

Fondements et enjeux de la science ouverte

Violaine Louvet

Journées Informatiques IN2P3/IRFU 2024



1 Quelques éléments historiques

2 Les enjeux

3 Les freins

4 Les acteurs

5 Et concrètement ?

Science ouverte ?

Définition

La science ouverte est la diffusion **sans entrave** des résultats, des méthodes et des produits de la recherche scientifique.



- Elle est rendue possible grâce aux opportunités offertes par la **mutation numérique**
- ouvrir le savoir scientifique en le rendant facile à trouver, librement accessible, réutilisable, échangeable dans un cadre scientifique, juridique, éthique et technique adapté.

Plan

1 Quelques éléments historiques

2 Les enjeux

3 Les freins

4 Les acteurs

5 Et concrètement ?

Bref historique

Avant le XVIIe siècle : la science des mécènes

Les scientifiques dépendent de patronages aristocratiques : leurs mécènes veulent exploiter leur travail (à des fins commerciales ou ludiques). Le savoir circule peu.

Emergences des académies

Pour pallier les déficiences du patronage aristocratique, les premières académies voient le jour au XVIIe. Les revues scientifiques en découlent directement.

L'avènement du numérique

L'arrivée d'internet et du numérique bouscule les processus scientifiques et les pratiques de publications.

1991 : création de l'archive ouverte de prépublications ArXiv

Les dates clés du mouvement pour la science ouverte

Les dates clés



Source : <https://enssib.libguides.com/science-ouverte>

Les textes fondateurs au niveau national

- **2016** : loi pour une république numérique
 - Article 30 : possibilité aux chercheurs de diffuser leurs articles en libre accès (HAL, etc.), si au moins 50% de financement public, et ce même s'ils ont au préalable cédé leurs droits d'exploitation de façon exclusive à un éditeur.
 - Article 38 : les données de la recherche produites par les chercheurs dans le cadre de leurs activités de recherche et qui ne portent pas atteinte à la protection de la vie privée ou au secret des affaires doivent être mises en ligne par les universités.
- **2018** : premier plan national pour la science ouverte
 - généraliser l'accès ouvert aux publications,
 - structurer et ouvrir les données de la recherche,
 - s'inscrire dans une dynamique durable, européenne et internationale.
- **2021** : deuxième plan national pour la science ouverte
 - généraliser l'accès ouvert aux publications,
 - structurer, partager et ouvrir les données de la recherche,
 - ouvrir et promouvoir les codes sources produits par la recherche,
 - transformer les pratiques pour faire de la science ouverte le principe par défaut.

Les textes fondateurs au niveau national

- **2016** : loi pour une république numérique
 - Article 30 : possibilité aux chercheurs de diffuser leurs articles en libre accès (HAL, etc.), si au moins 50% de financement public, et ce même s'ils ont au préalable cédé leurs droits d'exploitation de façon exclusive à un éditeur.
 - Article 38 : les données de la recherche produites par les chercheurs dans le cadre de leurs activités de recherche et qui ne portent pas atteinte à la protection de la vie privée ou au secret des affaires doivent être mises en ligne par les universités.
- **2018** : premier plan national pour la science ouverte
 - généraliser l'accès ouvert aux publications,
 - structurer et ouvrir les données de la recherche,
 - s'inscrire dans une dynamique durable, européenne et internationale.
- **2021** : deuxième plan national pour la science ouverte
 - généraliser l'accès ouvert aux publications,
 - structurer, partager et ouvrir les données de la recherche,
 - ouvrir et promouvoir les codes sources produits par la recherche,
 - transformer les pratiques pour faire de la science ouverte le principe par défaut.

Le cas particulier du logiciel

- L' **échange libre de logiciels** a existé depuis les débuts de l'informatique, sans qu'on ait défini les principes de ces partages.
 - En particulier dans la communauté des hackers (« bidouilleurs »), désignant originellement ces jeunes étudiants en informatique du MIT, génies de la programmation
- Avec le développement des ordinateurs personnels, l'importance économique des logiciels apparaît. C'est la naissance des **logiciels propriétaires**
 - Bill Gates et son Open Letter to Hobbyist en 1976.
 - Copyright Act adopté aux Etats-Unis en 1976 protégeant les logiciels par le droit d'auteur américain.
- Richard Stallman fonde la **Free Software Foundation** (FSF) en 1985. Cet organisme formalise les **4 grands principes du libre dans des licences**, textes juridiques qui fixent les conditions pour exploiter un logiciel ou une oeuvre intellectuelle.

Le cas particulier du logiciel

- Le logiciel libre est donc avant tout une philosophie du partage des communs, de la connaissance .
- Cette philosophie est naturellement née dans le milieu universitaire .
- Le mouvement de la science ouverte, en particulier autour des données, ne fait finalement que reproduire ce qui existe dans le monde du logiciel depuis le début.
- Mais n'apparaît dans le mouvement de la SO que dans le 2e PNSO en 2021 ...



