

Protection des données personnelles et recherche : constats et perspectives d'évolution



Sophie VULLIET-TAVERNIER

Directrice des relations avec les publics et la recherche -
Commission Nationale de l'Informatique et des Libertés (Cnil)

La recherche moderne a souvent besoin d'utiliser des données personnelles dont la confidentialité est protégée par la loi. Cette contradiction peut être surmontée. La Cnil encourage les dispositifs qui permettent aux chercheurs de travailler tout en fournissant les garanties indispensables.

Protection des données personnelles et recherche sont-elles inconciliables ? Non assurément. Mais l'application de la loi informatique et Libertés dans le monde de la recherche, c'est encore beaucoup de malentendus, d'idées reçues et d'incompréhensions à lever... Ce sont à l'évidence des notions à clarifier, une démarche à expliquer, une meilleure prise en compte des besoins spécifiques des chercheurs...

Ainsi, la frontière entre données personnelles, anonymat et pseudonymat n'est pas toujours perçue de façon claire, nonobstant les conseils et avis émis sur le sujet (cf. par exemple l'avis du groupe européen des autorités de protection des données mai 2014¹).

De même la démarche d'analyse « informatique et libertés » qui doit accompagner tout projet de recherche portant sur des données personnelles et qui s'apparente pourtant à une démarche éthique, n'est souvent pas ressentie comme telle, les chercheurs n'y voyant parfois qu'une procédure administrative de plus (il faut déclarer à la Cnil...).

Enfin, alors même que les spécificités de la recherche sont d'une certaine façon prises en compte par le corpus juridique de la protection des données, tant au niveau européen que national, leur caractère peu visible et complexe ne permet pas toujours leur réelle appréhension par les chercheurs et ce, quelle que soit la discipline.

Face à ces constats, quelles ont été et quelles sont les réponses apportées par la Commission Nationale de l'Informatique et des Libertés (Cnil) et le nouveau cadre de régulation qui se dessine avec l'entrée en application, en mai 2018 du règlement européen sur la protection des données personnelles, le RGDP ? Quels sont, aujourd'hui et demain, les nouveaux enjeux pour la protection des données dans le domaine de la recherche ?

La Cnil et la recherche

Depuis sa création, à la fin des années 70, la Cnil a beaucoup évolué...
Au-delà des chiffres d'activité (cf. encadré sur la Cnil en bref) et de l'exercice des missions

1. http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_fr.pdf

traditionnelles de contrôle a priori (via les formalités déclaratives) et a posteriori (via les contrôles et les sanctions), la Cnil s'est donnée aujourd'hui pour objectif de :

- mieux répondre aux besoins des usagers via notamment son service besoin d'aide, ses 500 FAQ et sa réponse en ligne, via aussi ses tutoriels et conseils pratiques disponibles sur ses sites www.cnil.fr et www.educnum.fr et ses fils Twitter et Facebook...
- accompagner la conformité au moyen des labels (+ de 90), des packs de conformité, du service dédié aux correspondants informatique et libertés (CIL)
- avoir une politique de contrôles et de sanctions plus ciblée en particulier en fonction des enjeux liés à la mondialisation du numérique et des sujets de préoccupation et d'inquiétudes remontant du public (spam, droit à l'oubli, vidéosurveillance...)
- accompagner l'innovation et la recherche.

La Cnil, la recherche et les statistiques : une longue histoire :

La Cnil a, très tôt, entretenu des relations avec le monde de la recherche, relations certes parfois tendues mais toujours passionnantes et riches d'enseignements.

Cette prise en compte des spécificités de la recherche s'est tout d'abord marquée dans l'organisation même de la Cnil.

Un membre de la Cnil est ainsi spécifiquement en charge des questions de recherche et des statistiques. Au sein des équipes de la Cnil, une direction² en lien avec d'autres services (et notamment le pôle innovation et prospective) assure l'interface avec les chercheurs qui sollicitent la Cnil. Par ailleurs, le service santé suit particulièrement les projets dans le domaine de la recherche médicale.

Enfin, il doit être relevé que la composition actuelle de la Cnil fait une assez large part au monde académique puisque 4 commissaires sur 17 en sont issus.

