

Baromètre français de la Science Ouverte 2021, 10 points en plus d'accès ouvert pour les publications scientifiques françaises

Selon l'édition 2021 du Baromètre de la Science Ouverte (BSO), 62 % des 166 000 publications scientifiques françaises publiées en 2020 sont en accès ouvert en décembre 2021. Ce taux progresse de 10 points en un an. Le niveau d'ouverture des publications varie de manière significative d'une discipline à l'autre. Certaines disciplines comme les sciences physiques et les mathématiques se sont engagées de longue date dans l'ouverture de leurs publications tandis que d'autres, comme la chimie, connaissent des processus rapides de rattrapage. L'édition 2021 du BSO suit aussi la transparence des résultats des essais cliniques menés en France. Il ressort que seulement 57 % des essais cliniques terminés ces 10 dernières années ont partagé publiquement leurs résultats.

Une hausse soutenue du taux d'accès ouvert aux publications entre 2018 et 2021

L'accroissement régulier du taux d'accès ouvert observé chaque année depuis 2018 est un indicateur de l'impact des politiques publiques en faveur de l'accès ouvert. Il passe de 38 % en 2018 à 52 % en 2020 et 62 % en 2021 soit + 24 points en 3 ans et + 10 points en un an. Cette progression témoigne de l'évolution des pratiques de publication des chercheurs, du renforcement des infrastructures de publication en accès ouvert et des stratégies des acteurs de l'édition scientifique. Cette hausse continue se retrouve aussi pour chaque discipline. Certaines affichent des taux d'ouverture plus élevés que la moyenne depuis 2018, comme les Mathématiques (57 % en 2018, 72 % en 2021). De plus, la progression est forte pour des disciplines qui avaient des taux plus bas en 2018, comme la chimie (+ 36 points d'ouverture depuis 2018).

Tableau 1 : Taux d'accès ouvert des publications françaises parues dans l'année précédant la date d'observation (en %)

Discipline	Accès ouvert en 2018 (en %)	Accès ouvert en 2021 (en %)	Évolution 2018-2021 (en points)
Biologie fondamentale	49	74	+ 25
Sciences physiques, Astronomie	52	73	+ 21
Mathématiques	57	72	+ 15
Sciences de la terre, Ecologie	41	66	+ 25
Chimie	28	64	+ 36
Informatique	36	58	+ 22
Recherche médicale	30	58	+ 28
Sciences humaines	18	50	+ 32
Sciences de l'ingénieur	27	50	+ 23
Sciences sociales	21	46	+ 25
Toutes disciplines	38%	62%	+ 24

Chaque voie d'accès joue un rôle important dans l'ouverture des publications

L'accès ouvert aux publications scientifiques peut se faire par plusieurs voies : la publication nativement en accès ouvert par l'éditeur sur une plateforme de publication ou le dépôt par l'auteur dans une archive ouverte. Ces deux voies ne sont pas exclusives, une publication pouvant être à la fois disponible sur une archive ouverte et sur la plateforme de publication de l'éditeur.

Ainsi, en 2021, 18 % des publications françaises parues en 2020 sont ouvertes seulement sur archive ouverte, 16 % seulement sur la plateforme de l'éditeur ou de diffusion et 28 % sont ouvertes via les deux voies. Cette simultanéité, qui tend à s'accroître au cours du temps, est un facteur de résilience puisqu'il permet à la fois d'offrir la qualité éditoriale et de garantir la pérennité de l'accès aux publications scientifiques françaises.

Tableau 2 : Taux d'accès ouvert des publications françaises parues dans l'année précédant la date d'observation (en %)

Discipline	Archive ouverte seulement	Archive ouverte ET éditeur / plateforme de diffusion	Éditeur / plateforme de diffusion seulement	Accès ouvert
Biologie fondamentale	18	44	12	74
Sciences physiques, Astronomie	36	28	9	73
Mathématiques	42	20	10	72
Sciences de la terre, Ecologie	21	31	14	66
Chimie	31	21	12	64
Informatique	33	15	10	58
Recherche médicale	9	32	17	58
Sciences humaines	6	10	34	50
Sciences de l'ingénieur	20	15	15	50
Sciences sociales	13	11	22	46
Toutes disciplines	18%	28%	16%	62%

Toutes les disciplines n'adoptent pas les mêmes vecteurs de publication en accès ouvert. Pour certaines, la pratique du dépôt en archive ouverte est historiquement ancrée. Ainsi, les mathématiques sont très adeptes des archives ouvertes et les sciences humaines confient plus volontiers leur ouverture aux éditeurs. Entre les deux existent de nombreuses situations, propres à l'organisation et à l'histoire des disciplines. Il y a aussi de plus en plus de disciplines qui adoptent les deux voies.

