

La recherche internationale et la protection des données

Par Gaëlle BUJAN

Déléguée à la protection des données, CNRS

La recherche évolue à l'échelle mondiale dans ses fondements, ses valeurs et ses réseaux d'experts. À l'ère du numérique, l'accélération des échanges et les perspectives de recherche accompagnent une diffusion des savoirs, dont les limites sont sans cesse repoussées. Simultanément, les risques sont nouveaux et démultipliés ; ils sont associés à des impératifs d'intégrité, d'éthique et de respect des personnes. La réglementation s'adapte à ce mouvement : elle vise à responsabiliser les acteurs, à contribuer à la confiance de la société dans la recherche et devient de plus en plus protectrice des droits et des personnes.

Dans cet environnement évolutif et d'apparence contraignant, les organismes de recherche ont mis en place des politiques favorisant la culture de la protection des données et la réalisation de leurs objectifs à l'échelle internationale. Recherche internationale et réglementation relative à la protection des données ne sont ainsi pas incompatibles.

Tout établissement impliqué dans des activités de recherche a besoin quotidiennement de ressources humaines, financières et matérielles. Ces dernières recouvrent notamment les données de la recherche issues d'expériences, de calculs complexes... ; elles sont constituées de textes, de logiciels... Selon les disciplines scientifiques, les natures et les supports des données sont variables. Leur sensibilité peut s'entendre dans une logique de concurrence industrielle, être liée aux intérêts fondamentaux de l'État, à la vie privée et à des dispositions réglementaires.

Parmi ces données de recherche, une partie constitue des données à caractère personnel (c'est-à-dire toute donnée permettant d'identifier directement ou indirectement une personne⁽¹⁾).

Dans cet article, nous centrons notre propos sur ces seules catégories de données qui requièrent, pour pouvoir être utilisées et/ou réutilisées à des fins de recherche scientifique, des mesures techniques et organisationnelles adaptées à la protection de la vie privée des personnes.

Le Règlement européen sur la protection des données personnelles (RGPD) consacre de manière large le traitement de données personnelles à des fins de recherche scientifique⁽²⁾. L'article 89 du RGPD prévoit les « garanties et dérogations applicables aux traitements à des fins de recherche scientifique ou historique. »

Il s'agit donc de faciliter la recherche, pour laquelle la donnée personnelle constitue le matériau essentiel. D'ailleurs, la Commission nationale de l'informatique

et des libertés (CNIL) accompagne cette évolution en mettant à la disposition de la communauté scientifique un ensemble de guides et de fiches pratiques.

Au niveau du Centre national de la recherche scientifique (CNRS), dont la mission est « d'identifier, d'effectuer ou de faire effectuer, seul ou avec ses partenaires, toutes les recherches présentant un intérêt pour la science ainsi que pour le progrès technologique, social et culturel du pays »⁽³⁾, la protection des données personnelles est associée à l'éthique, à la déontologie et à l'intégrité scientifique, qui sont des valeurs fortes de l'établissement. Il porte une réelle responsabilité sociale lorsque des données lui sont confiées par des citoyens pour conduire sa mission originelle de recherche.

Dans son expertise et son accompagnement, le Centre prend en compte l'environnement évolutif de la recherche et la sensibilité des données. La politique de protection des données mise en œuvre par le CNRS concerne tous les traitements, bien que cela peut parfois s'avérer complexe au regard du caractère international de la recherche.

L'environnement évolutif de la recherche

Les progrès de la connaissance et les changements technologiques, économiques, sociaux, environnementaux et culturels contribuent à l'évolution constante des conditions et du contexte d'exercice de l'activité de recherche.

Ces différents facteurs forment des opportunités ou, au contraire, apparaissent comme des contraintes.

⁽¹⁾ Article 4 du RGPD (Règlement UE 2016/679 du Parlement européen et du Conseil, du 27 avril 2016).

⁽²⁾ Considérants 156, 157 et 159 du RGPD.

⁽³⁾ Décret n°82-993 modifié du 24 novembre 1982.

Le numérique

Les technologies de l'information et de la communication ont connu une accélération conséquente au cours des vingt dernières années : le développement des réseaux sociaux et la frontière devenue extrêmement ténue entre les sphères privée et publique modifient profondément les comportements individuels ou collectifs. L'accélération de la diffusion et de la circulation de l'information ouvre de nouvelles opportunités scientifiques.

Les progrès de la connaissance prennent appui sur de nouveaux objets de recherche, de nouvelles modalités d'accès aux données et des modalités d'expérimentation différentes. La situation sanitaire récente a transformé les possibilités de réalisation des expériences, lesquelles s'opèrent à domicile et non plus uniquement au sein des laboratoires. Les protocoles sont plus aisés à élaborer du point de vue sanitaire, mais sont plus intrusifs en termes de collecte des données personnelles.

Le volume des informations disponibles et accessibles, les capacités des grands calculateurs et des infrastructures de recherche, ainsi que les potentialités des algorithmes ouvrent des champs nouveaux à l'échelle mondiale dans les domaines de l'intelligence artificielle, du quantique...

