

SCIENCE EUROPE

GUIDE PRATIQUE POUR UNE HARMONISATION INTERNATIONALE DE LA GESTION DES DONNÉES DE RECHERCHE

**Edition augmentée avec une rubrique pour l'évaluation
des Plans de gestion des données**

La traduction française a été réalisée par Manon Parrinello (Université de Haute-Alsace et Inist-CNRS) et Jean-François Nominé (Inist-CNRS). Cette traduction est publiée sous licence Creative Commons Attribution 4.0.

Contenu

Avant-propos de Thierry Damerval	3
Introduction.....	4
EXIGENCES MINIMALES POUR LES PLANS DE GESTION DES DONNÉES	6
Introduction.....	7
Exigences minimales pour les plans de gestion de données.....	8
CRITÈRES DE SÉLECTION DES ENTREPÔTS DE CONFIANCE	10
Introduction.....	11
Critères de sélection des entrepôts de confiance.....	12
CONSEILS DE MISE EN OEUVRE	13
Traduction de ces exigences minimales dans un modèle de DMP	14
Conseils pour la sélection des entrepôts dignes de confiance.....	20
DIRECTIVES POUR LES ÉVALUATEURS	23
Rubrique d'évaluation du plan de gestion des données	24
Rubrique d'évaluation des DMP.....	25
Notes et références	35
Annexe : Compatibilité avec les principes FAIR.....	36

Avant-propos de Thierry Damerval

Membre du conseil de direction de Science Europe et président de l'Agence nationale de la recherche

Janvier 2021

L'objectif de la science ouverte est de rendre les connaissances accessibles à tous. Elle encourage à faciliter la diffusion des résultats de la recherche et vise à améliorer son efficacité en favorisant la découverte, l'accessibilité, l'interopérabilité et la réutilisation des données. À l'Agence nationale de la recherche (ANR), nous apportons notre soutien à une harmonisation européenne et internationale des initiatives de partage des données de recherche, et ce en conformité avec le principe « aussi ouvert que possible, aussi fermé que nécessaire ». Nous encourageons nos chercheurs à prendre en considération la gestion des données de recherche (RDM) et le partage des données dès la phase d'élaboration d'un projet de recherche ainsi que tout au long de son cycle de vie.

La première édition du Guide pratique de Science Europe pour l'alignement international de la gestion des données de recherche a été publiée en janvier 2019 et a été rapidement reprise par de nombreux organismes de financement et opérateurs de recherche à travers toute l'Europe, jusqu'à la Commission européenne, qui l'utilisent pour définir leurs propres politiques de gestion de données de recherche et s'en servent également comme support de formations à l'adresse des chercheurs. La clé de ce succès a été l'engagement du Conseil néerlandais de la recherche (NWO), présidé par l'ancien membre du comité directeur de Science Europe, le Professeur Stan Gielen, et la promesse que ce dernier a fait en 2017 de défendre l'harmonisation de la gestion des données de la recherche entre les organismes de financement de la recherche en Europe.

Deux ans après avoir lancé la première édition de son guide, Science Europe en présente une nouvelle version enrichie. Suite aux demandes de nombreux acteurs de la recherche, un quatrième chapitre vient s'ajouter aux trois premiers et propose des conseils à ceux qui sont appelés à évaluer des plans de gestion de données.

Je suis fier que l'ANR ait été l'une des premières organisations à mettre en œuvre les recommandations de Science Europe et que nous ayons réussi à assurer son adoption au niveau national. Je suis convaincu que cette deuxième édition augmentée se révélera encore plus utile que la première. Alors qu'un nombre grandissant d'organismes qui ont pour mission de financer et d'opérer des travaux scientifiques s'engagent à améliorer la gestion des données, cet ouvrage de référence soutiendra véritablement l'harmonisation des politiques de gestion des données de recherche en Europe et au-delà.

Thierry Damerval

Introduction

Le secteur de la recherche connaît un important changement de paradigme vers la science ouverte et vise à rendre les résultats de la recherche disponibles pour être utilisés et réutilisés par d'autres chercheurs. Des données de recherche dont la qualité est garantie sont essentielles pour une bonne création de connaissances. Pour répondre à la nécessité de rendre le système de recherche dans son ensemble plus efficace vis-à-vis de l'utilisation des connaissances existantes, les données devraient être mises à disposition selon les principes FAIR, c'est-à-dire qu'elles devraient être trouvables, accessibles, interopérables et réutilisables. Des initiatives dans ce sens telles que l'European Open Science Cloud (EOSC) encouragent les données de recherche FAIR et nécessiteront des pratiques institutionnelles rigoureuses pour garantir un partage des données dans des conditions optimales. Il est donc essentiel que le secteur de la recherche publique joue un rôle de premier plan dans l'instauration et la mise en œuvre de politiques et de pratiques de gestion des données de recherche (RDM) pour les données FAIR.

Agences de financement de la recherche, organismes de rechercheⁱ, chercheurs individuels et examinateurs ont des exigences et des besoins différents au regard des politiques et pratiques de gestion des données de recherche (RDM). Ce guide a pour objet de présenter les exigences de base des plans de gestion des données (PGD), les critères de sélection d'entrepôts dignes de confiance, des conseils de mise en œuvre aux chercheurs pour satisfaire à ces exigences, et une rubrique avec des prescriptions pour l'évaluation des PGD. Par ailleurs, les lignes directrices présentées dans ce guide fournissent aux organisations et aux communautés de la recherche une base commune à partir de laquelle elles peuvent élaborer leurs propres politiques de RDM. Ce sont là des exigences minimales qui peuvent être aménagées pour s'adapter aux politiques ou aux pratiques institutionnelles ou disciplinaires.

Ce guide a été élaboré pour aider les chercheurs à s'assurer que les données sont FAIR. Il peut y avoir des raisons légitimes (liées à des spécificités du projet ou relatives à des données personnelles) pour en différer ou limiter l'accès, qui impose une approche équilibrée en matière d'ouverture des données de recherche. Notre guide va même au-delà des principes et leur respect pour couvrir d'autres aspects comme le stockage des données, leur sauvegarde pendant le temps du projet et leur préservation à long termeⁱⁱ.

Ce guide a été élaboré par des experts appartenant à des organisations membres de Science Europe, et des acteurs du milieu de la recherche ont également été consultés pour prendre compte les différents besoinsⁱⁱⁱ. Il a pour but d'être au service de tous organismes de financement, opérateurs de recherche et chercheurs. Il met l'accent sur les questions de fond et laisse aux lecteurs toute liberté d'adaptation vis-à-vis des politiques et procédures des organismes et des communautés disciplinaires^{iv}.

Ce guide est divisé en quatre parties :

Exigences minimales pour les plans de gestion des données : six points que tout DMP devrait couvrir tous les aspects, illustrés avec des questions clés.

Critères de sélection des dépôts dignes de confiance : quatre points détaillant les critères que tout entrepôt de confiance doit respecter.

Conseils de mise en œuvre pour les chercheurs : renseignements détaillés et exemples pour guider la mise en œuvre des exigences et des critères dans les politiques d'établissement par les chercheurs.

Directives pour les évaluateurs : conseils aux examinateurs pour évaluer des DMP, en application des exigences de base présentées dans les chapitres précédents.

COMMENT UTILISER CE GUIDE

Agences de financement de la recherche, organismes et communautés de recherche sont encouragés à utiliser **ces exigences minimales pour les plans de gestion des données** comme une base pour établir leurs propres modèles de DMP ainsi que la **rubrique consacrée à l'évaluation des DMP**.

Les **conseils de mise en œuvre** sont destinés aux chercheurs dans l'élaboration et la mise à jour de leurs DMP ainsi que dans la gestion de leurs données sur tout le cycle de recherche. Ils offrent aussi des informations complémentaires aux établissements désireux d'apporter un soutien à leurs chercheurs dans ce domaine.

Pour les autres acteurs du secteur de la recherche, ce guide sert de document de référence sur la façon dont un DMP devrait être structuré et utilisé.

Les **critères de sélection des entrepôts dignes de confiance** associés à des conseils spécifiques aideront les agences de financement, les organismes de recherche et les chercheurs individuels à identifier des entrepôts adéquats pour le stockage et le partage des données.

Les exigences minimales pour les DMP et les critères de sélection des entrepôts dignes de confiance peuvent être considérés comme deux documents autonomes et utilisés indépendamment. Il est toutefois recommandé de tenir compte de ces deux ensembles lors de l'élaboration ou de la mise à jour d'une politique de données institutionnelle ou disciplinaire, de façon à aligner au mieux les positions des institutions.

