Proposition de modèle de politique pour les revues et éditeurs quant aux données de la recherche

Chloée Fabre, Françoise Catherine Gouzi

▶ **To cite this version:**

Chloée Fabre, Françoise Catherine Gouzi. Proposition de modèle de politique pour les revues et éditeurs quant aux données de la recherche. 2020. hal-03026731

HAL Id: hal-03026731

https://univ-tlse2.hal.science/hal-03026731

Preprint submitted on 9 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Proposition de modèle de politique pour les revues et éditeurs quant aux données de la recherche

Chloée Fabre, Service Commun de la documentation (SCD), Université Toulouse – Jean Jaurès ; Françoise Gouzi, Direction d'Appui à la Recherche (DAR), Université Toulouse – Jean Jaurès

Sources: Hrynaszkiewicz, I, Simons, N, Hussain, A, Grant, R and Goudie, S. 2020. Developing a Research Data Policy Framework for All Journals and Publishers. *Data Science Journal*, DOI: https://doi.org/10.5334/dsj-2020-005 et https://datascience.codata.org/article/10.5334/dsj-2020-005/
Pour réutilisation de cette politique, utiliser la **licence Creative Commons Attribution (CC BY).**

Ce modèle s'articule avec les objectifs du Plan National pour la Science Ouverte, en particulier, avec le deuxième axe « **structurer et ouvrir les données de la recherche** » :

« Notre ambition est de faire en sorte que les données produites par la recherche publique française soient progressivement structurées en conformité avec les principes FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable), préservées et, quand cela est possible, ouvertes. […]. Dans le cadre du soutien public aux revues, la France recommandera l'adoption d'une politique de données ouvertes associées aux articles et le développement des data papers. Une politique similaire concernant les thèses sera également mise en place. […]. » (Source : https://www.ouvrirlascience.fr/plan-national-pour-la-science-ouverte/).

Ainsi qu'avec la **feuille de route du CNRS (axe 2, action 2)** :
« Développer la publication des données (data papers), le dépôt conjoint publications / données et accompagner les chercheurs et les chercheuses dans l'utilisation des outils de gestion des données. Cet objectif est une façon de s'engager concrètement dans la question du partage des données. Les données sur lesquelles s'appuient les publications doivent être rendues accessibles et réutilisables dès que possible. Cette exigence est déjà mise en œuvre dans le cadre du programme européen H2020 et doit être généralisée à toute donnée produite à partir de fonds publics. Il s'agira donc de faciliter le dépôt concomitant à la publication dans l'archive ouverte des jeux de données sur lesquels se basent la publication (ceci ne signifie en aucun cas qu'il faudrait déposer les données dans les archives ouvertes). Les jeux de données associées aux publications seront déposés dans un réservoir de données approprié, si possible un dépôt thématique qui assurera au mieux la qualité et la diffusion des données. Plus généralement le CNRS encourage la diffusion de toutes données structurées par leur dépôt dans des réservoirs de données thématiques ou généralistes, éventuellement accompagné de la publication de data papers. Pour ce faire, les instituts du CNRS en lien avec (les services de) la DIST, pourront identifier le cas échéant des structures ressources autour d'eux, comme les bibliothèques universitaires, les MSH (Maison des Sciences et de l'Homme) en SHS, ainsi que des personnes ressources à même d'accompagner les scientifiques dans ces nouvelles pratiques. Ces « voltigeurs » en charge des activités de curation de données préfigurent la reconnaissance effective du « data stewardship ». Les instituts, qui ont une politique de soutien aux revues, disposent d'un levier supplémentaire et pourraient inclure à l'avenir un volet sur les données liées aux publications. ». (Source : https://www.cnrs.fr/sites/default/files/press_info/2019-11/Plaquette_ScienceOuverte.pdf)

Résumé de la politique pour les auteurs
Les auteurs sont encouragés à partager en libre accès les données de la recherche et à fournir une « data availability statements » sans y être obligés. L'accord pour publication ne sera pas affecté par le fait que les auteurs partagent ou non leurs données de la recherche. **L'auteur doit :** ● Fournir une « Data availability statements » ● Citer les données sous-jacentes à l'article ● Partager les données via un entrepôt
Résumé de la politique pour les revues et éditeurs
La revue encouragera la diffusion en libre accès des données et publiera une « data availability statements » dans les articles publiés. L'accord pour publication ne sera pas affecté par le fait que les auteurs partagent ou non leurs données de la recherche. **La revue doit :** ● Répondre aux questions sur cette politique ● Publier une « data availability statements » lorsqu'elle est fournie par l'auteur ● Vérifier que les données sont citées ● Vérifier que les données sont partagées via un entrepôt

Définition des données de la recherche	Cette politique s'applique aux données de la recherche qui seraient nécessaires pour vérifier les résultats de la recherche présentés dans les articles publiés dans la revue. Les données de la recherche comprennent les données produites par les auteurs et les données provenant d'autres sources qui sont analysées par les auteurs dans leur étude. Les données de la recherche comprennent l'ensemble « 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 »[1].