Les données et les bases de données construites et exploitées pour des finalités de recherche sont de plus en plus des informations économiquement stratégiques et, de fait, qui s'inscrivent dans une compétition scientifique internationale accrue. Les risques de captation, de piratage, de potentielles réutilisations à des fins mercantiles sont réels et imposent la mise en œuvre de politiques de protection des données, voire des personnes. Les établissements publics déploient à cette fin la politique de protection des systèmes d'information (PSSI) élaborée par l'État. Comme dans les autres organismes de recherche, cette PSSI s'applique au CNRS, mais en tenant compte des spécificités de ses missions, notamment le nomadisme, la compétition internationale et la protection du patrimoine scientifique et technique. Pour l'ensemble de ses activités, la stratégie numérique de l'État est donc prise en compte.

L'ouverture de la science

La communauté scientifique est également confrontée au changement de paradigme induit par l'ouverture de la science. Cette nouvelle orientation vise à renforcer l'efficacité de la recherche, favoriser la transparence et la reproductibilité de la recherche et à l'intégrer dans la compétition internationale.

Impulsées par Frédérique Vidal, ministre de l'Enseignement supérieur, de la Recherche et de l'Innovation⁽⁴⁾, des initiatives conséquentes ont été prises pour accélérer l'évolution de la mise en place des plans d'action dédiés décidés au niveau du ministère précité, au CNRS (plans d'ouverture des données, science ouverte)...

⁽⁴⁾ <https://www.enseignementsup-recherche.gouv.fr/fr/le-plan-national-pour-la-science-ouverte-les-resultats-de-la-recherche-scientifique-ouverts-tous-49241>

Le double objectif d'ouverture et d'accessibilité est à intégrer dès le démarrage d'un projet et à introduire dans la préparation des plans de gestion de données et le cycle de vie des données notamment en vue de leur réutilisation.

Les évolutions réglementaires

Pour accompagner les évolutions dans l'organisation du travail et la plus grande circulation des biens et des services, la protection des biens et des personnes est renforcée. L'éthique, la déontologie et l'intégrité scientifique sont des valeurs fortes portées par les établissements en charge de la recherche et elles sont aujourd'hui inscrites dans les textes réglementaires.

La sensibilité des données

Les données et jeux de données présentent un intérêt économique réel si leur utilisation est détournée de leur finalité de recherche. Cette sensibilité est ici liée à leur valeur économique potentielle.

Des catégories particulières de données

Selon l'article 9 du RGPD, des traitements portant sur des catégories particulières de données personnelles sont possibles, mais sous certaines conditions.

Il s'agit des « données qui révèlent l'origine raciale ou ethnique, les opinions politiques, les convictions religieuses ou philosophiques ou l'appartenance syndicale, ainsi que le traitement de 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. »

Les données peuvent concerner le domaine de la santé : données relatives à la santé physique ou morale passée, présente ou future, qui donnent une indication sur l'état de santé de la personne. Leur traitement et les procédures dépendent de la finalité de la recherche ainsi que précisé dans le guide pour la recherche⁽⁵⁾ (Institut des sciences humaines et sociales du CNRS).

D'autres données font l'objet de dispositions spécifiques :

- le numéro de sécurité sociale est une donnée identifiante, dont l'utilisation est très encadrée ; elle est circonscrite à quelques finalités définies par la CNIL ;
- les données d'infraction ou liées à des condamnations relèvent pour leur exploitation de dispositions spécifiques prévues par la loi Informatique et libertés modifiée.

Des données hautement personnelles

Elles peuvent présenter une forte sensibilité pour la personne ; il peut s'agir de données liées aux revenus des ménages, au patrimoine individuel, aux activités professionnelles... Des dispositifs spécifiques et protecteurs de la vie privée permettent à la communauté scientifique d'utiliser ces données à des fins de

⁽⁵⁾ https://www.inshs.cnrs.fr/sites/institut_inshs/files/pdf/Guide_rgpd_2021.pdf

recherche. Par exemple, le Centre d'accès sécurisé aux données (CASD) permet d'accéder à distance à une infrastructure sécurisée assurant ainsi la confidentialité des données.

Des données concernant des personnes vulnérables

La notion de personne vulnérable n'est pas définie précisément par la réglementation. Elle peut concerner les mineurs, les personnes âgées ou des catégories particulières de population.

Pour leurs activités de recherche et en fonction de leur métier, les chercheurs, ingénieurs ou techniciens apportent une forte attention à la protection de leurs données : ce sont, par exemple, des personnes travaillant dans des animaleries, ayant accès à des équipements dangereux ou à risque, des spécialistes dans des domaines confidentiels, stratégiques ou liés à la défense des intérêts de la nation... Dans ces situations, une identification de la personne et des coordonnées de celle-ci deviennent des données sensibles à protéger (en ne mentionnant pas d'informations identifiantes dans les annuaires publics, notamment).

La politique de protection des données au CNRS

Le Centre national de la recherche scientifique⁽⁶⁾ est une institution de recherche parmi les plus importantes au monde ; elle œuvre dans tous les champs de la connaissance (voir la Figure 1 ci-dessous).

