

## **Améliorer les chances de succès de son projet ANR grâce à la Science Ouverte**

Guide produit par le Groupe Données du GTSO-Couperin

Avril 2020

Version 1

DOI : [10.5281/zenodo.3741666](https://doi.org/10.5281/zenodo.3741666)

## À propos des auteurs

**Romain Féret** (n°ORCID : [0000-0002-1527-1482](https://orcid.org/0000-0002-1527-1482)) est responsable des services en matière de gestion des données au SCD de l'Université de Lille. Il est aussi chargé de mission *Open Access*. Depuis 2017, il accompagne régulièrement le montage de projets ANR et H2020. Il suit aussi des projets en cours. Il pilote le groupe Données de recherche du GTSO-Couperin et il représente l'Université de Lille dans [OpenAIRE](https://openaire.eu).

**Laetitia Bracco** (n°ORCID : [0000-0002-2939-9110](https://orcid.org/0000-0002-2939-9110)) est *data librarian* à l'Université de Lorraine. Elle accompagne doctorants et enseignants-chercheurs dans la gestion des Données de la recherche et réalise de la *data visualisation* pour le suivi de la Science Ouverte dans son établissement.

**Stéphanie Cheviron** est *data librarian* en charge de la gestion des données scientifiques au sein du service des bibliothèques de l'Université de Strasbourg depuis 2015. Elle est cheffe de projet fonctionnel chargée du développement de bases de données scientifiques avec les chercheurs. Elle suit des projets ANR et H2020 depuis 2019.

**Élise Lehoux** (n°ORCID : [0000-0003-2929-6768](https://orcid.org/0000-0003-2929-6768)) est cheffe de projet Information Scientifique et Technique et Données de la recherche à Université de Paris. Elle accompagne doctorants et enseignants-chercheurs à la gestion des Données de la recherche et participe aux projets dédiés à la Science Ouverte portés par le service des bibliothèques de l'université.

**Cécile Arènes** (n°ORCID : [0000-0002-1839-3530](https://orcid.org/0000-0002-1839-3530)) est chargée de mission Données de la recherche et Humanités numériques à la bibliothèque de Sorbonne université. Elle accompagne les chercheurs dans leur rédaction de plans de gestion des données et assure des formations à la science ouverte à destination des chercheurs et doctorants.

**Ling Li** (n°ORCID : [0000-0003-0882-3266](https://orcid.org/0000-0003-0882-3266)) est chargée du traitement des données scientifiques à Université Lumière Lyon 2, où elle travaille pour la Direction de la Recherche et le SCD. Elle aide les chercheurs à améliorer leurs pratiques de gestion des données.

## Relecteurs

Nous tenons à remercier **Kenneth Maussang** (Université de Montpellier, [0000-0002-8086-8461](https://orcid.org/0000-0002-8086-8461)) et **Corentin Spriet** (Université de Lille, [0000-0002-5805-3426](https://orcid.org/0000-0002-5805-3426)) d'avoir accepté de relire ce document et d'y avoir apporté d'utiles modifications. Les auteurs de ce guide en assument néanmoins l'entière responsabilité.

## Traducteurs

Ce guide a été traduit en anglais grâce à la participation de **Laetitia Bracco** (Université de Lorraine), **Stéphanie Cheviron** (Université de Strasbourg), **André Dazy** (Couperin), **Marlène Delhaye** (Aix-Marseille Université), **Romain Féret** (Université de Lille) et **Ling Li** (Université Lumière Lyon 2).

La version anglophone de ce guide est disponible en ligne : <https://dx.doi.org/10.5281/zenodo.3749577>

## À propos de ce document

Ce guide a été rédigé par les membres du groupe Données du [GTSO-Couperin](https://www.couperin.org/services-et-prospective/open-access/gtao)<sup>1</sup>.

Le **consortium Couperin** est une association à but non lucratif. Il regroupe 242 membres, principalement des universités, organismes de recherche et grandes écoles. En plus de ses missions d'évaluation et d'organisation de l'achat des ressources documentaires numériques au bénéfice de ses membres, Couperin œuvre à l'amélioration de la communication scientifique.

3

Le **Groupe de Travail Science Ouverte** de Couperin travaille à la production d'outils et de ressources afin de promouvoir la Science Ouverte dans les établissements membres du consortium.

## Citer ce document

Féret, R., Bracco, L., Cheviron, S., Lehoux, E., Arènes, C., & Li, L. (2020, Avril). *Améliorer les chances de succès de son projet ANR grâce à la Science Ouverte*. Zenodo. <https://dx.doi.org/10.5281/zenodo.3741666>

---

<sup>1</sup> <https://www.couperin.org/services-et-prospective/open-access/gtao>

## Glossaire

**Accord de consortium** : contrat conclu entre les partenaires d'un projet pour préciser les modalités de leur collaboration.

**ANR** : Agence Nationale de la Recherche.

**AAP** : Appel à projet.

**AAPG** : Appel à projet générique. Cet instrument de financement représente environ les  $\frac{3}{4}$  des projets financés par l'ANR.

**Archive ouverte** : plateforme où les documents déposés sont en accès ouvert. Elle peut être nationale (comme [HAL](#)<sup>2</sup> - Hyper Article en Ligne), institutionnelle ou disciplinaire (comme [arXiv](#)<sup>3</sup>). Les chercheurs peuvent y déposer une notice décrivant leurs publications (titre, auteurs, résumé...) et éventuellement y associer un fichier selon les droits dont ils disposent.

**CES** : Comité d'Évaluation Scientifique. Il existe 48 CES, un par axe de recherche (35 axes de recherche disciplinaires et 13 axes transversaux). Les comités déterminent la liste des projets financés par l'ANR, en s'appuyant notamment sur les rapports remis par les experts scientifiques qui ont évalué les projets. Ces experts sont extérieurs aux CES.

**Coordinateur du projet** : personne responsable de la réalisation de la totalité d'un projet (à distinguer du Responsable scientifique).

**CoSO** : [Comité pour la Science Ouverte](#)<sup>4</sup>. Il impulse et soutient une politique nationale Science Ouverte.

**CPP** : Comité de protection des personnes. Il doit être sollicité pour avis en amont de toute recherche impliquant la personne humaine. Son avis est contraignant.

**Data paper** : article décrivant un ou plusieurs jeux de données, notamment leur potentiel de réutilisation. Il peut être publié dans des revues spécifiques (*data journal*) ou dans des revues scientifiques traditionnelles.

**Dépenses éligibles/inéligibles** : distingue les dépenses que l'ANR rembourse ou ne rembourse pas. Les dépenses éligibles doivent être réalisées dans les règles prévues par le Règlement financier de l'ANR. Les partenaires issus d'établissements de recherche sont généralement financés à coût marginal par l'ANR. Cela exclut notamment la rémunération des personnels permanents.

**DMP** : *Data Management Plan*. Voir Plan de gestion des données (PGD).

**DOAJ** : *Directory of Open Access Journals*. Répertoire de revues en libre accès.

**DOI** : *Digital Object Identifier*. Identifiant unique attribué à des jeux de données ou à des publications.

**Données de la recherche** : ensemble des informations scientifiques produites ou collectées dans le cadre d'un projet de recherche, les données de la recherche peuvent être des photos, des mesures, des sons, etc. Elles sont nécessaires comme éléments probants afin de valider les résultats de la recherche et doivent être accompagnées d'informations qui les documentent, telles que des protocoles

<sup>2</sup> <https://hal.archives-ouvertes.fr/>

<sup>3</sup> <https://arxiv.org/>

<sup>4</sup> <https://www.ouvrirlascience.fr/presentation-du-comite/>



expérimentaux, des descriptifs méthodologiques ou des métadonnées. Ces données peuvent être diffusées dans des entrepôts généralistes ou spécialisés.

**DORA** : [San Francisco Declaration on Research Assessment](https://sfdora.org/)<sup>5</sup>.

**DPO** : *Data Protection Officer*, ou DPD (Délégué à la Protection des Données).

**Entrepôts** : plateformes sur lesquelles sont déposés, décrits et conservés des jeux de données de la recherche. Les entrepôts sont généralistes ou disciplinaires.

**FAIR** : Principes d'ouverture des données, qui visent à les rendre Faciles à trouver (Findable), Accessibles (*Accessible*), Interopérables (*Interoperable*), Réutilisables (*Reusable*).

**FNS** : Fonds National Suisse. Principale agence de financement de la recherche scientifique en Suisse.

**Frais de publication en libre accès** : APC (Articles Processing Charges), BPC (Books Processing Charges), BCPC (Book Chapters Processing Charges).

**H2020** : Horizon 2020, programme-cadre européen pour la recherche et l'innovation (2014-2020).

**JCJC** : Jeune Chercheuse Jeune Chercheur. Instrument de financement individuel, dédié aux chercheurs et chercheuses en début de carrière (chercheur ou chercheuse ayant soutenu sa thèse depuis moins de 10 ans).

**Métadonnées** : informations nécessaires à la description de données, en général structurées selon une norme.

**Open Access / Accès ouvert** : mise à disposition immédiate, gratuite et permanente sur Internet des publications scientifiques issues de la recherche et de l'enseignement. On distingue plusieurs modèles ou voies de l'*Open Access* : la voie verte (*green Open Access*) et la voie dorée (*gold Open Access*).

