ECOLE ETHIQUE DU NUMERIQUE
RECHERCHE ET PROTECTION DES DONNEES

ARCACHON

29 septembre 2016

Sophie VULLIET-TAVERNIER, Directeur des relations avec les publics et la recherche

Protection des données personnelles et recherche: quels enjeux aujourd'hui?

- Comment concilier les besoins spécifiques de la recherche et la protection des personnes?
- Des gisements de données personnelles de plus en plus massifs: l'accès aux données et leur réutilisation
- Prendre en compte les droits des personnes et l'aspiration à une plus grande maitrise de ses données
- La recherche française dans le contexte international trouver un cadre de régulation adapté

La question de l'application des principes de protection des données personnelles à la recherche

- Accès aux données et protection des données personnelles sont-ils inconciliables?
- Malentendus, idées reçues et incompréhensions à lever.
- Points de clarification:
  - La frontière données personnelles-anonymat-pseudonymat
  - Expliquer la démarche d'analyse I et L
  - La prise en compte des spécificités de la recherche

PLAN

- Enjeux, contexte et cadre de régulation
- La CNIL: missions, perspectives d'évolution, la CNIL et la recherche
- Notions clés et grille d'analyse de la protection des données personnelles, la prise en compte des spécificités recherche
- Règlement européen: les grandes lignes

Enjeux, contexte et cadre de regulation

QUELS ENJEUX, QUELS DEFIS ?

- Puissance des moyens de stockage et des capacités de calcul, multiplication des capteurs (big data) ; baisse des coûts
- Personnalisation des services, dématérialisation des démarches administratives …

- Nouvelles technologies, nouveaux usages, nouveaux comportements sociaux…
- La mondialisation: cloud, GAFA, Privacy shield

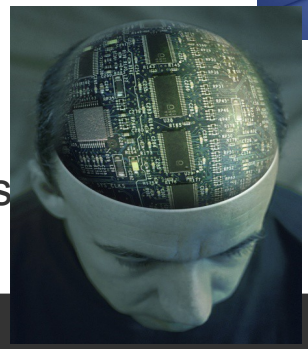
QUELS IMPACTS SUR LES LIBERTES ET LA VIE PRIVEE ?

- Traçage et surveillance, profilage et discrimination

- Des risques accrus : divulgation, failles de sécurité, utilisation détournée, usurpations d'identité…

- Commercialisation des données personnelles

- Vie privée - vie publique : où est la frontière?



- Comment assurer la maîtrise de ses données

QUELLES REPONSES ?

- **Autorégulation - déontologie - normes techniques - codes de conduite ?**
- **Lois sectorielles ?**
- **Une convention internationale ?**
  - Une voie possible : la convention n° 108 du Conseil de l'Europe– 28 janvier 1981(en cours de révision)
- **La régulation par la loi : le droit à la protection des données comme droit fondamental**
  - 80 Etats dotés de lois de protection des données –
  - L'approche de l'Europe : Charte européenne des droits fondamentaux, directive de 1995 et règlement européen 2016

LA PROTECTION DES DONNEES PERSONNELLES : une éthique du numérique appliquée aux données personnelles
- les 4 piliers-

- Reconnaître à toute personne des droits sur ses données

- Des règles pour encadrer les traitements de données personnelles

- Un régime de sanctions en cas de non respect des principes

- MISSIONS, PERSPECTIVES D'EVOLUTION, CNIL ET RECHERCHE

CNIL

LA CNIL EN BREF

- **Une autorité administrative indépendante**
  - 17 membres + le défenseur des droits
  - Services: 190 personnes
  - Budget 2016: 18 millions d'euros
- **Une triple mission**
  - **Contrôle** : déclarations et contrôles sur place et en ligne
  - **Sanction**
  - Information, **conseil**

- **+ de 90 000 déclarations/an;**
- **Correspondants informatique et libertés: +17000 organismes**
- **guides pratiques, tutoriels video**
- **6000 plaintes/an**
- **+ de 500 contrôles/an**
- **93 mises en demeure; 10 sanctions dont 3 financières en 2015**



QUI COMPOSE LA COMMISSION ?

La commission est composée de

17 membres

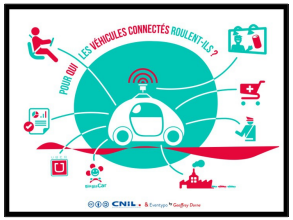
- **6** représentants des hautes juridictions
- **5** personnalités qualifiées
- **4** Parlementaires
- **2** membres du Conseil économique, social et environnemental

SÉNAT — ASSEMBLÉE NATIONALE — CONSEIL D'ÉTAT — COUR DE CASSATION

- 12 des 17 membres sont élus ou désignés par les assemblées ou les juridictions auxquelles ils appartiennent.
- La CNIL élit son Président parmi ses membres ; elle ne reçoit d'instruction d'aucune autorité.
- Les ministres, autorités publiques, dirigeants d'entreprises, publiques ou privées, ne peuvent s'opposer à son action.
- Le Président de la CNIL recrute librement ses collaborateurs.

