Amplitude du droit
ISSN : 2826-1305
Publisher : Université de Rennes

4 | 2025

La science ouverte appliquée aux ressources de la recherche scientifique sur le droit

Agnès Robin

https://lodelpreprod.univ-rennes2.fr/amplitude-droit/index.php?id=757

DOI : 10.56078/amplitude-droit.757

Electronic reference
Agnès Robin, « La science ouverte appliquée aux ressources de la recherche scientifique sur le droit », *Amplitude du droit* [Online], 4 | 2025, Online since 17 January 2025, connection on 22 September 2026. URL : https://lodelpreprod.univ-rennes2.fr/amplitude-droit/index.php?id=757

Copyright
Licence Creative Commons – Attribution 4.0 International – CC BY 4.0

La science ouverte appliquée aux ressources de la recherche scientifique sur le droit

Agnès Robin

OUTLINE

Introduction
1. L'ouverture de l'accès aux écrits scientifiques utilisés ou produits lors d'une activité de recherche juridique
  1.1. L'accès aux écrits scientifiques assuré par les éditeurs scientifiques
    1.1.1. L'accès restreint aux écrits scientifiques
    1.1.2. L'accès limité aux données associées aux écrits scientifiques
  1.2. L'accès aux écrits scientifiques renforcé par la science ouverte
    1.2.1 La consécration légale d'une dérogation d'ordre public
    1.2.2. La recommandation d'une diffusion immédiate du manuscrit
2. L'ouverture de l'accès aux données utilisées ou produites lors d'une activité de recherche juridique
  2.1. L'accès aux données publiques
    2.1.1. Les données utilisées et produites lors d'une recherche juridique, des données publiques
    2.1.2. Les données utilisées ou produites lors d'une activité de recherche juridique, des données ouvertes ?
  2.2. L'accès aux données privatives pour la recherche juridique
    2.2.1. Les bases de données couvertes par un droit de propriété intellectuelle
    2.2.2. Les pratiques, usages et savoir-faire couverts par le secret
Conclusion

TEXT

Introduction

1 Ce que l'on nomme « science ouverte » est le fruit d'une politique, initialement portée par l'Union européenne[1], d'ouverture de la production scientifique. Elle s'appuie sur l'opportunité que représente la mutation numérique pour encourager la diffusion la plus large possible des résultats, des méthodes ou des produits scientifiques. Reprise par certaines instances internationales (Unesco, 2021), cette politique s'est traduite en France par l'adoption de règles

légales, suivies de l'affirmation d'un cadre général, le Plan national pour la science ouverte (PNSO) [Ministère de l'Enseignement supérieur et de la Recherche, 2018, 2021]. Les règles résultent de deux lois principales. La première est la loi de programmation de la recherche du 22 juillet 2013[2] qui prévoit que « la recherche publique a pour objectifs : […] c) Le partage et la diffusion des connaissances scientifiques en donnant priorité aux formats libres d'accès ; […] e) L'organisation de l'accès libre aux données scientifiques ». La seconde est la loi pour une République numérique du 7 octobre 2016[3] dont l'article 30 contient le principe de l'ouverture de l'accès aux publications scientifiques, ainsi que celui de l'ouverture de l'accès aux données de la recherche[4]. Sur le plan européen, le règlement « Horizon Europe » définit la science ouverte comme « une approche du processus scientifique fondée sur le travail et des outils coopératifs ouverts et diffusant des connaissances[5] », comprenant les deux éléments suivants : l'accès ouvert aux publications scientifiques résultant de la recherche financée au titre du programme Horizon Europe et l'accès ouvert aux données de la recherche, y compris celles sur lesquelles s'appuient ces publications scientifiques, conformément au principe selon lequel les données de la recherche doivent être « aussi ouvertes que possible, aussi fermées que nécessaire[6] ». Formulé par le règlement pour encadrer la gestion des projets de recherche financés par des fonds européens, le principe de science ouverte fut ensuite repris, pour ce qui concerne les seules données de la recherche, par la directive (UE) 2019/1024 dans le cadre du dispositif légal relatif à l'ouverture des données publiques[7].

2 Au-delà de la démocratisation de l'accès aux connaissances, c'est un effort d'innovation que les pouvoirs publics entendent encourager et soutenir. Concernant les données de la recherche tout d'abord, la politique d'innovation de l'Union européenne repose sur l'idée qu'ouvrir les données de la recherche permettra de favoriser la mutualisation, le croisement et l'appariement des données pour produire de nouvelles connaissances et des innovations fondées sur l'utilisation de systèmes d'intelligence artificielle (Robin, 2022a). La science ouverte doit permettre de favoriser la réutilisation éventuelle par les pairs et/ou par le public (entreprises, associations, professions libérales, etc.) de ces données, afin que puissent être créées des solutions innovantes, éventuellement fondées sur des technologies

d'intelligence artificielle (Robin, 2022a). C'est en effet en considération des potentialités de réutilisation que les données de la recherche doivent être techniquement conçues dès le départ comme des « données ouvertes » (*open bydesign* ou *open by default*)[8]. Concernant les publications scientifiques ensuite, il a été observé que les pouvoirs publics contribuaient à l'économie de l'édition scientifique à plusieurs endroits du processus de recherche scientifique : par la rémunération des chercheurs[9], par le paiement des frais de traitement éditorial des articles (*Article Processing Charges* ou APC) et par l'abonnement aux revues et aux catalogues des éditeurs scientifiques. De même, la forte concentration du marché de l'édition scientifique autour de grands groupes éditoriaux proposant des bouquets de revues incontournables pour les chercheurs[10], la transformation du rôle de l'éditeur dont la prestation purement éditoriale se réduit en même temps que croît sa fonction de distributeur *via* des plateformes multiservices[11], le développement de la technologie numérique dans la diffusion d'informations ouvrant de nouvelles voies de publication pour les établissements publics de recherche, mais surtout l'augmentation des coûts d'abonnement aux revues scientifiques supportés par les établissements (Monvoisin, 2012)[12] ont fait souffler un vent de contestation important au début des années 2000, à l'initiative des documentalistes et bibliothécaires qui ont activement encouragé les chercheurs à publier leurs productions (articles, ouvrages, mémoires et thèses) *via* Internet. C'est officiellement à travers les déclarations des « trois B » (Budapest, 2002[13] ; Bethesda, 2003[14] ; Berlin, 2003[15]), suivies par l'Appel de Jussieu pour la science ouverte et la bibliodiversité de 2017, que se sont dessinées les véritables injonctions de mettre la ressource scientifique en accès libre et ouvert, et que la recommandation de la Commission européenne de 2012, ainsi que les dispositions de la loi pour une République numérique ont été entérinées.

3 L'étude de la science ouverte oblige dès lors à considérer, de façon réflexive, l'ensemble des processus de production scientifique, de même que l'environnement socio-économique dans lequel les recherches sont menées. S'agissant du processus de production scientifique, la perspective de mise à disposition des données de la recherche suppose, par exemple, généralement que les données aient été collectées, structurées, traitées, conservées, sécurisées, et ainsi

gérées par les chercheurs, selon des règles qui permettent ensuite à d'autres de les réutiliser. Cela impose ensuite de vérifier, sur le plan juridique par exemple, si la mise à disposition des données peut être réalisée sans qu'elle porte atteinte aux droits des personnes (droit des données à caractère personnel par exemple) ou aux droits sur les biens (droit de la propriété intellectuelle par exemple)[16]. L'étude de la science ouverte appliquée à la recherche sur le droit revient à s'interroger non seulement sur les pratiques et techniques de production scientifique (en et) sur le droit[17], mais aussi sur les modalités de diffusion des productions scientifiques elles-mêmes, pour en mesurer le degré d'accessibilité et donc vérifier leur conformité aux règles de science ouverte[18].

4 La présente étude se propose non pas d'analyser le processus de production et de diffusion scientifique en tant que tel, ni même l'environnement socio-économique de la recherche juridique – que seules des recherches en sociologie, en anthropologie ou en sciences de l'information et de la communication pourraient analyser en profondeur et dans toute leur complexité –, mais de présenter le droit positif de la science ouverte en tant qu'il s'applique aussi à la recherche (en et) sur le droit. Cette réflexion supposera donc, en tant que de besoin, une observation des pratiques de recherche en droit. Par exemple, l'on peut constater qu'une grande partie des ressources utilisées pour la recherche sur le droit est puisée dans le fonds commun de la documentation qui contient l'expression des sources du droit au sens large, c'est-à-dire les normes, règles et solutions, mais aussi toute la documentation y afférent constituée des ressources émanant de toutes les institutions publiques et participant à la publication, à la présentation et à l'évaluation des règles, normes et solutions[19]. L'ensemble de cette documentation, constituée des sources et ressources du droit, est mis à disposition du public afin d'assurer le fonctionnement d'un État de droit fondé sur des principes démocratiques. Doté d'une valeur constitutionnelle[20], le principe d'accès au droit commande de même que l'ensemble de ces documents doit, en effet, pouvoir être rendu accessible aux justiciables, c'est-à-dire à celles et ceux qui doivent la respecter et/ou qui peuvent s'en prévaloir, assistés ou représentés par les professionnels du droit. La recherche sur le droit présente de fait une singularité quant au matériau utilisé : l'accès au droit permet, dans le même

temps, l'accès aux ressources nécessaires à la recherche sur le droit manifestant une congruence entre le principe de science ouverte et le principe d'accès au droit. L'accès à la documentation juridique publique se trouve également assuré par les règles encadrant la diffusion des documents administratifs et des données publiques, contenues dans le Code des relations entre le public et l'administration (CRPA)[21]. Procédant à l'ouverture des données publiques, ces règles ont pour objectif de répondre à la demande, contemporaine de la révolution numérique, d'accéder à toutes les données produites par l'administration. La mise à disposition massive des décisions de justice par exemple, à certaines conditions néanmoins[22], permettrait certainement de procéder à l'analyse à grande échelle des données de la justice dans un but scientifique ou dans l'objectif commercial poursuivi par les entreprises (au sens large)[23] de produire et d'offrir des services innovants. Fruit d'une évolution législative importante, la politique d'ouverture des données publiques croise, sans se confondre avec elle, celle de l'accès au droit. L'accès à certaines ressources juridiques pour la recherche sur le droit s'avère donc être le résultat de l'application des règles relatives à l'accès au droit et à l'ouverture des données publiques, mais non de celle des règles relatives à la science ouverte.

5 De nombreuses autres ressources nécessaires aux recherches sur le droit appartiennent cependant au domaine privé. Il en est ainsi des documents contenant des informations couvertes par le secret, comme le secret professionnel ou encore le secret des affaires. Or, l'étude des contrats d'affaires, en particulier, peut s'avérer être une source précieuse d'enseignements, non seulement sur les pratiques du droit que l'étude de la jurisprudence ne permet pas, à elle seule, de mettre au jour puisque ne concernant que des situations contentieuses, mais aussi sur les éventuelles stratégies d'adaptation des comportements en regard de l'évolution du droit positif[24]. Ainsi en est-il également de la présentation des résultats scientifiques produits par la doctrine et ce, quelle que soit la forme d'expression scientifique, écrite ou orale, utilisée. Bien que souvent produites par des agents de l'État, ces ressources sont soumises, comme dans de nombreux autres domaines scientifiques, à des droits de propriété intellectuelle et, en particulier, au droit d'auteur lorsqu'il s'agit, par exemple, d'écrits scientifiques, de conférences ou encore de bases de

données. Comme c'est le cas dans d'autres domaines scientifiques où la recherche se nourrit des connaissances scientifiques antérieurement produites, la question de l'accès aux ressources qui alimentent la réflexion scientifique entrecroise donc celle de l'économie du processus de diffusion des productions scientifiques juridiques elles-mêmes. Aussi, en droit comme dans d'autres disciplines, les modalités et conditions de la diffusion des ressources produites au cours d'activités de recherche menées par les juristes conditionnent-elles l'accès à ces mêmes ressources pour la production de nouvelles connaissances scientifiques.

