



Licences libres et Science Ouverte

Célian Godefroid, pour le Groupe Juridique du GTSO Couperin

Une licence libre est un acte unilatéral (qui n'est pas pris entre deux personnes mais par une seule) qui se transforme en contrat juridique dès que la partie à laquelle il est destiné fait usage des droits octroyés. Elle s'applique à une œuvre de l'esprit ou à un logiciel. Par la licence libre, l'auteur concède tout ou partie des droits que lui confère le droit d'auteur à toute personne morale ou physique.

1. Les critères de choix d'une licence

Plusieurs facteurs peuvent influencer sur le choix d'une licence de diffusion plutôt qu'une autre, comme le fait de vouloir octroyer certains droits plutôt que d'autres aux éventuels réutilisateurs (1.), ou bien des exigences posées par la loi (2.).

1.1. Les droits octroyés par les licences

L'objectif d'une **licence libre ou de libre diffusion** est, avant tout, de **permettre un accès le plus large possible à la production intellectuelle visée par la diffusion**, en permettant notamment le partage et la reproduction d'une œuvre. C'est, par ailleurs, la motivation principale de celles et ceux qui font le choix d'« ouvrir » leur recherche. Ainsi, un jeu de données scientifiques placé sous licence libre peut être librement repartagé et réutilisé dans le cadre d'autres recherches, à condition de toujours en citer les auteurs initiaux (c'est le « droit moral » d'un auteur, qui est inaliénable).

Toutes les licences libres ou de libre diffusion permettent la réutilisation des données de recherche placées sous leur régime. Si l'ouverture des données de la part des organismes producteurs est de plus en plus répandue, le fait de chercher si des jeux de données disponibles pourraient correspondre à leurs besoins et leur éviter des expérimentations, n'est toutefois pas encore un réflexe dans les communautés scientifiques.

Beaucoup de chercheurs ont pris le pli d'une ouverture des données, mais beaucoup craignent encore une dépossession de leurs données ou de leurs résultats de recherche, par une appropriation frauduleuse, ou une commercialisation par une société privée qui leur semblerait injuste. Aussi, certains font le choix d'une licence de réutilisation qui n'autorise pas d'utilisation commerciale, comme les

¹ Ce travail est placé par ces auteurs sous Licence Ouverte 2.0 (<https://www.etalab.gouv.fr/licence-ouverte-open-licence/>) et sous licence Creative Commons-Attribution 4.0 (<https://creativecommons.org/licenses/by/4.0/>), concédant notamment le droit de le reproduire, copier, adapter, modifier, communiquer, traduire et exploiter, sous réserve de mentionner la paternité du travail et sa source.

licences Creative Commons portant l'option « Non-Commercial ». Cela n'aurait pas pour effet de prohiber définitivement toute utilisation commerciale autour de la production intellectuelle protégée, **mais l'ayant-droit s'en réserve la prérogative, et une telle cession ne serait autorisée qu'à la suite d'un accord à part, convenu entre lui et le tiers souhaitant en faire une réutilisation commerciale.** Si un contenu est placé, par exemple, sous une licence CC-BY-NC, alors son auteur autorise que le contenu soit repartagé, réutilisé, sauf à des fins commerciales, par une société privée par exemple.

On prendra garde au fait qu'une **licence restreignant les types d'utilisation en fonction de leur finalité, notamment commerciale, ne remplit pas les critères pour être qualifiée de « libre », et tombe dans la catégorie des licences « de libre diffusion »** ; cette distinction peut emporter des conséquences quand le financeur d'une activité de recherche requiert que celle-ci soit diffusée « librement ».

Aux choix personnels des chercheurs s'ajoutent des exigences légales ou contractuelles propres au mode de financement de la recherche.

1.2. Exigences juridiques autour des licences

En France, la Loi pour une République numérique est venue instaurer deux dispositions à l'article L. 533-4 du Code de la recherche : la possibilité pour les chercheurs de repartager leurs articles en accès ouvert après une période d'embargo et l'instauration d'un principe de libre réutilisation des données de la recherche publiées.

Les articles scientifiques disposent d'un régime juridique particulier en droit français, en ce qu'ils constituent une œuvre de l'esprit, dont la propriété intellectuelle n'échoit pas automatiquement à l'employeur (dans le cas présent, l'organisme de recherche de tutelle). Les droits moraux et patrimoniaux demeurent la propriété du chercheur, qui décide seul où publier son article : il est pour l'heure difficile de le contraindre à repartager son article dans une archive ouverte comme HAL.

Les données de la recherche répondent à une logique différente : elles ne constituent pas des œuvres de l'esprit, mais des bases de données, dont les droits patrimoniaux échoient au financeur majoritaire de l'activité de recherche (souvent, un organisme de recherche public). Elles constituent des documents administratifs, communicables au titre de la loi CADA. Leur mise à disposition doit être proactive, par les producteurs des jeux de données, et permettre une réutilisation « libre ». Comme vu précédemment, cela nécessite une licence la plus large possible, ne restreignant pas outre mesure les finalités de réutilisation, notamment d'éventuelles finalités commerciales.

