DANE INFO INFO N°7

Novembre 2015

LES DONNÉES À CARACTÈRE PERSONNEL

Des principes aux problèmes actuels

Quels usages de services collaboratifs sur l'internet ?

L'actualité est marquée par les problèmes complexes posés par les usages des réseaux sociaux et ces difficultés ont aussi des conséquences dans le cadre scolaire. Outre ceux qui sont utilisés dans le monde associatif ou universitaire, il existe aujourd'hui deux grands types de services collaboratifs disponibles via internet et susceptibles d'être utilisés dans un établissement scolaire, mais leurs finalités et les moyens de leur mise en œuvre sont clairement opposés :

– les réseaux sociaux proposés par les grandes entreprises multinationales (dont les GAFA : Google, Apple, Facebook, Amazon) dont le modèle économique est l'exploitation commerciale des données récoltées ;
– les E.N.T. mis en place par les collectivités et l'Éducation nationale, encadrés par un schéma directeur (SDET) assurant un espace de confiance conforme aux normes de la loi Informatique et libertés.

L'impératif de sécurité des données confiées à l'Éducation nationale impose que les enseignants utilisent avec leurs élèves l'E.N.T., partie intégrante du service public du numérique. Mais les usages de l'internet par les élèves dépassent ce cadre protégé, et l'éducation à la responsabilité que nous devons leur apporter ne saurait non plus s'y limiter. Le défi, difficile, que nous devons relever, est d'aider nos élèves à construire leur citoyenneté numérique pour que la liberté sur le web reste une réalité.

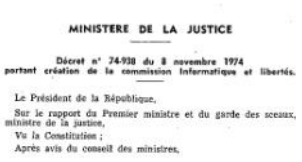
Nous espérons que ce document apporte aux professeurs et aux personnels de direction, responsables des traitements de données dans leur établissement, des éléments de réflexion pour œuvrer dans ce sens.

Pascal Faure, délégué académique au numérique

Un peu d'histoire

Le rapport Tricot, à l'origine de la loi de 1978

La mise en place du numéro « de sécurité sociale » ou Numéro d'Inscription sur le Registre des Personnes Physiques et le développement de l'informatique administrative ont fait naître au début des années 1970 un projet de fichier regroupant l'ensemble des informations disponibles sur chaque Français : SAFARI ou Système Automatisé pour les Fichiers Administratifs et le Répertoire des Individus. Un article du journal *Le Monde* du 21 mars 1974 (*SAFARI ou la chasse aux Français*) donne l'alerte à ce sujet et le débat public mène à la création de la commission informatique et libertés le 6 novembre 1974 puis à l'adoption de la loi *Informatique et libertés* en janvier 1978.


MINISTERE DE LA JUSTICE

Décret n° 74-938 du 8 novembre 1974 portant création de la commission informatique et libertés.

Le Président de la République,
Sur le rapport du Premier ministre et du garde des sceaux, ministre de la justice,
Vu la Constitution ;
Après avis du conseil des ministres,

Cette loi instaure un certain nombre d'obligations et d'interdictions relatives aux traitements informatiques de données à caractère personnel. Il s'agit essentiellement de protéger les individus d'éventuels abus de pouvoir de l'État ou d'entreprises.

Plus de détails sur le site des Correspondants informatique et liberté du C.N.R.S : http://www.cil.cnrs.fr/CIL/spip.php?article1871
Et sur le site de l'Institut National de l'Audiovisuel : http://www.ina.fr/video/CAB7600764601/informatique-et-liberte-rossi-video.html

SOMMAIRE

académie Nancy-Metz

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

Quelques principes de la loi Informatique et libertés

Qu'est-ce qu'une donnée à caractère personnel ?

C'est toute donnée relative à une personne identifiée ou pouvant l'être, directement ou indirectement, éventuellement par recoupement avec d'autres données.

Exemple : nom, prénom, date de naissance, immatriculation automobile, numéro de téléphone, adresse I.P., photographie, etc.

Certaines données ont un caractère plus sensible que d'autres, par exemple : « Il est interdit de collecter ou de traiter des données à caractère personnel qui font apparaître, directement ou indirectement, les origines raciales ou ethniques, les opinions politiques, philosophiques ou religieuses ou l'appartenance syndicale des personnes, ou qui sont relatives à la santé ou à la vie sexuelle de celles-ci. » (article 8, §1 de la loi du 6 janvier 1978). Ou encore celles qui font état de condamnations pénales, qui contiennent des appréciations sur les difficultés sociales des personnes ou qui utilise le « numéro de sécurité sociale », et dont les traitements sont soumis à des conditions plus restrictives que les autres données.

A quelles conditions un traitement de données est-il légitime ?

Rappelons les cinq principes fondamentaux, qui figurent sur les récépissés de déclaration délivrés par la CNIL.

1. Les données ne peuvent être recueillies et traitées que pour une finalité déterminée. Une utilisation ultérieure pour une finalité incompatible avec la finalité première, est illégale.