Les organismes désireux d'adapter ces conseils en fonction de leurs politiques d'établissement ou domaines disciplinaires pourront trouver des modèles correspondant aux différentes parties du présent guide et dans un format adaptable sur le site de Science Europe : <https://scieur.org/rdm>.

EXIGENCES MINIMALES POUR LES PLANS DE GESTION DES DONNÉES

Introduction

Les organismes de financement de la recherche et les organismes de recherche exigent de plus en plus souvent des plans de gestion des données. Ces plans aident le chercheur à prendre en compte tous les aspects pertinents de la gestion des données dès l'entame d'un projet de recherche. Un DMP devrait inciter les chercheurs à réfléchir à des solutions optimales en matière de traitement, d'organisation, de documentation et de stockage de leurs données.

À l'heure actuelle, les politiques de gestion des données de recherche varient beaucoup. Nombreux sont les agences de financement, organismes et communautés de recherche ayant élaboré leurs propres règles et modèles. Cela peut prêter à confusion pour les chercheurs et est d'autant plus problématique que de nombreux chercheurs obtiennent leur financement de sources différentes ; ils sont ainsi de plus en plus confrontés à des exigences différentes entre sources de financement et politiques institutionnelles. Il est urgent d'harmoniser les politiques de gestion des données afin d'apporter plus de clarté aux chercheurs. Les DMP ne devraient pas apparaître comme une tâche administrative supplémentaire, mais un moyen utile d'aide à la planification et la réalisation d'un projet de projet de recherche.

La suite de ce document présente six thèmes qui devraient être abordés dans les DMP, chacun d'eux accompagné de plusieurs questions d'orientation. Ces sujets et questions pour la mise en place d'un DMP constituent les exigences minimales que tout organisme de financement de la recherche doit demander au chercheur en vue d'élaborer un DMP utile. L'ordre des sujets à traiter peut être adapté en fonction de besoins spécifiques et de l'organisation souhaitée. Mais les six thèmes doivent être abordés dans le DMP.

Un exemple de modèle fournissant des conseils sur les aspects à approfondir dans le cadre d'un DMP se trouve à la page 14 du présent guide.

Exigences minimales pour les plans de gestion de données

Dans l'élaboration de robustes plans de gestion des données, il est demandé aux chercheurs de traiter les points suivants et de répondre aux questions suivantes :

1. Description des données et collecte ou réutilisation des données existantes

- a. Comment de nouvelles données seront-elles recueillies ou produites et/ou comment des données préexistantes seront-elles réutilisées ?
- b. Quelles données (types, formats et volumes par ex.) seront collectées ou produites ?

2. Documentation et qualité des données

- a. Quelles métadonnées et quelle documentation (par exemple méthodologie de collecte et mode d'organisation des données) accompagneront les données ?
- b. Quelles mesures de contrôle de la qualité des données seront mises en œuvre ?

3. Stockage et sauvegarde pendant le processus de recherche

- a. Comment les données et les métadonnées seront-elles stockées et sauvegardées tout au long du processus de recherche ?
- b. Comment la sécurité des données et la protection des données sensibles seront-elles assurées tout au long du processus de recherche ?

4. Exigences légales et éthiques, codes de conduite

- a. Si des données à caractère personnel sont traitées, comment le respect des dispositions de la législation sur les données à caractère personnel et sur la sécurité des données sera-t-il assuré ?
- b. Comment les autres questions juridiques, comme la titularité ou les droits de propriété intellectuelle sur les données, seront-elles abordées ? Quelle est la législation applicable en la matière ?
- c. Comment les éventuelles questions éthiques seront-elles prises en compte, les codes déontologiques respectés ?

5. Partage des données et conservation à long terme

- a. Comment et quand les données seront-elles partagées ? Y-a-t-il des restrictions au partage des données ou des raisons de définir un embargo ?
- b. Comment les données à conserver seront-elles sélectionnées et où seront-elles préservées sur le long terme (par ex. un entrepôt de données ou une archive) ?
- c. Quelles méthodes ou quels outils logiciels seront nécessaires pour accéder et utiliser les données ?
- d. Comment l'application d'un identifiant unique et pérenne (comme le DOI) sera réalisée pour chaque jeu de données ?

6. Responsabilités et ressources en matière de gestion des données

- a. Qui (par exemple rôle, position et institution de rattachement) sera responsable de la gestion des données (c.-à-d. le gestionnaire des données) ?
- b. Quelles seront les ressources (budget et temps alloués) dédiées à la gestion des données permettant de s'assurer que les données soient FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable) ?

CRITÈRES DE SÉLECTION DES ENTREPÔTS DE CONFIANCE

Introduction

L'accès aux données est l'un des piliers d'une science saine et reproductible. De plus en plus d'organismes, d'institutions et d'agences de financement de la recherche exigent que les chercheurs déposent les données de recherche dans des entrepôts. Les chercheurs ont besoin d'identifier des entrepôts dignes de confiance où ils peuvent stocker leurs données à des fins de partage. Il n'existe actuellement aucune liste de référence pour de tels entrepôts de confiance, alors que les catalogues d'entrepôts existants en répertorient plus de 2 000. Toutefois, la maturité et la fiabilité de ces dépôts sont difficiles à évaluer. De nombreux entrepôts n'ont pas encore demandé à être certifiés par un organisme de certification reconnu. Identifier un entrepôt approprié peut donc représenter une tâche difficile pour les chercheurs, leurs institutions d'appartenance, et les agences de financement.

Dans certaines disciplines, les chercheurs travaillent avec des entrepôts disciplinaires qui ont déjà mis en place certaines politiques et normes qui répondent aux besoins spécifiques d'une communauté. D'autres entrepôts de données servent un public de chercheurs plus large, et en conséquence leurs politiques et leurs normes sont plus génériques.

Certains dépôts ont été certifiés comme entrepôts dignes de confiance par un ou plusieurs organismes de certification reconnus. Afin de faciliter le repérage par les chercheurs des entrepôts de confiance, il est fortement recommandé que les entrepôts non encore certifiés demandent leur labellisation.

Il est toujours recommandé de se référer en premier lieu à des entrepôts disciplinaires ou certifiés largement reconnus. Quand de tels entrepôts ne peuvent être identifiés, alors les critères de sélection d'entrepôts présentés dans ce guide peuvent être utilisés ?

La liste de critères présentée dans ce guide se compose d'un certain nombre de critères minimaux, organisés autour de quatre thèmes majeurs, que tous les entrepôts de confiance doivent satisfaire. Cette liste n'accorde pas la priorité à un critère plutôt qu'à un autre.

Des explications plus détaillées sur les critères de sélection des entrepôts de confiance trouvent à la page 20 du présent guide.

Critères de sélection des entrepôts de confiance

Les entrepôts dignes de confiance doivent répondre aux critères minimaux suivants :

1. Attribution d'identifiants uniques et pérennes (PIDs)

- a. Permettre le repérage et l'identification des données
- b. Permettre la recherche, la citation et la récupération de données
- c. Permettant la gestion des versions des jeux de données

2. Métadonnées

- a. Permettre le repérage de données
- b. Permettre de référencer des informations pertinentes connexes, telles que d'autres jeux de données et publications
- c. Fournir de l'information librement accessible et mise à jour, y compris sur des données non publiées, protégées, rétractées, ou supprimées
- d. Utiliser des normes de métadonnées qui sont largement acceptées (par la communauté scientifique)
- e. S'assurer que les métadonnées sont récupérables par programme

3. Accès aux données et licences d'utilisation

- a. Permettre l'accès aux données dans des conditions bien définies
- b. Garantir l'authenticité et l'intégrité des données
- c. Permettre la récupération des données
- d. Fournir de l'information sur les licences et les autorisations d'accès aux données (idéalement sous forme lisible par programme)
- e. Garantir la confidentialité et respecter les droits des personnes concernées et des créateurs de données.

4. Préservation

- a. Assurer la pérennité des métadonnées et des données
- b. Faire preuve de transparence au sujet des missions, périmètre, politique de préservation, et des plans (y compris la gouvernance, la pérennité budgétaire, et plan de maintenance)

CONSEILS DE MISE EN OEUVRE

Traduction de ces exigences minimales dans un modèle de DMP

L'exemple suivant d'un modèle de plan de gestion des données est basé sur les exigences minimales pour les DMP¹. Ce minimum requis devrait être considéré comme un standard de base, tout en laissant la possibilité de formuler des spécifications additionnelles en fonction de besoins particuliers de certains domaines ou liées à des législations locale ou nationale.

