

# Science Ouverte

---

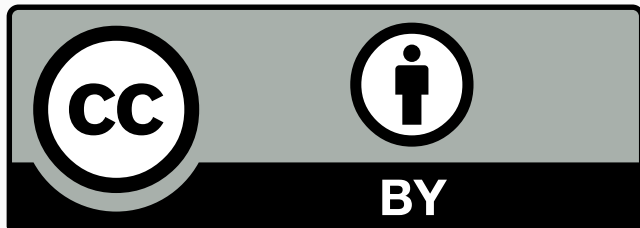
Nicolas Fressengeas

13 janvier 2023

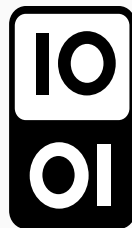
Université de Lorraine

# Licences (compatibles) d'utilisation de ce document

Creative Commons CC-BY



Etalab



Ouverture des publications

Gestion des données de la recherche

Le pilier logiciel

Évolution nécessaire de l'évaluation de la recherche



# Ouverture des publications

---

Enjeux

Effondrement, au XXI<sup>e</sup> siècle, du coût de la communication scientifique, mais...

<i>operating profit</i>	<i>company</i>	<i>industry</i>
7%	<u>Woolworths</u>	supermarkets, pokies
12%	BMW	automobiles
23%	Rio Tinto	mining
35%	Apple	premium computing
<b>34%</b>	<b>Springer</b>	<b>scholarly publishing</b>
<b>36%</b>	<b>Elsevier</b>	<b>scholarly publishing</b>
<b>40%</b>	<b>Wiley</b>	<b>scholarly publishing</b>

Source : *The Conversation, Free for all : ARC-funded research now open to the public*

## La Covid19, un cas d'école

- Pression des chercheurs
- Ouverture temporaire et partielle

## Le dépôt en archive ouverte

- Circulation rapide de l'information
- Accélération de la recherche
- Preuve d'antériorité

## South China Morning Post, 2020

China

### Will the coronavirus kill off the 'dinosaur' world of academic publishing?

- The deadly pandemic has brought back a long-running debate about companies profiting from the publication of research often freely supplied by the author
- As the biggest names in the business respond to academics' demands to bring down paywalls, new platforms are getting fresh studies out to the public



Linda Lew

Published: 11:00pm, 16 Mar, 2020 -

[www.you.com.trust.scmp](http://www.you.com.trust.scmp)



## Transparence des mécanismes de la recherche

3.

### Share trusted information

Visit the WHO website or chat to your doctor or nurse to find answers to common questions

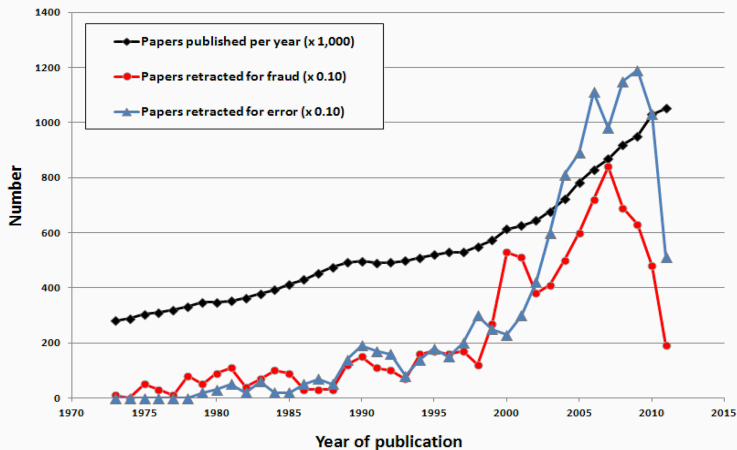
How do we know the vaccines are safe?

They've been thoroughly tested and reviewed. If you're interested, I know where we can find more information.



## Thousands of scientists publish a paper every five days

### Retractions as a function of total publications



Source : Steen RG, Casadevall A, Fang FC (2013) Why Has the Number of Scientific Retractions Increased?. PLOS ONE 8(7) : e68397. <https://doi.org/10.1371/journal.pone.0068397>



## Dissémination des articles

- Visibilité
- Accessibilité généralisée
  - PME
  - Pays du sud
  - Nous-même
- Facilitation des collaborations
  - Open Access Citation Advantage

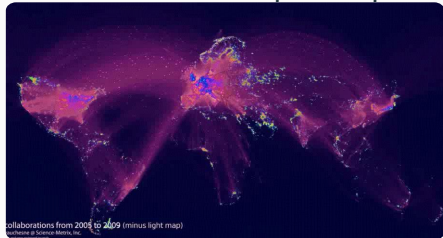
## Une alternative pérenne à SciHub



L'usage quasi-exclusif de l'anglais induit un biais international

- en défaveur de certaines disciplines
- en faveur d'une partie du monde

Collaborations vues par Scopus



Collaborations vues par Redalyc



Source : Arianna BECERRIL, deuxièmes Journées Nationales de la Science Ouverte

# Ouverture des publications

---

Politiques d'ouverture

## 100% en accès ouvert en 2030

- Généraliser l'obligation de publication en accès ouvert
  - articles
  - ouvrages
- Soutenir le modèle *diamant*
  - gratuit pour l'auteur
  - gratuit pour le lecteur
- Favoriser le multilinguisme
  - En s'appuyant sur la traduction automatique



cOAlition S

ANR



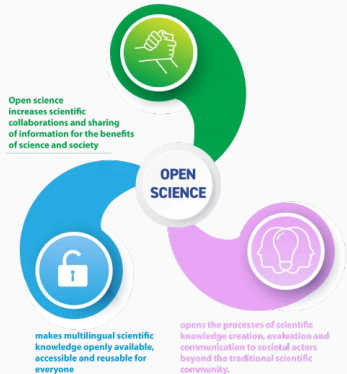
Commission européenne



Obligation dès maintenant d'un accès ouvert immédiat : *Stratégie de non-cession des droits*

The Recommendation calls on Member States to set up regional and international funding mechanisms for open science and to ensure that all publicly funded research respects the principles and core values of open science.

