



Open Science Pursuit

Fascicule animateur

BIBLIOMETRIE | BIBLIOMETRICS #01

[Vrai ou faux]

Mon H-index sera le même, qu'il soit calculé sur les bases Google Scholar, Web of Science ou Scopus.

My H-index will be the same whether it is calculated on Google Scholar, Web of Science or Scopus.

>> Faux | False

INDICE

Qu'est-ce qui différencie Google Scholar du Web of Science ?

What makes the difference between Google Scholar and the Web of Science?

Le H-index est calculé à partir des **citations repérées par chaque base**, donc les bases ayant **les couvertures les plus larges donnent des H-index plus élevés**.

Google Scholar agrège l'ensemble des publications, tandis que Scopus et le Web of Science font une sélection de revues suivant des critères de qualité, d'éthique, etc.

*The H-index is calculated from the **citations found by each database**, so the databases with the **largest coverage give higher H-indexes**.*

Google Scholar aggregates all publications, while Scopus and the Web of Science select journals according to criteria of quality, ethics, etc.

BIBLIOMETRIE | BIBLIOMETRICS #02

[Vrai ou faux]

Comparer un h-index de chercheurs venant de disciplines différentes est pertinent pour évaluer leur notoriété.

Comparing an h-index of researchers from different disciplines is relevant to assess their reputation.

>> Faux | False

INDICE

Les pratiques de citations sont-elles les mêmes en mathématiques, en chimie, en droit, en histoire ?

Are citation practices the same in mathematics, chemistry, law, history?

Les **pratiques de publications diffèrent beaucoup** entre les sciences humaines et sociales et les sciences techniques ou médicales, et même au sein des disciplines (ex : mathématiques/biologie), ce qui a un effet important sur les citations, donc les H-index. Pour éviter ce problème, des indicateurs dits pondérés ont été créés, comme le SNIP, Source Normalized Impact per Paper.

Publication practices differ greatly between the humanities and social sciences and the technical or medical sciences, and even within disciplines (e.g. mathematics/biology), which has an important effect on citations, and therefore H-indexes. To avoid this problem, weighted indicators have been created, such as the SNIP, Source Normalized Impact per Paper.

BIBLIOMETRIE | BIBLIOMETRICS #03

[Vrai ou faux]

Les indicateurs bibliométriques permettent d'évaluer une revue, un chercheur, un laboratoire, un financeur.

Bibliometric indicators enable the evaluation of a journal, a researcher, a laboratory or a funder.

>> Vrai | True

INDICE

Les bases de données les plus poussées permettent d'interroger de nombreux champs : citations reçues, pays des co-auteurs, participation à un brevet...

The most advanced databases make it possible to query numerous fields: citations received, country of co-authors, participation in a patent, etc.

Les indicateurs sont **variés et adaptés** à ce à quoi on les applique, ce qui peut être le travail d'un chercheur, un labo, une revue, un financeur...

Dans leur diversité, certains prennent en compte les auto-citations, d'autres sont pondérés en fonction des pratiques de la discipline, et ils sont calculés sur des corpus différents.

Lors de leur utilisation, **il est important de se pencher sur leurs méthodes de calcul et leur biais, et d'en utiliser plusieurs ensemble** pour avoir une vue pertinente.

*The indicators are **varied and adapted** to what they are applied to, which may be the work of a researcher, a lab, a journal, a funder, etc.*

In their diversity, some take into account self-citations, others are weighted according to the practices of the discipline, and they are calculated on different corpora.

When using them, it is important to consider their calculation methods and bias, and to use several of them together to get a relevant view.

BIBLIOMETRIE | BIBLIOMETRICS #04

[Question ouverte]

Citez un autre critère d'évaluation des chercheurs que le H-index.

Name one other criterion for evaluating researchers than the H-index.

Quelques exemples : score SIGAPS, index H-5, nombre de publications dans des revues cotées (Q1 ou disciplinaires), les récompenses obtenues pendant une carrière (médailles du CNRS, prix Nobel, médaille Fields, médailles et récompenses disciplinaires, distinction Highly Cited...)

>>

Some examples: SIGAPS score, H-5 index, number of publications in ranked journals (Q1 or disciplinary), awards obtained during a career (CNRS medals, Nobel Prize, Fields Medal, disciplinary medals and awards, Highly Cited distinction...)

INDICE

Certaines récompenses prestigieuses peuvent être utilisées en tant qu'indicateur.

Some prestigious awards can be used as indicators.

Il existe une très large variété d'indicateurs liés à la bibliométrie, mais leur bon usage nécessite d'utiliser plusieurs à la fois sur un même corpus pour compenser les biais des uns et des autres. Ils sont également à adapter en fonction du domaine et du volume de publications.

There is a very wide variety of bibliometric indicators, but to use them correctly, you need to use several at the same time for the same corpus to compensate for the biases of some and others. They also need to be adapted according to the field and the volume of publications.

BIBLIOMETRIE | BIBLIOMETRICS #05

[Question ouverte]

Citez une ressource (site internet, base de données) permettant de trouver des indicateurs d'évaluation bibliométrique des revues.

Name a resource (website, database) for finding bibliometric evaluation indicators for journals.

>>

Web of science (IF), Scopus (CiteScore), Scimago Journal Rank (SJR), Google Scholar

INDICE

L'une de ces bases le permet : CAIRN, ScienceDirect, Scopus.

One of these databases is one of them : CAIRN, ScienceDirect, Scopus.

Toutes les bases de données ne donnent pas accès à des indicateurs bibliométriques. Ils sont généralement **calculés sur des bases dédiées**, pour lesquelles les éditeurs sélectionnent seulement un certain nombre de revues. Les citations pour les articles issus de ces revues sont ensuite récoltées pour permettre les calculs. **La plupart des bases ont leurs propres indicateurs**, et continuent d'en développer de nouveaux.

Le saviez-vous ? Les calculs peuvent être similaires, mais les noms changent : l'Impact Factor du WoS devient le CiteScore chez Scopus.