**OpenAIRE** : consortium qui a pour objectif principal de soutenir le travail de recherche des scientifiques européens en créant et en exploitant une infrastructure d'accès ouvert. OpenAIRE a notamment un rôle d'agrégateur de productions de la recherche (publications, jeux de données, logiciels) et de mise en relation avec les projets financés par les financeurs européens de la recherche (H2020, ANR...).

**Open data** : données ouvertes, dont l'accès est libre et sans restriction.

**ORCID** : *Open Researcher and Contributor ID*. Identifiant unique pour les chercheurs.

**PGD** : Plan de Gestion des Données ou DMP (*Data Management Plan*). Document qui synthétise la description des données de recherche d'un projet et la manière dont elles seront gérées tout au long du projet, afin, notamment, de préparer leur partage, leur réutilisation et leur pérennisation.

**PRC** : Projet de Recherche Collaborative. Instrument de financement pour les projets multi-partenariaux n'impliquant pas d'entreprises ou de partenaires étrangers.

**PRCE** : Projet de Recherche Collaborative – Entreprises. Instrument de financement pour les projets multi-partenariaux impliquant une ou plusieurs entreprises.

**PRCI** : Projet de Recherche Collaborative – International. Instrument de financement pour les projets multi-partenariaux impliquant un ou plusieurs partenaires étrangers.

<sup>5</sup> <https://sfdora.org/>

**Responsable scientifique** : personne responsable de la réalisation des actions de l'un des partenaires du projet (à différencier du coordinateur scientifique).

**Revue hybride** : revue diffusée par abonnement mais dont certains articles peuvent être librement accessibles au lecteur (*Open Access*), moyennant le paiement de frais de publication en libre accès. La majorité des revues d'éditeurs comme Elsevier ou Springer sont des revues hybrides.

**Revue Open Access** : revue dont les articles sont immédiatement accessibles au public.

**RSSI** : Responsable de la sécurité des systèmes d'information.

**QOAM** : [Quality Open Access Market](http://qoam.eu/)<sup>6</sup>.

**SATT** : Société d'accélération du transfert de technologie, en charge de la valorisation économique de la recherche.

**TGIR** : Très Grande Infrastructure de Recherche

**Version acceptée pour publication** (*post-print* ou *author accepted manuscript*) : version comportant les révisions issues du processus d'évaluation par un comité de lecture (*peer-reviewing*). Fichier sans mise en page éditeur ou avec une mise en page partielle.

**Version éditeur** (*version of record, final version*) : article avec la mise en page finale. Version diffusée par l'éditeur.

**Version soumise pour publication** (*pre-print* ou *submitted manuscript*) : version envoyée par les auteurs à une revue, avant le processus de révision par les pairs.

**Voie dorée (gold Open Access)** : la voie dorée qualifie la publication d'articles dans des revues où les articles sont accessibles au lecteur en libre accès, c'est-à-dire sans barrières. Il existe plusieurs modèles économiques à l'intérieur de la « voie dorée ».

**Voie verte (green Open Access)** : la voie verte qualifie le dépôt et la diffusion des publications dans une archive ouverte, par un auteur ou par une personne tierce.

---

<sup>6</sup> <http://qoam.eu/>

## Table des matières

|  |    |
|--|----|
| 1. La Science Ouverte : pourquoi ? comment ?.....                                      | 8  |
| 1.1. Politique nationale en matière de Science Ouverte .....                           | 8  |
| 1.2. La politique Science Ouverte de l'ANR.....  | 10 |
| 1.2.1. Obligations des coordinateurs de projets.....                                   | 10 |
| 1.2.2. Recommandations de l'ANR .....  | 10 |
| 1.3. Du bon usage de ce guide .....  | 11 |
| 2. La proposition détaillée, étape par étape .....                                     | 12 |
| 2.1. Récapitulatif du projet.....  | 12 |
| 2.2. Contexte, positionnement et objectifs de la proposition .....                     | 14 |
| 2.2.1. Objectifs et hypothèses de recherche.....                                       | 14 |
| 2.2.2. Positionnement par rapport à l'état de l'art .....                              | 15 |
| 2.2.3. Méthodologie et gestion des risques.....  | 16 |
| 2.3. Organisation et réalisation du projet.....  | 18 |
| 2.3.1. Coordinateur scientifique et son consortium / son équipe .....                  | 18 |
| 2.3.2. Moyens mis en œuvre et demandés pour atteindre les objectifs .....              | 22 |
| 2.3.3. Les obligations pour les publications issues d'un projet financé par l'ANR..... | 26 |
| 2.4. Impact et retombées du projet .....   | 28 |
| 2.4.1. Pour tous les instruments de financement.....                                   | 28 |
| 2.4.2. Approches spécifiques selon les instruments de financement .....                | 30 |
| 2.5. Bibliographie et CV .....   | 31 |
| 3. Bibliographie .....   | 33 |
| 3.1. Politique publique en matière de Science Ouverte et réglementation.....           | 33 |
| 3.2. Guides et ressources pour mettre la Science Ouverte en pratique .....             | 33 |



# 1. La Science Ouverte : pourquoi ? comment ?

## 1.1. Politique nationale en matière de Science Ouverte

La notion de Science Ouverte s'est développée dans les dix dernières années afin de caractériser un double mouvement d'**ouverture des résultats de la science, mais également une nouvelle façon de faire de la recherche, passant par l'ouverture des processus, des codes, des méthodes ou des protocoles**<sup>7</sup>. Elle est issue d'un mouvement de contestation de la situation de monopoles des grands éditeurs par la promotion de l'*Open Access*. L'*Open Access*, ou accès ouvert, favorise la disparition de barrières d'accès, aussi bien économique que juridique, aux publications et, dans une moindre mesure, aux données de la recherche. Le **développement de la Science Ouverte répond également à des enjeux scientifiques** (efficacité, intégrité et reproductibilité des résultats), **sociaux** (contribution au dialogue science/société et plus grande transparence) et **économiques** (stimulation de l'innovation).

8



<sup>7</sup> <https://www.science-ouverte.cnrs.fr/le-mouvement-pour-la-science-ouverte/>



L'article 30 de la [loi pour une République numérique](#) (promulguée le 7 octobre 2016) garantit le droit aux chercheurs de diffuser la **version acceptée pour publication** (*post-print* ou *AAM*) de leur article au plus tard **6 mois** après la 1<sup>ère</sup> date de publication par l'éditeur en Sciences, Techniques et Médecine (**12 mois** pour les chercheurs en Sciences Humaines et Sociales). **Cette loi s'applique quel que soit le contrat signé avec l'éditeur de la revue, que la revue soit française ou étrangère.**

De plus, les données de la recherche qui ont été rendues publiques par les chercheurs ou leur établissement peuvent **être réutilisées librement** si elles ne sont pas protégées par un droit spécifique. Quelques mois plus tôt, la Loi Valter, [relative à la gratuité et aux modalités de la réutilisation des informations du secteur public](#) (promulguée le 28 décembre 2015), a établi un principe de gratuité dans la réutilisation des informations publiques. Ces deux lois consacrent *par défaut* le principe d'ouverture des données publiques (*Open data*), dont font partie les données de la recherche.

Voir : Lionel Maurel, « La réutilisation des données de la recherche après la loi pour une République numérique » [En ligne] <https://hal.archives-ouvertes.fr/hal-01908766/document>.

Le [Plan national pour la Science Ouverte](#), annoncé par le Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation en juillet 2018, est structuré selon trois axes. Il vise à **généraliser l'accès ouvert aux publications, à structurer et ouvrir les données de la recherche** et enfin, à **s'inscrire dans une dynamique durable, européenne et internationale**. Ce plan se décline en une série de mesures.

« La Science Ouverte n'est pas une mode, ce n'est pas une discipline, c'est un nouveau paradigme. Elle comporte donc ses nouvelles pratiques et ses nouvelles compétences. J'attacherai une attention particulière à ce que ces compétences soient considérées comme faisant partie du bagage initial des jeunes chercheurs. La Science Ouverte ne sera pas l'affaire d'un petit groupe de spécialistes, elle devra imprégner l'ensemble du monde de la recherche. » [Discours de Frédérique Vidal \(4 juillet 2018\)](#)<sup>8</sup>

<sup>8</sup> <https://www.enseignementsup-recherche.gouv.fr/cid132531/plan-national-pour-la-science-ouverte-discours-de-frederique-vidal.html>

## 1.2. La politique Science Ouverte de l'ANR

Dès 2013, l'Agence Nationale de la Recherche a initié une politique de Science Ouverte qui s'est intensifiée à la suite de la publication du [Plan national pour la Science Ouverte](#).

Voir : [La Science Ouverte à l'ANR](#)<sup>9</sup>

### 1.2.1. Obligations des coordinateurs de projets

10

En **2019**, l'ANR a décidé de mettre en place des obligations auprès des projets financés pour promouvoir et encourager à la mise en œuvre de la Science Ouverte. En cas de financement, le coordinateur ou la coordinatrice s'engage à :

1. **Déposer les publications scientifiques (texte intégral) issues du projet financé dans une archive ouverte de son choix**, dans les dispositions prévues par la loi pour une République numérique (embargo de 6 mois après la date de publication pour les STM et de 12 mois après la publication pour les SHS) ;
2. **À fournir dans les 6 mois qui suivent le démarrage du projet un Plan de Gestion de Données (PGD)**. Les coordinateurs et les partenaires devront mettre à jour leur PGD pour fournir une seconde version à mi-parcours (pour les projets de plus de 30 mois) et une version finale à la clôture du projet.

