

Les prépublications ouvrent-elles la voie à une science en temps réel ?

Rédigé par Andrea Chiarelli à l'intention de Knowledge Exchange (Research Consulting)

La traduction en français a été assurée par Alexandre Clément, traducteur-stagiaire dans le service traduction de l'Inist-CNRS.

Depuis son lancement en 1991, Arxiv, le plus connu des serveurs de prépublication, a reçu plus d'1,5 million de prépublications. BioRxiv, une plate-forme équivalente en plein essor dans le domaine des sciences de la vie qui a vu le jour en 2013, a recueilli quant à elle plus de 40 000 soumissions, soit les travaux de plus de 160 000 chercheurs provenant de plus de 100 pays différents. Incontestablement, les prépublications prennent de l'ampleur dans le paysage des communications savantes, mais qu'en est-il de leur avenir ?

Depuis 2018, Knowledge Exchange étudie cet écosystème passionnant en constante évolution. À sa demande, la société Research Consulting a mené une étude sur l'état et l'impact des prépublications. Knowledge Exchange a récemment publié un diaporama résumant les résultats obtenus en début de projet. Ces derniers se déclinent ainsi : une première analyse de la littérature, suivie de 38 entretiens avec des bailleurs de fonds de la recherche, des responsables d'organismes de recherche et de serveurs de prépublications, ainsi que d'autres chercheurs et fournisseurs de services. Dans ce billet, je vais vous faire part de quelques-unes des principales conclusions auxquelles nous sommes parvenus et vous présenter ce qui nous attend par la suite. Dans le cadre cette discussion, le terme prépublication est défini comme suit :

Version d'un travail de recherche, généralement avant évaluation par les pairs et publication dans une revue.

Manque de temps ? Faites défiler vers le bas pour accéder à un résumé contenant les principales idées à partager !

Là où tout a commencé

Les physiciens du Laboratoire national de Los Alamos ont mis en place arXiv en 1991. En économie, la diffusion électronique des working papers, aussi appelés « littérature grise » (leur propre version des prépublications) remonte à 1993, avec le lancement du projet Working Papers in Economics (WoPEc).

« La première prépublication que j'ai écrite date d'il y a 50 ans. Quand j'ai écrit mon premier article, il a été dactylographié et a fait l'objet de copies carbonées avant d'être envoyé aux personnes susceptibles d'être intéressées »

Manifestement, les prépublications ne datent pas d'hier : elles existent depuis un demi-siècle, d'abord sous forme de copie carbone, mais c'est seulement au cours des 20 dernières années qu'elles ont véritablement acquis leur force disruptive. Leur essor significatif a été accompagné par les développements technologiques et, en particulier, par la diffusion généralisée des bases de données et des archives ouvertes. Si ces dernières ont presque toujours été créées pour répondre à un besoin spécifique par le passé, il est maintenant relativement facile de se doter d'une base de données en ligne fonctionnelle.

Aujourd'hui, un ensemble d'autres disciplines suivent la voie tracée par les communautés des mathématiques, de la physique, et de l'économie, et l'un des facteurs clés de cette évolution réside dans la facilité de partage de contenu en ligne. BioRxiv (lancé en 2013), PsyArXiv (2016) et ChemRxiv (2017) ont constitué le point de départ de cette étude, mais une vaste gamme de serveurs de prépublication est désormais disponible, des nouveaux apparaissent chaque année, exploitant une palette de solutions techniques. Fait intéressant, nous avons constaté que le débat sur la valeur des prépublications semble de nouveau prendre de l'ampleur dans toutes les disciplines, les mérites de cette approche (re)naissante de la diffusion des savoirs étant fonction des usages disciplinaires et des méthodes existantes en matière de partage d'information.

Médias sociaux : les catalyseurs

L'un des résultats surprenant mais néanmoins révélateur de cette étude réside dans la manière dont Twitter a favorisé l'adoption des prépublications dans de nombreuses disciplines. Si Twitter est, par nature, un média ouvert et tout public, les communautés scientifiques « fermées » y ont recours pour discuter de leurs travaux, des dernières découvertes et tendances. Twitter permet en particulier aux utilisateurs de :

- découvrir des prépublications publiées par des pairs ;
- suivre des bots Twitter qui publient des prépublications depuis des serveurs dédiés ;
- partager leurs propres prépublications ;
- émettre et recevoir des commentaires ;
- faire la promotion de leurs travaux et communiquer avec les rédacteurs en chef des revues à fort impact après avoir connu un pic de popularité sur les réseaux sociaux.

