

Open Science

Outils, Perspectives et Enjeux des Numériques (OPEN)

Elise Hocquette (P52)

2023-02-13

Contents

1	Introduction : Qu'est-ce que l'Open Science ?	1
2	Les parties prenantes et acteurs en présence	2
2.1	Les éditeurs, les maisons d'édition, les comités éditoriaux	2
2.2	Les pouvoirs publics	2
2.3	Les organismes de recherche	3
2.4	Les chercheurs, à la fois auteurs et lecteurs	3
2.5	Les plateformes de diffusion	3
3	Les principes et les problématiques majeures	3
3.1	Les principes de l'Open Science	3
3.2	Publication des articles scientifiques en open access	4
3.3	Diffusion des données en open access	4
4	Avantages & intérêts	5
5	Critiques & limites à la mise en oeuvre	5
6	Conclusion générale	5
7	Bibliographie (Zotero)	6

1 Introduction : Qu'est-ce que l'Open Science ?

L'Open Science est une approche collaborative qui vise à rendre la recherche scientifique plus transparente, ouverte, et surtout accessible à tous et à tous les niveaux de la société. En effet, cela permet le libre accès aux données, méthodes et résultats de la recherche. Le but est de démocratiser l'accès aux savoirs, mais aussi d'améliorer l'efficacité de la recherche en offrant la possibilité à d'autres scientifiques de réanalyser des données déjà publiées afin de confirmer les résultats et les interprétations des premières analyses. L'Open

Science suppose donc des connaissances scientifiques transparentes, accessibles, partagées, et fondées sur la collaboration entre chercheurs (Molloy, 2011 ; Vicente-Saez, Martinez-Fuentes, 2018).

Ainsi, l'Open Science comprend notamment en accès libre : la publication d'articles scientifiques pour les lecteurs, ainsi que les données de la recherche pour que les chercheurs puissent les utiliser gratuitement. Les exemples de domaines d'application sont nombreux : en mathématiques, en astronomie, en génétique, en médecine... (Walport, Brest, 2011).

Cette synthèse s'orientera autour des principes et problématiques de l'open science, et étudiera comment promouvoir cette politique. Dans ce rapport, une première partie présentera les acteurs et parties prenantes, puis une deuxième partie expliquera les principes et les deux problématiques majeures de l'Open Science (articles et données en libre accès), ensuite une quatrième partie exposera les avantages et intérêts de l'Open Science, et enfin une cinquième partie détaillera les critiques et limites à la mise en oeuvre de cette politique.

2 Les parties prenantes et acteurs en présence

2.1 Les éditeurs, les maisons d'édition, les comités éditoriaux

Il existe deux modèles de publications scientifiques. Dans le premier modèle, ce sont les lecteurs qui s'abonnent aux journaux scientifiques et qui, de ce fait, absorbent les coûts de publication, gratuits pour les auteurs. Ce sont le plus souvent les institutions (organismes de recherche et universités) qui payent des abonnements collectifs aux éditeurs. En effet, les bibliothèques, surtout les bibliothèques universitaires, payent des abonnements coûteux pour rendre accessible des bouquets de publications scientifiques à leurs enseignants-chercheurs et leurs étudiants. Ce modèle est critiqué parce que le monde de la publication scientifique est entre les mains d'éditeurs privés puissants, et les auteurs doivent abandonner leurs droits à ces éditeurs.

Dans le second modèle, la logique est inversée. Ce sont aux auteurs de payer les frais de publication, l'accès étant gratuit pour les lecteurs. Cela permet une plus grande diffusion des articles scientifiques (McKiernan *et al.*, 2016), mais a conduit à l'émergence d'éditeurs peu scrupuleux qui pratiquent des tarifs élevés associés à une rigueur scientifique qui reste à désirer.

La frontière entre ces deux modèles de publications est parfois floue. Ainsi, certains auteurs recommandent aux éditeurs de clairement expliciter leur politique d'Open Science et de réduire les frais de publication afin de faciliter la dissémination des résultats (Fleming, Cook, 2022).

2.2 Les pouvoirs publics

Les pouvoirs publics, notamment la communauté Européenne, font la promotion du principe de 'Science Ouverte', ce qui implique plusieurs choses (Vicente-Saez, Martinez-Fuentes, 2018). Tout d'abord, les chercheurs sont incités à publier dans des journaux scientifiques en « open access ». D'autre part, les chercheurs sont invités à rendre accessibles les données issues de leurs recherches, en les déposant dans systèmes d'informations adaptés.