Le modèle présenté ci-dessous se réfère aux 15 questions couvrant les six points principaux exigés pour une bonne gestion des données. Les présents conseils de mise en œuvre et explications supplémentaires ont pour but d'aider les chercheurs à renseigner le modèle et pour s'assurer que tous les aspects pertinents de la gestion des données sont effectivement couverts. Le tableau ci-dessous est un exemple de la façon dont les exigences minimales définies auparavant peuvent être traduites en un modèle de DMP. Il appartiendra à chaque organisation et à chaque discipline d'élaborer des modèles qui correspondent à leurs besoins.

INFORMATIONS GÉNÉRALES

Renseignements administratifs

- Fournir des renseignements d'identification tels que le nom du déposant, numéro du projet, du programme de financement, version du DMP.

1 DESCRIPTION DES DONNÉES ET COLLECTE OU RÉUTILISATION DE DONNÉES EXISTANTES

1a Comment de nouvelles données seront-elles recueillies ou produites et/ou comment des données préexistantes seront-elles réutilisées ?

- Expliquer quelles méthodologies ou quels logiciels sera utilisés si de nouvelles données sont recueillies ou produites.
- Énoncer les éventuelles contraintes à la réutilisation des données préexistantes.
- Expliquer comment la provenance des données sera documentée.
- Indiquer brièvement le cas échéant, les raisons pour lesquelles l'utilisation de sources de données existantes a été envisagée mais écartée.

1b Quelles données (types, formats et volumes par ex.) seront collectées ou produites ?

- Donnez des détails sur le type de données : par exemple numérique (bases de données, tableurs), textuel (documents), image, audio, vidéo, et/ou médias composites.
- Détailler le format des données : la manière selon laquelle les données sont codées pour le stockage, généralement reflétée par l'extension du nom de fichier (par exemple PDF, XLS, DOC, TXT, ou RDF).
- Justifier l'utilisation de certains formats. Par exemple, les choix d'un format peuvent être guidés par l'expertise du personnel de l'organisme, ou par une préférence pour les formats ouverts, par les standards de format acceptés par les entrepôts de données, par l'usage

largement répandu dans une communauté de recherche ou par le logiciel ou l'équipement qui sera utilisé.

- Privilégier les formats standards et ouverts car ils facilitent le partage et la réutilisation à long terme des données (plusieurs catalogues fournissent des listes de ces « formats préférés »).
- Donnez des détails sur les volumes (qui peuvent être exprimés en espace de stockage requis (octets), et/ou en quantités d'objets, de fichiers, de lignes, et colonnes).

2 DOCUMENTATION ET QUALITÉ DES DONNÉES

2a Quelles métadonnées et quelle documentation (par exemple méthodologie de collecte et mode d'organisation des données) accompagneront les données ?

- Indiquez quelles métadonnées seront fournies pour aider à la recherche et à l'identification des données.
- Indiquez quelles standards de métadonnées seront utilisés (par exemple DDI, TEI, EML, MARC, CMDI).
- Utiliser les standards de métadonnées de communauté lorsque ceux-ci existent.
- Indiquez comment les données seront organisées au cours du projet, en mentionnant par exemple conventions de nommage, contrôle de version et structures des dossiers. Des données bien classées et gérées de façon cohérente seront plus faciles à retrouver, à comprendre et à réutiliser.
- Pensez à la documentation qui serait nécessaire pour permettre une réutilisation des données. Il peut s'agir notamment de l'information sur la méthodologie utilisée pour collecter les données, sur les procédures et méthodes d'analyse utilisées, sur la définition des variables, des unités de mesure, etc.
- Tenez compte de la façon dont ces informations seront obtenues et enregistrées par exemple dans une base de données avec des liens vers chacun des fichiers, dans un fichier texte de type « lisez-moi », dans les en-têtes de fichiers, dans un livre de référence (« code book ») ou dans les cahiers de laboratoire.

2b Quelles mesures de contrôle de la qualité des données seront mises en œuvre ?

- Expliquer comment la qualité et la conformité de la collecte des données seront contrôlées et documentées. Il s'agit là de préciser les processus comme la calibration, la répétition des échantillons ou des mesures, la capture standardisée des données, la validation de saisie des données, la revue par les pairs, ou la représentation basée sur des vocabulaires contrôlés.

3 STOCKAGE ET SAUVEGARDE PENDANT LE PROCESSUS DE RECHERCHE

3a Comment les données et les métadonnées seront-elles stockées et sauvegardées tout au long du processus de recherche ?

- Décrire l'endroit où les données seront stockées et sauvegardées au cours du processus de recherche et la fréquence à laquelle la sauvegarde sera effectuée. Il est recommandé de stocker les données dans au moins deux lieux distincts.
- Privilégier l'utilisation de systèmes de stockage robustes, avec sauvegarde automatique, tels que ceux fournis par les services informatiques de l'institution d'origine. Le stockage des

données sur des ordinateurs portables, des disques durs externes, ou des périphériques de stockage tels que des clés USB n'est pas recommandé.

3b Comment la sécurité des données et la protection des données sensibles seront-elles assurées tout au long du processus de recherche ?

- Expliquez comment les données seront récupérées en cas d'incident.
- Expliquer qui aura accès aux données au cours du processus de recherche et comment l'accès aux données est contrôlé, en particulier dans le cadre de recherches menées en collaboration.
- Tenez compte de la protection des données, en particulier si vos données sont sensibles (par exemple données à caractère personnel, politiquement sensibles des informations ou secrets commerciaux). Décrire les principaux risques et la façon dont ils seront gérés.
- Expliquer quelle politique institutionnelle de protection des données institutionnelles est mis en œuvre.

4 EXIGENCES LÉGALES ET ÉTHIQUES, CODES DE CONDUITE

4a Si des données à caractère personnel sont traitées, comment le respect des dispositions de la législation sur les données à caractère personnel et sur la sécurité des données sera-t-il assuré ?

- Lorsque vous manipulez des données à caractère personnel, veillez à ce que les lois sur la protection des données (par exemple, RGPD) soient appliquées, notamment :
 - Obtenir un consentement éclairé pour la préservation et/ou le partage de données personnelles.
 - Envisager l'anonymisation des données personnelles pour la préservation et/ou le partage (des données correctement anonymisées ne sont plus considérées comme des données personnelles).
 - Envisager la pseudonymisation des données personnelles (la principale différence avec l'anonymisation est que la pseudonymisation est réversible).
 - Envisagez le chiffrement des données, qui est considéré comme un cas particulier de pseudonymisation (la clé de cryptage doit alors être stockée séparément des données, par exemple par un tiers de confiance).
 - Expliquez si une procédure d'accès spécifique a été mise en place pour les utilisateurs autorisés à accéder aux données personnelles.

4b Comment les autres questions juridiques, comme la titularité ou les droits de propriété intellectuelle sur les données, seront-elles abordées ? Quelle est la législation applicable en la matière ?

- Expliquez qui sera le propriétaire des données, c'est-à-dire qui aura le droit d'en contrôler l'accès :
 - Expliquez quelles conditions d'accès s'appliqueront aux données ? Les données seront-elles librement accessibles, ou des restrictions seront-elles appliquées ? Dans Si oui, lesquelles ? Envisager l'utilisation de licences concernant l'accès et la réutilisation des données.

- Assurez-vous de couvrir, dans l'accord de consortium, ces questions de droits de contrôle d'accès aux données pour les projets multipartenaires et en cas de propriété partagée des données.
- Indiquez si les droits de propriété intellectuelle (par exemple la directive bases de données, droits *sui generis*) sont affectés. Dans l'affirmative, veuillez expliquer lesquels et comment cela sera-t-il traité.
- Indiquez s'il y a des restrictions sur la réutilisation des données fournies par des tiers.

4c Comment les éventuelles questions éthiques seront-elles prises en compte, les codes déontologiques respectés ?

- Déterminer si les questions d'éthique ont une incidence sur la façon dont les données sont stockées et transférées, qui peut les voir ou les utiliser, et quelle durée de conservation leur seront-elles appliquées. Démontrer que ces aspects sont bien pris en compte et planifiés.
- Adoptez les codes de conduite nationaux et internationaux et le code d'éthique institutionnel et vérifiez si une revue des pratiques (par exemple par un comité d'éthique) est requise pour ce qui concerne la collecte de données dans le cadre du projet de recherche.

