

Université Toulouse - Jean Jaurès: ANR - Modèle de PGD (français)

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 ?

Recommandations:

- Expliquer quelles méthodologies ou quels logiciels seront utilisés si de nouvelles données sont recueillies ou produites.
- Énoncer les éventuelles restrictions 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.

Exemple de réponse:

Données quantitatives issues d'un questionnaire :

Un questionnaire en ligne sera déployé pour lequel nous prévoyons 300 répondants. Ce questionnaire sera réalisé avec le logiciel Lime Survey et les données recueillies seront ensuite analysées avec le logiciel SPSS.

Données qualitatives issues d'entretiens :

L'équipe prévoit de mener 25 entretiens semi-directifs. Chaque entretien sera enregistré avec un enregistreur audio et retranscrit dans un second temps.

Données quantitatives issues de la base de données « X » :

L'équipe réutilisera les données issues de la base de données « X » qui a été développée par « X » et qui est accessible sous la licence CC-BY-NC.

Recommandations:

Vous devez mentionner dans cette section les méthodes et les protocoles utilisées pour la collecte de vos données concernant :

- **Les données générées dans le cadre du projet de recherche :** vous devez indiquer la méthode et les protocoles selon lesquels elles ont été produites.
- **Les données préexistantes réutilisées dans le cadre du projet de recherche :** vous devez indiquer la source des données réutilisées ainsi que leurs droits de réutilisation.

RGPD - La réutilisation des données personnelles issues d'une « recherche » est possible ponctuellement, à condition que le responsable scientifique, ou le nouveau, respecte ces dispositions :

- les participants ont été informés lors du recueil de données originel
- la finalité est précise et porte sur une question spécifique et ponctuelle
- la durée de recherche est limitée et connue

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

Recommandations:

- Donner 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").
- Donner 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).

Recommandations:

Cette section permet de préciser la nature des données qui seront générées ou réutilisées dans le cadre du projet. Nous vous recommandons de privilégier des formats ouverts (non propriétaires) et pérennes. Dans ce cadre, le CINES met à votre disposition un outil en ligne qui vous permet de vérifier que les formats choisis sont pérennes et archivables : [FACILE](#).

Il peut être assez difficile d'indiquer la volumétrie des données en début de projet. Dans un premier temps, vous pouvez indiquer une estimation qui pourra être affinée à chaque nouvelle version du PGD.

Exemple de réponse:

Données quantitatives issues d'un questionnaire :

Les données recueillies seront exportées dans un fichier texte au format .csv pour la conservation et dans un fichier SPSS Statistics au format .sav pour l'analyse statistique.

Ces fichiers devraient représenter un volume de X Go.

Les données analysées seront exportées dans un fichier texte au format .csv et dans un fichier SPSS Statistics au format .sav .

Ces fichiers devraient représenter un volume de X Go.

Données qualitatives issues d'entretiens :

Les entretiens seront enregistrés sous le format audio MPEG-4. Ils feront ensuite l'objet d'une retranscription dans des fichiers au format .txt.

Ces fichiers devraient représenter un volume de X Go.

Données quantitatives issues de la base de données « X » :

Les données réutilisées seront téléchargées au format .csv.

Le volume des données téléchargées est estimé à X Go.

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 ?

Recommandations:

- Indiquer quelles métadonnées seront fournies pour aider à la recherche et à l'identification des données.
- Indiquer quels standards de métadonnées seront utilisés (par exemple DDI, TEI, EML, MARC, CMDI).
- Utiliser les standards de métadonnées des communautés scientifiques lorsque ceux-ci existent.
- Indiquer comment les données seront organisées au cours du projet, en mentionnant par exemple les conventions de nommage, le contrôle de version et les structures des dossiers. Des données bien classées et gérées de façon cohérente seront plus faciles à retrouver, à comprendre et à réutiliser.
- Penser 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.
- Tenir 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.

Recommandations:

Dans cette section, vous devez expliciter comment vos données seront structurées et documentées. L'objectif est d'assurer une bonne compréhension des données afin de permettre une réutilisation ultérieure.