L'Agence Nationale de la Recherche (ANR) a une certaine politique en la matière, en donnant des consignes. Elle recommande de concevoir et de déployer une politique et une culture de la donnée, dans le cadre d'une démarche de performance, de transparence et d'ouverture, inscrite dans son contrat d'objectifs 2021-2025. Cette politique s'articule avec la politique de Science Ouverte Européenne (Garnier-Rizet *et al.*, 2022).

Les collectivités financent la recherche publique et subventionnent une partie de la recherche privée. Cependant, ces collectivités restent attachées à la diffusion des connaissances. En Novembre 2021, les Nations Unies ont publié une recommandation en faveur de la Science Ouverte pour l'intérêt général (Cudennec *et al.*, 2022).

2.3 Les organismes de recherche

Les organismes de recherche et les Universités du secteur public sont tenus de mettre en œuvre les politiques gouvernementales, et notamment celles concernant la politique de l’Open Science. Pour cette raison, ils conçoivent des outils permettant la diffusion libre des données et des articles, ou si ce n’est pas le cas, ils utilisent des outils collectifs déjà performants. Ainsi, l’Institut National de Recherche pour l’Agriculture, l’Alimentation et l’Environnement (INRAE) a mis en place une politique de science ouverte basée sur 10 actions phares (Regnier, 2021).

2.4 Les chercheurs, à la fois auteurs et lecteurs

Les chercheurs sont au cœur des processus de communication scientifique, qui assure la transmission des savoirs. En tant qu’auteurs de publications scientifiques, ils peuvent choisir ou non de publier dans des archives ouvertes. Ce critère n’est pourtant pas le seul pris en compte. En effet, les chercheurs choisissent le journal dans lequel ils veulent publier leurs articles, en fonction de leur discipline, de la notoriété du journal, et de la qualité estimée de leurs travaux. Toutefois, le choix de publier ou non dans une archive ouverte contribue à soutenir plus ou moins la politique d’Open Science. Il y a même des initiatives de chercheurs qui vont jusqu’à proposer des modèles de journaux scientifiques en accès libre innovants. Au final, il semblerait que la proportion d’articles en open access ait tendance à s’accroître (Piwowar *et al.*, 2018).

2.5 Les plateformes de diffusion

Les premières plateformes qui ont hébergé les publications scientifiques sont des actions à accès limités (par exemples, Web of Science et Scopus), car parfois liés à des éditeurs. Plus récemment, d’autres plateformes gratuites (par exemples, Google Scholar et Research Gate) assurent l’hébergement de publications scientifiques, mais leurs critères de qualité sont différents. Les institutions de recherche en France préfèrent l’entrepôt de publication « Hal » (Mirowski, 2018).

3 Les principes et les problématiques majeures

3.1 Les principes de l’Open Science

L’Open Science peut être déclinée selon 6 principes majeurs (Gallagher *et al.*, 2020) :

- Le principe de l’Open data correspond à la diffusion des données de la recherche de façon transparente, exhaustive et gratuite.
- Le principe de l’Open resources est de former les acteurs aux méthodes de l’Open Science et d’augmenter l’accès aux ressources telles que les bases de données.
- Le principe de l’Open source est de rendre accessible les méthodes, les logiciels et les protocoles d’analyse.
- Le principe de l’Open access est de rendre le travail de recherche publiée plus accessible afin d’en assurer la plus grande diffusion possible.
- Le principe de l’Open methods correspond au développement de méthodes standards reproductibles et partagées entre les acteurs de la recherche, de sorte que les données soient comparables car obtenues par les mêmes méthodes.
- Le principe de l’Open peer review repose sur une expertise des articles scientifiques transparente, afin de promouvoir des méthodes d’évaluation des données et des résultats de la recherche plus rigoureuse.



Figure 1: Les principes de l'Open Science (Gallagher et al., 2020)

3.2 Publication des articles scientifiques en open access

La première problématique concerne l'équilibre entre les deux modèles de publication déjà évoqués : celui où le lecteur est payeur, et celui où l'auteur est payeur. L'existence de ces deux modèles pose un grand nombre de questions. Ainsi, les instituts de recherche et les universités sont mises à contribution deux fois, par deux mécanismes différents : leur direction paye un accès à Web of Science ou Scopus, tandis que les chercheurs individuellement payent la publication de chacun de leurs articles lorsqu'ils choisissent un éditeur en open access (Louembe, 2022).

