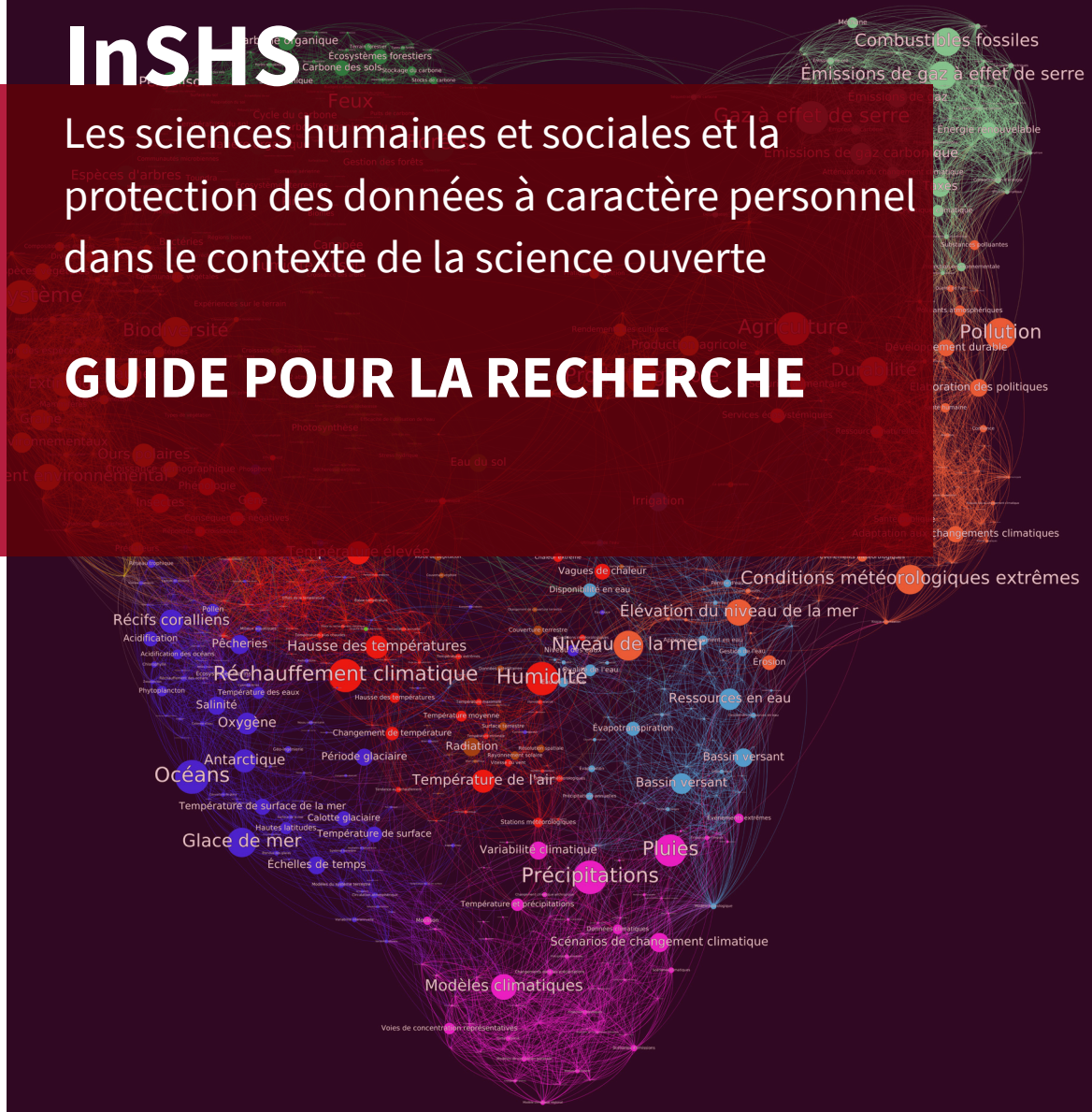
InSHS

Les sciences humaines et sociales et la protection des données à caractère personnel dans le contexte de la science ouverte

GUIDE POUR LA RECHERCHE



SOMMAIRE

Éditorial

“ L’informatique doit être au service de chaque citoyen […] Elle ne doit porter atteinte ni à l’identité humaine, ni aux droits de l’homme, ni à la vie privée, ni aux libertés individuelles ou publiques. ”

En 1978, la France adoptait la loi Informatique et Libertés, dont ces phrases, extraites de son premier article traduisent l’inspiration humaniste. Quarante ans plus tard, c’est le même esprit qui anime le Règlement Général de Protection des Données (RGPD) entrée en vigueur le 25 mai 2018.

Ce texte fait de l’Union européenne l’espace au monde où les données personnelles sont les plus fortement protégées, à un moment où certains excès de la révolution numérique rendaient nécessaires de revenir à des valeurs fondamentales centrées sur le respect de la personne humaine.

Parmi les nombreuses questions prises en compte par le RGPD, la recherche scientifique occupe une place importante et le législateur européen a su faire preuve d’un esprit d’équilibre à son égard. Si les grands principes de la protection des données personnelles s’appliquent aux activités de recherche, le texte reconnaît aussi leur légitimité et s’attache à mettre en place un régime spécifique offrant une latitude particulière aux chercheurs. Le RGPD affirme ainsi la compatibilité entre la protection des droits fondamentaux et la conduite des activités de recherche, sans opposer l’une à l’autre.

Le présent Guide a pour objectif de fournir aux communautés de recherche en Sciences Humaines et Sociales une ressource pour s’approprier ce nouveau cadre de la protection des données personnelles. Construit avec des chercheurs pour des chercheurs, il synthétise les règles applicables à chaque étape du cycle de vie des données et dégage les bonnes pratiques à mettre en œuvre en s’appuyant sur des exemples concrets.

Plus que toutes les autres disciplines scientifiques, les Sciences Humaines et Sociales mobilisent des matériaux (statistiques, enquêtes, interviews, archives, etc.) qui contiennent fréquemment des données personnelles appelant des précautions particulières. C’est la raison pour laquelle la Déléguée à la Protection des Données du CNRS, Mme Gaëlle Bujan, a souhaité d’abord travailler avec l’INSHS pour produire ce guide. Qu’elle en soit remerciée, ainsi que toutes les personnes ayant participé à son élaboration.

Le texte montre, en particulier, que les chercheurs ne sont pas seuls face à l’impératif de la protection des données. De nombreuses infrastructures comme Huma-Num, Progedo, le CASD (Centre d’Accès Sécurisé aux Données) ou encore le CINES (Centre Informatique National de l’Enseignement Supérieur) proposent d’ores et déjà des solutions aux chercheurs pour la gestion des données. Il s’agit de ressources précieuses dans lesquelles le CNRS est impliqué, avec la conviction que la mutualisation doit être privilégiée.

Parce que la protection des données et de la vie privée constitue un enjeu majeur pour le 21ème siècle, l’InSHS du CNRS veillera à épauler les unités de recherche dans cette période de transition. Le RGPD s’appuie sur un principe de responsabilisation des acteurs, et c’est avant tout une responsabilité collective qu’il s’agit de bâtir, par une collaboration étroite entre les chercheurs individuels, les directeurs d’unités, les infrastructures de recherche, les services de la Déléguée à la Protection des Données du CNRS et tous les professionnels intervenant au sein de cette chaîne.

Ce guide montre que le respect du RGPD n’est pas uniquement une question de conformité à la législation, même si cette dimension est essentielle. Les bonnes pratiques qu’il induit dans la collecte, le traitement, le stockage et la diffusion des données ont aussi une dimension épistémologique, appropriables par chaque discipline, et peuvent concourir à l’amélioration de la Science elle-même.

François-Joseph Ruggiu
Directeur de l’InSHS du CNRS

LE PRÉSENT GUIDE A ÉTÉ CONÇU PAR :

Isabelle André-Poyaud et Sandrine Astor, Ingénieures à Pacte, laboratoire de sciences sociales

Olivier Baude, directeur de la TGIR Huma-Num

Fabrice Boudjaaba, directeur adjoint scientifique à l'InSHS du CNRS

Gaëlle Bujan, Déléguée à la protection des données du CNRS

Béatrice Collignon, directrice de l'unité Passages

Frédéric Dubois, ingénieur de recherche, laboratoire d'ethnologie et de sociologie comparative

Emmanuel Kessous, laboratoire interdisciplinaire, sciences, innovations, sociétés

Lionel Maurel, directeur adjoint scientifique à l'InSHS du CNRS

Muriel Roger, professeure d'université, Centre d'Economie de la Sorbonne

Guide édité en juin 2019.

INTRODUCTION

Entré en application en mai 2018, le règlement européen sur la protection des données personnelles interroge la communauté scientifique, notamment dans les domaines des sciences humaines et sociales, quant à la compatibilité des travaux de recherche avec la réglementation.

Le droit est très protecteur et chaque personne peut librement disposer de ses données personnelles, à certaines conditions. Il s'agit d'un droit fondamental, la règlementation s'applique à tous (Cf. article 8 du Titre II « Libertés » de la charte des droits fondamentaux de l'Union européenne).

Ne pas se conformer à la règlementation sur la protection des données personnelles constitue une infraction pénale.

La communauté scientifique des sciences humaines et sociales utilise les données personnelles uniquement à des fins de recherche. Les législateurs l'ont compris et le règlement européen prend bien en compte les spécificités de l'activité scientifique : réutilisation possible des données à des fins de recherche, traitement de données sensibles (données de santé, sur les appartenances syndicales, les origines ethniques par exemple) à des fins de recherche en prenant des précautions adaptées, dérogations possibles sous certaines conditions à l'obligation d'information des personnes, … (Voir article 89 du RGPD).

L'objectif de ce guide est simple : il s'agit d'aider les chercheurs en SHS à comprendre ces dispositifs juridiques qui impactent leurs recherches et à leur fournir les bons réflexes et les bons outils dès lors qu'ils sont conduits à traiter des données à caractère personnel.

Il ambitionne d'être un outil d'accompagnement aux questionnements légitimes lors de la construction, la réalisation d'un programme de recherche, la publication des résultats et la potentielle réutilisation des données.

Les différents thèmes abordés dans le présent guide font référence à la règlementation sur la protection des données personnelles, à sa déclinaison pour les recherches en sciences humaines et sociales et présentent des exemples issus de situations rencontrées dans les laboratoires.