## Recommandation de l'UNESCO



# Ouverture des publications

---

Comment ouvrir ?

## Garantie par la Loi pour une République Numérique

- Obligation depuis 2018
- Embargo 6 ou 12 mois

## Stratégie de rétentation des droits

- ANR, cOAlition S & Europe
- Pour un dépôt sans embargo
- Sous licence Creative Commons

## Le dépôt en archive ouverte

**HAL**  
science ouverte

La connaissance libre et partagée

Accueil | Dépôt | Consultation- | Recherche | Documentation

**Coronavirus (COVID-19)**

Consulter et télécharger toutes les publications disponibles dans HAL.

➤ Accéder aux publications

**RECHERCHE**

Search terms

**NOMBRE DE DOCUMENTS SCIENTIFIQUES**

1 030 526

**NOMBRE DE RÉFÉRENCES**

2 932 581

**ACTUALITÉS**

**APPEL DE PARIS SUR L'ÉVALUATION DE LA RECHERCHE** (2020/2021)  
Ce texte a été préparé par le Comité pour la science ouverte et sera présenté aux Journées européennes de la science ouverte (Paris Open Science European Conference – OSEC 2022) qui se tiendront à Paris les 4 et 5 février 2022, organisées par la Présidence française du Conseil de l'Union européenne, suite à la publication de la recommandation de l'UNESCO sur la Science ouverte et à la publication par la Commission européenne du texte « Vers une réforme du système d'évaluation de la recherche ».

**LES DERNIERS DOCUMENTS DE HAL**

Nicolas Mérou. Etude de l'écologie de *Martalia retrigens* et *Bornania osseae*, deux parasites protozoaires de l'huître plate. *Ostrina* adults. Sciences agricoles, Université de La Rochelle, 2021. Français. (NNT : 2021LARO9006). (tel-03606057)

Adel Cherchali. Numerical methods for the ALM. General Mathematics [math.GM]. Université Paris-Est, 2021. English. (NNT : )

UNIVERSITÉ DE LORRAINE



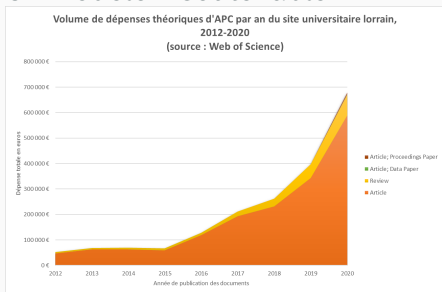
## La voie favorite des éditeurs

- Dépenses croissantes
- Conflits d'intérêt
- Revues prédatrices
- Revues hybrides

## La voie dorée n'est pas recommandée

- en général

## Un modèle insoutenable



## Gratuite pour le lecteur et l'auteur

- absence de conflit d'intérêt
- reprise en main de la partie éditoriale
- nouveaux modèles éditoriaux
- buts scientifiques et non lucratifs

HOME | ENSEIGNEMENT / RECHERCHE | RECHERCHE ET INNOVATION | DÉPÊCHE N°007304

### Publications : un "plan d'action" pour le modèle diamant est préparé par la Coalition S, Science Europe, Operas et l'ANR

La Coalition S, Science Europe, Operas, et l'ANR, travaillent à un "plan d'action" pour l'édition diamant avec le soutien du MESRI, présenté dans ses grandes lignes aux Journées européennes de la science ouverte le 4 février 2022. Johan Rooryck (Coalition S) précise qu'un "appel" sera lancé et une conférence organisée en juin sur les modalités de mise en œuvre. Pour Lidia Borrell-Damian (Science Europe), la voie diamant est "sous-estimée" alors qu'elle "peut s'appliquer à toutes disciplines". Pierre Mounier (OpenEdition et Operas) évoque un "idéal d'autonomie et d'autogestion des chercheurs".



## Publish & Read

- Apparenté à la voie diamant
- Nouvel eldorado des éditeurs
- Choix difficile pour les établissements



## Aucune intervention nécessaire

- Coût *apparent* nul
- *Alexandra Elbakyan* au secours (bénévole?) de tous les chercheurs de la planète
- Interrogations :
  - Légalité
  - Pérennité
  - Modèle économique



## Initiatives éditoriales diamant (ou presque)

### Ouvrages en accès ouvert

- modèle économique encore à affiner
- partenariat avec OpenEdition
- plateforme éditoriale partagée

### Revue en accès ouvert

- Open U Journals en gestation
- Pour accueillir vos initiatives éditoriales
- Partenariat Bordeaux et Inrae

# Gestion des données de la recherche

---

Publier des données?

