



HAL
open science

Publication et Données Reproductibles de la Recherche : Que Faire ?

Andrea Chiarelli, Lucia Loffreda, Rob Johnson, Jean-François Nominé, Manon Parrinello, Marc Rubio

► **To cite this version:**

Andrea Chiarelli, Lucia Loffreda, Rob Johnson, Jean-François Nominé, Manon Parrinello, et al.. Publication et Données Reproductibles de la Recherche : Que Faire ?. Knowledge Exchange. 2023, 82 p. <hal-04170637>

HAL Id: hal-04170637

<https://hal.science/hal-04170637v1>

Submitted on 25 Jul 2023

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.

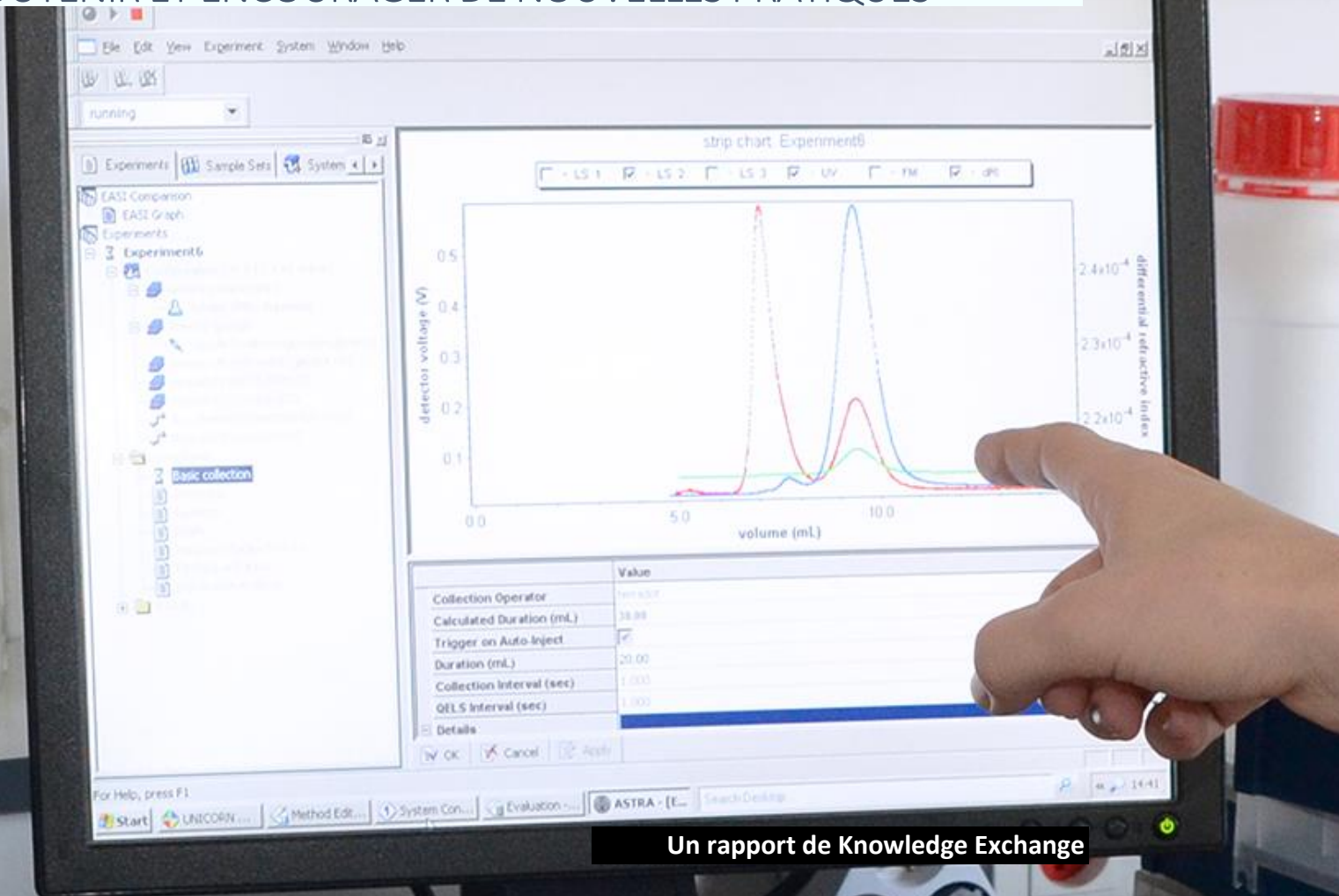
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



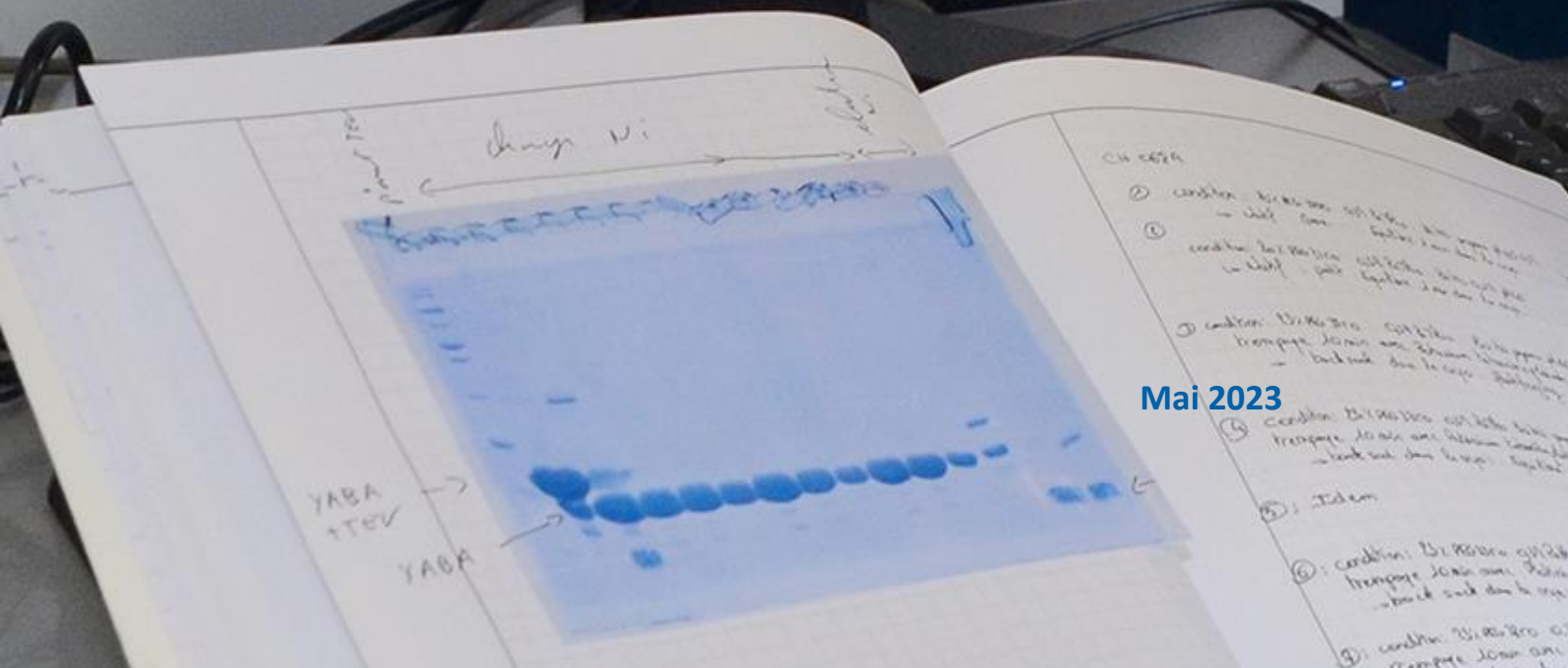
Distributed under a Creative Commons CC BY-SA 4.0 - Attribution - ShareAlike - International License

PUBLICATION ET DONNÉES REPRODUCTIBLES DE LA RECHERCHE : QUE FAIRE ?

SOUTENIR ET ENCOURAGER DE NOUVELLES PRATIQUES



Un rapport de Knowledge Exchange



Mai 2023

Remerciements

Cette étude a été réalisée grâce aux travaux du groupe de travail de Knowledge Exchange sur la publication de résultats de recherche reproductibles ci-après, animé par Juliane Kant et Anna Mette Morthorst et sous la houlette de Birgit Schmidt et Jeroen Sondervan en tant qu'experts pilotes :

- ▶ Anna Mette Morthorst (coopération danoise en matière d'e-infrastructure - DeIC).
- ▶ Birgit Schmidt (Université de Göttingen)
- ▶ Birte Christensen-Dalsgaard (Université d'Aarhus)
- ▶ Daniel Nüst (Université de Münster)
- ▶ Jeroen Sondervan (Université d'Utrecht)
- ▶ Juliane Kant (Deutsche Forschungsgemeinschaft - DFG)
- ▶ Matt Jaquery (Université d'Oxford)
- ▶ Pierre-Carl Langlais (Paris Sorbonne-CELSA)
- ▶ Saskia Woutersen (Université de Leiden)
- ▶ Verena Heise (Chercheuse indépendante en science ouverte, formatrice et consultante)
- ▶ Yrsa Neuman (Université Åbo Akademi)

Nous remercions également Bas Cordewener et Sarah James du Bureau de Knowledge Exchange pour leur soutien, ainsi que Laura Fortunato (Université d'Oxford), qui a procédé à une révision critique de notre recherche et nous a aidés à nous orienter dans les complexités de ce domaine en pleine évolution. Enfin, nous exprimons notre reconnaissance aux 51 contributeurs sans lesquels cette étude n'aurait pas été possible, pour leurs précieuses idées et pour le temps qu'ils nous ont consacré pendant une période particulièrement turbulente. On trouvera une liste complète des intervenants qui se sont impliqués dans ce projet à l'annexe A.

Version française :



Le présent ouvrage est la traduction du rapport original :

The Art of Publishing Reproducible Research Outputs: Supporting emerging practices through cultural and technological innovation.

rédigé par : Andrea Chiarelli, Lucia Loffreda, Rob Johnson (Research Consulting) email : andrea.chiarelli@research-consulting.com

et paru le 4/11/2021 : DOI: <https://doi.org/10.5281/zenodo.5521077>, tout le contenu publié peut être partagé (CC BY 4.0) creativecommons.org/licenses/by/4.0

La version française est une traduction de l'anglais post-éditée réalisée à l'aide du logiciel DeepL Pro, révisée et remaniée par les membres de l'équipe de Traduction de l'Inist-CNRS (UPA 076) :

- ▶ Jean-François Nominé (jean-francois.nomine@inist.fr - ) pour les chapitres 3, 6 et 7, et la révision générale de l'ouvrage ;
- ▶ Manon Parrinello (manon.parrinello@inist.fr - ) pour l'introduction, le chapitre 2 et les annexes ;
- ▶ Marc Rubio (marc.rubio@inist.fr) pour les chapitres 4 et 5.

Juillet 2023

Mêmes conditions que l'original (CC BY 4.0)

Crédit de page de couverture : © Vanessa CUSIMANO / IBCP / CNRS Images

Résumé

Introduction

Contexte et objectifs	<p><u>Knowledge Exchange</u> (KE) a commandé le présent rapport pour explorer les pratiques et les obstacles actuels dans le domaine de la reproductibilité de la recherche, en mettant l'accent sur la phase de publication et de diffusion. Nos conclusions visent à comparer et à inspirer les stratégies, les politiques et les pratiques opérationnelles et à mettre en commun les leçons tirées à partir d'un large éventail de groupes d'acteurs.</p>
Définitions	<p>Dans cet ouvrage, nous définissons la reproductibilité de la recherche comme les situations où les données et les procédures partagées par les auteurs d'une étude sont utilisées pour obtenir les mêmes résultats que dans leur travail original. Nous reconnaissons que d'autres concepts tels que la réplication, la robustesse et la généralisation des résultats de recherche sont des notions pertinentes dans un débat, mais nous avons considéré qu'ils dépassaient le cadre formel de notre recherche.</p>
Méthodologie	<p>Dans ce rapport, nous avons recueilli les points de vue d'organismes de financement de la recherche, d'institutions de recherche, de sociétés savantes, de chercheurs, d'éditeurs académiques et de fournisseurs d'infrastructures et de services dans le monde entier. Nous avons appliqué le modèle du Cadre de la science ouverte de <u>Knowledge Exchange</u> (<u>Knowledge Exchange Open Scholarship Framework</u>, ou KE OS Framework), qui permet d'aborder des aspects spécifiques de la science ouverte, afin d'étudier les pratiques reproductibles de publication : ce modèle a servi de base à la conception et à la réalisation de toutes les composantes de notre recherche, notamment une analyse documentaire complète et une série d'entretiens et de groupes de discussion avec un total de 51 contributeurs. Les résultats des entretiens et des groupes de discussion ont été transcrits et codés qualitativement pour une analyse thématique.</p>

Les tenants et aboutissants du discours sur la reproductibilité de la recherche

<p>Les infrastructures actuelles en pleine expansion peuvent être un atout pour les pratiques de reproductibilité.</p>	<p>L'essor du numérique a entraîné d'importantes mutations dans le paysage de la recherche, notamment de nouveaux outils et services technologiques, de nouvelles approches de recherche et la multiplication d'infrastructures techniques interconnectées. Dans ce contexte, diverses solutions sont plus largement que jamais à disposition pour documenter, partager et analyser les données et les résultats, ouvrant ainsi la voie à la reproductibilité dans les flux de travail et les pratiques de publication. Les principaux avantages de la recherche reproductible sont une confiance accrue dans les conclusions et les résultats et la possibilité de poursuivre son propre travail (ou celui d'un autre) à l'avenir. Au niveau du système, les pratiques de recherche reproductibles peuvent conduire à une plus grande transparence, ouverture et confiance dans la science.</p>
--	--

<p>Certains obstacles peuvent entraver les pratiques reproductibles</p>	<p>Certains obstacles peuvent entraver la publication de résultats de recherche reproductibles, notamment les structures d'incitation actuelles dans l'univers académique, les différences de capacités techniques entre chercheurs, la faible connectivité entre les solutions techniques et l'incohérence des normes de communication. Les méthodes de recherche, qui ont tendance à varier en fonction des disciplines scientifiques, ont également une incidence sur les efforts déployés que le travail d'un chercheur devienne reproductible. Par exemple, il est relativement simple de mettre en place des méthodes de recherche généralement en lien avec des disciplines quantitatives de manière à ce qu'elles deviennent reproductibles. En revanche, cela devient plus complexe à mettre en œuvre dans des processus opératoires où un élément qualitatif important est présent.</p>
---	---

Acteurs, rôles et responsabilités

<p>Au niveau micro : les chercheurs et les groupes de recherche ont un contrôle direct sur les pratiques quotidiennes.</p>	<p>Les chercheurs au niveau individuel et les groupes de recherche ont un rôle important à jouer, car c'est sur eux que repose la responsabilité de concevoir, réaliser et diffuser des travaux de recherche et ils sont les seuls à avoir facilement accès à tous les objets de recherche concernés. Ils peuvent favoriser les pratiques reproductibles de publication en encourageant et en mettant en place des phases de processus reproductibles et en intégrant les impératifs de reproductibilité lorsqu'ils évaluent les publications de leurs pairs.</p>
<p>Au niveau méso : les disciplines doivent afficher leurs exigences et les éditeurs doivent les mettre en œuvre.</p>	<p>Les communautés scientifiques et les disciplines peuvent influencer l'élaboration des politiques en devenant en définissant et en faisant connaître leurs approches de la reproductibilité à d'autres acteurs, tels que les éditeurs, qui ne souhaitent pas imposer de leur chef des prescriptions susceptibles de ne pas refléter les pratiques disciplinaires. Une fois que ces exigences sont claires, les éditeurs peuvent y contribuer en rendant obligatoire l'utilisation de listes de contrôle ou de directives appropriées auprès des réviseurs, des pairs examinateurs et des auteurs de sorte à encourager les pratiques <i>reprocompatibles</i>¹.</p>
<p>Au niveau méso : les organismes de recherche n'ont pas pour habitude d'imposer des pratiques de publication reprocompatibles.</p>	<p>De nombreux organismes de recherche n'ont pas de politiques spécifiques orientées vers la reproductibilité de la recherche. Cependant, ils consentent des investissements importants quand des directives (ou <i>mandats</i>) émanant de financeurs de la recherche ou des politiques se mettent en place. Du fait que les pratiques de recherche reproductibles ne sont pas actuellement un impératif ferme, il se peut que les organismes de recherche endossent la reproductibilité de la recherche par des approches ad hoc, en fonction de la stratégie de leur institution et de leur population de chercheurs.</p>

¹ Note de traduction (NdT): Pour lever toute ambiguïté dans la restitution d'une expression anglaise – *reproducible publishing practices* - dont la transposition pure et simple en français peut prêter à confusion, nous nous autorisons ce néologisme dans le présent ouvrage, pour ainsi marquer clairement l'objet de la reproductibilité : non la publication en tant qu'acte et objet en soi, mais les processus, méthodes, données qu'elle présente d'une manière qui autorise leur reproduction par les lecteurs. Si, pour des raisons stylistiques, l'expression de *pratiques reproductibles* pourra apparaître par la suite, le lecteur est invité à l'entendre dans le sens de cette « reprocompatibilité ».

<p>Au niveau macro : les organismes de financement de la recherche considèrent la reproductibilité comme se situant dans un débat plus large.</p>	<p>Peu d'organismes de financement de la recherche tiennent une position d'importance eu égard à la situation de la reproductibilité de la recherche. Celle-ci est souvent traitée dans le cadre de prescriptions plus larges telles que celles en lien avec les données, la science ouverte ou l'intégrité de la recherche. Tout comme les éditeurs, les organismes de financement de la recherche estiment qu'il est difficile de formuler des consignes de reproductibilité à des allocataires relevant d'un large éventail de disciplines, et ils ne ressentent pas d'urgence à développer de nouvelles politiques.</p>
---	---

Encourager et faciliter les pratiques de publication reprocompatibles

<p>Les incitations et le soutien actuels aux pratiques de publication reprocompatibles sont limités</p>	<p>Les efforts envers la reproductibilité ne sont pas actuellement encouragés à l'intérieur du cycle de la recherche, et les pratiques de publication en ce sens sont généralement perçues comme des activités en plus, non reconnues. Il est nécessaire de revoir systématiquement les dispositifs d'incitation scientifique actuels pour récompenser de manière plus logique les comportements propices aux pratiques reprocompatibles dans les publications. Les organismes de recherche peuvent apporter un soutien à chaque chercheur en facilitant l'allègement de certaines contraintes de temps et en complétant ses compétences lorsqu'elles font défaut. Ce genre de soutien peut se matérialiser dans la création de nouvelles fonctions dans leurs institutions comme <i>des intendants et des curateurs de données</i>², ainsi que des <i>bibliothécaires de spécialité disciplinaire</i>³.</p>
<p>De nouvelles filières de formation et de soutien se développent dans le monde entier.</p>	<p>Une série de filières de soutien et de formation, tant à l'intérieur qu'à l'extérieur des organismes de recherche, se développe dans le monde entier. Toutefois, un soutien plus structuré aux pratiques de publication reprocompatibles serait le bienvenu, car ces initiatives sont plus l'exception que la règle. On a observé que le statut de <i>champions</i> est un rouage important dans la sensibilisation, et des entretiens ont fait ressortir qu'il existe une marge d'amélioration de la formation à la reproductibilité dans les maquettes d'enseignements dispensés aux étudiants.</p>

Innovation technologique

<p>Il existe déjà de nombreuses infrastructures numériques pour porter les pratiques de publication reprocompatibles.</p>	<p>Un large éventail d'outils et d'infrastructures numériques existe dans le paysage actuel de la recherche, et les chercheurs ont généralement connaissance de la manière dont ceux-ci peuvent être employés pour des pratiques de publication favorables à la reproductibilité. Cependant, tous ne sont pas formés au même niveau à l'utilisation de ces outils (ceci souvent en fonction des habitudes de la discipline et des approches en recherche en vigueur), et le manque d'interopérabilité entre les infrastructures est considéré comme un obstacle matériel.</p>
<p>Les principes des données FAIR peuvent</p>	<p>Les problèmes décrits autour de la connectivité des services sont en partie recouverts dans le « I » des principes des données FAIR : trouvabilité, accessibilité, interopérabilité et réutilisabilité. De manière plus générale, les principes</p>

² NdT : Pour la préservation de la francophonie dans ce texte, nous restituons ainsi les vocables de *data steward* et *data curator* respectivement pour exprimer la hiérarchie entre deux fonctions aux champs d'intervention complémentaires.

³ NdT : Même remarque que précédemment concernant le vocable anglais de *subject librarian*.

sous-tendre ces pratiques dans la publication.	des données FAIR permettent la publication reprocompatible, car bon nombre des pratiques sous-tendues par la curation, le partage et la réutilisation des données favorisent également la reproductibilité.
--	---

Comment financer les coûts des pratiques de publication reprocompatibles

Le coût des contrôles de reproductibilité varie en fonction du moment et des responsabilités.	Le financement du temps et des autres coûts liés aux efforts de reproductibilité est une considération essentielle pour l'avenir, et de nombreux acteurs ont un rôle à jouer pour que cela soit possible. Nos recherches ont mis en évidence la nécessité de couvrir différents types de coûts : le temps et les efforts des chercheurs dans le cadre d'un projet de recherche ; ceux du personnel de soutien à la recherche employé dans les organismes de recherche ; les contrôles de reproductibilité internes au processus de publication ; et les contrôles de reproductibilité post hoc. Le premier de ces éléments est considéré comme une condition nécessaire pour faire de la reproductibilité de la recherche une priorité, tandis que les autres font actuellement l'objet de discussions et d'expérimentations encore en cours.
Trois voies principales sont disponibles pour financer le coût des contrôles de reproductibilité	Dans le contexte du processus de publication, nous avons identifié trois voies pour mettre en œuvre les contrôles de reproductibilité. Nous avons constaté que les éditeurs ont la possibilité de créer des fonctions en leur sein, tels qu'éditeurs de données ou de reproductibilité, et que des prestataires extérieurs pourraient exercer un rôle de vérification de la reproductibilité dans les articles. En outre, les pairs examinateurs pourraient assumer des responsabilités supplémentaires en testant la reproductibilité des articles, des données et du code lorsque ceux-ci sont envisagés pour la publication ; toutefois, nous notons que la communauté des chercheurs peut être réticente à envisager cette approche, car elle nécessite du temps et des compétences que tous les chercheurs ne possèdent pas nécessairement.
Les financements des infrastructures numériques peuvent être bénéfiques pour piloter de nouvelles solutions	Notre recherche a révélé que deux domaines des infrastructures numériques pourraient bénéficier de financements supplémentaires ou nouveaux pour que les pratiques reprocompatibles soient mieux accompagnées dans les publications. Premièrement, étant donné qu'une augmentation des pratiques de recherche ouverte et du partage nécessitera des fonctionnalités et des capacités améliorées, les financements peuvent permettre de renforcer le rôle des infrastructures existantes. Deuxièmement, les financements publics peuvent être envisagés comme un moyen de développer et de piloter les infrastructures numériques ab initio qui fournissent des fonctionnalités associées à la reproductibilité dans l'ensemble du processus de recherche, et dans l'objectif d'élaborer des modèles économiques durables à moyen et long terme.
Le contrôle de la conformité est complexe dans la pratique	Plusieurs complications apparaissent lorsqu'il s'agit de contrôler la conformité aux exigences de reproductibilité. En particulier, il est difficile de parvenir à un accord sur la responsabilité de la conduite des contrôles de reproductibilité. Cela s'explique en partie par le fait qu'une telle activité comprend l'examen d'objets de recherche connectés, ce qui, à son tour, peut nécessiter une com-

préhension des normes, des méthodologies ou des domaines (sous-)disciplinaires dont tous les acteurs ne sont pas dotés pour y veiller et/ou les faire respecter.

Conclusions

- La reproductibilité fait partie de la vision de la science ouverte
- Les pratiques reproductibles de publication évoluent dans le cadre d'un vaste processus de changement culturel dans le paysage de la recherche. Par conséquent, les progrès sont lents et s'étalent sur une période de temps considérable. D'autre part, l'innovation technologique progresse rapidement : le déséquilibre entre l'innovation technique et culturelle engendre un tableau complexe mais optimiste de l'avenir de la recherche reproductible. La grande majorité des chercheurs sont très exigeants : nous pensons qu'ils adopteront volontiers des pratiques de publication reprocompatibles, pour autant qu'un équilibre soit trouvé entre les attentes croissantes et les récompenses pratiques.
- La diversité sera la clé d'un changement positif
- Il est essentiel que l'accent mis sur la reproductibilité ne conduise pas à une « culture de la honte et du blâme », mais qu'il soit au contraire accueilli comme une occasion d'améliorer les pratiques de recherche. Le risque que les politiques et leur mise en œuvre laissent peu de place à la nuance existe. Pour certaines cultures épistémiques, la reproductibilité sera plus difficile à comprendre et à mettre en œuvre, ou ne sera peut-être même pas l'objectif ; dans d'autres cas, elle ne sera peut-être pas considérée comme la principale marque de qualité, mais simplement comme une option parmi d'autres. Il sera donc nécessaire de donner la priorité à la diversité lorsque nous repenserons les pratiques de recherche afin de préserver et de renforcer la confiance dans la science.

Cinq messages à retenir






	La reproductibilité fait partie de la vision de la science ouverte, aux côtés de concepts tels que la réplication, la robustesse et la généralisation des résultats de recherche. Il est difficile de poursuivre un changement de culture en matière de reproductibilité sans tenir compte de ce contexte plus large.
	La collaboration des acteurs est nécessaire pour continuer à développer des pratiques de publication reprocompatibles. Tous les acteurs, du chercheur individuel aux organismes nationaux et internationaux, ont un rôle à jouer, notamment dans le cadre de l'élaboration et de la mise en œuvre des politiques.
	Les incitations aux pratiques favorables à la reproductibilité dans les publications sont actuellement limitées. Les organismes de recherche commencent à aider les chercheurs à répondre à leurs attentes croissantes en matière de reproductibilité et une demande croissante se fait jour en direction de nouveaux parcours de formation et de soutien dans ce domaine.
	La gestion, la conservation et le partage des données et des méthodes de recherche sont des conditions nécessaires que les publications soient reprocompatibles. Il est essentiel que ces pratiques deviennent la norme pour faire avancer l'agenda de la reproductibilité, et certains rôles institutionnels dédiés, tels que les intendants des données, peuvent être nécessaires pour répondre à la demande de soutien.
	Les pratiques de publication reprocompatibles nécessitent une série de solutions technologiques, mais la plupart des contributeurs s'accordent à dire que celles-ci sont déjà disponibles dans le paysage actuel de la recherche. La principale lacune technique semble être l'interopérabilité entre les outils et les flux de travail disponibles. Cependant, nous constatons également que les programmes de formation ne couvrent pas les solutions technologiques pour la reproductibilité actuellement.

Table des matières

Résumé.....	3
Introduction.....	3
Les tenants et aboutissants du discours sur la reproductibilité de la recherche.....	3
Acteurs, rôles et responsabilités.....	4
Encourager et faciliter les pratiques de publication reprocompatibles.....	5
Innovation technologique.....	5
Comment financer les coûts des pratiques de publication reprocompatibles.....	6
Conclusions.....	7
Cinq messages à retenir.....	8
Table des matières.....	9
1. Introduction.....	11
Définitions.....	11
Méthodologie.....	12
Limitations.....	13
Structure du rapport.....	13
Données justificatives.....	13
3. Les tenants et aboutissants du discours sur la reproductibilité de la recherche.....	15
Comprendre l'importance de la reproductibilité de la recherche.....	15
La reproductibilité : difficile définition.....	16
Moteurs et avantages perçus.....	17
Obstacles potentiels.....	19
L'impact des différences disciplinaires.....	20
4. Acteurs, rôles et responsabilités.....	23
Exploration des rôles et des responsabilités.....	23
Niveau micro : chercheurs et groupes de recherche.....	23
Niveau méso : usages disciplinaires.....	24
Niveau méso : réseaux internationaux de reproductibilité.....	26
Niveau méso: attentes institutionnelles en matière de reproductibilité de la recherche.....	29
Niveau méso : le rôle des exigences de l'éditeur et leurs lignes directrices sur la reproductibilité.....	29
Niveau macro : le rôle des organismes de financement de la recherche dans la définition des politiques de reproductibilité.....	32
Niveau macro : l'intérêt croissant des pouvoirs publics.....	33
Contributions des niveaux micro, méso et macro dans le cadre de l'élaboration des orientations politiques.....	35
5. Encourager et faciliter les pratiques de publication reprocompatibles.....	38

Alléger les contraintes de temps pour que les pratiques de publication reprocompatibles se généralisent	39
Développer de nouvelles voies de soutien et de formation.....	40
6. Innovation technologique	43
Infrastructures numériques.....	43
Le déficit d'innovation : la connectivité des services	45
Évaluer et comparer les solutions techniques.....	47
L'impact potentiel des principes des données FAIR	47
Aperçu du discours sur la reproductibilité sur les réseaux sociaux.....	50
7. Comment financer les coûts des pratiques de publication reprocompatibles	52
Implications financières des flux de publications reproductibles	52
Financement du temps et des activités des chercheurs dans le cadre d'un projet de recherche	52
Financement du temps et des efforts du personnel d'appui à l'intérieur des organismes de recherche.....	53
Financement des contrôles de reproductibilité dans le cadre du processus de publication	54
Le coût des infrastructures numériques.....	59
Le contrôle du respect des préconisations et ses complexités	60
8. Conclusions.....	63
(Dés)équilibres entre infrastructures techniques et culturelles.....	63
Pratiques reprocompatibles de publication : quelle vision ?	64
Conduire le changement	65
9. Références.....	66
Données produites par l'étude de Knowledge Exchange	66
Bibliographie.....	66
10. Annexe A – Remerciements et participants	72
Tableau A1 : Groupe de travail de Knowledge Exchange sur la publication de résultats de recherche reproductibles	72
Tableau A2 : Contributeurs au projet	72
11. Annexe B – Structure des entretiens.....	77

1. Introduction

Ce travail a été commandé par Knowledge Exchange, un groupe d'organisations nationales de six pays européens qui souhaitent favoriser la science ouverte en créant une infrastructure d'information au niveau international. Dans cette étude, nous cherchons à examiner les pratiques et les obstacles actuels à la publication de résultats de recherche reproductibles et à déterminer comment les infrastructures techniques et sociales peuvent soutenir les développements futurs dans ce domaine.

