage en ligne) : icloud, Dropbox, Google Drive, etc
- livres sur supports numériques
- Mise en relations diverses : Airbnb, BlaBlaCar, Le Bon Coin, Booking
- musique en ligne
- assurance : ristourne si l'on est équipé d'un mouchard. Mais aussi : Plus de sport = rabais sur les tarifs de primes d'assurance !
- Authentification via d'autres sites : les sites récupèrent beaucoup plus de données que nécessaire. Ex : authentification par Google ou Facebook...

2.2. **Collecte « involontaire »**. D'autres données sont collectées à l'insu de l'utilisateur, qui n'a pas « le choix » dans les données qu'il transmet ou non.

- Traceurs sur les sites web de type **GoogleAnalytics** : IP, heure, pays, mots utilisés...
- Logs des serveurs, des DNS... (voir par exemple la commande **traceroute**) : IP, provenance, destination...
- Logs des fournisseurs d'accès à Internet (conservés pendant 1 an)
- Applications : toutes. Lisez vous les conditions ? Une partie des données collectée volontairement. Souvent, accès à BEAUCOUP plus de données que nécessaire.
- Cookies : sur la machine, souvent utilisés par plusieurs sites pour liens entre les services. Pub ciblée. Ex : publicité pour un produit recherché.

4. Nous ne parlons pas ici de la pression sociale qui pourrait laisser entendre que l'utilisation de certains services est obligatoire. Ainsi, l'usage d'une boîte email résulte de la volonté de chacun.

- GPS. Partout. Téléphones, tablettes sont équipés de GPS. Google conserve un historique des positions. Mais qui a réellement accès ? Officiellement, seulement l'utilisateur.
- Amazon : livres lu, consultés, achetés... Et pour chaque livre : pages lues, la vitesse de lecture etc. !
- Accès sans fil, et connexion automatique des appareils : bornes wifi, bornes téléphone
- toutes les cartes sans contacts, entrées des magasins, RFID très déployé : toutes les transactions de ces cartes sont loguées .

2.3. Collectées et créées par d'autres... Enfin, les tiers peuvent également générer et collecter des données sur d'autres tiers...

- réseau sociaux : images « taggées » par d'autres utilisateurs : Facebook, Copains d'avant etc.
- entreprise : le nom apparaît dans les contacts du site internet, les annuaires, les organigrammes etc.
- Vidéosurveillance. Nommée « vidéoprotection » depuis Brice Hortefeux. De plus, parfois gérée par des entreprises privées. Qui a accès à ces données ?
- Outils de surveillance divers : Hadopi...

3. DE L'UTILISATION DES DONNÉES PERSONNELLES

Pour cette partie, utilisation du livre *Economie des données personnelles et de la vie privée* [8], en particulier le chapitre *III. Exploitation des données personnelles, vie privée et externalités*.

Problème : données durables. Ne se détruisent pas avec le temps.

Données pas forcément à jours. Des conséquences économiques. Pertes de crédibilité des entreprises si demande trop intrusive. Mais beaucoup de contre exemple.

Question de la sécurité des données collectées.

3.1. Monétisation des données. Utilisateurs de plus en plus sensibles à la revente de leur données.

Véritable marché. En plus des biens vendus, l'entreprise collecte des données qu'elle peut revendre.

- Facebook a acheté l'application WhatsApp env. 20 milliard \$. Paie pour les données et presque pas pour l'application !

Les données d'achat sont les plus monétisées : cartes de fidélités etc... Mais elles rapportent sans doute plus que les bons de réductions auxquels elles donnent droit !

3.2. Publicité.

3.2.1. Marketing direct. Au téléphone, courrier, etc.

Très encadrée en France.

Cibles choisies pour leur vulnérabilité, la composition de leur foyer etc. Les données aident à cibler.

3.2.2. *Spamming*. Anonymat de l'expéditeur, coût technique négligeable = succès du spam.

Mais beaucoup d'aspects négatifs : demande du temps de lecture, de traitement etc. Peut nuire aux annonceurs au final.

3.2.3. *Sur le web*. Personnalisation des services : tout est personnalisable.

