

---

# DMP du projet "Pour une université de Sciences humaines et sociales responsable et influente : la transformation interne au service des grands défis internationaux"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "ANR - Modèle de PGD (français)" fourni par Agence nationale de la recherche (ANR).

## Renseignements sur le plan

<b>Titre du plan</b>	DMP du projet "Pour une université de Sciences humaines et sociales responsable et influente : la transformation interne au service des grands défis internationaux"
<b>Version</b>	Version initiale
<b>Objet/périmètre du plan</b>	<p>Ce plan de gestion de données (PGD) est la version initiale du PGD demandé par l'ANR, dans le cadre du financement attribué au projet Sorb'Rising en réponse à l'appel à projets ExcellencES 2021.</p> <p>Ce document, produit à partir de l'outil DMP OPIDoR, sera amené à évoluer au fur et à mesure que les activités du projet seront mises en place et que les données seront produites.</p>
<b>Domaines de recherche (selon classification de l'OCDE)</b>	Art (arts, history of arts, performing arts, music), Law, Social and economic geography, History and archaeology, Computer and information sciences, Languages and literature, Mathematics, Philosophy, ethics and religion, Political science, Economics and business, Sociology
<b>Langue</b>	fra
<b>Date de création</b>	2022-09-16
<b>Date de dernière modification</b>	2022-12-08
<b>Documents (publications, rapports, brevets, plan expérimental....), sites web associés</b>	<ul style="list-style-type: none"><li>PIA4 - Sélection 2021 appel à projets EXcellencES : <a href="https://anr.fr/fileadmin/aap/2021/selection/ia-excellences-2021-CR-jury.pdf">https://anr.fr/fileadmin/aap/2021/selection/ia-excellences-2021-CR-jury.pdf</a></li></ul>

## Renseignements sur le projet

<b>Titre du projet</b>	Pour une université de Sciences humaines et sociales responsable et influente : la transformation interne au service des grands défis internationaux
<b>Acronyme</b>	Sorb'RISING
<b>Résumé</b>	Sélectionné lors de l'appel 2021 « ExcellencES sous toutes ses formes », géré par l'Agence Nationale de la Recherche au titre du Plan France 2030 (ANR-21-EXES-0015), le projet Sorb'Rising est mis en œuvre par l'Université Paris 1 Panthéon-Sorbonne, en partenariat avec le CNRS et l'IRD, et avec le soutien d'institutions et réseaux internationaux (AUF, Una Europa, Unimed), d'organisations nationales et territoriales (Sénat, Ville de Paris, Région Île-de-France, CROUS), ainsi que d'organisations de formation et de recherche (Cité du Développement Durable, Fondation Saint-Cyr, IAE Paris - Sorbonne Business

School, INHA, Revue Défense Nationale).

Le projet Sorb'Rising a pour ambition de renforcer l'identité de l'Université Paris 1 Panthéon-Sorbonne, en tant qu'Université de rang mondial en sciences humaines et sociales, engagée et ouverte sur la société. Il est construit sur une durée de 8 ans (2022-2030) et s'articule autour de trois défis :

- former les étudiants aux enjeux du XXI<sup>e</sup> siècle, en particulier les transformations politiques, sociales, environnementales et numériques en cours ;
- développer une recherche transdisciplinaire et mieux diffuser les résultats des recherches en SHS en dehors du monde universitaire ;
- accompagner les personnels de l'université dans ces transformations.

Pour répondre à ces défis, ce projet structurant pour l'établissement se décline en quatre axes d'intervention :

- Axe 1 : Soutien à la recherche, incluant notamment : i) la promotion d'initiatives transdisciplinaires de recherche qui répondent aux grands enjeux sociétaux et ii) la mise en place d'un centre d'excellence numérique
- Axe 2 : Formation des étudiants à travers : i) des modules d'enseignements de compétences transverses et ii) des programmes de formation transdisciplinaires pertinents en termes de défis sociétaux
- Axe 3 : Renforcement de capacités des personnels en termes de compétences générales, telles la communication, la gestion de projet, les outils numériques, etc.
- Axe 4 : Lien Sciences – Sociétés, avec notamment : i) la création d'un incubateur entrepreneurial (pour accompagner la création d'entreprises en lien avec les SHS) et ii) d'une web-TV

Chacun de ces axes fait l'objet de production de données, dont les caractéristiques sont spécifiques aux types d'activités concernées. Sont donc présentés ci-dessous, par axes : i) la nature de ces données, ii) le type de données, ii) les producteurs de ces données, iii) et l'implication en termes de plan de gestion de données.