**Il n'est pas obligatoire de diffuser les données issues des projets ANR** (en particulier si cela représente un risque pour des personnes physiques ou morales) mais cela est recommandé.

Les obligations Science Ouverte de l'ANR sont identiques à celles émises par le programme de recherche européen H2020, depuis 2014 pour le libre accès aux publications et depuis 2017 pour la gestion des données.

### 1.2.2. Recommandations de l'ANR

L'ANR recommande l'utilisation de son [modèle de PGD](#)<sup>10</sup> qui est disponible dans [DMP OPIDoR](#)<sup>11</sup>, où il peut être rempli en ligne. Cependant, les coordinateurs sont libres d'utiliser une autre trame de PGD s'ils le souhaitent. Certains établissements proposent par exemple des [modèles institutionnels](#) de PGD, comme l'Observatoire de Paris, l'INRAE, le CIRAD ou l'Université de Paris. Aucun champ du PGD n'est obligatoire. Dans ce cas, il suffit d'indiquer : sans objet.

**Conseil** : Le répertoire [SOS-PGD](#)<sup>12</sup> recense les services qui accompagnent la rédaction des plans de gestion des données dans les universités et les établissements de recherche.

<sup>9</sup> <https://anr.fr/fr/lanr-et-la-recherche/engagements-et-valeurs/la-science-ouverte/>

<sup>10</sup> <https://anr.fr/fileadmin/documents/2019/ANR-modele-PGD.pdf>

<sup>11</sup> <https://dmp.opidor.fr/>

<sup>12</sup> <https://openaccess.couperin.org/sos-pgd/>

Par ailleurs, l'ANR recommande de privilégier la publication dans des revues ou ouvrages nativement en accès ouvert. Elle recommande également l'ouverture des données quand cela est possible, selon le principe : aussi ouvert que possible, aussi fermé que nécessaire. Quand les données sont diffusées, il est préférable qu'elles le soient selon les principes **FAIR** afin d'améliorer leur visibilité et de faciliter leur réutilisation. Les données peuvent être diffusées à certaines conditions, après un embargo ou accessibles uniquement sur demande par exemple.

Voir : [La politique Science Ouverte de l'ANR dans les AAP](#)<sup>13</sup>

### 1.3. Du bon usage de ce guide

Ce guide est mis à disposition des chercheuses et des chercheurs qui déposent un projet ANR. Il a pour objectif d'**aider les porteurs de projets à anticiper** des points dont il est parfois trop tard pour se préoccuper une fois que leur projet a été financé. Il s'agit d'**améliorer la faisabilité des projets** soumis et financés, de **faciliter les collaborations** entre partenaires en matière de Science Ouverte et de les aider à **bâtir une stratégie Science Ouverte** solide et ambitieuse.

Ce guide n'est pas :

Un recueil de phrases ou paragraphes tout rédigés, à copier-coller dans les propositions des projets.

Un document à appliquer dans son intégralité sans considération pour les spécificités des projets.

Un document validé par l'ANR qui remplace ou contredit la politique et les règles présentées par l'ANR dans ses documents officiels.

La rédaction de ce guide s'appuie néanmoins sur la documentation mise à disposition par l'ANR, ainsi que sur l'expérience de ses rédacteurs. Il s'inscrit aussi dans la lignée des recommandations émises en juin 2019 par le Collège Données du CoSO auprès de l'ANR<sup>14</sup>. Plusieurs de ces recommandations incitent à une prise en compte des aspects Science Ouverte dès la phase de sélection des projets :

Recommandations n°3 et 5 : mentionner la gestion des données dans la proposition détaillée ; l'évaluer dès la sélection des projets.

Recommandations n°6, 7 et 8 : encourager la valorisation des données via des *data papers* ; mentionner dans les publications et les jeux de données qu'ils sont issus de projets financés par l'ANR.

Recommandations n°12, 13 et 14 : anticiper les ressources nécessaires au respect des obligations Science Ouverte de l'ANR ; les traduire en réalisations concrètes.

<sup>13</sup> <https://anr.fr/fileadmin/documents/2019/ANR-Politique-science-ouverte-texte-AAP.pdf>

<sup>14</sup> Plan de gestion de données – Recommandations à l'ANR. Zoé Ancion, Francis André, Sarah Cadorel, Romain Féret, Odile Hologne, Kenneth Maussang, Marine Moguen-Toursel, Véronique Stoll. Juin 2019. Disponible à : <https://www.ouvrirlascience.fr/plan-de-gestion-de-donnees-recommandations-a-lanr/>. Consulté le 3 avril 2019.



## 2. La proposition détaillée, étape par étape

Certains aspects relatifs à la Science Ouverte peuvent être mentionnés dès la rédaction de la pré-proposition pour les projets dans lesquels la Science Ouverte est structurante. Cependant, la concision de ce premier document permet rarement de développer ces aspects. Les porteurs de projet qui souhaitent les anticiper dès l'étape 1 de l'AAPG pourront se référer aux sections correspondantes de la proposition détaillée (objectifs, état de l'art, partenariat et bibliographie).

### 2.1. Récapitulatif du projet

Il est important de faire figurer les rôles et responsabilités des membres de l'équipe qui vont mettre en œuvre la politique de Science Ouverte dans le tableau récapitulatif des personnes impliquées dans le projet. Selon le dimensionnement de l'équipe, ces rôles peuvent être cumulés par une personne ou bien répartis entre plusieurs. Les réponses aux questions suivantes permettent d'identifier les éléments clés à faire apparaître dans ce tableau.

- Gestion des données : y aura-t-il une personne responsable de la gestion des données au sein du projet (*data manager*) ? Est-ce que chaque partenaire aura un « correspondant données » qui travaillera en lien avec le *data manager* ? Le *data manager* peut être en charge de définir les bonnes pratiques, de coordonner la rédaction du PGD et sa mise en œuvre, et de faire lien avec les différents interlocuteurs (DPO, RSSI, juristes, informaticiens, etc.).
- Qualité et documentation des données : qui sera responsable de la qualité des données et de leur documentation (protocoles expérimentaux ou d'enquête, procédés de fabrication...) ? Ce rôle peut recouvrir la vérification de la qualité des saisies, de leur homogénéité, la production d'un *codebook* ou guide de saisie...
- Diffusion des productions scientifiques : qui gèrera la mise en ligne des publications scientifiques et des données sous-jacentes ? Qui va s'occuper de la communication auprès d'un public plus large ?

## Exemple de tableau récapitulatif des personnes impliquées dans le projet

| Partenaire        | Nom                 | Prénom  | Position actuelle | Rôle & responsabilités dans le projet (4 lignes max)                         | Implication sur la durée du projet (personne/mois) |
|-------------------|---------------------|---------|-------------------|--|--|
| Marvel University | DANVERS             | Carol   | Capitaine         | Coordinatrice scientifique<br>Tâches X, Y, Z                                 | 18   |
| Marvel University | OCTAVIUS<br>OCTOPUS | Otto    | Docteur           | Data Manager, chargé de la qualité des données et du PGD<br>Tâche Z          | 12   |
| Marvel University | GREY                | Jean    | Docteur           | Rédactrice du codebook<br>Tâche Z  | 3  |
| Marvel University | STRANGE             | Stephen | Docteur           | Bibliothécaire, chargé de la diffusion en Open Access des outputs<br>Tâche Y | 2  |

13

Si l'aspect Science Ouverte n'a pas été anticipé lors de la pré-proposition, il s'agit d'en préciser maintenant les conséquences principales sur le budget. Il y a des coûts à prendre en compte pendant et après la fin du projet. En effet, la dimension Science Ouverte d'un projet nécessite du temps de travail (à compter en personne/mois) et peut occasionner d'autres types de coûts.

### ***Pendant le projet***

Il faut tenir compte des coûts engendrés par la gestion des données, tant en moyens matériels (serveurs, sécurité, etc.) qu'en ressources humaines (temps de travail d'un *data manager*, documentation, nettoyage et vérification des données, etc.). La diffusion des résultats de la recherche et leur archivage pérenne peuvent également engendrer des frais, en particulier la publication en libre accès d'articles, de livres ou de jeux de données.

### ***Après la fin du projet***

Le stockage et l'archivage des données font aussi partie des dépenses liées à la Science Ouverte. Ces frais sont éligibles au remboursement pour une prestation allant jusqu'à 5 ans après la fin du projet, à condition d'avoir été payés avant la fin du projet.

L'impact de la Science Ouverte sur le budget d'un projet de recherche est précisé dans la partie 2.3.2 de ce guide.

**Conseil** : il convient de se rapprocher des services compétents au sein de son établissement (ingénieurs projet, juristes, bibliothécaires) pour bien analyser son projet afin de repérer les points de vigilance et d'éviter d'éventuels coûts cachés ou imprévus.

**Exemple** : une équipe de recherche photographie des objets archéologiques conservés par un musée. La coordinatrice a obtenu du musée l'autorisation de prendre ces clichés avant le dépôt de sa pré-proposition. Mais elle n'avait pas anticipé que des frais pourraient lui être demandés lors de la réutilisation des images dans des publications ou dans les thèses des doctorants impliqués dans son projet. Elle n'en a pris conscience que lors de la rédaction de la proposition détaillée. Il est alors nécessaire de préciser ce point parmi les modifications apportées au projet afin de justifier l'écart entre le budget de la pré-proposition et celui de la proposition détaillée.

