

DATAACC, UNE PLATEFORME POUR LES CHERCHEURS EN PHYSIQUE ET CHIMIE : METTRE EN LUMIÈRE LES RÉSULTATS NÉGATIFS

Entretien avec Marie-Emilia Herbet et Jérémie Léonard

Marie-Emilia Herbet

Conservatrice des bibliothèques et chef de projet en charge du développement de Dataacc

Jérémie Léonard

Ingénieur d'études chargé d'appui au projet Dataacc

Pôles documentaires de référence en chimie et en physique, les bibliothèques de Lyon-1 et de l'Université Grenoble Alpes se sont associées pour développer Dataacc, une plateforme soutenue par CollEx-Persée. Dataacc est conçue comme un service d'accompagnement des chercheurs en physique et chimie. Marie-Emilia Herbet et Jérémie Léonard, avec la contribution de Lucie Albaret, présentent au BBF le projet et ses dernières évolutions, centrées autour des résultats négatifs.

*

BBF : Qu'est-ce que Dataacc ?

M.-E. Herbet et Jérémie L. : Dataacc est un dispositif d'accompagnement des chercheurs¹ dans la gestion de leurs données de recherche en chimie et en physique financé par le programme CollEx-Persée. Le projet a bénéficié de l'appui de deux appels à projet CollEx-Persée pour lesquels il a été lauréat en 2019 puis en 2020, facilitant ainsi l'amorçage du service ainsi que sa consolidation.

La plateforme Dataacc.org, adossée au projet éponyme, a été lancée le 30 avril 2020.

Elle regroupe des articles, des ressources, des outils, des formations concernant la gestion des données appliquées de manière très concrète à la physique et la chimie. Elle propose notamment une sélection d'entrepôts pour chercher ou déposer des données. La plateforme est toujours en cours d'enrichissement. Une newsletter mensuelle

1 <https://www.dataacc.org/>

permet aux 300 abonnés d'avoir un point sur les actualités concernant la gestion des données et la science ouverte dans les deux disciplines.

Depuis lors, trois champs d'action ont été particulièrement investis : les cahiers de laboratoire électroniques (séminaires d'information, ateliers de formations et mise à disposition de logiciels), l'accompagnement à la rédaction de plans de gestion de données et la sensibilisation sur la diffusion des résultats négatifs².

L'approche disciplinaire découle de l'ancien réseau des Cadist³ dont les bibliothèques de Lyon-1 et Grenoble faisaient partie en étant respectivement responsables des collections d'excellence en chimie et en physique. Outre les deux BU partenaires, le projet s'articule autour d'une communauté de chercheurs intéressés par les thématiques couvertes. Thierry Billard, directeur de recherche en chimie (UMR CNRS 5246) appuie activement le projet, en tant que coordinateur scientifique. Plus récemment, un rapprochement a été opéré avec les chercheurs membres de l'association Science Unchained, dans le cadre du développement d'une plateforme initiée par l'association et spécialisée dans la publication d'expériences infructueuses, baptisée Gaffex⁴.

La conduite du projet Datacc mobilise une palette assez large de compétences : capacités rédactionnelles, connaissances en informatique, veille informationnelle, relations publiques et communication avec les chercheurs. La complémentarité nécessaire s'est construite entre les conservateurs de bibliothèques et deux ingénieurs d'études respectivement diplômés en informatique et physique.

Comment est née l'idée de travailler sur les résultats négatifs dans le cadre de Datacc ?

L'état des lieux mené sur l'utilisation des cahiers de laboratoire électroniques nous a naturellement conduits à réfléchir à l'enjeu des résultats négatifs. À plusieurs reprises, les développeurs de ce type de logiciels nous ont signifié l'avantage qu'auraient ce type d'outils pour publier « en un clic » le mode opératoire des expériences non concluantes méritant d'être diffusées, mais pour lesquelles les chercheurs ont très peu de temps à consacrer.