*Not all databases provide access to bibliometric indicators. They are usually **calculated on dedicated databases**, for which publishers select only a certain number of journals. Citations for articles from these journals are then collected to enable the calculations. **Most databases have their own indicators**, and continue to develop new ones.*

Did you know? The calculations may be similar, but the names change: the Impact Factor of WoS becomes the CiteScore at Scopus.

BIBLIOMETRIE | BIBLIOMETRICS #06

[Vrai ou faux]

La **bibliométrie** peut vous aider à repérer de futurs laboratoires partenaires.

Bibliometrics can help you identify future partner laboratories.

>> Vrai | True

INDICE

Les affiliations font-elles parties des métadonnées que l'on pourrait interroger ?

Are affiliations part of the metadata that could be queried?

Même si les rencontres interpersonnelles sont la manière la plus courante de commencer des collaborations, **l'analyse des collaborations passées, au niveau des auteurs et des laboratoires, et l'analyse des citations entre publications permet de repérer la complémentarité entre des équipes.** Cela aide à trouver des collaborations potentielles, par exemple en vue de la réponse à des appels à projets.

*Even if interpersonal meetings are the most common way of starting collaborations, **the analysis of past collaborations, at the level of authors and laboratories, and the analysis of citations between publications can identify the complementarity between teams.** This helps to identify potential collaborations, for example in response to calls for projects.*

BIBLIOMETRIE | BIBLIOMETRICS #07

[QCM]

L'Université Paris-Saclay propose son baromètre de la Science Ouverte chaque année. En 2022, le taux des publications de l'université parues en 2020 et en accès ouvert se situe entre : 60 et 70% | 70% et 80% | 80% et 90%

The University of Paris-Saclay proposes its Open Science barometer every year. In 2022, the rate of the university's publications published in 2020 and in open access is between: 60 and 70% | 70% and 80% | 80% and 90%

>> 70% et 80% | 70% and 80%

INDICE

En 2020, pour les publications de 2018, le taux était de 65%.

In 2020, for the publications of 2018, the rate was 65%.

Le baromètre de 2022 estime à 74% le taux d'ouverture sur les publications de 2020.

Développé par le MESRI, le Baromètre français de la Science Ouverte permet d'avoir une vue sur l'ouverture des publications à l'échelle nationale et à l'échelle des établissements. A ce jour, il ne prend en compte que les publications avec DOI mais de nouveaux développements sont à l'étude. D'année en année, son périmètre s'étend et son utilisation se simplifie.

The 2022 Barometer estimates that 74% of the publications of 2022 are open.

Developed by the MESRI, the French Open Science Barometer provides an overview of the openness of publications at national and institutional level. To date, it only takes into account publications with DOIs, but new developments are being studied. Year on year, its scope is being extended and it is becoming easier to use.

BIBLIOMETRIE | BIBLIOMETRICS #08

[Vrai ou faux]

La DiBiSO peut vous fournir une analyse sur des tendances dans votre domaine de recherche (trends).

The DiBiSO can provide you with an analysis of trends in your research area.

>> Vrai | True

INDICE

Les analyses des trends se font principalement à partir des bases Web of Science et/ou Scopus.

Trends are analysed mainly using Web of Science and/or Scopus databases

La Direction des Bibliothèques, de l'Information et de la Science Ouverte peuvent réaliser une telle analyse, **grâce aux outils de bibliométrie** comme SciVal. En effet, ces outils sont capables de **signaler les domaines « en vogue » et émergents**, mais demandent un temps de prise en main pour être utilisés efficacement.

*The Department of Libraries, Information and Open Science can carry out such an analysis **using bibliometric tools such as SciVal**. These tools **are capable of identifying 'hot' and emerging fields**, but require time to learn how to use them effectively.*

[Vrai ou faux]

Nature et Science sont les revues scientifiques ayant le plus haut facteur d'impact.

Nature and Science are the scientific journals with the highest impact factor.

>> **Faux | False**

INDICE

Pour quelles raisons le facteur d'impact est-il critiqué ?

Why is the impact factor being criticised?

Bien que conçu pour mesurer la renommée des revues, le Facteur d'Impact répond à des règles qui rendent cette corrélation inexacte. Le FI ne prend pas en compte les disparités de comportement entre domaines disciplinaires. De plus, certains types de publications ne sont pas pris en compte dans le dénominateur du FI, alors que leurs citations comptent au numérateur, permettant de « tricher ». Ces spécificités donnent l'avantage aux revues spécialisées dans certains types d'articles (éditoriaux ou méta-analyses par exemple) et certaines disciplines fortement pourvoyeuses de citations (médecine, biochimie – ex : The Lancet) sur les revues interdisciplinaires traditionnelles comme Nature et Science.

Although designed to measure the reputation of journals, the Impact Factor follows rules that make this correlation inaccurate. The IF does not take into account disparities in behaviour between disciplinary fields. What's more, certain types of publication are not taken into account in the IF denominator, whereas their citations count in the numerator, making it possible to 'cheat'. These specific features give an advantage to journals specialising in certain types of article (editorials or meta-analyses, for example) and certain disciplines with a high citation rate (medicine, biochemistry – e.g. The Lancet) over traditional interdisciplinary journals such as Nature and Science.

BIBLIOMETRIE | BIBLIOMETRICS #10

[Vrai ou faux]

Le H-index prend en compte toute la carrière d'un chercheur.
The H-index takes into account the whole career of a researcher.

>> **Vrai | True**

INDICE

Lequel semble être le plus intéressant : un indice qui prend en compte toute la carrière d'un chercheur ou les 5 dernières années ?

Which seems to be more interesting: an index that takes into account a researcher's entire career, or just the last 5 years?

Le H-index prend en compte toutes les publications repérées pour un chercheur, **et c'est l'un de ses défauts** : il favorise systématiquement les chercheurs plus avancés dans leur carrière et n'a aucune pertinence pour les chercheurs jeunes et prometteurs.