14

## 2.2. Contexte, positionnement et objectifs de la proposition

La partie 2.2 permet de détailler les objectifs principaux du projet, en mettant notamment en avant la potentielle réutilisabilité des résultats de recherche que l'on souhaite produire. En cela, il conviendra de souligner, en fonction des spécificités du projet, si l'ouverture de données représente un défi spécifique (par exemple, pour des données d'enquête nécessitant des techniques d'anonymisation complexes) ou un apport particulièrement notable aux données existantes. Il est aussi important de montrer la plus-value des données collectées par rapport aux données existantes et pourquoi les données existantes peuvent ou non être réutilisées.

### 2.2.1. Objectifs et hypothèses de recherche

Les produits de la recherche à décrire dans ce paragraphe peuvent prendre différentes formes :

- Jeux de données (exemple d'un jeu de données sonores sur le chant des baleines : <https://zenodo.org/record/1403351#.XoSEPnLgrIU>), éventuellement associés à un *data paper*.
- Bases de données structurées (exemple des ressources mises à disposition par l'équipement [Ortolang](https://www.ortolang.fr/)<sup>15</sup> sur la langue et son traitement).
- Code source d'un logiciel ou d'un algorithme (exemple du code du logiciel [Samvera Hyrax](https://github.com/samvera/hyrax)<sup>16</sup>).

Si les publications sont de toute manière concernées par l'obligation de dépôt dans une archive ouverte, les données de la recherche sont également un enjeu majeur de la Science Ouverte. Il est donc judicieux de souligner dans cette partie, le cas échéant, que l'un des objectifs structurants de ce projet de recherche est d'ouvrir les données produites afin d'encourager leur réutilisation.

<sup>15</sup> <https://www.ortolang.fr/>

<sup>16</sup> <https://github.com/samvera/hyrax>



**Exemple** : un laboratoire produit des statistiques de référence en hygrométrie. La création d'une base de données en accès libre, afin de valoriser ce travail et d'en permettre la réutilisation tout en pérennisant les données, peut constituer un axe Science Ouverte structurant du projet. Dans ce cas, il est nécessaire de prévoir les moyens humains, financiers et techniques pour mener à bien cet objectif.

**Conseil** : lister des exemples de réutilisation future des données produites.

### 2.2.2. Positionnement par rapport à l'état de l'art

L'un des principaux objectifs de la Science Ouverte est d'encourager les chercheurs à réutiliser les données existantes. Par conséquent, il convient de positionner son projet par rapport à l'existence ou non de données disponibles sur le sujet d'étude. Le cas échéant, il est intéressant d'une part de présenter ces jeux de données et d'expliquer leur intérêt (ce sont des références) ou les problèmes qu'elles soulèvent par rapport à l'état de l'art (ces données sont datées, il faut en produire de nouvelles) pour justifier éventuellement l'intérêt d'une nouvelle phase de collecte de données dans le cadre du projet. S'il est prévu s'appuyer sur des données existantes, il faut expliquer comment elles seront réutilisées : va-t-on enrichir ces données ? Vont-elles être comparées aux nouvelles données produites ?

**Exemple** : deux institutions ont chacune diffusé une base de données prosopographique en architecture en accès libre sur le web. Une équipe de recherche travaillant sur la pédagogie dans l'enseignement de l'architecture va réutiliser ces données via une extraction pour les enrichir avec les données que ses membres vont collecter dans un fonds d'archives. La base de données résultante sera moissonnable, c'est à dire que ses données pourront être automatiquement récupérées par d'autres bases de données, favorisant ainsi leur dissémination et leur découvrabilité.

Il est également possible de disposer de données préliminaires, produites lors de projets de recherche antérieurs, et de vouloir s'appuyer dessus pour poursuivre son hypothèse de travail. Ces informations doivent être mentionnées dans la proposition. On peut alors indiquer si ces données ont déjà été diffusées et ce qu'elles apportent à l'état de l'art.

**Exemple** : une chercheuse travaille depuis plusieurs années sur la constitution d'un corpus inédit de statuettes anciennes. Une partie des statuettes est photographiée. Sa proposition s'appuie sur ces données préliminaires pour demander un financement qui lui permettrait de finir les prises de vue, mais aussi de constituer une base de données en accès libre qui ferait référence dans son domaine.

### 2.2.3. Méthodologie et gestion des risques

Cette partie permet de détailler les grandes étapes du projet ainsi que les risques scientifiques liés. De nombreux enjeux de Science Ouverte peuvent intervenir ici.

#### Le découpage du projet en lots de travail (*work packages*) ou en tâches (*tasks*) :

1. Prévoir au sein du WP sur la gestion de projet une ou des tâches qui concernent spécifiquement la Science Ouverte, avec des livrables associés.

16

#### Exemple : *work package* intégrant des tâches et livrables liés à la Science Ouverte

|   |   |                    |                   |
|---|---|--------------------|-------------------|
| WP 1  | Coordination, Management and Dissemination            | Start: M1          | End: M48          |
| WP leader: Coordinator  |   |                    |                   |
| Contributions   | Coord. (X P/M)  | Partner 1 (X P/M)  | Partner 2 (X P/M) |
| Objectives: <ul style="list-style-type: none"><li>• Administrative and financial management</li><li>• Scientific management</li><li>• <b>Data management and dissemination activities</b></li></ul> |   |                    |                   |
| Tasks   |   |                    |                   |
| T1.1  | Administrative management, coordination (coord.)      |                    |                   |
| T1.2  | Financial management (coord.)                         |                    |                   |
| T1.3  | Promotion of internal communication (partner 1)       |                    |                   |
| T1.4  | <b>Writing and implementation of the DMP</b> (coord.) |                    |                   |
| T1.5  | <b>Project dissemination</b> (coord.)                 |                    |                   |
| Deliverables  | Deliverables  |                    |                   |
| D1.1  | Kick-off Meeting                                      | M1                 |                   |
| D1.2  | Consortium agreement                                  | M6                 |                   |
| D1.3  | <b>Data Management Plan</b>                           | M6, M30, M48       |                   |
| D1.4  | Periodic and final reports                            | M12, M24, M36, M48 |                   |
| D1.5  | <b>Dissemination report</b>                           | M48                |                   |
| D1.6  | Workshop  | M42                |                   |

**Conseil** : identifier un correspondant données par partenaire pour faciliter le travail de la personne en charge de coordonner la rédaction du PGD et de sa mise en œuvre. Préciser quel partenaire se chargera de cette coordination.

2. Mentionner, le cas échéant, si des actions spécifiques de diffusion des données et des résultats de recherche en *Open Access* sont prévues (exemples : publication d'un *data paper*, mise en ligne d'une base de données, diffusion de comptes rendus de séminaires ou de posters...) et les signaler sous forme de tâches et de livrables dans les *work packages* correspondants. Si un rapport de dissémination est produit, la stratégie de diffusion des données et de valorisation des productions scientifiques devra y figurer en incluant les publics visés.
3. Préciser si le projet nécessite l'accès à une infrastructure de recherche ou à une TGIR : des infrastructures seront-elles utilisées pour acquérir des données ou pour accéder à des données existantes ? En effet, ces infrastructures peuvent proposer des outils et des services afin d'améliorer la qualité des données et faciliter leur diffusion.

Exemples de TGIR : [HumaNum](https://www.huma-num.fr/)<sup>17</sup> dédiée aux humanités numériques, [Progedo](http://www.progedo.fr/)<sup>18</sup> pour PROduction et GEstion des DONnées en sciences humaines et sociales.

### La gestion des risques :

1. Expliciter quels types de données sensibles seront collectées (données à caractère personnel, données sensibles, données patients, données biologiques...) et quels traitements leur seront appliqués. Protection des données et diffusion ne s'opposent pas de manière binaire. Il est par exemple possible d'anonymiser des données ou de limiter leur réutilisation à certains usages.

**Exemple :** [beQuali](https://bequali.fr/fr/)<sup>19</sup>, une banque d'enquêtes qualitatives en SHS, recense des enquêtes qui ont été anonymisées de manière réversible ou irréversible. Les enquêtes, généralement réalisées sous forme d'entretiens ou d'observations, sont ensuite documentées et mises à disposition uniquement dans le cadre de réutilisations à des fins académiques, via un contrôle d'accès.

**Conseil :** Horizon 2020 propose un guide pour évaluer les questions éthiques liées à son projet, qu'il peut être utile de consulter : [How to complete the self-ethics assessment](https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-self-assess_en.pdf)<sup>20</sup>.

<sup>17</sup> <https://www.huma-num.fr/>

<sup>18</sup> <http://www.progedo.fr/>

<sup>19</sup> <https://bequali.fr/fr/>

<sup>20</sup> [https://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/ethics/h2020\\_hi\\_ethics-self-assess\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-self-assess_en.pdf)



2. Préciser si certains risques scientifiques sont liés à la gestion des données, comme l'incertitude sur l'obtention d'autorisations pour collecter des données sensibles (ou sur les délais), les difficultés liées au stockage ou à l'échange de données entre les partenaires du projet...

**Conseil** : identifier les interlocuteurs qui peuvent vous aider à gérer les données sensibles. Si les enjeux sont importants, il est vivement souhaitable de les faire intervenir dès la rédaction de votre proposition détaillée.