En l'absence de Twitter, bon nombre des participants à notre étude n'auraient même pas été au courant de l'existence des prépublications. Nous avons constaté à plusieurs reprises que les participants n'avaient jamais entendu parler de prépublications avant qu'un pair n'en fasse mention dans un tweet, ce qui a généralement amené les chercheurs à s'intéresser, creuser le concept et découvrir davantage de prépublications.

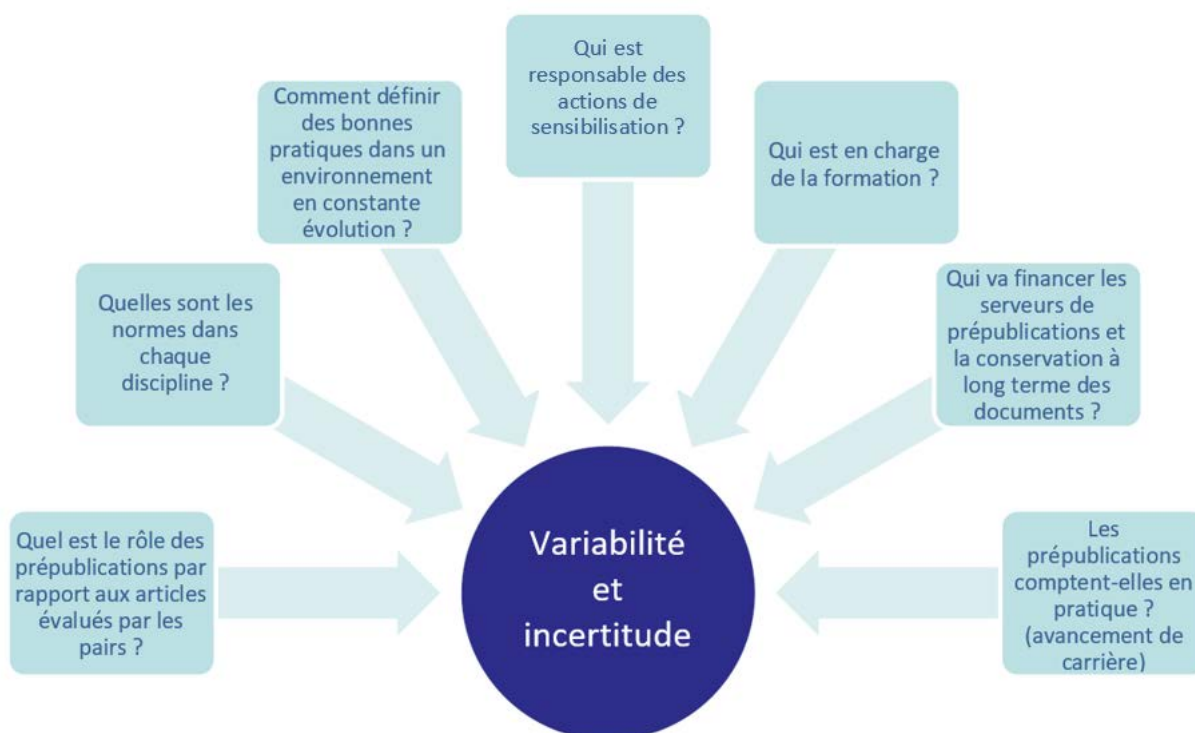
« Quelqu'un vient de tweeter sa récente publication, je clique et c'est une prépublication. Même si je ne suis pas nécessairement à la recherche de prépublications, je vais le savoir bien plus vite via Twitter. » - Chercheur

Qui est en charge des prépublications...

Au cours de nos recherches, nous avons pris conscience que l'écosystème des prépublications se caractérise par une incertitude et une variabilité fortes. De nombreuses questions en suspens (voir figure 1) n'ont pas encore trouvé de réponse à l'heure actuelle : elles portent notamment sur les rôles et les responsabilités de chacun en matière de promotion, de formation et de soutien, les effets des prépublications sur la progression de carrière et la relation entre prépublications et articles de revues évalués par des pairs. Il suffit de regarder ces exemples pour se rendre compte de la multitude d'acteurs à l'œuvre dans le monde des prépublications.