Axe	Nature des données produites	Type	Producteurs de données	Assujettissement à un plan de gestion de données
Axe 1	Données de recherche	Jeu de données	Equipes de recherche sélectionnées à l'issue des appels à projets Sorb'Rising (lancés de 2022 à 2029)	Chaque projet de recherche financé fait l'objet d'un PGD individuel
Axe 2	Contenus de formation	Ressources interactives	Enseignants-chercheurs de l'Université Paris 1	N/A
Axe 3	Contenus de formation	Ressources interactives	Enseignants-chercheurs de l'Université Paris 1	N/A
Axe 4	Vidéos et articles de vulgarisation scientifique	Audiovisuel et Texte	Services de communication de l'Université	N/A

Parmi les données qui sont produites dans le cadre de Sorb'Rising, seules les données de recherche, issues des projets financés, doivent présenter un plan de gestion de données. **Le présent PGD de Sorb'Rising, ou "PGD cadre", définit donc les recommandations et directives générales à suivre par chacun des porteurs de projets financés, pour la mise en place des "PGD individuels".**

Ces plans de gestion de données "individuels", relatifs spécifiquement à chacun des projets financés devront être élaborés conformément aux recommandations et règles présentées ici, ainsi qu'aux normes établies dans le cadre du RGPD concernant la production, le traitement, l'analyse, le stockage et l'archivage des données à finalité de recherche.

Ils seront à réaliser à l'aide de l'outil [DMP Opidor](#), mis à disposition de la communauté de l'Enseignement Supérieur et de la Recherche, sur la base du modèle de PGD de l'ANR (accessible sur DMP Opidor). En cas de difficultés, les responsables de projets pourront solliciter le soutien des services ad-hoc de l'Université : D2P, DIREVAL, le groupe de travail science ouverte de l'université (GTSO), le comité d'éthique, le DPO.

Les PGD individuels ainsi rédigés seront rattachés au présent PGD cadre, au fur et à mesure qu'ils seront mis en place.

### Sources de financement

- Agence Nationale de la Recherche : ANR-21-EXES-0015
- Institut de Recherche pour le Développement, IRD :
- Centre National de la Recherche Scientifique, CNRS :
- Université Paris 1 Panthéon Sorbonne, UP1PS :

**Date de début** 2022-03-15

**Date de fin** 2030-03-14

### Partenaires

- Institut de recherche pour le développement ()
- Centre national de la recherche scientifique ()

### Produits de recherche :

1. Données de recherche produites dans le cadre de l'Axe 1 du projet Sorb'Rising (Jeu de données)

### Contributeurs

Nom	Affiliation	Rôles
BORTOT Séverine	Université Paris 1 Panthéon-Sorbonne	
CHAUVET Stéphanie	Université Paris 1 Panthéon-Sorbonne	<ul style="list-style-type: none"> <li>• Personne contact pour les données</li> <li>• Responsable du plan de gestion de données</li> </ul>
NEAU-LEDUC Christine	Université Paris 1 Panthéon-Sorbonne	<ul style="list-style-type: none"> <li>• Coordinateur du projet</li> </ul>

Droits d'auteur :

Le(s) créateur(s) de ce plan accepte(nt) que tout ou partie de texte de ce plan soit réutilisé et personnalisé si nécessaire pour un autre plan. Vous n'avez pas besoin de citer le(s) créateur(s) en tant que source. L'utilisation de toute partie de texte de ce plan n'implique pas que le(s) créateur(s) soutien(nen)t ou aient une quelconque relation avec votre projet ou votre soumission.

# DMP du projet "Pour une université de Sciences humaines et sociales responsable et influente : la transformation interne au service des grands défis internationaux"

---

## 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 ?

Les données de recherche sont produites par des équipes de l'Université Paris 1 Panthéon Sorbonne, qui sont financées dans le cadre du projet Sorb'Rising. Celles-ci sont sélectionnées à l'issue d'appels à projets (AAP) visant à identifier des sujets de recherche transdisciplinaires répondant aux thématiques, aux enjeux et aux critères de sélection de Sorb'Rising.