La CNIL évolue…

- Mieux répondre aux besoins des usagers: **service besoin d'aide**, FAQ, refonte du site cnil.fr,
- Faire de **la pédagogie, sensibiliser, former**… (www.educnum.fr, tutoriels, …)
- Accompagner la **conformité**: labels (+ de 50), packs, et **PIA…Simplifier les formalités**
- Avoir une **politique de contrôles et de sanctions + ciblée**
- **Accompagner l'innovation et la recherche** (conseils, partenariats, chaire de recherche, ANR, prix…)
- Développer en concertation, la **réflexion prospective et éthique**: comité de la prospective, open data, vie privée 2020, santé connectée, voiture connectée… **cahiers IP et labo de la CNIL** : https://www.cnil.fr/fr/innovation-prospective/publications
- **http://linc.cnil.fr/**



La CNIL et la recherche

- **Partenariats**: CPU – CGE – INRIA - Institut Mines-Telecoms

- Membre de la Commission de réflexion sur l'ethique de la recherche en sciences du numérique (Cerna) de l'alliance Allistene

- Membre du comité du secret statistique du CNIS

- Actions de **sensibilisation** auprès des milieux universitaires et de la recherche

- Participation à des **travaux de recherche**: partenaire de **la chaire de recherche** de l'institut mines-telecoms VPIP sur valeurs et politiques des informations personnelles - projet **mobilitics** sur les smartphones avec Inria…appel à projets dans le cadre des PIA,

- Accompagnement de projets de recherche (big data, web social…)

- Participation à certains comités de l'ANR

- Projet de guide pratique, vade mecum pour les chercheurs

La CNIL, des missions élargies

- mission éthique du numérique
- Certification de méthodologies d'anonymisation

- Publicité de tous les avis de la CNIL

- Sanctions renforcées

NOTIONS CLES ET GRILLE D'ANALYSE DE LA PROTECTION DES DONNEES PERSONNELLES

LES NOTIONS CLES : LA DONNEE A CARACTERE PERSONNEL

DEFINITION

- Toute information concernant une personne physique identifiée ou identifiable, directement ou

EXEMPLES

Des risques de ré-identification accrus

La capacité de reconstruire le lien entre une donnée et une personne augmente

- avec le nombre et de la qualité des sources d'information disponibles
- les avancées mathématiques
- la puissance de calcul

- Massachusetts (2001) : croisement d'une base de données médicale pseudonymisée et une liste électorale (données nominatives). Le croisement a été effectué sur 3 valeurs : code postal, date de naissance et sexe. Latanya Sweeney est parvenues à réidentifier des individus (le gouverneur du Massachusetts lui même).

- AOL (2006) AOL a diffusé 20 millions de mots-clés figurant dans les recherches effectuées par plus de 650 000 utilisateurs sur 3 mois. L'adresse IP avait été remplacée par un numéro. Cet historique de recherche d'un individu fournissait des informations (âge, profession, goûts ou préférences)

- Etude MIT : une base de données d'horodatage des antennes-relais (GSM) considérée *a priori* comme anonyme. L'étude a montré qu'il suffisait de connaître quatre points de localisation d'une personne pour identifier le numéro d'un individu avec une très forte probabilité (90%). (2014 : relevés d'achats par carte bancaire)

Anonymat- pseudonymat

- **Définition**
  - un jeu de données est anonyme si on en a retiré tous les éléments permettant d'identifier la personne ;
  - et s'il est démontré qu'il est impossible de ré-identifier des personnes.
- **L'appréciation de l'anonymisation a évolué au fil du temps:**
  - La doctrine de la CNIL s'est élaborée de manière progressive : loi de 2004 ("bref délai"), évolution des technologies et des potentiels de ré-identification
  - Approche au cas par cas – libre choix du procédé d'anonymisation à utiliser .
  - L'avis du G29 du 10 avril 2014

Donnée personnelle- pseudonymat: directive, loi française, réglement

- La directive:

• Pour déterminer si une personne est identifiable, il convient de considérer l'ensemble des moyens susceptibles d'être **raisonnablement** mis en œuvre, soit par le responsable du traitement, soit par une autre personne pour identifier ladite personne.