3.3 Diffusion des données en open access

La deuxième problématique concerne la diffusion de données brutes et de codes informatiques associés à la recherche. Diffuser les données de recherches en accès libre est un principe louable mais pas suffisant. Ces données doivent répondre à des standards pour pouvoir être partagées, et surtout réutilisées. Le standard le plus connu est le standard FAIR (Findable, Accessible, Interoperable and Reusable), lié aux bonnes pratiques de gestion des données. Tout d'abord, les données doivent être faciles à trouver par les humains et les systèmes informatiques, en étant déposées dans des entrepôts, en leur attribuant un identifiant unique et pérenne, et en les indexant et en les décrivant par des métadonnées riches. Ensuite, les données doivent être accessibles, en étant stockées durablement, en facilitant leur accès et leur téléchargement et en spécifiant les conditions d'accès et les licences d'utilisation. Puis, les données doivent être interoperables, en privilégiant des formats ouverts ou largement utilisés, en utilisant des standards de métadonnées et de vocabulaire, en mettant à disposition le code source du logiciel nécessaire pour les lire, traiter et analyser, et en indiquant des liens vers d'autres ressources combinables. Enfin, les jeux de données doivent être réutilisables, en leur associant une licence de diffusion ainsi que de la documentation pour les décrire de façon détaillée, les contextualiser et les rendre facilement compréhensibles (Wilkinson *et al.*, 2016).

En complément, il est possible de faire une publication scientifique uniquement pour décrire des jeux de données, ce qui est appelé « Data paper ».

4 Avantages & intérêts

Le modèle en « open access » permet une diffusion plus grande et sans restriction des connaissances scientifiques, puisque les articles sont en libre accès pour n'importe qui. Il en est de même lorsque les données sont déposées sur un site d'information adapté. Ainsi, la réutilisation des données est grandement favorisée, permettant une nouvelle analyse éventuelle par d'autres laboratoires, afin de confirmer les résultats publiés.

Par ailleurs, la mise à disposition des données favorise les méta-analyses, c'est-à-dire les analyses conjointes de plusieurs jeux de données mis ensemble. Cette problématique de méta-analyse se développe énormément dans tous les champs scientifiques, et permet de dégager les lois générales qu'une seule étude ne peut pas dégager (Piwowar *et al.*, 2018). Plus généralement, l'Open Science est perçue comme un accélérateur de l'efficacité de la recherche, ce qui est particulièrement intéressant en médecine (Woelfle *et al.*, 2011).

5 Critiques & limites à la mise en oeuvre

Dans certains domaines scientifiques, la culture du partage des données n'est pas très répandue, prétextant la propriété intellectuelle. Ainsi, les chercheurs ne sont pas toujours enclins à diffuser leurs données, en tout cas avant leur publication dans des journaux scientifiques.

Par ailleurs, certaines entreprises du secteur privé sont invitées à rendre leurs données de recherche accessibles, mais elles sont souvent réfractaires à cela par souci de préserver le secret industriel ou commercial. C'est particulièrement le cas des technologies innovantes, comme les entreprises développant la viande de culture.

Certains chercheurs ont aussi du mal à changer leurs habitudes et préfèrent continuer de publier leurs travaux sur des plateformes payantes, qui confèrent une valeur plus élevée à leurs publications dans leur milieu de travail, et plus facilement reconnue dans le domaine de la recherche scientifique.

Par ailleurs, les scientifiques ont déjà une grande confiance dans le système traditionnel. De ce fait, une des faiblesses des plateformes est liée à un manque relatif de confiance par les chercheurs, compromettant ainsi leur réputation (Jamali *et al.*, 2016).

De plus, la réutilisation plus facile des données et des productions des chercheurs peut être jugée moins éthique (Eve, 2014). Compte tenu de ces limites, certains chercheurs proposent une stratégie de mise en oeuvre de l'Open Science, l'accès libre des articles et des données étant simplement une première étape vers une collaboration plus importante en toute transparence entre chercheurs (Dienlin *et al.*, 2021).