Les PGD individuels mis en place par chacun des responsables des projets financés doivent documenter :

- Le mode de recueil des données (dépouillements de sources primaires, enquêtes, mesures de terrain, expérimentations, analyses, modélisation, etc.)
- Les arbitrages technologiques effectués, qui devront être justifiés, pour chaque catégorie de données, en regard des possibilités : état de l'art et ressources dont dispose le projet (financières, techniques, humaines)
- L'éventuelle ré-utilisation de données préexistantes, avec mention impérative des sources d'où proviennent ces données, leur durée et conditions d'accessibilité, ainsi que leur lieu de stockage

---

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

Les projets de recherche sélectionnés et financés dans le cadre de Sorb'Rising pouvant être de nature variée, les données produites sont également plurielles, par exemple : données de dépouillements, données d'enquêtes, données statistiques, données démographiques, données géophysiques, données bibliométriques, données photographiques, etc.

La caractérisation de ces données doit être documentée dans chacun des PGD individuels, avec une attention particulière apportée, pour chacun des jeux de données :

- à l'utilisation d'un nom unique et explicite à définir, pouvant éventuellement être au format suivant : Sorb'Rising\_Acronyme du projet financé\_N° du jeu de donnée\_date (au format AAAA-MM-JJ)
- à la mise en œuvre d'outils et de standards pertinents (modernes, interopérables) pour chaque catégorie de données produites ou mobilisées (ex. données géomatiques, édition scientifique de sources, bases de données relationnelles, graphes de connaissances, données statistiques, données pour l'analyse de réseaux, sources numérisées, corpus bibliographique, thesaurus, etc.)
- à la finalité de la gestion des données : production de connaissances, science reproductible, préservation de données à valeur patrimoniale, diffusion auprès du public, etc.
- au périmètre thématique (philosophie, géographie, sciences économiques, etc.)
- au périmètre géographique
- à la nature des données (dépouillement de sources primaires, transcription d'entretiens, relevés de fouilles, relevés pédologiques, relevés de températures, images satellites, etc.)
- aux types de données (numérique, textuel, image, audio, modèle, ressource interactive, etc.)
- aux formats techniques (formats ouverts à privilégier)
- au volume de stockage requis pour ces données

Pour garantir l'accès aux données et leur pérennité, chacun des projets financés doit privilégier l'utilisation de formats standards et ouverts. Se référer lorsque possible, pour le choix du format de fichier, au [récapitulatif élaboré par DORANum](#) sur les formats ouverts et fermés.

Dans le cas où l'utilisation de formats ouverts n'est pas possible, les porteurs des projets le justifient dans leur PGD.

**Pour plus d'informations**, consulter le site du CINES qui présente les [différents types de format](#) et a développé un service de validation de formats, [FACILE](#), pour vérifier que les formats utilisés sont pérennes et peuvent être archivés.

## 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 ?

Les jeux de données issus des projets de recherche financés par Sorb'Rising sont de différentes natures, types et formats. C'est pourquoi, afin d'assurer une description uniforme, dans le but de faciliter le recensement des données, leur identification et leur ré-utilisation par la communauté scientifique, chaque responsable de projet financé doit fournir une notice de métadonnées en respectant la trame fournie par Sorb'Rising, en remplissant le formulaire suivant : LIEN A RAJOUTER ULTERIEUREMENT.

La compilation de ces informations, fournies par chacun des responsables de projets financés, permet d'établir un catalogue global des données de recherche produites dans le cadre de Sorb'Rising, au travers de l'ensemble des projets individuels de recherche.

Sont ainsi récapitulées, pour chacun des jeux de données :

- l'identifiant unique du jeu de données, de type DOI
- le nom du fichier de données, unique et explicite
- les finalités scientifiques de la collecte et de la conservation de ces données
- le périmètre thématique
- la nature des données
- le format des données
- le type de fichier correspondant
- la personne ou l'équipe contact, responsable de ces données
- le lieu de stockage du jeu de données
- les droits d'accès
- les principales publications associées au jeu de données (avec leur DOI)
- la documentation éventuellement disponible (en plus du fichier de type "Lisez-moi" présenté ci-dessous)

En plus des métadonnées, chacun des responsables de projets financés doit fournir un fichier de type "Lisez-moi" (au format .txt) comportant une présentation du jeu de données et la description de son contenu, incluant :

- l'identifiant unique du jeu de données, de type DOI
- le nom du fichier de données, unique et explicite
- l'ensemble des éléments pertinents pour la compréhension scientifique du contexte de production ou de collecte des données
- la méthodologie de collecte ou de production des données
- une documentation appropriée à la nature des données : par exemple, dans le cas des données statistiques, une documentation de type "codebook" explicitant les codes, variables, unités, etc.
- la méthode d'analyse et de traitement des données
- et toute autre information s'avérant nécessaire à la compréhension des données, comme par exemple :
  - les éléments liés aux traitements de données personnelles, qui documentent la fiche du registre des activités de traitement,
  - la présentation des procédures d'anonymisation et/ou de pseudonymisation de données personnelles,
  - la mention, lorsque nécessaire, du logiciel spécifique à utiliser pour la visualisation et/ou le traitement des données (avec indication du site où télécharger le dit-logiciel - Voir points 3a et 5c.)