Sur le fond, la commission a, quand elle l'estimait nécessaire au regard des enjeux soulevés par telle ou telle problématique de recherche, créé en son sein ou avec les acteurs concernés des groupes de travail permettant ainsi une réelle concertation : ce fut le cas, par exemple, pour les registres épidémiologiques sur le cancer dans les années 80, la recherche sur le SIDA dans les années 90, les « statistiques ethniques » dans les années 2000, l'open data au début des années 2010...les enjeux éthiques des algorithmes et de l'IA, les objets connectés, les smart cities, ...

Sur d'autres grands dossiers « historiques », le NIR, le recensement, les données génétiques,... elle n'a pas hésité à faire appel à des expertises extérieures (experts en statistiques, en génétique, comité national d'éthique...) pour mûrir sa réflexion.

La Cnil et la recherche aujourd'hui :

Aujourd'hui, la Cnil et la recherche, ce sont des partenariats : avec l'Inria³, l'Institut Mines-Télécoms, la Conférence des Présidents d'Universités (et le réseau « Supcil » des correspondants informatique et libertés dans les universités⁴), la Conférence des Grandes Ecoles...

C'est aussi l'accompagnement de la recherche, et en particulier, le conseil sur les projets de recherche (Big Data, web social...), la participation au comité du secret statistique du CNIS, à certains comités de l'ANR... et la conception, en cours, d'un guide pratique, vade-mecum pour

2. La direction des relations avec les publics et la recherche.
3. Institut national de recherche en informatique et automatique
4. <https://groupes.renater.fr/wiki/supcil/>

les chercheurs, en liaison notamment avec le CNRS.

C'est enfin la promotion de la recherche sur la protection des données. Ainsi, la Cnil est partenaire de la chaire de recherche de l'institut mines-télécoms sur « Valeurs et politiques des informations personnelles⁵ ». Elle participe également à des travaux de recherche (exemple du projet mobilitics sur les smartphones avec l'Inria).

Et depuis deux ans elle organise avec l'Inria un Prix européen « privacy » visant à récompenser un article scientifique paru sur cette thématique.

La contribution de la Cnil à la réflexion éthique et prospective

Même si la Cnil a toujours intégré la dimension éthique dans ses réflexions, la loi pour la République numérique⁶ a élargi son champ d'intervention en lui confiant expressément la mission de conduire une réflexion sur les problèmes éthiques et les questions de société soulevés par l'évolution des technologies numériques.

Cette mission a été engagée avec le lancement début 2017 d'un débat sur les enjeux éthiques des algorithmes qui s'est récemment conclu par la publication, en décembre 2017, d'un rapport⁷ intitulé « comment permettre à l'homme de garder la main ? ».

La publication de cet ouvrage s'inscrit dans une politique volontariste de la Cnil visant à assurer une plus grande visibilité à ses réflexions : c'est ainsi qu'a été créé en 2017, en collaboration avec la documentation française une nouvelle collection « le point Cnil » dont le premier numéro est consacré aux données génétiques.

Par ailleurs, la Commission participe aux travaux conduits respectivement par la Commission de réflexion sur l'éthique de la recherche en sciences du numérique (Cerna) de l'alliance Allistene et par le comité d'éthique de Teralab (GENES – institut Mines Telecoms)⁸.

Enfin, sur un plan plus prospectif, a été créé, en 2012, un Comité de la prospective, réunissant membres de la Cnil et personnalités extérieures pour développer la réflexion prospective sur les nouveaux enjeux sociétaux à venir à l'aune des développements technologiques et des usages numériques.

Ce nouveau dimensionnement s'est également concrétisé par l'ouverture d'un laboratoire d'innovation numérique au sein de la Cnil, le labo linc⁹, d'un site dédié <http://linc.cnil.fr/> et des publications : les cahiers IP dont le dernier numéro est consacré aux smart cities¹⁰.

Les actions ainsi entreprises sont autant de signaux en direction de la recherche et qui doivent être poursuivies et amplifiées à l'aune de la nouvelle régulation européenne.

