

Agnès Robin

Maître de conférences HDR en Droit privé à l'Université de Montpellier

Directrice adjointe du Laboratoire Innovation, Communication et Marché (LICeM)

Directrice du Master Droit de la propriété intellectuelle et du numérique de la Faculté de Droit

Responsable du DU Scientifique Data Management de l'Institut de Science des Données de Montpellier (ISDM)

Chargée de mission auprès du ministère de l'Enseignement supérieur et de la recherche (MESR)





Comment l'IA générative transforme-t-elle les pratiques de recherche ? Nouveaux enjeux d'intégrité scientifique

Le RIA : quelles conséquences pour les activités de recherche, quelles implications sur les données de recherche ?

15 avril 2025

- I. RIA et activités de recherche
- II. Gouvernance des données de la recherche
- III. Droit des données de la recherche

I. RIA et activités de recherche

Art. 3, 1° RIA

« Système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels »

+ Lignes directrices sur la définition des SIA, 6 févr. 2025

SIA comme outils de recherche

SIA comme produits de recherche

I. RIA et activités de recherche

A. SIA comme outils de recherche

Textes légaux

RIA

Considérant 25

« Toute activité de recherche et de développement devrait être menée conformément à des normes éthiques et professionnelles reconnues en matière de recherche scientifique et dans le respect du droit de l'Union applicable »

Textes déontologiques

Code de conduite européen (v. 2023)

2.3 Procédures de recherche

« Les chercheurs font connaître leurs résultats et leurs méthodes, y compris l'utilisation de services externes, d'outils d'intelligence artificielle ou automatisés, sous une forme qui respecte les normes faisant consensus de la discipline, et qui facilite, le cas échéant, la vérification et la reproductibilité »

2.8 Examen par les pairs (*reviewing*) et évaluation

« Les chercheurs, les institutions scientifiques et les établissements assurant une mission de recherche examinent et évaluent les demandes de publication, financement, nomination, promotion ou récompense d'une **manière transparente** et justifiable, et indiquent l'utilisation qu'ils ont faite d'outils d'IA et d'outils automatisés »

3.1 Manquement à l'intégrité scientifique et autres pratiques inacceptables

Cacher l'utilisation de l'IA ou d'outils automatisés dans la création de contenu ou la rédaction de publications

I. RIA et activités de recherche

B. SIA comme produits de recherche

Textes légaux

RIA

Exclusion de la recherche

Considérant 142

Dans la mesure où la recherche aboutissant à la production de SIA peut conduire à **l'obtention de résultats bénéfiques** « sur le plan social et environnemental, les États sont **encouragés à soutenir et à promouvoir la recherche et le développement de solutions d'IA propices à tels résultats**, telles que des solutions fondées sur l'IA destinées à renforcer l'accessibilité pour les personnes handicapées, à réduire les inégalités socio-économiques ou à atteindre les objectifs environnementaux, en y affectant des ressources suffisantes, y compris des financements publics et de l'Union, et, lorsqu'il convient et pour autant que les critères d'éligibilité et de sélection soient remplis, en envisageant des projets spécifiques qui poursuivent ces objectifs »

Art. 2, 6°

« Le présent règlement ne s'applique pas aux systèmes d'IA ou aux modèles d'IA **spécifiquement** développés et mis en service **uniquement** à des fins de recherche et développement scientifiques, ni à leurs sorties »

Immunité scientifique... mais tout est dans l'adverbe « spécifiquement »

Art. 2, 8°

« Le présent règlement ne s'applique pas aux activités de recherche, d'essai et de développement relatives aux systèmes d'IA ou modèles d'IA **avant leur mise sur le marché ou leur mise en service**. Ces activités sont menées conformément au droit de l'Union applicable (...) »

Limites à l'exclusion

Art. 2, 8°

« (...) Les essais en conditions réelles ne sont pas couverts par cette exclusion »

Comm. UE, Lignes directrices sur les pratiques interdites, 4 févr. 2025

« Une municipalité qui souhaite tester un logiciel de reconnaissance faciale à l'aide d'un système RBI dans les rues pendant le carnaval recrute des volontaires pour être identifiés par le système dans des conditions réelles. Étant donné que les tests en conditions réelles ne relèvent pas de l'exclusion de l'article 2, paragraphe 8, de la loi sur l'IA, les tests prévus doivent être entièrement conformes aux exigences de la loi sur l'IA pour les systèmes RBI, à moins que le système ne soit testé dans un bac à sable réglementaire de l'IA ou conformément au régime spécial pour les tests dans des conditions réelles en dehors du bac à sable, comme prévu aux articles 60 et 61 de la loi sur l'IA »

Limites à l'exclusion (suite)

Considérant 25

SIA conçus dans le cadre d'une R-D et destinés au marché

Considérant 97

Modèles d'IA à usage général dans le cadre d'une R-D et destinés au marché

Considérant 139 + art. 57

« bacs à sable réglementaires »

II. Gouvernance des données de la recherche

Quid ?