Exemple : on ne créera pas de comptes sur un outil de communication destiné au grand public ou sur un réseau social au moyen de données confiées aux établissement pour la gestion de la scolarité sans autorisation écrite du responsable de l'élève (voir l'article 5 de la délibération n°2012-184 du 7 juin 2012 : http://www.cnil.fr/documentation/deliberations/deliberation/delib/268/).

2. Seules doivent être enregistrées les informations pertinentes et nécessaires, non-excessives pour la finalité considérée.

Exemple : on fera particulièrement attention aux zones où il est possible de faire figurer des commentaires libres et on proscrira les commentaires excessifs, subjectifs, inappropriés, voire insultants, ou portant atteinte à la vie privée.

Une mention dans un logiciel de vie scolaire :

* **CNIL** : *Le motif enregistré pour l'absence, le retard ou la dispense ne doit pas consigner d'élément détaillé sur la vie privée des personnes*

3. Les données ne sont conservées que pour une durée limitée en fonction de la finalité de chaque fichier.

Exemple : il faut supprimer les comptes des usagers de l'ENT trois mois après leur départ de l'établissement (arrêté du 30 novembre 2006 sur le traitement ENT du 27 avril 2006, article 6), supprimer les données de prêt quatre mois après la fin du prêt dans un logiciel de gestion documentaire (article 4 de la norme simplifiée n°9 sur la gestion des prêts de livre).

4. Le principe de sécurité et de confidentialité : les données contenues dans les fichiers ne peuvent être consultées que par les personnes habilitées à y accéder en raison de leurs fonctions. Le responsable du traitement doit prendre toutes mesures pour empêcher que les données soient déformées, endommagées ou que des tiers non autorisés y aient accès.

5. Une exigence donne enfin sens à celles qui précèdent : le principe de loyauté, ou de transparence vis-à-vis des personnes dont les données sont collectées. Elles doivent être informées, en particulier :

- de la finalité du traitement ;
- des destinataires des données ;
- des modalités d'exercice de leur droit d'accès, de rectification et d'opposition aux traitements.

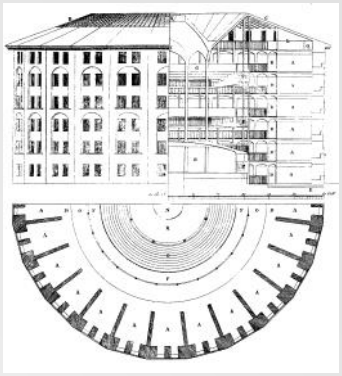
Pour d'autres précisions :

Le point sur les obligations déclaratives des établissements : http://www.cnil.fr/documentation/fiches-pratiques/fiche/article/traitements-de-gestion-scolaire-quelles-formalites-cnil-pour-les-chefs-detablissements/

Le guide Informatique et libertés pour l'enseignement du second degré : http://www.cnil.fr/fileadmin/documents/Guides_pratiques/CNIL_Guide_enseignement.pdf

Les nouvelles menaces

Un changement de problématique majeur



Le Panoptique de Bentham, ou comment surveiller tous les détenus d'une prison depuis un lieu central (gravure de 1791)

Le problème principal de la législation sur l'informatique et les libertés était donc de protéger le citoyen contre un État trop gourmand en données, jugé laxiste en ce qui concerne leur durée de conservation et n'hésitant pas à les interconnecter et à les centraliser : en quelque sorte, une version informatique du *Panoptique* de Bentham, où l'État exerce la surveillance permanente des citoyens depuis un lieu central d'où tout est visible en permanence.
L'évolution récente de l'informatique destinée au grand public fait apparaître un paradoxe, au moins apparent, qui porte sur la vie privée : le citoyen est à juste titre sourcilleux en ce qui concerne l'accès de l'Etat à certaines données, alors qu'il les expose sur des réseaux sociaux. Mais ce qu'il expose volontairement n'est qu'une partie des données qu'il produit.

Quelles sont les manifestations et les conséquences de ces évolutions ?

Un changement d'échelle : la masse de données exploitables produites par chaque individu connait une croissance exponentielle. Ces données sont produite de manière plus ou moins volontaire et consciente : publications sur un réseau social, courriers électroniques, photographies partagées, mais aussi : historique de navigation (ou ensemble des adresses visitées), historique des requêtes dans un moteur de recherche, achats effectués, données produites automatiquement par les objets connectés (géolocalisation par les téléphones par exemple et donc conservation des déplacements), etc.
Ainsi, une recherche sur un produit, une réservation d'hôtel ou de voyage, vont provoquer l'apparition sur les pages visitées par l'utilisateur de publicités ciblées vers ses centres d'intérêt. Vous visitez des sites en langues étrangères et vous vous voyez proposer des cours de langues : marketing ciblé ou « reciblage » publicitaire.
L'ensemble de ces données, objets d'échanges commerciaux entre différents opérateurs, constitue aujourd'hui un enjeu économique de première importance. Choisir d'utiliser un service gratuit, c'est accepter qu'il soit financé par le commerce de données et par de la publicité ciblée. En résumé :

Si c'est gratuit, c'est vous le produit : https://www.youtube.com/watch?v=8vLSf1i4E7A

Une mutation technique et un enjeu scientifique : volume, vitesse, variété

L'avènement de ce *Big Data* (données massives ou mégadonnées) constitue également un enjeu technique et scientifique majeur :

- les données ne sont pas structurées ou ont des structures très hétérogènes, au contraire des bases de données traditionnelles ;
- leur volume est beaucoup plus important ;
- leur traitement, pour être efficace d'un point de vue commercial, doit être rapide.