Contexte et objectifs

L'augmentation du nombre d'outils numériques dans le paysage actuel de la recherche offre plus que jamais la possibilité de documenter et d'ouvrir le processus de recherche et ses résultats⁴. En outre, les cultures de recherche ont évolué vers un désir de plus grande transparence, notamment dans le cadre de l'agenda de la science ouverte⁵ et des tendances actuelles de la communication savante. En particulier, la capacité à reproduire les résultats afin de confirmer leur fiabilité évolue pour devenir une part plus constante du processus de recherche : il est admis que les pratiques de recherche reproductibles peuvent permettre un développement scientifique continu et aider à garantir la qualité des résultats publiés.

Knowledge Exchange (KE) a donc lancé une activité intitulée « *Publishing Reproducible Research Output* » (Publier des résultats de recherche reproductibles), afin d'explorer les pratiques et les obstacles actuels dans ce domaine, et de déterminer comment les infrastructures techniques et sociales peuvent soutenir les développements futurs. Ce rapport fait partie du travail plus large de KE sur les infrastructures de recherche, la science ouverte et le libre accès : il cherche à comparer

et à donner des idées de stratégies, les politiques et les pratiques opérationnelles et à partager les leçons apprises par un large éventail de groupes d'acteurs.

Définitions

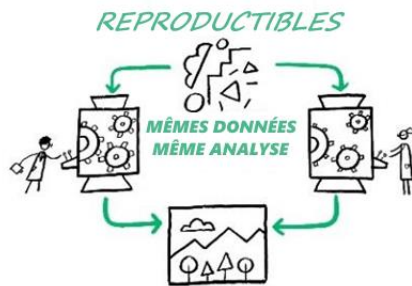
Nous reconnaissons que « la reproductibilité, du moins dans de nombreuses disciplines et cultures de recherche, est l'un des piliers des pratiques de recherche responsables »,⁶ et qu'elle est souvent associée au concept voisin de « répliquabilité ». Dans ce rapport, la reproductibilité de la recherche est définie comme le cas où les données et les procédures partagées par les auteurs d'une étude sont utilisées pour obtenir les mêmes résultats que dans leur travail original (voir figure 1). Cela peut nécessiter, par exemple, une description détaillée des méthodes utilisées pour traiter et analyser les données, l'accès à tout ensemble de données pertinent et la possibilité d'obtenir et d'exécuter un code informatique, le cas échéant.

⁴ Breuer, C., & Trilcke, P. (2021). Expanding academic publishing practices alongside the digital turn. Alliance of Science Organisations. 1-15. <https://doi.org/10.48440/ALLIANZOA.042>

⁵ Dans le présent rapport, nous utilisons le terme "science ouverte" au sens large, pour désigner la recherche ouverte et les communications sur la recherche, quelle que soit la discipline.

⁶ Tjeldink, J. K., Horbach, S. P. J. M., Nuijten, M. B., & O'Neill, G. (2021). Towards a Research Agenda for Promoting Responsible Research Practices. *Journal of Empirical Research on Human Research Ethics*, 16(4), 450-460. <https://doi.org/10.1177/15562646211018916>

FIGURE 1. NOTRE DÉFINITION DE LA REPRODUCTIBILITÉ (CC BY THE TURING WAY COMMUNITY ; SCRIBERIA)⁷



D'autre part, nous considérons que la réplication vise à valider la recherche, puisque de nouvelles données sont collectées et, parfois, de nouvelles analyses sont effectuées pour vérifier si des résultats cohérents peuvent être obtenus par rapport à l'étude en jeu. Cette considération, ainsi que d'autres, connexes, telles que la robustesse et la généralisation des résultats de recherche, n'ont pas été au cœur de notre recherche, mais sont mentionnées tout au long du rapport lorsqu'elles sont opportunes pour la discussion.

Méthodologie

Cette étude applique le *Cadre Knowledge Exchange de la Science Ouverte* (abrégé en KE OS Framework) - un modèle permettant d'aborder des aspects spécifiques de la science ouverte - pour étudier les pratiques reproductibles de publication (voir Figure 2). Le *KE OS Framework* comprend trois niveaux (micro, méso, macro), quatre phases de recherche (découverte, planification, phase de projet, diffusion) et quatre domaines (politique, économique, social, technologique).

FIGURE 2. CADRE KNOWLEDGE EXCHANGE DE LA SCIENCE OUVERTE⁸



Les dimensions ci-dessus ont été utilisées pour formaliser la conception et la réalisation d'une analyse documentaire complète, d'une série d'entretiens avec 41 parties prenantes du paysage de la recherche et de deux groupes de discussion, recueillant les points de vue de dix autres parties prenantes. Les résultats des entretiens et des groupes de discussion ont été transcrits et codés qualitativement à des fins d'analyse thématique.

Les groupes d'intervenants pris en compte comprenaient les organismes de financement de la recherche, les institutions de recherche, les sociétés savantes, les chercheurs (qu'ils soient établis ou en début de carrière), les éditeurs universitaires et les fournisseurs d'infrastructures et de services. Nos prises de contact ont donné la priorité à la collecte des points de vue des parties prenantes dans les pays représentés dans le partenariat Knowledge Exchange, à savoir le Danemark, la Finlande, la France, l'Allemagne, les Pays-Bas et le Royaume-Uni. Au cours de ce travail, nous avons également contacté des parties prenantes d'autres pays d'Europe et d'Amérique du Nord.

⁷ The Turing Way Community., & Scriberia. (2021). Illustrations from the Turing Way book dashes. Zenodo. <https://doi.org/10.5281/ZENODO.4906004>

⁸ Knowledge Exchange. (2017). Open Scholarship Framework. <https://bit.ly/30L9Sqv>

Enfin, notre étude comprenait une analyse des données Twitter portant sur une série de mots-dièse (hashtags) définis en collaboration avec le groupe d'experts de KE dédié à cette étude. Une méthodologie détaillée de cette analyse est disponible sous la forme d'un R notebook (voir les données justificatives), et des informations spécifiques sont incluses dans le rapport.

Limitations

Nous reconnaissons que la présente étude est soumise aux limitations suivantes :

- ▶ Les participants à l'étude ont été recrutés via un échantillonnage de convenance, c'est-à-dire que nous avons interrogé des parties prenantes qui étaient à la fois disponibles et désireuses de participer.
- ▶ Nos interactions avec ces acteurs se sont concentrées sur une série de pays spécifiques. Par conséquent, il ne convient peut-être pas de généraliser les conclusions de cette étude aux cultures et contextes de recherche qui n'ont pas été consultés.
- ▶ Notre analyse s'appuie sur un codage thématique, qui repose sur une certaine dose d'interprétation subjective. Les données de recherche et leur analyse ont été ouvertement partagées afin de garantir la disponibilité de nos observations de fond à toute partie intéressée.
- ▶ Les données des réseaux sociaux sont un élément instantané à un moment donné, et nous notons que Twitter n'offre pas un échantillon représentatif de la population générale, et ne nous permet pas de tirer des conclusions définitives.

Structure du rapport

Dans les parties qui suivent, notre rapport se structure ainsi :

- ▶ reproductibilité de la recherche : définitions du propos ;
- ▶ parties prenantes, rôles et responsabilités ;
- ▶ encouragement et soutien des pratiques de publication reprocompatibles ;
- ▶ innovation technologique ;
- ▶ couverture des coûts des pratiques reproductibles de publication ; et
- ▶ conclusions.

Toutes les sections ont été approvisionnées par les résultats de notre analyse documentaire et des interactions avec des parties prenantes. Des citations et des études de cas sont fournies tout au long du rapport pour illustrer les déclarations faites et les crédibiliser dans les observations recueillies.

Données justificatives

Les informations complémentaires suivantes ont été mises à disposition au cours de ce projet, dans le but d'améliorer la transparence et la reproductibilité de notre travail :

- ▶ plan de gestion des données ;⁹
- ▶ sources documentaires et codage thématique ;¹⁰
- ▶ diaporama de l'analyse documentaire (rapport intermédiaire) ;¹¹
- ▶ liste des personnes interrogées et questions d'entretien ;¹²
- ▶ résumés anonymes du codage des entretiens à partir de NVivo ; et¹³
- ▶ R notebook et données soutenant l'analyse des réseaux sociaux.¹⁴

⁹ Loffreda, L., & Chiarelli, A. (2021). Publishing Reproducible Research Outputs - Data Management Plan. Zenodo. <https://doi.org/10.5281/ZENODO.4697399>

¹⁰ Loffreda, L., & Chiarelli, A. (2021). Publishing Reproducible Research Outputs - Literature sources and Thematic coding (Version 1) [Data set]. Zenodo. <https://doi.org/10.5281/ZENODO.4748748>

¹¹ Chiarelli, A., Loffreda, L., & Johnson, R. (2021). Publishing Reproducible Research Outputs - Literature findings. Zenodo. <https://doi.org/10.5281/ZENODO.4675457>

¹² Chiarelli, A., & Loffreda, L. (2021). Publishing Reproducible Research Outputs - Interviewees and interview questions (Version 1) [Data set]. Zenodo. <https://doi.org/10.5281/ZENODO.5141665>

¹³ Loffreda, L., & Chiarelli, A. (2021). Publishing Reproducible Research Outputs - Thematic coding of interview findings (Version 1) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.5512420>

¹⁴ Chiarelli, A. (2021). Publishing Reproducible Research Outputs - Text and data mining code (Version 2) [Computer software]. Zenodo. <https://doi.org/10.5281/ZENODO.5512019>

Les éléments de cette étude faisant appel à des personnes ont été soumis à l'aval du comité d'éthique de la recherche de la *School of Anthropology and Museum Ethnography* .

2. Les tenants et aboutissants du discours sur la reproductibilité de la recherche

L'essor des technologies numériques a donné naissance à toute une série d'options permettant de documenter, de partager et d'analyser les méthodologies, les données et les résultats de la recherche plus largement que jamais : ces options peuvent être adoptées par la communauté des chercheurs afin d'accroître la transparence et le partage, et de faciliter la réutilisation future des méthodologies, des données et des résultats. Certains obstacles culturels et pratiques subsistent, notamment le fait que l'effort de reproductibilité des flux de travail tend à varier selon la méthode et la discipline de recherche.

Comprendre l'importance de la reproductibilité de la recherche

Le concept de reproductibilité de la recherche est connu et celui-ci fait l'objet de débats depuis longtemps dans la sphère académique^{15,16} car il est au cœur de ce que nous appelons les bonnes pratiques de la recherche.

Au fil du temps, l'évaluation par les pairs est devenue la principale modalité permettant de vérifier qu'une personne a mené ses recherches avec un niveau de qualité et conformément aux principaux usages et méthodologies disciplinaires. Cependant, à partir de l'avènement des ordinateurs modernes et des langages de programmation (comme Fortran dans les années 1950), de nombreuses disciplines se sont vues « propulsées à une vitesse vertigineuse »,¹⁷ l'éventail des possibilités offertes pour documenter, partager et analyser des données et des résultats augmentant considérablement. Ce tournant numérique n'a pas seulement affecté la recherche traditionnellement informatique et quantitative, mais a également modifié la forme de nombreuses disciplines qualitatives,

notamment dans des activités comme le codage thématique, la numérisation de documents, les annotations et les enregistrements.

En outre, de nouvelles infrastructures de recherche interconnectées continuent d'émerger,¹⁸ tandis qu'au cours des dix dernières années, « les éditeurs scientifiques traditionnels se sont métamorphosés sous nos yeux, abandonnant leurs anciennes méthodes de travail et se lançant dans une entreprise de vaste réinvention de leurs tâches et leur vocation ». ¹⁹

« L'article traditionnel, en deux dimensions, ne suffit plus. De nos jours, vous aimeriez avoir des graphiques dynamiques, des graphiques interactifs, peut-être même des données interactives lorsque vous allez plus en profondeur. » Un éditeur

La méthode de travail numérique d'aujourd'hui et l'éventail des moyens de partage et de connexion des objets de recherche (numériques) peuvent permettre à la reproductibilité de devenir une réalité, ce qui était impossible auparavant. Dans cette section, nous examinons les principaux moteurs et obstacles à l'aune de la reproductibilité de la

¹⁵ Popper, K. (2005; première édition 1934). *The logic of scientific discovery*. Routledge.

¹⁶ Dewald, W. G., Thursby, J. G., & Anderson, R. G. (1986). Replication in empirical economics: The journal of money, credit and banking project. *The American Economic Review*, 76(4), 587-603. https://econpapers.repec.org/article/aeaecrev/v_3a76_3ay_3a1986_3ai_3a4_3ap_3a587-603.htm

¹⁷ Perkel, J. M. (2021). Ten computer codes that transformed science. *Nature*, 589(7842), 344-348. <https://doi.org/10.1038/d41586-021-00075-2>

¹⁸ European Strategy Forum on Research Infrastructures (ESFRI). (2021). ESFRI announces new RIS for Roadmap 2021. <https://www.esfri.eu/latest-esfri-news/new-ris-roadmap-2021>

¹⁹ The Scholarly Kitchen. (2021). Digital Transformation Requires Cultural Evolution. <https://scholarlykitchen.sspnet.org/2021/07/14/title-digital-transformation-requires-cultural-evolution/>

recherche afin de préparer le terrain à l'analyse plus détaillée présentée dans le reste de notre rapport.

La reproductibilité : difficile définition

La littérature académique et nos recherches montrent que, bien que la reproductibilité soit une question qui dépasse les frontières disciplinaires, l'éventail des termes et des interprétations utilisés est incohérent, qui plus est à des niveaux de maturité différents. Cela n'est pas surprenant, car le discours dans ce domaine s'est développé depuis le bas dans différentes communautés. Le paysage fragmenté de la reproductibilité de la recherche rend la communication plus difficile. Il rend également les efforts moins efficaces qu'espéré en raison de la confusion à propos de la définition du terme et d'un manque apparent de points communs entre les disciplines.

« [La confusion autour des définitions] crée certainement des frictions dans le sens où elle ralentit la compréhension de ce que les gens veulent dire, et il y a donc beaucoup de terminologies qui sont générées, donnant l'impression que c'est plus compliqué que ça ne l'est... cette confusion complique tout bonnement le dialogue. » Un fournisseur d'infrastructure

Dans cette étude, nous avons tenu comme relevant de la reproductibilité de la recherche des situations où les données et les procédures partagées par les auteurs d'une étude sont utilisées pour obtenir les mêmes résultats que dans leur travail original. Cette définition, ainsi que notre intérêt spécifique pour la phase de publication et de diffusion du processus de recherche, ont permis aux contributeurs de comprendre immédiatement notre domaine d'intérêt et notre terminologie. Certains contributeurs ont souligné les avantages de disposer d'une définition pratique et spécifique de la reproductibilité, car celle-ci :

- ▢ est utile dans **les actions de sensibilisation et d'enseignement** ; et
- ▢ rend le concept plus **compréhensible** et les problèmes éventuels plus faciles à résoudre.

Cependant, de nombreux entretiens ont également permis de souligner que l'inclusivité est essentielle et qu'une définition trop étroite entraînera un désengagement dans certaines disciplines. En particulier, notre définition est plus parlante dans les sciences quantitatives que dans les sciences qualitatives, principalement parce que ces dernières peuvent s'appuyer sur une série de méthodologies qui font largement appel au jugement et à l'interprétation individuels. Dans ces cas, les concepts de « données » et d'« analyse » sont plus nuancés, et il existe une certaine résistance aux définitions universelles dans ce domaine de recherche.

« La plupart des personnes travaillant dans des domaines [qualitatifs] n'auraient pas l'impression que [la définition de la reproductibilité] les concerne, car le processus analytique est, dans une certaine mesure, intrinsèquement subjectif. Ainsi, tout le concept de reproductibilité ne leur semble pas pertinent si elles se fient à cette définition, à tort ou à raison. » Un chercheur

En outre, certains interlocuteurs ont fait remarquer qu'il est difficile de séparer la reproductibilité de la réplication, de la robustesse et de la généralisation des résultats de recherche. Il est possible d'étudier ces concepts séparément en termes d'infrastructures, d'exigences et de politiques, mais tous s'accordent à dire que tout changement significatif dans les cultures de recherche devra tenir compte de ces considérations (et d'autres encore,

comme l'ouverture, la maturité FAIR,²⁰ l'éthique, l'intégrité, la rigueur, la transparence) de manière concertée.

« Que signifie la reproductibilité, et ne parlons-nous pas en fin de compte d'intégrité et de fiabilité de la recherche ? Si nous nous attachons à une définition particulière de la reproductibilité (et nous pouvons amener les auteurs à le faire), il resterait des problèmes conséquents qui n'ont pas été abordés. » Un éditeur

Un point important soulevé par plusieurs contributeurs est que l'ouverture (comme le partage des données et du code, le libre accès) facilite la reproductibilité, mais ne la garantit pas nécessairement. Par exemple, pour que la recherche soit reproductible, toutes les données et tous les fichiers doivent être publiés ouvertement, mais aussi clairement séparés, étiquetés et documentés, en utilisant des approches automatisées autant que possible pour (i) éviter l'intervention manuelle ; et (ii) renforcer la standardisation.²¹

Bien que la promotion des meilleures pratiques sous l'égide de la science ouverte soit effectivement un point de départ important, nous soulignons la différence entre deux activités clés :

- ▶ **le partage** des résultats de la recherche, p. ex., un article et les données, le code ou les logiciels sous-jacents (le cas échéant) et une méthodologie documentée détaillée ; et
- ▶ **l'évaluation et le test** de ces résultats quant à leur reproductibilité.

Cette distinction montre clairement que les rôles et les responsabilités (voir section 3) ne sont pas évidents lorsqu'il s'agit de pratiques de publication reproductibles : p. ex., les chercheurs peuvent s'efforcer de garantir que leurs propres résultats sont reproductibles, mais aussi vérifier les articles de quelqu'un

²⁰ FAIR data are data which meet principles of Findability, Accessibility, Interoperability, and Reusability Voir : Go Fair. (n.d.). FAIR Principles. <https://www.go-fair.org/fair-principles/>

d'autre (p. ex., en tant que pairs examinateurs ou réutilisateurs de données ou de code).

« Si vous partagez un protocole et des données et que les deux sont totalement ouverts, alors il se peut qu'on ne puisse pas toujours reproduire l'étude... Vous ne pouvez pas tout mettre par écrit dans un protocole, c'est toujours une simplification de la réalité... Il y a des limites à ce qui peut être partagé, même si vous faites de votre mieux. » Un chercheur

Moteurs et avantages perçus

Au cours de la dernière décennie, un certain consensus s'est dégagé sur le fait qu'une bonne recherche doit être reproductible, ce qui est considéré comme une exigence minimale pour la science plutôt que comme un objectif abstrait. Notre étude a révélé que, d'une manière générale, il n'est pas toujours facile de reconnaître ou d'expérimenter immédiatement les avantages des pratiques de publication reprocompatibles, mais les contributeurs s'accordent à dire que de nouveaux efforts dans ce sens sont stratégiquement importants pour l'avenir du monde de la recherche et pour renforcer notre confiance envers elle.

« Je pense que si l'on enseigne aux gens à mettre en place des flux de travail reproductibles, cela leur profite aussi en premier lieu. Si je dois toucher au même projet trois ans plus tard, je risque d'avoir oublié ce que j'ai fait précisément, et s'il y a une bonne documentation, tout est là et je gagne beaucoup de temps. » Un chercheur

Parmi les chercheurs qui ont contribué à cette étude, la plupart sont motivés pour s'assurer que leur travail est reproductible par la conviction personnelle qu'il s'agit d'une caractéristique fondamentale de la recherche

²¹ Fortunato, L., & Galassi, M. (2021). The case for free and open source software in research and scholarship. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2197). <https://doi.org/10.1098/rsta.2020.0079>

universitaire. Il est toutefois largement reconnu que l'investissement en temps exigé par les pratiques de publication reproductibles n'est pas encore couramment encouragé ou récompensé : si cela ne devait pas changer, la reproductibilité risque de rester une pratique limitée et de n'intéresser qu'un petit sous-ensemble de personnes motivées (voir section 4).

« Je pense qu'il y a une incitation intrinsèque à vouloir que votre travail soit correct. Mais à l'heure actuelle, cela signifierait peut-être doubler le temps que nous consacrons à notre travail, et cela ne nous y encourage pas. » Un chercheur

« Cinq raisons égoïstes de travailler de manière reproductible » ont été émises comme hypothèses dans la littérature,²² et elles trouvent un écho dans les résultats de nos entretiens et groupes de discussion :

- ▶ la reproductibilité permet d' « éviter le désastre » : des flux de travail de recherche transparents et ouverts peuvent aider les chercheurs à éviter les erreurs d'analyse et d'interprétation des données ayant un impact, tant avant qu'après la soumission d'un article.
- ▶ La reproductibilité facilite la rédaction d'articles : un processus de recherche bien documenté signifie qu'il sera plus facile d'écrire à son sujet, car on aura davantage confiance en des chiffres, des figures et des tableaux entièrement à jour.
- ▶ La reproductibilité aide les évaluateurs à voir les choses avec le point de vue de l'auteur : en ayant accès à une gamme d'objets de recherche, les pairs examinateurs peuvent comprendre pleinement ce que les auteurs ont fait, ce qui conduit à une meilleure appréciation des résultats et des conclusions d'une étude.

- ▶ La reproductibilité permet la continuité du travail : une bonne documentation des méthodes et des données et des codes de recherche devrait résister à l'épreuve du temps et aider les auteurs et les réutilisateurs à exploiter les résultats publiés.
- ▶ La reproductibilité permet d'asseoir sa réputation : bien qu'il existe peu d'exemples concrets à cet égard, on commence à penser que la publication de recherches reproductibles peut permettre à des chercheurs d'être reconnus au sein de certaines communautés disciplinaires.

Il est important de noter que les avantages perçus énumérés ci-dessus s'appliquent au niveau individuel pour le chercheur, mais que leur impact positif potentiel va bien plus loin : la disponibilité d'objets de recherche ouverts et/ou FAIR, y compris la documentation détaillée, peut ouvrir les portes à des recherches nouvelles et inattendues que les auteurs n'avaient pas envisagées au départ^{23,24} Cela peut s'appliquer aussi à l'utilisation de techniques émergentes nécessitant de grandes quantités de données en entrée, comme dans le cas de l'intelligence artificielle et l'apprentissage automatique, qui se justifient par le fait que « le système scientifique est en train de passer d'un système de rareté de la donnée à celui de données en saturation ». ²⁵

« Si vous travaillez de manière reproductible, et que d'autres personnes dans votre domaine prennent ces données et montrent que vous travaillez de cette manière, ils vont davantage se fier à vos données. Cela peut se traduire par des offres de collaboration, peut conduire les gens à réutiliser vos données plus souvent que les données de quelqu'un d'autre... et dans ce cas, cela se traduira par plus de citations. » Un organisme de recherche

²² Markowetz, F. (2015). Five selfish reasons to work reproducibly. *Genome Biology*, 16(1). <https://doi.org/10.1186/s13059-015-0850-7>

²³ European Commission. (n.d.). Facts and Figures for Open Research Data. https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science/open-science-monitor/facts-and-figures-open-research-data_en

²⁴ Department for Business, Energy & Industrial Strategy. (2019). Realising the potential: Final report of the Open Research Data Task Force. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/775006/Realising-the-potential-ORDTF-July-2018.pdf

²⁵ Burgelman, J. C., Pasco, C., Szkuta, K., Von Schomberg, R., Karalopoulos, A., Repanas, K., & Schoupe, M. (2019). Open Science, Open Data, and Open Scholarship: European Policies to Make Science Fit for the Twenty-First Century. *Frontiers in Big Data*, 2(43), 1-6. <https://doi.org/10.3389/fdata.2019.00043>

TABLEAU 1. PRINCIPAUX OBSTACLES AUX PRATIQUES DE PUBLICATION REPROCOMPATIBLES.

Récompenses et incitations	
<input type="checkbox"/>	Les incitations actuelles ne sont pas propices à la recherche reprocompatible.
<input type="checkbox"/>	Les mesures d'évaluation actuelles sont axées sur la quantité et l'impact plutôt que sur la qualité et la reproductibilité.
Standards de publication et de rédaction de rapports	
<input type="checkbox"/>	Le partage du code, des données et des méthodes n'est pas obligatoire.
<input type="checkbox"/>	Les méthodes ne sont pas suffisamment détaillées, notamment en raison des limitations du nombre de mots/pages.
<input type="checkbox"/>	Les normes de rédaction des rapports pour les auteurs et les pairs examinateurs sont insuffisants ou incohérents, y compris entre les disciplines.
Compétences techniques et analytiques	
<input type="checkbox"/>	La formation et le mentorat sont limités, notamment en ce qui concerne l'analyse informatique, le codage et les nouvelles techniques.
<input type="checkbox"/>	Tous les chercheurs ne sont pas familiarisés ou compétents à un même niveau en informatique et en logiciels aptes à la reproductibilité.
Infrastructure technique	
<input type="checkbox"/>	La connectivité entre l'infrastructure de communication savante et les flux de travail et les outils des chercheurs est limitée.
<input type="checkbox"/>	La collaboration et la communication entre les réservoirs de recherche et entre les entrepôts et les revues ne sont pas encore bien développées. Les outils permettant la connexion des objets de recherche ne le sont pas non plus.
Obstacles pratiques et culturels	
<input type="checkbox"/>	Certains auteurs peuvent être réticents (position individuelle) ou dans l'incapacité (obstacles pratiques ou juridiques) de partager des données pour de nombreuses raisons, notamment la propriété des données, leur caractère sensible, l'éthique et la confidentialité, la propriété intellectuelle ou la crainte d'une utilisation abusive.

Obstacles potentiels

Nos travaux ont mis en lumière une série d'obstacles majeurs à la publication de résultats de recherche reprocompatibles (voir tableau 1). Bien que les chercheurs soient généralement conscients des attentes existantes en matière de qualité du travail scientifique et de pratiques de recherche ouvertes, ils sont confrontés à ces conflits de priorités et à un manque de temps et d'incitations pour porter leur attention sur les flux de travail de recherche reproductible. Les chercheurs interrogés ont souligné que, bien que les avantages de l'investissement dans les pratiques reproductibles soient appréciés et compris, leur mise en place est ralentie par la culture du « publish or perish » (publier ou périr), qui met les chercheurs sous pression pour publier régulièrement dans des revues savantes et pouvoir progresser dans leur carrière académique.

Cela les oblige souvent à donner la priorité au fait de publier rapidement, ce qui peut avoir pour conséquence de partager peu leurs données (voire pas du tout), donc de mettre à mal la reproductibilité des travaux en question.

« Le principal obstacle à l'heure actuelle est probablement le manque de temps. Au départ, il faut beaucoup de temps pour mettre en place un cycle de recherche reproductible et pour préparer les matériaux de manière à ce qu'ils aient un sens pour les autres et soient partageables. D'après mon expérience personnelle, il faut investir son temps. » Un chercheur

Sans surprise, nos entretiens ont également montré que les articles publiés continuent d'être les débouchés de la recherche les plus récompensés, ce qui ne fait qu'augmenter les préoccupations énoncées dans le tableau 1. Des initiatives visant à remédier à ce problème

sont en cours d'élaboration, et nous en parlons plus en détail dans la section 4.

« Les incitations à la recherche restent indéfectiblement l'article publié. Ainsi, le comportement, les habitudes, au cours des cent dernières années, ont consisté à se concentrer sur la publication de l'article, puis à passer à autre chose. Tant que l'on continuera à accepter cela, je pense que nos tentatives d'encourager la reproductibilité ou nos efforts pour qu'elle se réalise dans une énorme proportion de la littérature porteront peu. » Un éditeur

Sur le plan pratique, nous avons constaté que les chercheurs ont des compétences variables (souvent selon la discipline) lorsqu'il s'agit de publier des recherches reproductibles. Ces compétences se développent à des rythmes différents (toujours selon les disciplines), généralement sous l'effet d'exigences méthodologiques propres à chaque domaine.

L'adoption des outils et services numériques qui sous-tendent les pratiques de publication reprocompatibles varie considérablement, et nos constatations suggèrent que les chercheurs des disciplines quantitatives semblent plus familiers avec les produits existants car ils sont actuellement utilisés dans leur processus de recherche. Il convient de noter que les outils qui, en principe, peuvent permettre la reproductibilité tout au long du processus de recherche et au moment de la publication, ne sont pas toujours utilisés à ces fins. P. ex., GitHub peut servir de socle potentiel pour des cycles de recherche reproductibles, mais bien souvent il est utilisé comme une plateforme de collaboration plutôt que pour partager et vérifier ses données et son code au moment de la publication, parallèlement à l'intégration du dépôt de celle-ci dans Zenodo.²⁶ Dans la section 5, nous explorons une gamme plus large de technologies qui peuvent faciliter les pratiques de publication reprocompatibles.

« Si vous faites de la recherche informatique, les outils existent bien, mais ils sont également dispersés et évoluent différemment selon la discipline... Pour un autre type d'étude, p. ex. en archéologie, la recherche informatique est beaucoup moins structurée à cause d'outils moins développés. C'est donc spécifique à chaque discipline, mais la situation est définitivement en train d'évoluer. » Un organisme de recherche

Les difficultés liées au partage des données sont particulièrement importantes et ont des répercussions sur l'ensemble du cycle de vie de la recherche : l'absence de données (FAIR et, dans de nombreux cas, ouvertes) à la base des analyses constitue un obstacle à l'exécution des contrôles de reproductibilité et, plus largement, empêche la réutilisation. Dans certains cas, le partage des données est en effet complexe, p. ex. lorsque des considérations éthiques interviennent ou lorsque les données sont commercialement sensibles (voir section 2).

« Les données médicales présentent beaucoup plus d'obstacles à gérer. Certaines données dans certaines disciplines ne peuvent tout simplement pas être mises à disposition. Les données de différents domaines sont plus ou moins faciles à rendre reproductibles en raison de la dynamique des types de matériaux avec lesquels ils travaillent. » Un éditeur

L'impact des différences disciplinaires

Jusqu'à présent, notre discussion n'a fait que survoler la question de l'impact potentiel des différences disciplinaires sur la capacité d'un chercheur à mettre en œuvre des pratiques de recherche reproductible. Notre analyse documentaire et nos entretiens indiquent que nous pouvons être en présence de différentes problématiques en fonction (i) des méthodes de

²⁶ GitHub. (n.d.). Making your Code Citable. <https://guides.github.com/activities/citable-code/>

recherche utilisées ; et (ii) de la discipline d'un chercheur (bien que les deux soient souvent étroitement liées).

