



HAL
open science

Les revues SIC et les données de recherche Une étude empirique

Joachim Schöpfel, Eric Kergosien

► **To cite this version:**

Joachim Schöpfel, Eric Kergosien. Les revues SIC et les données de recherche Une étude empirique. Balisages, 2021, 3, 10.35562/balisages.641 . hal-03552220

HAL Id: hal-03552220

<https://hal.science/hal-03552220v1>

Submitted on 2 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

Les revues SIC et les données de recherche

Une étude empirique

Joachim Schöpfel

Maître de conférences, Université de Lille, GERiCO
joachim.schopfel@univ.lille.fr

Eric Kergosien

Maître de conférences, Université de Lille, GERiCO
eric.kergosien@univ.lille.fr

Dans le cadre de la politique de la science ouverte, les revues scientifiques jouent un rôle central pour l'ouverture et le partage des données de recherche. Notre étude porte sur 95 revues en sciences de l'information et de la communication (liste de la section CNU 71). Nous avons analysé leur politique éditoriale en matière d'intégrité et de transparence en appliquant les dix critères du *TOP Factor* du Center for Open Science aux recommandations et instructions aux auteurs. Plus de la moitié des revues n'affichent aucune politique dans ce domaine. Les autres encouragent les auteurs à partager leurs données et/ou d'autres matériels, avec parfois des consignes pour la citation des données. Très peu de revues vont plus loin et exigent une telle transparence. Les résultats sont discutés par rapport aux divergences d'interprétation et d'évaluation, par rapport à la représentativité de l'échantillon et par rapport aux choix des rédactions et éditeurs. Plusieurs pistes pour aller plus loin sont indiquées.

Revue scientifique, données de recherche, TOP factor, science ouverte, sciences de l'information et de la communication

Information and Communication Science

Journals and Research Data

An empirical study

In the context of open science policy, academic journals play a central role in opening and sharing research data. Our study focuses on 95 journals in information and communication sciences (CNU section 71 list). We analyzed their editorial policies on integrity and transparency by applying the ten criteria of the Center for Open Science's *TOP Factor* to recommendations and instructions to authors. More than half of the journals have no policy in this area. The rest encourage authors to share their data and/or other materials, sometimes with instructions for citing the data. Very few journals go further and require such transparency. The results are discussed in

relation to differences in interpretation and evaluation, in relation to the representativeness of the sample, and in relation to the choices made by editors and publishers. Several perspectives to go further are indicated.

Academic journals, research data, TOP factor, open science, information and communication science

Les revues SIC et les données de recherche

Une étude empirique

Introduction

La politique de la science ouverte met les revues scientifiques face à l'enjeu des données de recherche. Le Plan national de la science ouverte publié par le Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation en 2018 recommande, dans le cadre du soutien public aux revues, l'adoption d'une politique de données ouvertes associées aux articles. Cette annonce vise à « créer les conditions et promouvoir l'adoption d'une politique de données ouvertes associées aux articles publiés par les chercheurs » (MESRI, 2018). Le deuxième plan de juillet 2021 confirme et renforce cette politique pour la période 2021-2024 (MESRI, 2021). Concrètement, ce plan annonce en premier lieu la mise en œuvre de l'obligation de diffusion des données de recherche financées sur fonds public et la création d'une plateforme nationale pour les données de recherche (Recherche Data Gouv). Le plan intègre également plusieurs mesures en lien avec les publications :

- Encourager les bonnes pratiques en matière de citations des données ;
- Suivre la dynamique d'ouverture des jeux de données associés aux publications et déposés dans une sélection d'entrepôts grâce au baromètre de la science ouverte ;
- Développer une démarche proactive d'ouverture des données associées aux articles et publications sur des entrepôts de confiance thématiques ou sur Recherche Data Gouv ;
- Dans le cadre du soutien public aux plateformes d'édition scientifique et aux archives ouvertes, encourager leur articulation avec Recherche Data Gouv.

Par ailleurs, ce nouveau plan préconise plusieurs initiatives dans le domaine de la recherche médicale afin de réduire le « biais de publication » (tendance à ne publier que les études ayant obtenu un résultat positif). Ces initiatives visent entre autres à rendre obligatoire la déclaration de toute collecte prospective de données individuelles en biologie-santé, pour les essais cliniques et les études observationnelles, et à développer un portail de déclaration.

La politique de la science ouverte de l'Etat français n'est pas un phénomène isolé mais s'insère dans un écosystème de stratégies et d'initiatives de la part des établissements et organismes de recherche, des agences de financement, des éditeurs, de l'industrie et des services de l'information et, au niveau international, d'autres Etats et de la Commission Européenne.

Les éditeurs et les rédactions des revues scientifiques doivent donc s'adapter pour proposer aux chercheurs une offre de service appropriée aux incitations et obligations des autorités, des institutions et des agences.

En tant que chercheurs en sciences de l'information et de la communication (SIC), nous nous sommes posés la question quant à nos propres revues – à quel degré les revues scientifiques en SIC ont-elles commencé à modifier leur politique éditoriale afin d'intégrer les données de recherche ? Après un état de l'art, nous présenterons et discuterons ici les résultats d'une étude empirique avec un échantillon de 95 revues SIC. Cette étude fait suite à deux études sur l'appropriation de la science ouverte par les SIC : l'une sur le libre accès aux revues SIC (Schöpfel et al., 2018), l'autre sur les entrepôts de données de recherche en SIC (Prost & Schöpfel, 2019).

Etat de l'art

Comme vecteur principal de la communication des résultats de la recherche, les revues scientifiques jouent un rôle central pour la mise en œuvre de la politique de la science ouverte, aussi et notamment par rapport au partage des données de recherche. L'éditeur Ubiquity Press inclut l'accès aux « objets de recherche » (données, logiciels, bioressources, méthodologies...) dans ses critères d'intégrité¹ tandis que le rédacteur en chef de Molecular Brain se concentre sur les enjeux de l'évaluation des articles et de la crise de reproductibilité, « absence of raw data means absence of science » (Miyakawa, 2020).

Politiques publiques et autres initiatives

En France, trois documents récents précisent les développements attendus en ce qui concerne la mise en œuvre d'une politique de données par les revues :

- Les « Critères QUERO » (Gouzi et al., 2020) préconisent l'adoption d'un principe d'intégrité scientifique (18) et de lier les données, les codes et les autres résultats de recherche avec la publication qui en découle, dans la mesure du possible (30) ;
- Dans le cadre du soutien aux revues, l'INSHS du CNRS a établi un catalogue de critères de bonnes pratiques et de science ouverte² dont la diffusion ouverte des données de la recherche associées à la publication lorsque le cadre légal le permet (B4.b) et le développement d'une politique concernant la publication de données sous forme de data papers ou le dépôt conjoint publications/données (G5) ;

¹ Ubiquity Press research integrity <https://www.ubiquitypress.com/site/research-integrity/>

² INSHS <https://inshs.cnrs.fr/fr/soutien-aux-revues>

- Le Comité pour la science ouverte a publié sept recommandations détaillées à destination des revues qui veulent développer une politique de données (CoSO, 2021). Ce document propose un cadre de référence pour différentes politiques de données « selon la nature des incitations et prescriptions qu’elles proposent, en particulier selon qu’elles incitent ou obligent à diffuser tout ou partie des données (...), précisent des modalités de mise à disposition des données (et) soumettent les données au processus de revue par les pairs au même titre que les publications ».

