

INSA de Toulouse: 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:

Exemple : 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:

Dans cette section apparaissent les méthodes et les protocoles utilisés pour la collecte de vos données. Dans la 1^{re} version du PGD, vous pouvez simplement indiquer que cette question est bien prise en compte, mais qu'elle sera traitée au fil du projet.

Pour pouvoir compléter la version finale du PGD :

Si ce sont de nouvelles données, vous pouvez expliquer la méthode et les protocoles de génération de ces données.

Si ce sont des données préexistantes, comment ont-elles été constituées ? S'agit-il d'archives que vous avez dépouillées, d'une base de données que vous avez achetée ou encore de données issues du web ?

Vous pouvez vous demander : Quels sont les traitements que vous avez apportés à vos données ? Comment êtes-vous passé de données brutes aux données raffinées ? Quels sont les traitements que vous avez apportés aux données existantes (réutilisées) ?

Pour vous aider : <https://dmeg.cessda.eu/Data-Management-Expert-Guide/2.-Organise-Document/Documentation-and-metadata>

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).

Exemple de réponse:

Exemple : Les données consistent en des séries chronologiques numériques (date * paramètre * valeur) et sont téléchargeables au format CSV (format ouvert). La volumétrie de la base de données est de X Go en 2023.

Recommandations:

Dans cette section, on vous demande de décrire le type de données qui seront générées ou réutilisées dans le cadre de votre projet : par exemple numérique (bases de données, tableurs), textuel (documents), image, audio, vidéo, et/ou médias composites...

Le format demandé est celui sous lequel les données seront stockées et mises à disposition : PDF, xls, doc, txt, rdf... Nous vous recommandons de privilégier des formats ouverts (non propriétaires) et pérennes. Pour vérifier que les formats utilisés sont pérennes et archivables, vous pouvez utiliser l'outil du CINES : <https://facile.cines.fr/>
Enfin, la volumétrie de vos données peut être exprimée en espace de stockage requis (octets), et/ou en quantité d'objets, de fichiers, de lignes, de colonnes...

Il peut être assez difficile de donner la volumétrie des données en début de projet. Vous pouvez donc seulement fournir une estimation et faire évoluer cette partie à chaque nouvelle version du PGD.

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.

Exemple de réponse:

Exemple : Chaque dossier partagé est accompagné d'un fichier texte du même nom avec le suffixe "README" qui contient les métadonnées selon le format Dublin Core : Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifiant, Source, Language, Relation, Coverage, Rights. A ces métadonnées de base pourront s'ajouter des métadonnées propres au type de protocole ou de donnée. La liste de ces métadonnées sera établie au cours du projet.

Recommandations:

L'objectif de cette section est d'assurer une bonne compréhension des données afin de permettre une réutilisation ultérieure.

Si vous utilisez des standards de métadonnées (par exemple DDI (Data Documentation Initiative), TEI, EML (Ecological Metadata Language), MARC, CMDI, etc. ; le site [Disciplinary Metadata](#) donne les métadonnées par disciplines et par types de ressources) pour décrire les données brutes comme les données délivrables vous pouvez l'indiquer ici. Si vous ne pouvez pas utiliser de standards reconnus, il faut préciser les métadonnées que vous utilisez.

L'outil [Datacite Metadata Generator](#) vous permet de générer un fichier de métadonnées.

Pour accompagner vos jeux de données, il vous sera demandé de produire un fichier Readme, qui servira à décrire les fichiers, leurs conventions de nommage, le contrôle des versions, la structure des dossiers, éventuellement les logiciels à utiliser pour les consulter, etc. Vous pouvez également accompagner vos données de l'ensemble la documentation associée : questionnaire d'enquête, dictionnaire des variables, protocole, ...

L'objectif est que chaque jeu de données soit accompagné de toutes les informations nécessaires pour que les données puissent être accessibles, interprétables et réutilisables par une personne n'ayant pas participé à l'étude qui les a produites.

Pour vous aider, le guide "Décrire ses jeux de données dans les règles du FAIR" : <https://doi.org/10.5281/zenodo.8376581>

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

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.

Exemple de réponse:

Exemple : Tout fichier déposé sur l'espace partagé fait l'objet d'une double vérification : par la personne produisant le fichier et par le coordinateur qui vérifie le dépôt.

Les métadonnées sont renseignées sur la base de thesaurus. Les bonnes pratiques de laboratoire seront suivies pour le contrôle et la qualité des données.

Recommandations:

Pour décrire la démarche de contrôle qualité que vous allez adopter vous pouvez vous appuyer sur l'article « [Qualité des données](#) » sur la base Techniques de l'Ingénieur (<https://doi-org.gorgone.univ-toulouse.fr/10.51257/a-v2-h3700>) ou sur le guide « [Traçabilité des activités de recherche](#) » du CNRS (<https://qualite-en-recherche.cnrs.fr/gt/tracabilite-des-activites-de-recherche/>)

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.