5 PARTAGE DES DONNÉES ET CONSERVATION À LONG TERME

5a Comment et quand les données seront-elles partagées ? Y-a-t-il des restrictions au partage des données ou des raisons de définir un embargo ?

- Expliquez comment les données pourront être retrouvées et partagées (par exemple, par dépôt dans un entrepôt de données de confiance, indexé dans un catalogue, par l'utilisation d'un service de données sécurisé, par traitement direct des demandes de données, ou l'utilisation de tout autre mécanisme).
- Définissez le plan de préservation des données et fournir l'information sur la durée d'archivage pérenne des données.
- Expliquez à quel moment les données seront rendues disponibles. Indiquer les délais de publication prévus. Expliquez si une utilisation exclusive des données est revendiquée et, dans l'affirmative, pour quelle raison et pour combien de temps. Indiquez si le partage des données sera différé ou limité, par exemple pour des raisons de publication, pour protéger la propriété intellectuelle, ou dépôt de brevets.
- Indiquez qui pourra utiliser les données. S'il s'avère nécessaire de restreindre l'accès pour certaines communautés ou d'imposer un accord pour le partage de données, expliquez comment et pourquoi. Expliquez les mesures qui seront prises pour dépasser ou minimiser ces restrictions.

5b Comment les données à conserver seront-elles sélectionnées et où seront-elles préservées sur le long terme (par ex. un entrepôt de données ou une archive) ?

- Indiquez quelles données ne doivent pas être divulguées ou doivent être détruites pour des raisons contractuelles, légales, ou réglementaires.
- Indiquez comment il sera décidé quelles données garder. Décrire les données qui seront à préserver à long terme.

- Décrire les utilisations (et/ou les utilisateurs) prévisibles des données dans un cadre de recherche.
- Indiquez où les données seront déposées. Si aucun entrepôt reconnu n'est proposé, démontrer dans le DMP que les données pourront être prises en charge efficacement au-delà de la durée de financement du projet. Il est recommandé de démontrer que les politiques des entrepôts et les procédures de dépôts (y compris les normes de métadonnées, et coûts mis en œuvre) ont été vérifiés.

5c Quelles méthodes ou quels outils logiciels seront nécessaires pour accéder et utiliser les données ?

- Indiquez si les utilisateurs potentiels ont besoin d'outils spécifiques pour l'accès et la (ré)utilisation des données. Tenez compte de la durée de vie des logiciels nécessaires pour accéder aux données.
- Indiquez si les données seront partagées via un entrepôt, si les demandes d'accès seront traitées en direct, ou si un autre mécanisme sera utilisé ?

5d Comment l'application d'un identifiant unique et pérenne (comme le DOI) sera réalisée pour chaque jeu de données ?

- Expliquez comment les données pourraient être réutilisées dans d'autres contextes. Les identifiants persistants (PID) devraient être appliqués de manière à ce que les données puissent être localisées et référencées de façon fiable et efficace. Ces PID aident aussi à comptabiliser les citations et les réutilisations.
- Indiquez s'il sera envisagé d'attribuer aux données un PID. Typiquement, un entrepôt pérenne digne de confiance attribuera des identifiants persistants.

6 RESPONSABILITÉS ET RESSOURCES EN MATIÈRE DE GESTION DES DONNÉES

6a Qui (par exemple rôle, position et institution de rattachement) sera responsable de la gestion des données (c.-à-d. le gestionnaire des données) ?

- Décrivez les rôles et les responsabilités concernant les activités de gestion des données par exemple : saisie de données, production de métadonnées, qualité des données, stockage et sauvegarde, archivage et partage des données. Nommez les personne(s) responsable(s) impliquées dans la mesure du possible.
- Pour les projets menés en collaboration, expliquez comment s'effectue la coordination des responsabilités de gestion des données entre partenaires.
- Indiquez qui est responsable de la mise en œuvre du DMP, et qui s'assure qu'il est examiné et, si nécessaire, révisé.
- Envisagez des mises à jour régulières du DMP.

6b Quelles seront les ressources (budget et temps alloués) dédiées à la gestion des données permettant de s'assurer que les données soient FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable) ?

- Expliquez comment les ressources nécessaires (par ex, le temps) à la préparation des données pour le partage/préservation (curation des données) ont été chiffrées. Examinez et justifiez soigneusement toutes les ressources nécessaires pour distribuer les données.

- Il peut s'agir de frais de stockage, de coût matériel, de temps de personnel, de coûts de préparation des données pour le dépôt, de frais d'entrepôt et d'archivage.
- Indiquez si des ressources supplémentaires sont nécessaires pour préparer les données en vue de leur dépôt ou pour payer tous les frais demandés par les entrepôts de données. Si oui, précisez-en le montant et comment ces coûts seront couverts.

Conseils pour la sélection des entrepôts dignes de confiance

Le tableau suivant fournit des conseils pour la sélection des entrepôts dignes de confiance, structurés selon quatre thèmes principaux.

1 ATTRIBUTION D'IDENTIFIANTS UNIQUES ET PÉRENNES (PIDS)

Un dépôt digne de confiance devrait :

1a Permettre le repérage et l'identification des données

- S'assurer que des PIDs sont effectivement présents dans les métadonnées associées.

1b Permettre la recherche, la citation et la récupération de données

- affecter systématiquement des PID (par exemple un DOI^v, URN^{vi}, ARK^{vii}) aux données qu'il détient, ce qui permet aux données et métadonnées associées d'être localisées, référencées et récupérées, même en cas de changement de lieu de stockage des données.

1c Permettant la gestion des versions des jeux de données

- s'assurer que la version des données stockées dans l'entrepôt est clairement spécifiée et documentée par le biais d'un système d'audit permanent pour permettre d'en retracer la provenance.

Nota : Tous les entrepôts n'utilisent pas un système de PID accepté et universel, comme ceux mentionnés ci-dessus. En lieu et place, ils utilisent un identifiant local ou un numéro administratif que l'entrepôt lui-même administre. Cela augmente le risque que les données ne puissent pas être localisées si elles viennent à être déplacées, ou si l'entrepôt cesse d'exister, se réorganise ou modifie sa gouvernance.

2 MÉTADONNÉES

Les données doivent être décrites avec précision à l'aide d'un jeu métadonnées riche. Les métadonnées doivent renseigner sur la façon dont les données ont été générées, sous quelle licence et comment elles peuvent être réutilisées et doivent fournir le contexte pour une interprétation appropriée par d'autres chercheurs.

Un entrepôt digne de confiance devrait :

2a Permettre le repérage des données

- assurer l'interopérabilité et la réutilisation des données par d'autres en fournissant données et métadonnées dans un langage accessible, basé sur un formalisme bien établi. Données et métadonnées doivent être décrites en utilisant des vocabulaires standards et des formats permettant aux systèmes informatiques de les rechercher, les combiner de manière automatique et de distinguer métadonnées et fichiers de données.

2b Permettre de référencer des informations pertinentes connexes, telles que d'autres jeux de données et publications

- s'assurer qu'il est possible de déclarer dans les métadonnées des liens vers des informations pertinentes ou connexes en fournissant un PID et une description de la relation scientifique. Un type particulier d'information concernerait les chercheurs associés, pour lesquels des identifiants pérennes existent (tels que ORCID^{viii}, ISNI^{ix}, ou DAI^x).

2c Fournir de l'information librement accessible et mise à jour, y compris sur des données non publiées, protégées, rétractées, ou supprimées

- s'assurer que les métadonnées sont archivées sur le long terme et qu'elles restent toujours consultables, même si les données de recherche correspondantes ne sont pas ou plus disponibles (par exemple en raison d'une restriction de confidentialité, d'obligations légales, ou d'autres mesures de protection).
- s'assurer que les données rétractées en raison de mauvaises pratiques de recherche ou de méconduites peuvent encore être localisées grâce aux métadonnées et conservées dans des bases de données afin d'en permettre l'examen.

2d Utiliser des normes de métadonnées qui sont largement acceptées (par la communauté scientifique)

- s'assurer que les métadonnées conservées par l'entrepôt sont lisibles par programme et respectent des normes largement acceptées par la communauté scientifique.
- s'assurer que les normes des communautés ou les bonnes pratiques en matière de traitement des données sont observées, si elles existent. Notez que les entrepôts spécialisés dans un domaine de recherche particulier peuvent avoir des normes de communautés concernant les données et les métadonnées qui y sont déposées.

2e S'assurer que les métadonnées sont récupérables par programme

- encourager à la structuration des métadonnées de manière à permettre leur récupération par programme, par exemple en fournissant un formulaire avec des champs spécifiques à remplir.

