

## Conseil scientifique de l'Institut des sciences de l'information et de leurs interactions (INS2I)

### Recommandations suite au séminaire thématique "Science ouverte" (« Pratiques de publication »)

Le Conseil Scientifique de l'Institut des Sciences de l'Information et de leurs Interactions (INS2I) du CNRS a organisé le 6 juin 2021 un séminaire thématique "Science ouverte". Ce séminaire se place dans le contexte d'une mobilisation importante pour la science ouverte au CNRS et plus généralement en France, avec notamment l'annonce récente du deuxième Plan national pour la science ouverte.

Le sujet étant vaste et riche, le choix a été fait de se concentrer sur les publications et les logiciels, mais pas sur les données. Trois intervenants ont été invités par le CSI à s'exprimer sur ces différents aspects :

- Daniel Le Berre, Professeur Université d'Artois, Lens : « *Enquête sur les pratiques de publication de la SIF : focus sur la science ouverte* »,
- Roberto Di Cosmo, Professeur Université Paris Diderot, Paris : « *Construire le pilier logiciel de la Science Ouverte : opportunités et enjeux* »,
- Denis Bourguet, Directeur de recherche INRAE à Montpellier, et Thomas Guillemaud, Directeur de recherche INRAE à Sophia-Antipolis : « *Modèles de publication et l'expérience Peer Community In* ».

#### ***Séminaire de Daniel Le Berre : « Enquête sur les pratiques de publication de la SIF : focus sur la science ouverte ».***

Le conseil des associations de la société informatique de France (SIF) a élaboré **un référentiel des pratiques de publication en informatique** dans le but cartographier les pratiques de publication des enseignants-chercheurs et chercheurs en informatique en France. Un sondage eu lieu entre octobre 2018 et février 2019. Le document qui a vocation à évoluer au cours du temps, se décline en deux parties. Une analyse globale des pratiques de publication de la discipline informatique et une analyse plus détaillée par thématique opérée par les associations thématiques ou les GDR. Concernant l'archivage ouvert, la tendance générale est de déposer ses publications une fois l'article accepté. En algorithmique, recherche opérationnelle et en informatique théorique, l'utilisation des archives ouvertes se fait souvent avant la soumission, ou au moment de la soumission, ce que l'on ne retrouve pas dans les autres thématiques. Le choix de la revue ne dépend pas du fait que l'accès aux articles soit ouvert ou pas.

La question des publications en accès ouvert gratuit ou payant, et des pratiques vertueuses (ou pas) de la science ouverte, est développée dans le séminaire de Denis Bourguet et Thomas Guillemaud.

***Séminaire de Roberto Di Cosmo : « Construire le pilier logiciel de la Science Ouverte : opportunités et enjeux ».***

Le logiciel est désormais partout, dans toutes les disciplines si bien qu'on peut avoir l'impression que les informaticiens produisent moins de logiciels que les autres. L'élément essentiel est le code source qui encapsule toute la connaissance produite par ailleurs. Dans son livre paru en 1985, Harold Abelson énonçait que « les programmes doivent être écrits de sorte que les gens puissent les lire, et accessoirement pour que les machines les exécutent ». Dans les années 1980, très peu de codes source étaient disponibles, ce qui n'est plus le cas de nos jours. Les trois piliers de la science ouverte sont les archives ouvertes, les données ouvertes, les codes sources ouverts. La prise de conscience tardive mais grandissante de l'importance du pilier logiciel de la science ouverte pose des questions majeures, notamment pour ce qui concerne la reproductibilité des résultats et la reconnaissance des productions logicielles dans la recherche. L'infrastructure mutualisée « Software Heritage », laquelle a pour objectif de collecter, préserver et partager tous les codes sources publiquement disponibles, apporte un élément de réponse important pour le premier de ces enjeux, et il est souhaitable d'accélérer son adoption par notre communauté. Une réflexion en profondeur doit être engagée pour le deuxième enjeu, en concertation avec d'autres disciplines, notamment dans le cadre du Comité d'Orientation de la Science Ouverte (CoSO) au niveau national, des sociétés savantes et de l'EOSC à niveau international.

Un enjeu important concerne l'archivage à long terme : comment garantir que l'environnement logiciel de notre époque soit décrypté dans 100 ou 500 ans ? Préserver l'environnement nécessaire pour exécuter des logiciels est un défi majeur qui n'a pas encore de réponse définitive, il sera cependant toujours possible de lire et de comprendre les codes sources, d'où l'importance d'archiver ceux-ci.

***Séminaire de Denis Bourguet et Thomas Guillemaud : « Modèles de publication et l'expérience Peer Community In ».***

Les chercheurs publient pour que leurs travaux soient rendus publics. Quelques grands éditeurs publient 50% des publications, cela représente une industrie extrêmement lucrative, par exemple la publication d'un article dans la revue Nature coûte 9500 €. L'initiative Plan S vise à garantir un accès libre, complet et immédiat aux publications scientifiques évaluées par les pairs et financées par des institutions publiques ou privées. De nouveaux modes de publications se déploient, par exemple le modèle F1000, racheté par l'éditeur Taylor & Francis, reposant sur les archives ouvertes et le fait que les rapports soient publics. Le projet « Peer Community In (PCI) » a pour objectif de créer des communautés de chercheurs qui évaluent et recommandent les articles de leur champ scientifique déposés dans des archives ouvertes, par exemple PCI Ecology, PCI Evolutionary Biology, PCI Genomics. Ce mode de publication est gratuit et les décisions éditoriales, les recommandations ainsi que les rapports sont rendus publics. Au comité national, les sections 29, 30 et 52 et le conseil scientifique de l'INEE soutiennent ce projet.

