

## Baromètre français de la Science Ouverte

Selon le nouveau baromètre de la science ouverte, 41 % des 133 000 publications scientifiques françaises (suivant les adresses d'affiliation des auteurs) publiées en 2017 sont en accès ouvert. L'étendue et le type d'ouverture sont caractérisés par de fortes disparités entre disciplines.

### À RETENIR

ACCÈS OUVERT DES PUBLICATIONS PUBLIÉES EN 2017  
(observé en Septembre 2018)

41 %

des publications avec une affiliation française  
( soit 53 856 publications ) en accès ouvert

Source : Unpaywall et HAL  
traitement MESRI-SIES

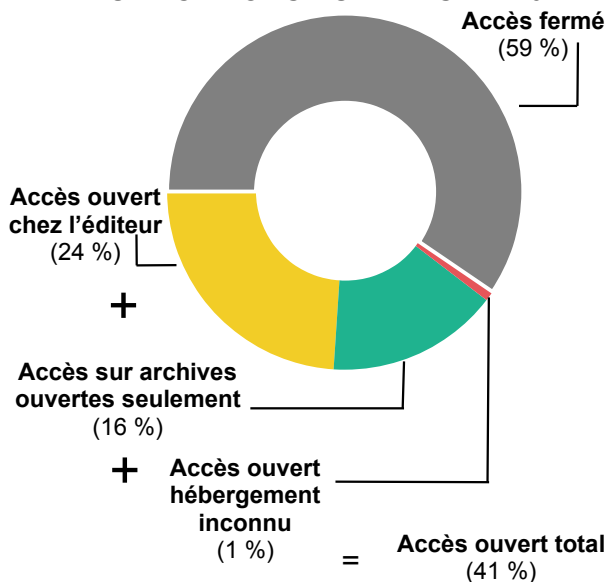
PUBLICATIONS DE L'ANNÉE 2017

132 970

publications avec un DOI et une affiliation française  
ont été identifiées à partir de données ouvertes

Source : Unpaywall et HAL  
traitement MESRI-SIES

ACCÈS EN 2018 POUR LES PUBLICATIONS PUBLIÉES EN 2017



### 1 | Mise en place du baromètre

En juillet 2018, le ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation a rendu public le Plan national pour la science ouverte. Ce plan vise à favoriser le développement d'un écosystème de la science renouvelé, caractérisé notamment par une circulation sans entraves des publications scientifiques et des données. La construction d'un **baromètre de la science ouverte** s'inscrit dans ce plan et constitue par ailleurs un des engagements du Plan d'Action National de la France au sein du **Partenariat pour un Gouvernement Ouvert** (PGO). Ce baromètre vise, à terme, à suivre l'étendue des ressources scientifiques accessibles librement : publications scientifiques, données, code. Pour sa première édition, le baromètre se concentre sur les publications en accès ouvert (mises à disposition librement sur l'internet public).

Le taux d'accès ouvert est mesuré comme le ratio du nombre de publications en accès ouvert rapporté au nombre total de publications sur un même périmètre. Les publications prises en compte sont celles qui disposent d'un « DOI », identifiant numérique d'objet, ce qui constitue une règle assez générale. Deux types d'accès ouvert peuvent être distingués, en fonction du type d'hébergement, chez l'éditeur ou sur une plateforme d'archives ouvertes (comme HAL par exemple). Le taux d'accès ouvert est estimé à partir des données figurant sur la plateforme **Unpaywall**, qui recense les publications au niveau mondial et leur caractère ouvert ou non. La méthodologie appliquée permet d'identifier les seules publications avec au moins une affiliation française et d'estimer leur taux d'ouverture, par année et par discipline.

Les résultats de cette note Flash se basent sur les données Unpaywall de Septembre 2018.

### 2 | De fortes disparités disciplinaires

132 970 publications scientifiques publiées en 2017 avec une affiliation française ont été identifiées dans le champ du baromètre. Parmi elles, 53 856, soit 41 %, étaient en accès ouvert en 2018. 31 876 étaient directement disponibles chez l'éditeur (soit 24 % du total).

D'une discipline à l'autre, non seulement la proportion d'accès ouvert varie largement mais aussi le type d'accès ouvert (hébergement chez l'éditeur ou sur archives ouvertes). Ces disparités sont le reflet de pratiques de publication très différentes entre les communautés scientifiques. Ainsi, en mathématiques, malgré un faible taux de publications hébergées chez les éditeurs, le taux d'accès ouvert global atteint 61 %, grâce à un recours massif aux archives ouvertes (48 %). Dans le domaine de la biologie fondamentale, le niveau global d'ouverture est de 49 %, porté principalement par l'ouverture chez l'éditeur. Les deux modes d'ouverture sont ainsi complémentaires pour s'adapter aux usages de chaque communauté.