Vous devez :

- **Indiquer les standards de métadonnées utilisés :**

Nous vous recommandons d'utiliser des standards de métadonnées reconnus. Il existe des standards génériques (Dublin Core, DataCite Metadata Schema) ainsi que des standards utilisés dans des domaines disciplinaires (Data Documentation Initiative (DDI) pour les données d'enquête, Ecological Metadata Language (EML) pour les données en écologie, ...). Vous pouvez consulter des répertoires qui proposent des standards :

- Digital Curation Center (DCC) : <http://www.dcc.ac.uk/resources/metadata-standards>
- Research Data Alliance (RDA) : <http://rd-alliance.github.io/metadata-directory/standards/>
- Fairsharing : <https://fairsharing.org/standards/>

Si vous ne pouvez pas utiliser de standards reconnus, il faut préciser les métadonnées que vous utilisez. Il est possible de créer un fichier de métadonnées (lien DataCite Metadata Generator).

- **Préciser si vous associez de la documentation à vos jeux de données :**

Nous vous recommandons de créer à minima un fichier « Read me » qui accompagnera vos jeux de données et qui permettra la compréhension et une bonne interprétation de ceux-ci. Un exemple de gabarit est disponible sur Doranum (lien – docs à télécharger dans vérifier ses données).

Vous pouvez également accompagner vos données de l'ensemble la documentation associée : questionnaire d'enquête, dictionnaire des variables, protocole, ...

- **Mentionner le mode d'organisation de vos fichiers et les règles de nommage utilisées :**

Si vous n'avez pas adopté de mode d'organisation de vos fichiers, nous vous recommandons de vous baser sur les bonnes pratiques proposées par [Doranum](#).

Vous pouvez également contacter mdrburo@univ-tlse2.fr pour un accompagnement dans l'élaboration d'une charte de nommage (modèle et recommandations).

Exemple de réponse:

Selon la nature et la diversité des jeux de données générées et/ou collectées dans votre projet de recherche, vous pouvez apporter :

Une réponse propre à chaque jeu de données :

- *Données quantitatives issues de questionnaires :*

Les fichiers rassemblant les données recueillies générées à partir du logiciel Lime Survey seront documentées selon les spécifications « Data Documentation Initiative » (DDI) qui est un standard pour décrire les données d'enquête quantitatives en sciences humaines et sociales. Une documentation sera produite, elle comprendra le protocole d'enquête et la description des variables.

- *Données qualitatives issues d'entretiens :*

Chaque retranscription d'entretien sera accompagnée du protocole d'enquête, du contexte de passation de l'entretien ainsi qu'une présentation de la personne interrogée.

Une réponse commune à tous les jeux de données :

Chaque fichier de données sera accompagné au minimum d'un fichier de type « Read Me » indiquant le contexte dans lequel les données ont été collectées, les métadonnées ou tout autre information permettant une réutilisation des données. La convention de nommage des fichiers et dossiers sera celle [recommandée sur le site de DORANum](#) (INIST – CNRS).

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

Exemple de réponse:

Les bonnes pratiques de terrain et de laboratoire seront suivies pour le contrôle de la qualité des données que ce soit lors de leur collecte ou de leur analyse.

Recommandations:

- 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.

Recommandations:

Vous devez mentionner dans cette section la démarche de contrôle qualité que vous allez adopter dans le cadre de votre projet de recherche. L'objectif est de présenter les processus mis en place quant à la rigueur des méthodes utilisées et la qualité des données.

Le [Réseau Qualité en Recherche du CNRS](#) a produit un guide sur la « traçabilité des activités de recherche et gestion des connaissances » ainsi qu'un ensemble de documents relatifs à la qualité dans les laboratoires.

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 ?

Recommandations:

- 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.

Recommandations:

STOCKAGE – Il s'agit de décrire les supports de stockage ainsi que les procédures de sauvegarde mises en place pour éviter tout risque de perte ou de faille de sécurité. Vous devez ainsi indiquer la localisation de ces supports ainsi que la fréquence de sauvegarde.

Nous vous recommandons d'utiliser des outils institutionnels Scout, MyCore, Sharedocs, Humanum-Box, Filesender, ...) et de ne pas stocker vos données seulement sur vos ordinateurs ou des espaces de stockage externe (disques durs, clés USB).

RGPD - Les données doivent être stockées sur un serveur institutionnel situé dans un datacenter sécurisé qui assure automatiquement les sauvegardes tous les jours. L'accès aux données doit être nominatif.

Pour un accompagnement sur les solutions de stockage : mdrburo@univ-tlse2.fr

Exemple de réponse:

Les données de ce projet sont stockées sur le serveur institutionnel sécurisé, Scout, dans le datacenter de l'UFTMiP. Les sauvegardes sont assurées tous les jours.