18

Ces services supports peuvent être : le Délégué à la protection des données ([DPO](#)<sup>21</sup>), le Comité de protection des personnes ([CPP](#)<sup>22</sup>), le comité éthique ou les services juridiques de l'établissement. Il peut aussi être intéressant de solliciter la bibliothèque universitaire pour chercher un équilibre entre protection et diffusion des données.

Pour savoir si les données de recherche produites dans le cadre du projet peuvent être diffusées ou non, le [guide d'analyse du cadre juridique](#)<sup>23</sup> en France sur l'ouverture des données de recherche est une aide précieuse.

## 2.3. Organisation et réalisation du projet

La partie 2.3 est l'occasion de présenter l'expérience de l'équipe ou du consortium qui portera le projet, en mettant en avant à la fois la complémentarité du collectif de travail et l'expertise de chacun de ses membres. Les obligations Science Ouverte émises par l'ANR nécessitent des porteurs de projets qu'ils identifient les moyens et compétences dont ils disposent déjà et ceux dont ils auront besoin.

### 2.3.1. Coordinateur scientifique et son consortium / son équipe

Qu'il s'agisse d'un projet individuel ou d'un projet collaboratif, le coordinateur ou la coordinatrice scientifique aura intérêt à valoriser son expérience et ses réalisations en matière de Science Ouverte.

#### *Le coordinateur scientifique*

**Expertise scientifique** : dans sa présentation individuelle, le coordinateur du projet peut signaler ses réalisations en matière de Science Ouverte, telles qu'un nombre de publications important déposées en archive ouverte, des jeux de données déposés dans des entrepôts généralistes ([Zenodo](#)<sup>24</sup>, [Dryad](#)<sup>25</sup>)

<sup>21</sup> <https://www.cnil.fr/fr/le-delegue-la-protection-des-donnees-dpo>

<sup>22</sup> <https://www.ars.sante.fr/comite-de-protection-des-personnes-1>

<sup>23</sup> <https://prodinra.inra.fr/ft?id=C7D38E14-877E-4883-AB91-0536DD4D63B9>

<sup>24</sup> <https://zenodo.org/>

<sup>25</sup> <https://datadryad.org/>

ou disciplinaires ([Protein Data Bank](https://www.rcsb.org/)<sup>26</sup>, [Cambridge Structural Database](https://www.ccdc.cam.ac.uk/solutions/csd-system/components/csd/)<sup>27</sup>...) ou la contribution à la création de bases de données importantes dans son domaine. Il est possible par exemple :

- D'indiquer le nombre de vos publications en *Open Access* grâce au site [Dissemin](https://dissem.in/)<sup>28</sup>.
- D'indiquer le nombre de jeux de données déjà diffusés grâce au moteur de recherche de [DataCite](https://search.datacite.org/)<sup>29</sup>.
- De préciser le nombre de jeux de données diffusés dans des bases de données internationales.

**Expérience en gestion de projet :** plus généralement, le coordinateur peut aussi indiquer s'il a déjà participé à des projets de recherche (ANR, H2020...) qui étaient couverts par des obligations Science Ouverte ou s'il a déjà contribué à la rédaction d'un PGD dans un autre contexte qu'un projet de recherche (projet non financé, thèse...).

### Focus : identifier les ressources disponibles au sein des institutions impliquées dans le projet

Au même titre que les coordinateurs s'appuient généralement sur une cellule d'appui pour les questions relatives aux aspects administratifs et financiers, il est essentiel qu'ils identifient les services pertinents pour les aider sur les aspects Science Ouverte. Ces services peuvent être :

- Service commun de documentation : *Open Access*, métadonnées, aide à la rédaction d'un PGD, appui au dépôt de jeux de données.
- SATT ou service de valorisation économique : identifier les données pouvant faire l'objet d'une valorisation économique et celles qui peuvent être diffusées librement.
- Délégué à la protection des données : identifier les données à caractère personnel qui devront faire l'objet d'un traitement particulier ; se former à des outils pour gérer ces données de manière responsable (cryptage des supports de stockage, anonymisation des données d'enquête ou d'entretien).
- La Direction des Systèmes d'Information : sécurisation des données, stockage et partage des données, hébergement de bases de données, développements à façon, aide à la budgétisation des prestations informatiques.

<sup>26</sup> <https://www.rcsb.org/>

<sup>27</sup> <https://www.ccdc.cam.ac.uk/solutions/csd-system/components/csd/>

<sup>28</sup> <https://dissem.in/>

<sup>29</sup> <https://search.datacite.org/>

**Conseil** : si vous ne savez pas qui peut vous aider à rédiger votre plan de gestion des données, vous pouvez consulter [SOS-PGD](https://openaccess.couperin.org/sos-pgd/)<sup>30</sup>. Ce répertoire recense les services qui accompagnent la rédaction des plans de gestion des données dans les universités et les établissements de recherche.

D'autres types de ressources peuvent être mentionnées, comme :

- Des formations ou ateliers dédiés à la Science Ouverte que les membres du consortium ont suivis ou qu'ils envisagent de suivre.
- Les infrastructures et services associés disponibles : espace de stockage de données sécurisé, équipement d'acquisition et de traitement de données...

20

### **Les projets individuels**

En complément des éléments mentionnés plus haut, le porteur d'un projet individuel a intérêt à expliciter les points suivants :

- Ses compétences Science Ouverte actuelles et celles qu'il compte développer au cours du projet.
- La répartition des rôles et responsabilités au sein de son équipe pour les activités qui relèvent de la gestion des données.

### **Les projets collaboratifs : coordination générale**

En plus des points ci-dessus, les coordinateurs de projets collaboratifs devront veiller à bien faire apparaître les interactions entre chaque partenaire, en particulier en matière de gestion des données. À la lecture de la proposition détaillée, il est essentiel de comprendre facilement si les partenaires travaillent sur les mêmes types de données, ce qui implique une mise en commun des données et une harmonisation des pratiques (convention de nommage, organisation, procédures de sauvegarde...), ou s'ils travaillent sur des données différentes ne nécessitant que le partage des résultats obtenus.

Cela affecte les solutions retenues pour stocker et partager les données, ainsi que le degré d'harmonisation nécessaire dans les pratiques de gestion des données entre les partenaires. Dans la proposition détaillée, il convient de faire apparaître :

- L'équilibre entre centralisation et décentralisation de la gestion des données en précisant la répartition des rôles et responsabilités entre les partenaires. Expliciter le choix de nommer explicitement un data manager ou non.
- Où seront stockées les données et comment elles seront partagées : expliciter les outils utilisés et justifier leur choix (robustesse de la solution, sécurité et localisation des données).
- Les règles principales de partage de propriété des données. Pour les données réutilisées : ont-elles été produites en dehors du consortium ou sont-elles mises à disposition par un des

<sup>30</sup> <https://openaccess.couperin.org/sos-pgd/>



partenaires ? Pour les données acquises en cours de projet : qui en sera propriétaire ? Ces informations seront ensuite affinées au sein de l'accord de consortium.

### Focus : Faut-il nommer un *data manager* ?

Un *data manager* a la charge de définir des bonnes pratiques au sein du consortium et de vérifier leur mise en œuvre par chacun des partenaires. Il n'est pas nécessaire que ce soit le coordinateur du projet, il peut s'agir d'un chercheur, d'un ingénieur ou d'un technicien. Il est souhaitable que cette personne ait une connaissance fine des pratiques de recherche des acteurs du projet ainsi qu'une expertise scientifique lui permettant de comprendre comment les données sont structurées.

En effet, son rôle implique de s'investir au niveau opérationnel auprès de l'ensemble des chercheurs du projet, et plus particulièrement des responsables scientifiques de chaque partenaire. Il n'est pas toujours nécessaire qu'une personne dédiée soit recrutée sur cette fonction. Une personne impliquée dans le projet peut se voir confier ce rôle. Il sera alors important de bien le positionner au sein de l'équipe afin qu'il puisse endosser son rôle.

21

**Exemple :** la bibliothèque de l'Université d'Utrecht met à disposition de ses chercheurs des *data managers* pour 35 € de l'heure. Ces *data managers* sont intégrés aux projets et apportent leur expertise en matière de gestion et de curation des données<sup>31</sup>.

### Les projets collaboratifs : partenaire par partenaire

En complément d'une stratégie générale, il est important que chaque partenaire précise un certain nombre d'éléments le concernant. Selon les projets, chaque partenaire pourra par exemple :

- Lister les principaux types de données qu'il produira.
- Identifier les expertises présentes au sein de son équipe en matière de gestion des données.
- Répartir les rôles et responsabilités au sein de son équipe.
- Décrire le soutien apporté par son institution : infrastructures ; accompagnement par des services spécifiques (bibliothèque, valorisation, DPO).

**Conseil :** la proposition détaillée doit rester un document concis. Vous ne pourrez pas détailler l'ensemble des points listés ci-dessus. Identifier les caractéristiques propres à votre projet pour cibler les points principaux à mentionner dans votre projet.

<sup>31</sup> Jacques Florès. Supporting researchers with RDM costs at Utrecht university. 2019. [www.dcc.ac.uk/webfm\\_send/2954](http://www.dcc.ac.uk/webfm_send/2954) (diapositive 12).