Au niveau international, l’initiative QUEST recommande l’accès aux données de recherche pour renforcer la crédibilité, l’utilité et l’éthique de la recherche biomédicale, avec le préenregistrement des protocoles d’études avant le début de la recherche (Strech et al., 2020). Une telle initiative, tout comme la création d’un « Retraction Watch Transparency Index » par le Center for Scientific Integrity³ révèle trois points :

1. Dans la mesure où transparence et intégrité sont liées, la question des données de recherche fait partie d’un écosystème plus large, qui inclut d’autres aspects comme la rétraction d’articles, le préenregistrement d’études etc. ;
2. Même si en France, l’intérêt porte surtout sur les revues en SHS, les vrais enjeux se trouvent surtout dans les domaines des sciences de la vie et des sciences médicales ;
3. Il y a une demande de critères explicites et objectifs afin d’assurer un suivi (monitoring) et une évaluation des publications.

Strech et al. (2020) ont publié l’ébauche d’un écosystème de recherche responsable au sein d’une institution dans le domaine biomédical qui met en relation les grands principes de qualité scientifique, les processus et les services de soutien. Une telle formalisation va dans le même sens que la National Information Standards Organization (NISO) quand elle définit un cadre pour la reconnaissance des pratiques de partage de données et de méthodes dans les revues scientifiques (NISO, 2021), à partir des projets de « badges » (Kidwell et al., 2016) de plusieurs éditeurs (Springer Nature, ACM, IEEE) et du Center for Open Science (COS)⁴.

Le TOP Factor du Center for Open Science

À ce jour, le COS semble être le seul organisme international à avoir développé, à partir de 2015, un outil et une procédure pour analyser, évaluer et comparer les revues scientifiques par rapport à leur transparence et leur ouverture (Nosek et al., 2015 ; Mellor, 2021). L’évaluation s’appuie

³ Center for Scientific Integrity <http://retractionwatch.org/>

⁴ Center for Open Science <https://www.cos.io/>

sur l'analyse de la politique éditoriale d'une revue (recommandations, instructions aux auteurs etc.) à partir de dix critères, et le résultat de l'évaluation – le *TOP Factor* – est défini comme un nouvel « indicateur de qualité », différent d'un indicateur de prestige et complémentaire au facteur d'impact (Kepes et al., 2020). TOP veut dire « Transparency and Openness Promotion » ; aussi, la mise en place d'un tel indicateur est accompagnée par des recommandations aux revues afin d'adapter leur politique éditoriale concernant la disponibilité des données, des méthodes d'analyse et/ou du code utilisé pour l'analyse des données et des matériaux physiques ou numériques utilisés pour mener l'étude, mais aussi des méthodes de recherche et le préenregistrement de l'étude (Aalbersberg et al., 2018).

L'objectif de cette démarche est de développer des normes communes pour les pratiques ouvertes dans les revues, en traduisant les normes et les valeurs scientifiques en actions concrètes, ce qui peut nécessiter l'ajustement disciplinaire (Nosek et al., 2015). Plusieurs études montrent des exemples pour un tel ajustement, en psychologie (Powell et al., 2020) comme en sciences de la vie et de l'environnement (Parker et al., 2016).

Lors de l'analyse de la politique éditoriale d'une revue, chaque critère doit être évalué indépendamment des autres (cf. tableau 1).

No	Critère	Description
1	<i>Data citation</i>	Référence de données dans un entrepôt en utilisant des moyens standard, y compris le DOI.
2	<i>Data transparency</i>	Mise à disposition des données dans le cadre de la soumission à la revue.
3	<i>Analyse code transparency</i>	Mise à disposition du code d'analyse dans le cadre de la soumission à la revue.
4	<i>Materials transparency</i>	Mise à disposition des matériaux dans le cadre de la soumission à la revue.
5	<i>Design & analysis reporting guidelines</i>	L'inclusion d'instruments décrivant formellement la conception et l'analyse de l'étude, tels que les normes <i>Preferred Reporting Items for Systematic Reviews and Meta-Analyses</i> (PRISMA) ou <i>Consolidated Standards of Reporting Trials</i> (CONSORT).
6	<i>Study preregistration</i>	Préenregistrement de la collecte des données avant la réalisation de la collecte.
7	<i>Analysis plan preregistration</i>	Préenregistrement de l'analyse des données avant l'analyse.
8	<i>Replication</i>	Volonté explicite de la revue d'inclure des articles ne reposant pas sur la nouveauté mais des répliqués de travaux antérieurs.
9	<i>Registered reports & publication bias</i>	L'examen prospectif par les pairs, c'est-à-dire l'évaluation d'un manuscrit soumis à une revue avant la collecte et/ou l'analyse des données.

10	<i>Open science badges</i>	Utilisation de ce que l'on appelle des badges, qui sont attribués si un article adhère à des normes spécifiques, fournissant ainsi une incitation à promouvoir la transparence et l'ouverture.
----	----------------------------	--

Tableau 1. Les dix critères du *TOP Factor*

L'idée est simple : chaque article devrait comporter une déclaration indiquant la disponibilité des produits et matériaux de recherche qui sous-tendent ses conclusions. L'évaluation de chaque critère se fait à partir d'une grille de quatre niveaux, relatifs à la mention, l'encouragement, l'exigence et l'application de pratiques spécifiques de transparence et d'ouverture, qui sont mises en œuvre dans un système de notation à ancrage verbal (Tableau 2).

Niveau d'engagement	Points attribués (valeur)
Simple mention, ou rien	0
Encouragement, incitation	1
Obligation, exigence	2
Evaluation, contrôle	3

Tableau 2. Les quatre niveaux du *TOP Factor*

Ainsi, l'analyse attribue à la revue pour chaque critère entre 0 et 3 points, à l'exception du critère 10 qui n'a que 2 points maximum. Ces évaluations individuelles créent une sorte de profil spécifique de transparence et d'ouverture pour chaque revue. La somme de tous les points – entre 0 et 29 points - est la valeur du *TOP Factor* de cette revue.