*The H-index takes into account all the publications identified for a researcher, and **this is one of its shortcomings**: it systematically favours researchers who are more advanced in their careers, and has no relevance for young, promising researchers.*

BIBLIOMETRIE | BIBLIOMETRICS #11

[Question ouverte]

Citez un indicateur bibliométrique construit sans utiliser le critère des citations.

Name a bibliometric indicator constructed without using the citation criterion.

Ex : volume de publications, collaborations avec des auteurs étrangers ou avec des auteurs issus d'entreprises privées (corporate), Altmetrics (nombre de consultations, de téléchargements, de mentions sur Twitter etc.)

>>

Examples: volume of publications, collaborations with foreign authors, collaborations with authors from private companies (corporate), Altmetrics (number of hits, downloads, mentions on Twitter, etc.).

INDICE

Quelle donnée chiffrée peut-on par exemple demander à un doctorant ou un chercheur lorsqu'il postule ? (réponse : nombre de publications)

For example, what figures can you ask a doctoral student or researcher to provide when applying for a job? (answer : number of publications)

Toute statistique mesurable, d'ordre quantitatif, concernant les publications scientifiques, fait partie de la bibliométrie. Les usages peuvent varier, mais multiplier les angles d'approche est important pour toute évaluation.

Any measurable, quantitative statistic concerning scientific publications is part of bibliometrics. The uses may vary, but multiplying the angles of approach is important for any evaluation.

[QCM]

Parmi les trois recommandations suivantes, laquelle est absente de la Déclaration de San Francisco sur l'Evaluation de la Recherche (DORA, 2012) ?

- 1- Diversifier les indicateurs d'évaluation ;
 - 2- Mettre les publications scientifiques en Open Access ;
 - 3- Cesser d'utiliser le facteur d'impact pour évaluer les chercheurs
- Which of the following three recommendations is missing from the San Francisco Declaration on Research Assessment (DORA, 2012)?*
- 1- Diversify evaluation indicators;
 - 2- Make scientific publications Open Access;
 - 3- Stop using the impact factor to evaluate researchers

>>

2- Mettre les publications scientifiques en Open Access

2- Make scientific publications Open Access;

INDICE

La réponse est dans le nom de la déclaration !

The answer is in the name of the declaration!

Bien que souvent associé à DORA, l'appel à ouvrir la science n'y est pas présent en tant que tel. En revanche, **cette déclaration appelle à tirer parti des avantages de la publication électronique pour améliorer les conditions de la communication scientifique et rendre transparents les critères d'évaluation.**

Although often associated with DORA, the call to open up science is not present as such. Instead, **the declaration calls for the advantages of electronic publication to be exploited to improve the conditions for scientific communication and make evaluation criteria transparent.**

BIBLIOMETRIE | BIBLIOMETRICS #13

[QCM]

A quelle entreprise appartient la base de données Web of Science ?

Which company owns the Web of Science database?

Clarivate Analytics | Elsevier | Thomson Reuters

>> Clarivate Analytics

INDICE

En 2016, la société Thomson Reuters a annoncé qu'elle arrêterait son activité autour de la propriété intellectuelle et de la science.

In 2016, Thomson Reuters announced that it was shutting down its intellectual property and science business.

Après avoir appartenu à Thomson Reuters, **le Web of Science fait désormais partie du groupe Clarivate**. L'un de ses arguments par rapport à ses concurrents (dont Scopus, de l'éditeur Elsevier) est que le Web of Science est indépendant par rapports aux éditeurs, ce qui permettrait une meilleure neutralité.

*After belonging to Thomson Reuters, **the Web of Science is now part of the Clarivate group**. One of its arguments in relation to its competitors (including Scopus, from the publisher Elsevier) is that the Web of Science is independent of publishers, which makes it more neutral.*

BIBLIOMETRIE | BIBLIOMETRICS #14

Classez les trois publications suivantes de la plus susceptible à la moins susceptible d'être fortement citée au bout de 2 ans :

- 1) un chapitre de l'ouvrage *Snapshot-Based Methods and Algorithms*, avec pour titre "Basic ideas and tools for projection-based model reduction of parametric partial differential equations" (mathématiques);
- 2) une review publiée dans le *New England Journal of Medicine*, avec pour titre "Pathophysiology of inflammatory bowel diseases" ;
- 3) l'ouvrage "French Defence Policy Since the End of the Cold War" publié chez Taylor & Francis

Rank the following three publications from most to least likely to be highly cited after 2 years:

- 1) a chapter in *Snapshot-Based Methods and Algorithms*, entitled "Basic ideas and tools for projection-based model reduction of parametric partial differential equations" (mathematics);*
- 2) a review published in the *New England Journal of Medicine*, entitled "Pathophysiology of inflammatory bowel diseases";*
- 3) the book "French Defence Policy Since the End of the Cold War" published by Taylor & Francis*

>> 2, 1, 3

INDICE

Quelle est la publication la plus rapide à écrire et à lire ?

Which publication is the quickest to write and read?

Les reviews (type de publications peu innovantes mais synthétiques et tournée vers la pratique) **sont un type de publications très souvent citées à court terme car adaptées à des modèles de publication rapides, surtout en médecine où elles sont utilisées et lues par les praticiens.** Inversement, **les ouvrages sont des types de publications longs à écrire mais aussi longs à lire et à utiliser**, souvent tournés vers la recherche fondamentale, donc ils ne sont souvent cités qu'au bout de plusieurs années. Ce constat est particulièrement vrai en sciences humaines, où les expérimentations et les processus de publication durent souvent très longtemps.

Reviews (which are not very innovative but are synthetic and practice-oriented) **are a type of publication that is very often cited in the short term because they are adapted to rapid publication models, especially in medicine where they are used and read by practitioners.** Conversely, **books are a type of publication that take a long time to write, but also take a long time to read and use**, and are often geared towards fundamental research, so they are often only cited after several years. This is particularly true in the human sciences, where experiments and publication processes often last a very long time.

DONNEES | DATA #01

Parmi ces sites, lequel n'est pas un entrepôt de données ?

Which of these sites is not a data repository?