6 Conclusion générale

Pour conclure, il faut beaucoup de temps pour diffuser les principes de l'Open Science. C'est un changement de culture qui implique une évolution des mentalités, pour que les chercheurs soient vraiment motivés à complètement partager leurs données. Dans certains domaines, comme l'astronomie ou la génétique où le partage des données est un pré-requis (car les analyses requièrent d'emblée un grand nombre de données), cette culture est largement diffusée. Dans d'autres domaines, comme les sciences vétérinaires qui analyse souvent des cas spécifiques (à l'exception de l'épidémiologie), il est plus difficile d'avoir accès à des grands jeux de données comparables.

Par ailleurs, l'Open Science induit également un changement dans les méthodes de travail, avec la volonté dès le départ de concevoir des expérimentations et d'acquérir des données selon des principes (méthodes standards) permettant la comparaison de données issues de différents laboratoires. Cela nécessite aussi en amont la construction d'outils de partage (plateformes de publications scientifiques, plateformes de données).

Dans le cadre de ces deux objectifs (évolution des mentalités et construction d'outils de partage), les principaux acteurs de la recherche, que ce soit au niveau politique (financeurs) ou au niveau institutionnel (directions des organismes de recherches), mettent en place des politiques adaptées.

7 Bibliographie (Zotero)

La bibliographie suivante a été générée par Zotero :

- CUDENNEC, Christophe, SUD, Megha et BOULTON, Geoffrey, 2022. Governing Open Science. *Hydrological Sciences Journal* [en ligne]. décembre 2022. Vol. 67, n° 16, pp. 2359-2362. [Consulté le 12 février 2023]. DOI 10.1080/02626667.2022.2086462. Disponible à l'adresse : <https://doi.org/10.1080/02626667.2022.2086462>
- DIENLIN, Tobias, JOHANNES, Niklas et BOWMAN, Nicholas David, 2021. Agenda for Open Science in Communication. *Journal of Communication | Oxford Academic* [en ligne]. 2021. [Consulté le 12 février 2023]. Disponible à l'adresse : <https://academic.oup.com/joc/article/71/1/1/5803422>
- EVE, Martin Paul, 2014. *Open Access and the Humanities: Contexts, Controversies and the Future* [en ligne]. Cambridge : Cambridge University Press. [Consulté le 7 février 2023]. ISBN 978-1-107-09789-6. Disponible à l'adresse : <https://www.cambridge.org/core/books/open-access-and-the-humanities/02BD7DB4A5172A864C432DBFD86E5FB4>
- FLEMING, J. I. et COOK, B. G., 2022. Open Access in Special Education: A Review of Journal and Publisher Policies. *Remedial and Special Education*, 43(1), 3-14. [en ligne]. 2022. [Consulté le 12 février 2023]. Disponible à l'adresse : <https://journals.sagepub.com/doi/10.1177/0741932521996461>
- GALLAGHER, Rachael V., FALSTER, Daniel S., MAITNER, Brian S., SALGUERO-GÓMEZ, Roberto, VANDVIK, Vigdis, PEARSE, William D., SCHNEIDER, Florian D., KATTGE, Jens, POELEN, Jorrit H., MADIN, Joshua S., ANKENBRAND, Markus J., PENONE, Caterina, FENG, Xiao, ADAMS, Vanessa M., ALROY, John, ANDREW, Samuel C., BALK, Meghan A., BLAND, Lucie M., BOYLE, Brad L., BRAVO-AVILA, Catherine H., BRENNAN, Ian, CARTHEY, Alexandra J. R., CATULLO, Renee, CAVAZOS, Britany R., CONDE, Dalia A., CHOWN, Steven L., FADRIQUE, Belen, GIBB, Heloise, HALBRITTER, Aud H., HAMMOCK, Jennifer, HOGAN, J. Aaron, HOLEWA, Hamish, HOPE, Michael, IVERSEN, Colleen M., JOCHUM, Malte, KEARNEY, Michael, KELLER, Alexander, MABEE, Paula, MANNING, Peter, MCCORMACK, Luke, MICHALETZ, Sean T., PARK, Daniel S., PEREZ, Timothy M., PINEDA-MUNOZ, Silvia, RAY, Courtenay A., ROSSETTO, Maurizio, SAUQUET, Hervé, SPARROW, Benjamin, SPASOJEVIC, Marko J., TELFORD, Richard J., TOBIAS, Joseph A., VIOLLE, Cyrille, WALLS, Ramona, WEISS, Katherine C. B., WESTOBY, Mark, WRIGHT, Ian J. et ENQUIST, Brian J., 2020. Open Science principles for accelerating trait-based science across the Tree of Life. *Nature Ecology & Evolution*. mars 2020. Vol. 4, n° 3, pp. 294-303. DOI 10.1038/s41559-020-1109-6.
- GARNIER-RIZET, Martine, RASSINOX, Laurent et ANCION, Zoé, 2022. La politique science ouverte de l'ANR. *Agence Nationale de la Recherche (Tour)*. 2022.
- JAMALI, Hamid R., NICHOLAS, David et HERMAN, Eti, 2016. Scholarly reputation in the digital age and the role of emerging platforms and mechanisms. *Research Evaluation* [en ligne]. janvier 2016. Vol. 25, n° 1, pp. 37-49. [Consulté le 12 février 2023]. DOI 10.1093/reseval/rvv032. Disponible à l'adresse : <https://doi.org/10.1093/reseval/rvv032>
- LOUEMBE, Kareen, 2022. Publier en accès libre à INRAE, informations pratiques et coûts. *Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement (INRAE)* [en ligne]. 2022. Disponible à l'adresse : <https://www6.inrae.fr/dipso/Nos-outils-pour/S-informer-Fiches-pratiques-et-recommandations/Publier-en-acces-libre-a-INRAE-informations-pratiques-et-couts>
- MCKIERNAN, Eric, BOURNE, Philip et BROWN, Titus, 2016. Point of View: How open science helps researchers succeed | eLife. *eLife Sciences, Biochemistry and Chemical Biology* [en ligne]. 2016. [Consulté le 12 février 2023]. Disponible à l'adresse : <https://elifesciences.org/articles/16800>
- MIROWSKI, Philip, 2018. The future(s) of open science. *Social Studies of Science* [en ligne]. avril 2018. Vol. 48, n° 2, pp. 171-203. [Consulté le 12 février 2023]. DOI 10.1177/0306312718772086. Disponible à l'adresse : <https://doi.org/10.1177/0306312718772086>
- MOLLOY, Jennifer C., 2011. The Open Knowledge Foundation: Open Data Means Better Science. *PLOS Biology* [en ligne]. décembre 2011. Vol. 9, n° 12, pp. e1001195. [Consulté le 12 février 2023]. DOI 10.1371/journal.pbio.1001195. Disponible à l'adresse : <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1001195>
- PIWOWAR, Heather, PRIEM, Jason, LARIVIÈRE, Vincent, ALPERIN, Juan Pablo, MATTHIAS, Lisa, NORLANDER, Bree, FARLEY, Ashley, WEST, Jevin et HAUSTEIN, Stefanie, 2018. The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ* [en ligne]. février 2018.