# De quelles données parle-t-on ?

## Ne sont pas concernées

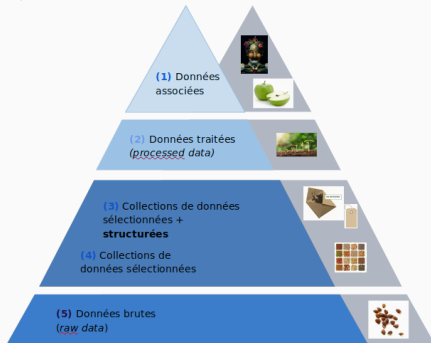
- Les données intermédiaires
- Les données en cours de traitement

## Les données sont, *per se*, un produit de la recherche

- Données validées
- Données citables

## Cycle de vie des données

### Pyramide des données



# Les différentes façons de déposer & publier des données

*Ouvrir autant que possible en fermant autant que nécessaire*

## **Données publiques**

- Métadonnées publiques
- Données publiques

## **Données sous embargo**

- Métadonnées publiques
- Données publiques après un temps d'embargo

## **Données fermées mais métadonnées publiques**

- Métadonnées publiques
- Données fermées mais accessibles sur demande





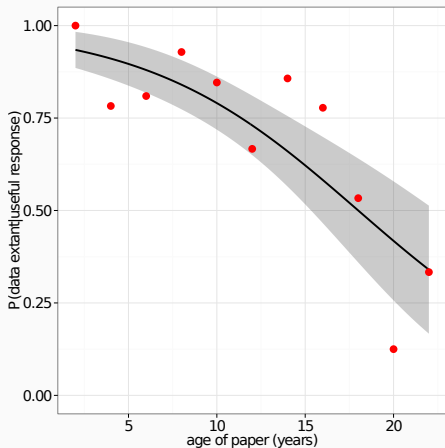
# Gestion des données de la recherche

---

Enjeux

## Les données sont une richesse

- Stockage pérenne
- Description
- Traçabilité



Source : The Crumbling Wall : Data Archiving and Reproducibility in Published Science, *North American DDI (NADDI) Conference 2014*

## 12 jeux de données à l'Université de Lorraine ?



RECHERCHE - OUVERTURE D'UTILISATION - SUPPORT - FRANÇAIS - SE CONNECTER

### Une production au même titre que les publications

- Témoins d'une activité
- Indicateur dans l'évaluation
- Appropriation très variable en fonction des disciplines

POUR DÉPOSER VOS DONNÉES, IL EST NECESSAIRE DE SE CONNECTER AVEC PREMIERE FILLE ET DE DEMANDER L'OUVERTURE DES DROITS ►► MENU SUPPORT ►►

**Déposer vos données en 30 minutes chrono !** Tous les guides à votre disposition ►► MENU GUIDE D'UTILISATION ►►

Cherchez cet endroit :  **Q Chercher** Recherche avancée

1 à 10 de 76 résultats ▶ 10

<b>Collection(s)</b> Jeu de données (12) Partenaires (1) (en)
<b>Catégorie de la collection</b> Laboratoire(s)
<b>Année de publication</b> 2021 (1) 2020 (10) 2019 (5)
<b>Auteur(s)</b> Gourdeau, Guillaume (2) Aïme-Pierre, Pierre-Corinne (1) Jégo, Bertrand (1) Aubert, Emmanuel (1) Gault, Aurélien (1)
<b> Sujet</b> Phytogéologie (1) Géologie (1) Minéralogie, cristal et sciences de la vie (4) Sciences de la terre et de l'atmosphère (12)
<b>Date de dépôt</b> 2021 (1) 2020 (6)

**EBSD Maps acquired on different steels with transformed microstructures**  
4 Nov. 2022 - LENS  
Brenauer, Simon; Martinez-Ostarragal, Tomas; Fritous, Bianca; Gey, Mathias; Coustaret, Audrey; Loukatchev, Natalia; Aïme-Pierre, Pierre-Corinne; Courtenot, Lionel. 2022. "EBSD Maps acquired on different steels with transformed microstructures". <https://doi.org/10.21203/rs.3.rs-2095517>. Université de Lorraine. V1. VREF: ean32p3q323d4703mg1 [Fractal]

EBSD Maps as CIF format is a JSON text format developed by Dorel, but not sufficiently explicit to be used outside of Oxford software). Acquisition parameters and steel grade are given in the CSV file.

**Soil methane oxidation in naturally-regenerated oak-dominated temperate deciduous forest stands responds to soil water status regardless of their age.**  
5 Nov. 2022 - [View](#)  
Dres, Nicolas; Pêan, Caroline; Epron, Daniel. 2022. "Soil methane oxidation in naturally-regenerated oak-dominated temperate deciduous forest stands responds to soil water status regardless of their age". <https://doi.org/10.1016/j.geoderma.2022.115744>. Université de Lorraine. V1. VREF: ean32p3q323d4703mg1 [Fractal]

The oxidation of atmospheric methane (CH<sub>4</sub>) by methanotrophic bacteria in upland forest soils is an important but overlooked ecosystem service. Our objective was to determine which factors drive variability in soil CH<sub>4</sub> oxidation spatially and seasonally in oak-dominated temperate deciduous forests.