Zenodo | ArXiv | Recherche Data Gouv

>> **ArXiv**

INDICE

L'intrus porte bien son nom ! Il s'agit de la toute première archive ouverte, créée en 1991.

We would rather store data than archive (ArXiv) them...

Zenodo est un espace pluridisciplinaire créé par le CERN en 2013, qui permet d'entreposer des logiciels, des publications, mais aussi des jeux de données. Ceux-ci pourront bénéficier d'un DOI au moment du dépôt. Il est recommandé dans le cadre des projets Horizon Europe.

ArXiv est une archive ouverte initialement spécialisée en physique, ouverte aujourd'hui à plus de disciplines (maths, informatique...), qui reçoit avant tout des **pré-publications**.

Enfin, **Recherche Data Gouv est un entrepôt pluridisciplinaire développé par le MESRI** dédié à la publication aux jeux de données. L'Université Paris-Saclay y a créé son propre espace institutionnel.

***Zenodo is a multi-disciplinary space created by CERN** in 2013 for storing software, publications and datasets. These can be assigned a DOI at the time of submission. It is recommended for Horizon Europe projects.*

***ArXiv is an open repository** initially specialising in physics, but now open to a wider range of disciplines (maths, computing, etc.), which primarily receives pre-publications.*

*Finally, **Recherche Data Gouv is a multi-disciplinary repository developed by the MESRI** dedicated to the publication of datasets. Université Paris-Saclay has created its own institutional space there.*

DONNEES | DATA #02

[Vrai ou faux]

Ouvrir les données veut dire partager toutes ses données, sans exception.

Open data means sharing all your data without exception.

>> Faux | False

INDICE

Des informations médicales sur un patient unique en son genre pourraient-elles être partagées ?

Could medical information about a unique patient be shared?

On ne partage que ce qui est partageable !

Par exemple, les données qui ne peuvent pas être anonymisées ne peuvent pas être diffusées. La règle est toujours la même : « aussi ouvert que possible, aussi fermé que nécessaire » !

Only share what is shareable!

For example, data that cannot be anonymised cannot be shared. The rule is always the same: "as open as possible, as closed as necessary"!

Que signifie l'acronyme FAIR ? *What does FAIR stands for ?*

1. Facile à trouver Accessible Interopérable Réutilisable
[Findable Accessible Interoperable Reusable]
2. Facile à trouver Accepté Intelligible Réutilisable
[Findable Accepted Intelligible Reusable]
3. Fraîchement trouvé Accessible Intelligible Recyclable
[Fresh Accessible Intelligible Recyclable]

>>

1. Facile à trouver Accessible Interopérable Réutilisable
[Findable Accessible Interoperable Reusable]

INDICE

Les machines aussi doivent pouvoir comprendre les données, grâce à des formats et vocabulaires adaptés !

Machines also need to be able to understand data, thanks to appropriate formats and vocabularies!

Afin de maximiser la possibilité d'être **réutilisées**, les données doivent être :

- **Facile à trouver et accessibles**, par exemple en les déposant dans un **entrepôt dédié** et en prenant le temps de les **décrire** ;
- **Compréhensibles par différentes machines** et logiciels, donc avec des formats/langages **interopérables** !

To maximise the possibility of re-use, data must be :

- ***Easy to find and accessible***, for example ***by depositing it in a dedicated repository*** and taking the time to ***describe it*** ;
- ***Understandable by different machines*** and software, and therefore ***interoperable!***

DONNEES | DATA #04

Remettez dans l'ordre le cycle de vie des données :

Diffuser – Analyser – Traiter – Réutiliser – Collecter – Stocker

Put the data life cycle in order:

Disseminate - Analyse - Process - Reuse - Collect - Store

>>

Collecter – Traiter – Analyser – Stocker – Diffuser – Réutiliser

Collect - Process - Analyse - Store - Distribute - Reuse

INDICE

Je récupère mes données, je les nettoie, regarde ce qu'elles ont à me dire, les sauvegarde, avant de les partager pour qu'elles puissent être utiles à d'autres recherches.

I collect my data, clean it up, see what it has to say to me and save it, before sharing it so that it can be useful for other research.

Le cycle de vie des données de recherche **décrit le processus d'utilisation des données de leur création à la publication et à leur réutilisation ultérieure** :

- **Création** ou collecte des données ;
- **Traitement** des données ;
- **Analyse** des données ;
- **Conservation** des données ;
- **Accès** aux données ;
- **Réutilisation** des données.

The research data lifecycle **describes the process of using data from creation to publication and subsequent re-use**:

- **Creating** data;
- **Processing** data;
- **Analysing** data;
- **Preserving** data;
- Giving **access** to data / data discovery;
- **Reusing** data.

Source : INIST

Qu'est-ce qu'un Data Paper ?

What is a Data Paper ?

« Un Data paper est un article scientifique évalué par les pairs et citable. Il décrit un jeu de données, la méthode ayant permis de l'obtenir et le potentiel de réutilisation de ce jeu. Le Data paper ne décrit pas de résultats de recherche et ne contient ni discussion, ni conclusion. » (source : CIRAD)

>>

"A data paper is a peer-reviewed and citable scientific article. It describes a dataset, the method by which it was obtained and the potential for reuse of the dataset. The data paper does not describe research results and does not contain any discussion or conclusions. (source: CIRAD)

INDICE

Il vient compléter le jeu de données lui-même.
It complements the dataset itself.

Les Data Paper permettent de **mettre en valeur le travail réalisé autour des données**. Ce sont des articles de **qualité**, qui vont permettre la compréhension et la réutilisation des données.

Data Papers are a way of showcasing the work done on data. They are high-quality articles that will help people to understand and re-use the data.

A quel moment doit-on travailler sur un DMP ?

En début de projet / A mi-projet / En fin de projet / Les 3

When should you work on a DMP?

At the beginning of the project | Mid-project | At the end of the project | All 3

>>

Les 3.

All 3.

INDICE

A quel(s) moment(s) est-il préférable de se soucier du stockage des données ? A quel moment le document est-il demandé par les financeurs ?