3 ACCÈS AUX DONNÉES ET LICENCES D'UTILISATION

Un entrepôt digne de confiance devrait :

3a Permettre l'accès aux données dans des conditions bien définies

- préciser clairement les conditions dans lesquelles les données peuvent être réutilisées. De telles conditions (licence) sont généralement incluses dans les métadonnées.

3b Garantir l'authenticité et l'intégrité des données

- s'assurer que les métadonnées contiennent de l'information détaillée sur la provenance des données, y compris la façon dont elles ont été générées, traitées, dans quel contexte elles peuvent être réutilisées, et quel est leur degré de fiabilité.

3c Permettre la récupération des données

- permettre la récupération de données ou au moins de métadonnées à l'aide d'un protocole standardisé et ouvert (pas un protocole de communication propriétaire).

3d Fournir de l'information sur les licences et les autorisations d'accès aux données (idéalement sous forme lisible par programme)

- permettre de faire référence de façon structurée à la licence d'utilisation, de façon à ce que les conditions d'utilisation soient claires, de préférence pour l'homme comme pour la machine.
Dans la mesure du possible, les licences communes ou largement acceptés (comme Creative Commons) devraient être utilisés et référencés par URL.

3e Garantir la confidentialité et respecter les droits des personnes concernées et des créateurs de données.

- fournir un moyen d'authentification et d'autorisation des personnes et des machines utilisatrices, permettant de définir des droits d'accès spécifiques à l'utilisateur (ou au groupe) pour tenir compte des données présentant des problèmes de confidentialité ou d'autres restrictions.

4 PRÉSERVATION

Un dépôt digne de confiance devrait :

4a Assurer la pérennité des métadonnées et des données

- assurer la préservation, la disponibilité et l'accès sur la durée aux données et métadonnées qui lui sont confiées par les utilisateurs.

4b Faire preuve de transparence au sujet des missions, périmètre, politique de préservation, et de l'organisation (y compris la gouvernance, la pérennité budgétaire, et le plan de maintenance)

- gérer la préservation des données et des métadonnées de façon documentée. En particulier, l'entrepôt devrait disposer d'une politique de préservation qui détaille les missions et périmètre de l'entrepôt, les aspects relatifs à la gouvernance, la viabilité budgétaire, les partenaires sous-traitants, et les durées de conservation (le calendrier de mise en œuvre de la préservation).
- disposer d'un plan d'urgence accessible publiquement et assurer la préservation des données et métadonnées au-delà de l'existence même de l'entrepôt (par exemple en permettant l'extraction et le transfert de données et de métadonnées vers un autre entrepôt).

DIRECTIVES POUR LES ÉVALUATEURS

Rubrique d'évaluation du plan de gestion des données

1 À QUOI SERVENT LES RECOMMANDATIONS POUR L'ÉVALUATION DES DMP ?

Lorsque les organismes de recherche (de financement) demandent aux chercheurs d'élaborer des DMP, ceux-ci doivent être en mesure d'assurer un suivi de leur évaluation et une remontée d'information vers les chercheurs. Les présentes recommandations sont conçues pour étayer et guider l'évaluation des DMP en incitant à opérer une analyse où il est vérifié que tous les aspects requis sont traités. Celles-ci sont rédigées de manière générique et délibérément en langage clair. En plus d'être destinée à conseiller les chercheurs, la présente rubrique d'évaluation des DMP vise également à garantir la « FAIRitude » des données, même si cette notion n'est pas toujours explicitement mentionnée dans toute cette rubrique.

2 POURQUOI LA PRÉSENTE MISE EN FORME ?

Ces conseils pour l'évaluation des DMP reprennent les exigences et lignes directrices des chapitres précédents. Elle propose des critères qui aident l'évaluateur dans son approche des informations fournies dans les DMP pour juger si elles sont suffisantes, donc garantir que l'équipe de recherche gèrera les données comme prévu. Ils se présentent sous la forme d'une rubrique qui énumère les différents critères et niveaux de performance qui indiquent dans quelle mesure ces critères sont respectés. Cette rubrique contient deux niveaux de performance : « suffisamment traité » et « insuffisamment traité ». L'expression « insuffisamment traité » fait référence à des informations insuffisantes ou jugées incorrectes.

3 CETTE RUBRIQUE PEUT-ELLE ÊTRE ADAPTEE ?

Suivant la même structure que les chapitres précédents, cette rubrique présente les critères de base pour l'analyse tout en laissant aux organismes une certaine souplesse pour s'adapter aux cadres législatifs, aux situations institutionnelles ou aux impératifs disciplinaires. Les différents éléments de ces recommandations peuvent et doivent être adaptés en conséquence. Ces modifications doivent être mentionnées à la fois dans les indications données aux chercheurs et aux évaluateurs des DMP. Il est donc important de souligner que cette rubrique doit être considérée comme des orientations, et non comme un outil prêt à l'emploi (comme une liste de contrôle). Les organisations qui souhaitent développer des listes de contrôle servant à l'évaluation des DMP peuvent utiliser la rubrique comme base et l'adapter en conséquence.

4 POUR QUI ?

Ces conseils seront utiles à toute personne amenée à évaluer un DMP. Il s'agit notamment des agents de la recherche, des évaluateurs ou des gestionnaires de données dans les institutions. Les chercheurs les trouveront également utiles comme source d'information complémentaire (en s'appuyant sur la section précédente). Les évaluateurs qui utilisent la rubrique sont fortement encouragés à ne pas l'utiliser uniquement comme un exercice où l'on se limite à cocher des cases, mais plutôt à s'en servir pour réunir leurs commentaires et leurs notes. Le fait de partager ces commentaires avec les chercheurs apportera à ces derniers un soutien et des éclaircissements supplémentaires pour améliorer leurs DMP.

5 QUAND PEUT-ON L'UTILISER ?

Les DMP sont examinés à différentes étapes du cycle de vie des projets de recherche, en fonction des politiques institutionnelles. La grille d'évaluation est conçue pour fonctionner parallèlement à ce processus et peut être utilisée chaque fois qu'un DMP est évalué. Les évaluateurs doivent garder à l'esprit qu'un DMP est un document qui change au fil du temps. Le niveau de détail fourni dans un DMP peut varier selon la version évaluée. Par exemple, selon qu'il s'agit d'une première version au moment d'une demande de financement ou une version ultérieure accompagnant le dépôt dans un entrepôt.

Rubrique d'évaluation des DMP

Question concernant les DMP	Recommandations applicables aux DMP	Niveaux de performance	
INFORMATIONS GENERALES			
Prescriptions pour les chercheurs		Suffisamment traité Le DMP...	Insuffisamment traité Le DMP...
Informations administratives	- Fournir des informations telles que le nom du porteur de projet, le numéro du projet, le programme de financement et la version du DMP.	- Contient les informations minimales requises pour identifier le déposant du projet et ses références.	- Fournit peu, voire aucune information, ce qui empêche d'identifier aisément le responsable du projet.
1 DESCRIPTION DES DONNÉES ET COLLECTE OU RÉUTILISATION DES DONNÉES EXISTANTES			
Prescriptions pour les chercheurs		Suffisamment traité Le DMP...	Insuffisamment traité Le DMP...
1a Comment les nouvelles données sont-elles collectées ou produites et/ou comment les données existantes sont-elles réutilisées ?	- Expliquez quelles méthodologies ou logiciels seront utilisés si de nouvelles données sont collectées ou produites. - Si la réutilisation des données existantes est contraignante, précisez-le. - Expliquez comment la provenance des données sera documentée. - Indiquez brièvement les raisons pour lesquelles la réutilisation de toute source de données existante antérieurement envisagée a été écartée.	- Indique clairement d'où proviennent les données existantes et comment les nouvelles données seront collectées ou produites. Il explique clairement les méthodes et les logiciels utilisés. - Explique comment avoir accès aux données existantes et les éventuelles contraintes liées à la réutilisation de ces données. - Explique clairement, le cas échéant, pourquoi de nouvelles données doivent être collectées au lieu de réutiliser les données existantes.	- Fournit peu, voire aucune information sur l'origine des données et sur celles qui seront collectées ou réutilisées. - Le cas échéant, ne fournit pas de justification suffisante pour la génération de nouvelles données.
1b Quelles données (par exemple le type, les formats, et les volumes) seront collectées ou produites ?	- Donnez des détails sur le type de données. Elles peuvent par exemple être numériques (bases de données, feuilles de calcul), textuelles (documents), iconographiques... Elles peuvent également se présenter sous forme audio, vidéo et/ou supports mixtes.	- Décrit ou énumère clairement les types de données qui seront générés (par exemple numériques, textuelles, audio ou vidéo) et leurs formats associés, y compris, si nécessaire, quel processus de conversion de ces données a été utilisé. - Explique le choix des formats et indique s'ils sont ouverts et	- Fournit peu, voire aucune information sur les types de données qui seront générés et ne donne pas de raison valable pour ce manque de détails (il n'indique par exemple pas pourquoi aucune donnée ne sera produite ou générée).