**Données de réplication pour : Raw data from mass spectrometry (MS and MS-MS) analysis of new genetically engineered stamobomycin analogues, the 37 membered mini-stamobomycin.**  
16 Oct. 2021 - [DataFile](#)  
Aïme, Bertrand; Weissenau, Kira. 2021. "Données de réplication pour : Raw data from mass spectrometry (MS and MS-MS) analysis of new genetically engineered stamobomycin analogues, the 37 membered mini-stamobomycin". <https://doi.org/10.21203/rs.3.rs-1533051/v1>. Université de Lorraine. V1

The modular organization of the type 1 polyketide synthase (PKS1) is proof that the modular engineering of desirable analogues. By combining multiple, state-of-the-art approaches including modification of docking domain, use of modules of varying domain complexity, alternative...

**Observation in situ de l'évolution de la microstructure d'un alliage de Titane (TA6V) au cours d'un traitement thermique en microscopie confocale**  
9 Nov. 2021 - [View](#)

*« Ce serait vraiment bien si on avait plus d'interactions ouvertes, pas seulement en termes de santé publique, mais aussi dans le partage des données de recherche »*



**Anthony Fauci**

White House Coronavirus Task Force

Huffington Post, le 28 janvier 2020



## Transparence des données recueillies

3.

### Share trusted information

Visit the WHO website or chat to your doctor or nurse to find answers to common questions

How do we know the vaccines are safe?

They've been thoroughly tested and reviewed. If you're interested, I know where we can find more information.



## Ouvrir les données pour

- la reproductibilité
- la transparence
- le contrôle par les pairs

## Données et publications

- Les éditeurs commencent à exiger des données accessibles

## Décret du 3 décembre 2021

- Lien intégrité & ouverture des données

### Essay

## Why Most Published Research Findings Are False

John P.A. Ioannidis

### Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lower production of tested relationships; when there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance. Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias in this way. Decisions on the conduct and interpretation of research,

Published research findings are sometimes refuted by subsequent evidence, with ensuing confusion and disappointment. Refutation and controversy is seen across the range of research designs, from clinical trials and traditional epidemiological studies [1–3] to the most modern molecular research [4,5]. There is increasing concern that in modern research, false findings may be the majority or even the vast majority of published research claims [6–8]. However, this should not be surprising. It can be proven that most claimed research findings are false. Here I will examine the key

The Essay section contains opinion pieces on topics of broad interest to a general medical audience.

factors that influence this problem and some correlates thereof.

### Modeling the Framework for False Positive Findings

Several methodologists have pointed out [9–11] that the high rate of nonreproduction (lack of confirmation) of research discoveries is a consequence of the coexistent, yet ill-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a  $p$ -value less than 0.05. Research is not most appropriately represented and summarized by  $p$ -values, but, unfortunately, there is a widespread notion that medical research articles

### It can be proven that most claimed research findings are false.

should be interpreted based only on  $p$ -values. Research findings are defined here as any relationship reaching formal statistical significance, e.g., effective interventions, informative predictors, risk factors, or associations. “Negative” research is also very useful. “Negative” is actually a misnomer, and the misinterpretation is widespread. However, here we will target relationships that investigators claim exist, rather than null findings.

As has been shown previously, the probability that a research finding is indeed true depends on the prior probability of it being true (before doing the study), the statistical power of the study, and the level of statistical significance [10,11]. Consider a  $2 \times 2$  table in which research findings are compared against the gold standard of true relationships in a scientific field. In a research field both true and false hypotheses can be made about the presence of relationships. Let  $R$  be the ratio of the number of “true relationships” to “no relationships” among those tested in the field.  $R$

is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is  $R/(R+1)$ . The probability of a study finding a true relationship reflects the power  $1 - \beta$  (one minus the Type II error rate). The probability of claiming a relationship when none truly exists reflects the Type I error rate,  $\alpha$ . Assuming that  $\alpha$  relationships are being probed in the field, the expected values of the  $2 \times 2$  table are given in Table 1. After a research finding has been claimed based on achieving formal statistical significance, the post-study probability that it is true is the positive predictive value, PPV. The PPV is also the complementary probability of what Wacholder et al. have called the false positive report probability [10]. According to the  $2 \times 2$  table, one gets  $PPV = (1 - \beta)R / (R - \beta R + \alpha)$ . A research finding is thus

**Caption:** Ioannidis JPA (2005) Why most published research findings are false. *PLoS Med* 2(8): e124.

**Copyright:** © 2005 John P.A. Ioannidis. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Abbreviation:** PPV, positive predictive value

John P.A. Ioannidis is in the Department of Clinical and Epidemiological Medicine, University of Athens School of Medicine, Ioannidis, Greece and Institute for Clinical Medicine, Tufts New England Medical Center, Tufts University School of Medicine, Boston, Massachusetts, United States of America. E-mail: jioannidis@cs.tufts.edu

**Competing interests:** The author has declared that no competing interests exist.

**DOI:** 10.1371/journal.pmed.0020124



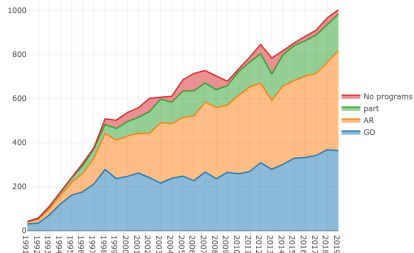
## Publication des données

- Dissémination rapide
- Réutilisation pour d'autres usages

## Naissance d'une pandémie

- Données non communiqués
- Attente de publication dans un journal prestigieux

## Hubble Space Telescope



Bleu : mêmes auteurs pour données et articles  
Jaune : articles sur données réutilisées

6 mars 2020

## Les données obtenues sur fonds publics devraient être publiques

- pour les entreprises
- pour les citoyens

## Ordre national du mérite

COVIDTRACKER

ACCUEIL | ENREGISTRÉS | OUTILS | AUTRES

### Statistiques et visualisations de données Covid19

CovidTracker est un outil permettant de suivre l'évolution de l'épidémie à Coronavirus en France et dans le monde. Pour des analyses quotidiennes des chiffres, vous pouvez suivre @publieumonitor sur Twitter, ainsi que @covidtracker\_fr.