Le site du COS contient une base de données avec les profils et les TOP Factors pour actuellement (juillet, 2021) presque 650 revues scientifiques de plusieurs dizaines d'éditeurs et couvrant un large éventail de domaines. Mais très peu de revues correspondent aux SIC, dont *Political Communication* (6 points), *Communication Research Reports* (5 points) et *Information, Communication & Society* (1 point). La revue avec le *TOP Factor* le plus élevé est *Meta-Psychology* ; mais plus de la moitié des revues évaluées n'ont aucun ou seulement 1 point.

En résumé, le *TOP Factor* couvre les principales dimensions de la science ouverte et fournit un outil utile et normalisé qui permet de comparer les revues selon la mesure dans laquelle elles encouragent ou exigent l'adhésion aux principes de la science ouverte. Afin de rendre l'évaluation plus fiable et faciliter la mise en place d'une politique éditoriale appropriée, l'équipe du COS vient de publier une procédure standardisée, le « TRUST process » (Mayo-Wilson et al., 2021).

L'application du TOP Factor

Depuis la mise en place du *TOP Factor*, plusieurs études ont appliqué la méthodologie de ce nouvel indicateur pour faire le bilan du degré d'ouverture et de transparence des revues d'un domaine particulier (Tableau 3). Ces études montrent un paysage hétérogène, plus ou moins avancé, avec une majorité de revues au tout début de ce processus.

Domaine	Auteurs et année	Nombre de revues	Moyenne	Médiane	Maximum
Biologie	Marshall & Strine (2020)	19	0,53	0	n/a
Biologie	Spitschan et al. (2020)	28	n/a	2,5	9
Médecine, psychologie	Cashin et al. (2020)	10	n/a	3,5	7
Informatique	Ballou et al. (2021)	51	2,5	0	15
Psychologie	Nutu et al. (2019)	60	n/a	n/a	n/a

Tableau 3. Etudes avec le *TOP Factor*

Malgré le petit nombre de revues, ces études confirment un paysage contrasté, avec quelques domaines et quelques revues plus ouverts et plus transparents que d'autres. En herpétologie (Marshall & Strine, 2020), sur les 19 revues évaluées seulement 3 ont obtenu un score. L'étude en chronobiologie (Spitschan et al., 2020) a constaté qu'il n'y a pas de corrélation significative entre le facteur d'impact et le *TOP Factor*.

D'autres études

D'autres études posent le même genre de question mais poursuivent une autre approche. Une publication récente applique seulement une partie des critères du TOP Factor à 30 revues dans le domaine du droit et constate que 14 revues ne contiennent aucune recommandation concernant la transparence des données, codes ou d'autres matériels (Zeiler & Chin, 2021). Pour un échantillon de 60 revues en psychologie, Nutu et al. (2019) décrivent une implémentation « asymétrique » d'une politique de données, avec le plus souvent un caractère non-obligatoire – et sans beaucoup d'effet sur les articles publiés.

Deux études évaluent l'ouverture et la transparence non pas au niveau des revues mais au niveau des articles. L'étude longitudinale de Serghiou et al. (2021) avec 2,75 millions d'article dans le domaine biomédical (source :

PubMed Central) révèle une augmentation du partage des données et codes et du préenregistrement des protocoles de recherche entre 1959 et 2020 mais constate aussi que d'une part, cela concerne peu de revues (<1% de revues avec une majorité d'articles de ce type) et peu d'articles (14,5% pour le partage des données, 2,5% pour le partage du code, 2,5% pour le préenregistrement). Ils observent tout de même la transparence de la recherche publiée s'est considérablement améliorée par rapport à deux autres aspects : la déclaration d'un conflit d'intérêt, et la déclaration des sources de financement. Dans le domaine des SIC, selon une étude de 10 517 articles 2010-2020 dans 26 revues de communication, seulement 5,1% des articles mentionnent des « pratiques de science ouverte », tel que le partage de données (Markowitz et al., 2021).

Toujours dans le domaine des SIC, Jackson (2021) vient de publier les résultats d'une analyse de contenu des recommandations et instructions aux auteurs de 201 revues (sources : Web of Science et Scopus), avec un constat mitigé : « (...) 10 (5%) exigeaient explicitement que les données de recherche originales soient rendues publiquement accessibles dans une archive ouverte, tandis que 6 (3%) exigeaient que les données soient rendues disponibles sous condition, sur demande. Par ailleurs, 96 revues (48%) recommandaient que les données soient archivées publiquement mais n'imposaient pas l'archivage ouvert comme condition de publication. Les informations sur les exigences de disponibilité des données étaient absentes des sites web de 89 revues (44%) ».

Bowman et al. (2021) ont mené une enquête internationale avec 330 chercheurs en communication. Résultat : une majorité connaît et soutient les principes et objectifs de la science ouverte mais seulement une (petite) minorité pratique ces principes, surtout en ce qui concerne le partage des matériels et la lecture des preprints. L'étude révèle également une opposition claire contre toute forme d'obligation : « optional but not mandatory ».

Plusieurs études de cas décrivent la mise en place d'une politique de données par une revue scientifique particulière. Ces études (dont une en SIC : Read et al., 2021) soulignent l'intérêt d'une telle politique éditoriale mais également la nécessité d'une période de transition, pour expliquer, argumenter et accompagner les auteurs-chercheurs (Grahe, 2020 ; Greiff et al., 2020). Une autre étude va plus loin dans la mesure où elle évalue l'impact de nouvelles instructions et recommandations sur les pratiques des auteurs – un impact décrit comme réel mais limité (Giofré et al., 2017).

Cependant, la pandémie du Coronavirus semble avoir accéléré ce processus. Ainsi, Besançon et al. (2020) constatent davantage de pratiques en faveur de la science ouverte, notamment dans le domaine médical : surtout par rapport au préenregistrement des études mais également au

niveau du partage des données et des codes. Une autre publication montre l'intérêt de ces pratiques dans les domaines de la biologie et de l'environnement et propose des outils de suivi (monitoring) (O'Dea et al., 2021).

Méthodologie

Comme pour l'étude de 2018 (Schöpfel et al., 2018), nous avons retenu la liste des revues qualifiantes (reconnues) de la section 71 du CNU comme représentative pour les revues scientifiques en SIC⁵. Nous avons vérifié et enrichi l'information pour chaque revue à partir de la base de données Ulrichsweb (ProQuest) et à partir de son site en ligne. Des 98 revues de la liste du CNU, un titre est un doublon et deux revues ont cessé de paraître. Les autres 95 revues constituent l'échantillon de cette étude.

Ensuite, nous avons évalué la politique éditoriale de chaque revue à partir des informations de son site web (instructions et recommandations aux auteurs etc.). L'évaluation a été faite à partir des dix critères de l'indicateur *TOP Factor* du Center for Open Science (COS), en attribuant une valeur de 0 à 3 en fonction des instructions et recommandations du site. Chaque revue a été évaluée par deux personnes différentes ; en cas de résultats divergents, il y a eu une troisième évaluation approfondie. La somme des dix critères correspond à l'indicateur *TOP Factor*.

