Quinze recommandations aux chercheurs

sur la protection des données dans le cadre de leurs activités de recherche

Synthèse :

Le réseau SupDPO propose à la communauté des chercheurs **15 recommandations pratiques** pour les aider dans la mise en conformité de leurs projets de recherche.

Les recherches scientifiques mobilisent des matériaux qui contiennent fréquemment des données personnelles, et qui, à ce titre, appellent des précautions particulières. Pour ne pas qu'elle soit un frein, **la réglementation en vigueur doit être anticipée au maximum** dans les projets de recherche.

Les pratiques de la recherche sont depuis toujours fondées sur l'utilisation, le partage, la réutilisation des documents et données. Les chercheurs sont soumis à des injonctions qui peuvent paraître contradictoires entre les volontés politiques d'ouverture des données de la recherche d'un côté et la nécessité de la protection des données personnelles de l'autre.

La prise en compte des données personnelles dans les projets de recherche permet d'en augmenter la qualité et de favoriser une éthique qui permet le progrès des connaissances tout en garantissant les droits fondamentaux des personnes qui confient leurs données aux chercheurs.

1. **Associer le Délégué à la Protection des Données** (DPD ou DPO) le plus en amont possible de la mise en œuvre du projet de recherche ; cela permet de porter des projets qui intègrent le principe de respect de la vie privée dès la conception (*privacy by design)*.

2. **Saisir dès que possible le comité d'éthique (ou comité d'éthique recherche) de l'établissement**, car les textes sont soumis à interprétation. Le comité d'éthique devrait avoir un rôle d'aiguillage des projets de recherche, ce qui permet ensuite d'avoir un parcours plus fluide pour le chercheur.

3. **Dès le début du projet de recherche, définir les objectifs du projet,** les buts poursuivis et s'assurer que le traitement des données personnelles se fasse en accord avec ceux-ci, dans un principe de minimisation des données.

4. Prendre en compte le fait que **la mise en conformité d'un projet de recherche peut être longue** et il convient de l'anticiper au maximum. Il est notamment nécessaire que le *rétroplanning* tienne compte des délais règlementaires et institutionnels (CPP, délais de procédures auprès de l'INDS, du CEREES ou encore de la CNIL, Comités d'éthique pour la recherche (CER) marchés publics, financements européens FEDER, analyse d'impact des traitements sur la vie privée, etc.).

Document élaboré dans le cadre du réseau des Délégués à la protection des données de l'Enseignement supérieur et de la recherche (SupDPO) en partenariat avec le Réseau des responsables recherche (R3Sup).
Ce document n'a pas été soumis ni validé par la Commission nationale de l'informatique et libertés (CNIL).

5. Prendre en compte le fait qu'**avant publication**, **les éditeurs vérifient de plus en plus la conformité au RGPD**. Cette dernière permet donc d'éviter les déconvenues lors de la soumission des articles à des revues.

6. La mise en conformité au RGPD d'un projet de recherche passe par plusieurs étapes de vérification de la protection des personnes et protection des données. Un protocole de recherche conforme aux règles, permet de passer toutes les étapes sur une seule et même base. Le protocole peut même servir de fondement pour **la réalisation du plan de gestion des données** (*data management plan*, DMP[1]), obligatoire dans les projets à financements européens et ANR (à partir de 2020). Si le protocole évolue sensiblement au cours du projet, il est nécessaire d'en faire une mise à jour en lien avec le DPO.

7. **Les captations audio ou vidéo** (entretiens, notamment) sont des traitements de données personnelles (voix, image) et nécessitent d'obtenir le consentement préalable des personnes concernées (ou de leurs tuteurs concernant les mineurs). Le consentement doit être éclairé et donné pour l'utilisation dans le cadre de la recherche. Il doit prévoir les utilisations ou réutilisations ultérieures, le cas échéant. Le consentement donné au début de l'étude peut être retiré à tout moment par la personne et **des processus doivent donc être définis pour gérer ces consentements** au cours du projet de recherche.

8. **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* ». La définition d'une donnée de santé ne doit donc pas être minimisée et la pratique montre que les données de santé ne sont pas uniquement utilisées par les sciences médicales et odontologiques et toutes les disciplines de recherche peuvent y avoir recours (et en premier lieu les sciences humaines et sociales).