« La plupart de mes travaux sont quantitatifs et consistent en des essais contrôlés randomisés, mais je comprends très bien les réticences de certains de mes collègues quand ils évoquent leurs recherches qualitatives. La rédaction de leurs travaux est extrêmement difficile, tout comme la préservation de l'anonymat tout en ménageant la transparence et en permettant aux autres de faire les mêmes déductions à partir de ces données. » Un chercheur

Nous avons constaté que certaines méthodes de recherche empirique se prêtent mieux aux flux de travail reproductibles,²⁷ comme le développement de logiciels et les expériences standardisées ou semi-standardisées. Ces méthodes sont typiques des disciplines quantitatives, où les objets de recherche tels que les tableaux de données sont courants et peuvent être générés par des moyens semi- ou entièrement automatisés pour l'analyse, le partage, l'inspection et la réutilisation.

En revanche, avoir des flux de travail reproductibles devient plus complexe lorsqu'un élément qualitatif important est présent, lorsque la recherche n'est pas standardisée ou lorsqu'elle repose sur des matériaux rares, uniques, non-pérennes ou inaccessibles. Dans ces cas, la collecte, la conservation, la documentation et/ou le partage des données prendront plus de temps. Ainsi, des considérations comme la protection des données et l'éthique peuvent entrer en ligne de compte. Ainsi, un

des obstacles courants consiste à « dépersonnaliser les données personnelles tout en [...] limitant la perte de richesse des informations fournies ».²⁸

« Dans mon domaine, nous devons faire preuve d'une grande prudence, car nous traitons des données provenant de personnes volontaires et, évidemment, nous les désidentifions autant que possible. Par exemple, un scanner cérébral peut identifier une personne de manière unique. Donc parfois, nous traitons des données qui sont intrinsèquement difficiles à anonymiser. » Une société savante

Nous reconnaissons que la présente section ne mentionne pas les méthodes non empiriques qui s'appuient sur des observations personnelles, des réflexions sur l'actualité, et/ou l'autorité ou l'expérience de l'auteur.²⁹ Dans ces cas, les personnes interrogées se demandent si la reproductibilité de la recherche est même souhaitable ou applicable en première intention. Certains contributeurs ont notamment souligné qu'une approche unique de la reproductibilité n'est pas appropriée et qu'il est essentiel de tenir compte de la diversité des méthodes et des disciplines de recherche pour faire avancer ce débat.

« Un article sur la religion ou la philosophie n'est pas nécessairement fondé sur des preuves ou des données spécifiques. L'auteur peut se contenter de partager sa nouvelle théorie et ne rien faire pour la prouver ou la réfuter. La signification du terme « reproductibilité » peut donc changer ou la question ne se pose même pas dans certains cas. » Un fournisseur d'infrastructure

²⁷ Penders, B., Holbrook, J., & de Rijcke, S. (2019). Rinse and Repeat: Understanding the Value of Replication across Different Ways of Knowing. *Publications*, 7(3), 52. <https://doi.org/10.3390/publications7030052>

²⁸ Kurapati, S., & Teperek, M. (2019). What does reproducibility mean for qualitative research? <https://openworking.wordpress.com/2019/02/11/what-does-reproducibility-mean-for-qualitative-research/>

²⁹ Dan, V. (2017). Empirical and Nonempirical Methods. *The International Encyclopedia of Communication Research Methods*. 1-3. <https://doi.org/10.1002/9781118901731.iecrm0083>

Les investissements en infrastructure de données du Conseil de la recherche économique et sociale (ESRC)

L'ESRC, qui fait partie de UK Research and Innovation, est la plus grande organisation publique du Royaume-Uni pour le financement de la recherche sur les sujets économiques et sociaux, centrée sur la recherche ayant un impact sur les entreprises, le secteur public et la société civile. L'une de ses missions est de « développer et soutenir l'infrastructure nationale de données qui sous-tend une recherche de haute qualité », ce qui se concrétise, notamment, dans des subventions accordées à un large éventail d'infrastructures de données.

L'ESRC a réalisé d'importants investissements dans des infrastructures de données (p. ex., l'Administrative Data Research UK, le UK Data Service, l'Urban Big Data Centre et le Consumer Data Research Centre), et celles-ci hébergent toutes des installations et des technologies qui protègent les identités et permettent un accès sécurisé et contrôlé aux données.

Ces infrastructures sont essentielles pour permettre une publication reprocompitable dans toutes les disciplines où la sensibilité et la confidentialité des données sont un enjeu important pour les particuliers comme pour les entreprises. Dans le cadre de la Covid-19, le UK Data Service a mis en place des mesures supplémentaires (avec l'accord des propriétaires des données) pour permettre l'utilisation de ses installations Secure Lab à domicile. Cela démontre que, grâce à l'innovation, il existe des solutions pour que des utilisateurs puissent accéder à distance à des données sensibles et confidentielles tout en maintenant une protection renforcée. Autre exemple : l'Administrative Data Research UK, qui a investi dans SafePods - un réseau d'environnements sûrs indépendants pour assurer un accès à des données sécurisées.

De manière plus générale, l'ESRC impose également des règles en matière de partage des données à ses bénéficiaires, en partant du principe que « les données sont les éléments principaux du patrimoine de la recherche économique et sociale ». Ces prescriptions portent aussi sur la documentation et les métadonnées afin de « fournir aux utilisateurs secondaires les informations essentielles pour comprendre les données de manière indépendante, permettre la découverte et la réutilisation scientifique » - des comportements qui sont très favorables à la recherche reproductible et aux pratiques de publication.

Enfin, nous notons le rôle important que les bailleurs de fonds peuvent jouer en investissant dans des infrastructures dédiées pour soutenir les utilisateurs de données et promouvoir des comportements positifs en matière de pratiques de publication reprocompatibles : le UK Data Service a organisé l'événement « Love Your Code » dans le cadre de la semaine Love Data 2020, axé sur l'échange d'informations et de connaissances autour du partage de codes et de ses avantages.

3. Acteurs, rôles et responsabilités

Les rôles et les responsabilités pour la reproductibilité sont complexes, et la coopération entre acteurs est essentielle. Les chercheurs et les groupes de recherche peuvent concourir activement à des pratiques reprocompatibles de publication en adoptant des processus adaptés, en étant soutenus par leurs institutions et en utilisant les infrastructures de recherche qui se déploient. Les milieux de la recherche peuvent favoriser l'élaboration et la transmission de leurs usages aux revues et aux organismes de financement de la recherche, au sein desquels des orientations, y compris concernant la phase de publication, sont en train d'être définies actuellement.

Exploration des rôles et des responsabilités

Dans le cadre de notre projet, nous avons tenté d'analyser les rôles et les responsabilités des différents acteurs en lien avec les pratiques de publication compatibles avec la reproductibilité. Le tableau 2 présente un récapitulatif de nos observations, et ce qui ressort de plus marquant, c'est la nécessité d'une large collaboration entre intervenants pour réaliser l'ensemble des activités identifiées.

Dans la présente partie, nous dressons un état des lieux des relations entre acteurs pour donner à voir les actions qu'ils pourraient entreprendre aux niveaux micro, méso et macro (selon le cadre de la Science ouverte de Knowledge Exchange) pour que la question de la reproductibilité aille au-delà du stade du discours.

Niveau micro : chercheurs et groupes de recherche

Nos consultations nous ont livré le constat d'un large consensus autour du rôle clé que les chercheurs ont à jouer pour la reprocompatibilité des pratiques de publication. Cela s'explique simplement par le fait que ce sont eux qui exercent la responsabilité de conce-

voir, réaliser et diffuser les travaux de recherche, et qu'ils sont les seuls à avoir facilement accès à tous les objets de recherche concernés (et à les comprendre dans le détail).

Individuellement, ou avec leurs groupes de recherche, les chercheurs peuvent porter les pratiques de publication reprocompatibles en encourageant et en appliquant des cycles de travaux qui correspondent à cela. Cela peut comprendre notamment une gestion des données de recherche, leur collecte et leur conservation et celles des codes, ainsi que le partage en parallèle des objets de recherche qui se rapportent à leurs articles. Ils peuvent également jouer un rôle lorsqu'ils interviennent en tant qu'évaluateurs entre pairs, p. ex. dans les cas où des revues affichent des consignes de reproductibilité (voir ci-dessous).³⁰

Ainsi que nous l'abordons au terme de la présente section, les microacteurs peuvent également avoir une place essentielle quand s'élaborent des politiques : ils sont les seuls à pouvoir aider les acteurs de niveaux méso et macro à comprendre les impératifs spécifiques qui se présentent dans des disciplines et des contextes de recherche différents.

³⁰ Turkyilmaz-van der Velden, Y., Dintzner, N., & Teperek, M. (2020). Reproducibility Starts from You Today. *Patterns*, 1(6). <https://doi.org/10.1016/j.patter.2020.100099>

« Dans l'idéal, la responsabilité revient aux auteurs de s'assurer que leur travail est reproductible. Et c'est ce qu'ils devraient faciliter. Dans une certaine mesure, la situation veut que ce soit les auteurs qui soumettent leurs travaux et tout le monde tient pour acquis qu'ils ont fait leur métier comme il faut. » Un éditeur

Niveau méso : usages disciplinaires

Nos entretiens ont fait ressortir la nécessité que les communautés de recherche et les disciplines s'approprient l'élaboration des prescriptions portant sur les pratiques reproductibles de publication. Il s'agit d'un aspect essentiel, car parmi nos contributeurs, éditeurs comme financeurs, partagent le

même sentiment de préoccupation à imposer d'en haut des injonctions qui ne correspondraient pas à des pratiques disciplinaires.

L'élaboration d'usages disciplinaires (à l'instar des initiatives issues des sociétés savantes), constitue un outil très puissant, car dans les faits ceux-ci font office d'orientations qui viennent de l'intérieur dans différents domaines de la recherche. Toutefois, pour qu'une discipline définisse de nouveaux usages, il faut que les avantages qui en découlent soient manifestes : en ce qui concerne la reproductibilité, identifier des avantages spécifiques et mesurables peut s'avérer difficile aux niveaux micro ou méso (voir section 2), ce qui peut entraver des avancées au sein des communautés disciplinaires.³¹³²³³

TABLEAU 2. RÔLES ET RESPONSABILITÉS POSSIBLES DANS LA PUBLICATION DE RÉSULTATS DE RECHERCHE REPRODUCTIBLES.

Niveau	Groupes d'acteurs	Responsabilités détaillées découlant de notre consultation
Micro	Chercheurs et groupes de recherche	<ul style="list-style-type: none"> • Favoriser et appliquer des flux de travail reproductibles, y compris la collecte et la curation de données et de codes • Partager des objets de recherche pertinents (numériques et physiques) parallèlement aux publications • Soumettre des articles à des tests de reproductibilité lorsque l'on intervient en tant que pair évaluateur (le cas échéant, notamment parce que les mécanismes de récompense actuels n'encouragent pas cette démarche)
	Communautés et disciplines	<ul style="list-style-type: none"> • Discuter, tester et affiner les coutumes et les standards disciplinaires relatifs aux pratiques de publication reproductibles • Faire connaître les préconisations relatives à ces pratiques de publication aux acteurs de niveau supérieur
Méso	Sociétés savantes	<ul style="list-style-type: none"> • Favoriser les forums de discussion sur la reproductibilité • Offrir une formation spécifique à la discipline afin de répondre aux attentes des éditeurs et des organismes de financement de la recherche
	Réseaux (inter)nationaux de reproductibilité	<ul style="list-style-type: none"> • Créer des passerelles entre les discussions qui se mènent entre membres de communautés (débat relativement récents et disparates d'une discipline à l'autre) et d'autres groupes d'intervenants • Favoriser la collaboration par-delà les frontières nationales • Pourvoir à un certain degré de coordination dans le contexte plus large des pratiques de science ouverte
	Opérateurs de recherche	<ul style="list-style-type: none"> • Définir des orientations politiques pour le personnel (p. ex. partage d'objets de données, étendue de la vérification attendue, différences disciplinaires) dans le contexte plus large des pratiques de science ouverte • Sensibiliser aux principales prescriptions découlant des attentes politiques

³¹ Association for Computing Machinery (ACM). (2021). Practices Working Group Blog. <https://reproducibility.acm.org/blog/>

³² The Academy of Medical Sciences. (n.d.) Reproducibility and reliability of biomedical research. <https://acmedsci.ac.uk/policy/policy-projects/reproducibility-and-reliability-of-biomedical-research>

³³ National Academies of Sciences, Engineering, and Medicine. (2019). Reproducibility and Replicability in Science. The National Academies Press. <https://doi.org/10.1016/j.patter.2020.100099>

Ni-veau	Groupes d'acteurs	Responsabilités détaillées découlant de notre consultation
		<ul style="list-style-type: none"> ● Apporter un soutien sous une forme associant de manière pertinente intendants de données, curateurs d'objets de recherche ou bibliothécaires de spécialité disciplinaire ● Offrir une formation générale comme spécifique à la discipline (à l'intention des étudiants et du personnel) pour répondre aux attentes des éditeurs et des organismes de financement de la recherche ● Offrir l'accès à une combinaison adéquate d'infrastructures numériques et physiques pour soutenir les flux de travail de recherche reproductibles ● Financer les outils liés à la reproductibilité pendant leur phase de démarrage ou pilote ● Élaborer et mettre en œuvre des mécanismes de récompense pour les pratiques de publication reprocompatibles en les inscrivant dans le cadre plus large des pratiques de science ouverte
	Éditeurs et revues académiques	<ul style="list-style-type: none"> ● Définir des orientations et attentes destinées aux auteurs (p. ex., partage d'objets de données, étendue des vérifications requises, différences disciplinaires) dans le contexte plus large des pratiques de science ouverte ● Sensibiliser aux principales prescriptions découlant des attentes politiques ● Élaborer et mettre en œuvre des mécanismes de récompense pour les pratiques de publication reprocompatibles ● Mettre en œuvre les orientations politiques, p. ex. au moyen de listes de contrôle, de lignes directrices et de standards relatifs à la reproductibilité de la recherche à l'intention des auteurs, des rédacteurs et des examinateurs ● Donner des responsabilités pour le contrôle de la reproductibilité des manuscrits soumis ou acceptés aux personnels de rédaction en charge des données ou de la reproductibilité et les y former ● Inciter les pairs examinateurs à définir des attentes claires concernant les rôles et les responsabilités ● Collaborer avec les fournisseurs d'infrastructures pour mettre en œuvre des solutions pratiques (p. ex., intégration par des revues des carnets de recherche Binder ou Jupyter, le cas échéant)
	Fournisseurs d'infrastructures ³⁴	<ul style="list-style-type: none"> ● Inciter un plus grand nombre de disciplines à participer à l'élaboration d'un ensemble plus exhaustif d'outils et de flux de travail (p. ex., au-delà des disciplines fortement informatisées) ● Procéder à des contrôles de reproductibilité en partenariat avec des auteurs, des éditeurs et des organismes de financement de la recherche ● Élaborer des modèles économiques durables pour servir le marché académique, y compris, le cas échéant, avec des subventions de la part d'organismes de financement de la recherche et/ou d'opérateurs de recherche ● Travailler avec les éditeurs et les revues pour mettre en œuvre des solutions pratiques (p. ex., faire en sorte que les revues utilisent des carnets de recherche Binder ou Jupyter, le cas échéant)
Macro	Organismes de financement de la recherche	<ul style="list-style-type: none"> ● Soutenir les communautés de recherche dans la discussion, l'expérimentation et l'affinement de prescriptions spécifiques en matière de reproductibilité de la recherche (p. ex., le développement d'instruments de financement à cet effet) ● Définir des orientations à l'intention des bénéficiaires de subventions (p. ex., partage d'objets de données, étendue des vérifications requises, différences disciplinaires) dans le contexte plus large des pratiques de science ouverte ● Sensibiliser aux principales prescriptions découlant des attentes politiques ● Élaborer et mettre en œuvre des mécanismes de récompense pour les pratiques de publication reprocompatibles ● Couvrir (une partie) des dépenses engendrées par des pratiques de publication reprocompatibles lorsque celles-ci sont en adéquation avec les orientations politiques ● Opérer un suivi et une évaluation de la mise en œuvre des politiques ● Financer les outils liés à la reproductibilité pendant leur phase de démarrage ou pilote
	Acteurs politiques	<ul style="list-style-type: none"> ● Définir les orientations politiques au niveau systémique et coordonner les initiatives nationales dans le contexte plus large des pratiques de la science ouverte ● Définir les infrastructures numériques et matérielles financées par les pouvoirs publics comme étant une priorité stratégique, afin d'assurer la durabilité à long terme et d'éviter la captivité commerciale auprès de fournisseurs

³⁴ Il convient de noter que les fournisseurs d'infrastructures sont examinés séparément dans la section 5.

Pour autant, certaines disciplines vont déjà très loin dans la mise en œuvre de flux de travail reproductibles tout au long du processus de recherche, y compris dans le domaine de la publication. Ainsi, la physique des particules est à l'avant-garde de cette philosophie, et il est intéressant de comprendre les principales raisons de cette situation :

- ▢ Les infrastructures comme celles construites par l'Organisation européenne pour la recherche nucléaire (communément appelée le CERN) sont très coûteuses et donc partagées. Le partage des données est rendu obligatoire par les politiques des organismes et la vérification des travaux des autres est un mode de fonctionnement habituel.
- ▢ Les expériences exigent une large collaboration et une grande confiance, ce qui implique que la transparence en ce qui concerne le code, les données et les méthodes soit un point fort essentiel.

« Une expérience pouvant solliciter plus de 5 000 physiciens implique des groupes qui manipulent les données et les traitent, ensuite nous n'avons plus qu'à utiliser tout ce matériau. Il y a tant de gens qui contribuent à différentes dimensions et qui partagent ouvertement leur travail : c'est une réalisation collaborative. » Un chercheur

Dans de nombreux autres domaines scientifiques, faire adopter des pratiques de publication reprocompatibles n'est pas chose aussi

facile : du moins pour certains, elles sont perçues comme exigeant un gros investissement sans aucun bénéfice immédiat. Bien que des champions isolés et de nombreuses communautés de pratique s'efforcent d'ouvrir la voie au changement culturel, leur essor est encore loin de se généraliser. C'est pourquoi les politiques qu'adoptent les éditeurs et les agences de financement de la recherche peuvent jouer un rôle essentiel pour créer un équilibre entre initiatives venant de la base et en local et pour instaurer des attentes fortes pour les acteurs à tous les niveaux.³⁵³⁶

« Ces initiatives sont importantes, en ce sens qu'elles favorisent les activités ascendantes et voilà qui est très important, mais in fine, ce n'est pas que d'activités venant de la base dont nous avons besoin... le rôle de ces activités venant du bas, c'est d'en susciter qui viennent d'en haut. » Un chercheur

Niveau méso : réseaux internationaux de reproductibilité

Comme nous l'avons déjà indiqué précédemment, la reproductibilité se présente sous des formes différentes selon les disciplines. En conséquence, non seulement les exigences varient, mais les débats n'en sont pas partout au même stade de maturité. En un sens, il est inévitable que ces discussions se déroulent en parallèle plutôt qu'en commun, mais des initiatives telles que les réseaux de reproductibilité en Grande-Bretagne³⁷, Allemagne³⁸,

³⁵ Willis, C., & Stodden, V. (2020). Trust but Verify: How to Leverage Policies, Workflows, and Infrastructure to Ensure Computational Reproducibility in Publication. *Harvard Data Science Review*, 2(4). <https://doi.org/10.1162/99608f92.25982dcf>

³⁶ Westwood, S. (Twitter). (2021). I'm curious, what got you involved in Open Science/ replication crises? <https://web.archive.org/web/20210708162443/https://twitter.com/westwoodsam1/status/1382694739644923915>

³⁷ UK Reproducibility Network. (n.d.). Welcome to the UK Reproducibility Network. <https://www.ukrn.org/>

³⁸ Réseau allemand de reproductibilité. (n.d.). The German Reproducibility Network. <https://reproducibilitynetwork.de/>

Suisse³⁹, Australie⁴⁰, Slovaquie⁴¹ et Italie⁴² pourraient ressortir comme déterminantes. La plupart de ces réseaux sont relativement jeunes et leur impact est difficile à évaluer pleinement à ce stade, mais ils ont le potentiel de pouvoir :

- ▶ créer des liens entre les débats qui se mènent entre membres de communautés (débats relativement récents et disparates d'une discipline à l'autre) et d'autres groupes d'intervenants ;
- ▶ favoriser la collaboration par-delà les frontières nationales ; et
- ▶ offrir un certain degré de coordination.

« Certains [groupes d'intervenants] n'y travaillent pas du tout, simplement faute d'intérêt, d'autres sont très progressistes, et essaient vraiment d'entraîner les autres parce qu'ils voient ce que cela apporte d'essayer de s'y mettre. Ceux qui sont les plus actifs fonctionnent en mode collaboratif, en particulier ceux qui travaillent en interdisciplinarité. » Un fournisseur d'infrastructure

Dans ce contexte, nous constatons qu'une ligne de pensée prend corps à des niveaux nationaux comme internationaux ces toutes dernières années. C'est ainsi que des contributeurs de projet ont souligné par leurs commentaires combien il importe d'attirer l'attention sur des tentatives de reproduction qui auraient échoué dans certaines disciplines, dans une démarche de sensibilisation dans des domaines où pareil aspect n'est pas reconnu et pour y initier un débat sur la reproductibilité.

Globalement, il convient de noter qu'il est difficile d'évaluer l'impact des activités axées sur la reproductibilité indépendamment des autres pratiques scientifiques ouvertes associées : p. ex., la reproductibilité et le partage de données (ouvertes et/ou FAIR) vont de pair, et il serait très difficile de parler de l'une sans aborder l'autre.

³⁹ Réseau suisse de reproductibilité. (n.d.). SWISSRN. <https://www.swissrn.org/>

⁴⁰ Réseau australien de reproductibilité. (n.d.). Welcome to the Australian Reproducibility Network. <http://www.aus-rn.org/>

⁴¹ Réseau slovaque de reproductibilité. (2021). RNSK. <https://www.slovakrn.org/>

⁴² Salvato, G. (Twitter). (2021) The Italian Reproducibility Network, a star is about to be born! Stay tuned for more information. <https://twitter.com/gerardosalvato/status/1417078428285775878>

Le réseau britannique de reproductibilité

Le réseau britannique de reproductibilité (UK Reproducibility Network - UKRN) a été créé en mars 2019 pour rassembler et coordonner les efforts visant à améliorer la qualité de la recherche au Royaume-Uni. Sa structure se traduit en plusieurs composantes, distinctes mais en interaction, à savoir (i) les réseaux locaux de chercheurs ; (ii) les institutions membres ; et (iii) les acteurs externes comme les organismes de financement de la recherche, les éditeurs et d'autres organisations sectorielles. L'UKRN s'est développé rapidement : au moment de la rédaction de ces lignes, il regroupe des réseaux locaux d'une bonne cinquantaine d'institutions britanniques et 20 membres institutionnels.

Une caractéristique essentielle des activités de l'UKRN se matérialise dans le soutien qu'il apporte aux initiatives existantes, dont beaucoup ont été mises sur pied par des chercheurs en début de carrière. En effet, l'UKRN collabore avec les instigateurs de ces initiatives pour accroître leur audience et les promouvoir auprès de ses membres. D'autres réseaux de reproductibilité, sur le modèle de l'UKRN, ont récemment vu le jour dans d'autres pays, et 11 d'entre eux ont été créés à ce jour ou sont en voie de l'être.

En septembre 2021, l'UKRN s'est vu attribuer 4,5 millions de livres sterling du fonds de développement pour la recherche (Research England Development - RED), « afin de permettre au réseau d'accroître ses activités et d'accélérer l'adoption de pratiques de recherche ouvertes dans l'ensemble du secteur ».⁴³

⁴³ UK Reproducibility Network. (2021). Major funding boost for UK's open research agenda. <https://www.ukrn.org/2021/09/15/major-funding-boost-for-uks-open-research-agenda/>

Niveau méso: attentes institutionnelles en matière de reproductibilité de la recherche

Les entretiens que nous avons eus avec les organismes de recherche ont montré qu'à l'heure actuelle, il n'est pas courant qu'ils affichent des orientations spécifiquement consacrées à la reproductibilité de la recherche. Il est toutefois de plus en plus fréquent que la reproductibilité soit mentionnée au passage, dans le contexte d'autres lignes politiques ou prescriptions institutionnelles (à l'instar des organismes de financement de la recherche, voir ci-dessous).

« En science ouverte, nous avons trois politiques différentes. D'abord, nous avons défini notre politique sur les données, puis nous avons adopté celle sur les publications, et ensuite, la dernière en date, celle sur la recherche ouverte. Celle-ci combine les deux autres, mais les associe sous l'angle du processus de recherche vu dans son ensemble. Par exemple, dans la politique de données, nous avons quelques principes FAIR, et c'est là que l'on y retrouve la notion de reproductibilité. » Un organisme de recherche

Les organismes de recherche sont portés à faire des investissements importants lorsque des prescriptions (mandats) sont formulées par les agences de financement ou au niveau politique : par exemple, c'était le cas au Royaume-Uni quand, en 2015, le conseil de recherche en sciences physiques et de l'ingénieur (Engineering and Physical Sciences Research Council - EPSRC) instituait de nouvelles prescriptions en matière de gestion des

données de recherche, ce qui a incité les universités à traiter ce domaine d'un point de vue plus stratégique.⁴⁴

Plus loin dans cette section, nous relevons que, dans la plupart des cas, les pratiques de recherche reproductibles ne figurent pas parmi les prescriptions des agences de financement. Il en résulte qu'il se pourrait que les organismes de recherche continuent à surveiller le paysage et à aborder la question de la reproductibilité de la recherche par le biais d'approches ad hoc et en fonction de leurs stratégies propres et de leurs effectifs en chercheurs. On peut citer comme exemple le Luxembourg Centre for Systems Biomedicine (LCSB) de l'Université du Luxembourg qui a décidé de s'associer à SciCrunch pour améliorer son contrôle interne de la qualité, dans le cadre de son initiative de recherche responsable et reproductible (Responsible and Reproducible Research - R3).^{45,46}

Niveau méso : le rôle des exigences de l'éditeur et leurs lignes directrices sur la reproductibilité

Les éditeurs jouent un rôle important dans l'économie de la réputation qui, pour une grande part, est le socle sur lequel repose l'univers académique. En particulier, ils agissent comme un passage obligé pour les auteurs avant qu'ils puissent faire état de leurs articles parus dans des revues, une monnaie d'échange particulièrement précieuse : étant donné que ce sont les éditeurs, et plus particulièrement les revues, qui détiennent les clés de la réussite d'un chercheur, ils jouissent d'une position de force pour imposer leurs exigences et, encore plus important, être structurants.

⁴⁴ Jisc. (2018). Meeting the requirements of the EPSRC research data policy. <https://www.jisc.ac.uk/guides/meeting-the-requirements-of-the-epsrc-research-data-policy>

⁴⁵ Research Information. (2020). SciCrunch announces Luxembourg collaboration. <https://www.researchinformation.info/news/scicrunch-announces-luxembourg-collaboration>

⁴⁶ Université Du Luxembourg. (n.d.). Responsible and Reproducible Research (R3) https://www.uni.lu/lcsb/research/bioinformatics_core/responsible_and_reproducible_research_r3

Les listes de contrôle ou lignes directrices à l'intention des rédacteurs de revues, des évaluateurs et des auteurs sont des « outils simples et néanmoins puissants » que les revues peuvent utiliser pour encourager des pratiques de publication reproductibles. Tous les acteurs concernés disposeraient ainsi d'un outil pour s'assurer qu'au moins un ensemble donné d'objets de recherche est disponible en parallèle d'un article de recherche, soit au moment de sa soumission, soit au moment de son acceptation. La principale difficulté est d'identifier à quoi ressemblerait un tel ensemble d'objets de recherche, et cela dépend en grande partie des disciplines.⁴⁷ En outre, pareilles listes de contrôle et lignes directrices peuvent être plus ou moins strictement appliquées, selon les us et coutumes en vigueur dans la discipline et les options de la revue : ainsi, la non reproductibilité n'apparaît guère comme un motif légitime de rejet d'un article que dans un nombre restreint de revues. Dans la pratique, cela pourrait se présenter sous la forme d'une décision de la rédaction du type « à réviser et à soumettre à nouveau » après que la revue, le rédacteur de données, le pair réviseur ou des agences externes de vérification ne soient pas parvenus à reproduire les éléments soumis.^{48,49}

« En fait, je suis défavorable aux listes de contrôle ou aux lignes directrices taillées sur mesure par chaque éditeur, parce que les choses seraient beaucoup plus faciles si on disposait de standards communautaires que tous pourraient adopter. » Un éditeur

Les listes de contrôle et les lignes directrices possèdent également un intérêt pédagogique et peuvent concourir à la sensibilisation envers les pratiques de recherche chaque fois qu'elles sont mises en œuvre et partagées avec les futurs auteurs et d'autres intervenants dans le processus de publication.

Bien qu'il soit difficile d'évaluer l'impact des listes de contrôle et des lignes directrices au niveau systémique, nous notons que certaines initiatives progressent à grands pas dans la bonne direction :

- ▶ les lignes directrices TOP fournissent aux organismes de financement de la recherche, aux éditeurs et aux institutions de recherche des options personnalisables pour que ceux-ci élaborent de nouveaux standards correspondant à leurs objectifs, et ces préconisations sont appliquées par plus de 1 100 revues ;^{50,51,52}
- ▶ le réseau EQUATOR comporte une base de données exhaustive et consultable de 470 lignes directrices de présentations des travaux et résultats dans le domaine de la recherche en santé ;⁵³
- ▶ les lignes directrices ARRIVE sont une liste de contrôle des informations qui doivent figurer dans les publications décrivant la recherche sur les animaux, afin de permettre aux lecteurs et aux réviseurs d'examiner les travaux de recherche de manière adéquate, d'évaluer leur rigueur méthodologique et d'en reproduire les méthodes ou les résultats.⁵⁴

⁴⁷ Goeva, A., Stoudt, S., & Trisovic, A. (2020). Toward Reproducible and Extensible Research: from Values to Action. *Harvard Data Science Review*, 2(4). <https://doi.org/10.1162/99608f92.1cc3d72a>

⁴⁸ Office of the AEA Data Editor. (2021). Guidance on testing replicability of code. <https://aeadataeditor.github.io/aea-de-guidance/reproducibility-checks>

⁴⁹ GitHub (n.d.). ReScience C: Frequently Asked Questions. <https://rescience.github.io/faq/>

⁵⁰ Center for Open Science. (n.d.). Top Guidelines. <https://www.cos.io/initiatives/top-guidelines>

⁵¹ Center for Open Science. (n.d.). Sample Implementation of Guidelines for Transparency and Openness Promotion (TOP). *Journal Policies and Practices*. <https://osf.io/edtxm/>

⁵² Yeston, J.S. (2021). Progress in data and code deposition. <https://blogs.science-mag.org/editors-blog/2021/07/15/progress-in-data-and-code-deposition/>

⁵³ EQUATOR Network. (n.d.). Enhancing the QUALity and Transparency Of health Research <https://www.equator-network.org/reporting-guidelines/>

⁵⁴ National Centre for the Replacement, Refinement & Reduction of Animals in Research. (n.d.). ARRIVE Guidelines.. <https://arriveguidelines.org/about>



Politique et protocole sur les vérifications par des tiers au sein de l'American Economic Association (AEA)

L'AEA est une association savante à but non lucratif qui se consacre à l'exposé et à la publication de la recherche en économie. Ses revues suivent une politique stricte de disponibilité des données et des codes qui est indicative de sa volonté de « publier des articles uniquement si les données et les codes utilisés dans l'analyse sont documentés de manière claire et précise et si l'accès aux données et aux codes n'est pas exclusif aux auteurs ». L'AEA emploie également un rédacteur de données en interne, dont la mission est de « concevoir et de superviser la stratégie des revues de l'AEA pour l'archivage et la curation des données de recherche ainsi que la promotion des travaux de recherche reproductibles ».⁵⁵

Pour une part essentielle, l'AEA effectue les vérifications de reproductibilité par elle-même, mais elle peut aussi faire appel à des répliqueurs externes (<https://bit.ly/3G5hNz6>) lorsque, pour diverses raisons, ces vérifications ne peuvent être effectuées en temps opportun. L'AEA fournit des directives claires (<https://bit.ly/3l8oisy>) à l'intention de ses répliqueurs, internes comme externes, qui se composent des étapes suivantes :

- télécharger ce que l'on appelle « l'archive de réplification » de l'auteur ;
- suivre une liste de contrôle pour effectuer une vérification de la reproductibilité ;
- rédiger un compte-rendu à partir des résultats de la vérification précitée (voir un **exemple** : <https://bit.ly/3l8opEu>) ;
- transmettre le rapport au rédacteur de données.