Le dépôt en archive ouverte constitue le moyen de contrebalancer la politique restrictive de certains éditeurs en matière d'accès ouvert et joue dans ce cas un rôle décisif dans l'accès de tous aux résultats de la recherche française. À l'inverse, lorsque l'éditeur publie nativement en accès ouvert, le dépôt en archive ouverte peut apparaître moins nécessaire aux auteurs. Ces différences disciplinaires sont couplées avec les politiques d'ouverture des éditeurs et plateformes de diffusion.

Tableau 3 : Modalité d'ouverture chez les éditeurs et plateforme les plus importants en volume de publications parues en 2020 (top 10) (en %)

Editeur / Plateforme	Accès ouvert sur archive ouverte seulement	Accès ouvert sur le site éditeur / plateforme	Accès ouvert (total)
Elsevier	13	41	54
Springer-Nature	18	34	52
Wiley	24	37	61
MDPI	0	100	100
OpenEdition	2	73	75
IEEE	41	7	48
Oxford University Press	24	46	70
Cairn.info	6	2	8
Informa	15	25	40
Frontiers	1	99	100

HAL est l'archive ouverte la plus utilisée

Les archives ouvertes sont des plateformes sur lesquelles sont déposées des publications scientifiques, qui peuvent être consultées par tous. Elles sont alimentées le plus souvent par un dépôt effectué par les auteurs, mais peuvent l'être dans certains cas par les éditeurs de revues eux-mêmes. Les archives ouvertes assument différentes

fonctions : elles rendent disponibles en accès ouvert des articles qui paraissent dans des revues sous abonnement, elles assurent la conservation pérenne de la littérature scientifique et facilitent l'identification de la production d'un laboratoire ou d'un établissement. HAL est l'archive ouverte la plus utilisée pour déposer les publications françaises, avec plus de 37 000 publications parues en 2020 qui y sont hébergées. PubMed Central (29 000 publications), arXiv (10 000 publications) et bioRxiv (4 300 publications) jouent aussi un rôle significatif. Sur le corpus des publications françaises parues en 2020 et déposées sur une archive ouverte, HAL en contient 48 %. Presque 40 000 publications parues en 2020 sont donc hébergées sur un autre archive ouverte sans être disponibles sur HAL. Sur la totalité des publications françaises de 2020, 22 % du total sont en accès ouvert sur HAL (15 % en 2019). Ainsi, la contribution de HAL est de plus en plus significative, mais HAL ne contient encore que moins d'un quart de la totalité de la production française.

Les résultats des essais cliniques sont encore peu partagés, notamment pour les promoteurs académiques

Le Baromètre de la Science Ouverte introduit le suivi de la transparence des résultats des essais cliniques et études observationnelles menés en France. La déclaration des essais cliniques et de leurs résultats dans des bases de données publiques contribue à une plus grande transparence de la recherche médicale. Elle permet une circulation rapide des résultats, y compris lorsque ceux-ci ont été infructueux et ne font pas l'objet d'une publication scientifique. Parmi les essais cliniques enregistrés et signalés comme terminés dans un registre public depuis 2010, des résultats ont été déclarés ou identifiés automatiquement pour 57 % d'entre eux (résumés des résultats, publications scientifiques associées, ...). Ce taux monte à 76 % pour les promoteurs industriels et n'est que de 31 % pour les promoteurs académiques.

Eric JEANGIRARD
MESRI-SIES

La méthodologie du Baromètre de la Science Ouverte a été revue en 2021. Elle est détaillée dans le document « A New Framework for the French Open Science Monitor », A. L'Hôte, E. Jeangirard, D. Torny and L. Bracco.

Cette refonte méthodologique et technique s'accompagne d'une révision des résultats annoncés dans les précédentes Notes Flash, entraînant des changements de quelques points sur certains indicateurs.

Sources : Le Baromètre de la Science Ouverte repose uniquement sur des sources ouvertes, en particulier Unpaywall, HAL, clinicaltrials.org et EU Clinical Trials Register (EUCTR). Unpaywall est une base de données mondiale, ouverte, recensant plus de 100 millions de DOI avec leur méta-données disponibles (titre, auteur, éditeur ...) et leur type d'accès. ClinicalTrials.org et EUCTR sont les deux principaux registres publics en ligne d'études cliniques.

Champ : Publications avec un DOI (Digital Object Identifier) et dont au moins un des auteurs a une affiliation française. Essais cliniques et études observationnelles menés (au moins en partie) en France, et enregistré dans clinicaltrials.org ou EUCTR. Les données du BSO sont disponibles en open data sur le portail Open Data du MESRI :

<https://data.enseignementsup-recherche.gouv.fr/explore/dataset/open-access-monitor-france/>

Pour plus d'informations, rendez-vous sur le site web du Baromètre de la Science Ouverte : <https://barometredelascienceouverte.esr.gouv.fr>