[1] 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, s.l., 2019, vol.OJ L.

Document de travail, juin 2020, Chloée Fabre (SCD, UT2J) et Françoise Gouzi (DAR, UT2J), licence CC BY 4.0

Définition des exceptions	Les données de la recherche qui ne sont pas nécessaires pour vérifier les résultats rapportés dans les articles ne sont pas couvertes par cette politique. Cette politique sera limitée par les exceptions légitimes encadrées par la loi, par exemple en ce qui concerne le secret professionnel, les secrets industriels et commerciaux, les données personnelles ou les contenus protégés par le droit d'auteur. Les alternatives à la diffusion en libre accès de données personnelles ou sensibles sont : ● Anonymisation ou pseudonymisation des données avant diffusion en libre accès ● Diffusion uniquement des métadonnées relatives aux données
Délai de mise à disposition	L'accès ouvert aux données est fortement recommandé sans embargo ou après un délai le plus court possible n'excédant pas celui de la publication.
Entrepôts de données	« Les chercheurs seront invités à déposer les données dans des entrepôts de données certifiés, dont la gouvernance et les règles de propriété intellectuelle seront conformes aux bonnes pratiques. À ce titre, les infrastructures nationales et européennes de recherche seront privilégiées, notamment via des centres de données thématiques et disciplinaires. »[2] Il est préconisé que la revue recommande l'utilisation d'un entrepôt de données certifié disciplinaire ou à défaut un entrepôt généraliste de type Zenodo.
Citation des données	La revue demande aux auteurs de citer toutes les données accessibles en libre accès dans leur liste de références. La citation de chaque jeu de données doit inclure un identifiant pérenne (par ex : DOI). Les citations de jeux de données, lorqu'elles apparaissent dans la liste des références, doivent inclure les informations minimales recommandées par DataCite[3] et prendre en compte le style bibliographique imposé par la revue (consignes aux auteurs). Les citations peuvent être générées automatiquement via l'outil : https://citation.crosscite.org/
Licence pour les données	La revue encourage la mise à disposition des données sous des licences ouvertes qui permettent leur libre réutilisation. Les auteurs doivent utiliser les licences préconisées par l'entrepôt où les données ont été déposées. L'éditeur de la revue ne revendique pas de droits d'auteur sur les données.
Soutien aux auteurs	Les questions concernant le respect de cette politique doivent être envoyées à [**email**].

[2] https://www.ouvrirlascience.fr/plan-national-pour-la-science-ouverte/
[3] Consortium international dont l'objectif est de faciliter l'accès aux données de la recherche et leur réutilisation

Document de travail, juin 2020, Chloée Fabre (SCD, UT2J) et Françoise Gouzi (DAR, UT2J), licence CC BY 4.0

« Data availability statements »	La revue demande aux auteurs d'inclure une « data availability statements » dans tout article faisant état de résultats dérivés de données. Les « data availability statements » doivent inclure des informations sur les modalités d'accès aux données. Les « data availability statements » doivent prendre l'une des formes suivantes (ou une combinaison de celles-ci quand plusieurs types de données) : · Les jeux de données générés et/ou analysés au cours de la présente étude sont disponibles dans l'entrepôt [Nom de l'entrepôt] ; [DOI] · Les jeux de données générés et/ou analysés au cours de la présente étude ne sont pas accessibles en libre accès en raison de [**spécifier la raison**] mais sont disponibles auprès de l'auteur sur demande justifiée. · Les jeux de données générés et/ou analysés au cours de la présente étude sont disponibles auprès de l'auteur sur demande justifiée. · Le partage de données ne s'applique pas à cet article car aucun jeu de données n'a été généré ou analysé au cours de la présente étude.

Document de travail, juin 2020, Chloée Fabre (SCD, UT2J) et Françoise Gouzi (DAR, UT2J), licence CC BY 4.0