6 Deux notions structurent les règles propres à la science ouverte. La première, la notion de « données de la recherche », est de loin la plus difficile à appréhender. La notion de donnée elle-même, tout d'abord, s'avère à la fois fuyante et totalisante. Fuyante, tout d'abord, en ce qu'elle n'emporte pas l'application d'un régime juridique particulier. Si les données ont été, en tant qu'objets de droit, appréhendées par différents dispositifs légaux en particulier depuis la fin des années 1990 [25], la notion n'a, pendant longtemps, fait l'objet d'aucune définition légale. Ce n'est que tout récemment, à l'aune de l'adoption du Règlement européen sur la gouvernance des données, que les données ont été définies de manière très large comme « toute représentation numérique d'actes, de faits ou d'informations et toute compilation de ces actes, faits ou informations, notamment sous la forme d'enregistrements sonores, visuels ou audiovisuels » [26]. Totalisante ensuite, dans la mesure où elle peut, sur le plan technique, englober tout type d'information dès lors que celle-ci a été formatée pour en faciliter le traitement [27]. Une donnée peut donc consister en n'importe quelle information et contenir un message d'une certaine complexité [28]. L'adoption progressive de règles de plus en plus nombreuses ayant pour objet les « données [29] », doublée des effets induits par la révolution numérique, a eu pour effet, collatéral mais important, de transformer tout type de ressource informationnelle en « donnée », obligeant alors à rebours à requalifier ces ressources afin de les distinguer entre elles et d'identifier les différents statuts dont elles relèvent [30]. S'agissant de la notion de « données de la recherche », elle est, quant à elle, définie par la directive (UE) 2019/1024, comme

> « des 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 [31] ».

7 Cette définition qui a déjà fait l'objet d'amples commentaires par ailleurs (Robin, 2022a, n° 11) [32], exclut les publications scientifiques – seconde notion qui soutient l'édifice législatif de la science ouverte – du champ d'application de la directive. Cette exclusion pourrait sembler ne pas converger avec la conception retenue par le législateur européen de la science ouverte qui commande l'ouverture tant des données de la recherche que des écrits scientifiques. La contradiction n'est cependant qu'apparente. De fait, si le champ d'application de la directive (UE) 2019/1024 embrasse tous les documents détenus par les organismes du secteur public des États membres et par certaines entreprises publiques [33], il exclut cependant les documents détenus par des organismes exerçant une activité de recherche et des organisations finançant une activité de recherche, y compris des organisations créées pour le transfert des résultats de recherche [34], ce qui permet, entre autres choses, d'écarter les écrits scientifiques. Cette exclusion qui, à défaut de transposition de la directive, n'a jamais été reprise dans la loi française, tient compte du fait que les productions de la recherche sont le fruit d'une expression libre et personnelle [35]. Parce qu'elles sont en effet, et avant tout, des œuvres de l'esprit susceptibles d'être protégées par le droit d'auteur, les publications scientifiques, ne sont soumises ni au régime juridique des documents administratifs, ni au régime juridique des données de la recherche dont il convient par conséquent de les distinguer.

8 Compte tenu de ce qui précède, nous aborderons la question à travers ces deux objets que sont les données de la recherche, d'un côté, et les écrits scientifiques [36], de l'autre. Bien que particulièrement restrictive [37], cette dichotomie fonde cependant désormais les analyses relatives à la mise en œuvre de la science ouverte, ces objets étant les seuls visés par les textes [38]. Les écrits scientifiques et les

données de la recherche seront, en outre, envisagés indifféremment sous l'angle de leur utilisation et de leur production. En effet, les règles relatives à la science ouverte concernent les productions scientifiques tant dans leur faculté à être produites par les chercheurs que dans celle à être utilisées par d'autres conformément au schéma circulaire du « cycle de vie des données » qui exprime l'action publique en matière de science ouverte[39]. Et, dans la mesure où, comme on l'a dit, les règles relatives à l'ouverture de la science reposent sur le principe de l'ouverture de l'accès à la science, seront abordées, dans un premier temps, l'ouverture de l'accès aux écrits scientifiques utilisés ou produits lors d'une activité de recherche juridique (1) et, dans un second temps, l'ouverture de l'accès aux données utilisées ou produites lors d'une activité de recherche juridique (2).

1. L'ouverture de l'accès aux écrits scientifiques utilisés ou produits lors d'une activité de recherche juridique

9 L'accès aux écrits scientifiques issus de la recherche juridique est assuré par les éditeurs privés ou publics (presses universitaires) qui les publient en leur qualité de cessionnaires des chercheurs (1.1). L'accès aux références et données associés aux écrits scientifiques est, en outre, limité en raison de la logique concurrentielle qui structure le marché de l'édition scientifique (1.2).

1.1. L'accès aux écrits scientifiques assuré par les éditeurs scientifiques

10 Majoritairement déployée dans le secteur privé, l'édition scientifique permet, même s'il existe des exceptions, d'assurer un accès restreint, c'est-à-dire non ouvert, aux écrits scientifiques (1.1.1). Cette logique concurrentielle a pour effet de limiter l'accès aux données accompagnant les écrits scientifiques (1.1.2).

1.1.1. L'accès restreint aux écrits scientifiques

11 Consacrée par les décrets révolutionnaires, la protection des écrits scientifiques a été affirmée tant par les conventions internationales[40] que par la loi du 11 mars 1957[41]. Malgré la référence légale aux seuls « écrits scientifiques », il faut considérer que les œuvres scientifiques protégeables par le droit d'auteur ne sont pas cantonnées à un genre d'expression particulier, mais trouvent à se déployer sous de nombreuses formes d'expression (écrits, entretiens, dessins, schémas, photos, vidéos, sites Internet, bases de données, conférences, cours…). Ainsi, les œuvres scientifiques peuvent-elles être exprimées à l'oral (conférences, colloques, etc.)[42] ou à l'écrit (littérature) ou encore sous la forme graphique ou photographique. L'écriture algorithmique peut en revanche être, à coup sûr, exclue du giron du droit d'auteur, seuls les logiciels étant admis à s'y loger. La reconnaissance d'un droit d'auteur sur une œuvre de l'esprit permet d'attribuer des droits à une ou plusieurs personnes, c'est-à-dire de désigner l'auteur ou les coauteurs, ce qui n'est pas toujours chose aisée dans certains champs scientifiques où les signatures figurant sur un même écrit scientifique peuvent être nombreuses (Pontille, 2016 ; Robin, 2022a).

12 Que l'œuvre soit le résultat du travail d'un ou de plusieurs auteurs, l'attribution des droits est réalisée par la loi aux auteurs personnes physiques[43], sauf qualification de l'œuvre scientifique d'œuvre collective[44]. Comme tous les auteurs, les chercheurs, auteurs d'écrits scientifiques, disposent en tout état de cause des droits moraux et patrimoniaux reconnus par la loi[45]. À ce titre, ils sont en mesure de revendiquer l'attribution (paternité) de leurs œuvres mais aussi le respect de leur intégrité[46] et de décider souverainement de leur divulgation[47]. Au titre du droit d'exploitation, les auteurs, qui ont par ailleurs la faculté de céder leurs droits à titre gratuit, peuvent demander à être rémunérés en contrepartie de la cession de leurs droits au profit d'un exploitant. Dans le domaine de l'édition scientifique, il s'agit donc pour le chercheur de céder ses droits, à titre gratuit ou à titre onéreux, afin d'autoriser l'éditeur à éditer l'écrit scientifique, c'est-à-dire à en assurer la mise en forme, la fabrication, la publication, la promotion et la distribution[48]. En pratique, cette cession est généralement consentie à titre exclusif, pour la durée du monopole légal d'exploitation (soixante-dix ans

post mortem) et pour le monde entier. Il faut enfin mentionner que le statut d'agents publics des auteurs d'écrits scientifiques n'est pas un obstacle à l'exercice de leurs droits. En effet, si les fonctionnaires sont, de manière générale, soumis à des dispositions dérogatoires emportant cession de plein droit de leurs droits au profit de l'État [49], tel n'est pas le cas des chercheurs qui peuvent être considérés comme des « agents auteurs d'œuvres dont la divulgation n'est soumise, en vertu de leur statut ou des règles qui régissent leurs fonctions, à aucun contrôle préalable de l'autorité hiérarchique », en raison de leur indépendance [50] et de la liberté d'expression qui caractérise l'exercice de leurs fonctions d'enseignement et de recherche (Cornu, Mallet-Poujol, 2005 ; Vivant, 2005 ; Cornu, 2006 ; Dreyfus, 2006) [51].

13 Ainsi, du fait de la nature juridique du droit d'auteur qui est, à juste titre selon nous, considéré comme un droit de propriété individuelle (Robin, Chatry, 2022) [52], les chercheurs, auteurs d'écrits scientifiques, sont engagés individuellement dans les relations contractuelles avec les éditeurs. L'économie de l'édition scientifique [53] repose donc tout entière sur le fait que le droit d'auteur est attribué aux chercheurs eux-mêmes [54]. Afin de remplir leur mission de diffusion des connaissances scientifiques [55], sur la base de laquelle leur travail est évalué (bibliométrie), les chercheurs ont une obligation de publier et, de préférence, dans les revues ayant acquis une forte notoriété. Leur capacité de négociation des conditions de cession de leurs droits s'avère donc de manière générale peu élevée, même si elle l'est davantage pour les chercheurs relevant de la recherche en droit (juristes) [56]. Dans ce contexte, le principe d'une gestion collective des droits sur les écrits scientifiques pourrait d'ailleurs s'avérer une piste intéressante (Boutron, Tremolière, 2023). En leur qualité d'ayants droit des chercheurs, les éditeurs scientifiques sont parvenus à constituer des portefeuilles importants de titres de droit d'auteur, qu'ils peuvent exploiter sur un marché composé non seulement des bibliothèques universitaires et des fonds dédiés à la recherche, mais aussi de nombreux professionnels (magistrats, avocats, entreprises, associations). Grâce au droit d'auteur, les éditeurs parviennent donc à assurer l'accès aux ressources produites par les chercheurs. Ce n'est cependant pas d'un accès ouvert qu'il s'agit, mais d'un accès restreint puisque conditionné au paiement

d'un prix, du moins si l'on s'appuie sur la conception retenue par le législateur européen qui dispose qu'un accès ouvert s'entend comme « la pratique consistant à fournir gratuitement l'accès en ligne à des résultats de recherche à l'utilisateur final, sans restriction sur l'utilisation et la réutilisation au-delà de la possibilité d'exiger l'indication de l'auteur[57] ». Il faut également constater que, si certaines revues ont fait le choix d'une diffusion gratuite[58], ce n'est généralement que grâce au soutien apporté par un financement public et à l'absence de rémunération, là encore, des chercheurs au titre de leur droit d'auteur. Dit autrement, le choix d'un accès ouvert généralisé des écrits scientifiques ne pourrait être le fait que de la puissance publique, qu'il s'agisse de l'évaluation, de la publication et de la diffusion des productions scientifiques[59].