Exemple de réponse:

Exemple : Les codes sources de logiciels sont sauvegardés sur des machines locales et en ligne sur des dépôts institutionnels eux-mêmes dupliqués sur les solutions publiques (forgemia.inra.fr, gitlab, github). Les données générées seront stockées sur les serveurs hébergeant le serveur Galaxy au sein de l'IFB. Les workflows Galaxy seront stockés sur le serveur qui hébergera la plate-forme Galaxy puis publiés sur le ToolShed (<https://toolshed.g2.bx.psu.edu>).

Recommandations:

Il vous est demandé ici de décrire l'endroit où les données seront stockées et sauvegardées au cours du projet et la fréquence à laquelle la sauvegarde sera effectuée (ici il est question du stockage des données chaudes pendant la durée du projet de recherche. Le stockage après la fin du projet sera abordé à la question 5). Afin d'éviter tout risque de perte ou de faille de sécurité il est recommandé de stocker les données dans au moins deux lieux distincts.

Il vous est recommandé de privilégier l'utilisation de systèmes de stockage robustes, avec sauvegarde automatique (outils institutionnels par ex.). Il n'est pas recommandé de stocker vos données sur des ordinateurs portables, des disques durs

externes, ou des périphériques de stockage tels que des clés USB.

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

Exemple : Les métadonnées sont archivées sur les serveurs de l'INSA et bénéficient du système de sauvegarde de l'Institut. Les serveurs sont sauvegardés quotidiennement et les sauvegardes conservées deux mois.

Recommandations:

Si vous traitez de données sensibles (à caractère personnel, politiquement sensibles, secrets commerciaux, etc.) comment seront-elles sécurisées au cours de l'étude : stockage sur un serveur protégé, encodage des données, anonymisation des données personnelles, etc.

Si vous traitez de données sensibles il faut mentionner ici le nom de la personne responsable de leur sécurisation et de leur conservation. Après sollicitation Le DPO (Data Protection Officer) de l'INSA Toulouse : contact-dpo@insa-toulouse.fr s'assurera de la conformité de votre traitement des données avec le RGPD, vous pourrez ensuite décrire dans le document les méthodes employées pour appliquer ses recommandations.

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.

Recommandations:

Si votre étude comporte plus spécifiquement des données à caractère personnel il vous faudra entrer en contact avec le DPO de l'INSA (contact-dpo@insa-toulouse.fr) pour qu'il certifie la conformité du projet avec le RGPD, notamment en ce qui concerne le consentement des participants pour la conservation et le partage de leurs données après anonymisation. Il vous sera également demandé d'expliquer de quelle façon les données seront anonymisées et éventuellement cryptées pendant le stockage.

Exemple de réponse:

Exemple : Le projet ne traite pas de données à caractère personnel.

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:

Exemple : Pour les données produites dans le cadre du projet de recherche, les partenaires de recherche sont conjointement propriétaires des données tel que mentionné dans l'accord de consortium. A la fin du projet, les données qui seront en libre accès seront sous licence X

Exemple : Les données réutilisées pendant ce projet sont la propriété de X. Elles sont disponibles ici : et placées sous licence X

Recommandations:

Pour les données générées au cours du projet, le propriétaire est en principe précisé dans l'accord de consortium. Vous pouvez donc reprendre les éléments définis dans le document en précisant les conditions d'accès qui s'appliqueront aux données : seront-elles librement accessibles, ou des restrictions seront-elles appliquées ? Si oui, lesquelles ? Pour permettre l'accès et la réutilisation des données, quelles licences leur seront appliquées ?

Pour les données préexistantes et que vous réutilisez dans le projet de recherche, il faut en préciser la source des données ainsi que les droits de propriété intellectuelle (données libres de droits ? sous licence ?)

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.

Exemple de réponse:

Exemple : Chacun des participants au projet est tenu de respecter les chartes d'éthique et de déontologie de son organisme d'affiliation dont les principes sont définis par la Charte nationale de déontologie des métiers de la recherche.

Recommandations:

Si les données recueillies font l'objet de questions d'éthique elles devront subir un traitement particulier en conséquence. Si l'approbation d'un comité d'éthique est nécessaire vous pouvez l'indiquer ici avec ses recommandations. Pour ces questions vous pouvez contacter la référente intégrité scientifique de l'établissement.

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:

Exemple : Le versement sur les bases de données publiques (ex : data.gouv.fr) se fera en cas de publication, valorisation et exploitation des données.

Recommandations:

Dans la première version du PGD vous pouvez renseigner ce champ a minima : indiquer que vous prenez en compte ces questions pour répondre aux exigences de l'agence de financement et que vous en préciserez les modalités dans les versions suivantes du PGD.