De plus, la licence de réutilisation sous laquelle doivent être placés les résultats de la recherche publique et les données ayant permis d'y parvenir n'est pas entièrement laissée au libre choix des chercheurs : une liste limitative de licences agréées est émise par décret et codifiée à l'article D. 323-2-1 du Code des relations entre le public et l'administration. De façon surprenante, les licences Creative Commons ne font pas partie des licences autorisées. **Il n'est donc, techniquement, pas autorisé de partager des articles ou des données de recherche sous licence Creative Commons, quand bien même cela serait la licence la plus répandue et connue du monde universitaire. L'État a élaboré une licence de réutilisation**

spécifique, la « Licence ouverte de réutilisation d'informations publiques », ou licence Etalab 2.0, dont les permissions sont les mêmes que la licence Creative Commons-Attribution. D'autres licences adaptées aux logiciels sont proposées, sans solution souveraine française.

Par ailleurs, des agences de financement, comme l'Agence Nationale de la Recherche française, présentent des exigences d'ouverture des données et des résultats de la recherche : depuis 2022, l'ANR exige, pour les publications issues de travaux financés par elle, une mise en ligne du texte de l'article sous licence Creative Commons-Attribution (CC-BY)², et son dépôt simultané à la publication dans une revue sur une plateforme ouverte comme HAL.

De plus, de nombreux établissements de recherche, organismes de recherche de tutelle, et même le Ministère de l'Enseignement supérieur et de la Recherche, via le comité « Ouvrir la Science ! », recommandent l'utilisation de la licence Creative Commons-Attribution pour le placement sous licence libre des publications et des données de la recherche.

2. Quelle licence choisir ?

2.1. Publications scientifiques et données de la recherche

D'après la loi les jeux de données partagés doivent permettre une « réutilisation libre », c'est-à-dire sans entrave, notamment au niveau des réutilisations possibles et plus particulièrement des réutilisations commerciales. La licence de réutilisation choisie doit également figurer parmi la liste de licences agréées par décret. Il est aussi possible d'utiliser la licence Open Database Licence, qui autorise la réutilisation d'un jeu de données, à la condition que le produit final soit lui aussi publié sous licence de libre partage. Cette licence n'étant pas « libre », il n'est pas certain qu'il soit souhaitable de la conseiller à grande échelle, pour lui préférer la licence Etalab.

Cela n'est pas applicable à tous les documents déposés dans HAL : le repartage après embargo permis par la Loi pour une République numérique n'autorise qu'un « accès ouvert », c'est-à-dire une libre diffusion par l'auteur, sans que cela ne vienne remettre en question une éventuelle cession de droits, même perpétuelle, consentie au profit de l'éditeur scientifique par les auteurs de l'article. Si l'éditeur n'a apposé aucune licence au moment de la publication, **il est indispensable que le document partagé sur HAL le soit sans licence** (ou « copyright », même si ce régime ne trouve aucune réalité en droit français, qui parle plutôt de « droit d'auteur »).

En revanche, si une stratégie de non-cession de droits est appliquée par l'équipe de recherche lors des négociations avec les éditeurs, alors le fichier auteur partagé dans HAL pourra l'être sous la même licence que celle appliquée sur les manuscrits auteur transmis à l'éditeur.

Une solution pour être en conformité avec la loi tout en continuant de bénéficier de la popularité et de l'identification à l'international des licences Creative Commons, est d'apposer les deux licences à la

² Suivant la stratégie de non-cession de droits, consistant à placer le texte des manuscrits auteur transmis aux éditeurs pour relecture, sous licence libre ou de libre réutilisation, permettant de contourner une clause de cession exclusive de droits contenue dans le contrat d'édition.

fois, dès que cela est permis : la licence CC-BY et la licence Etalab, dans la mesure où toutes deux sont compatibles et emportent les mêmes autorisations de réutilisation. Lorsqu'il n'est possible de ne faire figurer qu'une licence, il est préférable de faire figurer prioritairement la licence Etalab, quitte à faire mention de la licence Creative Commons ailleurs (page de garde, description du jeu de données...).

Pour les dépôts de preprint, on fera attention à la licence apposée : une licence CC BY-SA empêchera de changer de licence pour un CC BY lors du dépôt de la version publiée. Là encore, le choix d'une licence plus libérale est à privilégier.

2.2. Logiciels issus de la recherche

Les logiciels, contrairement aux publications scientifiques et aux jeux de données, peuvent être placés sous une licence avec obligation de réciprocité, sans contrevenir à une disposition législative. Il existe un outil en ligne à disposition des développeurs de logiciel, permettant de trouver la licence la plus adaptée à sa situation : <https://choosealicense.com/>.

Licences libres	
Licence Apache	Toute réutilisation autorisée Citer les auteurs originaux Mentionner les parties modifiées
Licence CeCILL (CEA, CNRS, INRIA Logiciel Libre)	Applicable dès le téléchargement du logiciel Litiges devant le juge français
Licence BSD (Berkeley Software Distribution)	« Domaine public », autorisation très large
Licence MIT	Incorporation du logiciel au sein d'un autre possible
Licences libres avec obligation de réciprocité	
Licence CeCILL-B/CeCILL-C	CeCILL-B : obligation de documenter le code CeCILL-C : obligation de placer le code réutilisé sous licence libre
Licence GNU GPL/LGPL/AGPL	Réutilisations sous licence identique LGPL : réutilisations sous licence libre, pas forcément LGPL AGPL : toutes les réutilisations, même privées, sous licence AGPL
Licence MPL (Mozilla Public License)	Mentionner le code modifié
Licence EPL (Eclipse Public License)	Incorporation possible, sous licence libre ou non, mais sans modification Si le code est modifié, placer sous licence EPL

Licence EUPL
(European Union Public License)

Vise une conformité avec les 27 États
Litiges devant la CJUE