### En un coup d'œil

Max à jour : 31er

137

**Taux d'incidence**

Nombre de cas par semaine pour 100k habitants. Le seul d'entre eux 50.

2.13%

**Taux de positivité**

Proportion des tests qui sont positifs parmi l'ensemble des tests.

0.79

**Taux de reproduction R1**

Nombre de personnes contaminées par 1 individu. Au-dessus de 1, l'épidémie progresse.

43.2%

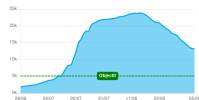
**Tension hospitalière**

Nombre de lits de réanimation occupés par les patients Covid19 par rapport au nombre de lits de 2014.

### Cas positifs

Plus

On préleve en moyenne **13 117** cas positifs au Covid19 chaque jour, **en baisse (-24%)** par rapport à la semaine dernière (par date de prélèvement, J-5).



Derniers chiffres : 14 534 tests positifs remontés le 07/09 (5pF), 2 725 tests positifs prélevés le 05/09 (SI-DEP).

### Adm. soins critiques

Plus

Il y a en moyenne **150 admissions en soins critiques** pour Covid19 chaque jour, **en baisse (-8%)** par rapport à la semaine dernière.



### Soins critiques

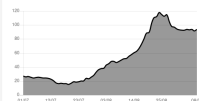
Plus

Il y a actuellement **2 224 personnes en soins critiques** pour Covid19, **en baisse (-3%)** par rapport à la semaine dernière.



### Décès hospitaliers

Il y a en moyenne **94 décès hospitaliers** pour Covid19 chaque jour, **en hausse (+1%)** par rapport à la semaine dernière.





# Gestion des données de la recherche

---

Politiques publiques

## Structurer, partager et ouvrir les données de la recherche

- Mettre en œuvre l'obligation de diffusion des données de recherche financées sur fonds publics
- Créer Recherche Data Gov
- Promouvoir l'adoption d'une politique de données pour les rendre faciles à trouver, accessibles, interopérables et réutilisables (FAIR)



## Article 6

Les établissements [...] **définissent une politique** de conservation, de communication et de réutilisation des résultats bruts des travaux scientifiques menés en [leur] sein. A cet effet, **ils veillent à la mise en œuvre par leur personnel de plans de gestion de données et contribuent aux infrastructures** qui permettent la conservation, la communication et la réutilisation des données et des codes sources.

## Gestion des données de la recherche

### Obligations



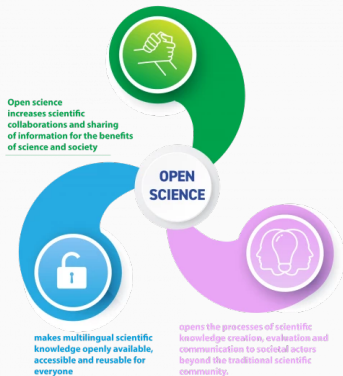
### Moyens



Deed of association of the new EOSC Association signed on 29 July 2020

Open research data that [...] that can be openly used, reused, retained and redistributed by anyone, subject to acknowledgement. Open research data are available in a timely and user-friendly, human- and machine-readable and actionable format.

## Recommandation de l'UNESCO



# Gestion des données de la recherche

---

Politiques institutionnelles

La plateforme nationale  
fédérée des données  
de la recherche

**LANCEMENT**  
Printemps 2022



**ACCOMPAGNER**  
LES ÉQUIPES DE RECHERCHE



**DÉPOSER & PUBLIER**  
DES DONNÉES DE RECHERCHE



**DÉCOUVRIR**  
LES DONNÉES DE RECHERCHE

## Plateforme nationale Recherche Data Gov

- Dépôt & publication de données
- Référencement / Catalogue
- Accompagnement / Formations



RECHERCHE ▾

GUIDE D'UTILISATION

SUPPORT

FRANÇAIS ▾

SE CONNECTER



## Plateforme lorraine **DOREL**

- Dépôt & publication de données
- Référencement / Catalogue
- Accompagnement / Formations
- Fusion à terme avec *Recherche Data Gouv*

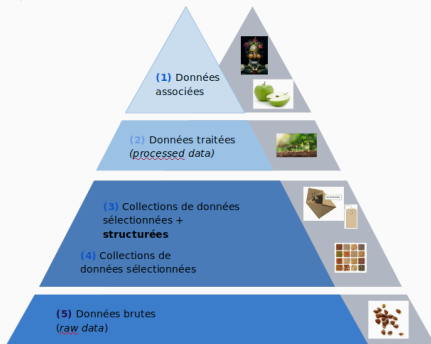




# Quelles données déposer/publier?

- Stratégie de valorisation des données à construire
- Toute donnée réutilisable peut être
  - déposée,
  - publiée,
  - partagée.
- Politique des données à définir