Les informations et exemples diffusés dans le guide sont centrés sur les données de recherche.

Il ne s'agit pas ici de traiter des questions de protection des données à caractère personnel et de la vie privée liées au fonctionnement ou l'administration de la recherche dans les laboratoires. Pour ces données, les textes applicables relèvent des règlementations européenne et française sans qu'il n'y ait spécifiquement de dérogations pour l'enseignement supérieur et la recherche.

Le guide sera mis à jour au moins une fois par an, dans la mesure du possible, par un comité d'experts similaires à celui qui a construit sa première version.

ELÉMENTS DE CONTEXTE

Dans l'environnement numérique qui caractérise la société actuelle, avec les larges possibilités de diffusion de l'information et des données, y compris les données à caractère personnel, l'utilisation quotidienne des réseaux sociaux et la faible frontière entre la sphère publique et la sphère privée, il est essentiel que la communauté scientifique conserve la confiance de chaque citoyen dans les activités et programmes de recherche initiés. La qualité de la recherche, une éthique de la recherche respectée constituent autant de principes simples qui participent au progrès de la connaissance.

L'ENVIRONNEMENT DE LA RECHERCHE

Le **numérique** ouvre des perspectives nouvelles, des accès à des données en masse et potentiellement réutilisables, de nouvelles techniques pour conserver, héberger, transférer des informations, … autant de ressources importantes pour la

recherche dont la fiabilité doit être préservée.

De plus en plus, le progrès de la connaissance, les avancées scientifiques et l'innovation sont en partie fondés sur l'utilisation/la réutilisation et le partage des données. La science ouverte, à savoir, la diffusion sans entrave des publications et des données de la recherche, constitue aujourd'hui « un nouveau paradigme » dans lequel chaque chercheur doit s'inscrire (Discours de Frédérique Vidal, Ministre de l'Enseignement supérieur, de la recherche et de l'innovation, du 4 juillet 2018 à l'occasion du lancement du Plan national pour la science ouverte).

La science ouverte s'appuie sur les opportunités du numérique pour développer l'accès à tous des publications et des données de la recherche ; elle contribue à l'efficacité de la recherche, ouvre les possibilités de s'intégrer dans la compétition internationale, favorise par la transparence de la recherche, la confiance des citoyens. Toutes les disciplines scientifiques sont concernées. ***Le Plan national pour la science ouverte*** s'inscrit dans les engagements internationaux pris par la France pour la transparence de l'action publique. Il répond également à l'ambition européenne de l'Amsterdam Call for Action on Open Science pour rendre accessibles, sans délai et sans paiement, les résultats de la recherche aux citoyens, aux entreprises, aux acteurs de la recherche.

Le plan, doté d'un budget de 5,4 millions d'euros, est décliné en trois axes :

- Généraliser l'accès à la science ouverte : publier obligatoirement en accès ouvert pour tout projet financé sur fonds publics; simplifier le dépôt des publications par les chercheurs ;
- Structurer et ouvrir les données de la recherche : développer la diffusion ouverte des données issues de la recherche financée sur fonds publics ; créer les conditions et promouvoir l'ouverture des données de la recherche ;
- S'engager dans une dynamique durable, européenne et internationale : développer les compétences en matière de science ouverte ; inciter les opérateurs de recherche à se doter d'une politique de science ouverte ; contribuer à la structuration européenne au sein du European Open Science Cloud.

L'évolution croissante des textes réglementaires dans tous les domaines de la vie des citoyens affecte également le travail scientifique. Ainsi, dans les domaines de la protection des données à caractère personnel, de nombreux autres textes s'appliquent aux traitements de données mis en œuvre, qu'il s'agisse de la loi sur la confiance numérique, du code de la santé publique, de la loi Jardé modifiée et de ses décrets d'application, du code du patrimoine, du code de la propriété intellectuelle, etc.

LES PRINCIPES DE LA RECHERCHE

La qualité de la recherche, l'éthique, l'intégrité scientifique sont les pratiques et comportements qui contribuent à la confiance des citoyens et des acteurs de la recherche.

Ces principes doivent naturellement être suivis pour toutes les recherches, y compris celles qui comportent des données à caractère personnel.

Une définition des données adaptée au projet, leur pertinence, leur volumétrie, leur mise à jour, leur stabilité dans le temps, la transparence de la méthodologie de leur construction contribuent à la fiabilité du travail scientifique, à la reproductibilité des résultats.

Pour tout projet, il est important de réunir, utiliser, analyser des données en lien avec la problématique de recherche. Une démarche objective, le respect des données collectées et du partage des résultats de la recherche sont compatibles avec les principes d'éthique et d'intégrité scientifique à suivre.

Chaque établissement d'enseignement supérieur et de recherche et de nombreuses institutions qui financent la recherche imposent ces pratiques et veillent à leur respect.

→ Le CNRS s'est doté d'un comité d'éthique et engage des réflexions sur les questions d'éthique générale suscitées par la pratique de la recherche et liées aux conséquences sociales et morales du progrès de la connaissance, aux principes

qui régissent les comportements individuels, à l'exercice de la science elle-même.

→ L'ANR a adopté en juin 2014 une politique d'éthique et d'intégrité scientifique qui décrit les principes fondamentaux que doivent respecter tous les acteurs de la recherche ainsi que les droits et les devoirs de ceux qui évaluent et soutiennent l'activité scientifique. Depuis 2019, l'ANR impose la réalisation d'un Plan de Gestion des Données (ou DMP : Data Management Plan) pour tous les projets retenus (voir page 9 du plan d'action de l'ANR)

→ L'Union européenne impose dans le programme H2020 une gestion des données de la recherche ; ces dernières doivent être « trouvables, accessibles, interopérables et réutilisables » (Principes FAIR pour « Findable, Accessible, Interoperable and Resable »). Les partenaires des projets soutenus par l'UE doivent ainsi réaliser un plan de gestion des données.

CHAPITRE 1 :

LES PRINCIPALES DÉFINITIONS ET LEUR APPLICATION POUR LES RECHERCHES EN SCIENCES HUMAINES ET SOCIALES

Dans cette première partie, les principaux concepts sont définis et complétés autant que possible d'exemples relevant des différentes disciplines des sciences humaines et sociales.
Les projets de recherche en SHS intégrant des données à caractère personnel sont quotidiens et il est important de protéger les informations des personnes impliquées dans les projets scientifiques.

1.1. LES DONNÉES À CARACTÈRE PERSONNEL

Les données à caractère personnel sont toutes informations qui permettent d'identifier directement ou indirectement la personne (Article 4 du RGPD) :

- Les données directement identifiantes : nom, prénom, adresse, photo, voix, etc
- Les données indirectement identifiantes : un numéro de téléphone, le croisement d'informations tel le fils du directeur de recherche, ce dernier habitant sur l'île de Batz, etc

Exemple : Un projet de recherche visant la réalisation d'un plan de déplacement d'entreprise ; les noms et prénoms des personnes ne sont pas collectés (ces informations ne sont pas nécessaires) mais des données sur les déplacements des personnes, leurs employeurs, leurs catégories socio-professionnelles et leur lieu de résidence permettent une identification des personnes physiques concernées. Ces informations sont donc des données à caractère personnel.

A noter

- Les données **anonymisées** de manière irréversible, qui ne permettent plus la réidentification d'une personne, ne sont pas soumises à la règlementation sur la protection des données personnelles.

- Les données **pseudonymisées** sont les données à caractère personnel qui ne peuvent plus directement être attribuées à la personne concernée. Mais le recours à des informations supplémentaires, par exemple une table de correspondance, permet de réidentifier cette dernière. Dans ce cas la réglementation sur la protection des données personnelles s'applique.

Parmi les données à caractère personnel, plusieurs sont **des données dites « sensibles »** dans la règlementation : les données qui révèlent la prétendue origine raciale ou l'origine ethnique, les opinions politiques, les convictions philosophiques ou religieuses, l'appartenance syndicale, l'orientation sexuelle, les données de santé, les données biométriques qui permettent d'identifier une personne, les données génétiques.

Les traitements de données dites sensibles ne sont pas autorisés (article 9 du RGPD) sauf exceptions explicitement listées dans la règlementation (par exemple, après consentement de la personne, données rendues manifestement publiques par la personne concernée, intérêt public important, sauvegarde de la vie humaine). L'utilisation de ces données sensibles est possible pour des finalités de recherche publique et lors de la préparation du projet, il faut obtenir l'avis préalable de la CNIL, dans certains cas, et organiser la sécurisation des données.

D'autres données font l'objet d'un encadrement spécifique :

- Le numéro de sécurité sociale (NIR) est une donnée directement identifiante dont l'utilisation est strictement encadrée. Cette donnée peut être utilisée si le traitement a une finalité exclusivement scientifique et sous réserve d'avoir fait l'objet d'une opération cryptographique préalablement au traitement de données.
- Les données d'infractions ou liées aux condamnations ne peuvent être traitées que par les juridictions et un certain nombre d'organismes limitativement énumérés dans la loi. Toutefois dans le cadre par exemple de conventions

avec le ministère de la justice, les établissements de recherche publique et les laboratoires associés peuvent parfois être conduits, sous certaines conditions strictement encadrées, à traiter ces données, et notamment si et seulement si le traitement n'a pas pour objet ou effet de réidentifier une personne.

1.2. LES ACTEURS ET LEUR RÔLE

La règlementation européenne en matière de protection des données personnelles modifie la logique de responsabilité sur les traitements de données.

D'une part, les obligations des acteurs, tels les sous-traitants, sont renforcées. D'autre part, le responsable du traitement doit mettre en œuvre les mesures techniques et organisationnelles adaptées pour assurer la conformité du traitement à la règlementation.

Pour les projets de recherche, plusieurs acteurs sont impliqués dans la conformité.

- **Le chercheur** qui conduit et mène tout projet de recherche financé ou non impliquant ou non plusieurs partenaires réalise les démarches pour s'assurer de la conformité du projet /traitement à la règlementation.

- **Le doctorant** réalise les démarches dans le cadre de son projet de recherche. Pour les sciences humaines et sociales, pour le traitement de données lié au projet conduit dans une unité mixte de recherche CNRS, le responsable du traitement est le directeur d'unité. Le doctorant réalise les démarches pour la conformité à la règlementation en lien avec son directeur de thèse.