Les répliqueurs sont tenus d'émettre une appréciation claire et fondée sur des preuves de la reproductibilité d'un article, qui peut aller de « reproduction complète » à « impossible de reproduire la plupart ou la totalité des résultats » : ceci va éclairer les décisions du rédacteur de données de l'AEA, qui peut exiger que le ou les auteurs effectuent des modifications et soumettent à nouveau leur travail jusqu'à ce que ladite « archive de réplification » soit acceptée.

En particulier, un rapport d'acceptation découlant du processus de vérification de la reproductibilité que nous venons de présenter constitue l'étape obligatoire pour lancer le processus d'édition qui mène à la publication finale.

Remarque : les formulations de l'AEA tendent à utiliser indifféremment « réplification » et « reproductibilité », mais, dans la pratique, ses préconisations utilisent la définition de la « reproductibilité » adoptée dans le présent rapport.

⁵⁵ Vilhuber, L. (2021). Report by the AEA Data Editor. AEA Papers and Proceedings, 111: 808-17. <https://doi.org/10.1257/pandp.111.808>

Certaines revues sont également réputées pour promouvoir le changement dans le domaine de la reproductibilité, et nous mettrons en avant, p. ex. :

- ▶ les consignes en la matière de l'American Economic Association ;⁵⁶
- ▶ la présence d'une rubrique « résultats reproductibles » dans le Journal of Water Resources Planning and Management (qui fait partie de l'American Society of Civil Engineers, ou ASCE) ;⁵⁷
- ▶ la création d'un nouveau type d'article appelé « Verification Reports » (rapports de vérification) dans la revue Cortex, dont la ligne est l'évaluation des faits avancés dans une recherche publiée grâce à une réanalyse des données originales d'une étude ;^{58,59}
- ▶ les recommandations sur la reproductibilité et sur les carnets de recherche Jupyter établies à destination des auteurs qui publient dans des revues de l'American Geophysical Union ;⁶⁰
- ▶ la revue Image Processing On Line (IPOL), où chaque article contient un passage sur un algorithme et son code source révisé par des pairs, avec une démonstration contrôlée en ligne ;⁶¹
- ▶ la revue ReScience a lancé un « Défi de reproductibilité décennale » (Ten Years Reproducibility Challenge) pour encourager les auteurs à vérifier les résultats produits grâce à du code développé par le passé, et même à s'assurer si celui-ci

peut s'exécuter sur du matériel ou du logiciel d'aujourd'hui et s'il retourne toujours exactement les mêmes résultats ; et⁶²

- ▶ la mise au point d'articles de recherche exécutables (Executable Research Articles - ERA) par eLife and Stencila, qui offrent à des auteurs la possibilité de publier des articles comportant des travaux reproductibles informatiquement qui utilisent des images Docker optimisées (V. section 5),^{63,64}

Niveau macro : le rôle des organismes de financement de la recherche dans la définition des politiques de reproductibilité

Nombre de nos interlocuteurs considèrent les prescriptions des financeurs comme l'instrument essentiel pour faire avancer le dossier de la reproductibilité de la recherche : les agences de financement ont entre leurs mains un pouvoir considérable et, dans la plupart des cas, les chercheurs ne peuvent pas travailler sans leur soutien. Dans ce domaine, l'exemple le plus connu à émaner d'un financeur est le programme « Études de réplication » (Replication Studies) mis en place par le Conseil néerlandais de la recherche (également connu sous le nom de NWO) de 2016 à 2019 et qui accordait des subventions à des études de réplication et de reproduction, sans

⁵⁶ American Economic Association. (n.d.). Policy and Protocol on Third-Party Verifications. <https://www.aeaweb.org/journals/data/policy-third-party>

⁵⁷ Rosenberg, D. E., Jones, A. S., Filion, Y., Teasley, R., Sandoval-Solis, S., Stagge, J. H., Abdallah, A., Castronova, A., Ostfeld, A., & Watkins, D., Jr. (2021). Reproducible Results Policy. Journal of Water Resources Planning and Management, 147(2). [https://doi.org/10.1061/\(asce\)wr.1943-5452.0001368](https://doi.org/10.1061/(asce)wr.1943-5452.0001368)

⁵⁸ Chambers, C. D. (2020). Verification Reports: A new article type at Cortex. Cortex, 129, A1–A3. <https://doi.org/10.1016/j.cortex.2020.04.020>

⁵⁹ Elsevier. (n.d.). Verification Reports – Guide for Authors. https://www.elsevier.com/_data/promis_misc/VR_GuideForAuthors.pdf

⁶⁰ Advancing Earth and Space Science. (n.d.). Data and Software for Authors <https://www.agu.org/Publish-with-AGU/Publish/Author-Resources/Data-and-Software-for-Authors>

⁶¹ Image Processing On Line (IPOL). (n.d.). Image Processing On Line. <https://reproducibilitynetwork.de/>

⁶² ReScience C. (n.d.). Ten Years Reproducibility Challenge. <https://rescience.github.io/ten-years/>

⁶³ eLife. (2021). eLife and Stencila announce roadmap for bringing reproducible publishing to more authors. <https://elifesciences.org/for-the-press/2b653a68/elifescience-and-stencila-announce-roadmap-for-bringing-reproducible-publishing-to-more-authors>

⁶⁴ Stencila. (n.d.). Executable document pipelines. <https://stenci.la/>

que l'on puisse pour autant le ranger parmi les injonctions politiques.⁶⁵

Au sein des organismes de financement de la recherche, le sujet de la reproductibilité vient s'inscrire actuellement dans des politiques et des positions plus larges, et ce thème n'est pas toujours explicitement mentionné. Ainsi, les orientations portant sur d'autres aspects de la recherche ouverte (p. ex., les données ouvertes) peuvent y faire référence. En effet, des contributeurs à des projets ont relevé que les prescriptions de reproductibilité devraient être intégrées dans d'autres choix voisins, car la publication reproductible est un sujet cha-peauté par les thèmes plus larges de la science ouverte et des pratiques de recherche ouvertes. D'autres orientations associées, comme celles liées à l'intégrité de la recherche, peuvent également mentionner la reproductibilité aux côtés de la rigueur et de la transparence.

« De notre côté, nous avons des guides de « bonnes pratiques de la recherche », nous essayons donc d'y intégrer presque tout. Cela passe par l'accès ouvert, la gestion des données, tout ce que nous voulons que les gens pratiquent. Cela nous dispense d'une théorie de volets d'orientations différents, comme c'était le cas par le passé. » Un organisme de recherche

Nos interlocuteurs dans le secteur du financement de la recherche n'ont pas perçu la nécessité d'une action immédiate au niveau macro-économique en ce qui a trait à la reproductibilité de la recherche. Ils ont évoqué que cela était dû en partie à la difficulté de fixer des exigences envers des bénéficiaires de subventions relevant d'une diversité de disciplines, mais beaucoup ont également reconnu

que l'objectif primordial est la science ouverte, la reproductibilité y étant l'une des diverses préoccupations. C'est ainsi qu'on ne sait pas très bien qui doit financer les activités de curation et de partage des données. Il s'agit en l'espèce d'une condition nécessaire (mais non suffisante) pour une publication reproductible, mais sur laquelle les acteurs de l'enseignement supérieur et de la recherche doivent encore se positionner.⁶⁷

Niveau macro : l'intérêt croissant des pouvoirs publics

À ce jour, les acteurs politiques ne se sont guère impliqués de manière significative dans les débats autour de la reproductibilité de la recherche. Toutefois, leur intérêt dans ce domaine prend de plus en plus d'importance consécutivement à des engagements généraux en faveur des pratiques de science ouverte.

C'est le cas, par exemple, du deuxième Plan national pour la science ouverte élaboré par la France qui élargit le champ des engagements antérieurs du pays autour de la science ouverte et « s'inscrit résolument dans une ambition européenne » à cet égard. Le plan comporte des allusions aux pratiques reproductibles dans la publication, notamment en ce qui concerne l'ouverture et la promotion des codes sources, et reconnaît qu'il s'agit là d'une obligation importante pour assurer la reproductibilité et permettre le partage des connaissances.⁶⁸

⁶⁵ Conseil néerlandais de la recherche (NWO). (n.d.). Replication Studies. www.nwo.nl/en/researchprogrammes/replication-studies

⁶⁶ Conseil néerlandais de la recherche (NWO). (2020). Replication Studies third round: repetition of important research. <https://www.nwo.nl/en/news/replication-studies-third-round-repetition-important-research>

⁶⁷ Digital Science, Hahnel, M., Borrelli, L. M., Hyndman, A., Baynes, G., Merce Crosas, Nosek, B., Shearer, K., Selm, M. V., Goodey, G., & Nature Research. (2020). The State of Open Data 2020. <https://doi.org/10.6084/M9.FIGSHARE.13227875.V2>

⁶⁸ Ouvrir la Science. (n.d.). Deuxième Plan national pour la science ouverte. <https://www.ouvrirlascience.fr/deuxieme-plan-national-pour-la-science-ouverte/>

Programme d'études de réplication de NWO

Entre 2016 et 2019, l'agence de financement NWO a mené un programme pilote intitulé « Études de réplication », comprenant des volets « reproduction-réplication à l'aide de données existantes » et « réplication à partir de données nouvelles », le tout doté d'un budget total de 3 millions d'euros à distribuer par le biais de trois appels à propositions. Le programme s'est concentré pour commencer sur la répétition de travaux « piliers » (dits « cornerstone »), c'est-à-dire des recherches publiées qui ont eu un rejaillissement significatif sur la science, la politique ou le débat public. Lors de son dernier cycle, ce programme ne finançait plus que la répétition de travaux de recherche (c'est-à-dire leur réplication), bien que les études portant sur la reproduction restaient éligibles. Ce programme est désormais en cours d'évaluation après achèvement de son dernier cycle.

Les objectifs stratégiques de NWO étaient les suivants : « acquérir une expérience qui permette de comprendre comment la recherche par réplication peut être intégrée efficacement dans les programmes de recherche » et « déboucher sur un éclairage et une réflexion sur les prescriptions que NWO applique à la recherche dans les domaines de la méthodologie et de la transparence ». Tout cela montre combien il importe d'établir des éléments probants et des impératifs auprès de la communauté scientifique et d'examiner attentivement comment ceux-ci peuvent être mis en œuvre dans le contexte des orientations prescrites par un financeur de la recherche.

Nos échanges avec NWO ont souligné que ce type de mécanisme de financement revêt son importance pour sensibiliser, en particulier à court et à moyen terme, et ce alors que les pratiques continuent d'évoluer dans différentes disciplines. Toutefois, il est à prévoir qu'un financement spécifique à la reproductibilité et la réplication devienne superflu à long terme, au fur et à mesure que ces activités se généraliseront et pourront recevoir un soutien par le biais des subventions attribuées aux projets de recherche.

Au Royaume-Uni, la feuille de route pour la recherche et le développement (UK Research and Development Roadmap) couvre les pratiques de publication reproductibles et adopte une position très similaire, prenant l'engagement de rendre la publication ouverte obligatoire et d'encourager le partage ouvert des données, « afin que la reproductibilité soit possible et que les connaissances soient partagées et diffusées de manière collaborative ». Il convient également de relever que la commission science et technologie du Parlement britannique (Science and Technology Committee) a récemment lancé une enquête sur la reproductibilité et l'intégrité de la re-

cherche afin d'examiner ces questions en profondeur et d'évaluer le rôle potentiel de l'agence UKRI (UK Research and Innovation).⁶⁹⁷⁰

Ces deux exemples font ressortir que les décideurs peuvent influencer sur la façon dont l'entreprise de recherche développe des pratiques liées à la reproductibilité. Compte tenu de leur éloignement du travail scientifique au quotidien, les acteurs politiques se concentrent principalement sur la formulation d'attentes de haut niveau et le soutien de leur mise en œuvre via un financement stratégique aux organisations de recherche, aux

⁶⁹ Department for Business, Energy & Industrial Strategy. (2021). UK Research and Development Roadmap. <https://www.gov.uk/government/publications/uk-research-and-development-roadmap/uk-research-and-development-roadmap>

⁷⁰ Parlement britannique. (n.d.). Reproducibility and research integrity. <https://committees.parliament.uk/work/1433/reproducibility-and-research-integrity/>

programmes scientifiques et aux infrastructures numériques clés.

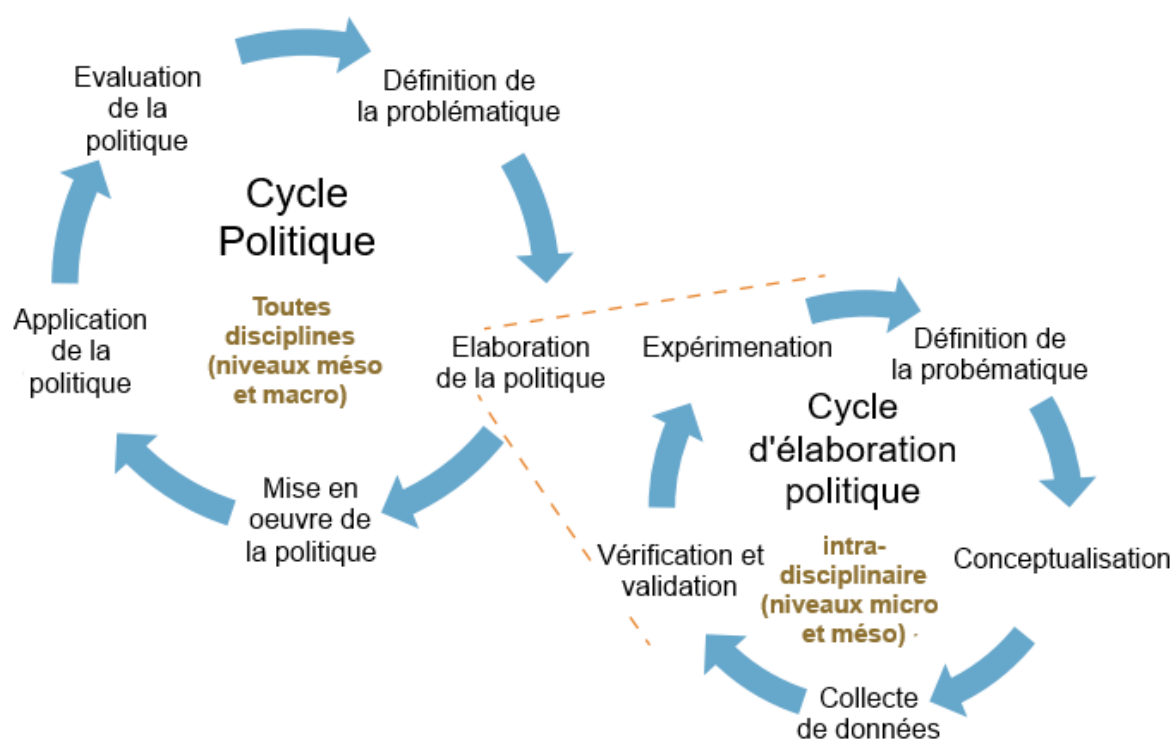
Contributions des niveaux micro, méso et macro dans le cadre de l'élaboration des orientations politiques

Sur le plan politique, le discours autour de la reproductibilité de la recherche, y compris au stade de la publication, en est actuellement à une phase d'élaboration. Le « cycle politique » est un outil qui peut servir à mieux comprendre les étapes nécessaires pour faire avancer le débat. Il peut se décomposer en deux volets principaux, l'un qui opère entre

disciplines et l'autre à l'intérieur de celles-ci (voir la figure 3) :⁷¹

- ▣ le cycle interdisciplinaire se déroule aux niveaux méso et macro et va de la définition d'un problème à l'élaboration d'une politique visant à y remédier, puis à sa mise en œuvre, son application et son évaluation ;
- ▣ d'autre part, les échanges de vues au sein des disciplines se déroulent au niveau micro, et comprennent la conceptualisation de solutions politiques possibles, la collecte des données, la vérification ou la validation et l'expérimentation, jusqu'à ce qu'une position politique appropriée soit atteinte.

FIGURE 3. LE DISCOURS SUR LA REPRODUCTIBILITÉ DE LA RECHERCHE VU SOUS L'ANGLE DU CYCLE POLITIQUE.



⁷¹ Janssen, M., & Helbig, N. (2018). Innovating and changing the policy-cycle: Policy-makers be prepared! *Government Information Quarterly*, 35(4), S99–S105. <https://doi.org/10.1016/j.giq.2015.11.009>

Nous pouvons affirmer que l'étape de la « définition de la problématique » dans le cycle des politiques a été franchie dans tous les domaines, mais que très peu de disciplines ont franchi le pas qui les mène à l'étape de l'« élaboration de politiques ». Notre propos ici fait ressortir le rôle important que les us et coutumes disciplinaires peuvent jouer dans l'encouragement de pratiques de publication reprocompatibles, et le cycle politique montre que les acteurs au niveau micro peuvent exercer une influence dans l'élaboration des lignes politiques. Aujourd'hui, pareille chose se produit avec lenteur et, souvent, en silos, ce qui freine les avancées aux niveaux méso et macro (avec quelques exceptions notables, comme nous l'avons déjà évoqué). En conséquence, le discours sur la reproductibilité reste quelque peu enlisé à l'intérieur de ce cycle politique.

« Des efforts sont déployés, comme dans le cas du Réseau britannique de reproductibilité, pour rassembler plusieurs initiatives afin d'harmoniser ou de faire converger des idées. Mais je pense que, pour le moment, [le paysage] est un peu flou. » Une société savante

Pour plus de clarté, il est intéressant de procéder à une comparaison entre les politiques en matière d'accès libre et celles relatives aux données ouvertes (voir la figure 4). En ce qui concerne le libre accès, toutes les étapes du cycle des politiques ont été franchies à ce jour et, sans doute, de multiples cycles ou itérations d'orientations en la matière ont été accomplis. On peut observer une grande part d'expérimentations en matière de politiques d'accès ouvert et de modèles économiques, en partie parce que des financements importants y sont consacrés. En revanche, la discussion va moins loin pour ce qui est des données ouvertes et FAIR. Certes, des orientations ont été préconisées dans certaines disciplines et par certains financeurs, mais elles sont rarement appliquées. Bien que nous venions de brosser un tableau simplifié à l'excès et que nous n'avons pas tenu compte de différences entre pays, disciplines et autres, il n'est pas surprenant que le discours sur la reproductibilité soit moins élaboré que ces éléments de comparaison - notamment parce qu'il est porté par une culture de la science ouverte qui n'a pas encore atteint sa pleine maturité.

FIGURE 4. COMPARAISON ENTRE LES SITUATIONS DES POLITIQUES DE LA SCIENCE OUVERTE.



4. Encourager et faciliter les pratiques de publication reproductibles

Les efforts de reproductibilité ne sont actuellement pas encouragés à l'intérieur du processus de recherche, et les pratiques de publication favorable à la reproductibilité sont généralement perçues comme des activités additionnelles sans contrepartie. Pour réduire la pression d'attentes toujours plus grandes envers les chercheurs, les organismes de recherche peuvent créer des postes pour soutenir la reproductibilité et faciliter la formation et le développement de compétences dans ce domaine. Incitations à la publication reproductible

Dans un monde idéal, « le travail de reproductibilité serait intégré au processus de recherche, de sorte que l'effort qu'il nécessite ne se distingue pas du travail de recherche en lui-même »⁷². Ce n'est pas le cas dans la situation actuelle, et les pratiques de recherche qui favorisent la publication reproductible sont perçues par les chercheurs comme un travail qui vient se surajouter aux tâches qu'implique classiquement un projet. Fort de ce que notre analyse fait ressortir jusqu'ici, il est clair que la reproductibilité ne peut pas être cet espèce d'« équipement en option » qui vient s'ajouter au moment de la publication, car elle recouvre un ensemble de pratiques qui imprègnent tout le processus de recherche : pour que la reproductibilité prenne une place plus importante, un changement culturel important est nécessaire.

À long terme, le meilleur bénéfice serait de passer d'une appréciation purement symbolique des efforts de reproductibilité (p. ex., promotions, augmentations de salaire, distinctions, citations) à une reconnaissance plus large pour que les travaux scientifiques soient reproductibles par défaut. Toutefois, nos interlocuteurs ont indiqué que les incitations actuelles à la publier de manière reproductible sont limitées et peu efficaces, sauf à se présenter sous la forme de consignes imposées par la revue que l'on choisit (ce qui constitue

cependant un moyen de pression extrinsèque plutôt qu'une motivation intrinsèque).

« Et une fois que [la reproductibilité] deviendra possible, une fois qu'elle deviendra facile et une fois qu'elle deviendra normative, alors il n'y aura plus aucune raison de ne pas la pratiquer. Et cela devrait certainement, en principe, améliorer la qualité de nos résultats de recherche, améliorer l'efficacité avec laquelle nous générons des connaissances qui peuvent bénéficier à la société. » Un chercheur

Il est donc essentiel d'identifier de nouveaux moyens d'encourager et de récompenser les comportements propices à la reproductibilité, et deux suggestions principales ont été avancées par nos interlocuteurs :

- ▢ rendre les étapes de travail et les contrôles associés à la reproductibilité aussi faciles que possible pour les chercheurs, en réduisant au minimum les obstacles à l'entrée, afin de garantir que toute nouvelle attente soit satisfaite avec un minimum d'efforts supplémentaires ; et
- ▢ redonner la priorité aux pratiques de science ouverte dans les commissions de recrutement et d'avancement, comme les attributions de subventions (en tenant compte, mais sans s'y limiter, de la

⁷² Association for Computing Machinery. (n.d.). Reproducibility PRINCIPLES: Taking the pulse. <https://reproducibility.acm.org/2021/04/08/reproducibility-principles-taking-the-pulse/>

reproductibilité), afin de les intégrer parallèlement aux indicateurs traditionnels ou autres.

« Je crois que cela relève des pratiques d'embauche et de financement, et un certain nombre d'activités sont en cours pour essayer de faire reconnaître les données, le code informatique et d'autres résultats dans le système d'évaluation de la recherche. » Un éditeur

Plusieurs acteurs du paysage de la recherche ouverte travaillent déjà dans ces directions, comme :

- ▶ la Déclaration de San Francisco sur l'évaluation de la recherche (DORA)⁷³ et le Manifeste de Leyde sur les indicateurs de la recherche⁷⁴ ont contribué à promouvoir des pratiques plus équilibrées dans l'évaluation de la recherche (p. ex., les organismes de financement britanniques Wellcome et UKRI^{75,76} font une promotion active de DORA et celle-ci a été signée par d'autres agences de financement comme la DFG en Allemagne)⁷⁷;
- ▶ le Conseil européen de la recherche (ERC), qui a récemment interdit toute mention du facteur d'impact des revues dans ses appels à projets⁷⁸ et
- ▶ L'université d'Utrecht, qui a décidé d'exclure le facteur d'impact dans les décisions de recrutement et de promotion.⁷⁹

Quoi qu'il en soit, à l'heure où nous écrivons ces lignes, le paradigme « publier ou périr » reste très vivace, ce qui influence grandement

les choix des chercheurs quand ils doivent définir leurs priorités dans leurs pratiques et phases de travail au quotidien.

En outre, le processus de recherche et l'ensemble des attentes des chercheurs deviennent de plus en plus complexes dans de nombreuses disciplines, ce qui laisse peu de latitude pour mettre toute la panoplie d'activités requises pour une recherche reproductible : pour tenir compte des nouvelles pratiques, il faudra gérer avec discernement les pressions subies par les chercheurs, p. ex. en ralentissant le rythme dans les projets de recherche ou en leur apportant un soutien.

« Pour les scientifiques, c'est vraiment une grande frustration. On leur demande de faire dix mille tâches différentes. Bien qu'on explique aux chercheurs comment faire ces tâches, on ne pense jamais qu'ils ont déjà trop de travail sur les bras. Et cela ne tient pas debout ». Un éditeur

Alléger les contraintes de temps pour que les pratiques de publication reprocompatibles se généralisent

Bien que notre définition de la reproductibilité soit relativement simple, les compétences requises pour réexécuter le code de quelqu'un d'autre ou évaluer la qualité d'un jeu de données sont loin d'être répandues. P. ex., un chercheur (à titre personnel ou en tant que pair examinateur) peut être un expert de premier plan dans son domaine, mais ne pas

⁷³ Déclaration sur l'évaluation de la recherche (DORA). (s.d.). Signer DORA. <https://sfdora.org/>

⁷⁴ Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics : Le manifeste de Leyde pour les métriques de la recherche. *Nature*, 520(7548), 429-431. <https://doi.org/10.1038/520429a>

⁷⁵ Wellcome. (n.d.). Guide pour les organismes de recherche sur la façon de mettre en œuvre des approches responsables et équitables pour l'évaluation de la recherche. <https://wellcome.org/grant-funding/guidance/research-organisations-how-implement-responsible-and-fair-approaches-research>

⁷⁶ UKRI. (n.d.). Research integrity. <https://www.ukri.org/our-work/supporting-healthy-research-and-innovation-culture/research-integrity/>

⁷⁷ DORA (Twitter). (2021) La Fondation allemande pour la recherche (@dfg_public) a signé DORA. <https://twitter.com/DORAssessment/status/1402273251502571531>

⁷⁸ Matthews, D. (2021). Le Conseil européen de la recherche interdit le facteur d'impact des revues dans les appels d'offres. *Times Higher Education*. <https://www.timeshighereducation.com/news/european-research-council-bans-journal-impact-factor-bids>

⁷⁹ Woolston, C. (2021). Facteur d'impact abandonné par une université néerlandaise dans les décisions d'embauche et de promotion. *Nature*, 595(7867), 462-462. <https://doi.org/10.1038/d41586-021-01759-5>

avoir les compétences nécessaires pour effectuer un contrôle de reproductibilité.

« Nous devons nous assurer que les gens aient la possibilité d'étudier un manuscrit reproductible et d'en tirer quelque chose. Je doute que la plupart des gens soient polyvalents au point que vous puissiez leur mettre n'importe quelle analyse sous les yeux et qu'ils voient de quoi il en retourne. Il se peut même qu'ils ne codent pas dans ce langage ». Un chercheur

Le soutien que les organismes de recherche apportent à leurs chercheurs peut contribuer à les soulager de certaines contraintes de temps et les aider à compléter leurs compétences là où ils en manquent. Les organismes de recherche sont bien placés pour aider leurs agents grâce à de nouveaux personnels d'appui comme des intendants et curateurs de données, des ingénieurs en logiciels de recherche ou les bibliothécaires de spécialité disciplinaire, car ceux-ci peuvent les conseiller et offrir un soutien technique tout au long de l'avancement des projets de recherche,^{80,81,82} et nous observons que les situations d'intervention de cet appui tendent à varier selon l'institution.

Nos entretiens ont fait ressortir que les actions visant à soutenir les pratiques de publication reprocompatibles sont souvent difficiles à financer (voir section 6), car elles nécessitent un engagement stratégique en faveur des pratiques de recherche ouverte et que l'importance de la curation des données et d'autres pratiques soient plus largement appréciées. Les choses sont toutefois en train

de changer dans plusieurs organismes de recherche,^{83,84,85} même si le processus de changement culturel dans cette direction a été lent.

Développer de nouvelles voies de soutien et de formation

Un rapport de la Research Data Alliance⁸⁶ se penche davantage sur le rôle potentiel des organismes de recherche et déclare qu'une augmentation des formations sur la reproductibilité serait souhaitable. Dans le même temps, nos entretiens ont mis en évidence les dangers d'une approche monolithique et trop standardisée pour soutenir les pratiques de publication reprocompatibles, ce qui montre l'importance de reconnaître la diversité quand on développe de nouveaux mécanismes dans ce domaine.

Une série d'opérations de soutien et de formation, tant au sein qu'en dehors des organismes de recherche, fleurissent actuellement dans le monde entier, et nous en présentons quelques exemples dans le tableau 3. Toutefois, il convient de noter que ces exemples sont l'exception plutôt que la règle, et plusieurs contributeurs ont souligné qu'un soutien plus structuré aux pratiques reproductibles de publication serait le bienvenu

⁸⁰ Kurapati, S. (2019). Devenir un gardien des données. LSE Impact Blog. <https://blogs.lse.ac.uk/impactofsocialsciences/2019/04/30/becoming-a-data-steward/>

⁸¹ Teperek, M., & Plomp, E. (2019). Le rôle et la valeur des intendants de données dans les universités : une étude de cas de TU Delft sur l'intendance des données. <https://doi.org/10.5281/ZENODO.2684278>.