1.1.2. L'accès limité aux données associées aux écrits scientifiques

14 Cessionnaires du droit d'auteur des chercheurs, les éditeurs scientifiques détiennent un nombre toujours plus important[60] de droits sur les ressources scientifiques. Sur un plan technique, ces ressources sont considérées comme des données et les références qui les accompagnent, comme des métadonnées. Ensemble, elles constituent aujourd'hui l'essentiel de la valeur ajoutée du capital des groupes d'édition. La constitution de vastes bases de données[61] permet de fait aux éditeurs d'offrir, grâce à la technologie numérique, des fonctionnalités intéressantes pour les utilisateurs par le croisement des données dont ils disposent grâce à la création d'hyperliens et au développement de systèmes d'intelligence artificielle (accès aux décisions, articles ou encyclopédies ou à la chronologie d'une affaire jugée à différents degrés de juridiction). L'accès à ces fonctionnalités permet ainsi aux abonnés de travailler dans un confort inédit grâce à une navigation, sinon fluide, du moins rapide, d'une ressource à l'autre. La logique monopolistique et concurrentielle de l'édition scientifique produit toutefois une limite en ce qu'elle ne permet pas d'accéder directement, par la technique de l'hyperlien, aux contenus édités par les sociétés d'édition concurrentes. Seules les données correspondant à des ressources éditées par l'éditeur publiant l'écrit scientifique s'avèrent en effet directement accessibles. Pour les autres, seules les métadonnées sont communiquées (auteur, titre,

source, date), obligeant les utilisateurs à reconstituer la base de données des sources utilisées par le premier chercheur pour vérifier les références et éventuellement les réutiliser, en se reportant aux autres catalogues dans l'espoir que son établissement de rattachement y soit abonné. La publication d'un écrit scientifique issue d'une recherche scientifique publique sur le droit ne permet donc pas d'accéder à l'ensemble des données (jeux de données) issues d'un même travail et référencées dans un même document. Outre le fait que toutes les revues juridiques ne figurent pas dans les catalogues des universités malgré la participation des universitaires à ces revues [62], on voit poindre la possibilité d'une inégalité de services rendus au public entre les différents établissements quant à l'accès à la documentation, que ce soit pour les chercheurs ou pour les étudiants. Pour l'ensemble de ces raisons, l'accès aux écrits scientifiques tel qu'assuré par les éditeurs ne peut être considéré comme produisant les effets escomptés par la politique publique de science ouverte, c'est-à-dire assurer un accès ouvert et direct à toutes les données collectées ou utilisées à l'occasion d'une production scientifique en droit et rassemblées sur un document à l'occasion de la publication d'un écrit scientifique.

1.2. L'accès aux écrits scientifiques renforcé par la science ouverte

15 Le principe de l'ouverture de l'accès aux écrits scientifiques est encadré par des dispositions légales d'ordre public (1.2.1). Des recommandations ont en outre été émises préconisant, de façon plus radicale, une diffusion immédiate des écrits en accès ouvert (1.2.2).

1.2.1 La consécration légale d'une dérogation d'ordre public

16 La publication des travaux des chercheurs passe, en droit comme dans de nombreux autres domaines scientifiques, par le prisme quasi exclusif [63] de l'exploitation économique des ressources par les éditeurs, y compris lorsque l'édition est le fait des presses universitaires. L'accès à ces travaux est donc conditionné par l'abonnement aux revues ou l'achat des ouvrages en version papier ou électronique (*via* des plateformes comme Cairn et Open Edition par

exemple). L'économie de l'édition ayant donné lieu, comme on l'a vu précédemment, à des bouleversements, elle est aujourd'hui remise en cause, et les solutions d'autoarchivage sont vivement recommandées aux chercheurs[64]. Ces recommandations institutionnelles ont été accompagnées en France, comme dans d'autres pays (Robin, 2022a, n° 563-564), de l'adoption de dispositions autorisant les chercheurs, auteurs d'écrits scientifiques, à déroger au principe de liberté contractuelle et à la force obligatoire des contrats, en leur offrant, malgré l'exclusivité du contrat d'édition, la faculté de diffuser leurs articles sur d'autres supports (blogs, archives ouvertes, etc.)[65]. La loi prévoit en effet que les écrits scientifiques qui sont le fruit de recherches financées majoritairement par des fonds publics, publiés dans des revues périodiques, peuvent être mis à disposition en format ouvert dans leur version finale acceptée par l'éditeur (avant mise en page par celui-ci), soit immédiatement si l'éditeur propose déjà une mise à disposition ouverte, soit après un certain temps (durée d'embargo de douze mois en droit), à condition que la mise à disposition soit gratuite, qu'elle n'emporte pas une exploitation commerciale de l'écrit et que l'ensemble des coauteurs y aient, le cas échéant, consenti. La liberté alors accordée aux chercheurs répond aux préoccupations économiques et sociales qui sont à l'origine du débat : les résultats issus d'une recherche publique doivent être mis au service de l'ensemble de la société, ce qui implique une plus large diffusion que celle qui résulte d'une diffusion jusqu'ici contrainte par l'exclusivité imposée par les revues scientifiques (Cornu, 2016). S'il est clairement facultatif, le procédé est cependant devenu obligatoire en pratique, du moins dans le cadre de certains appels à projets (de type ANR) pour lesquels la diffusion en accès ouvert est devenue une condition de financement.

17 Afin de consolider la règle, le législateur a pris soin enfin d'affirmer le caractère d'ordre public des dispositions[66]. Il s'agit d'empêcher le chercheur, auteur d'écrits scientifiques, considéré ici comme « partie faible », de renoncer par contrat aux prérogatives dont il dispose en vertu de la loi[67]. L'étude d'impact de la loi pour une République numérique mentionne que

> « une atteinte est portée à la liberté contractuelle des auteurs et des éditeurs, le caractère d'ordre public empêchant l'auteur ou son établissement de renoncer au bénéfice du droit qui lui est conféré en

> ce qui concerne les publications et de céder à l'éditeur des licences tendant à limiter la réutilisation des données de la recherche [68] ».

18 Les contrats d'édition ne peuvent donc avoir pour effet de limiter les conditions de publication des écrits scientifiques et de réutilisation des données au-delà du seuil autorisé (durée d'embargo, à des fins non commerciales, version finale du manuscrit, etc.). Le caractère d'ordre public de la règle devrait permettre, en tant que loi de police, de faire rayonner les règles tant à l'égard des éditeurs scientifiques établis sur le territoire national qu'à l'égard de ceux établis hors de celui-ci (Robin, 2022a, n° 588).

1.2.2. La recommandation d'une diffusion immédiate du manuscrit

19 Le Plan national pour la science ouverte (PNSO) préconise par ailleurs la mise en œuvre de la « stratégie de non-cession des droits sur les publications scientifiques » (Ministère de l'Enseignement supérieur et de la Recherche, 2023). Cette stratégie consiste dans le fait de rendre immédiatement accessible un article dans une archive ouverte de type HAL et ce, malgré la publication de l'article par un éditeur. Cette recommandation reprend les conditions applicables aux projets de recherche financés par des fonds européens en vertu du programme « Horizon Europe » [69] dont l'article 39, 3°, prévoit que :

> « Les bénéficiaires veillent à ce que l'accès ouvert aux publications scientifiques s'applique dans les conditions établies dans la convention de subvention. En particulier, les bénéficiaires veillent à conserver, ou à ce que les auteurs conservent, suffisamment de droits de propriété intellectuelle pour se conformer à leurs obligations en matière d'accès ouvert. »

20 Elle est également portée par la cOAlition S, regroupement international de 28 organisations de financement de la recherche dont font partie l'Agence nationale de la recherche (ANR) et la Commission européenne. S'agissant spécifiquement des publications financées sur fonds publics, la cOAlition S prévoit une ouverture généralisée [70] :

> « Après le 1er janvier 2021, les publications scientifiques sur les résultats de la recherche financée sur fonds publics accordés par des

> agences de recherche ou des organismes de financement nationaux ou européens, doivent être publiées dans des revues ou sur des plateformes en accès libre. »

21 Fruit de cette coalition, le Plan S prévoit, de façon radicale, que les articles de recherche issus des projets que ces organisations financent soient disponibles en accès ouvert et sous licence libre dès la date de publication. Aussi est-il recommandé à l'auteur, qui renonce alors *de facto* à exercer ses prérogatives patrimoniales[71], d'avertir l'éditeur qu'une licence libre, généralement une licence *Creative Commons* de type CC-BY, sera appliquée au manuscrit et à toutes les versions successives jusqu'au « manuscrit auteur accepté » (MAA) pour publication après relecture par les pairs. Cette dernière version du manuscrit pourrait alors être diffusée immédiatement dans une archive ouverte[72]. Juridiquement respectueuse du droit d'auteur, cette stratégie, qui replace l'auteur au centre de l'exercice de ses prérogatives, s'autorise, d'une part, de ce qu'en pratique les chercheurs auront le plus souvent consenti une cession de leurs droits à titre gratuit, et d'autre part, du principe selon lequel l'éditeur n'est titulaire d'un droit que sur la mise en page de l'écrit scientifique et non sur les versions antérieures. Comme pour toutes les publications empruntant la « voie diamant » (Bernault, 2016), elle pose néanmoins à terme la question du financement des publications scientifiques, mais également celle de la qualité de l'évaluation des publications sur les plateformes de libre accès. Ces écueils ont été relevés comme pouvant nuire à l'indépendance, à la liberté d'expression du chercheur, à la qualité du travail d'édition, ainsi qu'à la « bibliodiversité » (Boutron, Tremolière, 2023 ; Henriet, Darcos, Ouzoulias, 2022).

2. L'ouverture de l'accès aux données utilisées ou produites lors d'une activité de recherche juridique

22 Les données de la recherche sur le droit sont constituées d'éléments tout à fait disparates qu'il est possible de ranger en deux catégories, en tant qu'elles peuvent être qualifiées soit de données publiques, soit

de données privatives[73]. Les premières sont celles qui sont utilisées pour la recherche sur le droit et qui répondent aux définitions relevant du dispositif de l'ouverture des données publiques (2.1). Les secondes, qui constituent également un corpus important des ressources utilisées pour la recherche sur le droit, sont ainsi dénommées en ce que, sans être des objets de propriété à proprement parler, elles sont, à différents titres, réservées privativement (2.2).

2.1. L'accès aux données publiques

23 Parce qu'elles sont produites dans le cadre des organismes du secteur public, c'est-à-dire des administrations, de nombreuses données utilisées pour les besoins de la recherche sur le droit sont des données publiques (2.1.1). En tant que telles, elles devraient également répondre à la définition juridique de données ouvertes (2.1.2).

2.1.1. Les données utilisées et produites lors d'une recherche juridique, des données publiques

24 **Les données utilisées pour la recherche, des données publiques par nature.** La notion de « données publiques » n'est pas, à proprement parler, définie par les textes. Cela n'est pas surprenant dans la mesure où c'est la notion de « document » – moins fuyante que celle de « donnée » – qui structure tout l'édifice de l'ouverture des données publiques. Le document, objet des obligations de diffusion et de réutilisation, est défini en droit de l'Union comme « a) tout contenu quel que soit son support (papier ou forme électronique, enregistrement sonore, visuel ou audiovisuel) ; ou b) toute partie de ce contenu[74] ». Plus synthétique que son homologue française[75], la définition comprend cependant le document comme un contenu là où l'on pouvait plutôt l'envisager comme un contenant, ou tout au moins comme un support[76]. Le document peut donc être considéré comme le contenu qui peut lui-même être le tout mais aussi l'une de ses plus petites parties, c'est-à-dire une donnée[77]. Parmi les documents exclus du champ d'application de la directive[78] figurent les documents sur lesquels des tiers détiennent des droits de propriété intellectuelle et les documents contenant des données sensibles protégées par des dispositions légales nationales particulières (dont

la protection de la sécurité nationale, le secret des données statistiques, le secret d'affaires, les secrets professionnels) et les documents contenant des données à caractère personnel. Les documents détenus par les organismes publics sont régis par le principe général énoncé par la directive qui est celui de la réutilisation gratuite des documents concernés, que cette réutilisation soit faite à des fins commerciales ou non commerciales [79]. Cette obligation d'autoriser la réutilisation ne peut cependant concerner que des documents existants au sens où ils ne sont pas à produire [80].