L'évaluation a été réalisée par treize personnes formées aux critères du COS (dont huit étudiants en Master 2 GIDE « Gestion de l'Information et du Document en Entreprise » de l'Université de Lille). Sur deux points précis, nous avons pris contact directement avec le directeur du COS, David Mellor, pour clarifier et homogénéiser l'interprétation des critères.

Les résultats ainsi obtenus et validés ont été saisis sous forme d'un tableau Excel qui a servi de base pour les analyses. Les premières analyses ont été présentées et discutées le 21 mai 2021 dans le cadre du séminaire I2SP du LERASS de l'Université de Toulouse 3 – Paul Sabatier⁶.

Résultats

Le score global (TOP Factor)

Le score maximal (TOP Factor) des 95 revues est 10 sur 29. 55 revues ont un score de 0. Les autres 40 revues ont un score entre 1 et 10, avec une moyenne de 1,35 et une médiane de 0 (Figure 1).

⁵ Liste des revues scientifiques reconnues du domaine Sciences de l'information et de la communication (SIC) (mise à jour le 28/11/2018) <http://cpdirsic.fr/wp-content/uploads/2019/06/Revue-reconnues-SIC-28-11-18.pdf>

⁶ Laboratoire d'Études et de Recherches Appliquées en Sciences Sociales <https://www.lerass.com>

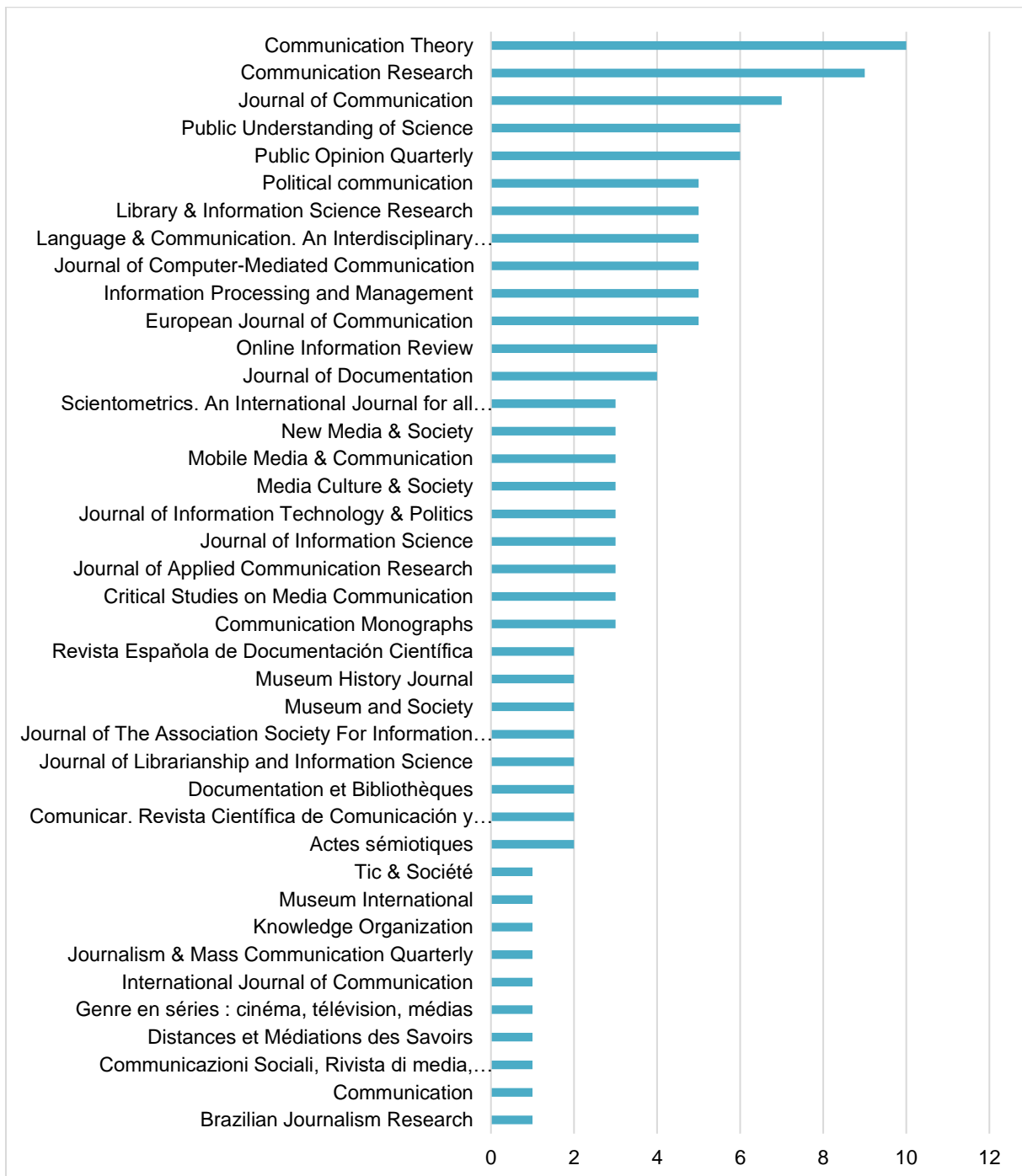


Figure 1. L'indicateur TOP Factor des revues avec un score positif (N=40)

Treize revues seulement ont un score >3. La revue avec le score le plus élevé est *Communication Theory* (Oxford University Press), suivi par *Communication Research* (Sage) et le *Journal of Communication* (Oxford University Press). Dans le domaine des sciences de l'information et de la documentation (*library and information sciences*), les revues avec le score

le plus élevé sont *Library & Information Science Research* et *Information Processing and Management* (les deux chez Elsevier).

La moyenne des dix critères

Chaque critère a été évalué sur une échelle allant de 0 à 3, à l'exception du critère 10 (cf. plus haut). Les trois critères avec les moyennes les plus élevées sont la transparence des données (0,34), la transparence d'autres matériels (0,31) et la citation des données (0,28) (Figure 2).