### 2.3.2. Moyens mis en œuvre et demandés pour atteindre les objectifs

Les moyens nécessaires pour mettre en place une stratégie Science Ouverte dépendent des objectifs que se fixent les membres du projet et de la complexité des données (sensibilité des données, volume...). Les moyens demandés sont dépendants des moyens mis en œuvre, c'est pourquoi nous avons opté pour une présentation de cette section selon la typologie de catégorie de coûts proposée par l'ANR. En ce qui concerne les frais d'*Open Access*, couramment appelés APC, les projets n'ont pas l'obligation d'y avoir recours. Cependant, c'est une pratique courante sur laquelle un focus a été fait. En matière financière, le [Règlement financier de l'ANR](#)<sup>32</sup> fait foi. La fiche sur les [dépenses éligibles](#)<sup>33</sup> est une ressource utile pour connaître les dépenses éligibles et celles qui ne le sont pas.

22

#### **Frais de personnel (catégorie a<sup>34</sup>)**

Ces dépenses représentent en général la majorité des dépenses liées à la Science Ouverte, même si elles sont généralement incluses dans le temps de travail des chercheurs à travers des activités courantes de la recherche : acquisition de données, organisation et documentation des données...

Cependant, la mise en œuvre d'une stratégie Science Ouverte exige un travail additionnel, plus ou moins important selon le degré de formalisation des pratiques de recherche des membres du projet. Ce temps de travail additionnel sera dédié à des activités telles que : anonymiser des données, améliorer la documentation des données (métadonnées supplémentaires, rédaction de fichiers ReadMe), harmoniser de manière systématique les conventions de nommage et d'organisation des fichiers, sauvegarder les fichiers dans différents formats, déposer les jeux de données dans un entrepôt prévu à cet effet, etc.

Ces bonnes pratiques sont une condition pour diffuser des données réutilisables par d'autres chercheurs selon les [principes FAIR](#)<sup>35</sup>. Elles ont aussi l'avantage de faciliter l'échange de données entre les partenaires du projet et de s'assurer de la continuité de l'accès et de la compréhension de ces données par les équipes qui les ont produites, même après le départ d'un collègue ou d'un doctorant. Le travail de curation de données peut être réparti entre les parties prenantes du projet ou être principalement confié à un *data manager*. Dans les deux cas, il conviendra d'estimer de la manière la plus juste possible le temps de travail nécessaire pour gérer les données en fonction des objectifs que s'est fixé le consortium.

<sup>32</sup> <https://anr.fr/fr/rf/>

<sup>33</sup> <https://anr.fr/fileadmin/documents/2017/ANR-RF-Fiche-COUTS.pdf>

<sup>34</sup> La catégorie fait référence aux catégories de coûts définies par l'ANR dans son règlement financier.

<sup>35</sup> <https://www.go-fair.org/fair-principles/>

Des outils existent pour faciliter le calcul du temps de travail à consacrer à la curation des données, par exemple :

- Le [Costing data management tool](#)<sup>36</sup> développé par le UK Data service
- L'Université d'Utrecht détaille les [coûts induits par la gestion des données](#)<sup>37</sup> à toutes les étapes du cycle de vie des données

### Focus : ouvrir des données d'entretien contenant des données à caractère personnel

23

Dans certains cas, l'ouverture des données est peu chronophage car elle n'implique pas de modifier les données, alors que dans d'autres cas, il est nécessaire de les transformer en profondeur. C'est le cas de l'ouverture des données obtenues lors d'entretiens par exemple.

Pour faciliter leur réutilisation et leur anonymisation, il peut être nécessaire de les transcrire puis de les anonymiser manuellement ou en utilisant des logiciels spécifiques. L'Université de Manchester a mis en ligne une dizaine de [fiches pratiques](#)<sup>38</sup> pour informer les chercheurs sur la manière de traiter des données qualitatives, notamment pour la transcription des données d'entretien. Pour calculer ce temps, il est important de tenir compte de nombreux facteurs comme le nombre et la durée des entretiens, le nombre de personnes interrogées en même temps, la qualité des enregistrements, l'expérience du transcripteur... Il faut au minimum 4h pour transcrire correctement 1h d'entretien.

### Coûts des instruments et du matériel (catégorie b)

Dans cette catégorie de coûts, les dépenses liées à la Science Ouverte peuvent par exemple couvrir les besoins suivants :

- Stocker et partager des données : que la solution utilisée soit proposée par le service informatique ou par un prestataire extérieur. Cela peut couvrir l'achat de supports physiques de stockage (serveurs NAS, disques durs externes) ou l'accès à un espace de cloud, ainsi que des services permettant de transférer d'importants volumes de données de manière sécurisée. Certains de ces coûts peuvent figurer parmi les coûts de prestation (catégorie d).
- Numériser des documents si cela nécessite l'achat ou l'utilisation intensive d'un numériseur.
- L'Université de Lausanne a développé un [calculateur de coût](#)<sup>39</sup> pour anticiper les coûts liés au stockage et à l'archivage des données.

<sup>36</sup> <https://www.ukdataservice.ac.uk/media/622368/costingtool.pdf>

<sup>37</sup> <https://www.uu.nl/en/research/research-data-management/guides/costs-of-data-management>

<sup>38</sup> <https://www.socialsciences.manchester.ac.uk/morgan-centre/research/resources/toolkits/toolkit-08/>

<sup>39</sup> <https://costcalc.epfl.ch/>



### **Coûts des bâtiments et terrain (catégorie c)**

Cette catégorie de coûts n'a pas d'incidence sur les aspects Science Ouverte.

### **Coûts du recours aux prestations de service (catégorie d)**

Dans cette catégorie de coûts, les coordinateurs de projets peuvent faire figurer des tâches qui sont externalisées telles que :

- Accès à un cloud
- Anonymisation
- Conseils juridiques
- Numérisation
- Transcription

Ils peuvent aussi mentionner les dépenses relatives à l'acquisition et à la maintenance de logiciels permettant de rendre des données diffusables et réutilisables, telles que des logiciels de transcription, d'anonymisation de données (ex. [Amnesia](#)<sup>40</sup>) ou de conversion de formats de fichiers.

C'est aussi dans cette catégorie de coûts que peuvent apparaître les frais de diffusion en libre accès des articles, ouvrages, jeux de données... (voir le focus sur les frais de publication ci-dessous).

### **Frais généraux (catégorie e)**

Cette catégorie de coûts n'a pas de grande incidence sur les aspects Science Ouverte, sauf si des formations sur ces thèmes sont nécessaires et impliquent le paiement de frais de déplacements pour les formateurs sollicités.

**Conseil** : en matière de Science Ouverte comme pour le reste du projet, il est essentiel que les moyens demandés soient en cohérence avec les objectifs du projet et les moyens dont le consortium dispose déjà.

<sup>40</sup> <https://amnesia.openaire.eu/amnesiaInfo.html>

## Focus : les frais de publication en libre accès (APC, BPC, BCPC, DPC)

Les APC (*Article Processing Charges*) sont des frais à régler pour publier un article en libre accès.

Pour savoir si une revue en *Open Access* demande des APC, il est possible de consulter le DOAJ (*Directory of Open Access Journals*<sup>41</sup>) et d'y saisir le nom de la revue. Exemple pour la revue *Ecology and Society*<sup>42</sup> : les APC sont de 975\$ minimum par article.

25

La base QOAM (*Quality Open Access Market*<sup>43</sup>) permet en outre de comparer qualité de la revue (apport du processus de revue par les pairs, éditorialisation des contenus, délais de publication) et montant des APC demandés, ainsi que de savoir quels montants d'APC ont été réellement payés par les auteurs. Exemple : la revue *Ecology and Society*<sup>44</sup> obtient le score de 4,4 sur 5.

Les frais d'APC ne sont éligibles que pour publier dans des revues complètement en *Open Access*. **Ce ne sont pas des coûts éligibles pour publier dans des revues hybrides.**

Les revues hybrides sont des revues qui offrent la possibilité de publier des articles en libre accès moyennant le paiement d'APC ; mais les autres articles demeurent accessibles uniquement via un abonnement. Le modèle économique de ces revues leur permet de générer un revenu via :

- Le paiement de l'abonnement à la revue par les services communs de documentation.
- Le paiement des APC par les auteurs.

Le paiement de frais d'*Open Access* sont éligibles pour d'autres productions scientifiques, telles que : Les ouvrages (BPC : *Book Processing Charges*) et chapitres d'ouvrages (BCPC : *Book Chapters Processing Charges*). Par exemple, l'éditeur Berghahn Books<sup>45</sup> demande 15 000\$ aux auteurs pour la publication en libre accès d'un ouvrage et 2 000\$ pour la publication d'un chapitre d'ouvrage. Les jeux de données (DPC : *Dataset Processing Charges*). Par exemple, l'entrepôt de données Dryad<sup>46</sup> demande 120\$ par jeu de données jusqu'à 50 Go et 50\$ de plus par dizaine de Go supplémentaires.

Comme pour l'ensemble des dépenses des projets ANR, les frais de publication ne sont éligibles qu'à la condition d'avoir été payés avant la fin du projet, après service fait.