25 Les ressources utilisées pour les recherches sur le droit peuvent généralement être qualifiées de documents administratifs et, en tant que tels, leur réutilisation gratuite par les chercheurs est autorisée. La mise à disposition des décisions de justice, par exemple, résulte ainsi non pas du fait qu'il s'agirait de données de la recherche dont il faudrait assurer l'ouverture, mais du fait qu'elles constituent avant tout des documents administratifs produits ou reçus par les juridictions de l'ordre administratif et judiciaire en tant qu'elles sont des administrations. L'ouverture de ces décisions n'est donc pas le résultat de la mise en œuvre des règles relatives à la science ouverte. Bien que leur mise à disposition soit soumise à des contraintes particulières [81], elles subissent le même sort que tous les autres documents rendus publics par l'administration ou ceux conservés au sein des services d'archives sous réserve de leur classement [82]. Il faut gager que l'ouverture des décisions de justice aura pour effet à terme de compléter la base de données publique Légifrance qui ne représente pour l'heure qu'une partie des décisions rendues par les juridictions françaises. De même, au titre des documents administratifs répondant à l'obligation de diffusion figurent les algorithmes [83], définis comme « l'étude de la résolution de problèmes par la mise en œuvre de suites d'opérations élémentaires selon un processus défini aboutissant à une solution [84] ». Ils peuvent constituer des ressources utiles pour les recherches sur le droit, notamment lorsqu'il s'agit de comprendre comment les règles légales ou réglementaires ont été implémentées dans le code informatique pour faire produire à un système d'intelligence artificielle une décision individuelle par exemple [85].

26 **Les données utilisées ou produites lors d'une recherche, des données publiques par destination.** Les données de la recherche ne

sont pas des données publiques par nature. Leur intégration dans la directive (UE) 2019/1024 ne résulte pas d'une définition commune qui engloberait tous les documents publics, y compris les données de la recherche. La notion qui bénéficie ainsi d'une autonomie conceptuelle [86] subit également l'application d'un régime juridique particulier. Ce statut juridique exceptionnel s'explique sur le fond par le fait qu'elles sont le fruit d'un écosystème particulier, celui de la recherche scientifique, et qu'elles sont détenues par des établissements qui, s'ils exercent bien une mission de service public (recherche scientifique), développent aussi toutefois une activité industrielle et commerciale (valorisation). La définition légale exprime l'idée que les données de la recherche n'acquièrent ce statut qu'à partir du moment où elles sont utilisées ou produites au cours d'activités de recherche. L'expression semble judicieuse à condition de définir l'activité scientifique sur la base du critère intellectuel et moral (et non institutionnel ou organique) qui caractérise, à notre sens, la démarche scientifique (Robin, 2022a, n° 12). La définition légale européenne des données de la recherche retient par ailleurs une acception restrictive de la notion. Outre le fait que la définition exclut les publications scientifiques de la notion de données, le champ d'application des dispositions est cantonné aux données « utilisées 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 », c'est-à-dire aux données qui permettent d'administrer la preuve scientifique afin de vérifier le sérieux et l'objectivité des résultats produits à partir de l'analyse des données, ce qui exclut de la définition un grand nombre de données (données brutes, données intermédiaires, etc.). L'exclusion est justifiée par l'objectif poursuivi par le législateur européen d'encourager la réutilisation des données [87], nécessitant concomitamment, pour répondre à la « crise de reproductibilité » que traversent certaines disciplines scientifiques, de rappeler et d'imposer le respect des règles relatives à l'intégrité scientifique (Robin, 2022a) [88].

27 Ainsi, dès lors que des ressources utilisées ou produites lors d'une activité de recherche juridique peuvent être qualifiées de données de la recherche au sens de la directive, elles doivent obéir au principe de science ouverte énoncé par l'article 10 de la directive [89]. Les États

doivent donc encourager 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. En outre, les données de la recherche ne sont soumises au principe d'ouverture et de libre réutilisation qu'à la double condition qu'elles soient issues d'une recherche financée au moyen de fonds publics et que les chercheurs les aient rendues publiques par l'intermédiaire d'une archive ouverte institutionnelle ou thématique. Dès lors qu'elles sont mises en accès libre, les données de la recherche sont considérées comme librement réutilisables à des fins commerciales ou non commerciales. La règle de l'article 10 ménage cependant les intérêts des tiers impliqués dans le processus de production et de diffusion scientifique. La mise à disposition des données doit de fait être réalisée en tenant compte des intérêts commerciaux légitimes, des activités de transmission des connaissances et des droits de propriété intellectuelle. Dans la recherche sur le droit, seule la documentation publique échappe à ces contraintes pour des raisons qui, comme on l'a vu, ne tiennent pas à l'application des règles de science ouverte mais de celle des règles relatives à l'accès au droit ou à l'ouverture des données publiques. Toutefois, l'application des règles de science ouverte peut s'avérer pertinente y compris sur des ressources issues de la documentation publique. La mise à disposition *via* une archive ouverte de décisions difficiles à trouver ou obtenues grâce à une demande particulière auprès d'un conseil (avocat) ou d'une juridiction pourrait utilement servir les intérêts des chercheurs eux-mêmes, et aurait ainsi pour effet de transformer la ressource en une donnée ouverte.

2.1.2. Les données utilisées ou produites lors d'une activité de recherche juridique, des données ouvertes ?

28 La directive (UE) 2019/1024 conditionne l'ouverture des données publiques, mais aussi des données de la recherche, au fait qu'elles soient ouvertes.[90] En tant que données de la recherche, ou en tant que données publiques, les ressources utilisées ou produites lors d'une activité de recherche doivent donc être formatées et stockées dans des conditions qui en permettent la conservation en même

temps que la mise à disposition, et constituer ce que l'on appelle des « données ouvertes ». Le législateur européen définit la notion de format ouvert comme un « format de fichier indépendant des plateformes utilisées et mis à disposition du public sans restriction empêchant la réutilisation des documents ». Pour ce qui concerne les données de la recherche en particulier, la directive précise qu'elles doivent être formatées selon les standards FAIR [91]. Selon les préconisations de la Commission européenne (2018, p. 19), également reprises par la Research Data Alliance (RDA) [92], les données doivent, pour être FAIR [93] sur le plan technique, répondre à certaines exigences [94] (Wilkinson *et al.*, 2016). Le respect de la norme FAIR peut être facilité par la rédaction d'un plan de gestion de données (PGD) dès le commencement d'un projet de recherche. Souvent imposé lors d'une demande de financement public national ou européen, l'établissement d'un PGD doit en principe permettre aux chercheurs d'adopter une attitude réflexive sur la gestion de leurs données, d'adopter les bonnes pratiques dans ce domaine et d'encourager par conséquent la diffusion de données réutilisables. De ce point de vue, le PGD joue en quelque sorte le rôle des registres en matière de gestion des données personnelles imposés par la loi dans une démarche de mise en conformité interne (*compliance*). De la même manière que leurs homologues, ils doivent être tenus à jour régulièrement pendant toute la vie du projet de recherche, voire au-delà, permettant d'assurer une gestion raisonnée et responsable des données de la recherche publique. Bien que les modèles de PGD soient nombreux et variés (Hodson, Molloy, 2014), ils sont généralement tous structurés suivant le cycle de vie des données.

29 Afin que les données de la recherche sur le droit puissent être diffusées ou partagées en format ouvert, encore faudrait-il qu'elles aient été organisées et classées de manière structurée en « jeux de données ». Plusieurs solutions, sur le plan technique, permettent aujourd'hui aux chercheurs de conserver mais aussi de partager ou de diffuser leurs jeux de données, comme les infrastructures de données (ou infrastructures de recherche selon les cas) ou encore les entrepôts de données (Robin, 2022a). De nombreux entrepôts institutionnels se sont développés à travers le monde afin de proposer un service de dépôt, d'archivage pérenne et un service d'assistance et d'accompagnement à la gestion des données (Catherine, 2020 ;

Schöpfel, Rebouillat, 2021). Ainsi en est-il de la plateforme HAL qui offre désormais la possibilité d'archiver les jeux de données sous forme de « *data paper* » (Reymonet, 2017)[95], ou encore l'entrepôt Zenodo[96], l'outil de gestion de ressources Zotero[97], etc. Quel que soit l'outil utilisé, les données doivent être décrites et documentées à l'aide des métadonnées. Dans le domaine de la recherche sur le droit, les jeux de données seraient donc constitués de l'ensemble des références traditionnelles des ressources utilisées dans la recherche sur le droit évoquées plus haut : références légales et réglementaires, références jurisprudentielles, articles de doctrine, rapports et autres documents des autorités publiques, etc. Si ces données sont déjà généralement très clairement identifiées et décrites dans le document présentant les résultats d'une recherche, soit en notes infrapaginales, soit en fin de document dans les tables alphabétiques de références bibliographiques, l'on pourrait toutefois imaginer, d'une part, que ces références soient publiées indépendamment du document de référence ou de la publication elle-même et, d'autre part, qu'elles contiennent outre les références, c'est-à-dire les métadonnées, le contenu des ressources lui-même, sous réserve des droits des personnes titulaires des droits d'auteur (auteurs ou éditeurs) sur les ressources. La diffusion de jeux de données structurés par le biais des entrepôts de données ne préjuge pas cependant du périmètre de leur ouverture. Si de nombreux entrepôts recommandent l'ouverture maximale par l'utilisation des licences permissives de type *Creative Commons* (généralement CC-BY 4.0), il n'est pas interdit de paramétrer l'accès aux données de manière, par exemple, que les droits des tiers soient préservés.

2.2. L'accès aux données privatives pour la recherche juridique

30 Par données privatives, il convient d'entendre toutes les bases de données réservées par un droit de propriété intellectuelle et, en particulier, par le droit *sui generis* des bases de données (2.2.1), et toutes les informations couvertes par le secret comme l'ensemble des pratiques, usages et savoir-faire déployés par les praticiens du droit (2.2.2)[98].

2.2.1. Les bases de données couvertes par un droit de propriété intellectuelle

31 Si les éditeurs détiennent, en tant qu'ayants droit des auteurs des droits d'auteur sur les écrits scientifiques, ils détiennent également, mais cette fois en tant que producteurs de bases de données, un droit *sui generis* sur les bases de données (systèmes d'information ou infrastructures) qu'ils conçoivent pour organiser l'accès aux ressources produites par la recherche sur le droit. Ce droit leur permet de monnayer, par exemple, l'utilisation de leurs bases de données lors d'opérations de fouille et d'exploration réalisées à des fins scientifiques. Ces données constituées des écrits scientifiques, sur lesquelles les sociétés d'édition capitalisent aujourd'hui, sont des données qu'il est en effet possible de collecter et d'analyser massivement afin d'en tirer des conclusions propres à nourrir l'activité scientifique des chercheurs, mais aussi l'activité économique d'acteurs privés. La consécration de l'exception légale de fouille de textes et de données [99], qui vient précisément répondre aux besoins générés par l'analyse massive de données, est assortie de plusieurs conditions dont l'une oblige les utilisateurs des bases de données à bénéficier d'un accès licite aux bases de données explorées [100]. Cette condition de licéité, qui figure également dans l'exception légale de copie privée depuis 2011 [101], permet de s'assurer que l'utilisation des données ne peut être mise en œuvre à partir d'un corpus auquel il aurait été accédé de manière illicite, c'est-à-dire sans autorisation des titulaires de droits. La notion de licéité de l'accès doit s'entendre, selon la directive, comme couvrant l'accès à des contenus fondé sur une politique de libre accès ou en vertu d'arrangements contractuels entre, d'une part, les titulaires de droits et, d'autre part, les organismes de recherche, comme des abonnements, ou en vertu d'autres voies légales. Ainsi, la constitution d'une base de données comme ISTEX [102] a-t-elle nécessité de la part du consortium qui en est à l'origine [103] de négocier l'accès aux données détenues par les éditeurs scientifiques avant de pouvoir l'ouvrir aux chercheurs pour qu'ils puissent y réaliser des fouilles (Bensamoun, Bouquerel, 2020 ; Robin, 2022a, n° 186 et suiv. ; Robin, 2022b). Il s'agira enfin d'observer que les bases de données conçues par le CNRS qui propulse les archives ouvertes HAL devraient pouvoir également être considérées comme

couvertes par un droit *sui generis* des bases de données détenu par l'organisme de recherche. La titularité des droits du CNRS sur la base de données HAL ne devrait cependant avoir aucune incidence sur l'utilisation qui peut en être faite dans la mesure où le CRPA interdit aux administrations, au sens de l'article L. 300-2 du même code, de se prévaloir du droit *sui generis* sur les bases de données pour faire obstacle à l'application du principe de réutilisation des données qu'elles contiennent[104].