	- Donnez des détails sur le format des données : la manière dont les données sont codées pour le stockage, souvent reflétée par l'extension du nom de fichier (par exemple PDF, XLS, DOC, TXT, ou).	standard. Si un format propriétaire est utilisé, il en explique la raison.	- Ne fait qu'énumérer/décrire les types de données sans en préciser le format.
--	--	--	--

2 DOCUMENTATION ET QUALITÉ DES DONNÉES			
Prescriptions pour les chercheurs		Suffisamment traité Le DMP...	Insuffisamment traité Le DMP...
2a	<ul style="list-style-type: none"> - Indiquez quelles métadonnées seront fournies pour aider les autres à identifier et à découvrir les données. - Indiquez les normes/standards de métadonnées (par exemple DDI, TEI, EML, MARC, CMDI) qui seront utilisés. - Utilisez les normes/standards de métadonnées éventuellement prescrites par des communautés académiques. - Indiquez comment les données seront organisées pendant le projet en mentionnant, par exemple, les conventions, le contrôle de version et les structures de dossiers. Des données de recherche cohérentes et bien ordonnées seront plus faciles à trouver, à comprendre et à réutiliser. - Réfléchissez aux autres documents nécessaires pour que la réutilisation des données soit possible. Il peut s'agir d'informations sur la méthodologie utilisée pour collecter les données, d'informations analytiques et 	<ul style="list-style-type: none"> - Décrit clairement les métadonnées qui accompagneront les données, en se référant aux bonnes pratiques de la communauté scientifique (par exemple, utiliser les normes de métadonnées lorsqu'elles existent). - Indique de manière concise quelles ressources documentaires seront nécessaires pour permettre la réutilisation des données, en indiquant où les informations seront enregistrées (par exemple, une base de données avec des liens vers chaque élément, un fichier texte de type "readme", des en-têtes de fichiers, des livres de codes ou des carnets de laboratoire). - Indique comment les données seront organisées pendant le projet (en mentionnant par exemple les conventions de nommage, la stratégie de contrôle de version et les structures de dossiers). 	<ul style="list-style-type: none"> - Fournit peu, voire aucune information sur les métadonnées complémentaires. - Fournit peu, voire aucune information sur les ressources documentaires, sans fournir aucun détail ni aucune explication supplémentaire.

	<p>procédurales, de définitions des variables, d'unités de mesure, etc.</p> <p>- Réfléchissez à la manière dont ces informations seront saisies et à l'endroit où elles seront enregistrées (par exemple dans une base de données avec des liens renvoyant vers chaque élément, un fichier texte de type "readme", des en-têtes de fichiers, des livres de codes ou des carnets de laboratoire).</p>		
<p>2b</p> <p>Quelles mesures de contrôle de la qualité des données seront utilisées ?</p>	<p>Expliquez comment la collecte des données est contrôlée et documentée afin que cette collecte soit cohérente et qualitative. Il peut s'agir de processus tels que l'étalonnage, la répétition d'échantillons ou de mesures, la saisie de données standardisées, la validation de la saisie des données, l'évaluation des données par les pairs ou la représentation à l'aide de vocabulaires contrôlés.</p>	<p>- Définit clairement l'approche adoptée pour assurer et consigner par écrit le contrôle de la qualité exercé sur la collecte des données pendant la durée de vie du projet.</p>	<p>- Fournit peu, voire aucune information sur la façon dont la qualité des données est contrôlée et consignée pendant la durée de vie du projet.</p>

3 STOCKAGE ET SAUVEGARDE PENDANT LE PROCESSUS DE RECHERCHE			
Prescriptions pour les chercheurs		Suffisamment traité Le DMP...	Insuffisamment traité Le DMP...
<p>3a</p> <p>Comment les données et les métadonnées sont-elles stockées et sauvegardées durant la période de recherche ?</p>	<p>- Décrivez où les données sont stockées et sauvegardées pendant les activités de recherche et à quelle fréquence la sauvegarde sera effectuée. Il est recommandé d'avoir au moins une autre sauvegarde des données dans un autre endroit.</p> <p>- Privilégiez l'utilisation d'un dispositif de stockage fiable qui sauvegarde automatiquement, comme le font les services d'assistance informatique de l'établissement</p>	<p>- Décrit clairement (même si c'est fait de manière sommaire) :</p> <p>> L'endroit où les données et les sauvegardes sont stockées pendant les activités de recherche.</p> <p>> La fréquence à laquelle les sauvegardes sont effectuées.</p> <p>> Si les données sont stockées sur un dispositif fiable qui opère des sauvegardes automatiquement (à l'instar du dispositif fourni par l'institution d'origine). ou</p>	<p>- Fournit peu, voire aucune information sur la manière dont les données seront stockées et sauvegardées pendant le projet.</p>

	<p>d'origine. Il n'est pas recommandé de stocker les données sur des ordinateurs portables, des disques durs autonomes ou des dispositifs de stockage externes tels que des clés USB.</p>	<p>- Explique pourquoi le dispositif de stockage institutionnel ne sera pas utilisé (en précisant les données concernées) et indique quels emplacements et supports de stockage (supplémentaires) seront utilisés, ainsi que les procédures employées pour stocker et sauvegarder les données pendant le projet.</p>	
<p>3b</p> <p>Par quels moyens la sécurité des données et la protection des données sensibles seront assurées au cours de la période de recherche ?</p>	<p>- Expliquez comment les données seront récupérées en cas d'incident.</p> <p>- Communiquez la liste des personnes ayant accès aux données pendant la recherche et comment l'accès à ces données est contrôlé, en particulier dans le cadre de partenariats de collaboration.</p> <p>- Intégrez la protection des données, en particulier si vos données sont sensibles (par exemple, si elles contiennent des données personnelles, des informations politiquement sensibles ou des données soumises à la confidentialité commerciale). Décrivez les principaux risques et comment les gérer.</p> <p>- Expliquez quelles politiques institutionnelles de protection des données sont mises en œuvre.</p>	<p>- Explique clairement :</p> <p>> Comment les données seront récupérées en cas d'incident.</p> <p>> Quelles politiques institutionnelles et/ou nationales de protection des données sont appliquées et fournit un lien qui redirige l'utilisateur vers l'endroit où elles peuvent être consultées.</p> <p>> Qui aura accès aux données pendant la durée des travaux de recherche.</p> <p>- Décrit clairement les mesures de sécurité complémentaires (sécurité physique, sécurisation du réseau, des systèmes et des fichiers informatiques) qui seront prises pour garantir la sécurité des données stockées et transférées lorsque des données sensibles sont concernées (par exemple des données personnelles, des informations politiquement sensibles ou des données soumises à la confidentialité commerciale).</p>	<p>- Fournit peu, voire aucune information sur la manière dont les données seront récupérées en cas d'incident, sur les politiques institutionnelles de protection des données mises en œuvre et ne mentionne pas qui aura accès aux données pendant la durée des travaux de recherche.</p> <p>- Fournit peu, voire aucune information sur la protection des données et la gestion des risques lorsque des données sensibles sont concernées (par exemple des données personnelles, des informations politiquement sensibles ou des données soumises à la confidentialité commerciale).</p>