## Pyramide des données



## Une politique nationale

- Un AMDAC *par ministère*
- Un ADAC *par établissement*

## Les codes sources entrent dans le périmètre

- Objectif de libération des codes produits par les chercheurs
  - Tout est à construire
  - Grande disparité disciplinaire

- Élaborer la politique de l'établissement
  - Accompagnement
  - Infrastructure
  - Reconnaissance
  - Incitation
  - ...
- Capital *données* de l'établissement
  - Identifier
  - Valoriser
- Partage des données
  - Risques
  - Opportunités
  - Coordination des services
- Réseau national des ADAC

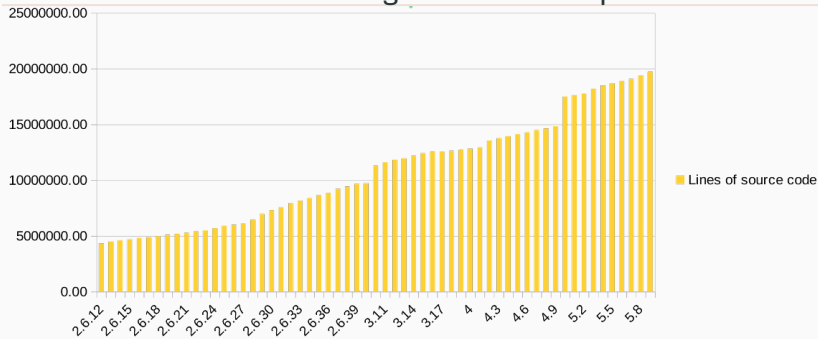
Quelques planches parmi celles qui suivent ont été inspirées de celles que Roberto Di Cosmo a présenté lors de la première Conférence Européenne sur la Science Ouverte (OSEC2022)

## Le pilier logiciel

---

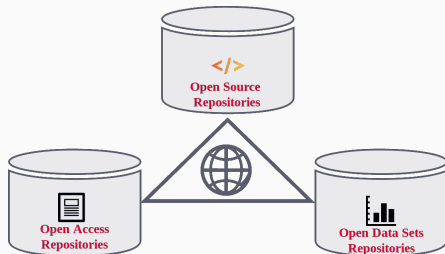
Spécificités logicielles

## Linux Kernel : 20 million de lignes dans votre poche



## Publications, données...& logiciel

- [...] software [...] essential in their fields. *Top 100 papers (Nature, 2014)*
- Sometimes, if you don't have the software, you don't have the data *Christine Borgman, Paris, 2018*

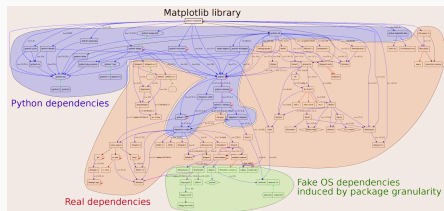


## Les logiciels évoluent

- sur des dizaines d'années
- leur *histoire* est importante

## Complexité logicielle

- Des *millions* de lignes
- Interdépendance extrême
  - facile à endommager
  - difficile à maintenir
- Des communautés de développement sophistiquées





# Le pilier logiciel

---

Enjeux

## Software Heritage

- Réalisation INRIA
- Archive logicielle mondiale
- Système de référencement ultra-fin : le *Software Heritage ID* (SWHID)
- Citation du logiciel
- Reproductibilité
- Traçabilité



We are building the universal software archive



Collect  
Preserve  
Share

We **collect** and **preserve** software in source code form, because software embodies our technical and scientific knowledge and humanity cannot afford the risk of losing it.

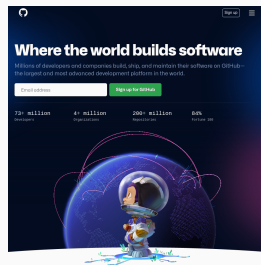
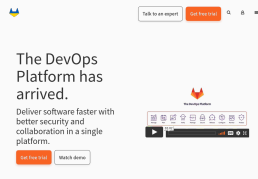
Software is a precious part of our cultural heritage. We curate and make accessible all the software we collect, because only by **sharing** it we can guarantee its preservation in the very long term.

[Discover our mission](#)



## Des communautés très organisées

- Forges logicielles
  - GitHub (Microsoft), GitLab (libre), sourceforge, code.gouv.fr...
- Workflow mondialisés
  - Bugzilla





# Le pilier logiciel

---

Politiques

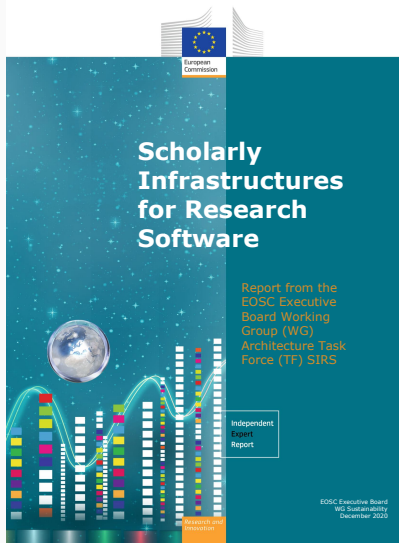
## Ouvrir & promouvoir les codes sources produits par la recherche

- Valoriser et soutenir la diffusion sous licence libre des codes sources
- Mettre en valeur la production des codes sources
- Définir et promouvoir une politique en matière de logiciels libres

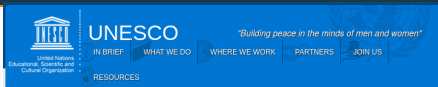


## EOSC SIRS report (2020)

*All research software should be made available under an Open Source license by default, and all deviations from this default practice should be properly motivated*



UNESCO collaborates to the project through its programs on documentary heritage preservation and access, including the Memory of the World programme and the Free and Open Source Software programme.