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

Recommandations:

RGPD - Si vous travaillez avec des données personnelles/sensibles, vous devez prendre contact avec rgpd-recherche@univ-tlse2.fr pour déclarer votre traitement.

Pour vous documenter sur le RGPD : cf. [lien vers l'ENT](http://ent-utm.univ-tlse2.fr/profils/accueil-personnel/recherche/rgpd-recherche/) (<http://ent-utm.univ-tlse2.fr/profils/accueil-personnel/recherche/rgpd-recherche/>)

Recommandations:

- Expliquer 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.
- Tenir 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 est mise en œuvre.

Exemple de réponse:

RGPD - Une analyse d'impact (ou une fiche de traitement) a été rédigée et visée par le DPO et le responsable sécurité de mon établissement. Elle encadre la conformité de mon projet au RGPD et un traitement respectueux de la vie privée des participant.e.s.

RGPD – Mon projet comprend la collecte de données sensibles. Elles seront pseudonymisées : une table de correspondance sera créée pour isoler les données directement identifiantes. Cette table sera chiffrée et accessible aux seuls responsables scientifiques du projet.

RGPD - Tous les membres du projet prendront connaissance de l'engagement de confidentialité afférent aux données personnelles.

Traitement anonyme – Le RGPD ne s'applique pas car aucune donnée identifiante n'est récupérée dans mon projet de recherche. Il est impossible, même pour les responsables scientifiques de ce projet, de savoir quels volontaires ont participé à cette étude. Il est également impossible de retrouver l'identité d'un participant par recoupement d'informations.

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 ?

Recommandations:

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).
- Envisager 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).
- Expliquer si une procédure d'accès spécifique a été mise en place pour les utilisateurs autorisés à accéder aux données personnelles.

Exemple de réponse:

Mon projet de recherche est conforme au RGPD et a été enregistré sous le numéro XXXXX dans le registre de conformité de mon établissement par le DPO.

Cette recherche est soumise à la Loi Jardé (Recherche Impliquant la Personne Humaine). Elle est menée conformément à la méthodologie de référence MR-00x homologuée par la Commission Nationale de l'Informatique et Libertés (Cnil) le xx/xx/20xx et à laquelle l'UT2J/CNRS s'est engagé(e) à se conformer (récépissé n° xxxx v 0 du xx xxxx 20xx).

Recommandations:

Nous vous accompagnons dans votre conformité au RGPD : rgpd-recherche@univ-tlse2.fr

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 ?

Recommandations:

- Expliquer qui sera le propriétaire des données, c'est-à-dire qui aura le droit d'en contrôler l'accès :
 - Expliquer quelles conditions d'accès s'appliqueront aux données. Les données seront-elles librement accessibles, ou des restrictions seront-elles appliquées ? Si oui, lesquelles ? Envisager l'utilisation de licences concernant l'accès et la réutilisation des données.
 - S'assurer 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.
- Indiquer si les droits de propriété intellectuelle (par exemple la directive bases de données, droits *sui generis*) sont affectés. Dans l'affirmative, expliquer lesquels et comment cela sera traité.
- Indiquer s'il y a des restrictions sur la réutilisation des données fournies par des tiers.

Exemple de réponse:

Données produites dans le cadre du projet de recherche :

Les partenaires du projet seront conjointement propriétaires des données tel que mentionné dans l'accord de consortium.

Pendant le projet de recherche, les différents partenaires pourront accéder à l'ensemble des données.

A l'issue du projet, les données qui feront l'objet d'une diffusion en libre accès seront sous licence CC-BY.

Données préexistantes réutilisées dans le projet :

Les données issues de la base de données « X » sont la propriété de « X ». La base de données est consultable sur [lien] sous la licence « X ».

Recommandations:

Vous devez mentionner dans cette section les droits de propriété intellectuelle pour :

• **Les données générées dans le cadre du projet de recherche :**

La propriété des données est en principe précisée dans l'accord de consortium. Vous devez donc reprendre les éléments qui ont été définis dans cet accord en précisant à qui appartiennent les données collectées, produites ainsi que les droits régissant les bases de données.

• **Les données préexistantes et que vous réutilisez dans le projet de recherche :**

Vous devez mentionner la source de ces données ainsi que les droits de propriété intellectuelle à considérer (données libres de droit ou données sous licence ou protégées par des droits spécifiques).