⁸² Collège impérial de Londres. (s.d.). Subject Support. <https://www.imperial.ac.uk/admin-services/library/subject-support/>

⁸³ Plomp, E., Andrews, H., Love, J., den Heijer, K., Dintzner, N., Ilamparuthi, S., Wang, Y., & Turkylmaz-van der Velden, Y. (2021). Communicating Open Science at

TU Delft. Open Working. <https://openworking.wordpress.com/2021/03/31/data-stewardship-at-tu-delft-2020-report/>

⁸⁴ Rice, R. (2021). Bienvenue aux Research Data Stewards : un nouveau rôle. L'Université d'Édimbourg. <https://www.ed.ac.uk/information-services/about/news/welcome-to-research-data-stewards-a-new-role>

⁸⁵ Université de Maastricht. (n.d.). Services de gestion des données. <https://library.maastrichtuniversity.nl/research-support/data-stewardship/>

⁸⁶ Peer, L., Arguillas, F., Honeyman, T., Miljković, N., Gehlen, K. & CURE-FAIR WG Subgroup 3. (2021). Les défis de la conservation des résultats de recherche reproductibles et FAIR. Research Data Alliance. <https://doi.org/10.15497/RDA00063>

« En médecine, p. ex., ils sont assez investis dans la promotion de la science ouverte, de la reproductibilité et de la répliquabilité, mais ils ont tendance à le faire de manière assez monolithique, en partant du principe qu'ils n'ont en leur sein qu'une et une seule culture épistémique. »
Un chercheur

Nos entretiens ont également porté sur le rôle potentiel des champions de la reproductibilité issus de la communauté universitaire, qui peuvent assurer la sensibilisation, la formation et le soutien au changement culturel dans des domaines spécifiques relevant de leur expertise disciplinaire.

Étude de cas : l'approche de l'Université de New York (NYU) pour soutenir les chercheurs en matière de reproductibilité de la recherche

La NYU offre un large éventail de services de gestion des données, comme de nombreuses universités dans le monde. Mais ce qui rend son approche originale, c'est la nomination d'un bibliothécaire spécialisé dans la gestion des données de recherche (DOR) et la reproductibilité, un poste présent dans peu d'autres universités (voire aucune).

Le poste de bibliothécaire DOR et reproductibilité de cette université est actuellement occupé par Vicky Rampin - la seule personne à occuper ce poste à l'université de New York, et la première bibliothécaire au monde à avoir la reproductibilité explicitement formulée dans l'intitulé et la description de son poste. Dans ses fonctions uniques dans leur genre, plusieurs domaines de responsabilité lui incombent : le conseil (p. ex., des interventions individuelles ou de groupe pour résoudre des problèmes spécifiques), la formation (p. ex., des ateliers, des sessions intégrées ou à la demande dans des cours, pour des départements et des programmes spécifiques), la sensibilisation (p. ex., lors de manifestations, sur des médias sociaux, des actions de proximité auprès d'utilisateurs), la création d'infrastructures (p. ex., soutien à l'université dans le développement d'une offre de services adaptée) et la recherche (p. ex., étude directe sur des sujets tels que la reproductibilité, la préservation des logiciels, les flux de travail de recherche). La bibliothécaire chargée de la gestion des données de recherche et de la reproductibilité travaille en étroite collaboration avec le bibliothécaire chargé de la gestion des données de recherche, en partant du principe qu'il existe un chevauchement important entre ces deux domaines.

Les conseils en matière de gestion des données fournis par l'université de New York couvrent également des solutions logicielles spécifiques qui permettent des pratiques reproductibles de publication, telles que l'**Open Science Framework** (<https://bit.ly/3oY8DOy>), les **carnets Jupyter** (<https://bit.ly/3FE47uS>), **GitHub**, **GitLab** (<https://bit.ly/3ApOeEt>) et **ReproZip** (<https://bit.ly/2Z8hh2P>) (voir la section 5 pour plus d'informations à ce sujet). Ce type de conseils n'est actuellement pas courant dans les organismes de recherche et manifeste un engagement fort envers la reproductibilité de la recherche vis-à-vis de la communauté universitaire locale.

Plus largement, la formation des étudiants aux pratiques de publication reprocompatibles pourrait également être améliorée dans les universités, notamment dans le cadre de modules existants. Dans certaines disciplines, les compétences en matière de reproductibilité sont déjà enseignées très tôt : p. ex., en physique des hautes énergies et dans certaines branches de l'économie et des sciences de la vie, il n'est pas rare de travailler avec des jeux de données et du code réels au niveau L/Mastère. Toutefois, la reproductibilité n'est pas un enseignement que l'on retrouve classiquement dans les cursus destinés aux étudiants.

« À ma connaissance, nous avons eu un cours obligatoire qui portait sur l'intégrité de la recherche, y compris le plagiat et la tricherie... Nous n'avons des cours obligatoires que lorsque c'est lié à des recours judiciaires... Pour cette problématique plus normative, cela ne se produira pas. » Un chercheur

87

TABLEAU 3 : EXEMPLES D'OPÉRATIONS DE SOUTIEN ET DE FORMATION

Objet/intitulé/dénomination de la formation	URL	Typologie
Bonne pratiques et écriture de code reproductible (Best Practices for Writing Reproducible Code)	https://bit.ly/3aZEM01	Cours et ateliers
Centre de compétences numériques (Digital Competence Centre)	https://bit.ly/3Ab3eWp	Offre de formation institutionnelle
Gain de compétences en science ouverte et reproductibilité de la recherche (Increasing Openness and Reproducibility in Research)	https://bit.ly/3lfX9Ux	Offre de formation institutionnelle
Entraînement à la science ouverte et aux pratiques scientifiques reproductibles (Openness and Reproducibility Research Practices Training)	https://bit.ly/3C64JqW	Offre de formation commerciale
La reproductibilité pour tous (Reproducibility for Everyone) ⁸⁸	https://bit.ly/3G6ofG6	Guides/ateliers animés par des pairs, une communauté
ReproducibiliTea	https://bit.ly/2XwCYZo	Sessions ou séquences communautaires
Reproducibility Hub	https://bit.ly/3iB1YpL	Offre de formation institutionnelle
ReproHacks	https://bit.ly/3acWjBK	Sessions ou séquences communautaires
Monter un plan de rigueur scientifique et de reproductibilité (Rigor and Reproducibility Planning)	https://bit.ly/2Ynblu	Offre de formation institutionnelle
Entraînement à la vérification de la reproductibilité (Training For Reproducibility Verification)	https://bit.ly/3iCU13d	Guides/ateliers animés par des pairs, une communauté
Manuel Turing Way (Turing Way Handbook)	https://bit.ly/3iA7fhc	Guides/ateliers animés par des pairs, une communauté

⁸⁷ Silver, A. (2017). Collaborative software development made easy. Nature. 550, 143-144. <https://doi.org/10.1038/550143a>

⁸⁸ Auer, S., Haeltermann, N. A., Weissberger, T. L., Erlich, J. C., Susilaradeya, D., Julkowska, M., Gazda, M. A., Schwessinger, B., Jadavji, N. M., Abitua, A., Niraulu, A., Shah, A., Clyburne-Sherinb, A., Guiquel, B., Alicea, B., LaManna, C., Ganguly, D., Perkins, E. & Jambor, H. (2021). A community-led initiative for training in reproducible research. eLife. <https://doi.org/10.7554/elife.64719>

5. Innovation technologique

On dispose d'un ensemble de solutions techniques établies et innovantes permettant de mettre en œuvre des pratiques reproductibles de publication dans le paysage de la recherche, y compris auprès de fournisseurs, non-commerciaux comme commerciaux. De nombreux acteurs connaissent les outils et les services existants, mais leur utilisation varie selon la discipline et dépend souvent de l'équilibre entre la part de travail qualitatif et quantitatif dans la recherche de chacun.

Infrastructures numériques

Dans cette section, nous examinons une série de solutions technologiques disponibles dans le paysage actuel de la recherche qui peuvent avoir un impact sur les pratiques reproductibles de publication. Nos entretiens ont mis en évidence une opinion partagée selon laquelle un grand nombre des outils nécessaires pour travailler et, en particulier, publier de manière reproductible, sont déjà disponibles. Les personnes interrogées n'ont pas le sentiment que quelque chose en particulier manque actuellement dans le paysage de la reproductibilité, mais elles constatent que des niveaux variables d'expertise technique parmi les utilisateurs et un manque de connectivité (entre les outils eux-mêmes, mais aussi entre les outils et les flux de travail de recherche) entravent les efforts.

« On a besoin de beaucoup d'innovation dans le paysage des infrastructures. Il ne s'agit pas d'inventer quelque chose de nouveau, il s'agit d'améliorer ce qui existe déjà... et de baisser la barre d'entrée pour des personnes à différents niveaux de connaissances... » Un fournisseur d'infrastructure

Compte tenu du nombre de (sous-)disciplines différentes, le consensus est qu'il n'est pas souhaitable de limiter la gamme d'outils utilisés par les chercheurs. Cela signifie également qu'il faut accepter que les outils commerciaux

aient un rôle à jouer au service de la communauté des chercheurs. Toutefois, nos entretiens ont indiqué que l'équilibre entre les infrastructures commerciales et les infrastructures financées par des fonds publics devrait être surveillé conformément aux principes de l'infrastructure universitaire ouverte POSI⁸⁹, afin d'éviter la captivité commerciale (le « *vendor lock-in* ») et d'autres obstacles tels que l'accès réduit à des outils coûteux dans les pays où les budgets de recherche sont limités.

« Nous voulons empêcher la captivité commerciale et une façon d'y parvenir est de construire des standards ouverts afin qu'il y ait toujours concurrence, qu'il y ait toujours d'autres acteurs qui puissent venir sur ce marché et élaborer un meilleur service avec ce standard minimal. » Un chercheur

La connaissance et l'adoption des outils et services axés sur la reproductibilité varient selon la discipline, et la maturité des pratiques au sein d'une discipline est influencée par (i) l'équilibre entre la recherche qualitative et quantitative, (ii) la mesure dans laquelle les approches informatiques sont utilisées, et (iii) la mesure dans laquelle une discipline travaille avec des « données » numériques ou physiques. En particulier, les chercheurs qui écrivent du code dans le cadre de leur travail ont tendance à reconnaître plus rapidement les objets de recherche qui pourraient rendre

⁸⁹ POSI (2020). The Principles of Open Scholarly Infrastructure. <https://openscholarlyinfrastructure.org/>

leurs résultats publiés reproductibles, car ils font déjà partie de leurs pratiques quotidiennes (bien que, dans de nombreux cas, leur curation n'est pas suffisante pour permettre la reproductibilité) : il n'est donc pas surprenant que de nombreux outils et services que nous avons identifiés dans cette étude concernent la reproductibilité *informatique*.⁹⁰ Nos conversations ont souvent porté

sur l'utilisation d'outils tels que Docker, Jupyter Notebooks, Binder et, plus largement, sur les environnements reproductibles (voir le tableau 4 pour plus d'informations sur ces outils). Au-delà de la reproductibilité informatique, nous notons également l'importance croissante des agents d'homologation par badges ou certificats, comme *Cascad* (particulièrement axé sur l'économie), qui pourraient jouer un rôle croissant à l'avenir.

TABLEAU 4. EXEMPLES D'OUTILLAGES EN APPUI À LA RECHERCHE.

Outil ou service	URL	Fonctionnalité principale
CASCAD	https://bit.ly/3nwsAtR	Vérification de la reproductibilité
CODECHECK ⁹¹	https://dockr.ly/3aW723R	Exécution indépendante des calculs sur lesquels reposent des articles de recherche
Docker	https://bit.ly/3oGCWcz	Conteneurisation de logiciels
eLife ERA	https://bit.ly/3pkCqRW	Articles de recherche exécutables
ISA Framework	https://bit.ly/2Z8Ay4b	Totalité d'un processus de recherche (sciences de la vie)
Jupyter Notebooks	https://bit.ly/2Yw4SoX	Partage du code et documentation
Binder	https://bit.ly/3vAU3hA	Environnements interactifs et reproductibles
Code Ocean	https://bit.ly/3GiRyFF	Plateforme de recherche informatique
Gigantum	https://bit.ly/3G7jlse	Plateforme de recherche informatique
LabArchives	https://bit.ly/3joLsW	Cahier de laboratoire électronique
LabFolder	https://bit.ly/3C6nESu	Cahier de laboratoire électronique
Octopus (en développement)	https://bit.ly/3ixJypO	Totalité d'un processus de recherche (partage et documentation des objets de recherche au fur et à mesure de l'évolution des travaux)
Observable ⁹²	https://bit.ly/3piFrSO	Plateforme de recherche informatique
ReproZip	https://bit.ly/3m2Ftfs	Logiciel de « colisage/empaquetage » pour archiver des données scientifiques reproductibles associées dans un « compendium » (corpus des données d'une même recherche).
Whole Tale	https://bit.ly/2Z7J6rF https://bit.ly/3BdU8Jj	Logiciel de « colisage/empaquetage » pour archiver des données scientifiques reproductibles associées dans un « compendium » (corpus des données d'une même recherche)
R Markdown et R Notebooks	https://bit.ly/2YqDojW	Partage du code et documentation
R-Squared	https://bit.ly/3DWuf24	Vérification de la reproductibilité
Renku (voir aussi l'étude de cas page 46)	https://bit.ly/3m3lnS8	Processus de recherche complet (science des données collaborative)
Stencila	https://bit.ly/3nwsAtR	Articles de recherche exécutables et pipelines de documents exécutables

⁹⁰ Obels, P., Lakens, D., Coles, N. A., Gottfried, J., & Green, S. A. (2020). Analysis of Open Data and Computational Reproducibility in Registered Reports in Psychology. *Advances in Methods and Practices in Psychological Science*, 3(2), 229-237. <https://doi.org/10.1177/2515245920918872>

⁹¹ Nüst, D. & Eglen S, J. (2021). CODECHECK : an Open Science initiative for the independent execution of computations underlying research articles during peer review

to improve reproducibility [version 2 ; peer review : 2 approved]. F1000Research, 10(253). <https://doi.org/10.12688/f1000research.51738.2>

⁹² Perkel, J. M. (2021). Reactive, reproducible, collaborative: computational notebooks evolve, 593(7857), 156-157. <https://doi.org/10.1038/d41586-021-01174-w>

Le champ d'observation du tableau 4 est limité et ne couvre pas les entrepôts en ligne (p. ex. [Zenodo](https://bit.ly/3nfdLLR) (<https://bit.ly/3nfdLLR>), [figshare](https://bit.ly/30HxRqt) (<https://bit.ly/30HxRqt>), [OSF](https://bit.ly/3C4IFg7) (<https://bit.ly/3C4IFg7>) (**voir aussi l'étude de cas page 46**), [re3data](https://bit.ly/3vzJX0s), [Dryad](https://bit.ly/3vzJX0s) (<https://bit.ly/3vzJX0s>), le [portail Open Data du CERN](https://bit.ly/2XwkF6A) (<https://bit.ly/2XwkF6A>), les services d'hébergement d'entrepôts de code source (p. ex. [Git Hub](https://bit.ly/3DYDn6z) (<https://bit.ly/3DYDn6z>), [Bitbucket](https://bit.ly/3jpGFYJ) (<https://bit.ly/3jpGFYJ>)), d'autres logiciels permettant de mettre en œuvre des pipelines analytiques reproductibles (p. ex. [Travis CI](https://bit.ly/3AWSzsf) (<https://bit.ly/3AWSzsf>)) et les infrastructures générales globales sur lesquelles reposent de nombreux outils et services en ligne (p. ex. [Amazon Web Services](https://go.aws/3DYDzmj) (<https://go.aws/3DYDzmj>), [Microsoft Azure](https://bit.ly/3vx79fy) (<https://bit.ly/3vx79fy>), [Google Cloud](https://bit.ly/3pp2JGI) (<https://bit.ly/3pp2JGI>)) : ces éléments sont importants pour faciliter les pratiques de publication compatibles avec la reproductibilité, mais leurs objectifs et leurs motivations sont différents sur le plan stratégique. C'est particulièrement le cas des entrepôts parce qu'ils peuvent servir dans un grand nombre de cas d'usage où les exigences de partage des données pour permettre pareilles publications sont simples - p. ex., celui où il peut suffire de mettre en commun un tableau ou un ensemble de fichiers multimédias (p. ex., des images, des documents audio, des vidéos), et le principal intérêt est de veiller à ce qu'ils soient hébergés sur une plateforme ayant une stratégie de curation numérique.

« Parce que j'ai beaucoup utilisé les ressources librement disponibles en ligne, je veux aussi promouvoir autant que possible la recherche ouverte dans mon propre travail, car je vois que les avantages sont énormes. » Un chercheur

Nous reconnaissons que les entrepôts physiques et numériques ont un rôle à jouer pour assurer la reproductibilité de la recherche. Cependant, nous n'abordons pas séparément les entrepôts physiques dans notre analyse : en effet, les contacts que nous avons eu avec les acteurs et notre analyse documentaire n'ont pas permis d'identifier la publication de résultats de recherche reproductibles comme une priorité généralisée dans les disciplines qui reposent sur des objets de recherche non numériques.

Le déficit d'innovation : la connectivité des services

Bien qu'il existe un large éventail d'outils et de services permettant de mettre en œuvre la reproductibilité dans les pratiques de publication, nombre de commentaires de la part de nos contributeurs ont fait ressortir le problème de la connectivité et de l'interopérabilité des services. En bref, il s'agit du fait que les chercheurs travaillent à l'aide d'un large éventail de solutions technologiques et de logiciels, mais que ceux-ci ne se « parlent » guère entre eux.

C'est dans ce domaine que l'innovation est la plus attendue et, dans certains cas, les fournisseurs d'infrastructures viennent à la rescousse en proposant des moyens de connecter les objets de recherche tout au long du processus de recherche ou a posteriori. Dans un scénario simple (et très courant) où les problèmes de connectivité découlent de la nécessité de rassembler des résultats provenant de différentes sources, des solutions comme [OSF Projects](https://zenodo.org/communities/)⁹³ ou [Zenodo Communities](https://zenodo.org/communities/)⁹⁴ peuvent apporter quelque chose. La seconde solution mentionnée ci-dessus consiste essentiellement à enregistrer toutes les informations qui concernent un projet sur une seule page.

⁹³ Bowman, S. (2020). Create a Project (OSF Projects). <https://help.osf.io/hc/en-us/articles/360019737594-Create-a-Project>

⁹⁴ Zenodo. (s.d.). Communities. <https://zenodo.org/communities/>

« C'est un chaos absolu, ces technologies. Une désorganisation désastreuse, une décentralisation catastrophique et sans mise en forme correcte... La seule chose qui m'intéresse, c'est de pouvoir trouver les choses dont j'ai besoin rapidement, en un seul endroit, sans devoir sauter d'un truc à l'autre ou à m'inscrire sur trente-six sites différents. Donc ça, c'est un énorme problème. » Un éditeur

Dans les disciplines à forte intensité de calcul, il est important non seulement de partager les

données et le code, mais aussi de « préserver la filiation »⁹⁵ de ses travaux entre eux. Cela inclut le contrôle des versions, les environnements de calcul et, potentiellement, les différents langages de programmation dans le contexte d'une même étude. Partager toutes ces informations pour qu'il soit possible d'opérer des contrôles de reproductibilité est plus compliqué et nécessite des approches plus spécifiques, appuyées par l'utilisation d'outils nouveaux comme Renku.

Étude de cas : approches visant à améliorer la connectivité des objets numériques

Si certaines des infrastructures techniques décrites dans cette section peuvent être considérées comme des outils autonomes, d'autres s'attachent à permettre et à améliorer la connectivité entre les objets numériques et les solutions techniques. L'une des principales caractéristiques de ces solutions est l'intégration avec des services et des outils externes tels que les services d'hébergement d'entre-pôts de données et de codes sources, les carnets Jupyter, protocols.io, docker et les gestionnaires de références tels que Zotero et Mendeley. Cela leur permet (i) de minimiser les complications pour les collaborateurs ou les réutilisateurs potentiels ; et (ii) de renforcer la confiance dans le processus de recherche en permettant la transparence.

Exemple 1 : Open Science Framework

L'Open Science Framework (OSF) réalise ce qui précède au moyen de projets. Ceux-ci permettent aux utilisateurs de rassembler des objets de recherche provenant de différentes sources (ou de les télécharger directement sur OSF) et comprennent différentes sections, p. ex. un Wiki, un journal des activités récentes, des enregistrements et des analyses. Par défaut, les projets sur OSF utilisent des URL pérennes, mais la plateforme propose également la génération d'identifiants d'objets numériques (DOI). Un exemple de projet créé par l'équipe d'OSF peut être consulté ici.

Exemple 2 : Renku

Renku est centré sur la reproductibilité informatique et est une plateforme open source auto-hébergée reliant un dépôt git en ligne à un environnement d'enregistrement et d'exécution de flux de travail et de carnets interactifs (Jupyterhub) ; il améliore la connectivité entre les objets numériques en se concentrant sur les étapes informatiques du processus scientifique. La plateforme cherche à conserver, enregistrer et utiliser la filiation des résultats générés par un projet et crée un graphe de connaissances sur la base duquel des étapes individuelles ou des pipelines entiers peuvent être répétés, recherchés et réutilisés.

OSF et Renku sont tous deux des logiciels en *open source* et misent fortement sur la collaboration et le partage et l'accessibilité des travaux » Il est important de noter que ces solutions permettent la création de projets privés ou à accès contrôlé (ou même de sous-ensembles de projets, dans le cas d'OSF), ce qui signifie qu'elles sont susceptibles de convenir à l'analyse et au traitement d'informations sensibles ou autrement confidentielles.

Enfin, les questions de connectivité sont également abordées par les compendiums de recherche, qui aident les chercheurs en informatique et fournissent « un moyen standard et facilement reconnaissable d'organiser

les matériaux numériques d'un projet pour permettre à d'autres d'inspecter, de reproduire et d'étendre la recherche ».⁹⁶ En substance, les compendiums de recherche sont des dossiers structurés comprenant tous les

⁹⁵ RENKU. (n.d.). Collaborative Data Science. <https://renkulab.io/>

⁹⁶ Nüst, D, Boettiger, C, Marwick, B. (2018). How to Read a Research Compendium. arXiv. <https://arxiv.org/pdf/1806.09525.pdf>

matériaux, le code et la documentation nécessaires pour reproduire et réutiliser le travail de quelqu'un d'autre. On peut en trouver des exemples sur Zenodo⁹⁷ ou GitHub.⁹⁸

Évaluer et comparer les solutions techniques

Étant donné que les pratiques de publication à visée reproductible dépendent des représentations disciplinaires, il est difficile de développer des critères généraux pour évaluer et comparer l'utilité et l'exhaustivité de différentes infrastructures techniques. L'équipe de Gigantum s'est efforcée de développer un schéma de haut niveau pour réaliser ce qui vient d'être énoncé, en se concentrant en particulier sur la reproductibilité informatique. Leur approche de l'évaluation des solutions de reproductibilité comprend quatre dimensions clés⁹⁹ :

- ▶ la personne ou le groupe qui a créé les objets de recherche à reproduire, y compris tout code et toute donnée de recherche pertinents (le « producteur ») ;
- ▶ l'ensemble des objets de recherche nécessaires pour reproduire les résultats et les méthodes ou procédures pour y parvenir (« le projet ») ;
- ▶ l'équipement, le stockage de données et les autres infrastructures numériques nécessaires à la reproduction du résultat (« l'infrastructure ») ; et
- ▶ la personne ou le groupe qui tentera d'effectuer le contrôle de reproductibilité (« le consommateur »).

À partir de ces dimensions, il est alors possible de formuler certaines questions pour évaluer et comparer les solutions, p. ex. :

- ▶ Est-il facile pour le producteur de créer un projet reproductible ?

- ▶ Dans quelle mesure est-il facile pour le consommateur de reproduire le projet ?
- ▶ Quelle est la flexibilité de l'infrastructure en termes d'emplacement, de configuration et de contrôle ?
- ▶ Combien de fois le consommateur doit-il réexécuter pour vérifier que les choses fonctionnent ?
- ▶ Jusqu'à quel point est-il facile pour le consommateur de réutiliser ou réadapter le projet ?

La principale limite de cette approche est qu'elle est subjective et peut conduire des personnes différentes à des conclusions différentes. Cependant, les auteurs de ce schéma notent à juste titre que les relations entre les quatre dimensions sont moins subjectives, et nous pensons qu'elles ont leur utilité pour informer sur le développement et l'optimisation des produits et aider à expliquer les choix des utilisateurs.

Dans ce contexte, nous soulignons également que les disparités économiques peuvent peser dans la comparaison entre différentes solutions techniques. Les disparités peuvent se manifester différemment dans le monde et entre groupes d'acteurs, et peuvent devenir plus visibles en termes d'accès (ou de non-accès) aux composantes des infrastructures.

L'impact potentiel des principes des données FAIR

Ce que nous avons appelé le « problème de la connectivité » est en partie contenu dans la lettre « I » des principes de données FAIR (facile à trouver, accessible, interopérable et réutilisable). Plus largement, nous notons que les principes FAIR sont très propices aux pratiques de recherche reprocompatibles, et le portail Open Research Europe énonce que

⁹⁷ Zenodo. (n.d.). Research Compendium. <https://zenodo.org/communities/research-compedium?page=1&size=20>

⁹⁸ GitHub. (n.d.). research-compedium. <https://github.com/topics/research-compedium>

⁹⁹ Whitehouse, T., Clark, D., & Tsang, E. (2019). Rebooting reproducibility : From re-execution to replication. eLife. <https://elifesciences.org/labs/6fb86e9c/rebooting-reproducibility-from-re-execution-to-replication>

« par extension, les mêmes pratiques qui permettent la réutilisation des données favorisent également la reproductibilité ». ¹⁰⁰

À l'instar des pratiques de publication à visées reproductibles, la mise en œuvre des principes FAIR tend à être liée à des considérations disciplinaires sur le plan pratique : nous nous attendons à ce que les progrès dans la globalité dans ce domaine suivent également le cycle politique, notamment en ce qui concerne la contribution que les différentes disciplines devront apporter dans l'analyse et la communication de leurs besoins (voir figure 3). Dans certaines disciplines, des progrès significatifs sont en cours, p. ex. dans le cas de la ressource développée pour les sciences de la vie, le *FAIR Cookbook*¹⁰¹.

« Nous savons que l'European Open Science Cloud est axé sur les principes FAIR. Nous savons que la Research Data Alliance en fait de même. Nous savons que les principes FAIR évoluent vers des éléments comme les méthodes logicielles. Les groupes d'intérêt commun convergent donc largement sur ces questions. » Un fournisseur d'infrastructure

En outre, la réflexion autour des principes FAIR évolue, faisant ressortir que « la planification de la gestion des données devrait être soutenue tout au long du cycle de vie de la recherche afin que les données puissent *naître FAIR* et rester *suffisamment FAIR au fil du temps* »¹⁰² - un concept qui pourrait être étendu à la reproductibilité de la recherche. Bien que la « FAIRitude » et la reproductibilité présente dans les publications relèvent de pratiques distinctes, cela renforce ce que nous affirmions plus haut comme quoi tout comportement impliquant la gestion, la documentation et le partage des données ou du code

ne peut être considérée comme une simple option apportée au processus de recherche, même si nous ne nous intéressons ici qu'à l'étape de la publication.

« Nous parlons aujourd'hui de données qui sont nées FAIR ou FAIR par conception . La mesure dans laquelle vous pouvez rendre vos données ou vos résultats analytiques FAIR dès le début du processus va conduire à des niveaux de reproductibilité beaucoup plus élevés. » Un chercheur

Par rapport aux principes FAIR, il faut également citer l'European Open Science Cloud (EOSC). Cette initiative vise à construire « un réseau de données FAIR et de services associés »,¹⁰³ qui réduira la fragmentation en fédérant les infrastructures de recherche existantes et améliorera leur interopérabilité. Il est toutefois important de reconnaître que l'écosystème EOSC est toujours en cours de développement au moment de la rédaction du présent rapport et qu'il n'est donc pas sur un pied d'égalité avec les autres infrastructures numériques examinées dans ce rapport (voir également l'étude de cas page 44).

¹⁰⁰ Commission européenne. (s.d.). Open Research Europe. <https://open-research-europe.ec.europa.eu/for-authors/data-guidelines/>

¹⁰¹ Rocca-Serra, P. & Sansone, S. A. (2021). The FAIR Cookbook - a guide for your data FAIRification journey. <https://www.oerc.ox.ac.uk/news/the-fair-cookbook-a-guide-for-your-data-fairification-journey/>

¹⁰² FairsFair. (2020). FAIRsFAIR Tables Preliminary Recommendations on Data Policy <https://fairsfair.eu/news/fairsfair-tables-preliminary-recommendations-data-policy>

¹⁰³ Luyben, K. (2021). EOSC: A Web of FAIR Data. <https://beopen-project.eu/storage/files/parti-2-eosc-a-web-of-fair-data-k-luyben.pdf>

Étude de cas : les principes de données FAIR et l'European Open Science Cloud (EOSC)

En juin 2021, la Commission européenne (CE) et l'association EOSC nouvellement formée ont établi un protocole d'accord marquant le début du partenariat européen coprogrammé sur l'EOSC dans le cadre du programme-cadre Horizon Europe.

L'EOSC est considéré comme l'infrastructure habilitante qui soutiendra « un processus de recherche plus ouvert et solide grâce au numérique non seulement pour les chercheurs, mais aussi pour la société dans son ensemble ». Le mémorandum note que les objets de recherche qui sont nés FAIR « assureront une véritable capacité à l'échelle européenne et un leadership mondial pour contribuer aux objectifs de développement durable (SDG), pour atteindre l'ambition de l'UE pour le Green Deal et pour mettre en œuvre d'autres politiques nationales ou sectorielles » : cela montre l'impact potentiel de la FAIRitude et l'importance significative accordée à cela au niveau de la politique européenne.

Selon l'agenda stratégique de recherche et d'innovation de l'EOSC, « l'EOSC offrira un environnement de recherche qui favorise la science ouverte et augmente la confiance et la reproductibilité des résultats de la recherche ». Ils notent en particulier que cet écosystème en développement « offrira une infrastructure soutenue et stable pour la recherche, avec une multitude de jeux de données et d'outils de recherche facilement disponibles, encourageant ainsi les chercheurs à développer leur propre environnement de recherche sur cette plateforme, en réutilisant les composants existants, plutôt que de construire des outils uniques et non réutilisables dans leur espace informatique personnel ». Cela devrait favoriser les pratiques reproductibles de recherche et de publication.

Le mémorandum signé comprend une série d'« objectifs spécifiques » (SO4, SO5) et d'« objectifs opérationnels » (OO5, OO6, OO8) qui traitent des principes des données FAIR et visent principalement à :

- augmenter le nombre de données de recherche financées par le secteur public qui sont FAIR par conception ;
- améliorer l'interopérabilité entre les objets de recherche ;
- travailler avec les communautés pour élaborer des exigences et des pratiques ;
- fournir des indicateurs et des outils pour mesurer l'adoption des principes FAIR ; et
- co-concevoir et adopter des récompenses et des systèmes de reconnaissance appropriés pour la mise en œuvre des pratiques de données ouvertes.