- L'article 2 de la LIL retient une définition plus extrensive:

• Pour determiner 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.

- Réglement: considerant 23 et definition du pseudonymat (art 4):Le traitement de données personnelles de telle façon qu'elles ne puissent plus être attribuées à une personne concernée précise sans avoir recours à des informations supplémentaires, pour autant que celles-ci soient conservées séparemment et soumises à des mesures techniques et organisationnelles afin de garantir cette non-attribution à une personne identifiée ou identifiable.

LE TRAITEMENT : UN CHAMP TRES LARGE

- Bases de données, dispositifs biométriques, réseaux sociaux, applications mobiles, RFID…

2. La grille d'analyse de la protection des données: une démarche éthique

Finalité

Les données sont recueillies et traitées pour un usage déterminé et légitime, préalablement défini
- la spécificité des finalités de recherche

Proportionnalité et pertinence

Les informations recueillies et traitées doivent être pertinentes et nécessaires au regard des objectifs poursuivis
Protection particulière pour certaines catégories de données (données dites sensibles, N° de sécu, biometrie…)

Durée de conservation

Une durée de conservation doit être définie en fonction de la finalité
Au-delà , archivage, effacement, anonymisation sauf recherche ultérieure

Sécurité et confidentialité

Des mesures de sécurité doivent être prises pour garantir la confidentialité et l'intégrité des données
Les guides pratiques de sécurité

Respect des droits des personnes

Les personnes concernées doivent être informées et disposent d'un droit d'accès, de rectification, de suppression et d'opposition/consentement sur leurs données, droit de connaitre et de contester les raisonnements utilisés
Les dérogations possibles – la question du consentement

GRILLE D'ANALYSE DE LA PROTECTION DES DONNEES

1. Finalité du traitement :

Principe: les données doivent être collectées pour des finalités déterminées, explicites et légitimes et ne doivent pas être traitées ultérieurement de manière incompatible avec ces finalités.

GRILLE D'ANALYSE DE LA PROTECTION DES DONNEES

2. Pertinence des données :

Principe : les données doivent être adéquates, pertinentes et non excessives

Protection particulière pour les données sensibles :

origines raciales ou ethniques, opinions politiques, état de santé… ►interdiction de collecte ou de traitement sauf exceptions (intérêt public, consentement exprès…)

Justifier

GRILLE D'ANALYSE DE LA PROTECTION DES DONNEES

3. Conservation limitée des données :

Principe : les données sont conservées en base active pendant une durée limitée, définie en amont

Archivage, effacement, conservation à des fins de recherche, statistiques

A justifier

GRILLE D'ANALYSE DE LA PROTECTION DES DONNEES

4. Obligation de sécurité et de confidentialité :

Principe : garantir la confidentialité et l'intégrité des données.

https://www.cnil.fr/sites/default/files/typo/document/Guide_securite-VD.pdf

https://www.cnil.fr/sites/default/files/typo/document/CNIL-Guide_Securite_avance_M

Anonymisation

chiffrement

GRILLE D'ANALYSE DE LA PROTECTION DES DONNEES

5. Respect des droits des personnes :

- **Droit d'information :** mentions sur les questionnaires d'enquete: https ://www.cnil.fr/fr/modeles/mention
- privacy by design (linc)
- **Droit d'opposition/ consentement**
- **Droit d'accès et de rectification**
- **Droit de connaitre et de contester les raisonnements utilisés**
- 

VIDÉO

VOUS AVEZ DES DROITS SUR VOS DONNÉES PERSONNELLES : PROFITEZ-EN !

erdues de vue…

Déclarations, exonérations et autorisations

- **Les applications et fichiers hors du champ de la loi/exonérés de déclaration par la loi ou par la CNIL**
  fichiers de personnes morales, activités exclusivement personnelles (ex. agendas)
  fichiers des membres des partis politiques, syndicats et églises

- *fichiers de paie , fichiers de fournisseurs, information et communication externes, …*

- **Les applications courantes soumises à déclaration, sauf CIL**

- *gestion RH, fichiers clients, **recherches hors données sensibles**…*

- **Les applications à risque soumises à autorisation ou avis de la CNIL**

- *traitements de données sensibles (y compris à des fins de **recherche**), dispositifs biometriques, fichiers de police, téléservices, interconnexions, flux transfrontières sous certaines conditions*

Le correspondant Informatique et Libertés

- Une désignation **facultative**, notifiée à la CNIL
- **dispense des déclarations courantes** mais tenue par le CIL d'un registre
- Assure une meilleure application de la loi, contacts privilégiés avec la CNIL (ateliers de formation, service dédié,…)
- CIL INRIA INRA INSERM CNRS…
- Reseau supcil https://groupes.renater.fr/wiki/supcil/