Par ailleurs, depuis le début du projet, nous avons opté pour un positionnement original en investissant des thématiques peu visibles en bibliothèque mais permettant d'agir là où des besoins existent sans qu'ils ne soient comblés par l'offre actuelle (les économistes parleraient de « faille de marché »). Il nous a donc semblé cohérent de dresser cette passerelle entre les cahiers de laboratoire (remplis d'expériences qui n'ont pas

2 [NDLR] Pour en savoir plus sur les résultats négatifs : Sabrina GRANGER. « Les oubliettes de la recherche ou "l'effet tiroir" ("file drawer effect") ». Faire passer les résultats négatifs de l'ombre à la lumière », *Bulletin des bibliothèques de France*. 25 mai 2021. En ligne : https://bbf.enssib.fr/matieres-a-penser/les-oubliettes-de-la-recherche-ou-l-effet-tiroir-file-drawer-effect-faire-passer-les-resultats-negatifs-de-l-ombre-a-la-lumiere_70065 [consulté le 15 novembre 2021].

3 Centres d'acquisition et de diffusion de l'information scientifique et technique.

4 [NDLR] Pour en savoir plus sur Gaffex : « Réhabiliter les impasses scientifiques : Gaffex, un projet de plateforme dédiée aux résultats négatifs – Entretien avec Gilmery Gallon et Estelle Rascol », in *2020. Effets et conséquences de la crise sanitaire*, Villeurbanne : Bulletin des bibliothèques de France, 2021, p. 110-115 (coll. L'Année des bibliothèques/BBF).

marché) et la diffusion de ces résultats dits « négatifs », qui ne trouvent pas facilement preneurs dans le système éditorial scientifique actuel.

Pouvez-vous nous fournir quelques éléments sur la méthodologie de l'enquête menée ?

L'enquête était destinée en priorité aux chercheurs et doctorants issus des domaines de la chimie, physique et disciplines connexes (sciences de l'ingénieur et sciences de l'environnement). Ce choix découle de la volonté de combler le déficit de connaissances sur le rapport entretenu par ces communautés vis-à-vis des résultats négatifs, quand la littérature scientifique actuellement disponible cible plus volontiers les chercheurs en biologie et sciences médicales. Nous avons opté pour des questions à choix multiples, avec la possibilité d'ajouter un commentaire. Cela permet à ceux qui disposent de peu de temps de ne pas passer plus d'une dizaine de minutes sur le questionnaire, et à ceux qui le souhaitent de développer un peu plus. De plus, les questions ouvertes sont généralement compliquées à traiter : nous l'avons fait manuellement afin d'extraire les verbatims des chercheurs susceptibles d'enrichir l'analyse globale.

Le questionnaire a été construit par l'équipe projet de Datacc, avant d'être soumis pour relecture et commentaires à une conservatrice des bibliothèques d'un autre établissement et à plusieurs chercheurs. La diffusion a été facilitée par l'envoi du questionnaire auprès des Groupements de recherche du CNRS des disciplines concernées, ainsi que par le biais des listes de diffusion chercheurs de nos établissements. Nous avons aussi bénéficié du relais des référents science ouverte des différentes universités.

On constate que les chercheurs interrogés ne sont pas rétifs à la diffusion de leurs résultats négatifs, bien au contraire. Les obstacles principaux identifiés sont : « Le manque de temps et la complexité de restitution et d'interprétation des résultats négatifs sont identifiés comme les freins principaux à leur diffusion. » Quelles sont les pistes de réflexion de l'équipe de Datacc pour apporter des réponses ?

La diffusion de résultats dits « négatifs » fait l'objet de multiples craintes. Outre le fait de s'attirer les foudres des pairs si le résultat en question contredit ce qui aurait été affirmé par un chercheur faisant autorité, leur publication soulève de multiples enjeux, dont la gestion de l'incertitude. Est-on bien certain que l'hypothèse ne peut être validée ? La méthodologie suivie était-elle bien correcte ? Ne risque-t-on pas de publier de « faux négatifs » ? Comment publier un tel résultat si l'on ignore la cause de l'échec ? Cette divulgation ne risque-t-elle pas de faire gagner du temps aux concurrents qui parviendraient, eux, à aboutir à une solution ?

Autant d'interrogations qui gravitent autour de la diffusion de ces résultats « atypiques » dans la littérature scientifique et pourtant si « typiques » du travail scientifique, dont les avancées sont l'aboutissement d'obstacles rencontrés puis surmontés.

Plutôt que de condamner l'hypothèse de la publication des résultats négatifs à la lumière des facteurs exposés ci-dessus, nous avons fait le pari du contraire à travers deux objectifs de court et de long terme.