- services personnalisés : recommandations, navigation, facilitation des achats : souvent à partir des données transmises « volontairement »
- services individualisés : données de contexte. Ex : pub d'un produit sur un produit en fonction du site ; pub sur les mails ; recommandation à partir de la navigation, des autres achats (cookies)

Exemple de Criteo :

« Cette société fait du retargeting : supposons que lundi je visite le site d'un de leurs clients, la compagnie A, mais que je n'y achete rien. Mardi, je visite le site d'une autre compagnie B, qui est aussi client de Criteo. Criteo s'en rend compte, et offre à A la possibilité de mettre une publicité sur le site de B, avec l'espoir de me faire consommer chez A. ⁵ »

3.3. Tarification dynamique.

- Amazon : tarification en fonction du comportement de l'internaute : achat, pages consultées, temps passé...
- Avions : *idem* + en fonction de(s) l'adresse(s) IP utilisée(s).
- Marketing du crédit, de l'assurance : données pour négocier avec le client et moduler les tarifs.

3.4. **Cas des données médicales.** Assurances, banques, employeurs potentiels.

Secret médical en plus des données personnelles.

Importance de la sécurité de ces données.

3.5. **Modélisation de la société et des comportements.** «On peut en tirer des modèles sur la façon dont vous vous comportez : à qui vous parlez, où vous allez, quels sites vous consultez... Il n'est pas nécessaire de savoir ce que vous dites. Ce qui importe, c'est d'être capable de cartographier les relations entre les gens et de savoir à qui vous êtes susceptible de le dire ⁶ »

5. Exemple cité dans l'enquête #MesDonnes : <https://www.facebook.com/groups/mesdonnees/permalink/1665758397000489/> [9].

6. Voir dans [4] : <http://rue89.nouvelobs.com/2015/05/12/nai-rien-a-cacher-tracking-mobile-explique-a-pere-259155>.

4. PROTÉGER LES DONNÉES

4.1. **Traces volontaires.**

- « Soigner » son/ses profil(s) (Problème des ressources humaines. . .)
- Utilisation de pseudonymes, mais certains sites imposent le vrai nom, et la vérification : Airbnb, . . .
- Consulter les politiques de confidentialité
- Utiliser les outils en connaissances de cause.
- Mais ces outils ont-ils besoin de toutes ces données ?
- Se poser des questions lorsqu'un outil requiert une authentification via Google ou FB. Pourquoi ? Recupération de beaucoup plus de données que nécessaire.
- Utiliser des outils respectant la vie privée : des logiciels libres et ouverts (accès au code source pour vérifier l'utilisation des données), outils de Framasoft, Tor, Protonmail, plugin Cryptocat, Qwant, DuckDuckGo etc
- Intérêt de la généralisation du *Sans contact* ? Pour qui ? Dans le commerce, pourquoi est-il moins cher pour les commerçants que le « avec contact » pour le même service ? Dans les transports ? Au Crous ?

4.2. **Involontaires.**

- Bloquer les indésirables avec des outils dédiés : Ghostery, Ad-block. . .
- Bloquer les cookies. Problème de fonctionnement de certains sites. . .
- Utiliser des proxy, des VPN, Tor, son propre serveur DNS. . .

4.3. **La loi nous protège-t-elle ?** « ça ne me concerne pas, je n'ai rien à cacher ! »

Mais ce n'est pas pour ça qu'on ne vous le reprochera pas un jour !

4.3.1. *Informatique et liberté.* CNIL Histoire rapide

- En 1974, projet d'un fichier pour identifier citoyens pour administration : le projet SAFARI⁷
- Inquiétude > création d'une commission en 1978.

Publication de la *loi relative à l'informatique, aux fichiers et aux libertés du 6 janvier 1978*.

Fondement de la protection des données personnelles en France.

Tout est déclaré à la Cnil sous peine de sanction. Pour une entreprise : le moindre fichier informatique contenant un nom est éligible.

4.3.2. *Loi renseignement.* Quels problèmes ?

Les opposants à la loi notent qu'elle permet :

- interception de données personnelles en temps réel sur Internet

7. Le monde : http://rewriting.net/wp-content/le_monde_-_21_03_1974_009-3.jpg

- extension des pouvoirs de services de renseignement, sans contrôle des citoyens
- légalisation de pratiques illégales par les services de renseignement
- contrairement aux services web dont le choix d'utilisation est libre, ce n'est pas le cas ici.
- renforcement de cette loi suite aux attentats du 13 novembre ?

4.4. #MesDonnées : accès à ses données chez des tiers.

4.4.1. *Une obligation en France.* Obligation de la loi informatique et liberté. Mentions légales obligatoires.