Home > Software Heritage

## Software Heritage



## Software Heritage

*"Part of our Heritage, Pillar of our Present, Enabler of our Future"* Paris Call on Software Source Code as Heritage for Sustainable Development

In today's digital world, software is everywhere: it allows our phones and computers to run, it powers our industries, it represents the basis of scientific research, it interfaces citizens with their states and enables the access to public services, among many other applications. In one word, software shapes every aspect of our lives.

However, software is often disregarded, and its preservation is at stake. Until recently, software source code was not stored in a reliable archive, but was either offline and not available to the public or it was stored on working platforms which may one day disappear. The lack of recognition that software is a unique form of human knowledge, a creative expression of problem solving and part of our heritage compounded the issue of preserving software. Software is fragile.

This common understanding and the resolve to protect software source code constitutes the essence of the collaboration between UNESCO and Inria, the French national research institute for the digital sciences. Inria first incubated the project [Software Heritage](#), which safely collects, curates and archives software source code and is the first worldwide archive of software. UNESCO collaborates to the project through its programs on documentary heritage preservation and access, including the Memory of the World programme and the Free and Open Source Software programme.

These were the key moments of the collaboration between Inria and UNESCO within the Software Heritage Project:

**April, 2017: Agreement signed between UNESCO and Inria**

**June, 2018: Software Heritage launch ceremony at UNESCO**



# Le pilier logiciel

---

Tout est à faire

## Tout est à faire dans les établissements

- L'ouverture des publications commence à être comprise
- La politique des données démarre
- Celle des logiciels reste à inventer

## Beaucoup est déjà fait dans certaines communautés

- Linux a eu 30 ans en 2021
- Le logiciel libre est partout
- Les communautés sont très organisées
- Elles ont leurs habitudes mais peuvent apporter beaucoup

Les Administrateurs des Données, des Algorithmes et des Codes Source en première ligne

## On oppose souvent logiciel libre et valorisation industrielle

- Un modèle économique à repenser
  - Ubuntu et Red Hat sont des sociétés privées
- Ce n'est pas le métier des chercheurs que de *vendre* un logiciel
  - ils produisent le moteur libre
  - une entreprise peut s'en emparer pour faire l'emballage
  - l'entreprise aura toujours besoin des chercheurs
- Les laboratoires peuvent également valoriser les services
- Réconciliation entre publication ouverte, dissémination et valorisation économique

# Évolution nécessaire de l'évaluation de la recherche

---

Enjeux

## D'une part : politique Science Ouverte

- ANR & Horizon Europe
  - Plan S
  - Plan de Gestion de Données
- Plan National pour la Science Ouverte
- CNRS
  - CRAC / RIBAC

## D'autre part, une évaluation quantitative : JIF, h index

- Promotion & recrutement
- Appels à projets
- I-SITE / IDEX
- International



## Recherche

- Articles
- Ouvrages
- Données
- Bases de données
- Algorithmes et codes sources
- Vulgarisation
- Valorisation
- Plateformes technologiques
- ...

## Enseignement

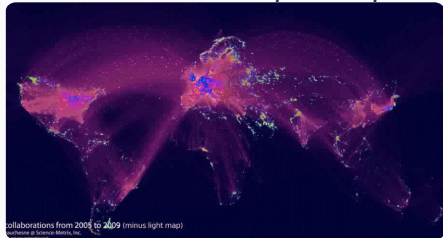
- Ouvrages
- Manuels
- Cours en ligne
- Ressources éducatives
- Scénarii pédagogiques
- ...

# L'enjeu du multilinguisme dans l'évaluation

L'usage quasi-exclusif de l'anglais induit un biais international

- en défaveur de certaines disciplines
- en faveur d'une partie du monde

Collaborations vues par Scopus



Collaborations vues par Redalyc



Source : Arianna BECERRIL, deuxièmes Journées Nationales de la Science Ouverte



## Processus de recherche

- Reproductibilité
- Intégrité scientifiques
- Traçabilité des résultats
- Protocoles expérimentaux
- ...

## Diffusion des résultats

- Ouverture de la science
- Initiatives éditoriales
- Diffusion des données de la recherche
- Diffusion et entretien de codes sources
- Organisation de conférences
- ...

## Reprendre en main les activités éditoriales

- Politique de dissémination des articles
- Travail éditorial de mise en forme
- De nouveaux modes éditoriaux à inventer

## Prendre soin de ses données et les valoriser

- Développer une politique des données
- Mettre en forme, décrire, publier et valoriser ses données

## Libérer les productions logicielles

- Comprendre les communautés du libre
- Formaliser une politique
- Prendre soin du code et le publier et l'entretenir

# Évolution nécessaire de l'évaluation de la recherche

---

Déclarations, intentions, actions

## Déclaration de San Francisco sur l'évaluation de la recherche

- La nécessité de mettre un terme à l'utilisation d'indicateurs basés sur les revues, comme les facteurs d'impact, dans le financement, les nominations et les promotions;
- celle d'évaluer la recherche sur sa valeur intrinsèque plutôt qu'en fonction de la revue où elle est publiée;
- celle encore d'exploiter au mieux les possibilités offertes par la publication en ligne (comme la levée de restrictions inutiles sur le nombre de mots, de figures et de références dans les articles et l'exploration de nouveaux indicateurs d'importance et d'impact).