La mise à disposition de ce fichier "Lisez-moi" se fait également à travers le formulaire de dépôt des métadonnées : LIEN A RAJOUTER ULTERIEUREMENT

En plus de la mise à disposition des métadonnées et du fichier Lisez-moi sur le formulaire Sorb'Rising ci-dessus, chaque responsable de projets doit également :

- sauvegarder ces deux fichiers ("métadonnées" au format .csv et "Lisez-moi" au format .txt), aux côtés de son / ses jeu(x) de données sur son espace de stockage / sauvegarde
- partager ces deux fichiers "métadonnées" et "Lisez-moi" sur l'entrepôt de données choisi, aux côtés de son/ses jeu(x) de données tel que présenté au point 5 ci-dessous.

---

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

Les jeux de données sont produits par différentes équipes de recherche, dans le cadre de leurs activités et sujets respectifs. Par conséquent, la qualité de chaque jeu de données est du ressort des responsables respectifs de chacune de ces équipes, tant pour la description de ces données (dans le fichier de métadonnées et le fichier "Lisez-moi"), que pour les données elles-mêmes (repérage des valeurs nulles, doublons, valeurs anormales, etc.).

Les outils utilisés et les procédures de contrôles mis en place doivent être précisés dans les PGD de chacun des projets financés.

---

### 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 ?

Chacun des PGD individuels des projets financés par Sorb'Rising doit détailler les lieux et supports de stockage ainsi que les procédures de sauvegarde qui sont utilisés pour les données et métadonnées : périodicité des sauvegardes, automatisation, nombre de copies et localisation. Un minimum de 2 copies en des lieux de stockage distincts doit être prévu.

Une attention particulière sera apportée aux systèmes de stockage qui doivent être robustes et pérennes, situés idéalement en France (et a minima dans l'Union Européenne), avec si possible des sauvegardes automatiques, tels que ceux qui pourraient être fournis par la Direction du Système d'Information et des Usages Numériques - DSIUN (en fonction des usages et de la volumétrie), par [la plateforme d'Huma-Num](#), ou encore par une autre infrastructure de recherche française ou européenne. Le stockage des données sur des ordinateurs portables, des disques durs externes, des périphériques de stockage tels que des clés USB ou encore des services de stockage "cloud" ne dépendant pas d'institutions françaises d'enseignement et de recherche (type Dropbox, Google Drive, etc.) sont interdits.

Il est par ailleurs attendu que chacun des jeux de données et les métadonnées correspondantes soient téléversés dans des entrepôts de données robustes et fiables (voir point 5). Outre de faciliter le partage des données, ces entrepôts permettent en effet également une sauvegarde automatique des données, ainsi que leur récupération / restauration en cas d'incidents informatiques.

Le stockage et la sauvegarde des données sensibles nécessitant une vigilance accrue (voir point ci-dessous), il convient dans le cas de la production de telles données, de contacter le Délégué à la Protection des Données (DPO) de l'Université Paris 1 : [dpo@univ-paris1.fr](mailto:dpo@univ-paris1.fr).

Le catalogue global des données de recherche, reprenant la compilation des métadonnées de chacun des projets financés, sera quant à lui sauvegardé par la Direction des Projets et de la Prospective, en tant que service coordinateur du projet Sorb'Rising : i) sur les serveurs de l'Université, ainsi que ii) sur une plateforme de partage à définir ultérieurement.

#### Cas particulier des logiciels et interfaces :

Si, dans le cadre d'un projet financé par Sorb'Rising, la visualisation et/ou l'analyse des données produites impliquent le développement d'applications ou de scripts, alors il est attendu que soient utilisés de préférence des technologies libres, que le code soit publié sous licence libre, et que celui-ci :

- soit partagé sur une plateforme de type [Github](#) ou [Gitlab](#)
- et sauvegardé / archivé sur le site [Software Heritage](#)

#### Pour plus d'informations :

- sur la sauvegarde et la sécurisation des données, se référer aux conseils de [DORANum](#)
- sur la sauvegarde et la sécurisation des codes et logiciels, se référer au passeport "[Codes et Logiciels](#)"

#### A noter :

Les responsables de projets sont invités à se rapprocher de la DSIUN pour communiquer leurs besoins en termes de stockage de données.