4 EXIGENCES LÉGALES ET ÉTHIQUES, CODES DE CONDUITE

Prescriptions pour les chercheurs	Suffisamment traité Le DMP...	Insuffisamment traité Le DMP...
<p>4a</p> <p>Si des données personnelles sont traitées, comment le respect de la législation et de la sécurité sur les données personnelles sera-t-il assuré ?</p>	<p>- Assurez-vous que la législation sur la protection des données (comme le RGPD) soit respectée lors du traitement des données personnelles :</p> <p>> Obtenez un consentement éclairé pour la préservation et/ou le partage des données personnelles.</p> <p>> Prévoyez de rendre les données personnelles anonymes pour la préservation et/ou le partage de ces données (celles qui sont anonymes ne sont plus considérées comme des données personnelles).</p> <p>> Envisagez la création de pseudonymes pour les données personnelles (contrairement à l'anonymat, les pseudonymes sont réversibles).</p> <p>> Envisagez d'utiliser le système de cryptage, considéré comme une forme de pseudonymisation (la clé de cryptage ne doit pas être stockée avec les données, mais par un tiers de confiance).</p> <p>> Précisez si une procédure d'accès géré pour les utilisateurs autorisés à accéder aux données personnelles existe.</p>	<p>- Fournit peu, voire aucune information démontrant que des données personnelles, le cas échéant, seront gérées en conformité avec la législation applicable.</p> <p>- Indique clairement si des données personnelles seront collectées/utilisées dans le cadre du projet et, le cas échéant, comment le respect de la législation applicable sera assuré (par exemple en obtenant le consentement éclairé, en envisageant le cryptage, l'anonymisation ou la pseudonymisation).</p> <p>- Décrit la procédure permettant de gérer l'accès aux seuls utilisateurs autorisés.</p>
<p>4b</p> <p>Comment les autres questions juridiques, telles que les droits de propriété, intellectuelle ou autres, seront-elles</p>	<p>- Expliquez qui sera le propriétaire des données, donc qui en contrôlera l'accès :</p> <p>> Quelles conditions d'accès s'appliqueront aux données ? Les données seront-elles librement</p>	<p>- Explique clairement, le cas échéant :</p> <p>> Qui aura le droit de contrôler l'accès aux données en indiquant celles qui sont concernées.</p> <p>> Les conditions d'accès et les licences de réutilisation</p> <p>- N'aborde pas les questions juridiques (ou seulement pour un sous-ensemble de données), sans motiver pareille absence.</p> <p>- N'aborde pas les questions relatives aux droits de contrôle d'accès aux données dans le cas d'un projet</p>

<p>traitées ? Quelle est la législation applicable ?</p>	<p>accessibles, ou y'aura-t-il des restrictions d'accès ? Si oui, lesquelles ? Pensez à l'éventualité d'utiliser des licences d'accès et de réutilisation des données.</p> <p>> Faites en sorte que ces questions soient intégrées dans l'accord de consortium en cas de projets à partenaires et propriétaires de données multiples.</p> <p>> Si des droits de propriété intellectuelle (comme la Directive sur les bases de données, en tant que droits <i>sui generis</i>) sont affectés, faites-en part. Dans l'affirmative, expliquez lesquels et comment ils seront traités.</p> <p>- Indiquez s'il existe des restrictions à la réutilisation des données de tiers.</p>	<p>qui s'appliqueront aux données.</p> <p>- Explique clairement, le cas échéant, comment les droits de propriété intellectuelle seront gérés.</p> <p>- Explique comment ces questions sont traitées dans l'accord de consortium pour les projets à partenaires et propriétaires de données multiples.</p> <p>- A défaut, il est explicitement indiqué que les données ne sont soumises à aucune restriction de ce type.</p> <p>- Indique, le cas échéant, s'il existe des restrictions sur la réutilisation des données de tiers.</p>	<p>comprenant plusieurs partenaires, sans motiver pareille absence.</p>
<p>4c</p> <p>Quelles sont les questions éthiques et les codes de conduite existants, et comment seront-ils pris en compte ?</p>	<p>- Étudiez si des questions éthiques peuvent affecter la manière dont les données sont stockées et transférées, qui peut les voir ou les utiliser, et combien de temps elles sont conservées. Montrez que vous avez connaissance de ces aspects et de leur intégration dans la planification concernée.</p> <p>- Suivez les réglementations nationales et internationales et les directives éthiques institutionnelles, et vérifiez si la collecte de données dans le cadre du projet de recherche doit faire l'objet d'une évaluation éthique (par un comité d'éthique par exemple).</p>	<p>- Fournit des détails sur les questions éthiques prises en considération qui peuvent affecter le stockage, le transfert, l'utilisation, le partage et/ou la conservation des données, et démontre que des mesures adéquates sont en place pour gérer les règles d'éthiques.</p> <p>- Mentionne, le cas échéant, si une évaluation éthique est en cours. Si une approbation éthique a été obtenue, il fait mention du comité et des pièces qui s'appliquent.</p> <p>- Renvoie aux directives éthiques et/ou aux réglementations qui s'appliquent ou explique de manière concise pourquoi les questions éthiques n'ont pas été prises en considération.</p>	<p>- Fournit peu, voire aucune information sur la prise en considération des implications éthiques et des réglementations et sans motiver pareille absence.</p>

5 PARTAGE DES DONNÉES ET CONSERVATION À LONG TERME

Prescriptions pour les chercheurs	Suffisamment traité Le DMP...	Insuffisamment traité Le DMP...	
<p>5a</p> <p>Quand les données seront-elles partagées et de quelle façon ? Des restrictions ou des embargos peuvent-ils subsister ?</p>	<ul style="list-style-type: none"> - Expliquez comment les données pourront être accessibles et partagées (par exemple, en les plaçant dans un entrepôt de données fiable, en les indexant dans un catalogue, en utilisant un service de données sécurisé, en traitant directement des demandes de données ou pas le biais d'un autre mécanisme). - Décrivez le plan de conservation des données et donnez des informations sur leur durée de conservation. - Communiquez la date de mise à disposition des données. Indiquez le délai de publication prévu. Indiquez si l'utilisation exclusive des données sera revendiquée et si oui, pourquoi et pour combien de temps. Précisez si le partage des données sera reporté ou restreint, par exemple pour publier, protéger la propriété intellectuelle ou obtenir des brevets. - Indiquez qui sera en mesure d'utiliser les données. S'il est nécessaire de restreindre l'accès à certaines communautés ou d'appliquer un accord de partage des données, expliquez comment et pourquoi. Précisez quelles mesures seront prises pour lever ou limiter les restrictions. 	<ul style="list-style-type: none"> - Décrit clairement comment les données et/ou les métadonnées seront rendues accessibles et partagées. - Précise quand les données seront partagées et sous quelle licence. - Mentionne le nom de l'entrepôt, du catalogue ou du registre qui partage les données. - Apporte des informations sur la durée de conservation des données et indique le début de leur accessibilité. - Explique clairement, le cas échéant, pourquoi le partage des données est limité ou impossible, qui peut y accéder et sous quelles conditions (par exemple, uniquement les membres de certaines communautés ou via un accord de partage). - Explique, si possible, les actions qui seront prises pour lever ou minimiser les restrictions du partage des données. 	<ul style="list-style-type: none"> - Fournit peu, voire aucune information sur le début de l'accessibilité des données et sur la stratégie utilisée pour rendre cela possible, ou alors l'explication n'est pas adéquate ou techniquement viable.
<p>5b</p> <p>Comment les données à préserver seront-elles désignées, et où seront-elles conservées à long terme (dans un entrepôt de données ou des archives par exemple) ?</p>	<ul style="list-style-type: none"> - Indiquez quelles données doivent être conservées ou détruites à des fins contractuelles, légales ou réglementaires. - Indiquez comment la conservation des données est décidée. Indiquez quelles 	<ul style="list-style-type: none"> - Fournit des détails sur les données collectées ou créées dans le cadre du projet qui seront conservées à long terme et en indique clairement la durée. Ces indications doivent respecter les politiques et/ou à la législation du bailleur de fonds, de l'institution ou du 	<ul style="list-style-type: none"> - Fournit peu, voire aucune information sur les dispositions prises pour la préservation des données.