Le projet PCI, encore assez méconnu dans la communauté informatique, suscite de nombreuses interrogations : y a-t-il un éditeur en chef qui définirait une ligne éditoriale ? les auteurs et les rapporteurs peuvent-ils décider d'être anonymes ? quel est le niveau d'internationalisation du projet ?

Denis Bourguet et Thomas Guillemaud apportent les réponses suivantes : le « managing board » fait office de comité éditorial. Les rapporteurs choisissent de signer ou non les rapports, il y a de très bons arguments pour ou contre l'anonymat des rapporteurs, on observe qu'environ 40% des rapporteurs apposent leur signature. Les éditeurs associés et les référés certifient qu'ils n'ont de conflit d'intérêt ni avec les auteurs

ni avec le contenu de l'article. Concernant l'internationalisation, les fondateurs sont français mais dans chaque communauté au moins 80% des éditeurs associés sont étrangers. En revanche, il y (encore) davantage de soumissions de la part d'équipes françaises que d'équipes internationales.

Il est fait remarquer qu'il serait plus judicieux de s'intéresser à la manière de mutualiser les coûts, notamment des infrastructures d'archives ouvertes, plutôt que de focaliser notre attention sur le prix des publications. Il faudra ensuite s'atteler à sortir les revues prestigieuses du giron des éditeurs.

Une observation générale est que les institutions ne doivent absolument pas pré-payer les frais de publication, car cela ne résoudra pas le problème du déséquilibre entre le coût véritable et le prix des publications.

## Recommandations du CSI

### À la communauté :

- Adopter les bonnes pratiques de citation et référencement des logiciels produits ou utilisés.
- Pérenniser la production logicielle de la recherche sur Software Heritage. – Archiver systématiquement les codes sources des logiciels de recherche et les référencer en utilisant les identifiants pérennes **SWHID**. – Promouvoir ces pratiques dans nos communautés internationales, en s'assurant que l'archivage et le référencement dans Software Heritage soit bien pris en compte, par exemple pour obtenir le badge "Artifact Available", en particulier par les Artifact Evaluation Committees.

### Aux sections du Comité national :

- Valoriser toute la production scientifique dans l'évaluation de l'activité des chercheurs et ingénieurs (notamment logicielle, voir, par exemple, la note de la CE Inria [<https://hal.inria.fr/hal-03110723>]), sans oublier la production de données).
- Valoriser la publication des résultats en version complète, ne serait-ce que comme preprint en archive ouverte, même dans les domaines ne publiant classiquement, qu'en conférences et pas en revues.
- Face à la profusion envahissante de revues prédatrices en open access, réaffirmer l'importance de privilégier les publications dans des canaux garantissant un processus de relecture impartial et reposant sur un comité éditorial qualifié.
- Encourager les chercheurs.euses à passer par des journaux ou plateformes, comme PCI par exemple, qui publient les rapports d'évaluation, les décisions des éditeurs et les réponses des auteurs aux remarques des évaluateurs. Les inciter à déposer leur données brutes, codes et scripts d'analyse sur des archives ouvertes, le développement de ces pratiques permettant d'assurer une plus grande fiabilité et reproductibilité des résultats.

### À l'institut :

- Soutenir les initiatives visant à basculer les conférences et revues prestigieuses dans un système Open Access de coût raisonnable : LIPIcs, épi-journaux, « Peer Community In (PCI) », etc.
- Rappeler l'obligation faite aux chercheurs.euses de publier systématiquement en accès libre leurs publications scientifiques évaluées par les pairs et soutenues pour plus de 50% par des financements publics, de manière à en garantir un accès complet et immédiat (ou après une période d'embargo).
- Résister à la dérive vers le fonctionnement auteur-payeur (que ce soit sous la forme laboratoire-payeur, organisme-payeur, ou autre) que les grands éditeurs commerciaux tentent d'imposer. Encourager les chercheurs.euses à publier dans des journaux diamants (i.e., des journaux libre d'accès et sans frais pour les auteurs), de manière à éviter le paiement de frais de publication (APC) et limiter ainsi l'impact financier sur nos institutions.

Gilles SASSATELLI  
Président du Conseil scientifique de l'INS2I



Recommandation adoptée le 7 décembre 2021

Vote : 16 oui, 0 non, 0 abstention

### Destinataires :

- M. Ali CHARARA, directeur de l'INS2I du CNRS
- Mme Sylvie ROUSSET, directrice des données ouvertes de la recherche (DDOR) au CNRS
- Directeurs des unités de l'INS2I
- M. Pierre SENELLART, président de la Section 6
- Mme Inbar FIJALKOW, présidente de la Section 7
- Présidentes et présidents des Sections et CID du Comité National

### et copie à

- M. Antoine PETIT, président directeur général du CNRS
- M. Alain SCHUHL, directeur général délégué à la science du CNRS
- Mme Dorothee BERTHOMIEU, présidente du Conseil scientifique du CNRS
- Mme Claudine GILBERT, présidente du CS INP ; M. Olivier DRAPIER, président du CS IN2P3 ; M. Serge SIMOENS, président du CS INSIS ; Mme Beatrice MARTICORENA, présidente du CS INSU ; M. Olivier SANDRE, président du CS INC ; Mme Nathalie VIENNE-GUERRIN, présidente du CS INSHS ; M. Yaël GROSJEAN, président du CS INSB ; M. Remi CARLES, président du CS INSMI ; Mme Patricia GIBERT, présidente du CS INEE