Le périmètre de la science ouverte aujourd'hui

Définition proposée par l'UNESCO, 2021

la science ouverte s'entend comme un concept inclusif qui englobe différents mouvements et pratiques visant à rendre les connaissances scientifiques multilingues, librement accessibles à tous et réutilisables par tous, à renforcer la collaboration scientifique et le partage des informations au profit de la science et de la société, ainsi qu'à ouvrir les processus de création, d'évaluation et de diffusion des connaissances scientifiques aux acteurs de la société au-delà de la communauté scientifique traditionnelle. Elle inclut toutes les disciplines scientifiques et tous les aspects des pratiques savantes, y compris les sciences fondamentales et appliquées, les sciences naturelles et les sciences sociales et humaines, et repose sur les piliers essentiels suivants : les connaissances scientifiques ouvertes ; les infrastructures de la science ouverte ; la communication scientifique ; la participation ouverte des acteurs de la société ; et le dialogue ouvert avec les autres systèmes de connaissances.

- Communication scientifique et publication des résultats,
- données de la recherche,
- logiciels libres,
- projets scientifiques,
- reproductibilité, réutilisation,
- transparence et intégrité,
- évaluation des résultats et pilotage,
- métriques,
- politiques européennes, nationales, locales.

Plan

- 1 Quelques éléments historiques
- 2 Les enjeux**
- 3 Les freins
- 4 Les acteurs
- 5 Et concrètement ?

Les enjeux de la science ouverte pour la société

- Le partage des connaissances scientifiques rend la recherche plus efficace , plus visible , moins redondante .
- La science ouverte modifie la façon dont la recherche s'inscrit dans la société en rétablissant la confiance entre le citoyen et le scientifique, et en permettant de renforcer l'intégrité scientifique.
- La publication des résultats scientifiques et le partage des données issues de recherches financées par des fonds publics en grande partie permet de reprendre le contrôle d'un système qui nous a échappé.
- La science ouverte est aussi un levier puissant pour réduire les inégalités et ouvrir la voie d'une plus grande égalité à l'échelle mondiale.

Des inégalités d'accès abyssales

Même dans de grandes universités, la mesure de l'ampleur de la fracture de l'accès aux productions scientifiques dans l'univers académique est énorme, et davantage hors des campus (petites et moyennes entreprises, dans les organisations non gouvernementales, dans les collectivités, dans les écoles ...)

Les enjeux de la science ouverte pour le chercheur

- **Accessibilité accrue des résultats** : diffusion plus large des connaissances. Renforcement de la visibilité des travaux des chercheurs, facilite la collaboration internationale, et permet à d'autres de reproduire et valider les expériences.
- **Collaborations facilitées** : favorise les échanges entre chercheurs, permet des collaborations plus rapides et plus efficaces
- **Transparence et reproductibilité** : encourage la reproductibilité des expériences, un enjeu majeur pour la validité scientifique.
- **Reconnaissance et évaluation des chercheurs** : remise en question (progressive) des méthodes traditionnelles d'évaluation basées principalement sur les publications dans des journaux à fort impact.

Plan

- 1 Quelques éléments historiques
- 2 Les enjeux
- 3 Les freins**
- 4 Les acteurs
- 5 Et concrètement ?

Les freins systémiques

- Le premier frein au libre accès aux articles académiques vient des **intérêts commerciaux** des maisons d'édition scientifique, dont le modèle économique repose sur l'abonnement.
- Pression sur les chercheurs **« publish or perish »**. Importance de prendre en compte de nouvelles productions (données, codes, ...) et de nouveaux critères dans l'évaluation (administration de la recherche, review, ...).
- **Infrastructures inégalement développées** : les plateformes de partage des résultats scientifiques ne couvrent pas tous les domaines, et ne sont pas forcément connues des chercheurs
- **Réticence culturelle et institutionnelle** : la transition vers la science ouverte nécessite un changement de méthodes mais aussi de mentalité au sein des communautés scientifiques et des institutions. Certaines disciplines sont plus conservatrices et peuvent être réticentes à abandonner le système actuel.