- **Le responsable du traitement (RT)** est la personne, l'autorité publique ou l'organisme qui détermine la finalité et les moyens du traitement mis en œuvre. (Article 4 du RGPD)

Au CNRS

Pour les unités mixtes de recherche, le directeur d'unité est responsable de traitement (RT). Il doit donc s'assurer du respect de la règlementation sur la protection des données à caractère personnel et désigner un délégué à la protection des données. Il s'appuie pour cela sur les responsables scientifiques des projets conduits au sein de l'unité.

Le plus souvent, lorsque le directeur d'unité est employé par le CNRS, il désigne la Déléguée à la protection des données du CNRS.

Chaque responsable du traitement a l'obligation de documenter les traitements de données personnelles et de tenir à jour un **registre des traitements** qui précise notamment :

- Les finalités du traitement
- Les catégories de personnes et catégories de données concernées
- Les destinataires des données
- Les informations sur l'utilisation des données, leur conservation et les droits des personnes concernées
- Les noms et coordonnées du responsable du traitement et du délégué à la protection des données

Au CNRS

Le registre de chaque responsable de traitement (soit de chaque unité) est tenu par la Déléguée à la protection des données (DPD) pour le compte des responsables de traitement.

La démarche formalisée est réalisée par le responsable scientifique du projet auprès de la DPD qui conseille, accompagne, valide l'enregistrement du traitement de données.

A leur demande et au moins une fois par an, la DPD transmet aux directeurs d'unité la liste des traitements mis à jour pour leur unité.

Les démarches d'enregistrement des traitements qui étaient antérieurement réalisées auprès de la CNIL ou des Correspondants Informatique et Liberté, sont donc maintenant généralement effectuées auprès du Délégué à la protection des données de l'unité.

A noter, toutefois que l'avis préalable de la CNIL est requis avant tout traitement qui présente un risque trop important pour les personnes concernées, risque mis en avant lors de l'analyse d'impact sur la vie privée effectuée (voir page 16) et que le chercheur ne peut pas diminuer sans porter atteinte à sa recherche.

Une autorisation de la CNIL peut également être nécessaire dans le cadre des recherches dans le domaine de la santé (voir le site de la CNIL et page 21).

Dans tous les cas, il est conseillé de prendre l'attache du ou de la Déléguée à la protection des données pour une saisine conjointe de la CNIL.

→ **Le sous-traitant** : « La personne physique ou morale, l'autorité publique, le service qui traite des données personnelles pour le compte, sur instruction et sous l'autorité d'un responsable de traitement » (Article 4 du RGPD). Il doit présenter des garanties adaptées pour assurer la sécurité et la confidentialité des données, précisées notamment dans le contrat qui lie obligatoirement le responsable de traitement et le sous-traitant. Ce dernier précise également leurs engagements respectifs pour le traitement de données.

Exemples de sous-traitant :

Offre de service de la TGIR Huma-Num pour l'hébergement de données

Société de sondage : passation d'une enquête par un institut de sondage. Le sous-traitant est en charge de la collecte des informations auprès des personnes concernées, le chercheur récupère les informations ainsi obtenues. Dans certains cas il est possible que ces données soient pseudonymisées par le sous-traitant.

→ Le Délégué à la protection des données

Chaque établissement public doit être doté d'un délégué à la protection des données chargé de conseiller et d'informer les responsables de traitement de leurs obligations pour l'application des règlementations en matière de protection des données à caractère personnel (Cf. articles 37 à 38 du RGPD). Il conseille sur le respect de la règlementation, coopère avec l'autorité de contrôle, s'assure du respect de la règlementation sur la protection des personnes.

Chaque directeur d'unité doit désigner un Délégué à la protection des données. La désignation est à effectuer auprès de la CNIL.

A noter au CNRS :

Pour les directeurs d'unité qui choisissent la DPD du CNRS, la démarche est la suivante : le directeur d'unité informe la DPD de son choix. Celle-ci prend contact avec la CNIL pour la formalisation et la désignation.

Dans tous les cas, un récépissé d'enregistrement est transmis par la CNIL au directeur d'unité avec copie au DPD désigné. Ce document doit être conservé par l'unité et est intégré dans le corpus des documents relatifs à la protection des données à caractère personnel de l'unité.

→ **La Commission nationale Informatique et Libertés (CNIL)** est l'autorité de contrôle et de conseil en France chargée de surveiller, d'informer, d'accompagner l'application du règlement européen et de la règlementation française en matière de protection des données à caractère personnel (Cf. article 51 du RGPD) et chapitre 2 de l'ordonnance 2018-1125 du 12 décembre 2018 relative à la protection des données personnelles (modifiant la loi Informatique et Libertés du 6 janvier 1978).

1.3. LE PÉRIMÈTRE DE LA RÈGLEMENTATION : LA TERRITORIALITÉ

Le règlement européen s'applique aux traitements de données réalisés dans le cadre d'activités conduites par un établissement situé sur le territoire de l'UE.
Il s'applique également aux traitements effectués par un responsable de traitement ou un sous-traitant établis hors de l'Union européenne mais qui visent des personnes qui se trouvent sur le territoire de l'Union européenne (Cf. article 3).

Exemple : des chercheurs français ont réalisé une étude portant sur la diversité génétique et linguistique de la population du Cap-Vert. La législation européenne a vocation à s'appliquer car le responsable de traitement est situé sur le sol européen, peu importe que la collecte ait lieu au Cap-Vert.

La protection des données à caractère personnel s'effectue de manière différente selon les pays. Hors de l'Union européenne, plusieurs ont adopté une réglementation reconnue par la Commission Européenne comme assurant un niveau de protection adéquat au regard du RGPD ; dans d'autres pays, il n'existe pas de protection. La CNIL tient à jour l'état de la législation dans chaque pays : la protection des données personnelles dans le monde.

Le transfert des données hors Union européenne est possible à condition d'assurer un niveau de protection suffisant et adapté. Ces transferts doivent être encadrés en utilisant différents outils juridiques (voir site de la CNIL).

Il est conseillé de se rapprocher du délégué à la protection des données de l'unité.

Pour un laboratoire de recherche international (UMI ou une UMIFRE) dont les partenaires appliquent différentes règlementations, le droit applicable doit faire l'objet d'une analyse précise.
Une clause sur la protection des données personnelles doit être intégrée dans les contrats de collaboration internationale.
Il est conseillé de prendre l'attache du Délégué à la protection des données de l'unité.

1.4. LE TRAITEMENT DE DONNÉES

Un traitement de données est toute opération qui porte sur des données à caractère personnel, quel que soit le procédé, le support utilisé, informatisé ou non. Les données sont utilisées pour répondre à des objectifs/des finalités. Le traitement de données au sens « protection des données à caractère personnel » dépasse l'analyse ou l'exploitation de la donnée, il concerne la collecte, l'analyse, la réutilisation des données, l'archivage …..Article 4.2 du RGPD.

Exemple :
L'hébergement de données par Huma-Num et l'archivage par le CINES sont des traitements de données personnelles.

La finalité du traitement fait partie des principes essentiels de la règlementation. Tout traitement de données se réalise en fonction d'une finalité déterminée, explicite et légitime. Les données ne peuvent pas être traitées d'une manière incompatible avec la finalité définie.

Les données peuvent néanmoins être utilisées ultérieurement à des fins de recherche en apportant des garanties pour protéger la vie privée des personnes concernées par les données collectées (Cf. Article 5 du RGPD et article 89 du RGPD).

Pour les recherches en sciences sociales, la problématique de recherche est souvent la finalité du traitement de données.

Exemple : Étude sociolinguistique de la variation du langage utilisé sur Twitter est la finalité pour des données collectées qui sont l'état civil, des données économiques, de géolocalisation.

1.5. LES PRINCIPES RELATIFS AU TRAITEMENT DE DONNÉES

Avant d'engager son projet de recherche et lorsque celui-ci contient des données à caractère personnel, le responsable du projet scientifique engage l'analyse sur :

- La licéité du traitement, soit le fondement du traitement (a)
- La finalité du traitement (b)
- La pertinence et la proportionnalité des données (c)
- La sécurisation et la protection des données (d)
- La conservation limitée des données (e)
- La transparence des informations sur l'utilisation des données (f)

Outre le fondement du traitement, il s'agit de s'assurer du respect des principes de la protection des données personnelles afin de conduire sa recherche.

(a) Le responsable du projet ou du traitement de données vérifie si le projet est licite, c'est-à-dire, respecte l'une des conditions suivantes : (Article 6 du RGPD)

- La personne a consenti au traitement de ses données
- Le traitement est fondé sur une base légale
- Le traitement est lié à l'exécution d'un contrat
- Le traitement est nécessaire à la sauvegarde des intérêts vitaux de la personne concernée
- Le traitement est nécessaire à l'exécution d'une mission d'intérêt public
- Le traitement répond à un intérêt légitime pour le responsable de traitement

Dans les domaines des sciences humaines et sociales, le fondement est le plus souvent lié au consentement, à une mission d'intérêt public ou à un intérêt légitime.

Exemples :

- *Les enquêtes de terrain en France métropolitaine se réalisent souvent sur le fondement d'un consentement donné à l'enquêteur.*
- *Une recherche en sociologie avec collecte de messages échangés sur Twitter peut avoir pour fondement la mission d'intérêt public.*

(b) La finalité du traitement correspond à l'objectif suivi

La finalité doit être en lien avec les missions de l'établissement, de l'entité.

Exemple :
Etude sur l'évolution des inégalités spatiales depuis 30 ans dans les zones urbaines et émergence des « trappes urbaines » avec exploitation de données personnelles de géolocalisation issues des bases de l'INSEE.

(c) La pertinence et la proportionnalité des données

Elles doivent être corrélées avec les finalités.

Exemple : dans le cadre d'un projet de recherche sur les loisirs des personnes, il peut être pertinent de collecter un certain nombre de données complémentaires telles que la religion. Cette donnée peut avoir une incidence sur le choix des loisirs en fonction des jours de pratique de ces activités de loisirs. La collecte de cette information est opportune car elle a des conséquences sur les résultats de la recherche.