La prise en compte des spécificités de la recherche dans les lois de protection des données

La loi informatique et Libertés, reprenant en cela les dispositions de la directive européenne de 1995¹¹ reconnaît expressément la possibilité de traiter ultérieurement les données initialement

5. <https://cvpip.wp.imt.fr/accueil/>

6. Loi du 7 octobre 2016.

7. <https://www.cnil.fr/fr/comment-permettre-lhomme-de-garder-la-main-rapport-sur-les-enjeux-ethiques-des-algorithmes-et-de>

8. <https://www.teralab-datascience.fr/fr/>

9. <https://linc.cnil.fr/propos-de-linc>

10. <https://linc.cnil.fr/la-plateforme-dune-ville-explore-les-enjeux-de-la-smart-city>

11. Directive 95/46CE du 24 octobre 1995 remplacée par le RGPD

recueillies pour une finalité donnée (par ex de gestion) à des fins statistique ou de recherche scientifique ou historique (art 6). De même les données peuvent être conservées en vue d'être traitées à de telles fins (art 36). En outre, des dérogations à l'obligation d'information sont également prévues (art. 32, III). Enfin, le droit d'accès n'est pas applicable aux données qui sont conservées sous une forme excluant manifestement tout risque d'atteinte à la vie privée des personnes concernées et ce sous contrôle de la Cnil.

Le recueil et le traitement, à des fins de recherche ou de statistiques, des données dites sensibles (origine raciale ou ethnique, opinions politiques, religieuses, philosophiques, appartenance syndicale, données de santé, vie sexuelle) est également possible sous certaines conditions.

De façon plus spécifique, le besoin d'un accès plus large des chercheurs aux bases de données administratives a progressivement été pris en compte à la suite de plusieurs interventions du législateur :

- 1988 : accès aux données de la statistique publique
- 1994 : accès aux données de santé pour la recherche
- 2013 : accès aux données fiscales
- 2016 : accès aux données de santé (SNDS)
- 2016-2017 : accès aux autres données produites par les administrations

S'agissant des évolutions récentes, la loi de modernisation de notre système de santé (art 193) a introduit de nouvelles modalités pour l'accès et la réutilisation des données de santé à des fins de recherche et d'évaluation en modifiant notamment dans la loi Informatique et Libertés, le circuit des procédures de formalités applicables en instaurant notamment un guichet unique via l'Institut National des Données de Santé (INDS) et sans entrer dans les détails des procédures plus ou moins allégées selon les types de recherches¹².

La loi pour une République numérique a quant à elle défini les conditions d'utilisation du NIR chiffré à des fins de recherche et de statistiques en prévoyant également une simplification des procédures.

S'agissant des droits des mineurs en matière de recherche médicale, cette loi prévoit sous certaines conditions, que l'information d'un des parents peut suffire ; elle introduit aussi la possibilité pour les mineurs de plus de 15 ans de s'opposer à l'information des parents.

Le règlement européen et la recherche :

De façon générale, le règlement général sur la protection des données (RGPD) qui entre en application en mai 2018¹³ poursuit un triple objectif :

- 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 (et en réduisant fortement les formalités) ;
- développer le rôle des autorités de protection des données (APD) et du groupe européen des APD en renforçant leur pouvoir de sanction et en instaurant un mécanisme de coopération au plan européen.

S'agissant de la recherche, les dispositions générales contenues dans la directive et dans la loi sont largement reprises dans le règlement européen (article 89 notamment) mais le champ est élargi puisque le RGPD ne vise pas seulement les réutilisations (« traitements ultérieurs ») de données à des fins de recherche, statistiques mais aussi les traitements ayant pour finalité première la recherche.

12. Pour plus d'information, cf <https://www.cnil.fr/fr/recherche-medicale-queelles-formalites-pour-les-theses-et-les-memoires>

13. Pour plus d'information, cf <https://www.cnil.fr/fr/comprendre-le-reglement-europeen>

Le règlement précise le régime juridique applicable aux traitements de données réalisés à des fins archivistiques dans l'intérêt public, ce qui constitue une nouveauté par rapport à la directive, à des fins de recherche scientifique ou historique ou à des fins statistiques. Reconnaisant les spécificités de ces types de traitements, il prévoit ainsi la nécessité de garanties appropriées et introduit des possibilités de dérogations à certaines dispositions du règlement. Les garanties appropriées insistent sur le principe de minimisation des données et consistent notamment en l'adoption de mesures techniques et organisationnelles portant notamment sur des méthodes de pseudonymisation.