Ne pas utiliser les indicateurs basés sur les revues, tels que les facteurs d'impact, comme succédané d'appréciation de la qualité des articles de recherche individuels, pour évaluer les contributions d'un scientifique en particulier ou pour prendre des décisions en matière de recrutement, de promotion ou de financement.

- Afficher explicitement les critères utilisés dans les décisions de recrutement, de titularisation et de promotion, en soulignant clairement, surtout pour les chercheurs débutants, que le contenu scientifique d'un article est beaucoup plus important que les indicateurs de publication ou l'image de marque de la revue dans laquelle il a été publié.
- Aux fins de l'évaluation de la recherche, tenir compte de la valeur et de l'impact de tous les résultats de travaux de recherche (y compris les jeux de données et les logiciels) en plus des publications scientifiques, et envisager un large éventail de mesures d'impact, y compris des indicateurs qualitatifs sur les retombées des travaux, comme leur influence sur les politiques et les pratiques.

## Appels pour réformer l'évaluation de la recherche

- 2015 : Leiden Manifesto
- 2016 : Amsterdam call for action
- 2018 : Appel de Jussieu
- 2019 : Hong Kong principles
- ...

Ce dont on ne veut plus est clair. Ce que l'on veut l'est moins.

Construisons une coalition européenne, sur le modèle de la cOAlition S, pour construire **ensemble** et changer les modalités d'évaluation, au niveau au moins européen.



# Construction d'une coalition européenne

- 26/11/2021 – *scoping report* de la Commission Européenne
  - Towards a reform of the research assessment system
- 16/12/2021 – Call for interest
  - Pour la construction d'une coalition
- 01/02/2022 – Appel de Paris sur l'évaluation de la recherche
- 4&5/02/2022 – les Journées européennes de la science ouverte (OSEC2022) : annonce politique de la démarche
- 10/06/2022 – Conclusions du Conseil de l'Union européenne sur l'évaluation de la recherche et la mise en œuvre de la science ouverte
- 01/12/2022 – Création de la coalition CoARA
  - Coalition for Advancing Research Assessment



[...] appellent à la **création d'une coalition** d'agences de financement de la recherche, d'institutions de recherche et d'autorités d'évaluation, volontaires et engagés pour une réforme du système actuel d'évaluation de la recherche selon des objectifs et des principes communs, et portant des actions communes (telles qu'un partage de documentation et de bonnes pratiques, ainsi qu'un dispositif commun de suivi de la mise en œuvre). **Le succès d'une telle coalition sera étroitement lié** à sa capacité à proposer des processus concrets de mise en œuvre, et **à sa capacité à associer et impliquer les chercheurs à tous les niveaux.**

1. Réforme des systèmes d'évaluation de la recherche en Europe
2. Approche européenne et capacités de publication universitaire et de communication scientifique
3. Développement du multilinguisme pour les publications universitaires européennes

RECONNAÎT que les systèmes d'évaluation de la recherche devraient être centrés sur la qualité et l'impact, et RAPPELLE que les systèmes d'évaluation de la recherche sont actuellement, dans une large mesure, trop axés sur l'utilisation de certains indicateurs quantitatifs fondés sur des revues et des publications et sur l'évaluation d'un éventail restreint de résultats de recherche; ESTIME qu'une telle approche peut avoir des effets négatifs sur la qualité, la reproductibilité et l'intégrité de la recherche; INSISTE sur le fait que l'évaluation de la recherche devrait inclure d'autres résultats et processus de recherche et promouvoir le partage des connaissances à un stade précoce et la collaboration afin d'accélérer la mise en œuvre des politiques et des pratiques en matière de science ouverte;

## CoARA : 4 engagements principaux — 4 core commitments

- **Recognise the diversity of contributions** to, and careers in, research in accordance with the needs and nature of the research
- **Base research assessment primarily on qualitative evaluation** for which peer review is central, supported by responsible use of quantitative indicators
- **Abandon inappropriate uses in research assessment of journal- and publication-based metrics**, in particular inappropriate uses of Journal Impact Factor (JIF) and h-index
- **Avoid the use of rankings of research organisations** in research assessment



- **Commit resources** to reforming research assessment as is needed to achieve the organisational changes committed to
- **Review and develop research assessment** criteria, tools and processes
- **Raise awareness of research assessment reform** and provide transparent communication, guidance, and training on assessment criteria and processes as well as their use

.../...

.../...

- **Exchange practices** and experiences to enable mutual learning within and beyond the Coalition
- **Communicate progress** made on adherence to the Principles and implementation of the Commitments
- **Evaluate practices**, criteria and tools based on solid evidence and the state-of-the-art in **research on research**, and **make data openly available** for evidence gathering and research

## Ne pas sous estimer les ressources nécessaires

- Une meilleure évaluation demandera un travail supplémentaire conséquent aux évaluateurs
- Des formations pour les évaluateurs et les évalués seront probablement nécessaires



Reconstruisons le ensemble