(d) La sécurisation et la protection des données

Le responsable de traitement est tenu de prendre toutes les dispositions pour protéger les données et empêcher qu'elles soient détournées, réutilisées à des fins non prévues, pour respecter l'intégrité et la confidentialité des données.
Des mesures de sécurité sont mises en place quelle que soit la nature de la donnée, à toutes les étapes du projet (voir page 24).

(e) La conservation limitée des données

Les données ne peuvent être conservées que pour une durée prédéfinie et limitée ; la finalité du traitement détermine la durée de conservation. A l'issue du traitement, les données sont soit anonymisées soit conservées pour une réutilisation ultérieure à des fins de recherche scientifique uniquement.

Pour la recherche, les données peuvent être archivées selon des dispositions spécifiques présentées au chapitre 2, page 26.

(f) La transparence du traitement des données

Les informations qui portent sur la finalité du traitement, le nom et les coordonnées du responsable du traitement, le nom et les coordonnées du délégué à la protection des données, les durées de conservation sont communiquées en toute transparence aux personnes concernées par le responsable du traitement des données.

Voir des exemples de mentions d'informations sur le site de la CNIL
Voir en annexe 2 un exemple de mention d'informations établi par l'unité mixte de recherche Pacte.

1.6. L'ANALYSE D'IMPACT SUR LA VIE PRIVÉE

Elle permet d'évaluer le risque d'un traitement de données sur la vie privée des personnes concernées. L'analyse est réalisée par le responsable du traitement (et par délégation, le responsable scientifique du projet) en lien avec le délégué à la protection des données et le responsable de la sécurité des systèmes d'information.

Elle est obligatoire lorsque le traitement est susceptible d'engendrer des risques élevés et doit notamment être réalisée dès lors que le traitement remplit au moins deux des critères suivants :

- Surveillance automatique,
- Données sensibles,
- Traitement à grande échelle
- Croisement de données
- Personnes vulnérables (patients, personnes âgées, enfants, etc)
- Evaluation/scoring (y compris profilage)
- Décision automatique avec effet légal
- Usage innovant ou utilisation NTIC
- Exclusion du bénéfice d'un droit, d'un contrat

La CNIL a publié la liste des traitements pour lesquels une analyse d'impact relative à la protection des données (AIPD) est obligatoire. Cette liste n'est pas exhaustive.

1.7. LES DROITS DES PERSONNES

Les droits des personnes sont renforcés par le RGPD.

→ L'information précise sur le traitement, la finalité, l'utilisation des données, la durée de conservation doit être faite aux personnes concernées par le traitement ; l'information doit être transparente et facilement accessible (article 12 du RGPD). Elle doit être faite directement auprès des personnes concernées. Une dérogation permet, lorsque la fourniture de ces informations est impossible ou exigerait des efforts disproportionnés, ou encore lorsque cette information serait susceptible de rendre impossible ou de compromettre gravement la réalisation des objectifs du traitement, de ne pas le faire auprès des personnes concernées mais de prendre des mesures appropriées pour protéger les droits et libertés des personnes, y compris en rendant les informations publiquement disponibles (article 14.5 du RGPD).

→ droit d'accès à ses données (Cf. article 15)

→ droit d'être informé d'une violation des données en cas de risque élevé pour les personnes concernées.

→ droit d'opposition (Cf. article 21) : une personne peut s'opposer, pour des motifs légitimes, à l'utilisation de ses données personnelles sauf si le traitement répond à une obligation légale. Une dérogation permet de ne pas donner droit à la demande lorsque le traitement a comme fondement l'exécution d'une mission d'intérêt public.

→ droit de rectification (Cf. article 16) : une personne peut demander à modifier les données la concernant.

→ droit à l'effacement (Cf. article 17) : une personne peut demander à avoir accès aux données la concernant et demander la suppression. Si l'exercice du droit est susceptible de rendre impossible ou de compromettre gravement la réalisation des objectifs du traitement, il est possible de ne pas accéder à la demande.

→ droit à la portabilité (Cf. article 20) : une personne peut demander à recevoir les données qui la concernent dans un format structuré et lisible par machine et de les transmettre à un autre responsable de traitement. Ce droit ne s'exerce pas au traitement nécessaire à l'exécution d'une mission d'intérêt public ou relevant de l'exercice de l'autorité publique dont est investi le responsable de traitement.

→ droit à une utilisation restreinte de ses données (Cf. article 23)

Toute personne peut aujourd'hui facilement exercer ses droits dès lors qu'elle connaît les noms et coordonnées du responsable de traitement et du Délégué à la protection des données, éléments qui sont obligatoires dans l'information faite aux personnes. Des délais de réponse sont prévus par la règlementation : à compter de la réception de la demande d'accès aux données, la transmission de celles-ci doit se faire dans un délai d'un mois.

Au CNRS, l'exercice des droits est traité par la Déléguée à la protection des données et par le responsable de traitement

Premier cas : demande effectuée auprès du responsable du traitement ou du responsable scientifique
Transmission de la demande à la DPD qui le conseille ; envoi de la réponse à la personne et transmission de la copie à la DPD

Deuxième cas : demande formulée auprès de la DPD
Le responsable du traitement ou le responsable scientifique est sollicité pour préparer la réponse qu'il envoie avec transmission de la copie à la DPD.

La demande d'exercice des droits, la copie de la réponse sont intégrées dans le registre du responsable du traitement tenu par la DPD.

L'une des principales évolutions du RGPD porte sur l'obligation, pour le responsable de traitement (et le sous-traitant) de définir et d'organiser les mesures permettant de démontrer à tout moment la conformité à la règlementation.

Cette responsabilisation des acteurs nécessite une analyse approfondie des données et de leur traitement. Lors de la construction puis de la réalisation des projets de recherche, cette analyse est importante, voire imposée par les financeurs.

CHAPITRE 2

LES PROJETS DE RECHERCHE, LE CYCLE DE VIE DES DONNÉES ET LA PROTECTION DES DONNÉES À CARACTÈRE PERSONNEL

Cette seconde partie s'attache à décrire de manière concrète les modalités d'analyse de la donnée, les questions liées à la règlementation sur la protection des données personnelles en suivant le cheminement de la vie de la donnée - dans le cadre de projets de recherche - , de sa collecte à sa diffusion voire son réemploi.

Dans le souci et le respect des principes d'une recherche éthique, la fiabilisation, la qualité de la recherche et des données induisent une démarche formalisée préalable à la constitution de tout projet de recherche. Elle permet de réfléchir de manière anticipée à toutes les étapes du projet et offre la possibilité de prendre les décisions adaptées au moment opportun facilitant ainsi la réalisation du projet : par exemple, identification des mesures de sécurité à prendre, sollicitation des autorisations préalables au projet (autorisation de tournage, utilisation des données de santé, des données sensibles dans certains cas …), anticipation des questions de conservation des données à la fin du projet et réflexion sur leur mise à disposition à l'issue du projet.

A lire : document d'Huma-Num sur les grandes étapes d'un projet de recherche à l'ère du numérique.

Tout projet engage désormais une réflexion préalable sur les données de recherche. Il s'agit de les définir et de décrire selon quelles modalités elles seront collectées, conservées, archivées, et ainsi d'envisager un plan de gestion des données.

La recherche de financement d'un projet est également facilitée voire conditionnée par un travail d'explicitation préalable de la nature des données qu'il utilise et de la manière dont il les traite. Dès lors, un projet scientifique doit formaliser et exposer l'ensemble des étapes de l'exploitation et du traitement des données. C'est pourquoi progressivement les pratiques évoluent et les établissements financeurs imposent de plus en plus des **plans de gestion de données**.

L'Union européenne depuis 2007, et depuis 2019 l'ANR, conformément à leurs objectifs de qualité et de libre accès aux données de la recherche demandent qu'un plan de gestion des données soit réalisé pour tout projet financé.

Un Plan de Gestion de Données (DMP) : Qu'est qu'un DMP ? Pourquoi un DMP ?

Un DMP est un document formalisé qui explicite la manière dont sont obtenues et traitées les données tout au long de leur cycle de vie, de leur collecte à l'archivage.
Il doit indiquer :

- quel est le traitement des données de recherche avant, pendant et après la fin du projet,
- les données qui seront collectées, traitées et/ou générées,
- si les données sont partagées, rendues accessibles, comment les données seront organisées et conservées (y compris après la fin d'un projet).

Un DMP :

- Garantit la qualité de la recherche
- Contribue à des données FAIR « facilement accessibles, identifiables, reproductibles » (pour les projets H2020)

- Est un outil de fiabilité à l'ère du numérique et de connaissance pour permettre la potentielle réutilisation des données liée à l'*Open access*
- Répond à une demande des financeurs : Union européenne, ANR, … Les frais associés peuvent être intégrés dans les dépenses éligibles.

Les éléments relatifs à la protection des données personnelles n'alimentent qu'une partie des informations dans un DMP, même si la conformité à la règlementation concerne toutes les étapes d'un DMP.

Plusieurs établissements proposent des méthodes et modèles pour réaliser un Plan de gestion des données :

L'INIST propose des modèles : voir OPIDOR.

L'université Paris Diderot propose une méthodologie pour la réalisation des DMP.

L'USR PROGEDO propose avec son réseau des plateformes universitaires de données, PUD, sur le territoire national, un accompagnement pour la réalisation des plans de gestion des données.

L'INRA propose un guide.

2.1. LA CRÉATION DE LA DONNÉE (COLLECTE)

2. 1. 1. LES CATÉGORIES DE DONNÉES

Le mode de collecte des données personnelles dans la règlementation ne doit pas être confondu avec le mode de production des données de recherche en sciences humaines et sociales.

A noter : la règlementation sur la protection des données personnelles distingue :

Les données collectées directement auprès des personnes concernées. Ces dernières doivent être informées précisément des données concernées, de leur utilisation, de l'objectif de leur traitement, de leur durée de conservation, et des modalités d'exercice de leurs droits associés, des noms du responsable de traitement et du Délégué à la protection des données.

Les données collectées indirectement font également l'objet d'une information sur le traitement des données comme précédemment. Pour ce type de collecte, il faut aussi informer les personnes concernées des catégories de données recueillies et de la source des données (en indiquant notamment si elles sont issues de sources accessibles au public).
Voir articles 13 et 14 du RGPD

A - Les données produites directement à des fins de recherche par le chercheur

Exemple :

Données issues d'une enquête sociologique, d'une enquête ethnologique, de terrain, d'archives orales, de questionnaires, de formulaires, réalisation d'interviews, données extraites du web, ….