Par exemple sur le site du journal *Le Monde* :

« Conformément aux dispositions de la loi du 6 janvier 1978 relative aux fichiers, à l'informatique et aux libertés, les clients du Monde interactif disposent d'un droit d'accès, de rectification et d'opposition aux données personnelles le concernant. Ce droit s'exerce soit en en modifiant ses informations personnelles en ligne soit auprès de :

LE MONDE INTERACTIF
 Direction marketing
 80, boulevard Auguste-Blanqui
 75707 Paris Cedex 13 »

4.4.2. *Des exemples d'accès.*

Exemple d'Airbnb. <http://rue89.nouvelobs.com/2015/10/07/99-pages-a4-airbnb-deballe-tout-quil-sait-261533>

Exemple de Criteo. <https://www.facebook.com/groups/mesdonnees/permalink/1665758397000489/>

Exemple d'Amazon. Exemple des données de Ph. Girard, RTS, opération *Donnez-moi mes données!*

Peu de données transmises : seulement coordonnées privées, date et prix d'achat, moyen de paiement. Pas de mention du produit acheté. Un autre fichier contenant les IP des opérateurs qui ont transmis les commandes sur smartphone.

« Moi qui m'attendais à en savoir un peu plus sur le profilage des clients ou sur l'algorithme d'Amazon qui permet au site de suggérer des "Recommandations pour vous", quelle déception ! »

Mais alors, où sont ces données personnelles ?

4.4.3. *La « mort numérique ».* Que deviennent nos données personnelles à notre décès⁸ ?

8. http://www.liberation.fr/futurs/2015/10/29/le-droit-francais-en-passe-de-reconnaitre-la-mort-numerique_1409717

- les données sont personnelles : impossible pour un tiers d’y accéder. Le droit d’accès terminé au décès.
- héritier pas de droit d’accès.
- loi : possible de contacter les différents services, entreprises, administrations, pour informer d’un décès et clore les comptes. Mais pas supprimer.
- plusieurs types de réponses, selon les services.
- la question devrait être tranchée dans la loi « République numérique »
 - on pourra, vivant choisir le sort des données après le décès (conservation, effacement etc.)
 - désigner une personne chargée de l’exécution de ces choix
 - déjà possible sur FB de choisir un légataire et de cocher une case pour suppression au décès.
- problème de l’application de la loi à des sites non français.
- rôle des notaires ? des testaments ?

4.4.4. *Prise de conscience médiatique : l’exemple de l’enquête ouverte*
 Donnez-moi mes données ! Enquête lancée par les journalistes et chercheurs de l’émission *On en parle* sur *La Première* (Radio Télévision Suisse).

But : faire le point sur l’accès, l’utilisation, l’échange, le commerce et la détention des données personnelles par tout organisme susceptible d’en collecter ou d’en utiliser.

Ouverte, car elle suggère une forte participation des citoyens dans les demandes d’accès aux données.

Publication de ces résultats sur

- Facebook : <https://www.facebook.com/groups/mesdonnees/>
- site de l’enquête : <http://www.rts.ch/mesdonnees>

Pas d’initiative équivalente en France.

5. CONCLUSION : VERS UNE SURVEILLANCE GÉNÉRALISÉE ?

Collecte de plus en plus importante, que les données soient transmises volontairement ou non.

La loi est-elle suffisante pour protéger ces données personnelles et leur utilisation ? Ou se situe l’utilisation abusive ?

Les données sont-elles en sécurité ?

Liens entre privé et les États par la sous-traitance de :

- Vidéosurveillance
- HADOPI
- Problème des *backdoor* dans les logiciels
- Traces involontaires de plus en plus nombreuses.
- Quel contrôle, pour, à terme, quels usages ?

Mais aussi :

- Accès facilité des services de renseignements aux données de connexion
- Question de l'état d'urgence post-attentats et des libertés individuelles.

Selon un sondage post-attentats⁹, 84% des personnes interrogées sont d'accord pour « limiter leurs libertés pour mieux garantir leur sécurité ».

Les données et informations personnelles sont bien entendu au coeur de la question.