Pour la diffusion des données après la fin du projet les agences de financements et l'INSA recommandent de passer par un entrepôt de données, de préférence labellisé entrepôt de confiance. Vous n'avez pas besoin de donner ici le nom de l'entrepôt choisi, seulement de préciser la solution de stockage pérenne choisie. Si vous faites un autre choix que l'entrepôt de données, il vous sera demandé de prouver que les données pourront être prises en charge efficacement au-delà de la durée du financement du projet. Si vos données feront l'objet d'un embargo avant diffusion il faut l'indiquer, en en précisant la durée et la raison (données sensibles ou personnelles, obligations contractuelles, etc.).

Dans le cadre du partage de vos données et de leur réutilisation le ministère préconise l'emploi de la « Licence Ouverte / Open License » d'Etalab.

Attention, depuis la Loi pour une République Numérique (Art. L312-1-1 CRPA), les données achevées produites au sein d'établissements publics doivent obligatoirement être publiées en ligne, selon le principe "aussi ouvert que possible, aussi fermé que nécessaire".

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:

Exemple : Les données et protocoles utilisés le long du projet, présents dans l'espace d'échange NextCloud, seront préservés le temps du projet. Ceux qui feront l'objet d'une publication, d'une valorisation ou d'une exploitation dans des outils de recherche seront préservés à long terme dans un dépôt de données public tel que data.gouv.fr.

Recommandations:

A la fin du projet toutes les données n'auront pas vocation à être conservées et diffusées, quelles données ne doivent pas être divulguées ou seront détruites (pour des raisons contractuelles, légales, réglementaires, etc.) ? Sur quels critères déciderez-vous quelles données vous conserverez et combien de temps ? Attention, la durée de conservation de certaines données peut être contrainte.

C'est dans cette section que vous pouvez préciser où les données seront déposées. Si aucun entrepôt de confiance (liste sur le site de CoreTrustSeal ou par le répertoire Re3data) ou entrepôt disciplinaire ne correspond à votre projet vous pouvez déposer vos données sur RechercheDataGouv, l'entrepôt national créé par le MESR pour répondre aux besoins des institutions françaises.

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.

Exemple de réponse:

Exemple : Les données de caractérisations physiques et chimiques seront dans des tableaux de type CSV accessibles librement. Les résultats des modèles seront dans des tableaux de type CSV. La reproduction des résultats nécessitera l'utilisation du code qui sera récupérable via une forge logicielle. Les données des analyses de cycle de vie (ACV) seront également des tableaux de type CSV. Les inventaires peuvent être utilisés dans un logiciel ACV qui nécessitera d'autres bases de données payantes.

Recommandations:

Si vos données ne peuvent pas être fournies dans des formats ouverts (cf. [CINES](#)) n'oubliez pas de préciser quel logiciel doit être utilisé pour y accéder ainsi que sa version.

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.

Recommandations:

La plupart des entrepôts de données attribuent un DOI aux jeux de données qui y sont déposés, c'est la solution à privilégier. Si ce n'est pas le cas vous pouvez faire appel à un service externe comme Crossref ou DataCite qui sont accessible via un abonnement payant (pour Datacite il faut passer par l'INIST représentant le CNRS).

Exemple de réponse:

Exemple : Un DOI sera attribué automatiquement aux jeux de données lors du dépôt sur l'entrepôt de données. SoftwareHeritage permet l'assignation d'un DOI (SHID) qui pourra être utilisé lors de la publication afin d'identifier de façon pérenne une version du logiciel.

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.

Exemple de réponse:

Exemple : Chaque membre du projet est responsable de la maintenance de son protocole ou de ses données, au moins pendant le temps du projet. Le responsable du PGD, s'attachera au cours du projet à ce que la gestion des données suive les principes FAIR et à ce que le plan de gestion des données soit révisé régulièrement en fonction de l'avancement du projet.

Recommandations:

Dans cette section, vous pourrez détailler les rôles des personnes impliquées dans la gestion des données du projet (ex : collecte, traitement, stockage, archivage et partage des données, rédaction du PGD, mise à jour du PGD, etc.).

Pour les projets menés en collaboration, il faut se poser la question de comment s'effectue la coordination des responsabilités de gestion des données entre partenaires ?

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.

Exemple de réponse:

Exemple : Une grande partie du projet est dédié à rendre FAIR les données et les outils. Une partie conséquente du budget et du temps sera donc allouée à cette tâche, par la nature même du projet. De plus, le budget du projet intègre de l'équipement informatique et des ressources contractuelles dédiées spécifiquement à la gestion des données produites au cours du projet.

Recommandations:

Si ces coûts sont estimés en amont ils peuvent être inclus dans la demande financière.

Les coûts nécessaires à la collecte, au stockage et à la conservation des données incluent : coûts RH environnés, coût du matériel si nécessaire, frais de stockage dans un entrepôt de données, etc.

Des outils existent pour vous aider à les estimer :

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

https://dorum.fr/enjeux-benefices/le-cout-de-la-gestion-des-donnees_10_13143_hch2-h207/