Dans l'hypothèse d'un recours à un prestataire, un contrat détermine les obligations et engagements de chaque partie. Le responsable de traitement doit communiquer toute information sur le traitement de données au sous-traitant et respecter les principes de la règlementation sur la protection des données. Un contrat encadre les obligations de chacun (Article 28 du RGPD).

A noter

Le sous-traitant a des responsabilités importantes pour la protection des données personnelles. Ainsi, il doit :

- Prendre en compte les aspects de sécurité, confidentialité et documentation des activités réalisées pour le responsable du traitement
- Aider le responsable du traitement dans la mise en œuvre de ses obligations (analyse d'impact sur la vie privée, notification de violation de données, sécurité)
- Tenir un registre des traitements effectués pour les responsables de traitement
- Désigner un Délégué à la protection des données dans les mêmes conditions qu'un responsable de traitement.

Dans le cadre de projets de recherche intégrant des données personnelles avec recours à un prestataire extérieur, il est recommandé de solliciter les services Partenariat et valorisation et/ou financier/achat de la tutelle qui gère les dépenses du projet pour adapter le contrat à la nécessaire protection des données personnelles.

A l'identique, lorsque le laboratoire traite des données dans le cadre de prestations de service, il est recommandé de prendre l'attache du service Partenariat et valorisation de la tutelle qui gère le contrat.

Pour les enquêtes réalisées par un/des étudiants : prévoir un engagement de confidentialité et veiller à la sécurité des systèmes d'information utilisés. Les étudiants respecteront les modalités prévues par le traitement enregistré auprès du DPD de l'unité au sein de laquelle la recherche est conduite.

B - Les données produites initialement à des fins autres mais ensuite utilisées à des fins de recherche

Exemple

Données provenant de la statistique publique, d'enquêtes d'instituts de sondage, d'établissements administratifs, données issues de fichiers administratifs

Dans le cas où les données sont collectées non directement auprès des personnes concernées, mais auprès d'un tiers ayant lui-même valablement collecté des données, le RGPD prévoit que la poursuite d'une finalité de recherche est compatible avec la finalité initiale pour laquelle des données ont été collectées. Dans ce cas, la règlementation sur la protection des données personnelles s'applique au nouveau traitement.

Dans le cas des données issues de la statistique publique, les chercheurs peuvent accéder à différents fichiers de données comportant des informations fines et individuelles sur les personnes ou les entreprises (composition ou revenu des ménages, bénéfices des entreprises, localisation, etc.). Ces fichiers sont de 3 types, et peuvent être issus de la même source. Ils se différencient par leur précision et les risques induits d'identification des personnes ou des entreprises. On distingue :

- les fichiers confidentiels donnant des informations précises sur les enquêtés qui peuvent permettre leur identification. L'accès à ces données est possible pour les chercheurs dans un cadre sécurisé, sur projet, après accord du Comité du Secret Statistique. Cet accès, payant, se fait via le Centre d'Accès Sécurisé aux données (https://www.casd.eu/)
- les fichiers de Production et de Recherche. Ces fichiers sont pseudonymisés. Les données mises à disposition des chercheurs sont non directement nominatives mais sont considérées comme pouvant être indirectement identifiantes (Cf. définition page 8). La production de ces fichiers est faite spécifiquement à des fins de recherche. Ils sont accessibles à la communauté scientifique via Quetelet Progedo Diffusion (http://quetelet.progedo.fr/)
- Les fichiers standards pour lesquels tous les traitements nécessaires ont été réalisés afin de les rendre anonymes (exemple : agrégation des professions et catégories socio-professionnelles et des zones géographiques

d'habitation pour les ménages). Ils sont accessibles en open data.

2. 1. 2. LES TYPES DE DONNÉES À CARACTÈRE PERSONNEL

Selon leur sensibilité, les données à caractère personnel peuvent faire l'objet de dispositions spécifiques afin de protéger la vie privée des personnes.

Pour les traitements à des fins de recherche, les données peuvent :

1- ne comporter aucune sensibilité spécifique
2- concerner des données dites sensibles (Cf. page 11).

Le traitement des données dites sensibles est normalement interdit sauf exceptions prévues par la règlementation : par exemple, consentement exprès ou pour des finalités de recherche. Et dans ce cas, le respect de la règlementation est particulièrement important (voir page 11).

Exemple

Étude comparative des habitudes alimentaires des lycéens résidant en France, aux Etats-Unis, au Sénégal et au Brésil ; suivi des enfants de 15 à 16 ans durant 3 ans : données sur le genre, la catégorie socio-professionnelle des parents, convictions religieuses …

3- concerner des populations vulnérables : les mineurs, les personnes âgées, les salariés auprès desquels des enquêtes sont menées, les prisonniers, les demandeurs d'asile. La règlementation ne définit pas la notion de population vulnérable. Mais des dispositions spécifiques sont à respecter : par exemple demander le consentement des personnes et garantir la sécurité et la confidentialité des informations. Pour des recherches impliquant des mineurs, l'information doit être adaptée et il est recommandé d'informer précisément les parents également (ou la personne disposant de l'autorité parentale).

Exemple

Étude visant à spécifier les évolutions des interactions soignants - patients, ces dernières peuvent susciter de nouvelles fragilités sociales, ou au contraire, participer à la réduction des inégalités sociales de santé.

4- concerner des recherches dans le domaine de la santé
Les données de santé sont définies de manière large comme étant des données relatives à la santé physique ou morale passée, présente ou future qui donne une indication sur l'état de santé de la personne.

Exemple

Suivi de la vie quotidienne d'hommes de moins de 30 ans atteints d'autisme : collecte de données sur la pathologie et son évolution, données sur la personne, ses habitudes de vie

Les démarches pour les recherches comportant des données de santé :

→ **Situation 1 :** la recherche implique des données de santé (grossesse, handicap, maladie chronique…) mais n'est pas de la recherche dans le domaine de la santé : la procédure est identique à celle comportant des données dites sensibles (voir page 11).

→ **Situation 2 :** recherche dans le domaine de la santé impliquant la personne humaine :
- Demande d'avis à un Comité de Protection des Personnes physiques (CPP)

Soit le traitement est conforme à une Méthodologie de référence (MR) définie par la CNIL. Dans ce cas, le responsable de traitement réalise une analyse d'impact sur la vie privée et signe un engagement de conformité à la MR auprès de

la CNIL. **Cet engagement de conformité sera inscrit au registre**. La liste des traitements mis en œuvre dans le cadre de la méthodologie de référence doit être inscrite dans le registre du responsable de traitement.

- MR-001 : impliquant la personne humaine pour des recherches interventionnelles. La MR-001 nécessite le consentement exprès et éclairé du patient ou celui de ses représentants légaux.
- MR-003 : impliquant la personne humaine pour des recherches non interventionnelles. Cette méthodologie n'implique pas le consentement du patient, mais l'information au patient est obligatoire.

Soit l'autorisation préalable de la CNIL est nécessaire et il faut également réaliser une étude d'impact sur la vie privée

→ **Situation 3 :** recherche dans le domaine de la santé n'impliquant pas la personne humaine

- Soit le traitement est conforme à la MR-004 : recherche n'impliquant pas la personne humaine, liées à des études, évaluations dans le domaine de la santé (**réaliser une étude d'impact sur la vie privée, un engagement de conformité auprès de la CNIL. Cet engagement de conformité sera inscrit au registre** ; la liste des traitements mis en œuvre dans le cadre de la méthodologie de référence doit être inscrite dans le registre du responsable de traitement.). Le traitement est également inscrit au répertoire public mis à disposition par l'Institut national des données de santé.

- Soit le traitement n'est pas conforme à cette Méthodologie de Référence et il faut demander un avis au CEREES via l'Institut national des Données de Santé (INDS), réaliser une étude d'impact sur la vie privée puis déposer une demande d'autorisation auprès de la CNIL.

Dans tous les cas, il est conseillé de conduire l'analyse d'impact sur la vie privée et l'ensemble des démarches avec le Délégué à la protection des données de l'unité.
Dans tous les cas, le traitement sera inscrit au registre des traitements et documenté avec l'analyse d'impact sur la vie privée, l'engagement de conformité et/ou l'autorisation de la CNIL et tout autre document utile.

5 - Les recherches incluant des données relatives aux condamnations pénales et aux infractions ou mesures de sûreté connexes doivent faire l'objet d'une attention particulière, et nécessitent notamment une analyse d'impact sur la protection des données.

2.1.3. LE FONDEMENT DU TRAITEMENT ASSOCIÉ À LA COLLECTE DE DONNÉES

Comme indiqué dans le chapitre 1, la règlementation prévoit six fondements parmi lesquels trois sont souvent retenus en recherche : le consentement, la mission d'intérêt public, l'intérêt légitime.

Dans le cadre des activités de recherche, les traitements devraient de manière préférentielle s'effectuer sur la base du consentement (respect du principe d'auto-détermination informationnelle), mais les traitements peuvent aussi s'appuyer sur le fondement de l'exercice d'une mission d'intérêt public.

Le choix revient toujours au responsable du traitement qui analyse la nature des données, la population concernée. Le choix doit être argumenté et le Délégué à la protection des données apporte les conseils pour déterminer le fondement du traitement.

FOCUS sur le consentement

Le consentement peut être donné sous des formes différentes par écrit ou par voie orale. Dans tous les cas, il est important d'assurer la traçabilité du recueil du consentement. Il doit comprendre les informations nécessaires pour que le consentement soit libre, spécifique, éclairé et univoque.

Voir sur le site de la CNIL, comment recueillir le consentement des personnes :
https://www.cnil.fr/fr/conformite-rgpd-comment-recueillir-le-consentement-des-personnes

Une fois donné, le consentement doit pouvoir être retiré aussi simplement qu'il a été accordé. Dans ce cas, les données associées ne peuvent plus être traitées dans le projet.

Le responsable de traitement est tenu de conserver la preuve que le consentement lui a été donné pour réaliser un traitement.

Voir annexe 1 : Exemple de consentement proposé par l'unité mixte de recherche Pacte.