On comprend dés lors l'intérêt pour les géants de l'Internet de recruter des mathématiciens de haut niveau, notamment français, rompus au maniement de l'algorithmique.

Un contrôle de plus en plus difficile

La masse et l'automatisation de la production, de l'enregistrement et du stockage des données rendent très difficiles le contrôle par les utilisateurs. A cela s'ajoute le caractère international des échanges de données : comment déterminer où est stockée telle ou telle donnée ? En fonction de cette localisation et de l'hétérogénéité des législations en matière de données, comment exercer le droit d'accès, de rectification ou de suppression reconnu par la loi française ? Quelle est la durée de conservation de ces données ? Comment s'assurer que le stockage n'en est pas indéfini ?

Des applications plus ou moins inquiétantes

L'exploitation des mégadonnées permet d'ores et déjà des applications multiples :

- en matière d'urbanisme, par l'analyse des déplacements, par exemple ;
- de santé publique, par la détection, plus ou moins efficace, d'épidémie en train de se déclencher.

Certaines de ces application sont susceptibles de toucher directement les individus : par exemple, pourquoi ne pas appliquer des tarifs d'assurance plus élevés à des individus :

- dont les montres connectées révéleraient qu'ils ne font pas ou pas assez de sport,
- dont les historiques de connexions montreraient qu'ils sont dans des situations à risque (parce qu'ils habitent ou fréquentent des zones dangereuses, qu'ils sont allés sur des forums consacrés à la violence conjugale, etc.).

D'autre part, 45% des recruteurs utilisent les résultats des moteurs de recherches pour des opérations d'embauche et 35% reconnaissent avoir refusé une embauche en raison de données trouvées par ce moyen. L'étape suivante consisterait à confier le recrutement à des algorithmes exploitant les traces laissées sur le réseau par les intéressés eux-mêmes et qui y demeurent pendant un temps indéfini. Si de telles pratiques peuvent difficilement passer pour loyales, comment établir un lien de causalité certain entre tel refus d'embauche et telle donnée accessible en ligne ?

Que faire ?

En conclusion, le refus d'une chasse au citoyen par l'Etat, problème originel de la loi *Informatique et liberté*, doit s'accompagner :

- *d'une éducation à la prudence* : la difficulté, voire l'impossibilité, de faire disparaître un contenu publié sur Internet doit inciter à la prudence lorsque l'on s'apprête à divulguer une donnée personnelle, même si c'est à destination de personnes considérées, temporairement, comme des proches.
- *d'une éducation à la liberté* : en même temps, la prudence ne saurait dégénérer en renoncement à la liberté d'expression. Un dernier aspect des réseaux sociaux est le développement de la surveillance latérale ou surveillance par les pairs. Cette surveillance débouche-t-elle sur l'autocensure ? Dans quelle mesure les réseaux sociaux influent-ils sur le mécanisme de la « spirale du silence » (http://www.pewinternet.org/2014/08/26/social-media-and-the-spiral-of-silence/) ?
- *d'une éducation à la discrétion* : apprendre à ne pas trop montrer devrait s'accompagner d'un autre apprentissage, tout aussi important : apprendre à ne pas chercher, à ne pas regarder, à ne pas espionner, à ne pas divulguer ce qui est du ressort de la vie privée d'autrui. Une information privée n'est pas nécessairement secrète. Qu'une information soit accessible ne rend pas légitime le fait de la recueillir à n'importe quelle fin. De même, on déconseillera à des professeurs de rentrer en relation avec des élèves ou des étudiants sur des réseaux sociaux destinés au grand public, pour ne pas être tenté d'empiéter sur la vie privée des élèves et pour leur donner l'exemple d'une saine distinction entre vie scolaire, vie professionnelle et vie privée.
- *d'une vigilance vis-à-vis de la commercialisation des traces et des données*, même devenues anonymes, laissées sur Internet et de ce qu'elles apportent aux pouvoirs susceptibles de les utiliser.

Bibliographie et sitographie sommaire

L'éducation au numérique par la CNIL : http://www.educnum.fr/
Le portail Internet responsable : http://eduscol.education.fr/internet-responsable/
Cahiers Innovation et prospective, n°1 : http://www.cnil.fr/fileadmin/documents/La_CNIL/publications/DEIP/CNIL-CAHIERS_IPn1.pdf