<sup>41</sup> <https://doaj.org/>

<sup>42</sup> <https://tinyurl.com/wabrydw>

<sup>43</sup> <https://www.qoam.eu/>

<sup>44</sup> <https://tinyurl.com/w8kqymh>

<sup>45</sup> <https://www.berghahnbooks.com/authors/>

<sup>46</sup> [https://datadryad.org/stash/publishing\\_charges](https://datadryad.org/stash/publishing_charges)

### 2.3.3. Les obligations pour les publications issues d'un projet financé par l'ANR



#### 1) Il n'est pas obligatoire de publier dans une revue en *Open Access*

L'ANR exige que les publications issues de projets qu'elle finance soient archivées de manière pérenne et diffusées dans une archive ouverte, soit directement dans [HAL](https://hal.archives-ouvertes.fr/)<sup>47</sup> soit par l'intermédiaire d'une archive institutionnelle, comme [Lilloa](http://lilloa.univ-lille.fr/)<sup>48</sup> ou [UnivOAK](https://univoak.eu/)<sup>49</sup>.

Pour rappel, la loi pour une République Numérique promulguée en octobre 2016 autorise les chercheurs à déposer la version finale acceptée pour publication de leurs articles dans une archive ouverte et à les rendre accessibles 6 mois après leur publication par l'éditeur (12 mois pour les chercheurs en sciences humaines et sociales, gestion et droit). **Cette loi s'applique quel que soit le contrat signé avec l'éditeur de la revue, que la revue soit française ou étrangère.**

<sup>47</sup> <https://hal.archives-ouvertes.fr/>

<sup>48</sup> <http://lilloa.univ-lille.fr/>

<sup>49</sup> <https://univoak.eu/>



**Les chercheurs peuvent donc respecter les obligations de l'ANR sans dépenser de frais d'APC.**

NB : n'importe quel chercheur couvert par le droit français, qui est co-auteur d'une publication, peut la déposer dans ces conditions dans une archive ouverte (avec l'accord de ses co-auteurs). Cependant, cela ne couvre pas les publications produites exclusivement par des partenaires étrangers dans le cas des PRCI, des projets transnationaux et des projets bilatéraux (France-Suisse, France Québec...). Dans ce cas, il est nécessaire de vérifier la politique de la revue, par exemple sur le site [Sherpa/Romeo](#)<sup>50</sup>.

27

Pour en savoir plus :

- [WillIO](#)<sup>51</sup>, outil permettant de vérifier vos droits et obligations de diffuser une publication en *Open Access*
- Le [guide d'application de la loi pour une République Numérique](#)<sup>52</sup>, publié par le Comité pour la Science Ouverte
- La [FAQ sur la loi pour une République numérique](#)<sup>53</sup>, produite par Couperin

## **2) Le dépôt dans une archive ouverte est obligatoire, quel que soit le mode de publication initial**

Même si l'article est publié dans une revue en *Open Access*, le dépôt dans une archive ouverte est obligatoire. Dans ce cas, il doit être fait dès la publication de l'article par l'éditeur.

En règle générale, lorsque des APC ont été payés, les auteurs demeurent propriétaires de leur article et/ou l'article est diffusé sous une licence [Creative Commons](#)<sup>54</sup>. La version finale, dite version éditeur, de l'article peut alors être archivée et diffusée en libre accès.

<sup>50</sup> <https://v2.sherpa.ac.uk/romeo/>

<sup>51</sup> <https://decadoc.typeform.com/to/W2ZZMV>

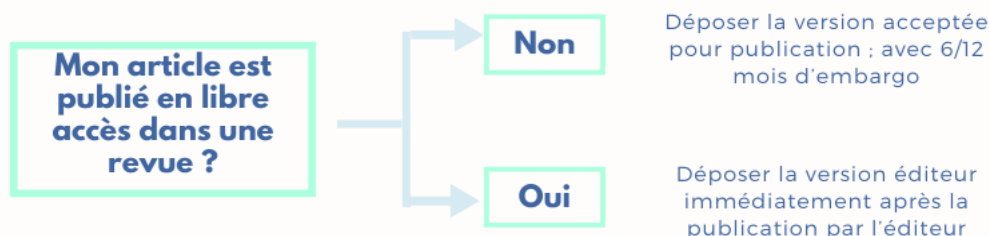
<sup>52</sup> <https://www.ouvrirlascience.fr/guide-application-loi-republique-numerique-article-30-ecrits-scientifiques-version-courte/>

<sup>53</sup> <https://openaccess.couperin.org/category/faq/>

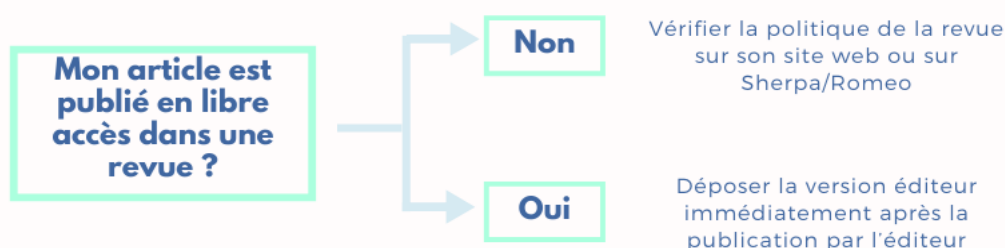
<sup>54</sup> <https://creativecommons.org/choose/>

## COMMENT DIFFUSER MON ARTICLE EN LIBRE ACCÈS DANS UNE ARCHIVE OUVERTE

### PUBLICATIONS DONT AU MOINS UN DES CO-AUTEURS EST COUVERT PAR LE DROIT FRANÇAIS



### PUBLICATIONS POUR LESQUELLES AUCUN DES CO-AUTEURS N'EST COUVERT PAR LE DROIT FRANÇAIS



**Conseil** : afin de trouver un équilibre entre publier dans des revues en Open Access (voie dorée), souvent moyennant APC, et le budget du projet, il est préférable d'adopter soit une stratégie d'archivage uniquement (déposer dans une archive ouverte, voie verte), soit une stratégie mixte entre voie verte et voie dorée.

## 2.4. Impact et retombées du projet

La partie III expose la stratégie de dissémination et de communication du projet. Si certains éléments sont communs à tous les projets, les coordinateurs de projets peuvent appuyer plus ou moins sur certains aspects selon l'instrument de financement dans lequel ils s'inscrivent afin de se différencier.

### 2.4.1. Pour tous les instruments de financement

En matière de Science Ouverte, les porteurs de projets ont intérêt à bâtir leur stratégie de dissémination autour de trois axes principaux : la diffusion des publications en libre accès, la diffusion des données et le lien entre Science Ouverte et valorisation de la recherche.

## Open Access

Cette section est l'opportunité de préciser la stratégie de diffusion qui est retenue pour les articles scientifiques : libre accès uniquement par l'intermédiaire de l'archivage en archive ouverte, ou une stratégie mixte qui inclut aussi la publication de quelques articles en libre accès dans des revues. Il est essentiel que la stratégie retenue par les différentes partenaires des consortiums soit cohérente.

Il est aussi important de préciser comment seront diffusées d'autres productions scientifiques, telles que : ouvrages ou chapitres d'ouvrages, actes de conférence, diaporamas présentés en conférence, posters... Même si l'ANR n'a pas émis d'obligations à ce sujet, les chercheurs ont intérêt à assurer une diffusion maximale de ces productions qui peuvent être déposées en archive ouverte en mentionnant le financement de l'ANR.

29

## Données de recherche

Diffuser des données fait désormais partie des attendus de l'ANR et c'est une bonne manière d'accroître l'impact d'un projet dans sa discipline mais aussi dans des disciplines connexes. C'est aussi un moyen d'assurer aux résultats obtenus par le projet une seconde vie après la fin du projet.

Il est préférable de distinguer les différentes manières de diffuser des données en valorisant en priorité les données diffusées dans des bases internationales, la création d'une base de données spécifique ou la publication d'un *data paper*.

Dans un second temps, il est possible de préciser comment les données qui sous-tendent les résultats présentés dans les publications, mais qui ne trouvent pas leur place dans des entrepôts disciplinaires, seront archivées et diffusées. L'utilisation d'un entrepôt généraliste comme Zenodo ou Dryad permettra d'obtenir un DOI pour les jeux de données et de faire un lien réciproque entre publications et jeux de données.

**Exemple** : des chercheurs publient un article dans la revue *Inorganic Chemistry*<sup>55</sup>. Le [jeu de données correspondant](#)<sup>56</sup> est déposé dans Zenodo. Un lien entre article et jeu de données est créé grâce au DOI.

Certaines données n'auront pas d'intérêt à être diffusées ou ne pourront pas l'être, selon le principe : aussi ouverte que possible, aussi fermé que nécessaire. Il peut être intéressant de préciser quelles données ne peuvent pas être diffusées et pourquoi, ainsi que ce qu'il est prévu d'en faire (suppression, accès confidentiel).

<sup>55</sup> <https://dx.doi.org/10.1021/acs.inorgchem.9b02096>

<sup>56</sup> <http://doi.org/10.5281/zenodo.3380352>



## De la Science Ouverte à la valorisation

La valorisation sociétale de la science est un enjeu grandissant pour les institutions de recherche et les agences de financement. Cette notion recouvre des réalités très différentes selon les projets et les disciplines.

L'essentiel est de retenir qu'il existe un continuum entre la Science Ouverte et la valorisation de la recherche, qui va de la diffusion des productions scientifiques à destination des communautés de recherche et des professionnels de son domaine (industriels, professionnels de santé...) jusqu'au grand public, en passant par des publics intermédiaires qui ont un rôle de médiation comme les enseignants du secondaire ou les journalistes scientifiques.