When is it best to worry about data storage? At what point is the document requested by funders?

La rédaction d'un DMP est de plus en plus demandée par les financeurs. **Ceux-ci peuvent demander un premier rapport au début, à mi-projet, et/ou à la fin.** Dans tous les cas, **il est vivement conseillé de se pencher dessus dès le début** : en effet, ces documents vous permettent de réfléchir à de nombreux points pour **éviter les mauvaises surprises** au moment du rendu final, et également **d'établir des bonnes pratiques**, par exemple sur la sauvegarde au long court et à long terme.

*Funding bodies are increasingly asking for a DMP to be drawn up. **They may ask for a first draft at the beginning, mid-project and/or at the end.** In any case, **it is highly advisable to work on it from the outset**: these documents allow you to think about a number of points to **avoid unpleasant surprises** at the time of final delivery, and also to **establish good practice**, for example on long- and short-term back-up.*

Qu'est-ce qui n'est pas une donnée ?

Des entretiens individuels, retranscrits par écrit | Des coordonnées géographiques concernant des monuments publiques | Des rapports détaillant une population étudiée

What is not a data ?

Individual interviews, transcribed in writing | Geographical coordinates of public monuments | Reports detailing a study population

>>

Des rapports détaillant une population étudiée

Reports detailing a study population

INDICE

Lequel des trois a déjà subi une analyse ?

Which of the three has already undergone an analysis?

La seule proposition qui n'est plus une donnée ici est le rapport : **c'est une production à part entière, écrite à partir de données** qui elles pourraient être des **données chiffrées comme l'âge de la population**, le nombre d'habitant, ou des données textuelles comme les noms.

*The only proposition that is no longer data here is the report: **it is a production in its own right, written from data** that could themselves **be numerical data such as the age of the population**, the number of inhabitants, or textual data such as names.*

DONNEES | DATA #08

[Vrai ou faux]

Attribuer un DOI à un jeu de données est obligatoire.

Assigning a DOI to a dataset is mandatory.

>> **Faux | False**

INDICE

Les identifiants pérennes comme le DOI sont là pour aider à identifier et rendre plus visible l'objet auquel ils sont attribués.

Persistent identifiers such as the DOI are there to help identify and make more visible the object to which they are assigned.

Il n'est pas obligatoire d'attribuer un DOI à un jeu de données, **néanmoins cela est fortement recommandé**. En effet, un DOI **assure la pérennité des données** puisqu'il ne changera jamais même si la ressource est renommée ou déplacée. Il permet une **meilleure identification de votre jeu**, et donc une meilleure visibilité, ainsi que de faciliter sa citation.

*It is not compulsory to assign a DOI to a dataset, but it is **strongly recommended**. A DOI **ensures that the data will last**, as it will never change even if the resource is renamed or moved. It also **makes it easier to identify your dataset**, and therefore more visible, as well as making it easier to cite.*

Comment lier un jeu de données à une publication ?

How to link a dataset to a publication ?

>>

En indiquant le PID du jeu de données dans la publication, et le PID de la publication dans le jeu de données.

By indicating the PID of the dataset in the publication, and the PID of the publication in the dataset.

INDICE

Il y a peut-être une piste à creuser avec les DOI...

There may be something to be explored with DOIs...

Un identifiant pérenne (le plus connu étant le DOI) permet la citation d'un élément de façon simple et précise. **Concrètement, au moment de la publication, vous pouvez déposer vos données dans un entrepôt qui vous fournira l'identifiant à mentionner dans la publication.** Une fois votre publication parue, vous pourrez **ajouter l'identifiant de l'article à la description des données.**

*A persistent identifier (the best known being the DOI) enables an item to be cited simply and accurately. **In practical terms, when you publish, you can deposit your data in a repository, which will provide you with the identifier to be mentioned in the publication.** Once your publication has appeared, you can **add the article identifier to the data description.***

DONNEES | DATA #10

[Vrai ou faux]

Des bibliothécaires de l'Université Paris-Saclay peuvent vous aider dans l'élaboration d'un Plan de Gestion de Données.

Librarians from the University of Paris-Saclay can help you draw up a Data Management Plan.

>> **Vrai | True**

INDICE

L'ouverture des données, comme l'ouverture des publications, fait partie du grand ensemble de la Science Ouverte.

The opening up of data, like the opening up of publications, is part of the wider Open Science movement.

L'Atelier de la donnée DatASaclay est là pour vous aider ! Il est constitué de bibliothécaires, documentalistes, professionnels de l'Information Scientifique et Techniques, personnels d'appui à la recherche, qui **proposent de relire les PGD de tous les chercheurs et doctorants** de l'UPSaclay. Ils assurent également régulièrement des **formations** auprès des doctorants et des **interventions** à ce sujet dans les laboratoires !

*The DatASaclay Data Workshop is here to help! It is made up of librarians, documentalists, Scientific and Technical Information professionals and research support staff, who **offer to review the DMPs of all the researchers and doctoral students at UPSaclay.** They also provide regular **training** for doctoral students and give **talks** on this subject in the laboratories.*

[Question ouverte]

Citez 3 types de documents que vous pouvez déposer dans HAL.

Name 3 types of documents you can submit in HAL.

Article, conférence, poster, ouvrage, chapitre d'ouvrage, article de blog scientifique, notice d'encyclopédie ou de dictionnaire, traduction, brevet, pré-publication, rapport, thèse, HDR, cours, image, vidéo, son, carte, logiciel...

>>

Journal article, conference, poster, book, book chapter, scientific blog article, encyclopaedia or dictionary entry, translation, patent, pre-publication, report, thesis, HDR, course, image, video, sound, map, software...

INDICE

HAL accepte une vingtaine de documents, la seule condition est qu'ils soient utiles dans le cadre de la recherche.

HAL accepts around twenty documents, the only condition being that they are useful in the context of research.

HAL accepte de nombreux types de documents, **à partir du moment où ils participent à l'activité de recherche**. Outre donc les traditionnelles publications et actes de congrès, on peut y retrouver des articles de blogs scientifiques, des brevets, des rapports, des cours, mais aussi des médias comme des images ou du son.

