



MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR, DE LA RECHERCHE ET DE
L'INNOVATION

Discours Science Ouverte à LIBER

Mme la Ministre Frédérique Vidal

Monsieur le Président de l'université de Lille, cher Jean-Christophe

Monsieur le Président Directeur Général de l'ANR, cher Thierry

Mesdames et Messieurs,

Dès ma prise de fonctions, j'ai décidé de faire la science ouverte une de mes priorités. La raison tient en trois points.

- Il y a tout d'abord l'importance de ce changement de paradigme qui ouvre la science dans tous les cercles académiques, et au-delà vers notre société, et constitue ainsi un élément indispensable pour progresser vers l'objectif fixé par le Président de la République concernant « la reconnaissance de la part que jouent la connaissance, la recherche, l'innovation, l'enseignement au cœur de la promesse de progrès qui est celle de la République ».
- Il y a ensuite le risque pour la France de rater le virage en cours au plan mondial vers la science ouverte.
- Il y a enfin la sous-exploitation de notre potentiel scientifique dans le cadre des pratiques fermées encore trop en vigueur.

Nous ne sommes pas suffisamment aujourd'hui dans une logique de science ouverte. Il faut payer pour avoir accès aux publications de nos chercheurs. Bien entendu, cela n'est pas un obstacle pour les universités les plus riches dans le monde et pour les grands groupes industriels. Mais cela limite fortement l'impact potentiel de nos recherches sur les campus de la planète, sur la société, sur l'industrie.

Peter Suber, le spécialiste mondial de l'accès ouvert, a montré que Yale, qui n'est pas exactement une institution démunie, avait 25% d'abonnements en moins que... Harvard. Et l'institution universitaire la plus riche en Inde, l'Indian Institute of Science, n'a accès qu'à 12% des abonnements que peut s'offrir Harvard. Cela signifie clairement que le système éditorial conçu au 19^e siècle et au 20^e siècle sur la base des contraintes et des modèles économiques du papier est devenu inadapté.

Ce système éditorial historique est de plus devenu obsolète à l'heure du web et de l'ubiquité numérique. Il a à l'origine été construit pour diffuser la connaissance. Il semble désormais fonctionner à front renversé, renforcer l'ignorance et construire des inégalités criantes d'accès. Les capacités d'accès aux publications varient de fait d'un lieu à l'autre, d'un pays à l'autre, d'une université à l'autre, voire d'un laboratoire de recherche à l'autre. C'est paradoxal, alors que les recherches publiques sont financées par tous, et alors que le Web offre

une infrastructure permettant d'accélérer la circulation des connaissances. La privation d'accès est encore plus vraie pour les organisations non gouvernementales, les médias, les petites et moyennes entreprises ou les enseignants, qui –tous– ont un besoin croissant d'accès aux produits de la recherche scientifique. Or les enjeux sont considérables et bien réels. J'en prendrai un exemple :

- Les barrières traditionnelles de la diffusion des résultats scientifiques auraient pu ralentir la lutte contre une maladie aussi rapide et mortelle qu'Ebola; heureusement les principes de la science ouverte se sont imposés pour ce type d'épidémie. Les anthropologues ont fait le pari d'une publication rapide et en accès ouvert sur Ebola. Cela a accéléré la prise en compte de la **culture des populations** pour le choix de la couleur des sacs mortuaires. Grâce à la diffusion large de ces recommandations, on a pu passer d'une couleur noire, culturellement inadaptée, à une couleur blanche, mieux acceptée et mieux comprise, et contribuer ainsi à réduire la vivacité de l'épidémie. Croyez-vous que la diffusion traditionnellement confidentielle d'articles aussi spécialisés en anthropologie aurait pu être à la hauteur de ce défi ?