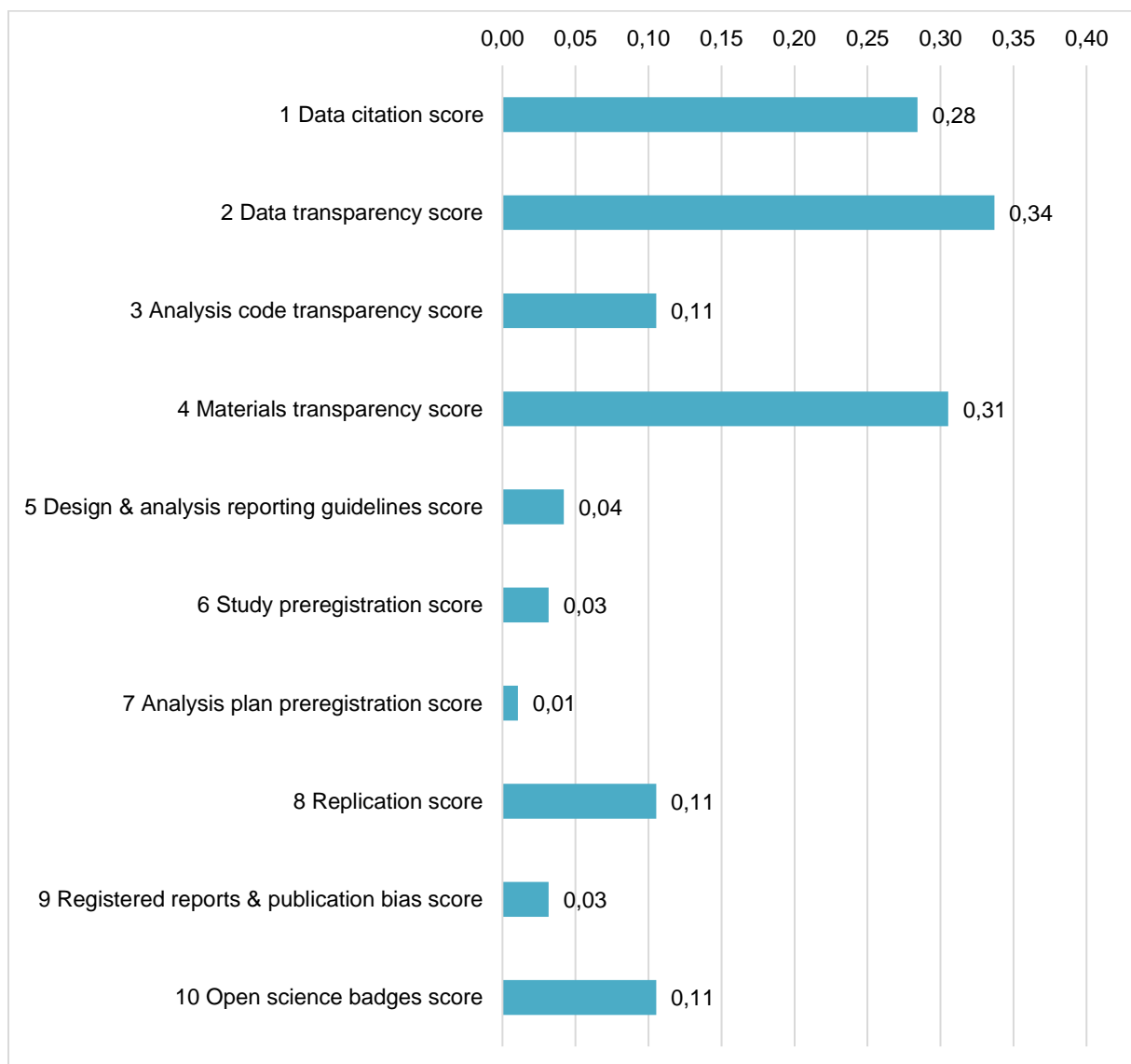


Figure 2. La moyenne des dix critères (N=95 revues)

Data transparency : A l'instar de la revue *Scientometrics* de Springer Nature, 30 revues SIC incitent ou rendent obligatoire la publication des données dans un entrepôt de données de recherche. Pour certaines revues,

il s'agit d'un encouragement de déposer les données dans un entrepôts (*data repository*) dans la mesure du possible (*whenever possible*). D'autres vont plus loin et demandent explicitement ce dépôt en ajoutant parfois que les auteurs sont priés de s'assurer que les données, matériels ou applications logicielles (codes) soutiennent leurs conclusions et sont conformes aux normes du domaine.

Materials transparency : 27 revues incitent ou rendent obligatoire le partage, d'une manière ou d'une autre, de tout matériel utilisé (logiciels, codes, modèles, algorithmes, protocoles, méthodes etc.) par le projet de recherche à l'origine de la publication, comme par exemple la revue *Library & Information Science Research* d'Elsevier. Parfois les instructions aux auteurs précisent la raison de ce choix, notamment pour faciliter la reproductibilité et la réutilisation des données.

Data citation : 25 revues incitent ou rendent obligatoire la citation des données de recherche utilisées pour la publication de l'article, avec des exemples et consignes explicites. La revue *Communication Theory* par exemple, publiée par Oxford University Press pour l'International Communication Association, précise sur son site qu'elle soutient les principes de citation des données de l'initiative internationale FORCE11⁷ et qu'elle exige que les jeux de données accessibles au public soient référencés dans la liste de référence avec un numéro d'accès ou un identifiant unique tel qu'un DOI et que les références de données doivent inclure les informations minimales recommandées par DataCite⁸ (auteur, année, titre, éditeur [ou nom de l'entrepôt], identifiant). La revue demande également d'ajouter la balise [dataset] au début d'une référence de données pour une meilleure identification et un balisage correct par l'éditeur.

Quant aux autres critères, la moyenne est encore plus basse et presque négligeable. La médiane pour l'ensemble des critères est 0, et leur valeur maximale est de 1 ou 2. Concrètement, cela veut dire qu'au moins la moitié des revues n'obtient aucun score pour l'ensemble des critères et qu'aucune revue n'annonce explicitement vouloir inclure les données, matériels, codes etc. dans l'évaluation des articles.

Discussion

Divergences d'évaluation

L'évaluation des 95 revues sur dix critères a produit 950 valeurs. Sur les 950 valeurs, il y a eu 63 incohérences (deux évaluations divergentes), ce

⁷ Joint Declaration of Data Citation Principles <https://www.force11.org/datacitationprinciples>

⁸ Data citation examples <https://datacite.org/cite-your-data.html>

qui correspond à 7%. Ces incohérences concernent surtout les critères 1 à 4 (Figure 3).