*HAL accepts many types of document, **as long as they contribute to research activity**. In addition to traditional publications and conference proceedings, this includes articles from scientific blogs, patents, reports and courses, as well as media such as images and sound.*

[Vrai ou faux]

Open Access veut dire payer des APC (Article Processing Charges) pour publier un article.

Open Access means paying APC to publish your article.

>> **Faux | False**

INDICE

Il existe plusieurs façons de publier en Open Access : la voie dorée, la voie verte, la voie diamant en sont les principales.

There are several ways of publishing in Open Access: the golden, the green and the diamond paths are the main ones.

Le paiement d'Article Processing Charges (APC) par l'auteur ou l'institution correspond au modèle de la voie dorée. Il s'agit d'un modèle très connu car adopté par la plupart des gros éditeurs pour faire la bascule en Open Access. **Cependant, il existe d'autres modèles, où la publication peut être ouverte sans pour autant que l'auteur ou le lecteur aient à payer.** Parmi ces modèles, on peut retrouver la voie verte avec le dépôt des publications sur des archives ouvertes (HAL, ArXiv...) et la voie diamant (revue avec un autre mode de financement, par exemple : fonds par des institutions, souscriptions, mécénat...).

*The payment of Article Processing Charges (APCs) by the author or institution corresponds to the golden road model. This is a very well known model as it has been adopted by most of the major publishers to make the switch to Open Access. **However, there are other roads, where the publication can be open without the author or reader having to pay.** Among these models, we can find the green road with the deposit of publications on open archives (HAL, ArXiv...) and the diamond road (journal with another mode of financing, for example: funds by institutions, subscriptions, patronage...).*

[Question ouverte]

Citez 3 caractéristiques d'un éditeur prédateur.

Cite 3 characteristics of a predatory publisher.

>>

Privilégie ses intérêts personnels (profit > science), donne des informations fausses ou trompeuses, a des pratiques éditoriales inadéquates, manque de transparence, sollicitations agressives.

INDICE

A l'inverse, quels sont les éléments qui caractérisent une revue de bonne qualité ?

Conversely, what characterises a good quality journal?

Parmi les éléments pour les repérer, on peut citer :

- **Communication agressive**, par mail ;
- **Pas de relecture** par les pairs, publié tel quel avec ses éventuelles coquilles (anecdote : des canulars ont même été publiés) ;
- Taux d'acceptation élevés, délai (très) rapide ;
- L'article peut au final **ne pas être publié / profusion de numéros, volumes...**
- Si comité éditorial affiché : **membres fictifs** (ou réelles, à leur insu), adresses non professionnelles ou coordonnées absentes, fausses affiliations ;
- Sites web qui peuvent être de faible qualité : orthographe, grammaire, images déformées, liens brisés ; publicité ; informations « à venir » ;
- Tentent **d'imiter l'apparence de revues légitimes** : noms proches, sites web, appropriation de l'ISSN ;
- Avancent de **fausses métriques** (indicateurs bibliométriques inventés)

Elements to spot them include:

- **Aggressive communication, by e-mail ;**
- **No peer review**, published as is with any typos (anecdote: hoaxes were even published);
- High acceptance rates, (very) fast turnaround time
- The article **may not be published in the end / profusion of issues, volumes...**
- If editorial committee displayed: **fictitious members** (or real ones, without their knowledge), non-professional addresses or missing contact details, false affiliations
- Websites that may be of poor quality: spelling, grammar, distorted images, broken links; advertising; "forthcoming" information
- **Attempt to mimic the appearance of legitimate journals:** similar names, websites, ISSN appropriation
- Put forward **false metrics** (invented bibliometric indicators)

[Question ouverte]

Où peut-on trouver une liste de revues en Open Access ?

Where can you find a list of Open Access journals ?

>> **DOAJ, Scopus, WoS...**

INDICE

Il peut s'agit de BDD intégrant ce critère en filtre.

These may be databases incorporating this criterion as a filter.

Avec la montée de l'Open Access, de plus en plus de bases de données en ont fait un critère de filtre, et donc indiqué sur les fiches revues cette modalité. C'est le cas des bases Scopus et Web of Science.

Il existe aussi **des moteurs de recherche spécialisés dans les revues en Open Access**. Le plus connu est le **DOAJ** – Directory of Open Access Journal – qui en recense bientôt près de 20 000, donc plus de 13 000 sans APC. Il couvre toutes les disciplines et ses possibilités de filtres en font le moteur idéal pour trouver où publier en OA.

With the rise of Open Access, more and more databases have made this a filter criterion, and therefore indicated this modality on their journal records. This is the case with the Scopus and Web of Science databases.

*There are also search engines specialising in Open Access journals. The best known is the **DOAJ** - Directory of Open Access Journals - which will soon list nearly 20,000 Open Access journals, i.e. more than 13,000 without APC. It covers all disciplines and its filtering options make it the ideal engine for finding where to publish in OA.*

[QCM]

Quels sont les options d'une licence Creative Commons ?

1. BY, NC, SA, ND : paternité, utilisation non-commerciale, partage à l'identique, non-modifiable
2. BY, ZM, ZS : paternité, zéro modification, zéro succession
3. DC, SA, NC : déposable sous condition, partage à l'identique, utilisation non-commerciale

What are the options of Creative Commons licences ?

1. *BY, NC, SA, ND: attribution, non-commercial use, share alike, no derivative*
2. *BY, ZM, ZS: attribution, zero modification, zero re-share*
3. *DC, SA, NC: conditionally depositable, share alike, non-commercial use*

1. BY, NC, SA, ND : paternité, utilisation non-commerciale, partage à l'identique, non-modifiable

>>

1. BY, NC, SA, ND: attribution, non-commercial use, share alike, no derivative

INDICE

Attribuer une licence à notre travail a-t-il un intérêt si elle n'oblige pas a minima d'être cité comme auteur ?

Is there any point in granting a licence to our work if it does not at least require us to be cited as the author?