Alors que les fausses nouvelles sont très facilement accessibles, les publications scientifiques sont protégées derrière des péages qui sont autant de barrières à l'accès au savoir. Il faut dix minutes pour inventer et diffuser une fausse nouvelle, croustillante, étonnante et facile à comprendre. Mais il faut dix ans pour produire une démonstration scientifique apportant une information scientifique de qualité. En tant que scientifiques, notre devoir est d'éclairer nos concitoyens. Nous devons pouvoir diffuser de façon rapide, transparente et complète nos résultats à tous les citoyens. La science ouverte est le véhicule idéal de la connaissance face aux rumeurs. Je ne me résigne pas à ce qu'il faille cantonner nos résultats scientifiques derrière une muraille de paiement, séparant les sachants des ignorants. J'attache ainsi une importance particulière à Wikipédia, qui est devenue une source d'information majeure pour toute la société. Il s'agit désormais du cinquième site le plus consulté au monde, et de loin la première source d'information encyclopédique au monde. La science ouverte est le seul moyen de rendre Wikipédia durablement fiable. Il y a actuellement des dizaines de milliers de citations dans la Wikipédia francophone vers des articles scientifiques. Je veux que nous allions plus loin encore, dans le tissage d'un lien étroit entre savoir profanes et savoirs savants, en allant au bout de la logique d'ouverture de nos publications, les transformant ainsi en biens communs.

Mais au-delà des progrès de société, la science ouverte est aussi la clé pour une meilleure recherche scientifique. En 1675, Isaac Newton écrit dans une lettre à Robert Hooke : « Si j'ai vu plus loin, c'est en montant sur les épaules de géants. ». Il reprenait une métaphore datant du 13^e siècle qui reste valide : on ne fait pas de bonne recherche sans s'appuyer sur les découvertes de chercheurs qui nous ont précédés. Mais cela est de moins en moins vrai, tant il devient impossible d'avoir accès à l'ensemble de la littérature scientifique mondiale. Publications hors de prix et données de la recherche inaccessibles sont de profonds obstacles aux progrès d'une science cumulative.

Faire le choix de la science ouverte, c'est autoriser notre production scientifique à trouver son lectorat, le lectorat large et universel qu'elle mérite. C'est ainsi que des dizaines d'études montrent une augmentation très significative à la fois des usages de lecture mais aussi des usages de citations pour les articles et les livres en accès ouvert. Et grâce à cette façon de publier, le potentiel d'interdisciplinarité est profondément renforcé. En effet, les frontières disciplinaires sont plus imperméables dans le monde des abonnements. Avec la science ouverte, la sérendipité permise par l'absence de frontière documentaire offre de nombreuses perspectives heuristiques.

Bien sûr, la tentation de la fermeture des publications et des données est forte. Pourtant,

la science est un bien commun, que nous devons partager le plus largement possible. Le rôle des pouvoirs publics est de rétablir la fonction initiale de la science, comme facteur d'enrichissement collectif. Car la diffusion des connaissances scientifiques a un impact direct en termes de développement économique, sanitaire, social. Je m'inspire de la stratégie qui a été mise en place dans le cadre du Human Genome Project, qui a coûté 3,8 milliards de dollars mais dont il a été décidé que les résultats seraient publics et considéré comme patrimoine de l'humanité. Cette décision a permis l'exploitation scientifique et médicale formidable que nous connaissons tous. D'ailleurs, on estime qu'en 2012 ce projet avait eu un impact économique de 796 milliards. Les données astronomiques sont traditionnellement diffusées après une année d'embargo. $\frac{2}{3}$ de la littérature scientifique en astronomie s'appuie sur des données ouvertes ! Ce potentiel énorme, exploité par les astronomes, ne l'est pas encore dans toutes les disciplines.

D'une certaine façon, la bibliothèque brûle.

Une vaste étude sur des jeux de données construits par les chercheurs en 1991 a montré qu'ils subissaient une déperdition rapide, atteignant 17% par an. Quand on sait qu'une partie importante des financements de la recherche est dédiée à la construction de corpus, de collecte de données, de réalisation d'expériences... un tel taux de disparition prive de découvertes et est inquiétant du point de vue de la solidité de la preuve scientifique :

- C'est une mauvaise gestion des fonds publics, car des données perdues ne sont plus réutilisables ni appariables avec de nouvelles données... Ce qui limite d'autant les nouvelles découvertes.

- Lorsqu'on revient sur une recherche menée il y a dix ans, et qu'il n'est plus possible de recomposer les données qui ont permis des conclusions importantes, il y a lieu de s'inquiéter. A un moment où l'intégrité scientifique est si importante pour défendre le rôle de la science dans la société, il est indispensable de maîtriser les données dans la durée.