Les freins pour les chercheurs

- Le risque de **perdre la priorité sur une découverte** est un frein important, en particulier dans les domaines très compétitifs.
- Problèmes de **propriété intellectuelle** et de reconnaissance de la paternité des travaux : crainte que les idées ou données soient utilisées sans consentement ou **crédits** approprié
- Implique souvent de **nouvelles compétences** pour les chercheurs (gestion des données, utilisation de nouvelles plateformes pour le partage ...). Investissement **en temps** et en formation.
- **Coûts de publication et d'accès** : la publication en accès ouvert dans des revues de qualité peut engendrer des frais de publication (Article Processing Charges, ou APCs) importants. Attention aux revues hybrides (qui cumulent abonnements et APC), modèle économique rejeté par les financeurs !
Les infrastructures de partage des données peuvent aussi nécessiter des financements.
- **Protection des données sensibles** : le partage de données dans certains domaines, comme la santé ou les sciences sociales, pose des défis éthiques liés à la protection de la vie privée et à la confidentialité des données.

Qui finance la SO ?

- Principalement des financements publics
 - Rôle essentiel des gouvernements qui ont mis en place des fonds dédiés à la science ouverte, en particulier pour financer les infrastructures nationales (publications, données)
 - La commission européenne joue également un rôle crucial de soutien d'infrastructures dédiées à travers des appels spécifiques, mais aussi à travers la prise en charge des coûts de science ouverte des projets financés.
 - Les établissements et organismes de recherche jouent un rôle majeur à travers les bibliothèques universitaires, le soutien au plateforme et la formation et l'accompagnement des communautés scientifiques.
 - Les agences de financement, à travers des exigences d'ouverture des résultats des projets financés, ou des appels dédiés à la science ouverte.
- Certains modèles économiques d'éditeurs : APC ou accords transformants (entre consortiums de bibliothèques et éditeurs par exemple)
- Certaines fondations privées jouent un rôle important en offrant des subventions (comme le Wellcome Trust ou la Bill & Melinda Gates Foundation par exemple).

Qui finance la SO ?

L'exemple d'ArXiv

Créée en 1991 par Paul Ginsparg, offre un accès libre à environ 1,5 million de prépublications. Hébergée à l'université Cornell, depuis 2001.

Programme de financement depuis 2010 pour alléger la charge financière et réduire sa dépendance à une seule institution : participation des plus gros établissements utilisateurs. Le principe est celui d'une adhésion sur 5 ans avec une cotisation annuelle dont le montant est calculé en fonction du nombre d'articles téléchargés.

L'exemple de HAL

Créée en 2001, permet l'accès ouvert à plus de 750 000 documents scientifiques.

Ressources pérennes (1 M / an) apportées par les tutelles du CCSD (CNRS, INRIA, INRAE et Université de Lyon), et un nouveau modèle de financement depuis 2021 qui fait évoluer la cotisation des établissements dotés d'un portail institutionnel.

Tarifification indexée sur le nombre d'EC, et ajustée pour tenir compte de la diversité des situations au sein des établissements, y compris des plus petits d'entre eux.

Plan

- 1 Quelques éléments historiques
- 2 Les enjeux
- 3 Les freins
- 4 Les acteurs**
- 5 Et concrètement ?

Acteurs internationaux

Mouvement issu des chercheurs, des communautés scientifiques, relayé par les établissements et conforté par les états et les financeurs.

Un large éventail d'acteurs parmi lesquels :

- **L'UNESCO** : un des principaux promoteurs de la science ouverte à l'échelle internationale qui a adopté en 2021 une recommandation sur la science ouverte
- **La coalition S**, regroupement d'agences de financement de la recherche dont le plan S exige que les résultats de recherches financées par des fonds publics soient publiés en accès ouvert immédiatement après leur publication.
- **La RDA (Research Data Alliance)**, organisation internationale visant à promouvoir et à faciliter l'interopérabilité des données scientifiques et à favoriser leur partage.

