

Séminaire Open Science

le point de vue d'un enseignant-chercheur

Nicolas Fressengeas

Laboratoire Matériaux Optiques, Photonique et Systèmes
Unité de Recherche commune à l'Université de Lorraine et à CentraleSupélec

27 janvier 2016



Points de vue

- 1 L'idéal de la Science Ouverte
 - Archives Ouvertes
 - Données Ouvertes
 - Outils Ouverts
- 2 Publish or Perish
 - Les Archives Ouvertes détournées
 - Pour un autre modèle d'évaluation des enseignants-chercheurs
- 3 La recherche sur projet
 - Mise en compétition des laboratoires
 - Pour une confiance retrouvée
- 4 Microsoft, Apple, Wolfram. . .
 - Les profits contre l'ouverture
 - Une politique volontariste nécessaire



Flash back

La collaboration scientifique au temps du XX^e siècle

Dissémination des résultats par la publication

- Édition, impression et distribution coûteuses
- Temps de communication important
- Processus de relecture aux mêmes échelles de temps
- Publication & collaboration rapide incompatibles

Collaboration par les congrès scientifiques

- Immédiateté de la communication
- L'oral privilégié
- Une organisation relativement inchangée



La collaboration par l'écrit

Rendue possible par la publication rapide sur Internet

ArXiv, HAL...

Les paroles s'envolent, les écrits restent

- Précision de l'écrit
- Exhaustivité de l'écrit
- Permanence de l'écrit
- Complément indispensable de l'oral

La question de la relecture, gage de qualité

- Question centrale
- Ralentissement de la publication, de la collaboration
- Le choix des Archives Ouvertes : pas de relecture
- Permanence de la publication, responsabilisation des auteurs
- Une question de crédibilité ?



Archives-Ouverte et crédibilité

La crédibilité hors le *peer-review*?

Deux propositions

Le modèle YouTube

- Publications permanentes et identifiées à l'auteur
- Commentaires publiés et identifiés
- Responsabilisation des auteurs et des commentateurs

Le modèle Wikipedia

- Publication initiale identifiée à l'auteur
- Modifications et commentaires identifiés
- Amélioration et correction continue des publications



Les données des publications

Premier pas : les données dans les publications

- Publications classiques ou en Archives Ouvertes
- Des **courbes enrichies** par les données brutes
- Certaines revues le permettent déjà

Deuxième pas : les données par elles-mêmes

- Les données brutes sont un matériau de choix
- La technologie permet aujourd'hui de les publier
- Pour que la communauté internationale s'en saisisse
- Certaines revues le permettent déjà
- Quid des Archives Ouvertes ?



Données et formats de données

L'ouverture de la science requiert l'ouverture des formats de données

Pérennité de la publication

- Les publications sont permanentes
- Les données doivent l'être aussi

Les formats de données doivent être pérennes

- Non liés à une entreprise
- Ouverts et documentés
- Libres !



Crédibilité scientifique et répétabilité

Un bon papier donne toutes les clefs pour reproduire les résultats qu'il présente

Répétabilité d'une expérience

- Conditions expérimentales
- **Matériel** de mesure
- **Logiciel** de traitement

Répétabilité d'un calcul

- Hypothèses de départ
- Cheminement logique
- **Logiciel** de calcul

Matériel et logiciel fermés ?

- Matériel de marque : comment fonctionne-t-il ?
- Logiciel fermé : quelle traçabilité pour les traitements ?
- Logiciel fermé : quels sont les calculs réellement faits ?



Ouverture du logiciel

Terreau existant de logiciels ouverts : les logiciels libres et la licence GPL

GNU General Public Licence

- Le code source est ouvert
- Chacun peut l'inspecter, le tester, et le modifier
- La garantie que le traitement des données ne cache rien
- Des calculs automatiques explicités

Une question de preuve mathématique

- Le calcul automatique introduit une faille dans la preuve mathématique
- Seule l'ouverture du logiciel peut la lever



Ouverture du matériel

Le matériel ouvert, un matériel que l'on peut reproduire

Matériel ouvert : des plans publics

- Objectif : un matériel reproductible
- Publication des plans

L'impression 3D ouvre des perspectives

- Un format descriptif 3D ouvert
- La publication des fichiers sous licence GPL
- Un écosystème matériel libre aussi riche que celui du logiciel libre ?



I have a dream

XXII^e siècle

- Les plans du matériel que j'utilise sont publics
- Il est piloté par du logiciel libre
- Mes calculs sont effectués par du logiciel libre
- Je publie mes données et mes résultats
- Mon article est modifiable en ligne
- La communauté mondiale s'empare des mes expérimentations

Pour le meilleur. . .

ou pour le pire

- Amélioration constante de l'article, des calculs et des expériences
- Ou vol des résultats, voire vandalisme de l'article. . .



HAL et l'évaluation HCERES

En ce moment, les validations formelles par l'équipe de HAL sont assez lentes...