32 Afin d'éviter que les chercheurs soient empêchés par les éditeurs, par le jeu des contrats de cession de droits dont l'objet peut être, outre l'article scientifique lui-même, les données associées à la recherche, le législateur français (depuis 2016[105]) et le législateur européen (depuis 2019[106]) ont adopté un principe de libre réutilisation des données de la recherche (Robin, 2022a). Selon ce principe[107], et sur le fondement de l'article L. 533-4, II, du Code de la recherche, dès lors que les données de la recherche ont été majoritairement financées par des fonds publics, qu'elles ne sont pas protégées par un droit spécifique ou une réglementation particulière et qu'elles ont été rendues publiques par le chercheur, l'établissement ou l'organisme de recherche, leur réutilisation est libre, sans que l'éditeur de l'article scientifique auquel les données sont associées ne puisse en limiter la réutilisation dans le cadre de la publication[108]. Le texte ne limite pas la réutilisation à des fins non commerciales, ce qui laisse ouverte la possibilité de procéder à une exploitation commerciale des données de la recherche dès lors qu'elles ont été rendues publiques, interprétation confortée par le texte européen qui prévoit que les données peuvent être réutilisées à des fins commerciales ou non commerciales conformément au régime juridique qui prévaut en matière de données publiques et prévu aux chapitres III et IV de la directive elle-même. Concernant la mise à disposition des données, le texte français ne désigne aucun outil pour l'opération de publication, là où le texte européen vise, quant à lui, les archives ouvertes institutionnelles ou thématiques. Les chercheurs et les établissements ou organismes publics peuvent donc décider, au cas par cas, quelles données ouvrir et à quelles conditions. Lorsqu'elles consistent en des références légales ou jurisprudentielles, c'est-à-dire en des données publiques, la plus-value d'une diffusion supplémentaire par les chercheurs peut sembler *a priori* peu élevée. Il faut cependant nuancer

cette affirmation car la diffusion par un chercheur des décisions de juridictions étrangères, par exemple, peut être précieuse pour d'autres chercheurs. De même, la diffusion de textes anciens ou oubliés, mais exhumés à l'occasion d'une recherche, constitue un atout important pour la recherche scientifique. Par ailleurs, compte tenu du fait que l'accès à l'ensemble des décisions de jurisprudence n'est encore pas assuré de façon exhaustive par Légifrance, imposant alors le recours à la base de données privée JurisData (accessible sur abonnement), la publication de nombreuses décisions par les chercheurs eux-mêmes pourrait s'avérer très utile à la communauté scientifique.

2.2.2. Les pratiques, usages et savoir-faire couverts par le secret

33 Sur le plan juridique, les données privatives utiles à la recherche sur le droit peuvent également consister en des modèles d'affaires, par exemple, conçus par les acteurs économiques et, en particulier, les contrats dont l'étude peut s'avérer une source précieuse d'enseignements, non seulement sur les pratiques du droit que l'étude de la jurisprudence ne permet pas, à elle seule, de mettre au jour puisque ne concernant que des situations contentieuses, mais aussi sur les éventuelles stratégies soit de contournement de la loi, soit d'adaptation des comportements. L'étude des pratiques, usages et savoir-faire déployés en droit s'avère d'autant plus intéressante que ces usages constituent précisément une source du droit à laquelle le législateur ou les juges réfèrent régulièrement. Si ces usages sont souvent recensés dans des codes ou chartes, un certain nombre d'entre eux propres à un secteur économique en particulier restent cependant encore non documentés (Mousseron, 2023)[109]. En outre, les savoir-faire, qui prennent la forme de contrats ou de modèles d'affaires sont généralement soumis au secret d'affaires auquel les principes de science ouverte ne dérogent pas[110]. Le déploiement du commerce électronique a cependant pour effet de produire un très grand nombre de contrats, conditions générales d'utilisation (CGU), dont l'analyse peut, en raison de leur diffusion (imposée par la loi[111]), être entreprise de façon systématique[112]. L'accès aux CGU n'équivaut pas aux modèles d'affaires mais peut, à coup sûr, révéler les choix et stratégies d'une entreprise. Leur étude s'avère donc éclairante, notam-

ment dans le cadre d'une recherche empirique en droit des contrats, en droit de la consommation, en droit du marché et de la concurrence, etc.

34 Par ailleurs, un mouvement se dessine aujourd'hui dans le sens d'une ouverture, ou plus exactement, d'un partage de données privées, c'est-à-dire de données détenues par les acteurs économiques et partagées, en accès restreint, avec des institutions publiques, des particuliers ou d'autres acteurs économiques dans des conditions susceptibles de préserver la confidentialité ou les droits des personnes. Tout d'abord, la directive (UE) 2019/1024 a inclus certaines entreprises publiques qui agissent dans des secteurs indirectement concurrentiels dans le périmètre des personnes concernées par les obligations d'ouverture des données[113], dès lors qu'elles exercent des activités dans les domaines définis et visés par la directive 2014/25/UE, qu'elles agissent en qualité d'opérateurs de services publics conformément à l'article 2 du règlement n° 1370/2007, qu'elles relèvent de certains secteurs (transports aériens, armateurs). La même directive encourage également la mise à disposition de données spécifiques de forte valeur, c'est-à-dire de données susceptibles de générer des avantages socio-économiques ou environnementaux importants et des services innovants, notamment pour les PME, et susceptibles d'être appariées à d'autres ensembles de données[114]. Ensuite, le partage de données privées entre les administrations et les opérateurs privés ou entre opérateurs privés est aujourd'hui encouragé par les règlements européens successifs que sont le Règlement sur la gouvernance des données[115], le Règlement sur les données[116] ou encore la proposition de règlement sectoriel sur l'espace européen des données de santé[117], qui ont pour objet d'améliorer les conditions du partage des données dans le marché intérieur. L'accès à ces données peut s'avérer utile aux recherches menées en droit dans la mesure où elles permettent, par exemple, de comprendre la régulation interne des entreprises ou encore l'intégration normes comme la norme ISO 26000 en matière de responsabilité sociétale des entreprises (RSE). L'accès à ces données, qui peuvent être partagées à des conditions restrictives[118], s'avère surtout intéressant pour les recherches menées dans d'autres disciplines (économie ou gestion, en particulier) dont les résultats peuvent ensuite utilement nourrir les analyses menées par les juristes.

Conclusion

35 La science ouverte interroge, de façon inédite, l'ensemble du processus de production scientifique. Se poser la question de l'application de la science ouverte à la recherche sur le droit invite donc à une réflexion d'ordre épistémologique. Si elle est intimement mêlée à l'ouverture des données publiques et au principe d'accès au droit, l'application de la science ouverte à la recherche (en et) sur le droit ne se réduit cependant pas à cette seule dimension en raison de la pluralité des ressources utilisées dans la discipline. L'analyse de la production scientifique en droit au prisme de la science ouverte montre surtout les potentialités non encore explorées de la constitution de corpus de données propres à un travail de recherche auquel il serait possible d'accéder directement et indépendamment de la présentation des résultats dans l'écrit scientifique. Un tel accès pourrait s'avérer précieux pour la vérification de la validité de certains résultats scientifiques obtenus et pour le gain de temps procuré pour d'autres recherches. Cet accès pourrait également s'avérer utile aux entreprises et cabinets de conseil souhaitant se saisir des données compilées et diffusées *via* une archive ouverte. Analyser la production scientifique (en et) sur le droit à travers la notion de « données » change en tout état de cause la perspective. La numérisation des ressources et leur compilation au sein des catalogues des éditeurs ou au sein d'archives ouvertes produisent un effet de massification réduisant par exemple lesdits écrits scientifiques en données dont l'exploration peut s'avérer à terme intéressante, ne serait-ce que pour observer les tendances, récurrences et biais de la recherche en droit. Tel est également l'objectif poursuivi par la science ouverte, justifiant ainsi que l'ouverture soit considérée comme le principe et la fermeture comme l'exception. Enfin, comme on l'observe d'ores et déjà dans les disciplines scientifiques pionnières en matière d'ouverture des données, la mise en œuvre de la science ouverte induit des changements dans les pratiques scientifiques, puisqu'elle implique de concevoir, dès le début du processus de recherche, la manière dont les données seront formatées, structurées, conservées et mises à disposition conformément au principe de l'ouverture « par défaut », dans la perspective qu'elles soient réutilisées un jour par d'autres.

BIBLIOGRAPHY

Les adresses Internet citées dans cet article ont été consultées le 17 décembre 2024.

Assemblée nationale, 2015, *Projet de loi pour une République numérique, étude d'impact*

Bensamoun A., Bouquerel Y., 2020, *Transposition des exceptions de fouille de textes et de données : enjeux et propositions*, rapport de mission, CSPLA

Bernault C., 2016, *Open access et droit d'auteur*, Bruxelles, Larcier

Boutron M., Tremolière A., 2023, *Rapport de mission sur la science ouverte et le droit d'auteur*, CSPLA

Boy L., 2007, « Normes techniques et normes juridiques », *Cahiers du Conseil constitutionnel*, n° 21, p. 119-126

Bulletin des bibliothèques de France, 1982, « Enrichissement du vocabulaire de l'informatique : arrêté du 22 décembre 1981 », n° 6, p. 355-358, [https://bbf.enssib.fr/consulter/bbf-1982-06-0355-009]

Catherine H., 2020, *Étude comparative des services nationaux de données de recherche. Facteurs de réussite*, MESRI, CoSO

Commission européenne, 2018, *Turning FAIR into reality*, rapport final et plan d'action du groupe d'experts de la Commission européenne sur les données FAIR

Cornu M., 2016, « À propos des productions intellectuelles de la recherche, entre logique privative et nécessités publiques », *Propriétés intellectuelles*, n° 20, p. 270-280

Cornu M., 2006, « Droit d'auteur des fonctionnaires : le périmètre contenu de l'exception de service public », *Dalloz*, p. 2185

Cornu M., Mallet-Poujol N., 2005, « Droit d'auteur des universitaires et des chercheurs : l'expropriation sans cause d'utilité publique », *Dalloz*, p. 3025

CNRS, 2016, *Livre blanc. Une science ouverte dans une république numérique*, DIST, [https://www.science-ouverte.cnrs.fr/wp-content/uploads/2019/07/Livre-Blanc-loi-numerique-Octobre-2016.pdf]

CNRS, 2015, « Résultats 2014 des grands éditeurs scientifiques : une croissante satisfaisante, des profits records. Un terrain favorable à de nouvelles concentrations ? », *DISTinfo14*, [https://www.science-ouverte.cnrs.fr/wp-content/uploads/2019/09/Distinfo14.pdf]