Ainsi, la première démarche est de structurer et conserver les données, en préalable à leur ouverture. Cela se fait très bien dans environ un cinquième de la communauté scientifique, mais le reste est largement laissé à l'appréciation individuelle et aux hasards de la vie des clés USB, des vols d'ordinateurs portables dans les coffres de voitures ou des défaillances des disques durs individuels...

En tant que Ministre de l'enseignement supérieur, de la recherche et de l'innovation un de mes objectifs est de faire en sorte que toutes les communautés et tous les acteurs, lorsque ce n'est pas déjà le cas, développent une démarche structurée de conservation de leurs données et fasse le nécessaire pour que ces données obéissent au principe FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable). Une fois que nous aurons fait cela, nous pourrons rendre la science plus ouverte, plus transparente et d'accès plus universel. **Ces données pourront servir comme instrument pédagogique, comme substrat scientifique et comme catalyseur d'innovation.**

Le problème n'est pas principalement technique, il est d'abord, et fondamentalement, humain. J'ai conscience des injonctions contradictoires que s'imposent aux chercheurs dans ce domaine, entre ouverture et fermeture. Pour relever les défis de la science ouverte, nous devons réformer en profondeur tout notre système pour que l'ouverture des publications et des données soit enfin reconnue comme une **bonne pratique**. Le système d'évaluation, trop marqué par une métrique datée, mécaniste et réductrice, doit évoluer lui aussi vers l'ouverture. Oui, nous devons sérieusement réfléchir aux évolutions des systèmes et méthodes d'évaluation, évaluation des chercheurs comme des institutions. C'est la raison pour laquelle je soutiens le mouvement des citations ouvertes, qui va dans la bonne direction. Nous devons aussi abandonner les excès d'évaluation quantitative, où le facteur d'impact règne en maître, pour développer une approche beaucoup plus qualitative. Je souhaite que la France

suive dans ce domaine les propositions de la Déclaration de San Francisco sur l'évaluation de la recherche et applique les principes du Manifeste de Leiden.

J'ai également conscience de la nécessité de faire évoluer les pratiques quotidiennes, et donc également les compétences. Nous avons trop souvent –permettez-moi l'expression– « appris sur le tas » les composantes numériques de nos métiers. Et nous avons souvent acheté des équipements de recherche sans penser à exiger ou sans obtenir que les données soient décrites dans des formats ouverts, interprétables en dehors de la machine et de l'écosystème du fabricant. Nous avons aussi de facto abandonné la propriété intellectuelle de nos revues à des éditeurs qui sont peu à peu passé en position de nous imposer des conditions financières inacceptables [120M€ d'abonnements pour la recherche publique en France par an]. Je souhaite ainsi que les négociations avec les grands éditeurs se finalisent à un niveau sensiblement réduit par rapport aux pratiques actuelles et plus en rapport avec la contribution réelle très significative apportée par la communauté de recherche.

La science ouverte n'est pas une mode, ce n'est pas une discipline, c'est un paradigme. Elle comporte donc ses nouvelles pratiques et ses nouvelles compétences. J'attacherai une attention particulière à ce que ces compétences soient considérées comme faisant partie du bagage initial des jeunes chercheurs. La science ouverte ne sera pas l'affaire d'un petit groupe de spécialistes, elle devra imprégner l'ensemble du monde de la recherche.

Vous l'avez compris, la science ouverte ne peut se concevoir que par une approche globale, qui intègre toutes les facettes de l'activité scientifique.

Nous pouvons à terme **atteindre 100% de publications scientifiques françaises en accès ouvert.**

Nous devons engager des **processus d'ouverture des données de recherche**, quand cela est raisonnable, conforme à l'éthique et au droit.

Nous devons **développer des formations, de nouveaux outils, de nouveaux services**, ou simplifier et améliorer ceux qui existent déjà.

Mais nous devons aussi **nous inscrire dans le mouvement mondial de la science ouverte.** Je souhaite que la France soit proactive et *leader* dans le domaine de la science ouverte, en participant fortement au paysage mondial dans ce domaine. Elle soutient en particulier les initiatives de l'union européenne qui, depuis 2012, a adopté des politiques volontaristes dans le domaine de la science ouverte. C'est pourquoi nous soutiendrons le « plan S » pour les publications ouvertes que ScienceEurope et Robert-Jan Smits ont développées et qui sera annoncé lors du congrès ESOF à Toulouse en présence du Commissaire Carlos Moedas. Ainsi nous serons en phase de la mise en œuvre des Conclusions du Conseil Compétitivité de Mai 2016 en soutien total de l'agenda Science ouverte du Commissaire Moedas.