Gouvernance = gouvernement / règles

Gouvernance = gestion



Le puzzle juridique de l'IA médicale DOSSIER

Le Règlement sur l'IA et la gouvernance des données ¹

Margo Bernelin
Chargée de recherche CNRS
UMR 6297 Droit et changement social, Nantes Université

Cela ne fait guère plus mystère, les systèmes d'intelligence artificielle (SIA) peuvent être gourmands en données, certains l'étant plus que d'autres. Ces données sont utilisées pour le développement des outils, leur entraînement, validation, test et mise en fonction, à des degrés divers selon le cycle de vie du système d'IA considéré. Dans le domaine de la santé, les systèmes d'IA traitent des données scientifiques, à l'image de publications pertinentes et autres données acquises de la science, mais également des données personnelles de santé afin de fournir une prédiction, un résultat, un score pour un patient ou un groupe de patients. Les données ainsi traitées sont diverses à la fois dans leurs sources mais également formats, contenus ainsi que dans les valeurs ou informations qu'elles peuvent exprimer. En ce sens, elles sont

soit bien identifiées et le droit en la matière tente, depuis un peu moins de dix ans maintenant, de les soutenir en élargissant le périmètre et le nombre des bases de données personnelles en santé disponibles. Quelques faits marquants jalonnent ce parcours dont voici les morceaux choisis :

2016 : création du système national des données de santé (SNDS) pour la mise à disposition de six grandes bases de données de santé dont celle de l'assurance maladie sur les feuilles de soin ² ;

2018 : lancement d'une politique en faveur du déploiement de l'IA en France avec la publication du rapport *Donner un sens à l'intelligence artificielle* de Cédric Villani ³ où la santé est considérée comme un secteur prioritaire

RDSS Sept.-oct. 2024

II. Gouvernance des données de la recherche

A. Fiabilité et conservation

> Fiabilité et robustesse des données

Textes légaux

Art. L. 211-2 C. rech.

« les travaux de recherche, notamment l'ensemble des activités de la recherche publique contribuant à ses objectifs mentionnés à l'article L. 112-1, respectent les exigences de **l'intégrité scientifique** visant à garantir leur caractère **honnête** et **scientifiquement rigoureux** et à consolider le lien de confiance avec la société »

Textes déontologiques

Code de conduite européen pour l'intégrité en recherche (v. 2023) (art. 1^{er})

« Les bonnes pratiques de recherche reposent sur des principes fondamentaux d'intégrité scientifiques. Ces principes servent de guide aux chercheurs, aux institutions et aux organisations dans leurs travaux ainsi que dans leur engagement envers les enjeux pratiques, éthiques et intellectuels inhérents à la recherche.

Ces principes comprennent :

- La **fiabilité** dans la conception, la méthodologie, l'analyse et l'utilisation des ressources.
- **L'honnêteté** dans l'élaboration, la réalisation, l'évaluation et la diffusion de la recherche, d'une manière transparente, juste, complète et objective.
- Le **respect** envers les collègues, les participants à la recherche, la société, les écosystèmes, l'héritage culturel et l'environnement.
- La **responsabilité** dans l'exercice d'activités de recherche, de l'idée à la publication, dans leur gestion et leur organisation, dans la formation, l'encadrement et le mentorat, et dans les implications plus générales de la recherche.

> Conservation et accessibilité des données

Textes légaux

Art. L. 211-2 C. rech.

« (...) Les établissements publics contribuant au service public de la recherche et les fondations reconnues d'utilité publique ayant pour activité principale la recherche publique au sens du même article L. 112-1 offrent les conditions du respect des exigences de l'intégrité scientifique pour les activités et travaux menés en leur sein. Ils mettent en place les dispositifs nécessaires pour promouvoir les valeurs de l'intégrité scientifique et favoriser le respect de ses exigences. Sans préjudice des dispositions du code du patrimoine sur les archives publiques, ils **conservent les résultats bruts des travaux scientifiques** réalisés en leur sein afin de permettre leur vérification »

Textes déontologiques

Code de conduite européen pour l'intégrité en recherche (v. 2023) (art. 2.5)

« Les chercheurs, les institutions scientifiques et es établissement assurant une mission de recherche assurent de façon adéquate la **gestion**, la **conservation** et la **préservation** sécurisée de toutes les données, métadonnées, protocoles, codes, logiciels et de tous autres matériels de recherche pendant une période raisonnable et clairement définie »

« les chercheurs, institutions et organismes de recherche reconnaissent que les données, métadonnées, protocoles, codes, logiciels et autres matériels de recherche sont des **produits légitimes de la recherche**, qui peuvent être cités »