⁽⁶⁾ <https://images.cnrs.fr/video/7343>

L'implication du CNRS dans l'apport d'un haut niveau de sécurité en matière d'utilisation des données est renforcée depuis plus de dix ans. Au regard des risques encourus en termes de disponibilité, de fiabilité et d'intégrité des données et, par conséquent, de l'impossibilité éventuelle pour lui de conduire des projets de recherche, le CNRS a internalisé des fonctions d'analyse de la conformité des traitements de ses données à la loi Informatique et libertés.

Outre les aspects liés aux potentielles sanctions pénales et financières associées à l'absence d'application de la réglementation, les enjeux relèvent pour le CNRS de l'éthique et de l'intégrité scientifique. Pour le CNRS et la communauté scientifique, la protection de la donnée traduit leur souhait de conserver la confiance de la société dans l'activité scientifique et de crédibiliser leur action mise en œuvre essentiellement avec des ressources publiques.

Le CNRS a décidé de consacrer des ressources conséquentes à la protection des données et développe une culture de la protection de la donnée auprès de l'ensemble des communautés de l'établissement.

Son expertise est forte dans l'analyse des traitements à finalité de recherche scientifique, s'appuyant sur un réseau de compétences réparties dans les différentes directions du CNRS et ses unités de recherche. Des compétences dans les domaines juridique et de la sécurité des systèmes d'information sont aussi mobilisées.

Le choix fait d'une forte mobilisation pour sensibiliser et former à la protection des données personnelles a favorisé la prise en compte de cette dimension dans l'élaboration des projets et des protocoles de recherche.



Figure 1 : Les chiffres clés du CNRS – Source : Rapport d'activité 2021 du CNRS.

L'approche par les risques est retenue pour chaque traitement de données conduisant, selon les situations, à des analyses d'impact sur la vie privée. Au CNRS, ces moments de réflexion sont menés de manière collaborative entre les chercheurs, les juristes et les spécialistes de la sécurité des systèmes d'information et se fondent sur la méthodologie proposée par l'autorité de contrôle qu'est la CNIL.

La recherche internationale et la protection des données

Une caractéristique intrinsèque de la recherche est qu'elle est, par essence, internationale. La science se nourrit des avancées faites par une communauté qui travaille en réseau, par-delà les frontières.

Chaque chercheur est confronté à des réglementations qui diffèrent selon les pays : protection des données, du patrimoine scientifique, accords internationaux, principes éthiques de la recherche, protocoles spécifiques aux disciplines, mesures de sécurité dont la sécurité des systèmes d'information.

Le RGPD est très protecteur de la vie privée et, plus largement, des personnes. Il comporte de nombreuses orientations qui ont nécessité une évolution des réglementations nationales. Ainsi, dans de nombreux pays, la protection des données est une véritable préoccupation qui se traduit par un nombre conséquent de textes, dont la communauté doit tenir compte⁽⁷⁾. En cas de conflit entre les réglementations, il est souvent préconisé de retenir les dispositions les plus protectrices des droits des personnes.

Les projets financés par l'Union européenne

La dimension protection des données est affirmée et exigée par l'UE pour tous les projets qu'elle accompagne. Ainsi, chaque acteur a l'obligation sur chacun de ses sujets de recherche d'apporter des

informations détaillées sur l'utilisation et le devenir des données personnelles⁽⁸⁾.

L'Union européenne a élaboré un guide sur l'éthique et la protection des données pour aider les chercheurs dans la définition de leurs données sensibles et à porter à celles-ci une grande attention selon leurs thématiques de recherche.

Les démarches de conformité guidées par le terrain de recherche

Quel que soit le sujet de recherche, l'objectif de protection des personnes s'appuie sur des principes généraux que sont les finalités des traitements, la proportionnalité des données, les droits des personnes, les durées de conservation et les mesures de sécurité adaptées.

Pour chaque projet et chaque collaboration de recherche, des questions similaires sont soulevées : quels acteurs, quels rôles, quelles responsabilités, quelles données, comment partager, quel hébergement, quelle organisation adopter ?

La réglementation et les dispositions à retenir sont déterminées par l'analyse du projet et par le profil des acteurs concernés (leur pays d'origine, leur structure d'appartenance...). Les critères de localisation des responsables des traitements des données et de ciblage de ces traitements permettent de déterminer si le RGPD s'applique ou non.

Conclusion

La protection des données personnelles pour les finalités de la recherche s'insère dans les bonnes pratiques et les valeurs fondamentales de la recherche, quel que soit le périmètre national, européen ou international sur lequel elle porte. Il est un fait que le RGPD diffuse largement et que son application est de plus en plus internationale. Cette dimension est devenue cruciale pour le secteur de la recherche.

⁽⁷⁾ <https://www.cnil.fr/fr/la-protection-des-donnees-dans-le-monde>

⁽⁸⁾ Guide européen, voir : https://cache.media.education.gouv.fr/file/2018/54/9/h2020_hi_ethics-data-protection_en_1046549.pdf