Aperçu du discours sur la reproductibilité sur les réseaux sociaux

Dans le cadre de cette recherche, nous avons étudié les données Twitter entre novembre 2020 et juin 2021. Les objectifs de cet exercice étaient (i) de mesurer comment le discours autour de la reproductibilité de la recherche et des bonnes pratiques de recherche se développe en ligne ; et (ii) d'évaluer si des différences notables sont apparues entre les résultats de nos entretiens et les conversations en ligne. Au cours de la période susmentionnée, nous avons recueilli plus de 10 000 tweets en utilisant les hashtags répertoriés dans notre R notebook et dans les notes de bas de page.¹⁰⁴¹⁰⁵

Notre analyse des médias sociaux montre que les contributeurs les plus fréquents des discours sur le réseau social en ligne à propos de la reproductibilité se situent au niveau micro : sur plus de 30 comptes ayant tweeté plus de 20 fois, 12 appartenaient à des individus et cinq à des initiatives communautaires (p. ex., des réseaux de reproductibilité, des initiatives de formation et de sensibilisation-promotion). Il est intéressant de noter que six des comptes les plus prolifiques étaient des robots (*bots*), et qu'un seul appartenait à un organisme de recherche (la bibliothèque de l'Université de Leyde). Cela semble confirmer les conclusions de notre analyse documentaire et le fait que l'intérêt récent pour la reproductibilité de la recherche prend son essor principalement comme un mouvement partant de la base.

D'autre part, si on examine les comptes dans le jeu de données (c'est-à-dire les comptes qui ont tweeté au moins une fois en utilisant l'un des hashtags surveillés) avec le plus grand nombre d'abonnés, ils comprennent un éventail d'entités décisionnaires (p. ex. la DG Recherche et innovation de l'UE, l'Institut national américain des normes et de la technologie, Research en Allemagne), des revues et des éditeurs (p. ex., Science Magazine, British Medical Journal, Nature Biotechnology), des sociétés savantes (p. ex., Association for Computing Machinery, Microbiology Society) et des organes d'information (p. ex., Phys.org, LSE Impact Blog). Cela suggère que les groupes d'acteurs aux niveaux méso et macro sont effectivement attentifs, mais ne mènent pas la discussion - p. ex., les trois comptes les plus suivis n'ont eu qu'un tweet chacun dans notre jeu de données, alors que leur audience combinée était de près de 3,5 millions d'abonnés.

Le tableau 5 présente les hashtags inclus dans les discussions liées à la reproductibilité. Il est à noter que ce tableau inclut à la fois les hashtags que nous avons surveillés directement (p. ex. #reproducibility et #reproducibleresearch) et d'autres hashtags partagés dans les tweets que nous avons récoltés à côté d'un hashtag surveillé (p. ex. #openscience). Le fait que #openscience - un hashtag que nous n'avons pas surveillé - figure si haut dans la liste suggère que la reproductibilité et la science ouverte sont considérées comme faisant partie d'un même ensemble sur les médias sociaux, ce qui confirme une fois de plus les résultats de nos échanges avec les acteurs.

¹⁰⁴ Les hashtags surveillés dans cet exercice comprenaient : #Reproducibility, #Replicability, #ReproducibleScience, #ResearchReproducibility, #ReproducibleResearch, #ResearchCredibility, #GoodResearchPractices, #RegisteredReports, #GoodScience, #ResearchCompendium, #ResearchCompendia, #ReproducibilityCrisis, #ReplicabilityCrisis, #ReplicationCrisis et #TuringWay.

¹⁰⁵ Chiarelli, A. (2021). Publishing Reproducible Research Outputs - Text and data mining code (Version 2) [Logiciel]. Zenodo. <https://doi.org/10.5281/ZENODO.5512019>

TABLEAU 5. HASHTAGS UTILISÉS ET OCCURRENCES DANS LE JEU DE DONNÉES. (†=HASHTAG SURVEILLÉ)

Hashtag	Occurrences dans le jeu de données	Classification
#reproducibility†	4,148	hashtag clé dans les discussions sur la reproductibilité
#reproducibleresearch†	1,568	hashtag clé dans les discussions sur la reproductibilité
#openscience	1,182	hashtag clé dans les discussions sur la reproductibilité
#methodsmatter	854	hashtag généré par robot
#registeredreports†	481	Principalement associé à la réplication
#rrid	350	hashtag généré par robot
#accelerateopenscience	297	hashtag généré par robot
#stmpublishing	270	hashtag généré par robot
#research	204	Hashtag générique
#science	159	Hashtag générique

Notons que notre exercice de fouille de données avait une portée assez large et incluait également des hashtags liés à la réplication. Le seul d'entre eux à figurer dans le top 10 est #registeredreports, tandis que d'autres, tels que #replicability et #replicationcrisis, n'ont été utilisés que 121 et 109 fois, respectivement. Il est à noter que le top 10 comprend également quatre hashtags générés par des robots : le plus souvent, ils ne sont pas tweetés par des humains et, dans une certaine mesure, ils faussent artificiellement la discussion (mais ils sont tout de même moins fréquents que les trois premiers). Certains hashtags génériques sont également inclus dans le tableau 5, mais nous pensons qu'ils seront plus efficaces lorsqu'ils seront utilisés en combinaison avec d'autres hashtags axés sur la reproductibilité.

Enfin, il est intéressant de commenter les comptes¹⁰⁶ les plus mentionnés dans notre jeu de données, car cela peut donner une indication du leadership intellectuel et de l'impact potentiel (voir tableau 6). La bibliothèque de l'université de Leyde arrive en tête de liste en raison d'un séminaire de formation sur la reproductibilité de la recherche qu'elle a organisé avec Elsevier. Le UK Reproducibility Network figure également en haut de la liste, tandis que d'autres réseaux de reproductibilité apparaissent en bas de la liste, probablement en raison de leur récente existence. Il est intéressant de constater un nombre relativement important d'acteurs de niveau méso dans le tableau 6, notamment en raison des événements de formation et de sensibilisation-promotion, leur rôle en tant qu'infrastructures de reproductibilité et des annonces générales autour des hashtags que nous avons surveillés.

TABLEAU 6. COMPTES LES PLUS MENTIONNÉS (NOTE : RESEARCH CONSULTING, L'AUTEUR DU RAPPORT, EST RESSORTI EN TANT QUE 9^E DANS LA LISTE ET A ÉTÉ RETIRÉ DE CE TABLEAU).

Compte	Occurrences dans le jeu de données	Type
@ubleiden	43	Bibliothèque universitaire
@ukrepro	38	Initiative communautaire
@ElsevierConnect	30	Éditeur
@repro4everyone	24	Initiative communautaire
@CodeOceanHQ	23	Fournisseur d'infrastructure
@OSFramework	19	Fournisseur d'infrastructure
@SciReports	19	Éditeur
@protocolsIO	16	Fournisseur d'infrastructure
@SwissRN	14	Initiative communautaire
@GermanRepro	13	Initiative communautaire

¹⁰⁶ Twitter. (s.d.). À propos des réponses et des mentions. <https://help.twitter.com/en/using-twitter/mentions-and-replies>

6. Comment financer les coûts des pratiques de publication reprocompatibles

Il est essentiel de déterminer les mécanismes de financement pour subvenir aux investissements en temps et en travail qu'impliquent les pratiques reproductibles de publication. Quatre activités peuvent entraîner des coûts sur l'ensemble d'un projet de recherche : la curation, la préparation et le partage d'objets de recherche par les chercheurs ; l'appui apporté par le personnel des organismes de recherche ; les contrôles de reproductibilité effectués dans le cadre du processus de publication ; et tout contrôle de reproductibilité a posteriori.

Implications financières des flux de publications reproductibles

Tous les groupes d'intervenants consultés sont d'avis que la reproductibilité de la recherche devrait être une affaire dont s'empare chaque chercheur au premier chef, en grande partie parce que ce sont les chercheurs qui conçoivent, distribuent et publient leurs travaux. Nos entretiens ont souligné le fait qu'il est extrêmement difficile (si ce n'est impossible) d'examiner la question de la reproductibilité seulement à l'étape de la publication et de la diffusion : cela indique que, pour publier de manière reproductible, les chercheurs devraient intégrer un éventail de pratiques et de comportements sur tout le spectre du processus de recherche.

Dans cette section, nous présentons une série de mécanismes de financement destinés à soutenir les efforts de reproductibilité des travaux et les contrôles associés, en portant l'essentiel de l'attention sur les moyens pour couvrir les dépenses occasionnées par :

- ▢ le temps et les efforts déployés par les chercheurs dans le cadre d'un projet de recherche ;
- ▢ les ressources en temps et moyens du personnel d'appui dans les organismes de recherche ;
- ▢ les contrôles de reproductibilité dans le contexte du processus de publication ;

- ▢ les contrôles de reproductibilité a posteriori.

Le premier de ces points est considéré comme une nécessité pour que la reproductibilité de la recherche devienne un sujet d'attention plus prioritaire : pour qu'il rencontre un niveau d'attentes plus fort, il faut que les chercheurs reçoivent reconnaissance et soutien de la part de leurs organismes. Sans cela, la reproductibilité de la recherche risque de rester une activité spécifique à un domaine ou une niche. Par ailleurs, il est nécessaire de poursuivre les échanges et les expérimentations du côté des activités de contrôle et de vérification, y compris les questions de rôles et de responsabilités, ainsi que de phasage (voir section 3).

Il convient de noter qu'en raison de la nature évolutive du discours sur la reproductibilité, les mécanismes abordés dans cette section sont quelque peu expérimentaux par nature, et que l'on dispose de très peu d'informations pour comparer leur efficacité.

Financement du temps et des activités des chercheurs dans le cadre d'un projet de recherche

Un moyen simple de soutenir financièrement la reproductibilité de la recherche consiste à abonder directement le temps et les autres dépenses associées (par exemple, d'infrastructure, d'abonnements à des outils logiciels

ou à des services) dans le cadre de subventions de recherche. C'est ce que font déjà certains financeurs pour la gestion des données de recherche¹⁰⁷, et des mécanismes de financement similaires pourraient être imaginés pour soutenir la reproductibilité.

Certains de ceux que nous avons interrogés ont bel et bien reconnu qu'il importe de financer les efforts de reproductibilité, et si aucun mécanisme de ce type n'est actuellement en place, il est admis que des attentes accrues devraient être assorties de financements supplémentaires.

« Les fondations acceptent que ces [activités de reproductibilité] constituent une dépense future potentielle - c'est même une dépense aujourd'hui... La seule chose que nous avons exigée, c'est que celles-ci ne peuvent pas être rangées au titre de frais généraux non identifiés et vagues. » Un financeur de la recherche

Il est difficile d'estimer avec précision le temps nécessaire à la réalisation des activités liées à la reproductibilité (de la conception du projet à la publication), car celles-ci ne peuvent être dissociées de façon artificielle des autres bonnes pratiques de recherche (comme l'éthique, l'intégrité) et à ce qui touche par ailleurs à son ouverture (partage ouvert et FAIRisation des données, du code et des articles). En outre, les choses se compliquent encore si l'on prend en compte les différences d'investissements en temps et activités motivées par les us et coutumes disciplinaires et leur philosophie de la recherche. Cela donne à penser qu'il faudra faire preuve d'empirisme avant qu'un « outil de calcul des coûts de reproductibilité » semblable ou complémentaire à celui dont on dispose pour la gestion des données de recherche¹⁰⁹ puisse

être envisagé, d'autant plus que les deux pratiques sont étroitement liées.

Enfin, nous observerons que le financement de chercheurs ou de projets de manière individualisée par le biais de subventions de recherche ne tient pas compte de la nature à long terme de certains aspects de la reproductibilité. Par exemple, les infrastructures numériques et la préservation des données nécessitent des mécanismes de financement différents, car elles sont d'une durée de vie qui dépasse celle d'un projet et servent généralement un public beaucoup plus large.

Financement du temps et des efforts du personnel d'appui à l'intérieur des organismes de recherche

Certaines des personnes interrogées ont estimé peu probable que les activités de reproductibilité mobilisent nécessairement un équivalent temps plein (ETP) ou même un demi-ETP sur toute la durée d'un projet : par conséquent, les chercheurs principaux se trouveront le plus souvent dans l'impossibilité d'embaucher du personnel spécialisé dans ce domaine. C'est pourquoi il serait utile que certaines activités liées à la reproductibilité soient prises en main par le personnel d'appui employé par les organismes de recherche (voir la section 4).

« Il est très difficile pour un projet de recruter quelqu'un à 0,1 ou 0,2 ETP. Et qui va se présenter à ce genre d'emploi ? Disons que ce n'est pas une perspective intéressante pour une personne qui dispose de bonnes compétences dans les données que de choisir un poste comme ça, sauf situation particulière. » Un organisme de recherche

¹⁰⁷ OpenAIRE. (n.d.). How to identify and assess Research Data Management (RDM) costs. <https://www.openaire.eu/how-to-comply-to-h2020-mandates-rdm-costs>

¹⁰⁸ University College London (UCL). (n.d.). Costing Research Data Management. <https://www.ucl.ac.uk/library/research-support/research-data-management/policies/costing-research-data-management>

¹⁰⁹ UK Data Services. (2015). Outil de calcul des coûts. <https://www.ukdata-service.ac.uk/media/622368/costingtool.pdf>

Si une certaine responsabilité se voit assignée aux organismes de recherche, il est important de songer que les budgets des universités¹¹⁰ (et des bibliothèques¹¹¹) sont en butte à des pressions croissantes, notamment en raison de la pandémie de COVID-19. Au cours de notre consultation, il est apparu clairement que financer durablement du personnel de soutien à la recherche est certes difficile, sans être impossible pour autant. Ainsi, les organismes de recherche peuvent entreprendre de financer une partie des ETP d'une équipe d'intendants et de curateurs de données ou de bibliothécaires spécialisés (par exemple 20 %), et projeter que le reste (par exemple 80 %) soit couvert par financements de projets. Ce faisant, les organismes de financement de la recherche subventionneraient bel et bien ces postes et traduiraient de la sorte leurs volontés politiques (le cas échéant) en faveur de bonnes pratiques et de la reproductibilité dans la recherche. Pour que pareils investissements soient satisfaisants pour les organismes de recherche, il faut un engagement stratégique fort de la part de la hiérarchie et une volonté de lancer des programmes pilotes pour évaluer ce qui fonctionne dans leur contexte et leurs combinaisons de disciplines.

« L'université a financé un certain nombre de postes au niveau central faculté par faculté. On pensait en fait qu'en l'espace de trois ans, on en récupérerait entièrement les coûts grâce à des appels à projets soumis au sein de la faculté [...]. Je pense qu'il faudrait récupérer 80 % des salaires correspondants de manière régulière pour

¹¹⁰ Banque Mondiale. (2020). Under Pressure: Covid 19 and funding European Higher Education. <https://blogs.worldbank.org/education/under-pressure-covid-19-and-funding-european-higher-education>

¹¹¹ Brinton, J. (2021). The impact of COVID-19 on the UK publishing industry: Findings and opportunity. *Learned Publishing*, 34(1), 43–48. <https://doi.org/10.1002/leap.1363>

que ces postes soient prolongés." Un organisme de recherche

Financement des contrôles de reproductibilité dans le cadre du processus de publication

Il existe plusieurs façons de mettre en œuvre des contrôles de reproductibilité, correspondant à différentes étapes du processus de publication (voir Figure 5). Notre consultation a notamment mis en évidence les trois trajectoires suivantes :

- ▣ les évaluateurs scientifiques pourraient assurer des responsabilités supplémentaires, en testant la reproductibilité dans les articles, les données et le code soumis pour publication ;
- ▣ les éditeurs pourraient mobiliser du personnel interne tel que des éditeurs de données ou de reproductibilité ;¹¹²¹¹³
- ▣ des prestataires indépendants pourraient être chargés de vérifier la reproductibilité des articles par les auteurs, les revues ou les bailleurs de fonds.

La première de ces trois options est la moins susceptible de réussir, notamment parce que l'examen par les pairs est actuellement une responsabilité non financée et non écrite des chercheurs. La communauté de la recherche peut être défavorable à l'ajout de tâches ou de consignes liées à la reproductibilité à la

¹¹² Office of the AEA Data Editor. (2021). Office of the AEA Data Editor.

<https://aeadataeditor.github.io/>

¹¹³ Twitter. (2021). Ben Greimer. Extrait de ses tweets. https://web.archive.org/web/20210708161656/https://twitter.com/bgreiner_tweets/status/1383070111544774661

fonction d'évaluateur, car cela implique un investissement en temps que tous les chercheurs ne peuvent assumer, ainsi qu'une expertise supplémentaire (p. ex. les questions de code) qui peut ne pas être à la portée de tous.

« Pour ce qui est du travail d'évaluation en tant que pair, essentiellement, on ne dispose pas souvent de beaucoup de temps pour le faire. Je pense donc que compter sur l'examineur pour qu'il vérifie vraiment si tout coche bien les cases sur le plan informatique dans un manuscrit, c'est un peu trop charger sa barque. » Un chercheur

En revanche, les deuxième et troisième points ménagent un meilleur équilibre entre les responsabilités des acteurs de la recherche et, à ce titre, passent généralement pour plus acceptables.

Certains éditeurs ont fait remarquer que des revues craignent peut-être d'ajouter des exigences de reproductibilité au processus de soumission, mais nos investigations ont montré que certaines ont mis en place des vérifications de cet aspect dans leurs flux de travail (voir la section 3). Dans ces cas, le rôle des personnels de rédaction en charge des données ou de la reproductibilité est essentiel¹¹⁴¹¹⁵¹¹⁶, et cela s'appliquera probablement à un éventail de responsabilités qui portent sur la conformité aux lignes directrices relatives à la reproductibilité, un contrôle des

données de recherche et une analyse du code. La manière dont ces activités sont financées en interne par les revues ne ressort pas tout à fait clairement, mais nous avons constaté que ces responsabilités peuvent être mises en commun par un éditeur entre différents titres, notamment sous la forme de petites équipes spécialisées dans les données de recherche et la reproductibilité.

« Dans la recherche aujourd'hui, il y a beaucoup, beaucoup d'articles où vous traitez des données qui sont d'une telle taille, et si multidisciplinaires, qu'il n'est pas nécessairement pertinent que ce soit les pairs examineurs qui le fassent [effectuer des contrôles de reproductibilité]. Vous ne pouvez en fait examiner ce type d'informations que si vous réutilisez les données vous-même. » Un éditeur

Enfin, les chercheurs, les éditeurs ou d'autres parties prenantes peuvent souhaiter externaliser les contrôles de reproductibilité. Cela diffère quelque peu des autres mécanismes, car cela réduit une partie de la charge de travail attendue de la part des chercheurs, de leurs institutions et de leurs éditeurs et la déplace vers d'autres fournisseurs de services. Bien entendu, cela a un coût, que l'un des acteurs concernés devra supporter.

¹¹⁴ American Statistical Association (ASA). (n.d.). JASA Editors Talk Reproducibility. <https://www.amstat.org/ASA/Publications/Questions/As/JASA-Editors-Talk-Reproducibility.aspx>

¹¹⁵ Elsevier. (n.d.). Information Systems. <https://www.journals.elsevier.com/information-systems/editorial-board>

¹¹⁶ Biometrical Journal. (n.d.). Consignes aux auteurs de Biometrical Journal. <https://onlinelibrary.wiley.com/page/journal/15214036/homepage/forauthors.html>

Le programme Régie des données de l'université de technologie de Delft (TU Delft)

En 2017, la bibliothèque de la TU Delft a lancé un projet pilote de régie des données afin de répondre aux nécessités inhérentes à leur gestion dans l'ensemble de l'université. Cette action s'appuyait sur le constat qu'« assurer un changement culturel durable, n'est pas seulement une question de technologie et d'expertise, mais aussi, et ce qui est peut-être plus important, de communication et de confiance. Les relations entre les chercheurs et ceux qui les conseillent sur les pratiques de gestion des données doivent être développées au fil du temps et en permettant aux gens de se connaître et de travailler en étroite collaboration. » Ce projet de régie des données a été financé au niveau central grâce à une ligne budgétaire du conseil d'administration de l'Université et placé sous la houlette du coordonnateur régisseur des données qui intervenait depuis la bibliothèque de la TU Delft.

Quatre ans plus tard, ce projet pilote est devenu une ressource importante pour l'université, et les responsabilités en matière de financement ont été transférées aux différentes facultés. Les postes de régisseur des données ont été pérennisés par les facultés, reconnaissance du rôle important qu'ils jouent dans le traitement des demandes de gestion des données dans l'ensemble de l'établissement.

Entre 2020 et 2021, cette équipe de régisseurs a affiché un certain nombre de résultats à son profit :

- la prise en charge de plus de 800 demandes de la part de chercheurs ;
- l'organisation et la prestation d'actions de formation en gestion des données propres à chaque faculté ;
- l'élargissement des interventions au niveau de l'institution entière, par exemple en offrant un atelier bénévole d'initiation aux données génomiques, un atelier de perfectionnement au code et un autre atelier bénévole d'introduction aux données en sciences sociales ;
- la mise au point de programmes de gestion des données propres aux facultés et l'apport de conseils localisés et spécifiques aux chercheurs.

Dans l'ensemble, le projet de régie de données a fait le constat que « la construction collective est une composante essentielle pour une mutation : les données peuvent interagir efficacement au sein de leur faculté, avec la bibliothèque centrale, les services centraux et individuellement auprès des chercheurs. » « étant donné qu'un régisseur de données ne peut être au fait de toutes les pratiques disciplinaires spécifiques en présence au sein de leur faculté »

Le cadre stratégique de la TU Delft pour la régie des données énonce qu'« une bonne activité de régie - assurant [...] la collecte, la documentation, le partage, la publication et la conservation - des données est essentielle pour garantir que l'argumentation scientifique et ses résultats soient reproductibles, et pour que la TU Delft soit perçue comme un partenaire digne de confiance autant par les organisations publiques que commerciales. »

Quantifier ce que peuvent coûter des contrôles de reproductibilité externalisés est une affaire complexe, car ceux-ci ne seront pas forcément les mêmes en fonction des exigences et des approches disciplinaires. Cependant, un document de travail publié en collaboration avec l'agence de certification française *casca*d est d'un bon secours¹¹⁷ : elle a modélisé l'impact des économies d'échelle

réalisées sur le coût des contrôles de reproductibilité et a fait part de conclusions précieuses (voir tableau 7)¹¹⁸.

Les analyses menées par cette agence ont abouti au chiffre de 320 € par article contrôlé, et cette estimation diminue à 155 € par article à mesure que le nombre de contrôles augmente et/ou que les dépenses d'infrastructure sont amorties sur une période plus

¹¹⁷ Casca.d. (n.d.) casca.d - Certification agency for scientific code & data. <https://bit.ly/3B2YzGw>

¹¹⁸ Hurlin, C., & Perignon, C. (2019). Reproducibility Certification in Economics Research. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3418896>

longue (dans ce modèle, les coûts varient selon le nombre d'« équipes de vérification », et les chiffres ici se réfèrent à un cas impliquant une seule équipe).

TABLEAU 7. ESTIMATION DES COÛTS À LA MISE EN PLACE ET RÉCURRENTS (INFRASTRUCTURE ET TEMPS) POUR EFFECTUER DES CONTRÔLES DE REPRODUCTIBILITÉ (CASCAD).

Catégorie	Nature	Estimation (€)
Infrastructure informatique	Préparation	50 K
Coût d'accès aux données (estimation prudente)	Préparation	5 K
Main d'œuvre (10 heures par vérification)	Récurrent	150 par vérification

Notons que *cascad* opère principalement dans le domaine de l'économie, et les chiffres présentés ci-dessous reflètent cette orientation. C'est ainsi que les auteurs reconnaissent que le temps d'examen à *cascad* peut être supérieur à la moyenne en raison de la nécessité d'accéder à des données confidentielles.

Bien qu'il ne soit pas possible de généraliser les coûts indiqués dans le tableau 7, « les coûts de la reproductibilité vérifiée [en externe] sont loin d'être négligeables, mais tout de même gérables si l'on s'efforce de réduire les dépenses de mise en œuvre et si l'on trouve le bon modèle économique ».¹¹⁹

Financement des contrôles de reproductibilité a posteriori.

Les organismes de financement de la recherche peuvent proposer des mécanismes permettant de vérifier la reproductibilité de recherches publiées (voir par exemple le passage sur les « études de réplication » de NWO dans la section 3).

Il importe de reconnaître que le financement direct de ces contrôles n'est pas perçu comme une solution durable à long terme, mais comme un outil de sensibilisation visant à populariser l'idée de reproductibilité et à la

mettre sous les projecteurs. Sur le long terme, les agences de financement sont davantage susceptibles (et désireuses) de consentir des montants de subventionnements plus modestes destinés à assumer le coût supplémentaire encouru pour la prise en charge de la reproductibilité, tant tout au long du déroulement des travaux de recherche qu'au moment de leur publication.

« Dans la plupart de nos appels, nous permettons déjà aux gens de prévoir un budget spécifique pour la publication en libre accès. Et je pense que [les pratiques de reproductibilité] sont quelque chose qui pourrait également en faire partie. À l'avenir, nous espérons fonctionner avec un système de financement dans lequel il existe des modules que vous pouvez demander pour différents aspects de votre publication, et [la reproductibilité] pourrait simplement être un de ceux-là. » Un organisme de financement de la recherche

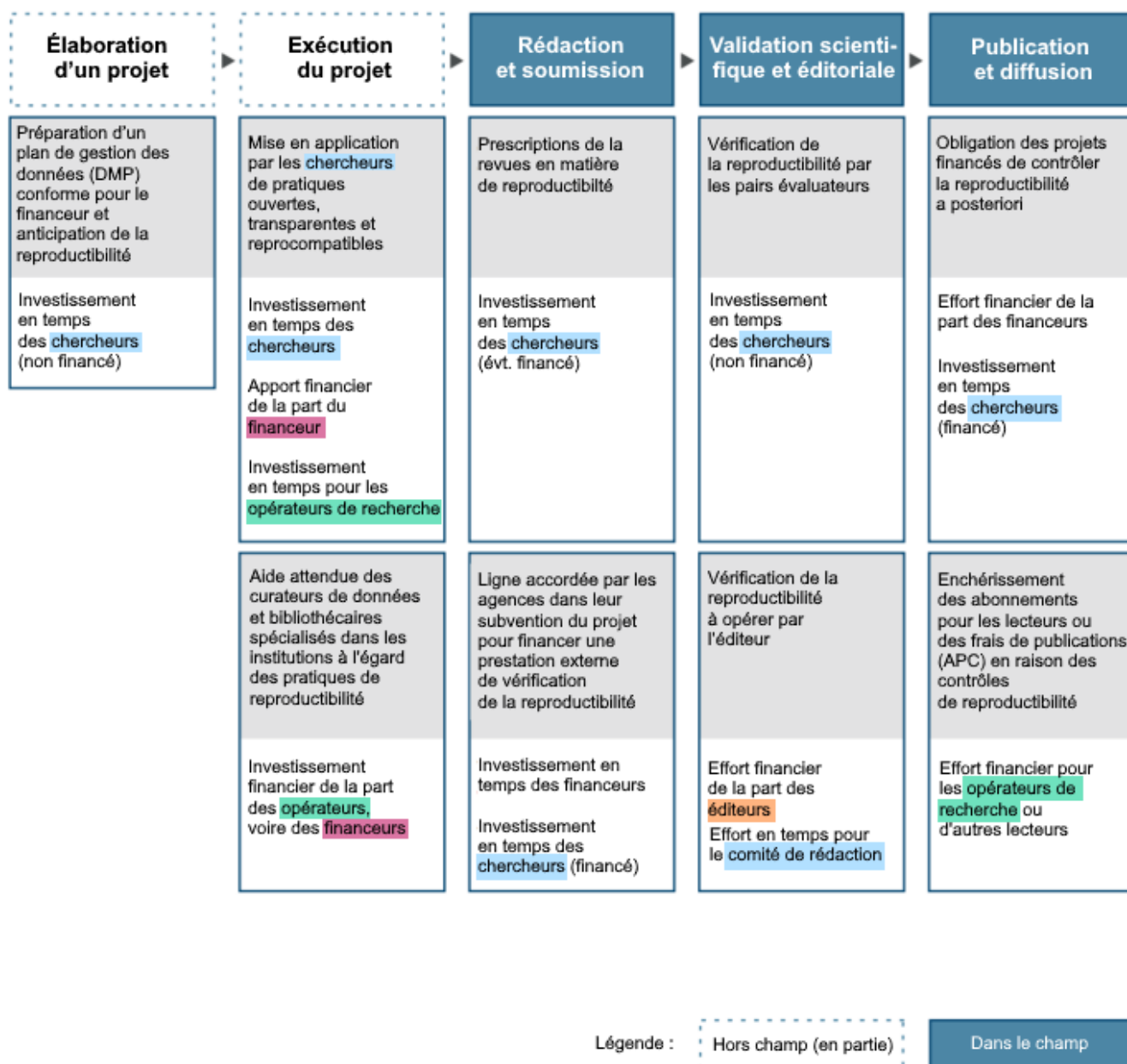
Ce qui précède ne veut pas dire qu'il n'y a pas de place pour la vérification a posteriori, mais simplement que cela ne semble pas être une stratégie pratique pour les financeurs de la recherche. Par exemple, les chercheurs qui cherchent à réutiliser les travaux publiés par d'autres peuvent décider de vérifier s'ils sont reproductibles avant de passer à l'étape suivante, mais dans le cadre des phases du déroulé de leurs propres travaux (et, éventuellement financées).

Enfin, au moins en principe, les organismes de financement pourraient eux-mêmes exercer un rôle de contrôle, en se concentrant sur la reproductibilité des travaux publiés en lien avec leurs subventions. Toutefois, au vu des difficultés importantes rencontrées pour vérifier le respect des politiques appliquées aux données de recherche¹²⁰, il est peu probable que cette stratégie soit poursuivie.

¹¹⁹ Hurlin, C., & Perignon, C. (2019). Reproducibility Certification in Economics Research. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3418896>

¹²⁰ Couture, J. L., Blake, R. E., McDonald, G., & Ward, C. L. (2018). A funder-imposed data publication requirement seldom inspired data sharing. PLOS ONE, 13(7), e0199789. <https://doi.org/10.1371/journal.pone.0199789>

FIGURE 5. INVESTISSEMENTS FINANCIERS, EN TEMPS ET MODÈLES DE FINANCEMENT DE LA RECHERCHE.