Acteurs européens

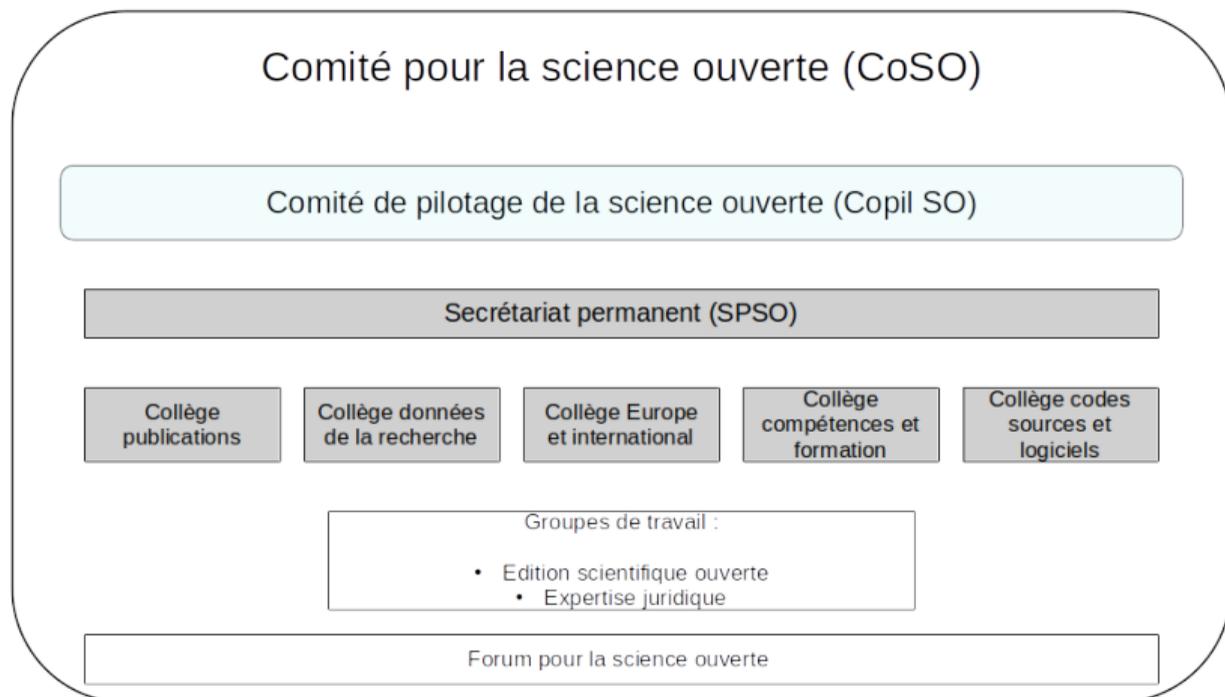
Beaucoup d'acteurs au niveau européen, en particulier :

- **La commission européenne**, un des acteurs les plus influents en matière de science ouverte en Europe. Elle intègre des principes d'ouverture dans ses programmes de recherche (Horizon Europe), qui impose que les résultats des recherches financées soient en accès libre.
- **EOSC (European Open Science Cloud)**, initiative phare de l'Union européenne visant à créer un environnement numérique fédéré pour faciliter l'accès, le stockage, le partage, l'analyse et la réutilisation des données de recherche à travers l'Europe.
- **COARA (Coalition for Advancing Research Assessment)** : initiative lancée en 2022, dont l'objectif est de réformer et d'améliorer les pratiques d'évaluation de la recherche, en promouvant des méthodes d'évaluation plus transparentes, inclusives, et alignées avec les valeurs de la science ouverte.

Acteurs nationaux

- Le MESR , qui joue un rôle clé de pilotage avec la publication des plans nationaux pour la science ouverte.
- Le COSO (Comité pour la science ouverte) qui coordonne et supervise la mise en œuvre des politiques de science ouverte en France.
- L'ANR qui impose désormais que toutes les publications issues de projets qu'elles finance soient déposées dans des archives ouvertes comme HAL et finance également des projets visant à développer les infrastructures de science ouverte en France.
- Couperin (Consortium Universitaire de Publications Numériques) acteur clé dans la négociation d'accords avec les éditeurs pour favoriser l'accès aux publications scientifiques. Il travaille également à la transition vers des modèles de publication en accès libre.
- Vous tous !!

Organisation du COSO



Plan

- 1 Quelques éléments historiques
- 2 Les enjeux
- 3 Les freins
- 4 Les acteurs
- 5 Et concrètement ?**

S'inscrire dans le mouvement de la science ouverte

- Anticiper les processus de recherche avec les plans de gestion de données (et de logiciels)
- Publier en accès ouvert (revues, préprints).
 - La loi de 2016 vous en donne le droit : déposer votre post print dans les archives ouvertes (ArXiv, HAL). Privilégier les revues émergeant au modèle diamant.
- Partager vos données dans des dépôts ouverts (FAIR).
 - Si pas de dépôts pour votre discipline : utiliser Recherche Data Gouv
- Partager vos codes et scripts en open source .
 - Idéalement sur une forge académique
- Publier vos protocoles et méthodes .
- S'engager dans des revues ouvertes (Open Peer Review).
- Sensibiliser et former vos collègues et étudiants à ces pratiques.

L'ouverture ne suffit pas : l'important c'est le chemin

La science ouverte est un changement culturel et méthodologique majeur qui nécessite des efforts et du temps

- Partager vos données : travail important de préparation, de curation, de documentation (métadonnées et descriptions claires)
- Rendre vos protocoles et méthodes accessibles : fournir des informations détaillées sur les outils, les équipements, et les logiciels utilisés afin que d'autres puissent répliquer vos travaux.
- Publier les codes sources, contribuer aux logiciels open source : documenter, commenter, structurer, faciliter l'installation et l'utilisation ...
- Former, sensibiliser, se former

Du temps et des efforts investis pour soi (pouvoir réutiliser ses propres données, codes ...), pour les autres (permettre la reproductibilité et la science cumulative)