Plusieurs articles du règlement comportent également des dispositions spécifiques à la recherche ou aux statistiques. Il en est ainsi de la dérogation à l'obligation d'information, dès lors que cette dernière serait susceptible de rendre impossible ou de compromettre gravement la réalisation des objectifs dudit traitement (à cet égard, le considérant 62 apporte un éclairage : devraient être pris en considération le nombre de personnes concernées, l'ancienneté des données).

Un projet de loi modifiant la loi informatique et liberté¹⁴ sera discuté au Parlement au début de l'année 2018. Il permettra l'application effective du RGPD.

Un article est consacré aux traitements à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique et historique, ou à des fins statistiques (art 12). Il s'agit de préciser les marges de manœuvre relatives aux mesures concernant de tels traitements permises par l'article 89 du règlement, lequel autorise le droit national à prévoir des dérogations à certains droits des personnes concernées (droits d'accès, de rectification, d'opposition, à la limitation du traitement, portabilité), sous certaines conditions, dans la mesure où ces droits rendent impossible ou entravent sérieusement la réalisation des finalités spécifiques et où de telles dérogations sont nécessaires pour atteindre ces finalités.

Des dispositions particulières sont également prévues pour les traitements à des fins de recherche, d'étude ou d'évaluation dans le domaine de la santé.

En tout état de cause, à l'aune de ce nouveau cadre de régulation, il apparaît d'autant plus indispensable de concevoir et de diffuser, en concertation avec les milieux de la recherche, un guide de bonnes pratiques permettant d'expliquer au mieux les règles de protection des données applicables à ce secteur. Il importe aussi de développer des actions de formation et de sensibilisation adaptées tant auprès des délégués à la protection des données désignés au sein des organismes de recherche qu'auprès des chercheurs intéressés à se faire les relais de cette sensibilisation.

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

Au-delà, les gisements de données personnelles de plus en plus massifs qui se constituent via nos pratiques numériques (web social, objets connectés...) suscitent des convoitises aussi de la part des équipes de recherche désireuses d'y accéder à des fins, légitimes, de connaissance.

Bien que partagées et souvent publiées de façon ouverte, ces données n'en conservent pas le moins le caractère de données personnelles dont l'exploitation à des fins de recherche appelle des garanties adaptées. La Cnil y réfléchit. Dans le contexte international, il s'agit là d'enjeux majeurs pour la recherche française.

14. <https://www.cnil.fr/fr/la-cnil-publie-son-avis-sur-le-projet-de-loi-relatif-la-protection-des-donnees-personnelles>

De même il est essentiel de développer plus fortement les travaux de recherche sur le chiffrement, les méthodes d'anonymat et de pseudonymat ; les dispositifs d'accès sécurisé aux données à l'égal du centre d'accès sécurisé aux données (CASD), dispositif exemplaire à bien des égards.

Enfin, mieux garantir les droits des personnes et l'aspiration de chacun à une plus grande maîtrise de ses données constitue un impératif que la recherche devra sans doute plus prendre en compte à l'avenir pour assurer ainsi une réelle participation citoyenne.

La CNIL en bref

- **Une autorité administrative indépendante**

- 18 membres + le défenseur des droits
- Services: 195 personnes
- Budget 2017: 17 millions d'euros

- **Une quadruple mission**

- Contrôle : déclarations, contrôles sur place et en ligne
- Sanction
- Information, conseil, éducation
- Réflexion éthique et veille prospective

- + de 90 000 déclarations/an;
- 800 autorisations recherche santé
- Correspondants Informatique et libertés: +18000 organismes
- guides pratiques, tutoriels
- 7000 plaintes/an
- + de 400 contrôles/an



CNIL.