Le plan national pour la science ouverte met en place les conditions du développement d'une science ouverte en France. Il est composé de **trois axes** majeurs, qui inscrivent la recherche française au cœur du mouvement mondial d'ouverture des données et de transparence de l'action publique. Il s'agit de contribuer au partenariat pour un gouvernement ouvert (OGP), coalition internationale qui porte les principes de transparence, de participation citoyenne et d'innovation démocratique. Dans son discours à l'ONU en tant que coprésident de l'OGP, Emmanuel Macron a déclaré vouloir « faire de la révolution numérique une opportunité pour transformer en profondeur les démocraties, les rendre plus efficaces, responsables et participatives. ». Ce plan s'inscrit parfaitement dans cette dynamique

[Premier axe : généraliser l'accès ouvert aux publications]

Le premier axe de ce plan national pour la science ouverte consiste à généraliser l'ouverture des publications.

1. Développer l'accès ouvert permettra d'augmenter le rayonnement et l'accessibilité de la recherche française. C'est pourquoi j'ai décidé que **nous mettrons en place une obligation de diffusion en accès ouvert des articles et ouvrages issus des recherches financées par appel à projets sur fonds publics.**

2. La communauté scientifique a perdu des pans entiers de contrôle du système éditorial. C'est pourtant un système essentiel pour le bon fonctionnement de la science. Car il faut le répéter : la science ouverte ne se conçoit pas sans éditeurs. Nous disposons en France d'une communauté éditoriale scientifique diverse et de grande qualité, comme par exemple en mathématiques et en sciences humaines et sociales. Il faut cultiver cette diversité, dans un rapport de confiance et de dialogue. Il est indispensable d'aller vers une plus grande diversité et un plus grand équilibre du paysage éditorial. Pour développer cette biodiversité, **j'ai décidé de créer un fond pour la science ouverte.** Celui-ci contribuera au montage de solutions françaises et internationales innovantes, tant du point de vue technique que du point de vue économiques. Par exemple, la mise au point de modèles de financement participatifs mobilisant toutes les bibliothèques de la planète constitue une piste à creuser. La France contribuera ainsi à la reprise du contrôle du système éditorial par la communauté scientifique.

3. Nous disposons en France d'un atout unique, avec l'archive ouverte nationale HAL. Elle permet notamment de déposer en accès ouverts des articles qui sont publiés dans des revues fermées. Elle permet aussi de garantir une conservation à long terme de la production nationale : c'est bien une archive ouverte. Nous avons déjà considérablement simplifié les procédures de dépôts des publications par les chercheurs, divisant par trois le temps de dépôt. **J'ai décidé de soutenir HAL et de simplifier encore les processus d'alimentation** en développant les fonctions d'alimentation semi-automatique

[Deuxième axe : structurer et ouvrir les données de la recherche]

Le deuxième axe de ce plan national pour la science ouverte consiste à structurer et, autant que possible, à ouvrir les données de la recherche.

4. Les **données** de la recherche sont le pilier des connaissances futures. Structurer, conserver et ouvrir les données de la recherche, permettra de favoriser la mutualisation, le croisement et l'appariement des données pour produire de nouvelles connaissances. **C'est pourquoi le Président de la République a annoncé lors de son grand discours sur l'Intelligence artificielle qu'il allait mettre en place une obligation de diffusion en accès ouvert des données issues de recherches financées par appels d'offres sur fonds publics.** Il va de soi que ce processus d'ouverture sera conditionné par le respect de la réglementation, notamment en ce qui concerne les données personnelles ou les données protégées par le secret professionnel, les secrets industriels et commerciaux, et que les bonnes pratiques seront définies par la communauté scientifique. Un **appel FLASH de l'ANR** sera publié prochainement pour accélérer la structuration, la citation et l'ouverture des données de la recherche des équipes françaises.