Depincé M., Robin A. (dir.), 2024, *Outil d'analyse de conformité des stipulations* (OACS), IERDJ, rapport n° 20.44, [https://ierdj.fra1.digitaloceanspaces.com/media_library/2024/04/20.44_OACS_rapport.pdfhttp://www.gip-recherche-justice.fr/publication/outil-danalyse-de-conformite-des-stipulations-oacs/]

Dreyfus J.-D., 2006, « Brèves remarques sur le droit d'auteur des agents publics

après la loi du 1er août 2006 », *AJDA*, p. 2179

Galvez-Behar G., 2020, *Posséder la science. La propriété scientifique au temps du capitalisme industriel*, Aubervilliers, Éditions de l'EHESS

Henriet P., Darcos L., Ouzoulias P., 2022, *Pour une science ouverte, réaliste, équilibrée et respectueuse de la liberté académique*, rapport n° 5154 (Assemblée nationale) / n° 573 (Sénat)

Hodson S., Molloy L., 2014, *Current best practice for research data management policies. A memo for the Danish e-Infrastructure Cooperation and the Danish Digital Library*, CODATA

Leclerc O., 2021, « Jalons prospectifs sur l'exigence de reproductibilité dans la recherche juridique », dans *Mélanges en l'honneur de P. Ancel*, Bruxelles, Larcier, p. 175-186

Ministère de l'Enseignement supérieur et de la Recherche, 2023 [2022], *Mettre en œuvre la stratégie de non-cession des droits sur les publications scientifiques. Guide pour les chercheuses et les chercheurs*

Ministère de l'Enseignement supérieur et de la Recherche, 2021, *Plan national pour la science ouverte* (2021-2024)

Ministère de l'Enseignement supérieur et de la Recherche et de l'Innovation, 2018, *Plan national pour la science ouverte* (2018-2021)

Monvoisin R., décembre 2012, « Recherche publique, revues privées », *Le Monde diplomatique*, p. 27, [https://www.monde-diplomatique.fr/2012/12/MONVOISIN/48501]

Mousseron P., 2023, *Droit des usages*, Paris, LexisNexis

Pontille D., 2016, *Signer ensemble*, Paris, Economica

Reymonet N., 2017, « Améliorer l'exposition des données de la recherche : la publication de *data paper* », [https://archivesic.ccsd.cnrs.fr/sic_01427978v1]

Reymonet N., Moysan M., Cartier A., Délémontez R., 2018, « Réaliser un plan de gestion de données "FAIR" : modèle », [https://archivesic.ccsd.cnrs.fr/sic_01690547v2]

Robin A., 2022a, *Droit des données de la recherche. Données publiques, innovation, science ouverte*, Bruxelles, Larcier

Robin A., 2022b, « Exploitation de l'innovation », *Cahiers droit, sciences & technologies*, n° 14, p. 225-243, [https://doi.org/10.4000/cdst.6214]

Robin A., 2020, « L'ouverture des données publiques scientifiques : de l'examen de la règle "*open as possible, closed as necessary*" », *Communication. Commerce électronique*, n° 9, étude 15, p. 5-10

Robin A., 2017 « Les données scientifiques au prisme du dispositif *open data* », *Communication. Commerce électronique*, n° 9, étude 14, p. 7-14

Robin A., Chatry S., 2022, *Introduction à la propriété intellectuelle. Unité et diversité*, 4e éd., Bruxelles, Bruylant

Schöpfel J., Rebouillat V., 2021, *Les entrepôts de données de recherche. Études de cas et enjeux*, Londres, ISTE Group

UNESCO, 2021, *Recommandation de l'Unesco sur une science ouverte*, [https://unesdoc.unesco.org/ark:/48223/pf0000379949_fre]

Vivant M., 2005, « De l'art de faire de la propriété intellectuelle un art de démobilisation », *Revue Lamy Droit de l'immatériel*, n° 9, p. 3

Westeel A., Claerr T. (dir.), 2010, *Numériser et mettre en ligne*, Villeurbanne, Presses de l'Enssib, [http://books.openedition.org/pressesenssib/414]

Wilkinson M. D. et al., 2016, « The FAIR guiding principles for scientific data management and stewardship », *Scientific Data*, *Nature*

NOTES

1 Commission européenne, *Pour un meilleur accès aux informations scientifiques : dynamiser les avantages des investissements publics dans le domaine de la recherche*, 17 juillet 2012, COM(2012) 401 final ; *Recommandation de la Commission du 17 juillet 2012 relative à l'accès aux informations scientifiques et à leur conservation* (points 1 et 2), JOUE L 194/39, 21 juillet 2012 ; *Recommandation de la Commission du 25 avril 2018 relative à l'accès aux informations scientifiques et à leur conservation*, JOUE L 143/12, 31 mai 2018.

2 Loi n° 2013-660 du 22 juillet 2013 relative à l'enseignement supérieur et à la recherche, JORF, 23 juillet 2013, suivie de l'ordonnance du 17 février 2014 modifiant la partie législative du Code de la recherche, JORF, 18 février 2014.

3 JORF, 8 octobre 2016.

4 Ces règles ont été intégrées dans le Code de la recherche par la création d'un nouvel article (art. L. 533-4 C. rech.) au sein du chapitre III dédié à la valorisation des résultats de la recherche par les établissements et organismes de recherche, qui contient l'intégralité des dispositions relatives aux résultats de la recherche scientifique.

5 Règlement (UE) 2021/695 du Parlement européen et du Conseil du 28 avril 2021 portant établissement du programme-cadre pour la recherche et l'innovation « Horizon Europe » et définissant ses règles de participation et de diffusion, et abrogeant les règlements (UE) 1290/2013 et (UE) 1291/2013 (art. 2, 5°, consid. 8), JOUE L 170/1, 12 mai 2021. Voir également dans le précédent règlement avec une autre formulation : Règlement (UE) 1291/2013 du Parlement européen et du Conseil du 11 décembre 2013 portant établissement du programme-cadre pour la recherche et l'innova-

tion « Horizon 2020 » (2014-2020) et abrogeant la décision n° 1982/2006/CE (art. 18), *JOUE* L 347/104, 20 décembre 2013.

6 Règlement (UE) 2021/695 « Horizon Europe », art. 14.

7 Directive (UE) 2019/1024 du Parlement européen et du Conseil du 20 juin 2019 concernant les données ouvertes et la réutilisation des informations du secteur public (art. 10), *JOUE* L 172/56, 26 juin 2019. Le fait que le principe soit énoncé au sein d'un dispositif relatif à l'ouverture des données publiques tend à brouiller les frontières entre le droit des données publiques et le droit des données de la recherche, c'est-à-dire entre les principes d'ouverture des données publiques et les principes d'ouverture des données de la recherche.

8 Voir par exemple le considérant 16 de la directive (UE) 2019/1024, précit., qui indique, s'agissant de l'ensemble des données publiques, que : « Les États membres sont donc encouragés à promouvoir, en ce qui concerne tous les documents relevant du champ d'application de la présente directive, la création de données fondées sur le principe d'"ouverture dès la conception et par défaut". »

9 Nous utiliserons le terme de « chercheur » dans un sens générique pour désigner toute personne qui exerce une activité de recherche soit à titre professionnel, soit de façon occasionnelle (avocats ou magistrats par exemple), quel que soit son sexe, son statut (chercheur, enseignant-chercheur) et ses fonctions ou son grade (professeur, maître de conférences, docteur, doctorant).

10 En 2014, cinq groupes détenaient 40 % de parts du marché en valeur de l'édition scientifique (CNRS, 2015 ; Assemblée nationale, 2015, p. 57). La concentration se construit autour d'une concurrence dite imparfaite, qui n'a pas pour effet de diminuer les prix d'accès aux revues, mais au contraire d'accroître leur niveau de rentabilité, celles-ci n'étant pas substituables entre elles du fait de la variabilité de leur notoriété (ou facteur d'impact). C'est donc la notoriété qui fixe le prix, l'économie de l'édition scientifique étant alors davantage une économie « d'exploitation de marque ». Le niveau de prix élevé est également et fortement lié au fait que les éditeurs sont ayants droit des auteurs chercheurs sur la base d'une cession, le plus souvent, à titre gratuit de leurs droits.

11 Agrégateurs de fonds numérisés munis d'outils d'exploration des données et d'outils d'analyse aidant au pilotage et à l'administration de la recherche.

12 Cette inflation fut l'élément déclencheur de la mobilisation de 2001, date de la publication de la pétition PLOS (Public Library of Science) des bibliothécaires universitaires, puis des chercheurs américains : « Le contribuable finance une recherche que le scientifique publiera – parfois à ses frais – dans une revue adossée à une entreprise privée, que d'autres chercheurs devront relire gratuitement et que les universités devront ensuite racheter à prix d'or. »

13 Initiative de Budapest pour l'accès ouvert, 14 février 2002.

14 Déclaration de Bethesda pour l'édition en libre accès (recherche biomédicale), 11 avril 2003.

15 Déclaration de Berlin sur le libre accès à la connaissance en sciences exactes, sciences de la vie, sciences humaines et sociales, 22 octobre 2003.

16 Le programme de recherche « CommonData » (« Les données de la recherche, des communs scientifiques ? ») s'attelle depuis 2017 au traitement de ces deux questions principales dont découlent de nombreuses autres (MSH SUD-LICeM), voir Robin (2022a).

17 Type de matériau utilisé, méthode retenue (empirique ou hypothético-déductive), etc.

18 [22] Les règles de science ouverte reposent en effet sur le principe d'ouverture de l'*accès* aux ressources utilisées ou produites par les chercheurs lors d'une démarche scientifique, c'est-à-dire un accès gratuit et ouvert à tous. Parce qu'elle revêt une importance capitale, la question de l'accès aux ressources en droit est généralement cantonnée à l'accès aux sources du droit et, en particulier, aux décisions de justice, alors qu'elle est en réalité plus vaste. Pour cette raison, la question de l'ouverture des données de la justice que d'autres contributeurs à cet ouvrage auront développée de manière approfondie, ne sera abordée que de manière ponctuelle.

19 Sans prétendre à l'exhaustivité, mentionnons les traités, constitution, lois, décrets, règlements, arrêtés, circulaires, journaux et bulletins officiels, délibérations et avis des autorités de régulation nationales ou européennes, instructions budgétaires et comptables, attributions, rapports, débats parlementaires et travaux législatifs, accords collectifs, avis des avocats généraux, rapports des conseillers, rapports d'activité, dictionnaire de l'Académie française, comptes de campagne des partis politiques, déclarations de patrimoine des personnes publiques, statuts des établissements publics, etc.

20 Loi n° 2016-1547 du 18 novembre 2016 de modernisation de la justice du XXIe siècle, JORF, 19 novembre 2016.

21 Les règles contenues dans le CRPA expriment le droit commun de la communication, de la diffusion et de la réutilisation des documents administratifs et des données publiques. Il est complété par d'autres dispositifs légaux, comme le Code de la justice administrative et le Code de l'organisation judiciaire, lorsque les données publiques sont des données de la justice, ou encore comme le Code de la santé publique, lorsque les données publiques sont des données de santé.

22 Par exemple, à l'occasion de l'adoption de la loi de programmation 2018-2022 et de réforme pour la justice, des règles dérogatoires au principe de l'ouverture des données de la justice (*open data* des décisions de justice) ont été perçues comme contraires au principe d'égalité entre les justiciables et une méconnaissance du principe d'égalité devant la loi et du droit à un procès équitable, dont découle le principe d'accès au droit. Les règles prévoyaient que les données d'identité des magistrats et des membres du greffe ne peuvent sous peine de sanctions pénales ou administratives, faire l'objet d'une réutilisation ayant pour objet ou pour effet d'évaluer, d'analyser, de comparer ou de prédire leurs pratiques professionnelles réelles ou supposées (art. L. 10 du Code de justice administrative ; art. L. 111-1 du Code de l'organisation judiciaire). L'argument n'a pas été retenu par le Conseil constitutionnel qui a estimé que l'objectif poursuivi par ces dispositions était d'éviter qu'une telle réutilisation permette de réaliser un profilage des professionnels de la justice, pouvant conduire à des pressions ou des stratégies de choix de juridiction de nature à altérer le fonctionnement de la justice (Cons. const., n° 2019/778 DC et n° 2019-779 DC, 21 mars 2019).