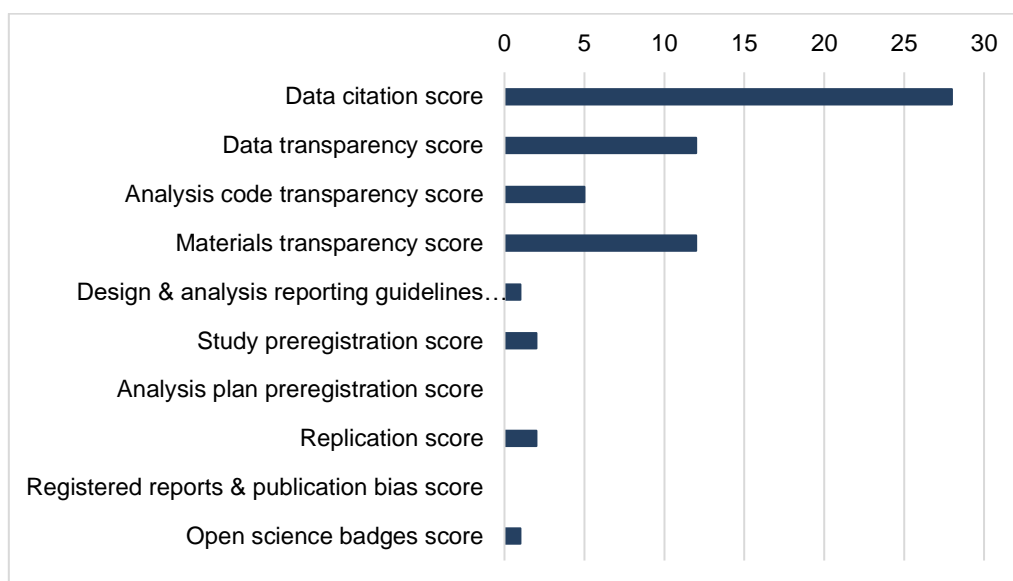


Figure 3. Nombre d'incohérences dans l'évaluation des critères du TOP factor (N=95 revues)

Toutes ces incohérences ont été corrigées par une troisième analyse approfondie. La différence d'interprétation du critère 1 (Data citation score) portait essentiellement sur le renvoi du site de la revue vers un site de standard bibliographique (Chicago, APA etc.) ; c'est pour ce point précis que nous avons sollicité l'avis du COS (David Mellor).

Parfois, ce problème est lié au fait de ne pas avoir pu identifier la bonne page ou le bon document. Mais dans la plupart des cas, il s'agit d'une divergence d'interprétation des consignes, un problème qui est partiellement lié aux ambiguïtés des formulations et de la terminologie des consignes eux-mêmes. Il n'est pas toujours évident de savoir si l'intention de la politique éditoriale est d'encourager ou exiger des pratiques de transparence, et il est possible que l'interprétation et la compréhension par les rédacteurs, éditeurs et auteurs divergent (Christian et al., 2020 ; Spitschan et al., 2020).

Représentativité de l'échantillon

Dans l'étude présentée dans ce le cadre de ces travaux, nous nous sommes limités aux revues qualifiantes sélectionnées par le Conseil National des Universités (CNU) de la section 71. Bien que représentatives de la section, ces revues ne couvrent pas l'ensemble des revues du champs, et n'intègrent pas certaines revues nouvelles qui prônent l'ouverture des données, et la

mise à disposition des articles en open access. Cette liste fermée est actuellement en cours de révision, et nous prévoyons de mettre à jour les résultats de notre étude lorsque celle-ci sera publiée. Nous pouvons tout de même affirmer dès à présent que les révisions seront mineures, et que les tendances mises en avant dans les résultats que nous présenterons seront sensiblement les mêmes.

Toujours concernant l'échantillon, une autre piste de travail est d'étendre les revues à celles répertoriées dans des bases internationales, et dans lesquelles des chercheurs en 71 publient, alors même qu'elles ne sont pas reconnues comme qualifiantes en 71. Nous pensons notamment aux bases du Web of Science (WoS) ou encore Scopus (avec 722 revues en *library and information sciences* et en *communication*), et dans ce cas, il faudrait sélectionner les revues supplémentaires selon un ou plusieurs critères précis, nombre de publications des auteurs de la section dans la revue sur une ou plusieurs années, thématiques des publications à travers les mots clés, titres, etc.

Une politique prudente

L'analyse des informations présentes sur le web décrivant la politique éditoriale de chacune des 95 revues fait ressortir différentes formes de demande aux auteurs pour publier leurs données. Pour la plupart, les revues ne demandent rien aux auteurs. Quant aux autres, les revues sont parfois dans l'incitation, et plus rarement dans l'obligation. Ainsi, le *Journal of Applied Communication Research* incite les auteurs à déposer leurs données dans un entrepôt reconnu : « *Authors are encouraged to deposit the dataset(s) in a recognized data repository that can mint a persistent digital identifier, preferably a digital object identifier (DOI) and recognizes a long-term preservation plan* ». A mi-chemin entre incitation et obligation, *Communication Theory* encourage fortement (« *strongly encourages* ») un tel dépôt quand c'est possible d'un point de vue éthique. Encore plus strict, les Actes Sémiotiques exigent explicitement que « *Les données sous-jacentes doivent être représentées avec précision dans le manuscrit ou rendues accessibles par l'indication d'un lien vers une base de données, ou d'un autre type de lieu de dépôt* ».

Mais l'obligation reste à ce jour une exception. Dans l'ensemble, les rédactions et les éditeurs font preuve d'une certaine prudence, préférant l'incitation et l'encouragement à une obligation qui risquerait d'être mal reçu par les auteurs (cf. Markowitz et al., 2021 ; Bowman et al., 2021). En comparaison avec d'autres domaines et en particulier, avec les disciplines biomédicales (cf. tableau 3), les revues SIC se démarquent surtout par deux aspects : relativement peu de revues ont formulé une telle politique (médiane = 0), et cette politique fait abstraction de plusieurs critères qui

semblent avoir davantage de valeur dans d'autres domaines (préenregistrements, code, badges...). Mais la situation peut changer rapidement dans un contexte dynamique et évolutif. Nous avons eu plusieurs retours de la part des rédactions qui vont dans le même sens : il y a une prise de conscience par rapport à la politique de la science ouverte des autorités, organismes et institutions, et il y a un début de réflexion et de discussion au sein des équipes éditoriales.

Pertinence du TOP Factor pour les revues SIC

Cependant, pour bien comprendre les enjeux et les différentes stratégies, il faudrait compléter cette étude par une analyse qualitative sur la base d'une enquête en ligne ou d'entretiens semi-directifs menés auprès des comités de rédaction des différentes revues de la section 71 afin de mieux cerner leurs objectifs et choix par rapport aux différents aspects d'intégrité et de transparence. Il serait particulièrement intéressant d'évaluer l'intérêt de ces différents aspects pour les revues en SIC, autrement dit, d'interroger l'universalité de la pertinence de ces critères d'intégrité et de transparence en général et du *TOP Factor* en particulier ; il serait également intéressant d'analyser l'impact d'une mutualisation de procédures et de dispositifs et du rôle des grands éditeurs et plateformes.

Mais pour revenir à la question de l'universalité des critères d'intégrité et de la transparence pour les revues SIC : cette question se pose à plusieurs niveaux, aussi bien par rapport aux contenus des revues (la part des études empiriques) et à l'arsenal de méthodologies disciplinaires (la part des données propriétaires ou à caractère personnel, la nature des « autres matériels » etc.) qu'aux enjeux réels et aux procédures habituelles de la démarche scientifique en SIC (pas de préenregistrement, peu d'entrepôts de données à ce jour etc.). La question se pose si la grille du *TOP Factor* est en adéquation avec les revues SIC, s'il faut une adaptation ou une autre grille. En attendant, elle permet la comparaison avec d'autres revues et d'autres domaines, sans porter un jugement de valeur.