[détails page suivante]

Les Creative Commons sont gérées par une association à but non lucratif au nom éponyme et reconnues dans de nombreux pays. **Les 4 options permettent la création de 6 licences. La seule option obligatoire est la paternité (BY),** et plus on ajoute d'options, plus la licence devient restrictive. **Les éditeurs qui ont adopté les CC conseillent généralement de mettre uniquement l'option « BY »** car les autres options peuvent venir gêner la réutilisation de notre travail, et donc sa citation. Attention : si on ajoute une option, par exemple le non-modifiable (ND, no-derivative), **cela ne signifie pas qu'on en interdit la modification,** seulement qu'il faudra demander à l'auteur l'autorisation.

*The Creative Commons are managed by a non-profit association with the eponymous name, and are recognised in many countries. **The 4 options allow the creation of 6 licences. The only compulsory option is authorship (BY),** and the more options you add, the more restrictive the licence becomes. **Publishers who have adopted CC generally advise you to add only the "BY" option,** as the other options may make it more difficult to re-use your work and quote from it. Please note: adding an option such as "no modification" (ND, no-derivative) **does not mean that modification is prohibited,** only that the author's permission must be sought.*

[Question ouverte]

Citez 3 avantages à publier vos travaux en Open Access.

Name 3 advantages of publishing in Open Access.

>>

Augmenter l'impact scientifique (OA = davantage cité) ; gagner en visibilité (indexation moteurs de recherche), saisie unique pour les rapports (HAL => HCERES, CRAC...), bénéficié de services à valeur ajoutée (pages web, CV en ligne, statistiques...), protection juridique aux travaux (horodatage, embargo), garantie de la pérennité des fichiers (CINES), garantir un accès libre et gratuit à la recherche

INDICE

Quels sont les problèmes rencontrés lorsqu'une publication est derrière un paywall pour lequel nous n'avons pas accès ?

What are the problems encountered when a publication is behind a paywall to which we have no access?

Plusieurs études ont démontré **qu'un article en Open Access était plus cité**. En effet, nombre d'établissements ne sont pas en mesure de payer tous les abonnements nécessaires pour leurs chercheurs. De ce fait, un grand nombre d'articles restent inaccessibles.

La voie verte particulièrement, représentée par les archives ouvertes, a **vu l'éclosion de services complémentaires autour d'elles** : pour HAL par exemple, on peut éditer des listes de travaux par chercheur, laboratoire ou établissement, car il s'agit d'une des rares archives où tous les types de travaux peuvent être signalés. Pour ArXiv, on peut présenter son travail en pre-print, s'assurer la paternité des idées présentées dedans et récolter quelques premiers retours. **Etablies par les universités ou des organismes de recherche publics**, il y a un véritable souci de conservation

des écrits derrière, et donc **l'assurance d'un archivage à long-terme**, contrairement à des organismes privés qui pourraient ne plus être en mesure d'en assurer le stockage un jour.

*Several studies have shown that **an Open Access article is more cited**. Many institutions are unable to pay for all the subscriptions their researchers need. As a result, a large number of articles remain inaccessible.*

***The green route in particular**, represented by open archives, **has seen the emergence of complementary services around them**: for HAL, for example, it is possible to publish lists of works by researcher, laboratory or institution, as it is one of the rare archives where all types of work can be reported. For ArXiv, you can present your work as a pre-print, ensure the authorship of the ideas presented in it and gather some initial feedback. **Established by universities or public research bodies**, there is a genuine concern for the conservation of written material, and therefore **the assurance of long-term archiving**, unlike private bodies which may one day no longer be in a position to store it.*

[Vrai ou faux]

Quand je suis financé par un projet européen ou ANR, je dois publier en Open Access.

When a project is funded by a European or ANR funding, I have to publish in Open Access.

>> Vrai | True

INDICE

On parle d'argent public !

We're talking about public money!

Depuis Horizon 2020 pour l'Union Européenne et depuis 2022 pour l'ANR, **les publications doivent impérativement être en accès ouvert**. Il y a différentes possibilités suivant les cas, avec la voie verte, la voie dorée, la voie diamant. Les revues hybrides sont vivement déconseillées, voire parfois non acceptées.

*Since Horizon 2020 for the European Union and since 2022 for the ANR, **publications must be open access**. There are different possibilities depending on the case, with the green path, the golden path and the diamond path. Hybrid journals are strongly discouraged, and sometimes even not accepted.*

! A ne pas poser si la question #01 a été posée !

[Vrai ou faux]

Puis-je déposer un chapitre d'ouvrage dans HAL ?

Can I deposit a book chapter in HAL ?

>> Vrai | True

INDICE

HAL a pour objectif de recueillir tous les documents utiles dans le cadre de la recherche.

HAL aims to collect all documents useful in the context of research.

HAL accepte une trentaine de types de documents, comme : des articles, actes de conférence, articles de blogs scientifiques, notices de brevets, HDR, logiciels (en lien avec Heritage), posters, rapports, médias... **et bien sûr, des chapitres d'ouvrage** ! L'obligation est **qu'ils doivent avoir un lien avec l'activité de recherche**. On y trouve donc aussi des cours et des articles vulgarisés.

HAL accepts around thirty types of document, including: articles, conference proceedings, scientific blog articles, patent notices, HDRs, software (linked to Heritage), posters, reports, media... and, of course, book chapters! The only requirement is that they must be related to research. So you'll also find courses and popular articles.

[Vrai ou faux]

Le modèle APC est-il un modèle économiquement viable et rentable pour les organismes de recherche ?

Is the APC model an economically viable and profitable model for research organisations?

>> **Faux | False**

INDICE

La fourchette d'un APC se situe entre 800 et 8000€.

The price range for an APC is between €800 and €8000.

La moyenne d'un APC, environ 1500€, rend ce modèle intenable économiquement pour les institutions. Des revues comme Nature demandent 5000€ par article. Sur OpenAPC, à qui les institutions peuvent faire remonter les montants dépensés, en 2021 uniquement, l'INSERM a dépensé 1 600 000€, l'INRAE 930 000€, le CEA 220 000€, CentraleSupélec 23 000€. L'Université Paris-Saclay travaille à recenser elle aussi ces dépenses spécifiques aux APC, c'est pourquoi les montants ne sont pas encore disponibles publiquement.