Détournement via les notices bibliographiques

- Incitation par les instances à utiliser HAL
- HAL, une base de donnée de notices bibliographiques
 - sans les documents associés
 - de publications déjà publiées dans des revues à comité de lecture
- La Science Ouverte est bien loin

L'instantanéité dévoyée

la preuve en ce moment

- Seules les notices bibliographiques sont mises en lignes
- Le plus souvent pour l'évaluation HCERES



HAL, une base de donnée bibliographique pratique

Seules les publications à comité de lecture comptent

Ouvrir la science prend du temps

- Publier un article dans HAL demande **un peu** d'investissement
- Le temps de l'enseignant-chercheur est compté
- Le temps est utilisé pour les *vraies* publications

Pour l'HCERES

- Une base de donnée bibliographique pratique
- mise à jour avant l'évaluation



Mon expérience personnelle

Un défenseur frustré de la Science Ouverte

- La science ouverte, une nécessité. . . un peu de militantisme
- Je publie la plupart des mes articles en Archives Ouvertes
- Mais seulement une fois qu'ils sont publiés dans des revues à comité de lecture



Des chercheurs évalués à la publication à comité de lecture

Stratégie et activité de recherche toutes entières vers cet objectif

Confidentialité

Une publication en Archive Ouverte pourrait compromettre une *vraie* publication

Gestion du temps

Ouvrir la science prends du temps, mais ne bénéficie pas à la carrière

Ouvrir la science

Adapter le mode d'évaluation des chercheurs



ANR, ERC, Région, Interreg, H2020. . .

Crédits récurrents évanescents

Multiplication des appels à projet

En 15 ans, un budget récurrent par chercheur divisé par 4

- Budget du Laboratoire dominé par les projets publics et les contrats privés
- Une manne financière partagée entre les laboratoires
- Mise en compétition de fait



Un laboratoire est-il une PME ?

Le plus grand secret d'une PME est son savoir-faire, ses données

Laboratoire = PME

- Un budget majoritairement sur contrats
- Une mise en **compétition** de fait
- Ouvrir ses **données** est bien trop risqué

Pire

Parfois

- La mise en **compétition** a percolé à l'intérieur des laboratoires
- Les **données** sont le secret des équipes



Les budgets récurrents et l'ouverture de la science

L'ouverture des données, une question éminemment politique

Pour que les chercheurs retrouvent confiance en leurs financements

- Des financements récurrents indépendants des projets
- Une collaboration sur des bases scientifiques uniquement
- Une éthique scientifique débarrassée des questions pécuniaires

Il faut que l'État retrouve confiance en ses chercheurs

- La Science Ouverte est une nécessité
- Le financement par projet nuit à la Science Ouverte
- Le manque de confiance dans les chercheurs est inefficace à long terme



Ouvrir ses données prends du temps

Au delà des questions politiques, la gestion des priorités

Ouvrir ses données, les difficultés

- Format ouvert et inter-opérable
- Architecture centralisée
- Des problèmes d'informaticien !
- Une vraie motivation est nécessaire

Ouvrir ses données peut-il profiter à soi même ?

- Dynamisation de la recherche via les collaborations ?
- Un risque assumé que cela ne profite qu'à d'autres ?



Mon expérience personnelle

Je n'ai jamais tenté d'ouvrir mes données

- Aucune donnée ouverte à ma connaissance dans mon domaine
- Une concurrence exacerbée sur certains thèmes
- Prudence, mère de sûreté, est sœur de confidentialité



L'ouverture des outils, logiciels et matériels

L'ultime défi de l'ouverture de la science

Des outils ouverts pour des méthodes transparentes

- Traitements de données
- Calculs littéraux
- Démonstrations via l'informatique
- Seule l'ouverture des outils permet la preuve mathématique

Un outil logiciel ouvert ?

- Code source publié et réutilisable
- Logiciel **libre**
- Nécessairement hors du marché du logiciel
- Que vont dire Microsoft, Apple, Google et Wolfram ?



L'écosystème libre freiné par le marché

Seule une politique volontariste pourrait permettre l'ouverture des outils

Des progrès lents mais constants

- Le grand public accède de plus en plus au monde libre
- Certains domaines scientifiques sont en pointe
- D'autres ne l'imaginent même pas !

Une volonté politique ?

- Des alternatives libres plutôt que des licences logicielles ?
- Des sessions de sensibilisation ?
- Des formations ?
- Une incitation de la direction
- Voire un séminaire sur la Science Ouverte utilisant des logiciels libres pour son matériel de présentation ?



L'ouverture du matériel ?

Elle semble bien lointaine...

Le futur de l'impression 3D ?



Mon expérience personnelle

J'utilise (quasi) uniquement du logiciel libre

- Ce diaporama est préparé avec un logiciel libre
- Tous mes calculs sont faits avec du libre
- Et nous sommes quelques uns. . .
- C'est un argument dans certains de mes papiers