   ⇨ La CNIL a rédigé une page d'aide à la définition de données de santé[2] ;

9. **L'hébergement des données de santé** est un domaine spécifique et sensible, qui peut être soumis à une certification[3]. A ce titre, il faut distinguer notamment la sous-traitance (dans le cadre où un prestataire assure l'hébergement des données de santé, par exemple) et la co-responsabilité (dans le cadre où les données sont hébergées par l'un des partenaires du projet, qui en détermine les modalités). Dans les deux cas, un accord doit être signé des deux parties pour définir les obligations respectives.

10. **L'anonymisation des données personnelles** permet de s'affranchir des règles prévues par le RGPD. Néanmoins, il s'agit d'un processus irréversible ne permettant plus de ré-identifier les personnes concernées par quelque moyen que ce soit, dont l'efficacité est discutée[4]. Elle doit en tout état de cause être réalisée dans les règles de l'art (exemples d'utilisation de techniques d'anonymisation des données : tranches d'âge plutôt qu'âge précis, cohortes plutôt qu'individus, localisation plutôt qu'adresse, etc.).

    ⇨ Attention ! **Une donnée pseudonymisée ou codée n'est pas anonyme** et donc constitue une donnée personnelle indirectement identifiante soumise au respect des règles de sécurité et de protection des données. La pseudonymisation doit être vue comme une mesure de sécurité.

[1] https://doranum.fr/wp-content/uploads/FicheSynthDMP.pdf
[2] https://www.cnil.fr/fr/quest-ce-ce-quune-donnee-de-sante
[3] https://esante.gouv.fr/labels-certifications/hebergement-des-donnees-de-sante
[4] Rocher, L., Hendrickx, J.M. & de Montjoye, Y. (2019) https://www.nature.com/articles/s41467-019-10933-3

11. **L'ouverture des données (*OpenData*)** doit se faire dans le respect de la protection des données personnelles (« *ouvrir autant que possible, fermer autant que nécessaire* »[5]).

    ⇨ Un Guide pratique de la publication en ligne et de la réutilisation des données publiques élaboré par les services de la CNIL et de la CADA en association avec les services d'Etalab[6].

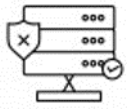
    ⇨ Un Guide d'analyse du cadre juridique en France sur l'ouverture des données de recherche (décembre 2017)[7].

12. **Les établissements proposent des services d'hébergement ou de travail collaboratif qui assurent la sécurité, la confidentialité et le respect de la vie privée**. Utilisez-les de préférence, conformément aux chartes informatiques de vos établissements.

13. **L'hygiène numérique** participe de la protection des données personnelles (bonnes pratiques : définition de mots de passe forts, chiffrement des supports physiques, utilisation d'écrans de confidentialité ; pratiques dommageables : utilisation du Wi-Fi public, partage de données sans gestion des habilitations,…).

    ⇨ l'ANSSI propose une documentation riche[8] : voir notamment leur guide d'hygiène informatique[9] et recommandations de sécurité relative aux mots de passe[10].

14. **Attention à l'usage des adresses de courriels personnels dans le cadre professionnel**. Ne pas mêler les usages publics / privés (notamment les mails professionnels envoyés depuis une adresse personnelle). Privilégiez toujours l'utilisation des adresses institutionnelles dans le cadre des projets de recherche (ex : *nom_projet@nom_etablissement.fr*).



15. **Se former et s'informer** permet de gagner en autonomie et de faciliter les démarches (atelier de la CNIL[11], MOOC du CNAM[12], etc.).

[5] https://education.newstank.fr/fr/tour/news/135985/science-ouverte-dacos-detaille-strategie-entrainer-tous-acteurs-esr.html
[6] https://www.cnil.fr/sites/default/files/atoms/files/guide_open_data.pdf
[7] https://www.ouvrirlascience.fr/wp-content/uploads/2018/11/Guide_Juridique_V2.pdf
[8] https://www.ssi.gouv.fr/administration/bonnes-pratiques/
[9] https://www.ssi.gouv.fr/administration/guide/guide-dhygiene-informatique/
[10] https://www.ssi.gouv.fr/administration/guide/mot-de-passe/
[11] https://atelier-rgpd.cnil.fr/
[12] https://www.fun-mooc.fr/courses/course-v1:CNAM+01032+session01/about