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

Recommandations:

- Déterminer si les questions d'éthique auront une incidence sur la façon dont les données seront stockées et transférées, qui pourra

les voir ou les utiliser et quelles durées de conservation leur seront appliquées. Démontrer que ces aspects sont bien pris en compte et planifiés.

- Adopter les codes de conduite nationaux et internationaux et le code d'éthique institutionnel et vérifier 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.

Recommandations:

Pour les données soulevant des questions éthiques, vous devez indiquer comment vous allez traiter ces questions et si vous avez sollicité l'approbation d'un comité éthique.

Exemple de réponse:

Le protocole de recherche a été évalué et autorisé par le comité d'éthique « X » en date du JJ/MM/AAAA.

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 ?

Recommandations:

- Expliquer comment les données pourront être retrouvées et partagées (par exemple, par le dépôt dans un entrepôt de données de confiance, l'indexation dans un catalogue, par l'utilisation d'un service de données sécurisé, par le traitement direct des demandes de données, ou l'utilisation de tout autre mécanisme).
- Définir le plan de préservation des données et fournir l'information sur la durée d'archivage pérenne des données.
- Expliquer à quel moment les données seront rendues disponibles. Indiquer les délais de publication prévus. Expliquer si une utilisation exclusive des données est revendiquée et, dans l'affirmative, pour quelle raison et pour combien de temps. Indiquer si le partage des données sera différé ou limité, par exemple pour des raisons de publication, pour protéger la propriété intellectuelle ou le dépôt de brevets.
- Indiquer 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, expliquer comment et pourquoi. Expliquer les mesures qui seront prises pour dépasser ou minimiser ces restrictions.

Exemple de réponse:

Selon la nature et la diversité des jeux de données générés et/ou collectés dans votre projet de recherche, vous pouvez apporter :

Une réponse propre à chaque jeu de données :

- *Données quantitatives issues de questionnaires* : Les données issues de l'analyse des questionnaires seront regroupées dans un jeu de données qui sera diffusé à l'issue du projet sur l'entrepôt de données Zenodo sous la licence « Licence Ouverte 2.0 ».

Une réponse commune à tous les jeux de données :

Pendant la durée du projet, les données seront accessibles aux personnes extérieures au projet uniquement sur demande auprès du porteur de projet, le temps que les membres du projet publient les résultats de leurs analyses.

A l'issue du projet, les jeux de données anonymisées seront déposés sur l'entrepôt de données Zenodo sous la licence « Licence Ouverte 2.0 ».

RGPD & Opendata - Le CHU nous fournit un fichier de données de santé de nos participants. Au vu de la sensibilité de ces données, il est impossible de prévoir une diffusion en libre accès.

Recommandations:

Il est à noter que cet aspect est souvent difficile à traiter en début de projet. Vous pourrez affiner votre réponse dans les différentes versions du PGD. En début de projet, vous devez, à minima, indiquer que vous prenez en compte ces questions notamment si vous devez répondre à des obligations de diffusion en libre accès de la part de l'agence de financement et que votre stratégie de diffusion sera précisée au fur et à mesure de l'avancée du projet.

A terme, vous devez :

- **Indiquer les jeux de données qui seront partagés en précisant :**
 - *Où seront diffusées les données* : Vous pouvez déposer vos données dans un entrepôt de données. Nous attirons votre attention sur le fait de déposer vos données sur un entrepôt de confiance, vous trouverez une liste d'entrepôts certifiés sur le site [CoreTrustSeal](https://www.data.gouv.fr/fr/entrepots/). Vous pouvez également consulter le répertoire re3data.org qui recense les entrepôts de données par domaine disciplinaire. Nous vous recommandons de déposer vos données dans l'entrepôt de données « [Zenodo](https://zenodo.org) » qui est l'entrepôt de données préconisé dans le cadre des appels à projet H2020.
 - *Quand seront diffusées les données* : Vous devez indiquer si vous diffuserez vos données avec un embargo.
 - *Sous quelle licence seront diffusées les données* : Vous devez indiquer sous quelle licence vous diffuserez vos données. Vous trouverez les différentes licences que vous pouvez utiliser sur le site de [data.gouv](https://www.data.gouv.fr/fr/licences) (<https://www.data.gouv.fr/fr/licences>)
- **Justifier les raisons pour lesquelles certains jeux de données ne seront pas partagés** : données sensibles ou confidentielles (cf. RGPD & Opendata), données contractuelles, ...