II. Gouvernance des données de la recherche

B. Ethique par conception (*ethics by design*)

design des données (FAIR)

design des codes source et algorithmes

design des infrastructures (archives ouvertes)

design des licences d'utilisation

design des contrats d'édition

design des contrats de collaboration de recherche



formation
information

III. Droit des données de la recherche

A. Principe général : l'ouverture

L'accès aux données doit être aussi ouvert que possible, pas plus fermé que nécessaire

Textes légaux

Union européenne

Dir. PSI (2019) + Règl. exécution (2022)

Communication Comm. UE (2012)

Prop. Règl. sur l'espace européen des données de santé (3 mai 2022)

France

Loi 22 juill. 2013 dite « Loi Fioraso » (loi sur la recherche)

Loi du 28 déc. 2015 dite « Loi Valter » (loi sur les informations publiques)

Loi du 7 octobre 2016 (loi pour une République numérique)

Définition des données de recherche

Directive PSI (art. 2, 7°)

« documents se présentant sous forme numérique autres que des publications scientifiques, qui sont **recueillis ou produits** au cours d'activités de recherche scientifique et utilisés comme **éléments probants** dans le processus de recherche, ou dont la communauté scientifique admet communément qu'ils sont **nécessaires pour valider** des conclusions et résultats de la recherche »

Régime juridique des données de la recherche

Directive PSI (art. 10)

« 1. Les États membres encouragent la mise à disposition des données de la recherche en adoptant les politiques et en prenant les mesures nécessaires à l'échelon national afin de rendre librement accessibles les **données résultant de la recherche financée au moyen de fonds publics** («politiques de libre accès») qui respectent le principe d'ouverture par défaut et sont compatibles avec les principes FAIR. Dans ce contexte, il y a lieu de tenir compte des préoccupations liées aux **droits de propriété intellectuelle, à la protection des données à caractère personnel et à la confidentialité, à la sécurité et aux intérêts commerciaux légitimes** dans le respect du principe « aussi ouvert que possible, mais aussi fermé que nécessaire ». Ces politiques de libre accès visent les organismes exerçant une activité de recherche et les organisations finançant une activité de recherche »

Art. L. 112-1 C. rech.

La recherche publique a pour objectifs :

- a) Le développement et le progrès de la recherche dans tous les domaines de la connaissance ;
- b) La valorisation des résultats de la recherche au service de la société, qui s'appuie sur l'innovation et le transfert de technologie ;
- c) Le partage et la diffusion des connaissances scientifiques en donnant priorité aux formats libres d'accès ;
- c bis) Le développement d'une capacité d'expertise et d'appui aux associations et fondations, reconnues d'utilité publique, et aux politiques publiques menées pour répondre aux défis sociétaux, aux besoins sociaux, économiques et du développement durable ;
- d) La formation à la recherche et par la recherche ;
- e) L'organisation de l'accès libre aux données scientifiques.**

III. Droit des données de la recherche

A. Exception : la fermeture

Principe d'articulation

L'accès aux données doit être aussi ouvert que possible, pas plus fermé que nécessaire

15 L'ouverture des données publiques scientifiques : de l'examen de la règle « *open as possible, closed as necessary* »



Agnès ROBIN,

maître de conférences HDR à l'université de Montpellier (ERCIM/UMR 5835),
directrice du Master 2 Droit de la propriété intellectuelle et TIC de la faculté de Droit de Montpellier,
chargée de mission à l'institut de la science des données de Montpellier,
responsable du DU « Scientific Data Management »

L'étude se propose de revenir sur le principe d'articulation qui préside à la politique européenne et nationale d'ouverture des données de la recherche scientifique « *open as possible, closed as necessary* ». La mise en œuvre de cette politique balance ainsi entre le respect de l'intérêt public qui commande la diffusion ouverte des données (diffusion) et la prise en compte des intérêts privés fondés sur le secret ou les droits de propriété intellectuelle (valorisation), y compris lorsque les résultats de la recherche sont majoritairement issus d'un financement public.

Qu'est-ce qui fait nécessité ?

Tout ce qui préempte les données en raison d'intérêts légitimement protégés

=

secrets

propriété intellectuelle

protection de la biodiversité

essais cliniques

données personnelles

SECRETS

Secret statistique

Secret d'affaires

Secret de la défense nationale

Secrets professionnels (santé notamment)

Potentiel scientifique et technique de la Nation (PPSTN)

Secret sur les données des essais cliniques

Contrats (partenariat public/privé)

etc.

DROIT DES DONNÉES DE LA RECHERCHE

Science ouverte, innovation, données publiques

Agnès Robin

Création Information Communication

 LORCIER



DU SCIENTIFIC DATA MANAGEMENT GESTION DES DONNÉES DE LA RECHERCHE



Responsable : Agnès Robin
Maître de conférences HDR, Université de Montpellier

Inscriptions :
<https://sdm.edu.umontpellier.fr/>



Merci de votre attention