23 Dont l'exemple emblématique en France est, dans le domaine de la diffusion des décisions de justice, la société Forseti qui édite le site Doctrine.fr.

24 La recherche juridique s'avère, de ce point de vue, tout à fait limitée et ne permet d'observer que ce qui émerge lors des contentieux et encore à condition que le contentieux lui-même ne soit pas organisé de manière que les informations demeurent confidentielles (contentieux de l'arbitrage par exemple).

25 Notamment avec l'introduction des bases de données dans le giron du droit d'auteur et la création du droit *sui generis* des bases de données.

26 Règlement (UE) 2022/868, du Parlement européen et du Conseil du 30 mai 2022 portant sur la gouvernance européenne des données et modi-

fiant le règlement (UE) 2018/1724 (art. 2, 1°), JOUE L 152/1, 3 juin 2022.

27 « Donnée : représentation d'une information sous une forme conventionnelle destinée à faciliter son traitement » (*Bulletin des bibliothèques de France*, 1982).

28 Par exemple, la conception d'une base de données composée d'œuvres cinématographiques oblige à considérer ces œuvres comme des données.

29 Données publiques, données de la recherche, données privées, donnés de santé, données personnelles, données environnementales, etc.

30 Cette conception totalisante est d'ailleurs aggravée par le choix du législateur européen d'englober également dans la définition précitée les bases de données elles-mêmes (« compilations »), Règlement (UE) 2022/868, précit.

31 Directive (UE) 2019/1024, précit., art. 2, 9°.

32 À l'occasion d'une réflexion plus théorique, nous avons proposé une autre définition des données de la recherche, plus large que la définition légale applicable en matière de données publiques, en ce qu'elle n'exclut pas les données qui ne permettraient pas de valider les résultats de la recherche, c'est-à-dire d'administrer la preuve scientifique. Les données de la recherche seraient, sur le plan théorique, considérées comme toutes les données utilisées et produites par les chercheurs dans le cadre d'une activité de recherche scientifique. Cette définition permet d'inclure tout type d'opération scientifique, de la plus automatisée (collecte de données brutes issues de capteurs environnementaux par exemple) à la plus sophistiquée (production de données triées, annotées et analysées dans le cadre d'une recherche en matière de textes juridiques ou de décisions de jurisprudence par exemple).

33 Directive (UE) 2019/1024, précit., art. premier, 1°.

34 Directive (UE) 2019/1024, précit., art. premier, 2°, l).

35 Voir *infra*, point 1.1.1.

36 La notion d'« écrits scientifiques » sera préférée à celle de « publications scientifiques » que la directive n'explicite pas. Tout d'abord, la notion d'« écrits scientifiques » est celle retenue par le législateur français à l'article L. 112-1, mais aussi à l'article L. 533-4 C. rech. Ensuite, elle s'avère plus large que la seconde qui n'appréhende la production scientifique que par sa capacité à être publiée. Or, la protection des œuvres de l'esprit par le

droit d'auteur n'est pas conditionnée par le fait de la publication, mais seulement par le fait de l'expression.

37 La distinction issue de la loi n'englobe pas en effet les ressources scientifiques présentées sous la forme orale. Bien qu'il eût été intéressant de procéder à l'analyse de ces corpus, en particulier depuis que ces ressources sont quasi systématiquement enregistrées, puis diffusées sur Internet, nous prenons le parti de ne pas les aborder dans le cadre de cette étude en raison du fait que leur exploitation ne présente pas les mêmes enjeux sur le plan économique que les productions écrites.

38 Directive (UE) 2019/1024, précit., art. 2, 9° ; Règlement (UE) 2021/695 « Horizon Europe » ; art. L. 112-1 et L. 533-4. C. rech.

39 Le cycle de vie des données traduit les différentes étapes de la vie des données : la collecte, le traitement, le stockage et l'archivage, la diffusion, la réutilisation.

40 Convention de Berne, 9 juillet 1886 ; Convention universelle de Genève, 6 septembre 1952.

41 Art. L. 112-2, 1°, du *Code de la propriété intellectuelle* (CPI).

42 Cette protection est admise de longue date : la Cour royale de Paris admettait par jugement du 27 août 1828 que « sans doute un professeur doit à ses élèves dans son cours le tribut de ses études, de ses travaux, de ses méditations mais […] il ne leur doit que pour leur instruction personnelle et non pour qu'ils puissent s'en emparer et les publier en cours d'ouvrage pour en recueillir le bénéfice pécuniaire ; […] ces leçons envisagées sous cet aspect sont la propriété du professeur, le fruit de ses veilles, de ses recherches, de ses réflexions, de son génie […], elles forment pour lui une propriété littéraire […] et nul n'a le droit de s'en emparer et de les publier contre sa volonté » (Galvez-Behar, 2020, p. 59).

43 Art. L. 111-1 CPI.

44 La qualification d'œuvre collective est généralement retenue pour les encyclopédies, ouvrages collectifs, actes de colloque, etc.

45 Art. L. 121-1 et suiv. et L. 122-1 et suiv. CPI.

46 Art. L. 121-1 CPI.

47 Art. L. 121-4 CPI.

48 La distribution électronique suppose elle-même une structuration des données au sein de vastes bases de données (catalogues) et l'appariement de

ces bases entre elles.

49 Art. L. 121-7-1 et L. 131-3-1 à L. 131-3-3 CPI.

50 L'indépendance des enseignants-chercheurs est garantie par un principe fondamental reconnu par les lois de la République : Cons. const., 20 janvier 1984, n° 83-165 DC et 6 août 2010, n° 2010-20/21 QPC.

51 Art. L. 952-2 du Code de l'*éducation*.

52 Même s'il peut être exercé en indivision (en cas de création plurale, par exemple, lorsqu'un ouvrage est co-écrit par plusieurs auteurs) auquel cas l'œuvre est soumise à l'application des dispositions relatives aux œuvres de collaboration de l'article L. 113-3 CPI. Pour une analyse de la structure juridique du droit de propriété intellectuelle et du droit d'auteur en particulier.

53 Pour les publications dans des périodiques, voir le site du Baromètre de la science ouverte, piloté par le ministère de l'Enseignement supérieur et de la Recherche.

54 Et non à leur employeur par exemple, ce qui pourrait être le cas si l'on considère le cas des États-Unis dont la loi prévoit le principe d'une attribution des droits au profit de l'employeur en cas de création salariée (Bernault, 2016). On pourrait par suite imaginer que les employeurs des chercheurs, auteurs d'écrits scientifiques, négocient collectivement les redevances dues aux auteurs, tout comme ils négocient les conditions d'accès aux œuvres par abonnement.

55 Art. L. 112-1 C. rech.

56 L'absence de rémunération des chercheurs est souvent justifiée, en France du moins, par leur statut de fonctionnaire qu'il est censé compenser.

57 Directive (UE) 2019/1024, précit., consid. 27.

58 Par exemple la revue *Cahiers Droit, science et technologies*, éditée par le GDR NoST ou encore la revue *Questions constitutionnelles*, éditée par le laboratoire Cercop de l'Université de Montpellier.

59 Sur la question, voir également *infra*, point 1.2.2.

60 Compte tenu des effets de la concentration du marché, voir *supra* « Introduction ».

61 Sur lesquelles ils bénéficient également d'un droit *sui generis*. Les ressources textuelles utilisées pour la recherche juridique deviennent des données dont l'exploitation est valorisée par les éditeurs scientifiques. Voir *infra*, point 2.2.1.

62 Par exemple, de la revue *Concurrences*.

63 Il faut exclure du champ économique les mémoires et les thèses non publiés par des éditeurs, mais conservés dans des bibliothèques et diffusés par elles.

64 Les plateformes d'archives ouvertes sont multiples et variées. La principale est, en France, celle proposée par le CNRS : Hyper Article en Ligne (HAL).

65 Art. L. 533-4, I., C. rech.

66 Art. L. 533-4, III, C. rech.

67 De ce point de vue, l'ordre public peut être qualifié d'ordre public de protection, même s'il présente également, à notre avis, les caractéristiques d'un ordre public de direction (Robin A., 2022a, n° 582).

68 Cette règle vaut également pour la diffusion des données de la recherche, voir *infra* point 2.2.1.

69 Règlement (UE) 2021/695 « Horizon Europe », précit.

70 En ce sens, elle soutient la Déclaration de San Francisco sur l'évaluation de la recherche selon laquelle la recherche doit être évaluée sur ses propres mérites plutôt que sur la base du mode de publication de la recherche.

71 Prérogatives auxquelles il renonce de toute façon dans de nombreux champs de la recherche scientifique à travers les contrats de cession de droits qui stipulent généralement une cession à titre gratuit. Le domaine de la recherche juridique fait, à ce sujet, figure d'exception.

72 Le ministère préconise d'informer l'éditeur selon une formule comme suit : « Cette recherche a été financée, en tout ou partie, par l'Agence nationale de la recherche (ANR) au titre du projet ANR-nn-XXXX-nnnn. À des fins de libre accès, une licence CC-BY a été appliquée par les auteurs au présent document et le sera à toute version ultérieure jusqu'au manuscrit auteur accepté pour publication résultant de cette soumission » ou encore « À des fins de diffusion en accès ouvert, une licence CC-BY a été appliquée par les auteurs au présent document et le sera à toute version ultérieure jusqu'au manuscrit auteur accepté pour publication résultant de cette soumission ». Une liste des éditeurs refusant la stratégie NCD (non-cession de droits) en cas de financement par l'ANR ou l'un des membres de la cOAlition S est fournie, voir [https://journalcheckertool.org/exception-lists/].

73 Cette dichotomie exclut d'autres qualifications comme celle de données à caractère personnel. Cette exclusion est justifiée par le fait que le régime juridique des données publiques prévoit les conditions dans lesquelles les données publiques à caractère personnel doivent être diffusées. La loi prévoit l'obligation de les anonymiser, obligation que l'on retrouve à propos de la mise à disposition des décisions de justice.

74 Directive (UE) 2019/1024, précit., art. 2, 6°.

75 C'est la notion de « document administratif » qui est retenue en droit français. L'article L. 300-2 CRPA prévoit-il que : « Sont considérés comme documents administratifs, au sens des titres Ier, III et IV du présent livre, quels que soient leur date, leur lieu de conservation, leur forme et leur support, les documents produits ou reçus, dans le cadre de leur mission de service public, par l'État, les collectivités territoriales et par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission. Constituent de tels documents notamment les dossiers, rapports, études, comptes rendus, procès-verbaux, statistiques, instructions, circulaires, notes et réponses ministérielles, correspondances, avis, prévisions, codes sources et décisions. »

76 Dans les dispositions relatives à la réutilisation, le droit français distingue, par exemple, le document des informations qu'il contient pour affirmer le principe de l'autorisation qui doit être donnée au public par les administrations de réutiliser lesdites informations. L'article L. 321-1 CRPA dispose que « les informations publiques figurant dans des documents communiqués ou publiés par les administrations mentionnées au premier alinéa de l'article L. 300-2 peuvent être utilisées par toute personne qui le souhaite à d'autres fins que celles de la mission de service public pour les besoins de laquelle les documents ont été produits ou reçus. »

77 Où l'on voit comment la notion de données « noyaute » à rebours et, sans dire son nom, l'ensemble des règles légales applicables aux documents administratifs. Cette confusion notionnelle est aggravée par le fait que les données de la recherche sont précisément elles-mêmes définies comme des « documents » recueillis ou produits au cours d'activités de recherche, voir *infra*, point 2.1.1, § 3.