Vol. 6, pp. e4375. [Consulté le 12 février 2023]. DOI 10.7717/peerj.4375. Disponible à l'adresse : <https://peerj.com/articles/4375>

REGNIER, Elodie, 2021. *La Politique de Science Ouverte d'INRAE*. 2021. Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement (INRAE), www.inrae.fr, Accueil > Toutes les actualités > Ouvrir les sciences, plus que jamais.

VICENTE-SAEZ, Ruben et MARTINEZ-FUENTES, Clara, 2018. Open Science now: A systematic literature review for an integrated definition. *Journal of Business Research* [en ligne]. juillet 2018. Vol. 88, pp. 428-436. [Consulté le 7 février 2023]. DOI 10.1016/j.jbusres.2017.12.043. Disponible à l'adresse : <https://www.sciencedirect.com/science/article/pii/S0148296317305441>

WALPORT, W et BREST, P, 2011. Sharing research data to improve public health. *The Lancet* 377: 537–539 [en ligne]. 2011. [Consulté le 12 février 2023]. Disponible à l'adresse : [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(10\)62234-9/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(10)62234-9/fulltext)

WILKINSON, MD, DUMONTIER, M et AALBERSBERG, J, 2016. The FAIR Guiding Principles for scientific data management and stewardship | Scientific Data. *Scientific Data* 3, 160018 [en ligne]. 2016. [Consulté le 12 février 2023]. Disponible à l'adresse : <https://www.nature.com/articles/sdata201618>

WOELFLE, Michael, OLLIARO, Piero et TODD, Matthew H., 2011. Open science is a research accelerator. *Nature Chemistry* [en ligne]. octobre 2011. Vol. 3, n° 10, pp. 745-748. [Consulté le 7 février 2023]. DOI 10.1038/nchem.1149. Disponible à l'adresse : <https://www.nature.com/articles/nchem.1149>