30

Présenter ses résultats de recherche sur des plateformes librement accessibles et destinées au « grand public » comme [The Conversation](#)<sup>57</sup> relève, par exemple, à la fois de la Science Ouverte et de la valorisation de la recherche.

En ce qui concerne la valorisation économique des productions de recherche, par l'intermédiaire de brevets par exemple, il faut s'assurer que la publication des données ne risque pas de mettre en péril la protection d'une invention. En cas de doute, il est préférable d'attendre le dépôt initial du brevet pour fixer l'antériorité de l'invention puis d'ouvrir les données dans un second temps, après un délai d'embargo. Les SATT ou les services de valorisation peuvent être sollicités sur ce sujet.

### 2.4.2. Approches spécifiques selon les instruments de financement

#### PRCI

Dans le cas des PRCI, il est intéressant de montrer la complémentarité des stratégies retenues par les partenaires des différents pays, tant pour la diffusion des publications que pour celle des données. L'utilisation de plateformes ou agrégateurs européens comme [Zenodo](#)<sup>58</sup> ou [OpenAIRE](#)<sup>59</sup> peut faciliter la valorisation des productions communes.

Il est important de bien anticiper cette partie, car chaque agence nationale peut avoir des obligations Science Ouverte spécifiques auxquelles sont soumis vos partenaires. Par exemple, le Fonds National Suisse (FNS<sup>60</sup>) demande un DMP dès la soumission du projet.

<sup>57</sup> <https://theconversation.com/fr>

<sup>58</sup> <https://zenodo.org/>

<sup>59</sup> <https://www.openaire.eu/>

<sup>60</sup> [http://www.snf.ch/fr/leFNS/points-de-vue-politique-de-recherche/open\\_research\\_data/Pages/default.aspx](http://www.snf.ch/fr/leFNS/points-de-vue-politique-de-recherche/open_research_data/Pages/default.aspx)

## PRCE

Il est important de bien tenir compte des aspects Science Ouverte dès la rédaction de l'accord de consortium qui lie les partenaires du projet car cela peut avoir un impact sur certains partages de propriété ou sur la conception de produits finaux comme des logiciels. Trouver le bon équilibre entre la protection, la valorisation des résultats (brevets, logiciels, bases de données, modélisations, etc.) et leur diffusion peut relever d'une stratégie mixte qu'il ne faut pas hésiter à développer.

**Exemple** : un logiciel est développé par un chercheur en collaboration avec une start-up lors d'un projet de recherche PRCE. Ses fonctionnalités de base seront gratuites et il sera placé sous une licence GPL, tandis que les développements complémentaires produits par la start-up seront payants. C'est un exemple de commercialisation d'un produit de recherche.

31

## PRC et JCJC

La demande ici s'apparente à un plan de dissémination, souvent demandé dans les AAP européens. La stratégie Science Ouverte doit être mentionnée comme un élément de dissémination, en précisant comment s'articulera la diffusion des résultats de la recherche à destination des différents acteurs du monde académiques (enseignants-chercheurs, étudiants, etc.) et d'autres cercles (science citoyenne, presse, décideurs, etc.). L'ouverture des résultats vers le grand public, ou vers d'autres publics extra-académiques, peut demander plus de temps que les formes traditionnelles de communication académiques, ce qu'il est nécessaire d'anticiper pour assurer une cohérence au projet.

## 2.5. Bibliographie et CV

La bibliographie et le CV ont une fonction d'annexe, mais ils peuvent aussi intégrer une dimension Science Ouverte.

### La bibliographie

- L'ANR recommande d'indiquer le **lien vers une version librement accessible des articles** quand elle existe. En plus de faciliter l'accès aux articles par les experts scientifiques et les membres des CES, cela montre une sensibilité à la Science Ouverte. Pour identifier rapidement s'il existe des versions en libre accès des références de la bibliographie, l'extension de navigateur [Unpaywall](https://unpaywall.org/products/extension)<sup>61</sup> est très utile.
- L'ANR demande à ce que **le facteur d'impact des revues ne figure pas dans la bibliographie**, comme préconisé par la Déclaration de San Francisco sur l'évaluation de la recherche ([DORA](https://sf-dora.org/read/)<sup>62</sup>). D'une manière générale, il est préférable d'éviter ce type de métriques dans la proposition détaillée. Cependant, certains CES y sont encore sensibles, et il est possible

<sup>61</sup> <https://unpaywall.org/products/extension>

<sup>62</sup> <https://sf-dora.org/read/>

de mentionner des métriques dans la présentation des membres de l'équipe ou du consortium. Il est préférable de privilégier des métriques à l'échelle des articles et des individus, comme le h-index, plutôt qu'à l'échelle des revues (facteur d'impact, CiteScore).

## Le CV

- Quand c'est possible, il est préférable que **les 5 publications majeures retenues par les responsables scientifiques soient accessibles en libre accès**. Si elles n'ont pas été publiées dans une revue en *Open Access*, c'est l'occasion de les déposer dans une archive ouverte. Une mention *Open Access* ou un lien vers la version librement accessible peuvent être ajoutés.
- Si les **données** qui sous-tendent ces publications ont été diffusées, elles peuvent aussi être signalées par un lien. Parmi les apports majeurs de ces publications, il ne faut pas négliger l'apport que peuvent représenter les données qui accompagnent les publications.
- La partie **valorisation** peut permettre de présenter des actions de valorisation des productions scientifiques comme la diffusion d'une base de données ou l'archivage et la diffusion de logiciels ouverts (dans [Software Heritage](https://www.softwareheritage.org/)<sup>63</sup> par exemple).

32

## Identifiant ORCID

- L'identifiant [ORCID](https://orcid.org/)<sup>64</sup> est composé de 16 chiffres. Il permet d'**identifier les chercheurs de manière univoque** même s'ils ont des homonymes ou plusieurs formes auteurs (ex. changement de nom en cours de carrière). Il peut être connecté avec d'autres identifiants chercheurs tels que l'IdHAL, le ScopusID ou le ResearcherID.
- Ce numéro n'est pas demandé dans le CV ni dans la proposition détaillée mais il est fourni pour les principales personnes impliquées dans le projet, lors de la complétion du formulaire en ligne de l'ANR. **Il est essentiel que le coordinateur et les responsables scientifiques disposent d'un identifiant ORCID et que leur profil soit complété**, à minima avec une liste de leurs publications. Les bibliothèques universitaires et les services d'information scientifique et technique offrent généralement une aide à la création et à la complétion des profils ORCID.

<sup>63</sup> <https://www.softwareheritage.org/?lang=fr>

<sup>64</sup> <https://orcid.org/>



## 3. Bibliographie

### 3.1. Politique publique en matière de Science Ouverte et réglementation

- ANR. *Règlement financier*. 2019. <https://ANR.fr/fr/rf/>
- Couperin, Science Ouverte France. FAQ sur la loi République numérique. 2018. <https://openaccess.couperin.org/category/faq/>
- LOI n° 2015-1779 du 28 décembre 2015 relative à la gratuité et aux modalités de la réutilisation des informations du secteur public. 2015-1779 <https://www.legifrance.gouv.fr/eli/loi/2015/12/28/2015-1779/jo/texte>
- LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique. 2016-1321. <https://www.legifrance.gouv.fr/eli/loi/2016/10/7/2016-1321/jo/texte>
- MESRI. *Guide d'application de la loi pour une République numérique (art. 30) - Écrits scientifiques*. 2018. <https://www.ouvrirelascience.fr/guide-application-loi-republique-numerique-article-30-ecrits-scientifiques-version-courte>
- MESRI. *Ouverture des données de recherche. Guide d'analyse du cadre juridique en France*. 2017. <https://prodinra.inra.fr/ft?id=C7D38E14-877E-4883-AB91-0536DD4D63B9>
- MESRI. *Plan national pour la Science Ouverte*. 2018. [https://cache.media.enseignementsup-recherche.gouv.fr/file/Actus/67/2/PLAN\\_NATIONAL\\_SCIENCE\\_OUVERTE\\_978672.pdf](https://cache.media.enseignementsup-recherche.gouv.fr/file/Actus/67/2/PLAN_NATIONAL_SCIENCE_OUVERTE_978672.pdf)

### 3.2. Guides et ressources pour mettre la Science Ouverte en pratique

- Commission Européenne. *How to complete your ethics self-assessment*. 2019. [https://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/ethics/h2020\\_hi\\_et\\_hics-self-assess\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_et_hics-self-assess_en.pdf)
- GOFAIR. *Fair principles*. 2016. <https://www.go-fair.org/fair-principles/>
- UK data service. *Data management costing tool and checklist*. 2015. <https://www.ukdataservice.ac.uk/media/622368/costingtool.pdf>
- Université de Lille. *WillO : droits et obligations des chercheurs de diffuser leurs publications en libre accès*. 2017. <https://decadoc.typeform.com/to/W2ZZMV>
- University of Manchester, School of Social Sciences. *Toolkit 08: Transcribing your own qualitative data*. 2010. <https://www.socialsciences.manchester.ac.uk/morgan-centre/research/resources/toolkits/toolkit-08/>
- Utrecht University. *Costs of Data Management*. 2018. <https://www.uu.nl/en/research/research-data-management/guides/costs-of-data-management>