Les autres procédures d'accès aux données pour la recherche

- L'accès libre aux données agrégées anonymes (ex via le site portail du réseau Quételet, le site INSEE…)

- L'accès aux données individuelles de la statistique publique (enquêtes et recensement) accessibles sous certaines conditions en fonction de leur degré de granularité:

  - via le CNIS (comité du label – avis du comité du secret statistique)

  - et via le CASD= centre d'accès sécurisé à distance aux données (sans récupération de celles-ci), géré par le GENES pour certains jeux de données

- L'accès aux données pour les sciences humaines et sociales: le guichet unique du GIS réseau Quetelet (centre Maurice Halbwachs- CNRS- plusieurs universités, INED, CASD), membre du CESSDA réseau européen

RECHERCHE MEDICALE ET LOI I ET L: UN REGIME DEROGATOIRE

- L'exception recherche dans la loi: finalité compatible, possibilités de dérogation à l'obligation d'information, de traitement des données sensibles
- Une prise en compte des spécificités de la recherche en santé dès les années 80
- Des formalités particulières: avis CCTIRS+ autorisation CNIL
- Un nombre toujours croissant de dossiers (800 autorisations recherche en santé en 2015) et des délais à maîtriser

LES EVOLUTIONS EN COURS

- **La loi de modernisation de notre système de santé (art 193):de nouvelles modalités pour l'accès et la réutilisation des données de santé à des fins de recherche et d'évaluation:**
  - Modification de la loi I et L pour les traitements de données à des fins de recherche, étude et d'évaluation en santé: **réorganisation du circuit des demandes d'autorisation CNIL:**
  - INDS: guichet unique + avis sur intérêt public de la recherche+ recommandations:
  - Avis, selon les types de recherche, des CPP ou du CEREES
  - Possibilité de procédures simplifiées et déclaration pour les traitements de données de santé en cas d'alerte sanitaire

REGLEMENT EUROPEEN

Le règlement européen: les objectifs (entrée en vigueur mai 2018)

- Renforcer les droits des personnes pour développer la confiance et contribuer à l'essor de l'économie numérique

- Assurer une plus grande harmonisation des règles de protection des données tout en renforçant la responsabilité des entreprises

- Renforcer le rôle des autorités de protection des données (APD) et du groupe européen des APD, le G29

Des principes précisés

Renforcement global des droits

Le renforcement des droits existants

- obligation générale de faciliter l'exercice des droits (fourniture d'une information claire, intelligible et aisément accessible)
- information renforcée (ex. transferts hors de l'UE)
- droit d'accès précisé (ex. : possibilité d'introduire une réclamation devant une « CNIL »)
- droit de rectification maintenu
- droit à l'effacement et à l'oubli numérique confirmé
- clarification de l'expression du consentement

Les nouveaux droits

- la portabilité des données
- la limitation du traitement
- conditions particulières pour le traitement des données des mineurs

Moins de formalités, plus de responsabilisation

Les responsables de traitements (et sous-traitants) ont une obligation générale de mettre en place des mesures appropriées et de démontrer cette conformité à tout moment : **c'est l'** ***accountability***.

- l'application des principes de ***privacy by design*** et ***privacy by default***
- la conduite **d'analyses d'impact**, ou « DPIA »
- la tenue d'un **registre** des traitements mis en œuvre
- la **notification de failles** de sécurité (aux autorités et personnes concernées)
- la **consultation de la CNIL** pour certains traitements présentant des risques élevés
- la **certification** de traitements et l'adhésion à des **codes de conduite**

Prise en compte des spécificités de la recherche et des statistiques dans le reglement

- Le règlement : une définition du champ et des dispositions particulières (art 89)
- Réutilisation des données à des fins de recherche: finalité licite et compatible, …
- Possibilité de traiter des données sensibles moyennant des garanties légales
- Possibilités de dérogations en matière d'information, d'exercice des droits des personnes, de durée de conservation…
- Garanties à prendre pour les droits et libertés des

En conclusion

- Harmonisation nécessaire entre droit national et droit européen
- Au-delà des formalités, la nécessaire prise en compte de la démarche « ethique » I et L
- L'importance des travaux de recherche sur le chiffrement, les méthodes d'anonymat et pseudonymat; le développement des dispositifs d'accès sécurisé aux données type casd…vers une recherche plus interdisciplinaire?
- Transparence, participation citoyenne et retours vers les personnes
- Les enjeux majeurs: big data, cloud, objets connectés, IA, modèles économiques du numérique, GAFA, …

Questions/réponses.

Merci de votre attention.