RÉFÉRENCES

- [1] Philippe Boucher, « *safari* » ou la chasse aux français, *Le Monde* (21/03/1974), URL [\url{http://rewriting.net/wp-content/le_monde_-_21_03_1974_009-3.jpg}](http://rewriting.net/wp-content/le_monde_-_21_03_1974_009-3.jpg).
- [2] Guillaume Champeau, *Déjà une nouvelle loi sécuritaire menaçant encore la vie privée*, *Numerama* (02/12/2015), URL [\url{http://www.numerama.com/politique/133397-vers-une-nouvelle-loi-securitaire-menacant-la-vie-privee.html}](http://www.numerama.com/politique/133397-vers-une-nouvelle-loi-securitaire-menacant-la-vie-privee.html).
- [3] _____, *Des extensions Chrome envoient toute votre navigation à des tiers*, *Numerama* (23/11/2015), URL [\url{http://www.numerama.com/tech/132027-ces-extensions-chrome-qui-envoient-toute-votre-navigation.html}](http://www.numerama.com/tech/132027-ces-extensions-chrome-qui-envoient-toute-votre-navigation.html).
- [4] Zineb Dryef, « *je n'ai rien à cacher* » : le tracking sur mobile expliqué à mon père, *Rue89* (12/05/2015), URL [\url{http://rue89.nouvelobs.com/2015/05/12/nai-rien-a-cacher-tracking-mobile-explique-a-pere-259155}](http://rue89.nouvelobs.com/2015/05/12/nai-rien-a-cacher-tracking-mobile-explique-a-pere-259155).
- [5] Ifop, *Les réactions des français aux attentats du 13 novembre*, *Le Figaro et RTL* (17/11/2015), URL [\url{http://www.youscribe.com/catalogue/tous/actualite-et-debat-de-societe/politique/les-reactions-des-francais-aux-attentats-du-13-novembre-2675072}](http://www.youscribe.com/catalogue/tous/actualite-et-debat-de-societe/politique/les-reactions-des-francais-aux-attentats-du-13-novembre-2675072).
- [6] Romain Ligneul, « *Changez votre photo de profil* » : Facebook et le bleu-blanc-rouge en un clic, *Rue89* (17/11/2015), URL [\url{http://rue89.nouvelobs.com/2015/11/17/changez-photo-profil-facebook-bleu-blanc-rouge-clic-262146}](http://rue89.nouvelobs.com/2015/11/17/changez-photo-profil-facebook-bleu-blanc-rouge-clic-262146).
- [7] Robin Prudent, *Sur 99 pages a4, airbnb déballe tout ce qu'il sait de moi*, *Rue89* (07/10/2015), URL [\url{http://rue89.nouvelobs.com/2015/10/07/99-pages-a4-airbnb-deballe-tout-quil-sait-261533}](http://rue89.nouvelobs.com/2015/10/07/99-pages-a4-airbnb-deballe-tout-quil-sait-261533).
- [8] Fabrice Rochelandet, *Économie des données personnelles et de la vie privée*, La Découverte, Repères, Paris, 2010.
- [9] Radio Télévision Suisse, *Enquête ouverte : Mes données*, 2015, [En ligne, le 04-11-2015], URL [\url{https://www.facebook.com/groups/mesdonnees/}](https://www.facebook.com/groups/mesdonnees/).
- [10] Wikipedia, *Cnil- Wikipedia, the free encyclopedia*, 2015, [En ligne, le 04-11-2015], URL [\url{https://fr.wikipedia.org/wiki/Commission_nationale_de_l'informatique_et_des_libertés}](https://fr.wikipedia.org/wiki/Commission_nationale_de_l'informatique_et_des_libertés).
- [11] _____, *Données personnelles — Wikipedia, the free encyclopedia*, 2015, [En ligne, le 04-11-2015], URL [\url{https://fr.wikipedia.org/wiki/Donnees_personnelles}](https://fr.wikipedia.org/wiki/Donnees_personnelles).
- [12] _____, *Echelon — Wikipedia, the free encyclopedia*, 2015, [En ligne, le 04-11-2015], URL [\url{https://fr.wikipedia.org/wiki/Echelon}](https://fr.wikipedia.org/wiki/Echelon).

9. Sondage Ifop pour Figaro/RTL publié le 17 novembre : *Les réactions des Français aux attentats du 13 novembre*. Détails dans le sondage complet : [5].

- [13] ———, *World wide web — Wikipedia, the free encyclopedia*, 2015, [En ligne, le 04-11-2015], URL `\url{https://fr.wikipedia.org/wiki/World_Wide_Web}`.

INSTITUT FOURIER, 100 RUE DES MATHS, 38402 SAINT MARTIN D'HERES
E-mail address: `romain.vanel@ujf-grenoble.fr`