L'objectif de court terme vise à faciliter le déploiement d'une plateforme de diffusion des résultats négatifs, avec possibilité de commenter les résultats exposés. Ce dialogue ouvert entre pairs semble primordial afin de désamorcer tout problème de méthode qui réduirait le résultat négatif à une « erreur ». Il vise également à participer au débat scientifique entre pairs de la même manière que pour les résultats dits « positifs ».

L'objectif de long terme est d'offrir aux chercheurs un système d'information cohérent en construisant une interopérabilité entre les solutions de cahiers de laboratoire électroniques proposées et la plateforme de dépôt en ligne des résultats négatifs.

Quelles ont été les réactions des chercheurs impliqués dans le projet à la lecture des résultats de l'enquête ?

Nous avons reçu plusieurs témoignages de chercheurs exprimant leur intérêt pour la thématique des résultats négatifs lors de la mise en ligne du pré-print. Certains ont jugé encourageant le plébiscite affiché en faveur de la publication de ce type de résultats. D'autres ont indiqué que les chercheurs ayant accepté de répondre étaient peut-être eux-mêmes plus « motivés » que la moyenne par le sujet, ce qui pouvait biaiser les résultats.

Pourriez-vous citer des exemples de projets étrangers dédiés aux résultats négatifs ?

Tout dépend de ce que l'on entend par projet. Parmi les canaux toujours actifs de diffusion de résultats négatifs, seule une poignée de revues a choisi de se spécialiser dans ce type de contenus. Notre étude en cite quatre (*Journal of Trial and Error*, *Journal of Articles in Support of the Null Hypothesis*, *Negative Results*, *Journal of Pharmaceutical Negative Results*).

Nous recensons aussi huit revues (*ACS Omega*, *OSA Continuum*, *BMC Research*, etc.) dont la politique éditoriale traduit clairement la possibilité de soumettre des articles décrivant des expériences non concluantes.

Enfin, les chercheurs interrogés ont également été amenés à citer 25 titres de revues dans lesquelles ils sont parvenus à publier des résultats négatifs, sous couvert, la plupart du temps, de contrebalancer le propos par la mise en avant de résultats positifs, qui restaient le cœur du sujet.

Quel peut être l'apport des professionnels de l'information à cette réflexion ?

Les bibliothécaires réalisent un travail de sensibilisation des chercheurs et doctorants aux enjeux de la science ouverte. Par « science ouverte », on désigne l'accès à la connaissance dans son ensemble, ce qui laisse un espace de réflexion pour l'enjeu de la diffusion des résultats négatifs, qui constituent des résultats malgré leur épithète et ont

une valeur intrinsèque. Il relève donc du ressort du bibliothécaire de faire apparaître cet enjeu dans le cadre des formations dispensées auprès des communautés scientifiques. •

Présentation des auteurs

Marie-Emilia Herbet est conservatrice des bibliothèques et chef de projet en charge du développement de Datacc, dispositif d'accompagnement à la gestion des données en chimie et en physique.

Jérémy Léonard est ingénieur d'études chargé d'appui au projet Datacc. Il a notamment contribué à la production et l'analyse des résultats de l'enquête avec R.

Contributeurice

Lucie Albaret est conservatrice des bibliothèques; elle pilote la Direction d'appui à la recherche et à la science ouverte (Darso) au sein de la Direction générale déléguée « Bibliothèques et appui à la science ouverte » de l'Université Grenoble Alpes.

Focus sur le TOP Factor, un indicateur pour promouvoir la transparence des revues

par Sabrina Granger

Si la transparence n'est pas le synonyme de qualité, il n'en demeure pas moins qu'elle constitue un prérequis majeur pour évaluer la qualité de la production scientifique. Les initiatives de la part des chercheurs eux-mêmes se multiplient et le *BBF* vous propose un focus sur l'une d'elles : le TOP Factor.

• **Quoi ?**

Transparency Openess Promotion (TOP) Factor

• **Quand ?**

Le TOP Factor a été lancé le 10 février 2020.

• **Pourquoi ?**

Le TOP Factor se veut une alternative au facteur d'impact. C'est un outil de notation des revues.