Pour compliquer les choses, un certain nombre d'acteurs et d'initiatives supplémentaires gravitent autour des prépublications et serveurs de prépublications. Il s'agit tout particulièrement des épi-revues (Peer Community in, preLights), des clubs de revues de prépublications (PREreview) et des infrastructures techniques qui permettent une indexation efficace des prépublications (Crossref, Web of Science).

Figure 1 - Complexité de l'écosystème des prépublications



Mais la question la plus urgente est peut-être de savoir qui est en charge de *poster* les prépublications ? Si toutes les autres interrogations peuvent être résolues au moyen de consultations, de discussions et de développements ultérieurs dans le paysage de la communication savante, cette question prend racine au cœur du problème :

- Une approche centrée sur le chercheur, où celui-ci poste lui-même ses prépublications, favoriserait la transparence et la liberté de choisir, ainsi qu'une certaine indépendance par rapport aux éditeurs universitaires dans un marché de l'édition toujours plus concentré.
- Une conception axée sur l'éditeur se traduirait probablement par des flux de travail plus fluides et une plus grande homogénéité en termes de prépublications, un éditeur pourrait poster des

prépublications pour la totalité des articles qui lui ont été soumis, au lieu de voir quelques auteurs seulement assumer cette tâche.

Aujourd'hui, dans la plupart des cas, les prépublications sont postées de manière individuelle par des chercheurs travaillant au sein de communautés disciplinaires engagées. En outre, les serveurs de prépublication sont en phase d'expérimentation avec la technologie et les flux de travail, pour ne citer qu'eux. Une vision focalisée sur l'éditeur n'est donc pas encore d'actualité en raison de l'incertitude élevée qui caractérise le monde des prépublications. À ce jour, certaines solutions comme [F1000 Research](#) ou [PeerJ](#) ont intégré les prépublications à leurs flux de publication, mais les possibilités de diffusion de ces solutions à grande échelle demeurent floues.

Quelle que soit l'approche choisie, il est essentiel de clarifier, sinon d'améliorer la communication quant à la question suivante : les prépublications postées avant la soumission à une revue pourraient-elles entraîner un refus d'évaluation par les pairs pour certains articles ? De nombreuses revues acceptent les prépublications et les informations à ce sujet sont bien souvent disponibles auprès des éditeurs et sur [Wikipédia](#) (qui met également en évidence les éditeurs qui n'acceptent *pas* cette pratique). Cependant, nous avons constaté que, dans certains cas, les auteurs restent tout de même méfiants.

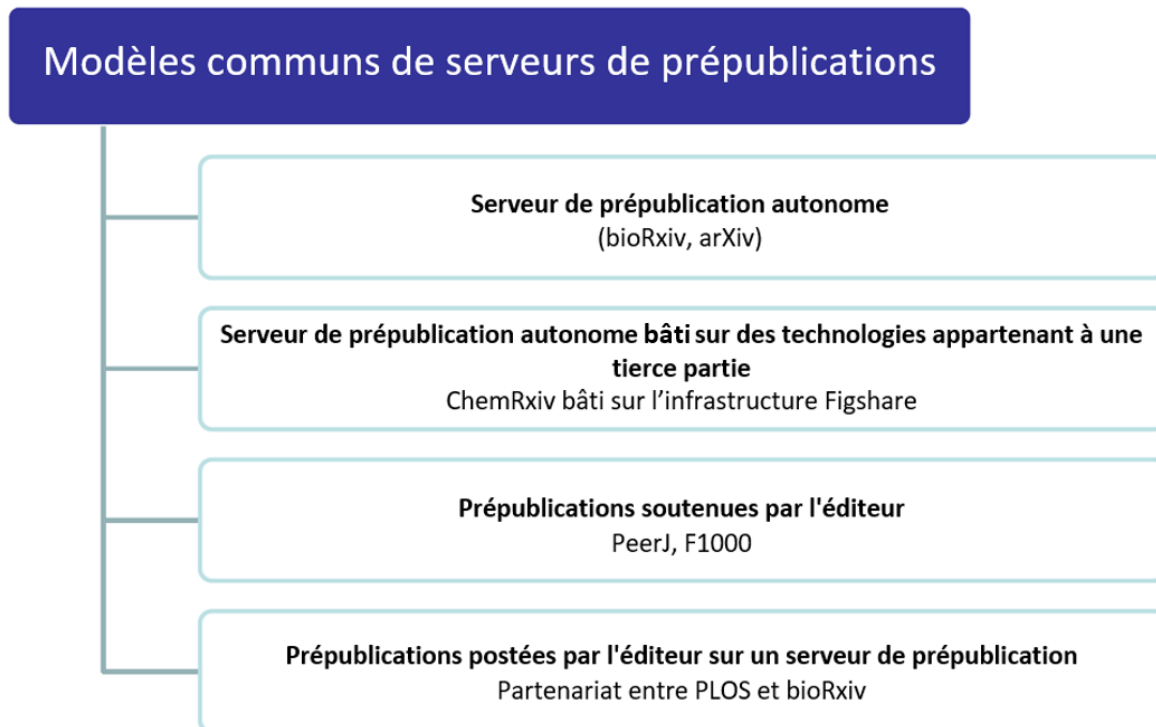
... et qui paye l'addition ?