78 Directive (UE) 2019/1024, précit., art. premier, 2°.

79 Directive (UE) 2019/1024, précit., art. 3 et 6.

80 Directive (UE) 2019/1024, précit., art. premier, 1°, a) et b).

81 Depuis l'adoption de la loi n° 2016-1321 du 7 octobre 2016 pour une République numérique, les données de la justice sont soumises au principe d'ouverture par défaut, comme toutes les autres données des administrations et des collectivités publiques. L'arsenal législatif a été complété par la loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice (Décret n° 2020-797 du 29 juin 2020 relatif à la mise à disposition du public des décisions des juridictions judiciaires et administratives, *JORF*, 30 juin 2020 ; Décret n° 2021-1276 du 30 septembre 2021 relatif aux traitements automatisés de données à caractère personnel dénommés « Décisions de la justice administrative » et « Jurilibre », *JORF*, 1er octobre 2021). L'ensemble des décisions rendues par les juridictions de l'ordre administratif (hors juridictions spécialisées) et les décisions rendues par la Cour de cassation et par les cours d'appel en matière civile, sociale et commerciale sont mises à disposition. En 2023, cette mise à disposition concerne également les décisions rendues en matière civile, sociale et commerciale de neuf tribunaux judiciaires.

82 Les archives publiques sont, en effet, soumises aux dispositions applicables aux données publiques (art. L. 211-4 du Code du patrimoine).

83 Art. L. 312-1-3 CRPA.

84 *Arrêté du* 27 juin 1989 relatif à l'enrichissement du vocabulaire de l'informatique, *JORF*, 16 septembre 1989, p. 11725, annexe I.

85 Sur le principe d'explicabilité et de transparence des algorithmes, voir proposition de règlement du 21 avril 2021 établissant des règles harmonisées concernant l'intelligence artificielle, COM(2021) 206 final.

86 Voir *supra* « Introduction ».

87 L'article 1 de la directive dispose en effet qu'« afin de favoriser l'utilisation des données ouvertes et de stimuler l'innovation dans les produits et les services, la présente directive fixe un ensemble de règles minimales concernant la réutilisation et les modalités pratiques destinées à faciliter la réutilisation […] c) de données de la recherche conformément aux conditions définies à l'article 10 ».

88 Il s'agit en particulier des sciences expérimentales et, parmi elles, la recherche biomédicale. Néanmoins, certains auteurs ont proposé, de façon innovante et originale, d'étendre la question à la recherche juridique (Leclerc, 2021).

89 « Aussi ouvert que possible, aussi fermé que nécessaire » (Robin, 2020).

90 « Le concept de données ouvertes s'entend généralement comme désignant des données présentées dans un format ouvert qui peuvent être librement utilisées, réutilisées et partagées par tous quelle qu'en soit la finalité […] », voir directive (UE) 2019/1024, précit., consid. 16.

91 Directive (UE) 2019/1024, précit., consid. 27.

92 L'alliance compte aujourd'hui 8 418 membres à travers le monde.

93 Pour « Facile à trouver, Accessible, Interopérables, Réutilisables ». Voir [https://zenodo.org/records/6623556].

94 La norme FAIR, du fait de sa formulation, est l'illustration de ce que la norme technique peut explicitement contenir une norme juridique et signifier un « devoir-être ». Sur la dialectique entretenue entre norme technique et norme juridique, voir Boy (2007), pour qui « la norme technique relève incontestablement du devoir-être puisqu'elle impose un modèle à observer ». L'autrice montre que les normes techniques bénéficient d'une juridicité et participent depuis longtemps au pluralisme juridique. Nous ajouterons qu'il est intéressant d'observer qu'un standard technique porte le nom, en forme d'acronyme, d'une valeur, le *fair*, dont la traduction renvoie aux notions d'équité, de justice, mais aussi de raisonnable. La *fair* attitude (ou parfois utilisée en mot-valise la « fairitude ») consiste donc à assurer la diffusion et la réutilisation des données de la recherche scientifique.

95 Un *data paper* consiste dans le fait de publier un ou plusieurs jeux de données brutes de manière à les rendre lisibles et réutilisables, notamment à l'aide d'informations précises, appelées métadonnées. Il est publié sous forme d'un article examiné par les pairs soit une revue scientifique classique, soit *via* une plateforme de stockage et d'édition de données. Le *data paper* peut être produit en accompagnement ou indépendamment de la publication des résultats de la recherche.

96 Entrepôt produit par le CERN (Organisation européenne pour la recherche nucléaire) et financé, entre autres, par la Commission européenne.

97 L'outil est aujourd'hui très recommandé auprès des jeunes chercheurs, y compris dans le domaine de la recherche juridique, car elle permet non seulement de partager les jeux de données, mais aussi de créer des listes de références bibliographiques et de les importer dans n'importe quel autre document.

98 D'autres types d'information peuvent être couverts par le secret, mais seuls les modèles d'affaires seront ici envisagés.

99 Cette exception de fouille de textes et de données à des fins de recherche scientifique constitue une exception au droit d'auteur et au droit *sui generis* des bases de données et a été consacrée par la directive (UE) 2019/790, du Parlement européen et du Conseil du 17 avril 2019 sur le droit d'auteur et les droits voisins dans le marché unique numérique et modifiant les directives 96/9/CE et 2001/29/CE (art. 3), *JOUE* L 130/114, 17 mai 2019, et transposée par l'ordonnance n° 2021-1518 du 24 novembre 2021 complétant la transposition de la directive 2019/790 sur le droit d'auteur et les droits voisins dans le marché unique numérique et modifiant les directives 96/9/CE et 2001/29/CE (art. premier), *JORF*, 25 novembre 2021.

100 Après avoir défini la fouille de textes et de données comme « la mise en œuvre d'une technique d'analyse automatisée de textes et données sous forme numérique afin d'en dégager des informations, notamment des constantes, des tendances et des corrélations », l'article L. 122-5-3 CPI dispose que « des copies ou reproductions numériques d'œuvres auxquelles il a été accédé de manière licite peuvent être réalisées sans autorisation des auteurs en vue de fouilles de textes et de données menées à bien aux seules fins de la recherche scientifique par les organismes de recherche, les bibliothèques accessibles au public, les musées, les services d'archives ou les institutions dépositaires du patrimoine cinématographique, audiovisuel ou sonore, ou pour leur compte et à leur demande par d'autres personnes, y compris dans le cadre d'un partenariat sans but lucratif avec des acteurs privés. »

101 Loi n° 2011-1898 du 20 décembre 2011 relative à la rémunération pour copie privée (art. premier), *JORF*, 21 décembre 2011.

102 La plateforme ISTEX est une « plateforme numérique multi-usages » (base de données de bases de données) « aux meilleurs standards internationaux, accessible à distance par l'ensemble des communautés scientifiques et offrant "tous les moyens accessibles de consultation et d'analyse aujourd'hui disponibles dans toutes les communautés de la Science" » (CNRS, 2016, p. 9).

103 Résultat d'un partenariat entre le CNRS, l'ABES, le consortium Couperin et l'Université de Lorraine, l'ISTEX est une plateforme qui offre à l'ensemble de la communauté de l'enseignement supérieur et de la recherche française un accès à 28,2 millions d'articles scientifiques en 2025 (publiés de 1473 à

2023), issus de revues d'une vingtaine d'éditeurs internationaux et francophones.

104 Art. L. 321-3 CRPA.

105 Aujourd'hui, art. L. 533-4, II, C. rech.

106 Directive (UE) 2019/1024, précit., art. 10. Voir *supra* « Introduction ».

107 Qui est le pendant des dispositions relatives à l'ouverture des écrits scientifiques, voir *supra*, point 1.2.1.

108 Art. L. 533-4, III, C. rech.

109 Voir cependant le travail mené par l'Institut des usages, [https://institutdesusages.com], Faculté de droit de Montpellier.

110 Le principe d'ouverture des données de la recherche impose de tenir compte de la « nécessité », c'est-à-dire de toutes les normes prescrites pour tenir compte d'intérêts légitimement protégés. Sur ce point, voir Robin (2020). Le secret, quelle qu'en soit la nature, constitue l'une des grandes limites au principe d'ouverture de la science avec – mais dans une moindre mesure – les droits de propriété intellectuelle, la protection des données à caractère personnel, ainsi que les contraintes liées à la sécurité des données ou des informations.

111 Cette diffusion est imposée en raison du fait qu'il s'agit de contrats conclus grâce à une technologie de communication électronique (art. 1125 et suiv. du Code civil) et/ou de contrats conclus à distance en droit de la consommation (art. L. 221-5 du Code de la consommation).

112 Pour un exemple de recherche croisant la science juridique et l'intelligence artificielle, voir le projet OACS, Outil d'analyse de conformité des stipulations (Depincé, Robin, 2024).

113 Directive (UE) 2019/1024, précit., art. premier, 1°, b).

114 Directive (UE) 2019/1024, précit., art. 14. Voir également Règlement d'exécution (UE) 2023/138 du Parlement européen et du Conseil du 21 décembre 2022 établissant une liste d'ensembles de données de forte valeur spécifiques et les modalités de leur publication et de leur réutilisation, JOUE L 19/43, 20 janvier 2023.

115 Règlement (UE) 2022/868 du Parlement européen et du Conseil du 30 mai 2022 portant sur la gouvernance européenne des données et modifiant le règlement (UE) 2018/1724, JOUE L 152/1, 3 juin 2022.

116 Règlement (UE) 2023/2854 du Parlement européen et du Conseil du 13 décembre 2023 fixant des règles harmonisées pour l'équité de l'accès aux données et de l'utilisation des données, *JOUE* série L, 22 décembre 2022.

117 Proposition de règlement du Parlement européen et du Conseil relatif à l'espace européen des données de santé, COM(2022) 197 final, 3 mai 2022.

118 Le Centre d'accès sécurisé aux données (CASD) permet, par exemple, d'accéder, à des fins de recherche scientifique, aux données statistiques produite par l'INSEE et couvertes par le secret. Le CASD est un groupement d'intérêt public rassemblant l'État représenté par l'INSEE, le GENES, le CNRS, l'École polytechnique et HEC Paris créé par arrêté interministériel du 29 décembre 2018. Il est membre du réseau européen IDAN (Allemagne, France, Pays-Bas, Royaume-Uni) qui permet d'accéder à l'ensemble des données hébergées par chacun des centres grâce à l'utilisation d'une *safe room* répondant aux critères de sécurité nécessaires.

ABSTRACT

Français

Bien que non fondée sur l'observation des faits plutôt réservée à d'autres disciplines scientifiques, la science juridique utilise et produit un grand nombre de données, les « données de la recherche juridique ». Cet ensemble regroupe tant les écrits scientifiques dont la publication est assurée par les éditeurs que l'ensemble de la documentation publique contenant les sources du droit au sens large, mais aussi, le cas échéant, les ressources privées (données, savoir-faire, contrats). Il s'agit alors de mesurer le degré d'ouverture de ces données au regard des règles de science ouverte légalement consacrées successivement par le législateur français et par le législateur européen, obligeant ainsi à réinterroger l'ensemble du processus de production et de diffusion scientifique dans cette discipline.

INDEX

Keywords

science ouverte, données publiques, innovation, intelligence artificielle, recherche en droit, édition scientifique, droit d'auteur, secret

AUTHOR

Agnès Robin
Maître de conférences HDR, Université de Montpellier, LICeM (UR_UM213) ;
agnes.robin@umontpellier.fr