RGPD & Opendata - Dans les cas où la diffusion des données à caractère personnel est autorisée, la publication des données personnelles en « *open data* » doit se faire dans le respect du RGPD. Elle ne peut en aucun cas permettre l'identification directe ou indirecte des personnes se prêtant à la recherche. Vous pouvez prendre contact avec rgpd-recherche@univ-tlse2.fr pour vous accompagner dans cette démarche.

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) ?

Recommandations:

- Indiquer quelles données ne doivent pas être divulguées ou doivent être détruites pour des raisons contractuelles, légales, ou réglementaires.
- Indiquer 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.
- Indiquer où les données seront déposées. Si aucun entrepôt reconnu n'est proposé, démontrer dans le plan de gestion des données 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 standards de métadonnées, et coûts mis en œuvre) ont été vérifiés.

Exemple de réponse:

L'ensemble des jeux de données anonymisés sera conservé pendant 15 ans sur les serveurs de l'université. Les données nominatives seront détruites à l'issue du projet.

Recommandations:

SÉLECTION POUR L'ARCHIVAGE - Vous devrez indiquer dans cette section les critères que vous mettrez en œuvre pour sélectionner les données que vous conserverez.

DURÉE DE L'ARCHIVAGE - Nous attirons votre attention sur le fait que la durée de conservation des données peut être contrainte :

- **Comités d'éthique (CER, CPP)** - Les comités d'éthique recommande de conserver les données de recherche pendant 15 ans.
- **RGPD** - La durée de conservation recommandée des données personnelles est de 2 ans après la dernière publication de la recherche. Après cette période, il est possible d'anonymiser les données en vue d'un archivage ou d'effacer les données identifiantes (table de correspondance). Cette méthode d'archivage respecte le cadre légal des méthodologies de référence de la CNIL.

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

Recommandations:

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

Recommandations:

Vous devez indiquer dans cette section si des logiciels spécifiques sont nécessaires pour lire les données en les précisant nominativement.

Exemple de réponse:

L'accès aux données nécessite l'utilisation du logiciel SPSS ou tout autre logiciel de type Stat Transfert qui permettra la conversion des données dans un autre format.

5d. Comment l'attribution d'un identifiant unique et pérenne (comme le DOI) sera-t-elle assurée pour chaque jeu de données ?

Recommandations:

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

Exemple de réponse:

Les données seront déposées sur Zenodo qui attribue automatiquement un DOI à chaque jeu de données.

Recommandations:

Nous vous recommandons de déposer vos données dans un entrepôt attribuant automatiquement un DOI (critère rentrant dans la certification des entrepôts de données de confiance). Pour plus d'information, se référer à la section 5.A.

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'est-à-dire le gestionnaire des données) ?

Recommandations:

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

Recommandations:

Dans cette section, vous devez indiquer l'ensemble des membres du projet qui auront une responsabilité dans la gestion des données (collecte, traitement, stockage, rédaction du PGD, ...).

Selon l'organisation des projets, les responsabilités peuvent être définies par work-package, localisation géographique (ex : terrain de recherche),...

Exemple de réponse:

Chaque chercheur collectant des données sera responsable :

- Du stockage des données collectées sur le terrain sur le serveur institutionnel
- Du respect des questions liées au RGPD
- ...

Le coordinateur du projet sera responsable :

- Rédiger le plan de gestion des données et de procéder à ses mises à jour
- Des démarches à mettre en œuvre auprès du DPO
- ...

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

Recommandations:

- Expliquer comment les ressources nécessaires (par exemple le temps) à la préparation des données pour le partage/préservation (curation des données) ont été chiffrées. Examiner et justifier soigneusement toutes les ressources nécessaires pour diffuser 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.
- Indiquer 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 le montant et comment ces coûts seront couverts.

Recommandations:

Dans cette section, vous devez évaluer les coûts financiers (techniques et humains) nécessaires à une bonne gestion des données (collecte, stockage, conservation,). Il est à noter que les coûts liés à la gestion des données peuvent être éligibles à remboursement selon les agences de financement.

Exemple de réponse:

En termes de ressources humaines : nous allons procéder au recrutement d'un ingénieur de recherche pour assurer...

En termes de matériel: nous allons procéder à l'achat X PC portables, X tablettes...