Conclusion

Dans le cadre de la politique en faveur de la science ouverte, les revues scientifiques occupent une place centrale, aussi bien pour la diffusion libre des publications que pour l'accès aux données de recherche. Notre étude porte sur ce dernier aspect. Nous avons analysé la politique éditoriale des revues qualifiantes en sciences de l'information et de la communication (SIC) en ce qui concerne la transparence et l'intégrité de la publication des résultats, en appliquant les critères du *TOP Factor* du Center for Open Science (COS).

Nos résultats montrent un paysage contrasté, une dynamique enclenchée, certes, mais plus ou moins prudente encore, au début d'un processus qui ne sera pas nécessairement similaire au développement et aux exigences dans d'autres domaines. Peut-être l'idée même que toutes les revues devraient suivre le même chemin est une double erreur, idéologique et bureaucratique. Erreur idéologique, car cette approche paraît contraire à la bibliodiversité. Erreur bureaucratique, car elle ne tient pas compte de la réalité du terrain.

Pour aller plus loin dans l'analyse et dans la mise en œuvre d'une politique de la science ouverte réaliste et raisonnable, plusieurs pistes sont ouvertes.

Elargir l'échantillon des revues. Dans la poursuite de cette étude, nous prévoyons d'enrichir l'échantillon d'analyse à davantage de revues en SIC, notamment à travers la base Scopus, l'idée étant de vérifier dans un premier temps si l'échantillon des revues qualifiantes de la section 71 est représentatif de l'ensemble des revues de la section en termes de politique d'ouverture des données.

Ajouter une analyse qualitative. Dans un second temps, nous comptons mener une enquête qualitative auprès des comités de rédaction des revues de la section afin de mieux cerner leur politique éditoriale concernant l'ouverture des données, et dans ce sens d'affiner notre analyse concernant le niveau d'incitation à l'ouverture de ces revues à la science ouverte.

Evaluer la particularité des revues SIC. Limiter une telle étude à un champs disciplinaire de la taille des SIC ne permet ni généralisation ni comparaison. Il serait souhaitable, par la suite, de contribuer à une analyse similaire d'autres domaines en SHS, comme par exemple de la sociologie ou de la linguistique, afin de bien comprendre la particularité des revues SIC par rapport aux données de recherche. Au moment de la rédaction (septembre 2021), nous avons commencé une étude sur les 138 revues françaises de la base Scopus dans les domaines archéologie, culture, géographie, linguistique, psychologie, sociologie et sciences politiques ; les résultats sont attendus pour début 2022.

A partir de ces analyses et sur le terrain de l'édition, il serait souhaitable de faire participer la communauté des parties prenantes (éditeurs, rédactions, hébergeurs de plateformes de revues, auteurs/chercheurs) à l'élaboration et à la mise en œuvre des politiques en matière de données, à aligner les exigences de la politique sur les normes et les meilleures pratiques et à exprimer les exigences de la politique en matière de données dans un langage clair et cohérent. Peut-être les SIC auront la capacité de développer une « approche modèle » propre à leur discipline et leurs revues, à l'instar d'initiatives (« *data sharing policy working group* ») dans d'autres domaines (cf. Read et al., 2021 ; Christian et al., 2020 ; Besançon et al., 2020). Une telle approche disciplinaire pourrait inclure une concertation avec des

experts en entrepôts de données afin de soutenir la mise en œuvre des politiques de données ; peut-être la mise en place d'une infrastructure nationale de données de recherche (Recherche Data Gouv) sera aussi l'opportunité pour un centre thématique pour les données de recherche en SIC.

Remerciements

Nous remercions les collègues qui ont contribué à l'évaluation des revues : Stéphane Chaudiron, Marianne Duquenne, Bernard Jacquemin et Hélène Prost. Nous remercions également les étudiants du Master GIDE qui ont pris en charge l'évaluation de plusieurs revues dans le cadre de leur formation : Tiffany Blanplain, Delphine Bréhéret, Simon Crépieux, Marion Hébert, Achille Planck, Camille Thieffry, Cymoril Vanoorenberghe et Estelle Waselinck.

Bibliographie

Aalbersberg, Ij., Appleyard, T., Brookhart, S., Carpenter, T., Clarke, M., Curry, S., ... Vazire, S. (2018). *Making Science Transparent By Default; Introducing the TOP Statement*. The TOP Statement Working Group. OSF Preprints. <https://doi.org/10.31219/osf.io/sm78t>

Ballou, N., Warriar, V. R., & Deterding, C. S. (2021). Are You Open? A Content Analysis of Transparency and Openness Guidelines in HCI Journals. *CHI Conference on Human Factors in Computing Systems (CHI '21)*, May 8-13, 2021. Association for Computing Machinery (ACM). Retrieved from <http://eprints.whiterose.ac.uk/170362/>

Besançon, L., Peiffer-Smadja, N., Segalas, C., Jiang, H., Masuzzo, P., Smout, C., ... Leyrat, C. (2020). Open Science Saves Lives: Lessons from the COVID-19 Pandemic. *BioRxiv*, 2020.08.13.249847. <https://doi.org/10.1101/2020.08.13.249847>

Bowman, N., Rinke, E. M., Lee, E.-J., Nabi, R., & Vreese, C. de. (2021). How Communication Scholars See Open Scholarship: A survey of International Communication Association scholars. *SocArXiv*. <https://doi.org/10.31235/osf.io/fwpxd>

Cashin, A. G., Bagg, M. K., Richards, G. C., Toomey, E., McAuley, J. H., & Lee, H. (2020). Limited engagement with transparent and open science standards in the policies of pain journals: a cross-sectional evaluation. *BMJ Evidence-Based Medicine*, bmjebm-2019-111296. <https://doi.org/10.1136/bmjebm-2019-111296>

Christian, T.-M., Gooch, A., Vision, T., & Hull, E. (2020). Journal data policies: Exploring how the understanding of editors and authors

corresponds to the policies themselves. *PLOS ONE*, 15(3), e0230281.
<https://doi.org/10.1371/journal.pone.0230281>

CoSO (2021). *Recommandations aux revues souhaitant définir une « politique de données » liées aux publications*. Paris : Comité pour la science ouverte, Collège Données de la recherche
<https://www.ouvrirlascience.fr/recommandations-aux-revues-souhaitant-definir-une-politique-de-donnees-liees-aux-publications/>

Giofrè, D., Cumming, G., Fresc, L., Boedker, I., & Tressoldi, P. (2017). The influence of journal submission guidelines on authors' reporting of statistics and use of open research practices. *PLOS ONE*, 12(4), e0175583. <https://doi.org/10.1371/journal.pone.0175583>