Les critères du TOP Factor s'appuient sur les pratiques inhérentes à la science ouverte. Il s'agit de :

- fournir des critères pour évaluer la qualité du travail conduit par les revues ; en d'autres termes, il s'agit de répondre à la question : « Les processus mis en œuvre par la revue permettent-ils de diffuser des résultats vraiment fiables ? » Le TOP Factor mesure à quel degré les politiques des revues facilitent ou non le processus d'évaluation de la qualité de la recherche ;
- promouvoir un changement de culture dans le monde académique. Les revues les plus cotées survalorisent par exemple les résultats spectaculaires au détriment des résultats crédibles.



Ce que N'EST PAS le TOP Factor :

- un indicateur de qualité des articles publiés ;
- un classement de revues ;
- un critère d'évaluation de la production des chercheurs.

• Par qui ?

Le Center for Open Science (COS) est une organisation à but non lucratif fondée en 2013. Son objectif est d'accroître l'ouverture, l'intégrité et la reproductibilité de la recherche scientifique. Le COS s'appuie sur des communautés de pratiques dédiées à la science ouverte dans différentes disciplines. Le COS développe des outils tel que la plateforme Open Science Framework. Il soutient les logiciels libres et gratuits. En savoir plus sur le comité de pilotage¹.

• Pour qui ? À qui le TOP Factor s'adresse-t-il ?

- Auteurs : en savoir plus sur le fonctionnement des revues et leur capacité à tenir compte des pratiques en faveur de la transparence de la recherche.
- Rapporteurs : identifier les éléments devant être pris en compte lors de l'analyse des textes soumis.
- Éditeurs : trouver des exemples de bonnes pratiques de revues dans leur domaine ou dans d'autres disciplines ; comparer les étapes de traitement des articles proposées par les autres revues.
- Lecteurs : savoir sur quels processus la revue s'appuie pour évaluer les articles.
- Financeurs, décideurs : identifier les revues dont le fonctionnement est cohérent avec les valeurs de la science ouverte afin d'encourager les chercheurs à publier dans ces revues.

Et pour les professionnels de l'information ?

Au-delà des informations sur les revues elles-mêmes, le site du TOP Factor² constitue un outil intéressant pour des professionnels de l'information car il leur permet de mieux comprendre les différences de pratiques entre les disciplines : « *TOP Factor makes that diversity visible and comparable across research communities. For example, economics journals are at the leading edge of requiring transparency of data and code whereas psychology journals are among the most assertive for promoting preregistration.* »³ [« *Le TOP Factor rend cette diversité visible et comparable entre les communautés de recherche. Par exemple, les revues d'économie sont à la pointe de l'exigence de transparence des données et du code alors que les revues de psychologie sont parmi les plus affirmées pour la promotion du préenregistrement.* » (NDLR : toutes les traductions sont obtenues via DeepL.com, version gratuite.)]

• Comment ?

Le TOP Factor permet de classer les revues, mais son objectif n'est pas de hiérarchiser les revues. Il s'agit d'un ensemble modulaire d'indicateurs pour promouvoir des bonnes pratiques éditoriales tout en tenant compte des spécificités des disciplines.

Les revues peuvent mettre en œuvre 8 standards, composés chacun de 3 niveaux.

Les données sont mises à jour et alimentées par les utilisateurs du site et les éditeurs. Elles sont ensuite vérifiées par le COS.

Une base de 8 standards :

- Data Citation
- Data Transparency
- Analytic Methods Transparency (Code Transparency)
- Research Materials Transparency
- Design (reporting guidelines) and Analysis Transparency⁴



- Preregistration of Studies
- Preregistration of Analysis Plans
- Replication

2 standards supplémentaires :

- Registered Reports et Publication Bias
- Open Science Badges

3 niveaux de mise en œuvre par standard :

- 1) Disclose / Divulguer
c.-à-d. : les auteurs indiquent si les données sont disponibles dans un entrepôt.
- 2) Verify / Vérifier
c.-à-d. : la revue exige que les données soient accessibles, sous réserve des exceptions à prévoir.
- 3) Require / Exiger, rendre obligatoire
c.-à-d. : un article ne peut être publié si les données ne sont pas accessibles.