Le coût des infrastructures numériques

Il peut paraître surprenant que notre propos n'ait pas abordé jusqu'à présent le coût des infrastructures numériques permettant d'exercer des pratiques de publication scientifique reprocompatibles. Notre étude a montré que les chercheurs utilisent déjà la plupart des outils dont ils auraient besoin pour publier de la sorte, tels que des entrepôts de codes et de données ou des logiciels spécialisés, et que l'accès à ces outils est soit gratuit à la source¹²¹, soit payé par des subventions de recherche ou les organismes de recherche. En outre, l'équilibre entre les infrastructures commerciales et publiques n'est pas considéré comme un problème, tant que les chercheurs disposent des outils dont ils ont besoin pour mener à bien leurs travaux et que le risque de captivité commerciale est limité (voir section 5).

« Si vous pouvez d'une manière ou d'une autre garantir que la captivité commerciale est évitée, ou du moins que le marché n'est pas verrouillé par un acteur en particulier, alors je pense que nous pouvons, d'une certaine manière, ne pas nourrir d'inquiétude envers le rôle du secteur privé. Je ne pense pas que nous voulons éliminer ces acteurs ; je pense que pour finir ils vont nous fournir les solutions les plus robustes de toute façon. » Un chercheur

Toutefois, les infrastructures numériques pourraient bénéficier d'un financement accru ou nouveau dans deux domaines :

- ▢ les efforts visant à renforcer et à améliorer la capacité des infrastructures existantes nécessaire au stade de la publication : nous avons constaté que les outils disponibles dans le paysage sont souvent adaptés à l'objectif, mais qu'une

plus grande généralisation des pratiques de recherche ouvertes et du partage nécessitera des améliorations fonctionnelles et de capacités (p. ex. stockage des données, rapidité d'accès, préservation numérique) ;

- ▢ les développements d'infrastructures numériques supplémentaires offrant des fonctionnalités liées à la reproductibilité sur l'ensemble du processus de recherche, notamment en améliorant l'interopérabilité et en intégrant mieux les solutions dans les flux de travail existants : ce type de solution est actuellement conçu et testé dans différents domaines disciplinaires, l'accent étant mis sur le suivi et la description des objets de recherche au fur et à mesure de l'évolution d'un projet.

Bien que le deuxième point dépasse légèrement le cadre du présent rapport, il mérite d'être examiné plus en détail. La question de la reproductibilité étant encore à un stade précoce de son évolution et en raison de son caractère fortement expérimental, les actions visant à la soutenir tout au long de la vie d'un projet ne sont pas encore financièrement viables, même si elles sont perçues comme telles et souhaitables par certaines communautés disciplinaires¹²². Dans ces scénarios, les agences de financement et les organismes de recherche peuvent jouer un rôle par l'apport d'un soutien financier à court et à moyen terme pour s'assurer que des flux de recherche novateurs axés sur la reproductibilité et la transparence soient imaginés et mis en œuvre. C'est le cas, par exemple, du modèle de financement adopté par Renku, qui est porté par le Swiss Data Science Center (une initiative commune de l'EPFL et de l'ETH de Zurich), et par le projet CONQUAIRE (Continuous Quality Control for Research Data to Assure Reproducibility) de l'Université de Bie-

¹²¹ Nous notons que de nombreux services sont gratuits au moment de la rédaction de ces lignes, par exemple Zenodo, l'OSF ou GitHub. Toutefois, leurs modèles de financement pourraient nécessiter un examen plus approfondi à l'avenir pour assurer leur viabilité à long terme.

¹²² Black, S., Gardner, D. G., Pierce, J. L., & Steers, R. (2019). Organisational Behaviour. <https://opentextbc.ca/organizationalbehavioropenstax/>

lefeld, qui a été soutenu par l'agence de financement allemande, la DFG, et dont le but a été de développer des services et des outils que les chercheurs peuvent employer lors de la création, la collecte ou le versionnage de données et pour les rendre réutilisables¹²³. Un exemple plus récent s'offre à nous avec la plate-forme Octopus, à laquelle UKRI a attribué un financement de 650 000 livres sterling et dont l'objectif est « de proposer une nouvelle ressource d'archivage primaire des travaux de recherche » destinée à collecter et évaluer les travaux scientifiques « au fil de l'eau ».¹²⁴

Enfin, nous faisons ressortir combien il est avantageux de faciliter les discussions et échanges entre projets à l'initiative du monde académique et des personnes ou organisations qui possèdent un savoir-faire commercial (sauf à ce que les deux se confondent). Les premiers disposeront d'une connaissance et d'une compréhension approfondies de leur public et de ses points névralgiques, alors que les seconds sauront apporter des stratégies de déploiement à grande échelle, de croissance viable des activités et d'adaptation au marché des produits. Pareilles discussions sont importantes, qu'un outil ou un service soit développé avec un but lucratif ou non, en particulier dans les cas où la subvention n'a été accordée que pour la phase de démarrage ou pilote et qu'aucun plan à long terme n'a été mis en place.

Comme nous l'avons mentionné plus haut, nous n'avons pas enquêté sur les infrastructures non numériques. Nous recommandons toutefois de surveiller leurs besoins à l'avenir, car des exigences de financement peuvent apparaître à mesure que le discours sur la reproductibilité continue à prendre de l'ampleur.

Le contrôle du respect des préconisations et ses complexités

Bien qu'une large panoplie de solutions technologiques soit en place pour porter les pratiques de publication reprocompatibles et qu'un éventail de modèles économiques existe, cela ne signifie pas que le contrôle et l'application des prescriptions de reproductibilité seront simples.

Il est utile de tirer parti de l'expérience acquise dans le paysage des données ouvertes et FAIR, où plusieurs exigences sont fixées par les financeurs et, dans une moindre mesure, les éditeurs.¹²⁵ Dans ces cas précis, on est en présence de lacunes importantes du côté des applications. En effet, malgré l'existence de politiques et de flux de travail (voir la figure 6 pour un exemple de flux de travail qui intègre l'examen des données), le coût de leur application et celui du suivi de leur observance sont souvent trop élevés pour être réalistes. Procéder à un contrôle des prescriptions de reproductibilité nécessite un examen encore plus complexe des objets de recherche associés, y compris une compréhension des questions (sous-)disciplinaires : par conséquent, il peut être difficile pour les financeurs d'assumer une responsabilité de suivi de ce type.

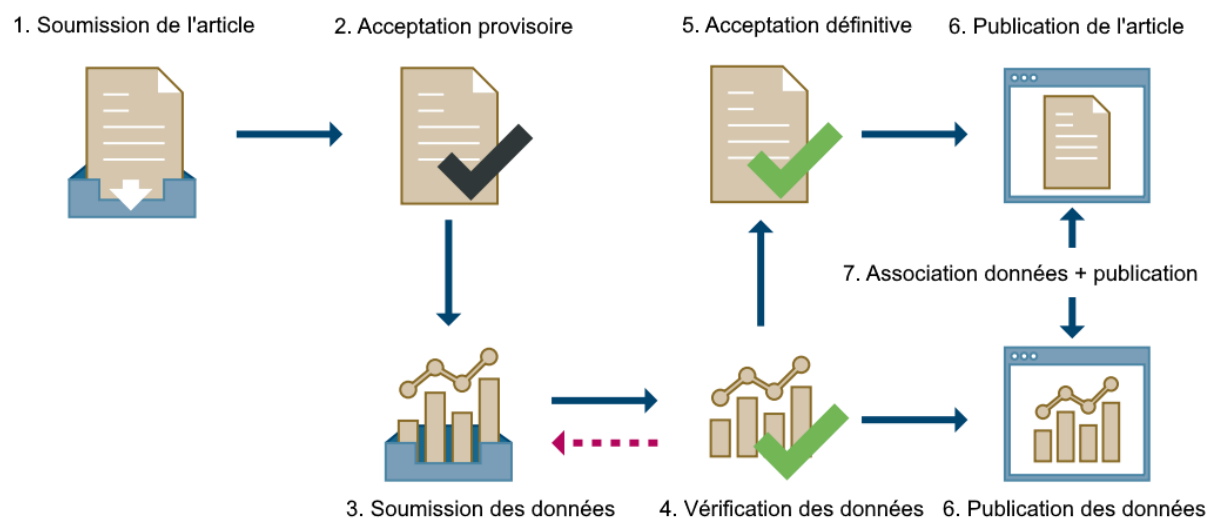
¹²³ Ayer, V., Wiljes, C., Cimiano, P., Pietsch, C., Vompras, J., Schirrwagen, J., & Jahn, N. (2017). Conquaire: Towards an Architecture Supporting Continuous Quality Control to Ensure Reproducibility of Research. *D-Lib Magazine*, 23(1/2). <https://doi.org/10.1045/january2017-ayer>

¹²⁴ Anderson, R. (2021). "Positively Disrupt(ing) Research Culture for the Better": An Interview with Alexandra Freeman of Octopus. *The Scholarly Kitchen*.

<https://scholarlykitchen.sspnet.org/2021/08/18/positively-disrupting-research-culture-for-the-better-an-interview-with-alexandra-freeman-of-octopus/>

¹²⁵ Stodden, V., Seiler, J., & Ma, Z. (2018). An empirical analysis of journal policy effectiveness for computational reproducibility. *Proceedings of the National Academy of Sciences*, 115(11), 2584–2589. <https://doi.org/10.1073/pnas.1708290115>

FIGURE 6. SCHÉMA-TYPE DE CONTRÔLE ÉDITORIAL ET DE VÉRIFICATION PAR UNE REVUE SCIENTIFIQUE (CC BY CHRISTIAN ET AL.)¹²⁶



« Les politiques prescrites par les agences de financement sont utiles, mais je ne suis pas sûr qu'elles soient efficaces... Le financeur ne vérifie généralement pas si vous avez fait ce que vous avez dit que vous alliez faire, et par défaut, cela laisse beaucoup le champ libre pour que tout bonnement les gens s'en abstiennent. »
Un chercheur

Le choix du moment où sont opérés ces contrôles de conformité, que nous avons abordé plus haut dans cette section, est également un point essentiel. Les vérifications après la soumission d'un article exigent que les pairs évaluateurs et/ou le personnel de rédaction évaluent soigneusement un certain nombre d'objets de recherche, y compris si la combinaison appropriée de ressources a été soumise d'emblée : cela peut être difficile en raison des contraintes de temps, subies par les évaluateurs, et des lacunes de connaissances disciplinaires, dans le cas des membres de la rédaction.

« Tout plaide à brève échéance à ce que nous ayons besoin de personnels spécialisés qui comprennent parfaitement ce que cela signifie de rendre les données et le code accessibles et reproductibles. » Un éditeur

Par ailleurs, pour ce qui est de la conformité a posteriori (c'est-à-dire après la publication), ce contrôle intervient, dans de nombreux cas, trop tard dans le processus de publication. En effet, bien qu'il puisse être valable comme moyen de sensibiliser à la reproductibilité, il est peu probable qu'il constitue un outil significatif pour favoriser l'observance à long terme.

Comme notre rapport se focalise sur la reproductibilité en lien avec les pratiques de publication, nous n'avons pas tenté d'examiner les contrôles "ex ante" (les tests que les auteurs ou d'autres collègues peuvent accomplir au cours du cycle de recherche et avant soumission de l'article). Pour compléter, nous précisons que des contrôles ex ante peuvent être

¹²⁶ Christian, T.-M. L., Lafferty-Hess, S., Jacoby, W. G., & Carsey, T. (2018). Operationalizing the Replication Standard. *International Journal of Digital Curation*, 13(1), 114–124. <https://doi.org/10.2218/ijdc.v13i1.555>

pratiqués de différentes manières, par exemple par l'examen du code au sein du groupe de recherche, par des contrôles de reproductibilité effectués par des pairs extérieurs ou par le fait de distribuer des préprints - une pratique qui a connu une popularité croissante au cours des dix dernières années.¹²⁷¹²⁸¹²⁹

« Je pense que l'une des façons d'assurer la reproductibilité est de s'assurer que les erreurs sont détectées le plus tôt possible, et ce qu'il y a de bien dans le fait de détecter une erreur à l'étape du préprint, c'est que vous pouvez la corriger avant qu'elle ne se retrouve dans la version d'enregistrement ou celle soumise à une revue. » Un chercheur

¹²⁷ Puebla, I. (2020). Preprints: a tool and a vehicle towards greater reproducibility in the life sciences. *Journal for Reproducibility in Neuroscience*, 2. <https://doi.org/10.31885/jrn.2.2021.1465>

¹²⁸ Baker, S (2021). Has the Pandemic Changed Research Culture- and is it for the better? *Times Higher Education*. <https://www.timeshighereducation.com/features/has-pandemic-changed-research-culture-and-it-better>

¹²⁹ Chiarelli, A., Johnson, R., Pinfield, S., & Richens, E. (2019). Accelerating scholarly communication: The transformative role of preprints. *Zenodo*. <https://doi.org/10.5281/ZENODO.3357727>

7. Conclusions

Des pratiques de publication propices à la reproductibilité prennent leur essor dans le contexte d'un vaste phénomène de changement culturel dans l'ensemble du paysage de la recherche. Cependant, l'évolution en cours est lente par rapport aux rapides innovations technologiques auxquelles nous assistons. Dans l'ensemble, nous pensons que les chercheurs seront prêts à adopter des pratiques de publication reprocompatibles, pour autant qu'un équilibre soit trouvé entre des attentes croissantes, des contreparties concrètes et des infrastructures adaptées, et à condition que la diversité tienne lieu de principe directeur.

(Dés)équilibres entre infrastructures techniques et culturelles

Nos travaux sur les pratiques reproductibles de publication brossent un tableau optimiste mais complexe. Nous avons constaté que les technologies nécessaires existent déjà pour l'essentiel et que tous les groupes d'acteurs sont attentifs à ce sujet. En même temps, nous avons entendu deux messages forts :

- ▶ la reproductibilité de la recherche ne devrait pas être envisagée seulement au stade de la publication, principalement

en raison de préoccupations liées à la préparation et à la curation des données ou des codes, mais aussi parce qu'elle nécessite un cadre d'ensemble qui englobe les mécanismes de récompenses, les structures d'appui et les financements ; et

- ▶ la reproductibilité de la recherche s'inscrit dans un débat beaucoup plus large sur la science ouverte et il serait très difficile de considérer cet objectif indépendamment des autres bonnes pratiques de recherche ouverte.

TABLEAU 8. INFRASTRUCTURE TECHNIQUE

Dimension	Paramètres essentiels
Infrastructure technique	<ul style="list-style-type: none">● Priorité sur des objets de recherche et des flux de travail spécifiques● Mise en œuvre technique, relativement simple● Manque d'interopérabilité entre les services et les outils● Niveau d'expérimentation significatif
Infrastructures culturelles	<ul style="list-style-type: none">● La reproductibilité fait partie d'un processus de changement culturel beaucoup plus large● Interrelations avec la réplication et d'autres concepts voisins, y compris certains recouvrements en termes de définitions, de cultures et de flux de travail

Nos travaux ont mis en lumière un déséquilibre visible entre les infrastructures culturelles et techniques, montrant que les premières peinent à suivre en raison de l'argument déjà évoqué, sur le fait que la reproductibilité et son évolution s'inscrivent dans un mouvement d'ensemble beaucoup plus large (voir tableau 8). Notre consultation a mis en évidence des solutions et des idées concrètes

pour faciliter les pratiques reproductibles dans la publication, qui sont le plus souvent techniques de par leur nature, ainsi que la complexité du changement culturel, qui est un processus à long terme et ne peut pas être aussi strictement contraint que le champ actuel de nos travaux.

Pratiques reprocompatibles de publication : quelle vision ?

On perçoit chez les chercheurs le désir de travailler et de publier de la meilleure manière possible et on attend d'emblée de leur part qu'ils travaillent au mieux de leurs capacités et respectent les exigences de qualité, d'éthique et d'intégrité. Par conséquent, nous nous attendons à ce qu'ils adopteront volontiers des pratiques reproductibles de publication, à condition que le système dans lequel ils opèrent offre la bonne palette de dispositifs d'appui, de récompense et de reconnaissance.

Les contributeurs au projet ont décrit les pratiques reproductibles de publication comme étant hautement souhaitables pour l'avenir

de la science : à condition d'éviter une culture de culpabilisation, toutes les parties consultées ont estimé qu'il s'agissait d'une trajectoire digne d'intérêt qui entraînera une amélioration à tous les niveaux de la collaboration, et de la transparence, ainsi qu'un partage de l'information. Nous avons toutefois identifié quatre obstacles stratégiques qui devraient entraver les efforts futurs : c'est ce qui est ressorti de notre rapport et qui se trouve synthétisé dans le tableau 9. Certains de ces obstacles sont de nature structurelle (p. ex., les rôles et les responsabilités, la familiarité avec les données de la recherche), tandis que d'autres sont plus opérationnels et peuvent être résolus avec un certain degré d'expérimentation et de compromis (comme la formation et l'offre de soutien, les modèles économiques).

TABLEAU 9. OBSTACLES PRATIQUES À LA REPRODUCTIBILITÉ DES PRATIQUES DE PUBLICATION.

Les pratiques de gestion des données de recherche n'ont pas encore été largement adoptées, ce qui entrave l'adoption de processus de recherche reproductibles.
Afin de permettre la publication de résultats de recherche reproductibles, une certaine forme de partage de données (publiques ou non) est requise. Les pratiques, les prescriptions et les contreparties en matière de gestion des données de recherche ne sont pas assez mûres dans de nombreux domaines, ce qui rend la publication reprocompatible difficile à concrétiser à un niveau général.
Il est nécessaire d'élaborer des listes de contrôle, des lignes directrices et des standards spécifiques à la publication reproductible.
Les financeurs et les éditeurs ne pensent pas être en mesure d'imposer des prescriptions détaillées aux communautés et aux disciplines, car ils interviennent à un niveau plus en amont. La nécessité se situe auprès des communautés et des disciplines qui doivent mettre en commun les prescriptions avec les acteurs de niveaux méso et macro afin que celles-ci puissent se retrouver dans les orientations, les listes de contrôle, les lignes directrices et les standards.
A l'heure actuelle, les modèles économiques et de financement pour couvrir les coûts de publication de recherche reproductible sont expérimentaux et rares.
Bien que certains financeurs subventionnent les coûts de curation et de partage des données, la publication reproductible engendre des contraintes supplémentaires. À ce jour, NWO est la seule agence à le faire explicitement envers les engagements de reproductibilité, mais cette initiative n'est pas considérée comme une approche à long terme. Des réflexions supplémentaires doivent être menées. Toutefois, le présent rapport propose une série de pistes de départ pour d'autres débats à venir autour des modèles économiques.
La formation à la publication reproductible est une affaire complexe, et où situer son intervention reste flou.
La formation et le soutien aux pratiques de publication reproductibles relèvent actuellement d'un large éventail de groupes d'intervenants, comme les sociétés savantes, les organismes de recherche et les éditeurs, notamment. Au sein des organismes de recherche, la création de fonctions comme celles d'intendants de données, de curateurs ou de bibliothécaires de spécialité disciplinaire pourrait montrer leur utilité, mais cela implique une volonté stratégique de la part des dirigeants.

Aujourd'hui, le plus grand obstacle pratique à la publication reprocompatible tient à ce que, dans certaines disciplines, la notion de par-

tage des données reste quelque peu incomprise (notamment dans certaines branches de recherche qualitative en sciences humaines et

sociales, mais à l'exclusion notable de certaines d'entre elles comme en économie ou en psychologie). Dans ces cas, partager des « données » pour valider ses publications pose un problème de taille et paraît être une exigence en partie prématurée.

Conduire le changement

Nos recherches font ressortir fortement le fait que les communautés et les disciplines de recherche auront forcément un rôle à jouer en faisant état de leurs besoins, ce qui permettra de s'assurer que de nouvelles préconisations seront fondées sur la réalité de leurs pratiques de recherche au quotidien. Une interprétation sous forme de cycle politique (voir [figure 3](#), page 36) montre clairement qu'il s'agit d'une condition nécessaire pour pouvoir passer à la mise en œuvre, l'application et l'évaluation des politiques.

La médaille a pour revers que les chercheurs des différents horizons disciplinaires ne seront pas intrinsèquement portés à faciliter le mouvement s'ils ne perçoivent pas les avantages apportés par les pratiques de publication reproductibles et n'en font pas l'expérience par eux-mêmes. Nous pensons que les organismes de financement de la recherche peuvent jouer un rôle important en facilitant d'autres débats, en premier lieu en examinant et en donnant crédit aux réflexions allant vers la reproductibilité dans le contexte des demandes de subvention. En outre, les mécanismes de financement (potentiellement au niveau international) pourraient porter de vastes consultations disciplinaires pour aider à élaborer des listes de contrôle, des lignes directrices et des standards qui seraient opérants pour la plupart des acteurs (p. ex., couvrir l'éventail approprié d'objets de recherche dans chaque discipline). Idéalement, des politiques concrètes et adaptées devraient émerger organiquement à partir de la base vers le haut et commencer par être adoptées par les revues et les fournisseurs d'infrastructures, puis généralisées et mises en œuvre par les organismes de financement de la recherche et les acteurs politiques.

En outre, il est largement admis que les organismes de financement de la recherche et les éditeurs disposent de possibilités pour favoriser l'observance de pratiques reproductibles de publication. Les premiers peuvent orienter les exigences et les pratiques de recherche par le biais des préconisations que stipulent leurs orientations, tandis que les seconds peuvent imposer des listes de contrôle, des lignes directrices et des standards applicables par les auteurs, leurs personnels de rédaction et les pairs évaluateurs. Dans l'idéal, il serait avantageux que les préconisations des agences de financement et des revues puissent évoluer de concert, au moins dans une certaine mesure. Ainsi, certaines revues pourraient craindre que des exigences trop strictes ne découragent les auteurs et définir des règles dans la concertation pourrait contribuer à assurer une harmonisation entre politiques et préconisations en matière de reproductibilité pratiquée dans les publications.

Nous tenons à souligner que les politiques et leur application risquent de laisser peu de place à la nuance. Pour certaines cultures épistémiques, par exemple, la reproductibilité sera plus difficile à comprendre et à mettre en œuvre, ou peut-être n'est-elle même pas un but en soi ; dans d'autres, la reproductibilité peut ne pas être considérée comme la marque de qualité essentielle, mais simplement comme une option parmi bien d'autres. Il sera donc nécessaire de prendre en compte la diversité lorsque nous repenserons les pratiques de recherche pour préserver et renforcer la confiance dans la science.

Dans l'ensemble, nous croyons que nos constatations dessinent une perspective à la fois prudente et optimiste : en s'attaquant aux obstacles identifiés et en tirant parti des possibilités existantes, on dispose d'une marge importante pour que les pratiques de publication compatibles avec la reproductibilité jouent un rôle de premier plan dans l'évolution actuellement en cours vers la science ouverte.

8. Références

Données produites par l'étude de Knowledge Exchange

- Chiarelli, A., & Loffreda, L. (2021). Publishing Reproducible Research Outputs - Interviewees and interview questions (Version 1) [Data set]. Zenodo. <https://doi.org/10.5281/ZENODO.5141665>
- Chiarelli, A., Loffreda, L., & Johnson, R. (2021). Publishing Reproducible Research Outputs - Literature findings. Zenodo. <https://doi.org/10.5281/ZENODO.4675457>
- Chiarelli, A. (2021). Publishing Reproducible Research Outputs - Text and data mining code (Version 2) [Computersoftware]. Zenodo. <https://doi.org/10.5281/ZENODO.5512019>
- Loffreda, L., & Chiarelli, A. (2021). Publishing Reproducible Research Outputs - Data Management Plan. Zenodo. <https://doi.org/10.5281/ZENODO.4697400>
- Loffreda, L. & Chiarelli, A. (2021). Publishing Reproducible Research Outputs - Thematic coding of interview findings (Version 1) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.5512420>
- Loffreda, L., & Chiarelli, A. (2021). Publishing Reproducible Research Outputs - Literature sources and Thematic coding (Version 1) [Data set]. Zenodo. <https://doi.org/10.5281/ZENODO.4748748>

Bibliographie

- Advancing Earth and Space Science. (n.d.). Data and Software for Authors. <https://bit.ly/3mH8Umz>
- American Economic Association. (n.d.). Policy and Protocol on Third-Party Verifications. <https://bit.ly/3lsmJ90>
- American Statistical Association (ASA). (n.d.). JASA Editors Talk Reproducibility. <https://bit.ly/3iQnpmU>
- Anderson, R. (2021). "Positively Disrupt(ing) Research Culture for the Better": An Interview with Alexandra Freeman of Octopus. The Scholarly Kitchen. <https://bit.ly/3Dwlmfu>
- Association for Computing Machinery (ACM). (2021). Practices Working Group Blog. <https://bit.ly/3njmXz3>
- Association for Computing Machinery. (n.d.). Reproducibility PRINCIPLES: Taking the pulse. <https://bit.ly/3jazl39>
- Auer, S., Haeltermann, N. A., Weissberger, T. L., Erlich, J. C., Susilaradeya, D., Julkowska, M., Gazda, M. A., Schwessinger, B., Jadavji, N. M., Abitua, A., Niraulu, A., Shah, A., Clyburne-Sherinb, A., Guiquel, B., Alicea, B., LaManna, C., Ganguly, D., Perkins, E. & Jambor, H. (2021). A community-led initiative for training in reproducibleresearch. eLife. <https://doi.org/10.7554/elife.64719>
- Australian Reproducibility Network. (n.d.). Welcome to the Australian Reproducibility Network. <https://bit.ly/2Z5pEvv>
- Ayer, V., Wiljes, C., Cimiano, P., Pietsch, C., Vompras, J., Schirrwagen, J., & Jahn, N. (2017). Conquaire: Towards an Architecture Supporting Continuous Quality Control to Ensure Reproducibility of Research. D-Lib Magazine, 23(1/2). <https://doi.org/10.1045/january2017-ayer>
- Baker, S. (2021). Has the Pandemic Changed Research Culture- and is it for the better? Times Higher Education. <https://bit.ly/3BwXWWV>
- Biometrical Journal. (n.d.). Biometrical Journal [Author Guidelines](https://bit.ly/3BvGUbK). <https://bit.ly/3BvGUbK>
- Black, S., Gardner, D. G., Pierce, J. L., & Steers, R. (2019). Organisational Behaviour. <https://bit.ly/2ZaOaMi>
- Bowman, S. (2020). Create a Project. OSF. <https://bit.ly/3mHd3a7>

Breuer, C., & Trilcke, P. (2021). Expanding academic publishing practices alongside the digital turn. »Scientificpractice« working group of the Priority Initiative »Digital Information« by the Alliance of Science Organisations. <https://doi.org/10.48440/ALLIANZOA.042>

Brinton, J. (2021). The impact of COVID-19 on the UK publishing industry: Findings and opportunity. *LearnedPublishing*, 34(1), 43–48. <https://doi.org/10.1002/leap.1363>

Burgelman, J. C., Pascu, C., Szkuta, K., Von Schomberg, R., Karalopoulos, A., Repanas, K., & Schoupe, M. (2019). Open Science, Open Data, and Open Scholarship: European Policies to Make Science Fit for the Twenty-First Century. *Frontiers in Big Data*, 2(43), 1-6. <https://doi.org/10.3389/fdata.2019.00043>

Center for Open Science. (n.d). Top Guidelines. <https://bit.ly/3jp5sfh>

Center for Open Science. (n.d.). Sample Implementation of Guidelines for Transparency and Openness Promotion(TOP). *Journal Policies and Practices*. <https://osf.io/edtxm>

Chambers, C. D. (2020). Verification Reports: A new article type at Cortex. *Cortex*, 129, A1–A3. <https://doi.org/10.1016/j.cortex.2020.04.020>

Chiarelli, A., Johnson, R., Pinfield, S., & Richens, E. (2019). Accelerating scholarly communication: The transformative role of preprints. Zenodo. <https://doi.org/10.5281/ZENODO.3357727>

Christian, T, M. L., Lafferty-Hess, S., Jacoby, W. G., & Carsey, T. (2018). Operationalizing the Replication Standard - A Case Study of the Data Curation and Verification Workflow for Scholarly Journals. *International Journal of DigitalCuration*, 13(1), 114–124. <https://doi.org/10.2218/ijdc.v13i1.555>

Conrad, L. Y. (2021). Digital Transformation Requires Cultural Evolution. *The Scholarly Kitchen Blog*. <https://bit.ly/3mJUzFH>

Couture, J. L., Blake, R. E., McDonald, G., & Ward, C. L. (2018). A funder-imposed data publication requirementseldom inspired data sharing. *PLOS ONE*, 13(7), e0199789. <https://doi.org/10.1371/journal.pone.0199789>

Dan, V. (2017). Empirical and Nonempirical Methods. *The International Encyclopedia of Communication ResearchMethods*. 1-3. <https://doi.org/10.1002/9781118901731.iecrm0083>

Declaration on Research Assessment (DORA). (n.d.). Sign DORA. <https://bit.ly/3m1bNiQ>

Department for Business, Energy & Industrial Strategy. (2019). Realising the potential: Final report of the OpenResearch Data Task Force. <https://bit.ly/3FLjrG2>

Department for Business, Energy & Industrial Strategy. (2021). UK Research and Development Roadmap (PolicyPaper). <https://bit.ly/3BA7gJq>

Dewald, W. G., Thursby, J. G., & Anderson, R. G. (1986). Replication in empirical economics: The journal of money, credit and banking project. *The American Economic Review*, 76(4), 587-603. <https://bit.ly/3pqB2gr>

Digital Science, Hahnel, M., Borrelli, L. M., Hyndman, A., Baynes, G., Merce Crosas, Nosek, B., Shearer, K., Selm, V., Goodey, G., & Nature Research. (2020). The State of Open Data 2020. <https://doi.org/10.6084/M9.FIGSHARE.13227875.V2>

DORA (Twitter). (2021) The German Research Foundation (@dfg_public) has signed DORA. <https://bit.ly/3DvehvP>

Dutch Research Council (NWO). (2020). Replication Studies third round: repetition of important research. <https://bit.ly/3mLFJOW>

Dutch Research Council (NWO). (n.d.). Replication Studies. <https://bit.ly/3v3u568>

eLife. (2021). eLife and Stencila announce roadmap for bringing reproducible publishing to more authors. <https://bit.ly/3mHdLUP>

Elsevier. (n.d). Verification Reports – Guide for Authors. <https://bit.ly/3oRuDuD>

Elsevier. (n.d.). Information Systems. <https://bit.ly/3apAFdF>

EQUATOR Network. (n.d.). Enhancing the QUALity and Transparency Of health Research <https://bit.ly/3AxXGFA>

European Commission. (n.d.). Facts and Figures for Open Research Data. <https://bit.ly/3lqleH7>