Gouzi, F., Ochandiano, J.-L. De, & Thomas, A. (2020). *Critères QUERO - Préconisations pour la qualité des revues scientifiques*. Document de travail. <https://hal-univ-bourgogne.archives-ouvertes.fr/hal-03060601>

Grahe, J. (2021). The necessity of data transparency to publish. *The Journal of Social Psychology*, 161(1), 1–4.
<https://doi.org/10.1080/00224545.2020.1847950>

Greiff, S., van der Westhuizen, L., Mund, M., Rauthmann, J. F., & Wetzel, E. (2020). Introducing New Open Science Practices at EJPA. *European Journal of Psychological Assessment*, 36(5), 717–720.
<https://doi.org/10.1027/1015-5759/a000628>

Jackson, B. (2021). Open Data Policies among Library and Information Science Journals. *Publications*, 9(2), 25.
<https://doi.org/10.3390/publications9020025>

Kepes, S., Banks, G. C., & Keener, S. K. (2020). The TOP factor: An indicator of quality to complement journal impact factor. *Industrial and Organizational Psychology*, 13(3), 328–333.
<https://doi.org/10.1017/iop.2020.56>

Kidwell, M. C., Lazarević, L. B., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L.-S., ... Nosek, B. A. (2016). Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency. *PLOS Biology*, 14(5), e1002456.
<https://doi.org/10.1371/journal.pbio.1002456>

Markowitz, D. M., Song, H., & Taylor, S. H. (2021). Tracing the Adoption and Effects of Open Science in Communication Research. *PsycArXiv*.
<https://doi.org/10.31234/osf.io/dsf67>

Marshall, B. M., & Strine, C. T. (2021). Make like a glass frog: In support of increased transparency in herpetology. *Herpetological Journal*, 31(1), 35–45. <https://doi.org/10.33256/31.1.3545>

- Mayo-Wilson, E., Grant, S., Supplee, L., Kianersi, S., Amin, A., DeHaven, A., & Mellor, D. (2021). Evaluating implementation of the Transparency and Openness Promotion (TOP) guidelines: the TRUST process for rating journal policies, procedures, and practices. *Research Integrity and Peer Review*, 6(1), 9. <https://doi.org/10.1186/s41073-021-00112-8>
- Mellor, D. (2021). Improving norms in research culture to incentivize transparency and rigor. *Educational Psychologist*, 56(2), 122-131. <https://doi.org/10.1080/00461520.2021.1902329>
- MESRI (2018). *Plan national pour la science ouverte*. Paris : Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation <https://www.ouvrirlascience.fr/plan-national-pour-la-science-ouverte/>
- MESRI (2021). *Deuxième Plan national pour la science ouverte*. Paris : Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation <https://www.enseignementsup-recherche.gouv.fr/cid159131/le-plan-national-pour-la-science-ouverte-2021-2024-vers-une-generalisation-de-la-science-ouverte-en-france.html>
- Miyakawa, T. (2020). No raw data, no science: another possible source of the reproducibility crisis. *Molecular Brain*, 13(1), 24. <https://doi.org/10.1186/s13041-020-0552-2>
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... Yarkoni, T. (2015). Promoting an open research culture. *Science*, 348(6242), 1422–1425. <https://doi.org/10.1126/science.aab2374>
- NISO (2021). *NISO RP-31-2021, Reproducibility Badging and Definitions*. Recommended Practice. <https://doi.org/10.3789/niso-rp-31-2021>
- Nutu, D., Gentili, C., Naudet, F., & Cristea, I. A. (2019). Open science practices in clinical psychology journals: An audit study. *Journal of Abnormal Psychology*, 128(6), 510–516. <https://doi.org/10.1037/abn0000414>
- O'Dea, R. E., Parker, T. H., Chee, Y. E., Culina, A., Drobniak, S. M., Duncan, D. H., ... Nakagawa, S. (2021). Towards open, reliable, and transparent ecology and evolutionary biology. *BMC Biology*, 19(1), 68. <https://doi.org/10.1186/s12915-021-01006-3>
- Parker, T. H., Forstmeier, W., Koricheva, J., Fidler, F., Hadfield, J. D., Chee, Y. E., ... Nakagawa, S. (2016). Transparency in Ecology and Evolution: Real Problems, Real Solutions. *Trends in Ecology & Evolution*, 31(9), 711–719. <https://doi.org/10.1016/j.tree.2016.07.002>
- Powell, D. M., Spence, J. R., & Stanley, D. J. (2020). I-O psychology's top journals at the bottom of the TOP ranking: Should we consider openness

and transparency when ranking journals? *Industrial and Organizational Psychology*, 13(3), 316–320. <https://doi.org/10.1017/iop.2020.52>

Prost, H., & Schöpfel, J. (2019). Les entrepôts de données en sciences de l'information et de la communication (SIC). Une étude empirique. *Études de Communication*, (52), 71–98. <https://doi.org/10.4000/edc.8604>

Read, K., Campbell, A., Kitchin, V., MacDonald, H., & McKeown, S. (2021). Embracing the value of research data: introducing the JCHLA/JABSC Data Sharing Policy. *Journal of the Canadian Health Libraries Association / Journal de l'Association des Bibliothèques de la Santé du Canada*, 42(1). <https://doi.org/10.29173/jchla29536>

Schöpfel, J., Prost, H., & Fraisse, A. (2018). Plus ou moins open : les revues de rang A en Sciences de l'information et de la communication. *RFSIC Revue Française des Sciences de l'Information et de la Communication*, (15). <https://rfsic.revues.org/4706#ftn22>

Serghiou, S., Contopoulos-Ioannidis, D. G., Boyack, K. W., Riedel, N., Wallach, J. D., & Ioannidis, J. P. A. (2021). Assessment of transparency indicators across the biomedical literature: How open is open? *PLOS Biology*, 19(3), e3001107. <https://doi.org/10.1371/journal.pbio.3001107>

Spitschan, M., Schmidt, M. H., & Blume, C. (2020). Transparency and open science principles in reporting guidelines in sleep research and chronobiology journals. *Wellcome Open Research*, 5, 172. <https://doi.org/10.12688/wellcomeopenres.16111.1>

Strech, D., Weissgerber, T., & Dirnagl, U. (2020). Improving the trustworthiness, usefulness, and ethics of biomedical research through an innovative and comprehensive institutional initiative. *PLOS Biology*, 18(2), e3000576. <https://doi.org/10.1371/journal.pbio.3000576>

Zeiler, K., & Chin, J. (2021). Replicability in Empirical Legal Research. *Annual Review of Law and Social Sciences*, 17 (submitted). https://scholarship.law.bu.edu/faculty_scholarship/1053/