Exemple appliqué à la citation de données :

- 1) « *Journal describes citation of data in guidelines to authors with clear rules and examples.* » [« La revue décrit la citation des données dans des directives aux auteurs avec des règles et des exemples clairs. »]
- 2) « *Article requires appropriate citation for data and materials used consistent with the journal's author guidelines.* » [« L'article nécessite une citation appropriée des données et du matériel utilisés, conformément aux directives de la revue à l'intention des auteurs. »]
- 3) « *Article is not published until providing appropriate citation for data and materials following journal's author guidelines.* »⁵ [« L'article n'est pas publié tant que la citation appropriée des données et des matériaux n'a pas été fournie conformément aux directives de la revue à l'intention des auteurs. »]

Exemple appliqué à la réplication :

- 1) « *Journal encourages submission of replication studies.* » [« La revue encourage la soumission d'études de réplication. »]
- 2) « *Journal will review replication studies blinded to results.* » [« La revue examinera les études de réplication en aveugle des résultats. »]
- 3) « *Registered Reports for replications as a regular submission option.* » [« Rapports enregistrés pour les réplications en tant qu'option de soumission régulière. »]

• Contenus de la base

Plus de 1 000 revues mettent en place les recommandations TOP⁶.

Éditeurs :

AAAS, American Economics Association, BioMed Central, Cambridge University Press, Cell Press, Elsevier, Emerald, Hogrefe, IEEE, Nature⁷, Oxford University Press, Plos, Peer Community In, Routledge, Royal Society Publishing, Sage, Springer, Taylor & Francis, Ubiquity Press, University of Chicago Press, Wiley.

• Autres initiatives pour une meilleure reconnaissance des bonnes pratiques de recherche

Reconnaître les pratiques de la science ouverte :

JONES, Phill et Fiona MURPHY. *Openness Profile: Modelling research evaluation for open scholarship*. Zenodo. 2021. En ligne : <https://zenodo.org/record/4581490> [consulté le 10 novembre 2021].

Synthèse en français sur le rapport⁸ : « *Le profil d'ouverture (PO) a vocation à être une ressource numérique dans laquelle les résultats et les activités d'un contributeur à la recherche qui concourent à son ouverture seraient accessibles en un seul endroit.* »



[Note de traduction : le terme « profil d'ouverture » renvoie à la locution « openness profile », une création lexicale qui a pour ambition de désigner un référentiel d'actions de recherche diversifiées que la numérisation des travaux de recherche permet et que l'avènement de la science ouverte favorise. Les composantes de ce référentiel correspondraient à des paramètres nouveaux de signalement et d'appréciation de la recherche accomplie par un acteur individuel ou collectif. Le terme de « profil science ouverte » a pu apparaître dans d'autres documents en langue française.]

Notes

1. <https://www.cos.io/about/board>
2. <https://topfactor.org/>
3. CENTER FOR OPEN SCIENCE. *New Measure Rates Quality of Research Journals' Policies to Promote Transparency and Reproducibility*. 2020. En ligne : <https://www.cos.io/about/news/new-measure-rates-quality-research-journals-policies-promote-transparency-and-reproducibility> [consulté le 10 novembre 2021].
4. Les critères varient très fortement selon les disciplines ainsi qu'en témoigne la documentation produite par le NIH et le réseau EQUATOR.
5. CENTER FOR OPEN SCIENCE. *TOP Resources - Evidence and Practices*. 2016. En ligne : <https://osf.io/kgvna/> [consulté le 10 novembre 2021].
6. <https://www.cos.io/initiatives/top-guidelines>
7. NATURE. «Announcement: Transparency upgrade for Nature journals», *Nature*. Mars 2017, vol. 543 n° 7645. p. 288288.
8. <https://www.ouvrirelascience.fr/profil-douverture-de-knowledge-exchange-un-modele-de-reference-pour-levaluation-de-la-science-ouverte/>

Résumé

Pôles documentaires de référence en chimie et en physique, les bibliothèques de Lyon-1 et de l'Université Grenoble Alpes se sont associées pour développer Datacc, une plateforme soutenue par CollEx-Persée. Datacc est conçue comme un service d'accompagnement des chercheurs en physique et chimie. Marie-Emilia Herbet et Jérémie Léonard, avec la contribution de Lucie Albaret, présentent au BBF le projet et ses dernières évolutions, centrées autour des résultats négatifs.