5. Afin de coordonner l'action nationale, **j'ai décidé de créer la fonction d'administrateur des données (Chief data officer - CDO) au ministère de l'enseignement supérieur, de la recherche et de l'innovation.** Il coordonnera l'action publique en matière de données scientifiques, afin d'optimiser leur structuration, leur conservation et leur circulation. Cet administrateur des données travaillera avec les autres administrateurs des données ministériels regroupés autour de l'Administrateur général des données. Il animera également un réseau d'administrateurs des données dans les établissements.

6. Enfin, **j'ai décidé d'encourager l'association de données structurées et ouvertes aux articles publiés par les chercheurs.** C'est une première étape vers une meilleure

association des données avec les articles. Nous encouragerons aussi le développement des « articles de données » (*data papers*), qui décrivent la façon dont ont été construites les données.

[Troisième axe : S’inscrire dans une dynamique durable, européenne et internationale]

Le troisième et dernier axe vise à inscrire la France dans une dynamique durable, européenne et internationale.

Faire de la France un pays de la science ouverte, c’est à la fois transformer les pratiques scientifiques pour qu’elles intègrent et banalisent la science ouverte au quotidien, et contribuer à la structuration du paysage international de la science ouverte.

Pour cela, nous développerons les compétences de la science ouverte. Il s’agit, d’une part, de généraliser les pratiques quotidiennes de la science ouverte, notamment dans le domaine des publications, des données, de la propriété intellectuelle et de l’évaluation par les pairs.

Il s’agit, d’autre part, de contribuer à un écosystème à la fois résilient, régulé et transparent, œuvrant dans le sens des intérêts de la communauté scientifique.

7. Nous nous engageons dans un changement culturel profond, à la fois pour les institutions mais aussi et surtout pour les chercheurs. Pour que les chercheurs publient en accès ouvert leurs publications, leurs données, leur code source, il est nécessaire de développer de nouvelles compétences, notamment dans les écoles doctorales. C’est la raison pour laquelle nous encouragerons les formations à la science ouverte, notamment à travers un financement dédié porté par le ministère qui encouragera la mise au point et la mutualisation des supports de formation.

8. J’invite les universités et les organismes de recherche à se doter d’une politique de science ouverte combinant la question des publications, des données, des compétences et de l’évaluation. Je vous invite à transposer dans vos établissements et dans vos laboratoires les mesures de ce plan pour la science ouverte, en les adaptant à vos réalités organisationnelles et disciplinaires.

9. Il faut également articuler nos actions nationales avec le niveau européen et international. C’est la raison pour laquelle nous participons à l’EOSC (European open science cloud) porté par la Commission européenne ainsi qu’à GO FAIR, initiative conjointe des Pays-Bas, de l’Allemagne et de la France pour impliquer l’ensemble des communautés scientifiques dans le mouvement de transformation des données pour qu’elles soient Faciles à trouver, Accessibles, Interopérables et Réutilisables (FAIR).

Je voudrais conclure en vous disant qu’on parle de science ouverte depuis trop longtemps. Il est temps de **passer à l’action.** Ce plan rend obligatoire l’accès ouvert pour les publications et pour les données issues de recherches financées sur projets. Il met en place un Comité pour la science ouverte et soutient des initiatives majeures de structuration du paysage concernant les publications et les données. Enfin, il est doté d’un volet formation et d’un volet international qui sont essentiels à la mobilisation des communautés scientifiques et à l’influence de la France dans ce paysage en cours de constitution. Il nous faut le courage et la volonté de financer cette politique, qui ne se nourrit pas que de belles intentions. Ce plan est par conséquent doté d’un budget de 5,4M€ la première année et de 3,4M€ les années suivantes.

Ce plan réussira parce qu’il est nécessaire ! Comme des collègues l’ont si bien écrit dans

un texte publié par *Le Monde* en 2013, « un savoir enfermé est un savoir stérile ». Avec ce plan, je vous confie aujourd'hui les clés pour ouvrir les portes, pour abaisser les barrières. Je vous demande de vous les approprier, de les utiliser, de les diffuser. Car la réussite de cette politique de science ouverte est dans vos mains ! Ce n'est qu'ensemble, vraiment tous ensemble, que nous ferons, de la science ouverte un véritable outil d'enrichissement collectif.

Je vous remercie.