Pour ce qui est du lancement et de la gestion des serveurs de prépublications, nos recherches ont fait apparaître quatre méthodes différentes (voir figure 2 ci-dessous) :

- Quelques serveurs de prépublication autonomes, tels que bioRxiv et arXiv, ont développé leurs propres solutions techniques et fonctionnent indépendamment des autres acteurs ou solutions technologiques.
- D'autres serveurs de prépublication autonomes, comme ChemRxiv, fonctionnent à l'aide d'une infrastructure technique et d'une technologie appartenant à une tierce partie (dans le cas de ChemRxiv, on note l'utilisation de l'infrastructure Figshare).
- Certains éditeurs comme PeerJ et F1000 Research intègrent les prépublications à leurs flux de publication. De plus, PeerJ met également à disposition sa plate-forme PeerJ Preprints, qui peut être utilisée de manière autonome (c'est-à-dire sans évaluation par des pairs).
- D'autres éditeurs tels que PLOS ont commencé à poster des prépublications d'articles ayant fait l'objet d'une soumission à une revue sur des serveurs de prépublication (dans le cas de PLOS, les prépublications sont postées sur bioRxiv).

Notez que, sous les deux rubriques « autonomes » figurent également des services tels que [Preprints.org](#) et [SSRN](#), qui sont exploités respectivement par MDPI et Elsevier. Ces derniers n'entrent pas dans la catégorie « soutenu par l'éditeur », ce label étant utilisé dans le cadre d'un processus plus global pour décrire les prépublications pouvant être postées et évaluées par les pairs via la même plate-forme.

Figure 2 - Modèles pour les serveurs de prépublication



Le choix du modèle est manifestement lié à l'évolution du monde des prépublications, se dirige-t-on vers une approche centrée sur le chercheur ou sur l'éditeur ? Les participants à notre projet ont mis en avant un désir profond de modèles à but non lucratif et autonomes à l'égard des éditeurs (p. ex. par le biais de consortiums), mais nous observons que la viabilité de cette approche dépend en grande partie de la disponibilité de financements provenant d'autres sources.

Je pense qu'il est très important que les archives fonctionnent comme une entité indépendante, sans but lucratif. - Bailleur de fonds de la recherche

L'avenir des prépublications

Pour en revenir au titre de ce billet et compte tenu de notre discussion ci-dessus, il semble que les prépublications nous rapprochent bien un peu plus de la science en temps réel. Toutefois, si elles devaient acquérir un poids important dans un plus grand nombre de domaines de recherche, trois interrogations essentielles devraient être posées, à savoir :

1. Comment intégrer les prépublications dans les flux de recherche et de publication existants ?
2. Les prépublications doivent-elles être conservées à long terme, et dans quelle mesure ?
3. La création de versions XML des prépublications, accessibles numériquement, peut-elle présenter une utilité, et est-ce viable financièrement ?

Une considération purement financière vient en effet compléter les questions ci-dessus : quel est le coût d'opportunité des prépublications et des serveurs de prépublication ? Il s'agit de se demander si l'argent dépensé pour les prépublications et les serveurs de prépublication peut produire de meilleurs résultats s'il est investi ailleurs, dans le cadre d'autres initiatives de science ouverte ou d'accès ouvert. Par exemple, les serveurs de prépublication commencent à envisager des solutions de conservation et des flux de travail sur le long terme : si une prépublication a été publiée dans sa version finale en tant qu'article en accès ouvert, existe-t-il une raison concrète de conserver la version antérieure ?