European Commission. (n.d.). Open Research Europe. <https://bit.ly/3DwgMy0>

European Strategy Forum on Research Infrastructures (ESFRI). (2021). ESFRI announces new RIs for Roadmap 2021. <https://bit.ly/2XyQOur>

FAIRsFAIR. (2020). FAIRsFAIR Tables Preliminary Recommendations on Data Policy. <https://bit.ly/3mBhPWw>

Fortunato, L., & Galassi, M. (2021). The case for free and open source software in research and scholarship. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2197). <https://doi.org/10.1098/rsta.2020.0079>

German Reproducibility Network. (n.d.). German Reproducibility Network. <https://bit.ly/3m0KuFe>

GitHub (n.d.). ReScience C: Frequently Asked Questions. <https://bit.ly/3C2z7lV>

GitHub. (n.d.). Making your Code Citable. <https://bit.ly/3b0xB86>

GitHub Topic: research-compedium. <https://bit.ly/3B1RaYe>

GO FAIR (n.d.). Fair Principles. <https://bit.ly/3njodSN>

Goeva, A., Stoudt, S., & Trisovic, A. (2020). Toward Reproducible and Extensible Research: from Values to Action. *Harvard Data Science Review*, 2(4). <https://doi.org/10.1162/99608f92.1cc3d72a>

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429–431. <https://doi.org/10.1038/520429a>

Hurlin, C., & Perignon, C. (2019). Reproducibility Certification in Economics Research. SSRN. <https://doi.org/10.2139/ssrn.3418896>

Image Processing On Line (IPOP). (n.d.). Image Processing On Line. <https://bit.ly/3E25YrD>

Imperial College London. (n.d.). Subject Support. <https://bit.ly/3lt4CzA>

Janssen, M., & Helbig, N. (2018). Innovating and changing the policy-cycle: Policy-makers be prepared! *Government Information Quarterly*, 35(4), S99–S105. <https://doi.org/10.1016/j.giq.2015.11.009>

Jisc. (2018). Meeting the requirements of the EPSRC research data policy. <https://bit.ly/3mCHVbE>

Knowledge Exchange. (2017). Open Scholarship Framework. <https://bit.ly/3GbkzCP>

Kurapati, S. & Teperek, M. (2019). What does reproducibility mean for qualitative research? <https://bit.ly/3v1RlVl>

Kurapati, S. (2019). Becoming a Data Steward. LSE Impact Blog. <https://bit.ly/3v02mmF>

Loffreda, L., & Chiarelli, A. (2021). Publishing Reproducible Research Outputs - Data Management Plan. Zenodo. <https://doi.org/10.5281/ZENODO.4697399>

Loffreda, L., & Chiarelli, A. (2021). Publishing Reproducible Research Outputs - Literature sources and Thematic coding (Version 1) [Data set]. Zenodo. <https://doi.org/10.5281/ZENODO.4748748>

Luyben, K. (2021). EOSC: A Web of Fair Data. EOSC. <https://bit.ly/3ltyls7>

Maastricht University. (n.d.). Data stewardship services. <https://bit.ly/3BzMObH>

Markowetz, F. (2015). Five selfish reasons to work reproducibly. *Genome Biology*, 16(1). <https://doi.org/10.1186/s13059-015-0850-7>

Matthews, D. (2021). European Research Council bans journal impact factor from bids. *Times Higher Education*. <https://bit.ly/3v2Rggw>

National Academies of Sciences, Engineering, and Medicine. (2019). *Reproducibility and Replicability in Science*. The National Academies Press. <https://doi.org/10.17226/25303>

National Centre for the Replacement, Refinement & Reduction of Animals in Research. (n.d.). About. <https://bit.ly/3jqy5J5>

Nüst, D., Boettiger, C., & Marwick, B. (2018). How to Read a Research Compendium. arXiv. <https://bit.ly/2ZjMHmF>

Nüst, D. & Eglén, S. J. (2021). CODECHECK: an Open Science initiative for the independent execution of computations underlying research articles during peer review to improve reproducibility [version 2; peer review: 2 approved]. *F1000Research*, 10(253). <https://doi.org/10.12688/f1000research.51738.2>

Obels, P., Lakens, D., Coles, N. A., Gottfried, J., & Green, S. A. (2020). Analysis of Open Data and Computational Reproducibility in Registered Reports in Psychology. *Advances in Methods and Practices in Psychological Science*, 3(2), 229–237. <https://doi.org/10.1177/2515245920918872>

Office of the AEA Data Editor. (2021). Guidance on testing replicability of code. <https://bit.ly/3FEbgen>

Office of the AEA Data Editor. (2021). Office of the AEA Data Editor. <https://bit.ly/3jorAq5>

OpenAIRE. (n.d.). How to identify and assess Research Data Management (RDM) costs. <https://bit.ly/2Ylh7yA>

Openly Scholarly Infrastructure. (2020). *The Principles of Open Scholarly Infrastructure*. <https://bit.ly/3G76vKK>

Ouvrir la Science. (n.d.). Second National Plan for Open Science. <https://bit.ly/30Dsokq>

Peer, L., Arguillas, F., Honeyman, T., Miljković, N., Gehlen, K. & CURE-FAIR WG Subgroup 3. (2021). Challenges of Curating for Reproducible and FAIR Research Output. *Research Data Alliance*. <https://doi.org/10.15497/RDA00063>

Penders, B., Holbrook, J., & de Rijcke, S. (2019). Rinse and Repeat: Understanding the Value of Replication across Different Ways of Knowing. *Publications*, 7(3), 52. <https://doi.org/10.3390/publications7030052>

Perkel, J. M. (2021). Reactive, reproducible, collaborative: computational notebooks evolve. *Nature*, 593(7857), 156–157. <https://doi.org/10.1038/d41586-021-01174-w>

Perkel, J. M. (2021). Ten computer codes that transformed science. *Nature*, 589(7842), 344–348. <https://doi.org/10.1038/d41586-021-00075-2>

Plomp, E., Andrews, H., Love, J., den Heijer, K., Dintzner, N., Ilamparuthi, S., Wang, Y., & Turkyilmaz-van der Velden, Y. (2021). Communicating Open Science at TU Delft. *Open Working*. <https://bit.ly/3mKQ8dM>

Popper, K. (2005; originally published in 1934). *The logic of scientific discovery*. Routledge.

Puebla, I. (2020). Preprints: a tool and a vehicle towards greater reproducibility in the life sciences. *Journal for Reproducibility in Neuroscience*, 2. <https://doi.org/10.31885/jrn.2.2021.1465>

RENKU. (n.d.). Collaborative Data Science. <https://bit.ly/3vwvk6f>

ReScience C. (n.d.). Ten Years Reproducibility Challenge. <https://bit.ly/3GblBPd>

Research Information. (2020). SciCrunch announces Luxembourg collaboration. <https://bit.ly/3awHrhq>

Rice, R. (2021). Welcome to Research Data Stewards A New Role. The University of Edinburgh. <https://bit.ly/3Dw6ZYQ>

Rocca-Serra, P. & Sansone, S, A. (2021). The FAIR Cookbook – a guide for your data FAIRification journey. <https://bit.ly/3ltxxn3>

Rosenberg, D. E., Jones, A. S., Filion, Y., Teasley, R., Sandoval-Solis, S., Stage, J. H., Abdallah, A., Castronova, A., Ostfeld, A., & Watkins, D., Jr. (2021). Reproducible Results Policy. *Journal of Water Resources Planning and Management*, 147(2).
[https://doi.org/10.1061/\(asce\)wr.1943-5452.0001368](https://doi.org/10.1061/(asce)wr.1943-5452.0001368)

Salvato, G. (Twitter). (2021) The Italian Reproducibility Network, a star is about to be born! Stay tuned for more information. <https://bit.ly/3DsslFW>

Silver, A. (2017). Collaborative Software Development Made Easy. *Nature*. 550, 143-144.
<https://doi.org/10.1038/550143a>

Slovak Reproducibility Network. (2021). Slovak Reproducibility Network. <https://bit.ly/3jiZ9tR>

Stencila. (n.d.). Executable document pipelines. <https://bit.ly/3aXfkZ0>

Stodden, V., Seiler, J., & Ma, Z. (2018). An empirical analysis of journal policy effectiveness for computational reproducibility. *Proceedings of the National Academy of Sciences*, 115(11), 2584–2589. <https://doi.org/10.1073/pnas.1708290115>

Swiss Reproducibility Network. (n.d.). SWISSRN. <https://bit.ly/2ZkpCAe>

Teperek, M., & Plomp, E. (2019). The role and value of data stewards in Universities: a TU Delft case study on data stewardship. <https://doi.org/10.5281/ZENODO.2684278>

The Academy of Medical Sciences. (n.d.) Reproducibility and reliability of biomedical research. <https://bit.ly/2YCCnpc>

The Turing Way Community., & Scriberia. (2021). Illustrations from the Turing Way book dashes. Zenodo. <https://doi.org/10.5281/ZENODO.4906004>

Tijdink, J. K., Horbach, S. P. J. M., Nuijten, M. B., & O’Neill, G. (2021). Towards a Research Agenda for Promoting Responsible Research Practices. *Journal of Empirical Research on Human Research Ethics*, 16(4), 450–460. <https://doi.org/10.1177/15562646211018916>

Turkylmaz-van der Velden, Y., Dintzner, N., & Teperek, M. (2020). Reproducibility Starts from You Today. *Patterns*.1(6). <https://doi.org/10.1016/j.patter.2020.100099>

Twitter. (2021). Ben Greiner Tweet Status. <https://bit.ly/306JugE>

Twitter. (n.d.). About replies and mentions. <https://bit.ly/3Cfp7Wk>

UK Data Services. (2015). Costing Tool. <https://bit.ly/3B27Oqz>

UK Parliament. (n.d.). Reproducibility and research integrity. <https://bit.ly/3mFJdmh>

UK Reproducibility Network. (n.d.). Welcome to the UK Reproducibility Network. <https://bit.ly/3josvqx>

UK Reproducibility Network. (2021). Major funding boost for UK’s open research agenda. <https://bit.ly/3ByvW55>

UKRI. (n.d.). Research integrity. <https://bit.ly/2YFvmnL>

Université Du Luxembourg. (n.d). Responsible and Reproducible Research (R3) <https://bit.ly/3mJVFBj>

University College London (UCL). (n.d.). Costing Research Data Management. <https://bit.ly/3v8zSqT>

Vilhuber, L. (2021). Report by the AEA Data Editor. *AEA Papers and Proceedings*, 111: 808-17. <https://doi.org/10.1257/pandp.111.808>

Wellcome. (n.d.). Guidance for research organisations on how to implement responsible and fair approaches for research assessment. <https://bit.ly/3Aw2E5O>

Westwood, S. (Twitter). (2021). I’m curious, what got you involved in Open Science/ replication crises? <https://bit.ly/3arjKan>

Whitehouse, T., Clark, D., & Tsang, E. (2019). Rebooting reproducibility: From re-execution to replication. *eLife*. <https://bit.ly/3Dsuems>

Willis, C., & Stodden, V. (2020). Trust but Verify: How to Leverage Policies, Workflows, and Infrastructure to Ensure Computational Reproducibility in Publication. *Harvard Data Science Review*, 2(4). <https://doi.org/10.1162/99608f92.25982dcf>

Woolston, C. (2021). Impact factor abandoned by Dutch university in hiring and promotion decisions. *Nature*, 595(7867), 462–462. <https://doi.org/10.1038/d41586-021-01759-5>

World Bank. (2020). Under Pressure: Covid 19 and funding European Higher Education. <https://bit.ly/3uZi95g>

Yeston, J.S. (2021). Progress in data and code deposition. <https://bit.ly/2YBxugl>

Zenodo. (n.d.). Communities. <https://bit.ly/30P9cQX>

Zenodo. (n.d.). Research Compendium Community. <https://bit.ly/3iTmKB1>

9. Annexe A – Remerciements et participants

Tableau A1 : Groupe de travail de Knowledge Exchange sur la publication de résultats de recherche reproductibles

Nom	Affiliation	Fonction	Pays
Anna Mette Morthorst	Danish e-infrastructure cooperation (DeiC)	Chargée de mission	Danemark
Birgit Schmidt	Université de Göttingen	Responsable de Knowledge Commons	Allemagne
Birte Christensen-Dalsgaard	Université d'Aarhus	Analyste en chef	Danemark
Daniel Nüst	Université de Münster	Chercheur	Allemagne
Jeroen Sondervan	Université d'Utrecht	Conseiller en édition	Pays-Bas
Juliane Kant	Deutsche Forschungsgemeinschaft (DFG)	Chargée de programme – Services aux bibliothèques scientifiques et systèmes documentaires	Allemagne
Matt Jaquierey	Université d'Oxford	Chercheur	R.-U.
Pierre-Carl Langlais	Paris Sorbonne-CELSA	Chercheur associé	France
Saskia Woutersen	Université de Leyde	Bibliothécaire en sciences numériques	Pays-Bas
Verena Heise	Indépendante	Chercheuse en science ouverte, formatrice et consultante	Allemagne
Yrsa Neuman	Université Åbo Akademi	Coordinateur en chef – Science responsable/Science ouverte	Finlande

Tableau A2 : Contributeurs au projet

Nom	Affiliation	Fonction	Pays
Alexandra Freeman	Winton Centre for Risk and Evidence Communication, Université de Cambridge	Directeur exécutif	Royaume-Uni
Amy Orben	Université de Cambridge, founder of ReproducibiliTea	Fellow	Royaume-Uni

Nom	Affiliation	Fonction	Pays
Andreas Bleuler	Swiss Data Science Center Renku	Informaticien principal	Suisse
Anita Bاندrowski	UC San Diego SciScore	Spécialiste en neurosciences	Etats-Unis
Bart Penders	Université de Maastricht University	Professeur associé	Pays-Bas
Bertram Ludaescher	Center for Informatics Research in Science and Scholarship, Université de l'Illinois à Urbana-Champaign Whole Tale	Directeur, Professeur	Etats-Unis
Brian Nosek	Center for Open Science	Cofondateur, directeur exécutif	Etats-Unis
Carlien Hillebrink	Dutch Research Council	Conseiller politique	Pays-Bas
Carsten Bergholtz	Université d'Aarhus	Professeur associé	Danemark
Catherine Bromley	Economic and Social Research Council	Directrice adjointe de la stratégie et de l'infrastructure des données	Royaume-Uni
Catriona MacCallum	Hindawi	Directrice de la science ouverte	Royaume-Uni
Christophe Perignon	HEC Paris	Professeur de finance, vicedoyen pour la recherche	France
Daniël Lakens	Université de Technologie d'Eindhoven	Professeur associé	Pays-Bas
Emma Ganley	Protocols.io	Directrice des initiatives stratégiques	Etats-Unis
Erik Schultes	GO FAIR International Support and Coordination Office	Coordinateur scientifique international	Pays-Bas
Gordon Feld	Central Institute of Mental Health	Chef du groupe de recherche sur la psychologie et la neurobiologie du sommeil et de la mémoire	Allemagne
Heidi Seibold	Johner Institut	Chercheuse	Allemagne
Jana Lasser	TU Graz	Chercheur postdoctoral	Autriche

Nom	Affiliation	Fonction	Pays
Jean-Sébastien Caux	Université d'Amsterdam SciPost	Professeur, président de SciPost	Pays-Bas
Jennifer Gibson	Interviewed in her role at eLife, currently at Dryad	Ancienne responsable de la communication ouverte de la recherche (eLife) Directrice exécutive (Dryad)	Royaume-Uni
Karel Luyben	EOSC Association Université de Technologie de Delft	Président de l'association EOSC, Rector Magnificus Emeritus	Pays-Bas
Katja Mankinen	CSC – IT Center for Science	Data scientist	Finlande
Kelly Lloyd	Université de Leeds	Étudiant en doctorat	Royaume-Uni
Kristina Hettne	Bibliothèques de l'Université de Leyde	Bibliothécaire en sciences numériques	Pays-Bas
Ksenija Bazdaric	Faculté de Médecine de Rijeka European Science Editing	Professeur assistant au département d'informatique médicale Rédacteur en chef	Croatie
Liz Bal	Jisc	Directeur des services de recherche ouverts	Royaume-Uni
Lorena Barba	Université George Washington	Professeur en génie mécanique et aérospatial	Etats-Unis
Malvika Sharan	The Turing Institute	Community manager	Royaume-Uni
Marcus Munafò	Université de Bristol	Professeur	Royaume-Uni
Mari Riipinen	Université de Turku	Chef de l'unité de développement de la recherche	Finlande
Marin Dacos	Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation	Conseiller en science ouverte	France
Mark Hahnel	Figshare	Fondateur, PDG	Royaume-Uni
Martina Grunow	Leibniz Information Centre for Economics	Chef de projet (services de publication)	Allemagne
Martine Garnier-Rizet	Agence nationale de la recherche	Chef du département de technologie numérique et de	France

Nom	Affiliation	Fonction	Pays
		mathématiques	
Mary Donaldson	Université de Glasgow	Service Coordinator (Research Data Management Service)	Royaume-Uni
Matti Ruuskanen	Université de Turku	Chercheur postdoctoral	Finlande
Philipp Cimiano	Université de Bielefeld	Professeur, Chef du Groupe de l'informatique sémantique	Allemagne
Philippe Bonnet	Université IT de Copenhague	Professeur	Danemark
Rémy Mosseri	Centre national de la recherche scientifique	Chargé de l'intégrité scientifique	France
Rik Henson	British Neuroscience Association	Président élu	Royaume-Uni
Sarah Callaghan	Patterns - Cell Press	Rédacteur en chef	Royaume-Uni
Simone Sacchi	Institut universitaire européen	Bibliothécaire pour les sciences ouvertes	Italie
Sonja Ochsenfeld-Repp	Deutsche Forschungsgemeinschaft (DFG)	Chef de division - Égalité des chances, intégrité de la recherche et développement inter-programmes	Allemagne
Sven Vlaeminck	Leibniz Information Center for Economics	Chef de projet (données de recherche)	Allemagne
Tim Smith	CERN Zenodo	Chef du groupe Collaboration, Appareils et Applications, Membre du comité directeur de Zenodo	Suisse
Tyler Whitehouse	Gigantum	Cofondateur, PDG	Etats-Unis
Ulrich Dirnagl	Hôpital universitaire de la Charité de Berlin	Directeur du département de neurologie expérimentale	Allemagne
Ulrik Nicolai de Lichtenberg	Fondation Novo Nordisk	Senior Scientific Manager	Danemark
Valerie McCutcheon	Université de Glasgow	Responsable de l'information sur la recherche (services aux bibliothèques)	Royaume-Uni
Vicky Rampin	Université de New York	Bibliothécaire - Gestion des données de recherche et reproductibilité	Etats-Unis

Nom	Affiliation	Fonction	Pays
Victoria Moody	Jisc	Directeur de la stratégie du secteur de la recherche et de l'innovation	Royaume-Uni

10. Annexe B – Structure des entretiens

Cette annexe comprend une liste de questions d'entretiens utilisées pour les besoins de cette étude. Les questions ont été posées en fonction du groupe d'appartenance des interlocuteurs et de leur temps disponible, ce qui signifie que toutes les questions n'ont pas été posées à toutes les personnes interrogées.¹³⁰

Légende:

- ▢ Organismes de financement de la recherche : OFR
- ▢ Organismes de recherche : OR
- ▢ Fournisseurs d'infrastructures : FI
- ▢ Éditeurs académiques : EA
- ▢ Chercheurs et groupes de recherche : CGR

Questions posées	OFR	OR	FI	EA	CGR
Dans quelle mesure notre définition de la reproductibilité vous est familière ? Connaissez-vous d'autres contextes dans lesquels ce terme est utilisé ?	✓	✓	✓	✓	✓
Pour quelles raisons la communauté de la recherche devrait considérer la reproductibilité comme faisant partie des pratiques quotidiennes de la recherche ?	✓	✓	✓	✓	
Sentez-vous un besoin d'agir au sujet de la reproductibilité de la recherche, surtout durant la phase de publication ?	✓				✓
Comment abordez-vous la question de la reproductibilité de la recherche dans vos politiques en tant que financeur de la recherche ?	✓				
Quelles récompenses et incitations sont propices aux pratiques de recherche reproductibles ?	✓	✓		✓	✓
Qui devrait se charger des vérifications de reproductibilité, et à quelle étape des processus de soumission, d'évaluation et de publication ?	✓	✓	✓	✓	✓
Comment les financeurs peuvent-ils permettre et contrôler les pratiques de recherche reproductibles par rapport à leurs politiques ?	✓				
Quels modèles économiques pourraient être appropriés pour couvrir les coûts de la reproductibilité (par exemple, du « chercheur payeur » au « financeur payeur ») ?	✓	✓		✓	✓
Dans quelle mesure la communauté a-t-elle besoin d'infrastructures publiques soutenant la reproductibilité, étant donné que	✓			✓	✓

¹³⁰ Chiarelli, A., & Loffreda, L. (2021). Publishing Reproducible Research Outputs - Interviewees and interview questions (Version 1) [Data set]. Zenodo. <https://doi.org/10.5281/ZENODO.5141665>

Questions posées	OFR	OR	FI	EA	CGR
les services commerciaux/revues développent leurs propres solutions ?					
Considérez-vous la reproductibilité comme un objectif souhaitable à long terme ?	✓	✓	✓	✓	✓
Les chercheurs connaissent-ils la gamme d'outils qu'ils peuvent utiliser pour permettre la publication de résultats de recherche reproductibles ?		✓	✓		
Comment avez-vous développé votre offre autour de la reproductibilité ? Par exemple, comment avez-vous identifié et délimité les besoins ?			✓		
L'infrastructure de recherche actuelle manque-t-elle de fonctionnalités ou de solutions techniques permettant la reproductibilité ?		✓	✓	✓	✓
Connaissez-vous les principes FAIR, et votre infrastructure les applique-t-elle ?			✓		
Quel est votre modèle de financement actuel, et dans quelle mesure est-il durable ?			✓		
Comment voyez-vous l'évolution du paysage de la reproductibilité en termes de services disponibles pour les utilisateurs finaux ? Ces services proviendront-ils principalement de services commerciaux/revues ou d'infrastructures publiques ?		✓	✓		
Selon vous, quel est le rôle des éditeurs en matière de reproductibilité ? Veuillez considérer le rôle des éditeurs et des coordinateurs de publication comme fournisseurs d'infrastructure.				✓	
Quels rôles (potentiellement nouveaux) sont nécessaires pour permettre la vérification de la reproductibilité des manuscrits ?				✓	
Quels outils ou infrastructures pour la reproductibilité connaissez-vous ?				✓	
La reproductibilité de la recherche intéresse-t-elle l'ensemble de votre institution ?		✓			
Dans quelle mesure les liens entre les outils et logiciels existants utilisés par les chercheurs influent-ils sur la publication de résultats de recherche reproductibles ?		✓			
Les chercheurs reçoivent-ils le niveau approprié de formation et de soutien dont ils ont besoin pour publier de manière reproductible et vérifier la reproductibilité du travail des autres ?		✓			
Votre organisation est-elle préoccupée par les risques potentiels découlant d'une adoption plus large des pratiques de partage de		✓			

Questions posées	OFR	OR	FI	EA	CGR
données et de codes (par exemple, le partage d'informations sensibles ou confidentielles) ?					
Comment voyez-vous l'évolution du paysage de la reproductibilité en termes de services disponibles pour les utilisateurs finaux ? Ces services proviendront-ils principalement de services commerciaux/revues ou d'infrastructures publiques ?		✓	✓		
Dans quelle mesure vos pairs et collaborateurs sont-ils sensibilisés à la reproductibilité de la recherche, notamment au stade de la publication ?					✓
Dans quelle mesure les efforts actuels en matière de reproductibilité de la recherche sont-ils suffisants pour susciter le changement ? Veuillez prendre en compte les groupes d'acteurs que vous considérez comme importants.					✓
Qu'est-ce qui doit changer dans le processus de la recherche pour pouvoir assumer financièrement les coûts associés à la reproductibilité ? Par exemple, une recherche plus lente, un financement plus important, des équipes de recherche plus grandes, la réduction d'un autre aspect (spécifique) de la recherche... ?					✓
Quelles sont, selon vous, les évolutions principales et les plus importantes en matière de reproductibilité au stade de la publication et de la diffusion ? Veuillez prendre en compte les évolutions technologiques et culturelles.					✓
Qui, selon vous, est le principal moteur de l'innovation en matière de pratiques de recherche reproductibles ?					✓
Quels sont, selon vous, les principaux obstacles à la publication de résultats de recherche reproductibles ?	✓	✓	✓	✓	✓
Recevez-vous le niveau approprié de formation et de soutien dont vous avez besoin pour publier de manière reproductible et vérifier la reproductibilité du travail des autres ?					✓
Connaissez-vous l'éventail d'outils que vous pouvez utiliser pour publier des résultats de recherche reproductibles ?					✓
Exigez-vous que les demandes de subvention contiennent des informations sur la reproductibilité des projets futurs et précédents ? Si oui, comment est-ce évalué ?	✓				
Par exemple, cela est-il mentionné dans les politiques de science ouverte, de libre accès ou de données ouvertes ? Qu'en est-il de votre politique d'intégrité de la recherche ?	✓				✓
Pensez-vous que la question de la reproductibilité pourrait ou devrait être un motif de rejet d'un article ?	✓	✓	✓	✓	✓

Questions posées	OFR	OR	FI	EA	CGR
Dans quelle mesure le contrôle de la conformité après publication est-il nécessaire ?	✓		✓	✓	
Dans quelle mesure les fonds de recherche peuvent-ils couvrir les pratiques de reproductibilité (comme c'est le cas pour la gestion des données de la recherche) ?	✓				
Seriez-vous prêt à supporter financièrement les coûts d'infrastructure (par exemple, les serveurs, les personnes chargées de la maintenance) directement (plutôt que par le biais d'une subvention) ?	✓				
Envisagez-vous de développer ou avez-vous développé votre propre infrastructure pour soutenir la science ouverte et la reproductibilité ? Si oui, comment est-elle financée, et les chercheurs sont-ils tenus de l'utiliser ?	✓				
Il arrive que des chercheurs utilisent des plateformes commerciales lorsqu'il en existe. Cela vous dérange-t-il en tant que financeur de la recherche ?	✓				
Quelle est votre vision à long terme de la reproductibilité de la recherche, et que se passera-t-il une fois que les pratiques de publication reproductibles seront devenues monnaie courante ?	✓	✓	✓	✓	✓
Quelle est la mission du service que vous fournissez, et pourquoi l'avez-vous créé ?			✓		
Votre plateforme connaît-elle un succès important ? Disposez-vous des statistiques succinctes que vous pourriez nous communiquer ?			✓		
Y a-t-il certaines disciplines ou certains pays où les infrastructures de reproductibilité sont particulièrement courantes ou usitées ?			✓		
Comment les services offerts aux chercheurs pourraient-ils être mieux intégrés ? Par exemple, comment les solutions et les flux de travail axés sur la reproductibilité pourraient-ils être adoptés dans les pratiques de recherche au quotidien aujourd'hui ?			✓		
Quels sont les points communs entre la gestion des données de la recherche et les infrastructures de reproductibilité ?		✓	✓		
Certains résultats sont-ils plus faciles à publier de manière reproductible que d'autres ? Connaissez-vous des exemples de difficultés spécifiques ?			✓	✓	✓
De quelle manière pourrait-on faciliter la vérification de la reproductibilité des recherches publiées ?			✓	✓	✓

Questions posées	OFR	OR	FI	EA	CGR
Connaissez-vous les articles de recherche exécutables, et quelle est leur valeur ?			✓	✓	
Quelles sources de financement avez-vous envisagées ou mises en place par le passé ? Quels sont vos attentes ou vos besoins pour l'avenir ?			✓		
Quel pourcentage des membres actuels de votre équipe de rédaction devrait adopter rapidement de nouveaux flux de travail pour rendre possible l'évaluation de la reproductibilité ?				✓	
De quelles manières le partage de documents et matériaux sous-jacents à une recherche pourrait-il être récompensé ? Par exemple, le code informatique, les données...				✓	
Quelles sont les infrastructures sur lesquelles vous comptez pour obtenir des données et des codes supplémentaires, indépendamment de la reproductibilité ?				✓	
Attendez-vous des chercheurs qu'ils respectent des consignes de rédaction ? Et les consignes de vos revues comportent-elles des éléments sur la reproductibilité ?				✓	
Vos actions en faveur de la reproductibilité reposent-elles sur un modèle économique durable ?				✓	
Plus généralement, comment soutenez-vous la science ouverte/la recherche reproductible dans votre organisation (infrastructure, enseignement, personnel de soutien, etc.) ?		✓			
Qu'avez-vous mis en place pour votre personnel en ce moment ?		✓			
La reproductibilité fait-elle partie des procédures d'évaluation éthique pour les travaux qui doivent être examinés ?		✓			
Comment assurez-vous l'adoption de pratiques reprocompatibles chez vos chercheurs ?		✓			
Quelle est la proportion de chercheurs qui, selon vous, prennent activement fait et cause pour la reproductibilité de la recherche ? Constatez-vous des fluctuations au sein de votre organisation et, plus largement, au sein de votre discipline ?					✓
Selon vous, où se situe actuellement votre domaine de recherche par rapport aux autres ?					✓
Les politiques actuelles des organismes de financement et des organismes de recherche sont-elles efficaces en matière de reproductibilité ?					✓

Questions posées	OFR	OR	FI	EA	CGR
Par exemple, cela est-il mentionné dans les politiques de science ouverte, d'accès ouvert ou de données ouvertes ? Qu'en est-il de votre politique d'intégrité de la recherche ?	✓				✓
Quelles sont les tentatives de changement de la culture actuelle dont vous avez connaissance ? Lesquelles ont fonctionné/échoué ?					✓
Pensez-vous que ces développements sont progressifs ou révolutionnaires ?					✓
Avez-vous le sentiment que ces parties prenantes travaillent ensemble et de concert ou en silos ?					✓
Selon vous, quels avantages pourraient présenter le fait de publier des résultats de recherche de manière reproductible ?					✓
La reproductibilité de la recherche est-elle un sujet de discussion fréquent entre vous et vos pairs ou collaborateurs ? Et dans votre discipline en général ?					✓
Quels conseils, par exemple de la part des revues ou des bailleurs de fonds, pourraient aider vos collègues et vous-même à vous conformer aux attentes en matière de reproductibilité ?					✓
Par exemple, quel est votre niveau de connaissance à propos des compendiums de la recherche et des articles de recherche exécutables ?					✓
Lorsque vous agissez en tant qu'évaluateur, quelles informations pourraient vous aider à évaluer la reproductibilité à l'intérieur d'un manuscrit ?					✓