	<p>données sont à conserver à long terme.</p> <ul style="list-style-type: none"> - Expliquez quelles sont les éventuelles utilisations (et/ou utilisateurs) des données dans le cadre de la recherche. - Indiquez où les données seront déposées. Si aucun dépôt établi n'est proposé, démontrez dans le DMP que les données peuvent être conservées efficacement au-delà de la durée de la subvention. <p>Il est recommandé de démontrer que les politiques et procédures des dépôts (y compris les normes de métadonnées et les coûts associés) ont été vérifiées.</p>	<p>pays, ou les normes communautaires.</p> <ul style="list-style-type: none"> - Détaille les (versions des) données et les documents annexes qui seront conservés ou détruits, et en explique la raison (par exemple, à des fins contractuelles, légales ou réglementaires). - Expose la manière dont la sélection est faite, et sur l'intérêt éventuel d'une réutilisation (ou non). - Décrit comment dont les données, les documents annexes et toute autre technologie requise, comme des copies de logiciels dans des versions spécifiques, seront archivées à long terme. - Explique comment les données seront gérées de manière durable au-delà de la durée de la subvention. - Fournit le nom de l'archive ou de l'entrepôt fiable (ou la manière de conserver et de préserver les données) qui sera exploité pour permettre la réutilisation des données. 	
<p>5c</p> <p>Quelles méthodes ou outils informatiques sont nécessaires pour accéder aux données et les utiliser ?</p>	<ul style="list-style-type: none"> - Indiquez si les utilisateurs éventuels ont besoin d'outils spécifiques pour accéder aux données et les (ré)utiliser. Tenez compte de la pérennité des logiciels nécessaires pour y accéder. - Indiquez si les données seront partagées par le biais de demandes de dépôt traitées directement, ou si un autre système sera utilisé. 	<ul style="list-style-type: none"> - Indique clairement les outils ou logiciels spécifiques (par exemple les scripts, codes ou algorithmes spécifiques développés pendant le projet, la version du logiciel) utiles pour les utilisateurs éventuels afin d'accéder aux données, les interpréter et les (ré)utiliser. - Fournit des informations, le cas échéant, sur tout protocole d'accès aux données (par exemple, si une authentification est nécessaire ou s'il existe une procédure de demande d'accès aux données). 	<ul style="list-style-type: none"> - Fournit peu, voire aucune information sur les logiciels développés au cours du projet et nécessaires pour accéder aux données et les interpréter, sur la manière dont ils seront mis à disposition, sans motiver pareille absence.
<p>5d</p> <p>Comment sera assurée l'attribution d'un identifiant unique et pérenne (tel qu'un identifiant d'objet numérique, ou DOI) à chaque jeu de données ?</p>	<ul style="list-style-type: none"> - Expliquez comment les données pourraient être réutilisées dans d'autres contextes. Des identifiants pérennes (PID) doivent être appliqués afin que les données puissent être localisées et référencées de manière fiable et efficace. Les PID permettent 	<ul style="list-style-type: none"> - Précise comment les données peuvent être réutilisées dans d'autres contextes. - Indique clairement si des identifiants pérennes sont attribués pour tous les jeux de données individuels, à des sous-ensembles, à des collections et si oui, lesquels. Si les PID ne sont pas fournis, il en explique la raison. 	<ul style="list-style-type: none"> - Ne fait aucune mention des identifiants pérennes, sans motiver pareille absence. - Ne fournit pas d'informations claires sur le type de PID qui sera attribué aux données et n'indique pas si des jeux de données individuels et/ou des collections

	<p>également de suivre les citations et la réutilisation.</p> <ul style="list-style-type: none"> - Indiquez si un PID pour les données sera envisagé. En général, un identifiant pérenne est fourni par un entrepôt digne de confiance et de conservation à long terme. 	<ul style="list-style-type: none"> - Présente clairement l'approche utilisée. Le choix des identifiants est justifié et fait référence à des normes internationales. 	entières seront dotés de PID.
--	--	---	-------------------------------

6 RESPONSABILITÉS ET RESSOURCES EN MATIÈRE DE GESTION DES DONNÉES

Prescriptions pour les chercheurs	Suffisamment traité Le DMP...	Insuffisamment traité Le DMP...
<p>6a</p> <p>Qui (le rôle, la fonction et l'institution, etc.) sera responsable de la gestion des données ?</p>	<ul style="list-style-type: none"> - Décrit clairement les rôles et les responsabilités en matière de gestion et d'intendance des données, ou <i>stewardship</i>, (comme la saisie et la qualité des données leur stockage, leur archivage et leur partage, ainsi que la production de métadonnées) en désignant si possible la ou les personnes responsables. - Indique clairement le responsable de la mise en œuvre et des ajustements quotidiens du DMP. - Dans le cas de projets de collaboration, explique comment coordonner les responsabilités de gestion des données entre les partenaires. 	<ul style="list-style-type: none"> - N'aborde pas la question de la responsabilité des activités de gestion des données et/ou n'indique pas le responsable de la mise en œuvre et des ajustements quotidiens du DMP. - Dans le cas d'un projet de collaboration, ne fournit aucune information sur la manière dont les responsabilités en matière de gestion des données seront coordonnées entre les partenaires.
<p>6b</p> <p>Quelles ressources (par exemple financières et temporelles) seront consacrées à la gestion des données et à la garantie de leur « FAIRitude » (Faciles à trouver, Accessibles, Interopérables et Réutilisables) ?</p>	<ul style="list-style-type: none"> - Expliquez comment les ressources nécessaires (comme le temps) pour préparer les données en vue de leur partage/préservation (conservation des données) ont été chiffrées. - Examinez et justifiez de manière concise toutes les ressources nécessaires à la livraison des données. Celles-ci peuvent inclure les coûts de stockage, le matériel, le temps du 	<ul style="list-style-type: none"> - Fournit peu, voire aucune information sur les ressources requises pour la gestion des données et la garantie de leur « FAIRitude » (les ressources ne sont par exemple pas énumérées ou leur coût est inapproprié), et/ou n'indique pas comment les coûts seront couverts.

	<p>personnel, les coûts de préparation des données pour le dépôt et les frais d'entrepôt.</p> <p>- Indiquez si des ressources supplémentaires seront nécessaires pour préparer les données en vue de leur dépôt ou pour couvrir les frais d'entrepôts de données. Dans l'affirmative, indiquez le montant nécessaire et comment ces coûts seront couverts.</p>	<p>aucune ressource supplémentaire n'est nécessaire.</p>	
--	--	--	--

Notes et références

- I. Le terme « organismes de recherche » désigne les organismes qui effectuent des recherches, les universités et les instituts de recherche.
- II. Pour plus d'informations sur la manière dont les principes FAIR sont traduits dans les exigences minimales et les critères de sélection, veuillez consulter l'annexe
- III. Le concept a été présenté lors d'une manifestation publique le 30 janvier 2018, et deux consultations ont été organisées en avril 2018 et en août/septembre 2018 en vue de la première version de ce guide. Une autre consultation a été organisée entre juillet et septembre 2020 concernant la rubrique devant la compléter dans la présente version et portant sur l'évaluation des DMP.
- IV. Pour des éléments de procédure pour la mise en œuvre des DMP, voir les informations du groupe de travail de l'Alliance intitulé *RDA DMP Common Standards Working Group* : <https://www.rd-alliance.org/groups/dmp-common-standards-wg>
- V. Digital Object Identifier
- VI. Uniform Resource Name
- VII. Archival Resource Key
- VIII. Open Researcher and Contributor ID
- IX. International Standard Name Identifier
- X. Digital Author Identifier

Annexe : Compatibilité avec les principes FAIR

		Exigence	Critère de sélection
Pour être Facile à trouver			
F1	Les données et les métadonnées sont identifiées par un identifiant global unique et pérenne	E 5d	Critère 1
F2	Les métadonnées décrivant les données sont riches	E 2a	Critère 2
F3	Les données et les métadonnées sont enregistrées et indexées dans un dispositif permettant de les rechercher	E 5d	Critère 1, Critère 2
F4	Les métadonnées spécifient l'identifiant de la donnée		Critère 2

Pour être Accessible			
A1	Les données et les métadonnées sont accessibles par leur identifiant via un protocole de communication standardisé	E 5c	Critère 1, Critère 2
A1.1	Le protocole utilisé est ouvert, libre et peut être implémenté de manière universelle	E 5c	Critère 2
A1.2	Le protocole utilisé permet l'authentification et l'autorisation si besoin	E 4b, E 5a, E 5c	Critère 3
A2	Les métadonnées sont accessibles même quand les données ne le sont plus	E 4c, E 5a, E 5d	Critère 2c

Pour être Interopérable			
I1	Les données et les métadonnées utilisent un langage formel, accessible, partagé et largement applicable pour la représentation des connaissances	E 1b, E 2a	Critère 2d
I2	Les données et les métadonnées utilisent des vocabulaires qui respectent les principes FAIR	E 2a, E 2b	Critère 2
I3	Les données et les métadonnées incluent des liens vers d'autres (méta)données	E 2a, E 5a, E 5c	Critère 2b

Pour être Réutilisable			
R1	Les données et les métadonnées ont des attributs multiples et pertinents	E 2a, E 2b	Critère 2
R1.1	Les données et les métadonnées sont mises à disposition selon une licence explicite et accessible	E 4b, E 5a	Critère 3d
R1.2	Les données et les métadonnées sont associées à leur provenance	E 1a, E 1b, E 2b	Critère 1c, Critère 2, Critère 3b, Critère 4a
R 1.3	Les données et les métadonnées correspondent aux standards des communautés indiquées	E 1b, E 2a	Critère 2d