The average APC is around €1,500, making this model economically untenable for institutions. Journals such as Nature charge €5,000 per article. On OpenAPC, to which institutions can report the amounts spent, in 2021 alone INSERM spent €1,600,000, INRAE €930,000, CEA €220,000 and CentraleSupélec €23,000. The University of Paris-Saclay is also working to identify these specific APC expenses, which is why the amounts are not yet publicly available.

[Question ouverte]

A quoi sert un identifiant ORCID ?

What is the use of an ORCID id ?

Visibilité, identité professionnelle, suivi tout au long de carrière, rôle pivot des identifiants, valorise l'ensemble des travaux, met en valeur le rôle de reviewer

>>

Visibility, professional identity, career-long follow-up, pivotal role of identifiers, enhances the value of all work, highlights the role of reviewer

INDICE

Quels problèmes se posent si on est identifié avec seulement son nom ?

What are the problems if you are only identified by name?

ORCID est une organisation mondiale à but non-lucratif qui a développé l'identifiant pérenne du même nom. De plus en plus apprécié des éditeurs, **il permet de vous identifier efficacement grâce à un numéro à 16 caractères** (permet de pallier aux homonymes, changements de noms, transcription de noms étrangers...) et **de regrouper vos publications**. Il est devenu également **central dans l'alignement des différents identifiants** qui peuvent exister.

ORCID is a global not-for-profit organisation that has developed a permanent identifier of the same name. Increasingly popular with publishers, it allows you to be identified effectively using a 16-character number (to compensate for homonyms, name changes, translation of foreign names, etc.) and to gather all your publications. It has also become central to the alignment of the various identifiers that may exist.

[Question ouverte]

Quels sont les avantages pour un Chief Editor à basculer le modèle de sa revue en modèle diamant ?

What are the advantages for a chief editor to switch the model of his journal to a diamond model?

Proposer un modèle économique plus vertueux, se détacher du modèle oligopole, accélérer l'accès libre sans contrainte pour les auteurs et lecteurs.

>>

Propose a more virtuous economic model, move away from the oligopoly model and accelerate free access without constraints for authors and readers.

INDICE

Pour quelles raisons le modèle traditionnel d'édition scientifique a été décrié ?

Why has the traditional model of scientific publishing been criticised?

En passant à un modèle diamant, on offre tout d'abord **un accès libre aux articles pour tous les lecteurs**, sans paywall, et **sans entraver la publication** côté auteur en demandant des APC. On vient également mettre sa pierre à un « rééquilibrage » du modèle de l'édition scientifique, puisque **le modèle traditionnel n'était plus soutenable financièrement** pour les organisations.

*By moving to a diamond model, we are first of all **offering free access to articles for all readers**, without a paywall, and without hindering publication by requiring APCs. We are also contributing to a 'rebalancing' of the scientific publishing model, **since the traditional model was no longer financially sustainable** for organisations.*

Quelles sont les 3 voies principales de l'Open Access ?

What are the 3 main routes to Open Access?

Voies verte (archives ouvertes), dorée (éditeur auteur-payeur)
et diamant (modèles où ni l'auteur ni le lecteur ne paient)

>>

*Green (open archives), golden (author-pays publisher) and
diamond (neither the author nor the reader pays) paths*

INDICE

Chacune a une couleur, la plus connue est la dorée.

Each has its own colour, the best known is the gold road.

Quand on parle d'Open Access, **on pense souvent au modèle auteur-payeur**, où le chercheur s'acquitte d'une certaine somme (les APC – Article Processing Charges) pour mettre l'article en libre accès. C'est un modèle répandu chez les grands éditeurs, **mais il en existe d'autres**. Par exemple, le « **Subscribe to Open** » : l'accès aux revues est ouvert si un certain montant de participation a été atteint. On peut également parler des **plateformes Publons ou Peer Community In**. Enfin, **la voie verte correspond aux archives ouvertes**, adossées à des institutions publiques.

When we talk about Open Access, **we often think of the author-pays model**, where the researcher pays a certain sum (the APC - Article Processing Charges) to make the article freely accessible. This model is widespread among the major publishers, **but there are others**. For example, "Open to Read": access to journals is open if a certain amount has been paid. There are also the **Publons and Peer Community In platforms**. Finally, **the green route is open archives**, backed by public institutions.

[Vrai ou faux]

Le CNRS encourage-t'il les auteurs à ne pas céder leurs droits à l'éditeur ?

Does the CNRS encourage authors not to transfer their rights to the publisher?

>> **Vrai | True**

INDICE

Rappel : HAL a été créé par le CNRS.

Reminder: HAL was created by the CNRS.

Depuis 2022, le CNRS demande à ses chercheurs d'appliquer la stratégie de non-cession des droits d'auteur. Le chercheur doit apposer une licence Creative Commons sur toutes les versions de son article jusqu'à la version acceptée pour publication. L'éditeur doit être mis au courant dès soumission de l'article. Cela permet à l'auteur de garder la main sur son article, à l'exception de la version qui sera publiée par l'éditeur. Selon le choix de l'auteur, les différentes versions peuvent alors être déposées en archives ouvertes, jusqu'au MAA inclus (manuscrit auteur accepté).

Since 2022, the CNRS has been asking its researchers to apply a non-licensing strategy. Researchers must affix a Creative Commons licence to all versions of their article up to the version accepted for publication. The publisher must be informed as soon as the article is submitted. This allows the author to retain control of the article, with the exception of the version that will be published by the publisher. Depending on the author's choice, the different versions can then be deposited in open repositories, including the MAA (accepted author manuscript).

Direction des Bibliothèques, de l'Information et de la Science Ouverte -
Université Paris-Saclay
Contact : amandine.lustrement@centralesupelec.fr

Version du 11/09/2024