2. 1. 4. LA FINALITÉ

L'article 5 du RGPD prévoit que les données personnelles ne peuvent être collectées que pour des « finalités déterminées, explicites et légitimes » qui doivent en principe être définies en amont du traitement et être portées à la connaissance des personnes concernées.

Néanmoins, le considérant 33 du RGPD admet aussi une certaine indétermination des finalités pour les traitements réalisés à des fins de recherche :

> *« Souvent, il n'est pas possible de cerner entièrement la finalité du traitement des données à caractère personnel à des fins de recherche scientifique au moment de la collecte des données. Par conséquent, les personnes concernées devraient pouvoir donner leur consentement en ce qui concerne certains domaines de la recherche scientifique, dans le respect des normes éthiques reconnues en matière de recherche scientifique. Les personnes concernées devraient pouvoir donner leur consentement pour ce qui est de certains domaines de la recherche ou de certaines parties des projets de recherche, dans la mesure où la finalité visée le permet » (considérant 33 RGPD).*

Bien que la finalité ne soit pas totalement déterminée, le traitement de données personnelles est possible. L'information des personnes doit être adaptée aux finalités connues au début de la recherche.

2. 1. 5. LA PROPORTIONNALITÉ DES DONNÉES

Il s'agit de savoir si les données collectées sont pertinentes, nécessaires à la réalisation du traitement.

Exemple

La recherche nécessite de collecter l'âge des personnes, mais il n'est pas besoin pour autant de collecter le jour et le mois de la naissance. Pour le domicile, est-il besoin de recueillir l'adresse précise ou seulement la ville de résidence ? Pour l'activité, est-il besoin de connaître seulement la catégorie socio-professionnelle ou le fait que la personne exerce tel métier dans telle entreprise, etc.

De même, des données personnelles imprévues peuvent être collectées.
Il s'agit des données collectées lors d'entretiens par exemple qui ne sont pas liées au traitement de données.

Le chercheur doit apprécier si les données doivent être conservées, supprimées pour son enquête initiale ou une réutilisation plus tard (idem pour les données dites sensibles).

2 . 2. LE STOCKAGE DES DONNÉES

La sécurisation de l'accès aux données, de leur stockage, de leur hébergement est essentielle pour protéger les données personnelles. Le responsable du traitement s'appuie en cela sur les outils proposés par ses établissements tutelles et respecte/fait respecter les mesures internes à l'établissement.

Les règles de base (liste non exhaustive) pour la sécurisation des systèmes d'information, de l'utilisation des outils numériques, des échanges de données et du stockage de données s'inscrivent dans le respect de la politique de sécurité des systèmes d'information des établissements de rattachement de l'unité et des chercheurs :

- L'authentification des utilisateurs des outils numériques : les certificats numériques, les mots de passe
- La gestion des habilitations : accès aux sites et données aux seules personnes habilitées par le responsable de traitement ou du projet de recherche
- La sécurisation des outils : chiffrement des ordinateurs et des smartphones
- La protection des réseaux informatiques internes
- La sécurisation des échanges entre organismes, entre unités et chercheurs
- L'utilisation d'outils sécurisés pour les visioconférences (au CNRS : recours à Skype entreprise).

Des outils accessibles
L'offre de services numériques du CNRS
Edition de la CNIL : la sécurité des données personnelles, édition 2018

FOCUS : Exemples de mauvaises pratiques :

Utilisation d'une messagerie non sécurisée pour les échanges d'informations
Interactions messageries privées et professionnelles
Les outils pour réaliser des enquêtes en ligne dont les serveurs sont à l'étranger (type Google Forms).
L'enregistrement de fichiers sur les postes de travail alors qu'ils sont accessibles uniquement sur des espaces de travail sécurisés
Echanges de fichiers contenant des données personnelles dites sensibles par messagerie sans chiffrement des messages

FOCUS : Les conseils de l'Etat dans le cadre du programme de sécurité numérique :

Gérer ses mots de passe
La sécurité des appareils mobiles
La distinction usages professionnels, usages privés

(source : GIP Action contre la Cybermalveillance)

Plusieurs entités proposent des solutions pour le stockage des données. Quel que soit le support, il est important de prévoir le volume des données, la durée de la conservation et d'estimer le coût du stockage (celui-ci peut être parfois pris en compte dans les dépenses éligibles des projets de recherche).

L'offre de service d'Huma-Num

2. 3. L'EXPLOITATION DES DONNÉES

Dans cette section, l'exploitation des données se comprend comme l'étape de l'analyse des données, de leur synthèse, du travail sur les données contribuant à éclairer la problématique de recherche.

Les données peuvent avoir des statuts différents.

1 – les données identifiantes portant sur un faible nombre de personnes nécessaires pour des analyses qualitatives (fréquent en ethnologie, géographie sociale ou culturelle, en sociologie).
Il est important de supprimer le caractère identifiant pour l'étape de publication et/ou à la fin du projet de recherche. Selon les situations, l'anonymisation ou la pseudonymisation sont requises.

2 – les données anonymisées

Le lien avec les données personnelles est rompu de manière irréversible.
Si, l'identification d'une personne n'est possible d'aucune manière, la règlementation sur la protection des données personnelles ne s'applique pas.
Dans le cadre des enquêtes qualitatives, l'anonymisation ne sera généralement pas possible, dans la mesure où existe un besoin avéré d'identifier des personnes.
Lorsque les objectifs de la recherche nécessitent de mentionner l'identité de l'interviewé (personnalité, expert,…), il convient de leur préciser que des données identifiantes seront publiées et de leur garantir l'accès à la retranscription.

Pour des études quantitatives, l'anonymisation devrait être opérée au nom du respect du principe de proportionnalité, dès le moment de la collecte des données.

3 – Les données pseudonymisées consistent à séparer les données directement identifiantes (exemple : nom et prénom) des autres données non identifiantes (exemple, en attribuant un numéro aux personnes évitant de faire apparaître leur nom, mais en conservant une table de correspondance permettant de remonter à l'identité de la personne).

FOCUS sur l'anonymisation
L'anonymisation des données nécessite que l'identification des personnes devienne impossible, que ce soit de manière directe ou indirecte, opération qui suit un processus spécifique.
Quelle que soit la technique utilisée, l'anonymisation doit conduire au respect de trois critères :
- Impossibilité totale d'isoler un individu
- Impossibilité totale de relier entre eux les enregistrements relatifs à deux individus
- Impossibilité de déduire des informations concernant un individu

Exemples de techniques d'anonymisation (avis du G29) :

- *Ajout de bruit : altérer la justesse de l'information en ajoutant de l'aléa*
- *Permutation : Mélanger les valeurs d'attributs au sein du jeu de données*
- *Généralisation : Changer la granularité des valeurs pour former des groupes*
    - *k-anonymat : au moins k personnes ont le même profil*
    - *L-diversité : au moins l valeurs ont le même attribut*

Big Data et Intelligence Artificielle

Le recours à des procédés de traitement des données innovants, comme le Big Data ou l'Intelligence Artificielle (IA), nécessite de prendre des précautions particulières, eu égard aux caractéristiques de ces technologies.

Parce qu'il implique le traitement – et souvent le croisement – de grandes masses de données, le Big Data comporte des risques plus élevés d'identification indirecte des personnes, même lorsqu'il s'opère à partir de données a priori anonymes.

La constitution d'une base de données à partir de plusieurs sources requiert des précautions particulières. Il faut s'assurer que le consentement des personnes concernées a bien été recueilli ou – dans les cas où les traitements s'effectuent sur un autre fondement – que le droit à l'information des personnes est respecté.

La mise en œuvre de technologies comme les Big Data ou l'Intelligence Artificielle nécessitera souvent de réaliser au préalable une étude d'impact sur la vie privée (PIA. Voir page 16). Une étude d'impact est en effet nécessaire lorsque sont présents plusieurs critères parmi lesquels le profilage des personnages, le croisement de données, des décisions automatiques ou des usages innovants impliquant le recours à de nouvelles technologies.

S'agissant de l'Intelligence Artificielle et de procédés comme le Machine Learning, il faut garder à l'esprit que les opérations de profilage (au sens de traitements de données personnelles à des fins d'analyse et de prédiction des comportements), de même que les décisions entièrement automatisées qui peuvent en découler, font l'objet d'un encadrement particulier dans le RGPD.

Les personnes peuvent notamment faire valoir des droits à la transparence et à l'intervention d'une personne humaine, lorsque de telles décisions ont un effet légal.

2.4. L'ARCHIVAGE DES DONNÉES

En principe, les traitements de données à caractère personnel doivent avoir une durée déterminée dans le temps, liée à la réalisation de la finalité pour laquelle elles ont été collectées. Le RGPD indique que cette durée de conservation des données soit fixée au « strict minimum », à l'issue de quoi les données doivent normalement être archivées conformément à la règlementation en matière d'archives publiques.
Néanmoins, le RGPD prévoit aussi que les données peuvent être conservées pour des durées plus longues lorsqu'elles font l'objet d'un traitement *« à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques »* (Cf. Article 5 du RGPD et informations sur le site de la CNIL).

La conservation des données, selon la règlementation sur les archives prévoit un cycle en trois temps.

Phase 1 : la base active
Elle correspond à l'utilisation courante des données, soit le temps de la recherche.

Phase 2 : l'archivage intermédiaire
Les données à caractère personnel peuvent, dans certaines conditions, être conservées à l'issue du traitement des données mais avec un accès restreint. Les données personnelles à des fins de recherche font souvent l'objet d'un archivage dit intermédiaire sous réserve que les droits des personnes et l'information associée soient respectés.

Au CNRS, la durée d'archivage intermédiaire est souvent de deux années après la dernière publication des résultats de la recherche.

Phase 3 : l'archivage définitif
Les données personnelles qui ne font pas l'objet d'une destruction peuvent être archivées selon les dispositions du Code du Patrimoine (livre 2). L'archivage définitif ne peut pas être réalisé dans le laboratoire. Il est effectué avec les services d'Archives départementales ou nationales en lien avec les établissements de rattachement du laboratoire.
Une fois archivées, des données de recherche comportant des informations à caractère personnel peuvent être consultées dans le respect des règles générales liées à la communication des archives déterminées par le Code du Patrimoine (délais de communicabilité, dérogations possibles pour les chercheurs, etc.).