## REPÈRES MÉTHODOLOGIQUES

**Sources** | Le Baromètre de la Science Ouverte repose uniquement sur des sources ouvertes, en particulier Unpaywall, HAL et la base de données bibliographiques Pascal et Francis.

**Champ** | Publications scientifiques avec un DOI (Digital Object Identifier) et dont au moins un des auteurs a une affiliation française.

**Unpaywall** | Base de données mondiale, ouverte, recensant plus de 100 millions de DOIs avec leur méta-données disponibles (titre, auteur, éditeur ...) et leur type d'accès (<https://unpaywall.org>)

### Étapes de la méthodologie

#### 1. Identification des publications

Les publications du champ du baromètre sont identifiées à l'aide d'un processus d'analyse automatique des pages web qui repère les affiliations françaises.

#### 2. Enrichissement des méta-données

Le champ disciplinaire des publications est inféré par un algorithme de machine learning supervisé entraîné sur la base bibliographique Pascal et Francis.

#### 3. Détection de l'accès ouvert

Les données Unpaywall (snapshot 2018) sont utilisées pour déterminer le type d'accès.

## Pour en savoir plus

- Le site web du baromètre de la science ouverte : <http://bso.enseignementsup-recherche.gouv.fr>

- La méthodologie complète est détaillée dans l'article suivant (en accès libre) Monitoring Open Access at a national level: French case study, Eric Jeangirard. ELPUB 2019 23rd edition of the International Conference on Electronic Publishing, Jun 2019, Marseille, France. <hal-02141819>

- Le code source du baromètre est en ligne : <https://github.com/MinistereSupRecherche/bso>

**Tableau 1 : Accès ouvert pour les publications publiées en 2017 (observé en Septembre 2018)**

Discipline	Accès ouvert chez l'éditeur	Accès ouvert sur archives uniquement	Accès ouvert (type inconnu)	Accès ouvert (global)
Mathématiques	13 %	48 %	0 %	61 %
Sciences physiques, Astronomie	24 %	30 %	1 %	55 %
Biologie fondamentale	37 %	12 %	0 %	49 %
Sciences de la terre, Ecologie	27 %	19 %	2 %	48 %
Informatique	10 %	33 %	0 %	43 %
Sciences de l'ingénieur	18 %	17 %	0 %	35 %
Sciences humaines	27 %	6 %	1 %	34 %
Chimie	16 %	14 %	1 %	31 %
Recherche médicale	23 %	7 %	1 %	31 %
Sciences sociales	21 %	8 %	1 %	30 %
<b>Toutes disciplines</b>	<b>24 %</b>	<b>16 %</b>	<b>1 %</b>	<b>41 %</b>

Source: Unpaywall et HAL, traitement MESRI-SIES

Inversement, en recherche médicale, qui représente environ un tiers de la production annuelle, seuls 31 % des publications sont en accès ouvert, et la plupart hébergées chez l'éditeur. De même, concernant les sciences humaines et les sciences sociales, le taux global d'ouverture est plus faible que la moyenne, avec un recours aux archives ouvertes relativement faible.

## 3 | Tendances de l'accès ouvert 2013-2017

Le niveau global d'accès ouvert des publications françaises augmente légèrement de 2013 à 2016, en évoluant de 40 % à 43 %, avec une dynamique similaire sur les deux types d'hébergement.

**Tableau 2 : Proportion d'accès ouvert pour les publications publiées entre 2013 et 2017 (observé en Septembre 2018)**

Année de publication	Accès ouvert chez l'éditeur	Accès ouvert sur archives uniquement	Accès ouvert (type inconnu)	Accès ouvert (global)
2013	25 %	14 %	1 %	40 %
2014	26 %	14 %	1 %	41 %
2015	28 %	15 %	0 %	43 %
2016	27 %	15 %	1 %	43 %
2017	24 %	16 %	1 %	41 %

Source: Unpaywall et HAL, traitement MESRI-SIES

La dernière année de publication analysée (2017) présente un léger repli (-2 pts), porté par la contraction observée pour les publications ouvertes hébergées chez l'éditeur (-3 pts par rapport à 2016). Plutôt qu'un réel changement de tendance, cette baisse constatée en 2018 pour les publications de 2017 est probablement liée au délai de passage en accès ouvert (potentiellement plusieurs mois s'il y a un embargo ou une barrière mobile).

**Eric JEANGIRARD**  
**MESRI-SIES**