La collaboration mentionnée ci-dessus entre PLOS et bioRxiv soulève également la question suivante : un plus grand nombre d'éditeurs pourraient-ils soumettre par défaut les articles reçus par les revues pour examen à des serveurs de prépublication ? La réponse à cette question dépend probablement des usages disciplinaires et des avantages potentiels dont pourraient bénéficier les diverses parties prenantes impliquées.

Nos cinq points clés sur les prépublications

1. **Twitter joue un rôle essentiel dans la diffusion des prépublications.** Même si les prépublications ne datent pas d'hier, les technologies offrent aujourd'hui à la communication savante de nouvelles manières d'exploiter cette approche en plein essor. À l'heure actuelle, les lecteurs prennent souvent connaissance des prépublications grâce à Twitter, ce qui est utile à la fois pour leur recherche d'informations et parce qu'ils ont la possibilité de fournir un retour aux auteurs.
2. **Une solution unique qui fonctionnerait pour toutes les disciplines n'est probablement pas envisageable pour le moment :** les prépublications évoluent dans un environnement complexe en raison de la grande diversité des approches techniques, des modèles économiques, des rôles et responsabilités de chacun, des usages disciplinaires, etc.
3. **Qui devrait être en charge de poster les prépublications : les chercheurs ou les éditeurs ?** Cette question renvoie aux modèles sous-jacents des serveurs de prépublication, ces derniers étant généralement autonomes (p. ex. à but non lucratif, subventionnés), ou soutenus par l'éditeur. Les serveurs de prépublication autonomes auraient tendance à se reposer sur les chercheurs pour poster eux-mêmes leurs prépublications (cela nécessite un haut niveau d'implication), tandis qu'une approche axée sur l'éditeur simplifie les flux de travail mais alimente les préoccupations concernant la concentration du marché dans l'écosystème de l'édition universitaire.
4. **Certains chercheurs craignent que les revues ne rejettent leurs soumissions si une prépublication a déjà été publiée :** c'est un obstacle majeur à la diffusion des serveurs de prépublication. Les informations sont souvent disponibles sur les sites Web des éditeurs et sur Wikipédia, mais en pratique, les préoccupations demeurent largement répandues.
5. **Des modèles économiques adaptés doivent être élaborés** si l'on veut que les prépublications portent leurs fruits à long terme. Dans ce processus, il sera important d'évaluer le coût d'opportunité des serveurs de prépublication et de tous les flux de travail connexes (notamment la conservation à long terme). Les mêmes fonds pourraient-ils être investis différemment et avoir un meilleur impact sur la communication savante et la science ouverte ?

Poursuivre les recherches

Les résultats de ces recherches seront présentés à l'[Open Repositories Conference](#) (Hambourg) le 12 juin prochain. Nous allons faire participer l'auditoire et voir quelles semblent être ses principales observations et préoccupations. Knowledge Exchange interviendra également lors de la [conférence LIBER](#) à Dublin quelques semaines plus tard et visera à intégrer les commentaires de la communauté dans un rapport résumant ses conclusions et leurs implications pour les différents publics. La publication du rapport final est attendue à l'automne, l'objectif est de profiter de conférences sur le sujet telles que COASP et FORCE11 pour échanger sur ces questions.

Parallèlement, nous préparons un article évalué par des pairs contenant une analyse plus approfondie de nos résultats, ainsi qu'un récapitulatif de la façon dont nous avons analysé nos résultats de recherche. Nous avons l'intention de poster ce document en prépublication au cours des prochains mois dans un premier temps, alors suivez Knowledge Exchange ([@knowexchange](#)) et Research Consulting ([@rschconsulting](#)) sur Twitter pour ne rien manquer des nouveautés en matière de prépublications !

Si vous souhaitez commenter nos travaux, veuillez contacter Karin van Grieken de Knowledge Exchange par courriel à Karin.vanGrieken@surfmarket.nl ou adressez-vous à moi : andrea@research-consulting.com et nous nous ferons un plaisir de vous répondre.

Veillez créer un widget/bouton

Partager nos diaporamas :

#Preprints = 'science in real time'? Découvrez ce que pensent vraiment les bailleurs de fonds, les chercheurs et les fournisseurs de services au sujet des prépublications. Les diaporamas des études de [@knowexchange](#) et [@rschconsulting](#) sont maintenant disponibles en anglais sur <http://bit.ly/Keprintsides>