Les durées de conservation doivent être définies et indiquées de manière transparente lors de l'enregistrement d'un traitement. Ces informations doivent également être transparentes pour toute personne concernée par un traitement de données personnelles. Les durées peuvent être modifiées en cours de traitement.

Le CINES et la TGIR Huma-Num (en collaboration avec le CINES et les Archives nationales) proposent des services pour

l'archivage des données.

2. 5. LE PARTAGE DES DONNÉES DANS LA RECHERCHE PARTENARIALE

Les activités de recherche sont parfois conduites dans le cadre de partenariats impliquant plusieurs entités, relevant de tutelles différentes, associant parfois des acteurs publics et des acteurs privés.

Dans ces hypothèses, il est impératif de prévoir en amont, dans le cadre d'une convention de partenariat (type accord de consortium) quelles qualités, au sens du RGPD, auront les partenaires impliqués dans le projet : responsable de traitement, co-responsable de traitement ou sous-traitant. Ces contrats de collaboration doivent permettre d'identifier les rôles et les obligations de chacun, notamment en matière de sécurisation des données.

Le porteur scientifique du projet est le responsable naturel du traitement, sous couvert de son directeur d'unité.

Lorsqu'une recherche partenariale est conduite entre plusieurs acteurs déterminés, le périmètre des accréditations pour l'accès et la manipulation des données doit être fixé en amont au niveau des conventions de collaboration.

Dans tous les cas, pour la finalité de recherche définie pour le projet, les financeurs n'auront pas le statut de responsables de traitement, notamment dans la mesure où ils ne sont généralement pas destinataires des données brutes, mais uniquement de la recherche finalisée.

2. 6. LA DIFFUSION DES DONNÉES ET LA PUBLICATION

Plusieurs hypothèses sont à distinguer :

- La diffusion des données anonymisées (toujours possible lorsque les données sont réellement anonymisées de manière irréversible) ;
- La transmission à d'autres chercheurs de données non-anonymisées ;
- La publication de données non-anonymisées dans des papiers de recherche ;
- La diffusion des données non-anonymisées sur la base du consentement des personnes concernées.

Pour ce qui est de la transmission de données non-anonymisées à d'autres chercheurs, elle est rendue possible sous autorisation du responsable de traitement en vertu du décret du 1er août 2018 qui prévoit (art. 100-1) :

> *« Les données issues de ces traitements conservées par le responsable du traitement ou son sous-traitant ne sont accessibles ou modifiables que par des personnes autorisées. Ces personnes respectent les règles de déontologie applicables à leurs secteurs d'activités. »*
>
> *Les autorisations accordées par les responsables de traitement à ces personnes sont strictement encadrées, elles respectent les finalités spécifiques et les garanties prévues par le décret pré-cité.*

Les recherches à partir de données personnelles publiques collectées en ligne ou via les services des réseaux sociaux. La règlementation s'applique si les données ne sont pas anonymes.

- définir la base juridique : généralement, le consentement ou la mission d'intérêt public
- définir la finalité du traitement et la pertinence des données collectées
- prévoir les modalités d'information des personnes (exemple : sur les sites web des laboratoires, par voie de presse) et notamment les modalités d'exercice de leurs droits
- recourir à l'anonymisation des données, dès que cela est possible
- déterminer les durées de conservation des données en fonction des finalités du traitement et de l'étape du projet

Porter une attention au mode de collecte des données, à la sécurisation du stockage, au partage, …

A lire : la délibération du 3 mai 2018 de la CNIL autorisant l'université de Lorraine à mettre en œuvre un traitement automatisé de données à caractère personnel ayant pour finalité une recherche sur les impacts pour la vie privée des publications d'informations librement accessibles sur les réseaux sociaux

2. 7. LE RÉEMPLOI DES DONNÉES

Le réemploi des données permet de partager la ressource « données personnelles » avec d'autres chercheurs, notamment dans le cadre de la science ouverte.

Depuis 2016, la règlementation (Loi République numérique, Loi Valter) a fixé un principe général d'ouverture et de libre réutilisation des informations publiques (Open Data par défaut). Les données de la recherche constituent en principe de telles informations publiques. Guide d'ouverture des données de recherche du CoSo

Néanmoins, la loi articule ces obligations d'ouverture avec la protection des données personnelles, en précisant que lorsqu'elles comportent des données personnelles, les informations publiques ne peuvent être rendues publiques « *qu'après avoir fait l'objet d'un traitement permettant de rendre impossible l'identification de ces personnes* » (anonymisation) ou avec le consentement des personnes concernées.
Les données peuvent être réemployées à des fins de recherche lorsqu'elles ont été anonymisées. Elles peuvent également l'être si les personnes ont donné leur consentement ou si le réemploi a été prévu par le traitement initial.

Sauf lorsque les données ont été anonymisées, le réemploi ne dispense pas des procédures qui réactualisent la conformité au RGPD : fondement licite du traitement, finalité explicite, légitime, proportionnalité des données, sécurisation des données, information des personnes ….
Voir : Guide pratique de la CNIL et de la CADA sur la publication en ligne et la réutilisation des données publiques (Open data)

Exemple : Quetelet Progedo Diffusion et le CASD donnent accès à des données par définition réutilisables pour des projets à finalité de recherche.

Les données peuvent être réemployées pour une exploitation commerciale, par exemple dans les situations de valorisation de la recherche. Dans tous les cas, un nouveau traitement devra être défini et les données à caractère personnel anonymisées. A priori, hors les situations de valorisation de la recherche accompagnées par les tutelles des laboratoires, la finalité des traitements de données ne peut avoir un objectif commercial. Des précautions doivent être prises et les conseils du Délégué à la protection des données, des services de valorisation de la recherche sont nécessaires.

ANNEXE 1 : EXEMPLE DE CONSENTEMENT

Proposé par l'UMR Pacte, Laboratoire de sciences sociales (CNRS, université Grenoble Alpes, Institut d'Etudes Politiques de Grenoble)

FORMULAIRE DE CONSENTEMENT DANS LE CADRE DE LA COLLECTE DE DONNEES PERSONNELLES

(*à remplir sur un papier à entête du laboratoire, en deux exemplaires à signer par le répondant. L'enquêteur remettra un exemplaire au répondant et l'autre au responsable du projet*).

Ce formulaire est destiné à recueillir votre consentement pour la collecte des données vous concernant, dans le cadre du projet *XXX* piloté par « *préciser équipe / laboratoire* ».

En signant le formulaire de consentement, vous certifiez :

- que vous avez lu et compris les renseignements communiqués dans la notice d'information,
- qu'on a répondu à vos questions de façon satisfaisante
- qu'on vous a informé que vous étiez libre d'annuler votre consentement ou de vous retirer de cette recherche en tout temps, sans préjudice.

Informations sur le participant :
Nom :
Prénom :
Adresse :

A remplir par le participant : *(à adapter selon le cas de figure)*
• J'ai lu et compris les renseignements fournis dans la fiche d'informations et j'accepte de plein gré de participer à cette recherche.
☒ OUI ☒ NON

Cas d'une enquête par entretien :
• J'accepte que mes propos soient enregistrés et exploités par l'équipe du projet XXX
☒ OUI ☒ NON
• J'accepte que mon image et mes propos soient filmés et exploités par l'équipe du projet XXX
☒ OUI ☒ NON
• J'accepte que mon image et mes propos soient diffusés dans le cadre de colloques scientifiques, séminaires ou dans toute forme de valorisation du projet XXX
☒ OUI ☒ NON

Cas d'une enquête par questionnaire :
• J'accepte que mes réponses aux questions posées soient exploitées par l'équipe du projet XXX
☒ OUI ☒ NON

Cas particuliers :
• J'accepte l'utilisation d'un système embarqué [ou d'objet connecté] pour collecter des données [géo-localisées, de pratiques de…] et que ces données [géo-localisées, de pratiques] soient exploitées par l'équipe du projet XXX
☒ OUI ☒ NON
• J'accepte que « les données sensibles » de type (énumérer les données concernées) soient collectées, conservées et exploitées par l'équipe du projet XXX.
☒ OUI ☒ NON
• J'accepte que mes données personnelles soient réutilisées dans le cadre de projets de recherche ayant les mêmes objectifs que celui du projet XXX.
☒ OUI ☒ NON

Nom, Prénom – Date – Signature

Un exemplaire de ce document vous est remis, un autre exemplaire est conservé dans le dossier.

ANNEXE 2 : EXEMPLE DE MENTION D'INFORMATION

Proposé par l'UMR Pacte, Laboratoire de sciences sociales (CNRS, université Grenoble Alpes, Institut d'Etudes Politiques de Grenoble)

NOTICE D'INFORMATION DANS LE CADRE DE LA COLLECTE DE DONNEES PERSONNELLES
(à remettre au participant sur un papier à entête du laboratoire).

(Vocabulaire à adapter à la cible : par exemple, les notices d'information destinées aux enfants doivent être rédigée dans un langage clair avec des mots utilisé par les enfants)

(Responsable du traitement)
Les informations recueillies [vous concernant] (si collecte directe) vont faire l'objet d'un traitement dans le cadre du projet *XXX* piloté par « *Nom, prénom, titre et rattachement institutionnel du responsable du projet, adresse postale et mail* »

(Si collecte indirecte par ex. données du web, préciser les personnes concernées)
Le traitement de données concerne :…

(Finalités du projet)
Le traitement a pour objet : « préciser l'objectif principal de la recherche et le cas échéant, détailler les sous-finalités ». Préciser ce qu'on attend de la personne…

Ex

Cas d'une enquête par entretien

Nous attendons de vous que vous participiez à un entretien durant lequel nous vous poserons des questions sur « *rappeler les finalités du projet* ». ….. L'entretien durera « ***préciser la durée*** ».
Préciser les modalités de la collecte :
Option 1 : L'entretien ne sera pas enregistré.
Option 2 : Les informations recueillies au cours de cet entretien font l'objet d'un enregistrement.
Option 3 : Les informations recueillies au cours de cet entretien font l'objet d'une prise vidéo / photo
Option 4 : Cas des parcours commentés avec collecte GPS : Le parcours que nous allons entreprendre fera l'objet d'un enregistrement par capteur GPS/GSM.

Cas d'une enquête par questionnaire

Nous attendons de vous que vous participiez à une enquête par questionnaire durant laquelle nous vous poserons des questions sur « *rappeler les finalités du projet* ». Le questionnaire durera « *préciser la durée de passation* ». *S'il s'agit d'une enquête longitudinale, préciser la durée de participation et les périodes de collecte.*

Recueils complémentaires (de type carnet de bord, traces GPS, avec objets connectés,…) :
A adapter selon les cas :
Nous souhaitons également vous faire remplir un carnet de bord pour connaître vos pratiques de « *préciser les types de pratiques* » pendant une durée de « *préciser la durée* ».
Nous souhaitons également utiliser un capteur GPS *(ou autre objet connecté)* pour comprendre vos pratiques de « *Préciser le type de pratique* » dans « *préciser le territoire* » pendant une durée de « *préciser la durée* ». Il vous est possible « *d'éteindre le GPS/ de désactiver l'objet connecté* » à tout moment.

(Nature des données collectées)
Seules les données strictement nécessaires à la réalisation de notre recherche seront collectées et traitées :
Lister les types de données personnelles collectées, par exemple :
Données d'identification
Données sur la vie personnelle (habitude de vie, situation familiale)
Données sur la vie professionnelle (CV, scolarité, formation, distinctions, publications….)
Données d'ordre économique et financière (revenu, situation financière)
Données de connexion (IP, logs…)

Données de localisation (données GSM, GPS…)
Données sensibles (opinions religieuses ou philosophiques, appartenances syndicales ou politique, orientation ou vie sexuelles, infractions ou condamnations, numéro de sécu, données de santé, biométriques ou génétiques)
Source des données (si collecte indirecte)
Ces informations sont recueillies auprès de (source à préciser et indiquer si elles sont issues ou non de sources accessibles au public) entre (mettre la période de collecte).

(Base légale du traitement)
La base légale du traitement repose sur :
A adapter selon les cas :

- l'exécution d'une mission de recherche publique => si projets financés sur fonds publics uniquement
- le consentement des participants => obligatoire si données sensibles, consentement conjoint de l'enfant et du représentant de l'autorité parentale si les enquêtés sont des mineurs de moins de 15 ans.

(Participation libre)
Votre participation au projet « préciser le nom du projet » est entièrement libre et volontaire.

(Retrait du consentement)
Vous êtes libre de vous retirer ou de cesser votre participation à ce projet à tout moment. Ce retrait n'aura aucune conséquence.
[Si l'on travaille avec des étudiants, des élèves on peut préciser que le retrait n'aura aucune incidence sur la suite de leurs études].
[Cas de collecte de données longitudinales : Dans le cas de collecte de données sur plusieurs périodes, le retrait du consentement sera effectif à partir de la date où il a été reçu par le responsable de traitement.]

(Pseudonymisation/ confidentialité)
Cas d'une enquête par entretien
Le projet « *préciser le nom du projet* » prend les engagements suivants :

- Votre identité sera dissimulée à l'aide d'un numéro aléatoire dans tous les écrits produits sur la base de vos propos (comptes rendus d'entretien, notes d'observation, notes d'analyse échangées entre les chercheurs, publications…).
- Aucune autre information ne sera conservée qui puisse révéler votre identité : les notes d'entretien, comptes rendus d'entretien, notes d'observation, notes d'analyses et publications seront complètement anonymisés.

Cas d'une enquête par questionnaire

- Votre identité sera dissimulée à l'aide d'un numéro aléatoire pour tous les types d'informations collectées (*liste à adapter selon le projet : questionnaires, données GPS, carnet de bord,…*).
- Seul le responsable de projet détient la table de correspondance qui permet de faire le lien entre votre identité et le numéro aléatoire attribué dans les différents fichiers (*liste à adapter selon le projet questionnaires, données GPS, carnet de bord,…*).

(Destinataires des données personnelles)
Le destinataire ou catégories de destinataires de ces données sont : « *indiquer qui a besoin d'y accéder ou de les recevoir selon les finalités définies ; préciser noms des organismes, partenaires, services….* »

(Transferts de données)
Option 1 : Toutes les données seront gardées en France ;
Option 2 : / Les données recueillies seront transférées / conservées par un des partenaires du projet dans un pays de l'Union Européenne, ce pays est soumis aux mêmes règles de protection de la vie privée que la France.
Option 3 : Les données recueillies seront transférées / conservées par un des partenaires du projet dans un pays hors Union Européenne. Le transfert est fondé sur des clauses contractuelles type de la Commission européennes ou encadré par des clauses contractuelles spécifiques, un code de conduite, une certification…etc approuvés

(Durée de conservation)
Vos données personnelles sont conservées en base active jusqu'à « préciser la date ou la durée ».
Option 1 : Après cette date, elles seront définitivement archivées (si intérêt scientifique, statistique ou historique important)
Option 2 : Après cette date, elles seront définitivement archivées de manière anonymisée (si pas d'intérêt à garder les

données personnelles).

(Mesure de sécurité)
Afin de garantir la confidentialité de vos données et éviter leur divulgation, les dispositifs suivants ont été mis en place :

- Seuls les services « préciser lesquels » sont autorisés à accéder aux données.
- (Le cas échéant) Le prestataire externe « préciser son rôle » est soumis à des garanties contractuelles protégeant vos données.
- Les mesures de sécurité, tant physique que logique, suivantes sont prises. (par ex : Protection anti-incendie, copies de sauvegarde, installation de logiciel antivirus, changement fréquent des mots de passe alphanumériques d'un minimum de 8 caractères, chiffrement des ordinateurs)

(Diffusion)
Les résultats de cette recherche seront diffusés de façon anonyme dans des colloques professionnels et scientifiques, dans des rapports destinés aux autorités, dans des revues professionnelles et académiques et dans des médias destinés au grand public (liste à adapter selon le projet).

(Droits des personnes)
Vous pouvez poser des questions au sujet de ce projet à tout moment en communiquant avec le responsable du projet par courrier électronique: « *préciser l'adresse*»
Vous pouvez accéder et obtenir copie des données vous concernant, vous opposer au traitement de ces données, les faire rectifier ou les faire effacer. Vous disposez également d'un droit à la limitation du traitement de vos données. Vous pouvez exercer ces droits en vous adressant à : « *indiquer les coordonnées du service ou de la personne chargé du droit d'accès – adresses postales et électroniques* ».
Vous pouvez contacter également le Délégué à la Protection des Données du laboratoire Pacte à l'adresse suivante : DPD – 17 rue Notre Dame des Pauvres – 54519 – Vandoeuvre lès Nancy Cedex -dpd.demandes@cnrs.fr

Après nous avoir contactés, si vous estimez que vos droits Informatique et Libertés ne sont pas respectés, vous avez la possibilité d'introduire une réclamation en ligne auprès de la CNIL ou par courrier postal. CNIL, 3 Place de Fontenoy, TSA 80715 – 75334 Paris Cedex 07 (https://www.cnil.fr/)

ANNEXE 3

Les principales questions en vue de la conformité à la règlementation sur la protection des données personnelles

Situation 1 :

Les données utilisées pour la recherche sont des données anonymes (non identifiantes et ceci de manière irréversible)	La règlementation sur la protection des données personnelles ne s'applique pas

Situation 2 :

Les données utilisées pour la recherche sont des données à caractère personnel : application de la règlementation avec des aménagements lorsque les traitements sont à des fins de recherche

		Voir pages
Qui est responsable ?	• Le responsable de traitement • Le sous-traitant • Les partenaires, les co-responsables de traitement	12, 27
Quelles sont les données	• Données non sensibles • Données dites sensibles • Numéro de sécurité sociale • Données d'infractions et de condamnations • Données des populations vulnérables	11, 21
Comment sont collectées les données, quels sont les destinataires ?	• Collecte directe • Collecte indirecte	19
Quelle est la finalité du traitement ?		23
Les données correspondent-elles à la finalité et sont-elles suffisantes pour le projet ?	Principes relatifs au traitement de données	15, 23
Combien de temps sont conservées les données ?		26
Comment sont informées les personnes ? quels droits ont les personnes ?		16
Quelles dispositions sont requises pour assurer la confidentialité et la sécurité des données ?	• Hébergement • Stockage • Partage de fichiers, de dossiers	24, 27
Une étude d'impact sur la vie privée est-elle nécessaire ?	• logiciel open source PIA de la CNIL	16

Quelles démarches doivent être réalisées ? Auprès de qui doivent-elles être réalisées ?	• Le registre des traitements • Le Délégué à la protection des données • La CNIL	12 13
Existe-t-il des dispositifs, des aides méthodologiques, des services pour la communauté scientifique ?	• Les techniques d'anonymisation, de pseudonomysation • Les offres d'hébergement des données, d'archivage	25, 26
Quelles données puis-je publier ?		27
Les données peuvent-elles être réutilisées ?		28

ANNEXE 4

LISTE DES SIGLES

ANR : Agence Nationale de la Recherche

CASD : Centre d'Accès Sécurisé aux Données

CEREES : Comité d'Expertises pour les Recherches, les Etudes et les Evaluations dans le domaine de la Santé

CINES : Centre Informatique National de l'Enseignement Supérieur

CNIL : Commission Nationale de l'Informatique et des Libertés

CPP : Comité de Protection des Personnes

CNRS : Centre National de la Recherche Scientifique

DPD : Délégué à la Protection des Données

INDS : Institut National des Données de Santé

INIST : Institut de l'Information Scientifique et Technique

INRA : Institut National de la Recherche Agronomique

MR : Méthodologie de Recherche

OPIDOR : Optimisation du Partage et de l'Interopérabilité des Données de la Recherche (portail mis en place et hébergé par le CNRS – INIST)

Programme H2020 : programme Horizon 2020

PUD : Plateformes universitaires des Données

TGIR : Très Grande Infrastructure de Recherche

TGIR Huma-Num : TGIR des Humanités Numériques

TGIR Progedo : TGIR Production et Gestion des Données en Sciences Sociales

UE : Union européenne

Photo de couverture : Visualisation interactive des thèmes relatifs au changement climatique générée par le Tweetoscope climatique.

INSTITUT DES SCIENCES HUMAINES ET SOCIALES
3, rue Michel-Ange 75016 Paris
www.inshs.cnrs.fr
@INSHS_CNRS
Abonnez-vous à la lettre d'information de l'InSHS :
inshs.com@cnrs.fr

Mise en page : InSHS Communication

Scannez le code pour nous rejoindre sur Twitter en express :