---

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

Le caractère sensible ou personnel des jeux de données produits par les projets financés par Sorb'Rising dépend de la nature de ces données et du sujet de recherche correspondant. Il convient donc que cela soit précisément décrit dans chacun des PGD individuels correspondant aux projets financés.

Dans le cas des projets concernés par la production ou le recueil de données sensibles ou de données à caractère personnel, la protection de ces données doit être assurée en respect de la [charte informatique de l'Université](#), ainsi que du [RGPD](#), au travers notamment de l'application de différentes mesures de protection :

- le choix de l'espace de stockage doit être adapté au caractère sensible ou non des données collectées, de sorte que le niveau de partage des données soit fonction du niveau de protection à appliquer

- des droits d'actions différentiels peuvent être attribués aux différentes personnes/fonctions : en lecture seule, édition, suppression
- des conditions d'accès différentiels : à distance ou en local uniquement, avec un accès complet ou restreint aux données en fonction des personnes/fonctions
- d'éventuelles mesures techniques spécifiques à la sécurité des données les plus sensibles peuvent être envisagées : protection des fichiers par mot de passe, chiffrement des fichiers ou du disque dur, utilisation de conteneurs sécurisés, etc.
- les données à caractère personnel peuvent être pseudonymisées ou anonymisées dès la collecte initiale ou lors d'un retraitement ultérieur
- d'un point de vue organisationnel, l'attribution ou le retrait de droits d'accès, ainsi que leur revue régulière, doivent faire l'objet d'une attention particulière
- etc.

Pour s'assurer de la mise en place des bonnes pratiques en matière de protection de ces données sensibles ou à caractère personnel, les porteurs de projets concernés peuvent se rapprocher du Délégué à la Protection des Données (DPO) de l'Université Paris 1 ([dpo@univ-paris1.fr](mailto:dpo@univ-paris1.fr)) et [du comité d'éthique](#) de l'Université Paris 1. Il convient de mener avec lui une analyse des risques encourus et des mesures techniques et organisationnelles à mettre en œuvre pour protéger ces données. (Mesures qui sont à décrire avec précision dans les PGD individuels.)

**Pour plus d'informations** sur la gestion des données sensibles, se référer aux conseils de la CNIL sur la [sécurité des données](#) et sur la [recherche scientifique](#), ou aux [recommandations du CNRS](#) concernant la protection des données sensibles en SHS dans un contexte de science ouverte.

## 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 ?

Les projets financés par Sorb'Rising donnant lieu à la production ou au recueil de données personnelles doivent faire l'objet d'une analyse en lien avec le DPO ([dpo@univ-paris1.fr](mailto:dpo@univ-paris1.fr)) et [le comité d'éthique](#) de l'Université Paris 1. L'objectif de cette analyse est de déterminer si les traitements sont en parfaite conformité avec le RGPD. Par ailleurs, des données sensibles peuvent nécessiter d'être protégées et d'échapper aux principes de libre diffusion, dans le cas notamment de : données scientifiques à risques (sécurité d'Etat, sécurité des populations, etc.), données personnelles, données sensibles, données de santé, données liées à l'intelligence économique (secret industriel et commercial).

Si des données à caractère personnel sont produites, les équipes devront se mettre en conformité avec le RGPD et la Loi Informatique et Libertés, en veillant notamment à :

- Obtenir un consentement éclairé en préalable du recueil de données personnelles
- Communiquer en toute transparence sur l'usage fait de ces données et leur partage éventuel
- Envisager l'anonymisation, la pseudonymisation ou le chiffrement au plus tôt des données personnelles
- Sécuriser ces données personnelles via des procédures d'accès spécifiques pour les utilisateurs autorisés à y accéder
- S'organiser pour pouvoir répondre aux demandes d'exercice de droits d'accès ou de suppression
- Définir une durée de conservation de ces données

**Pour plus d'informations**, se référer au [guide élaboré par l'InSHS](#) sur la protection des données personnelles et aux « [Quinze recommandations aux chercheurs](#) » formulées par les réseaux Sup DPO et R3Sup. La CNIL a de même mis en ligne des pages spécifiques destinées à la recherche : <https://www.cnil.fr/fr/recherche-scientifique-hors-sante>.

### 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 ?

Dans le cas où plusieurs équipes partenaires sont impliquées dans la genèse des jeux de données, les droits de chacune de ces équipes sur les données et le matériel doivent être spécifiés dans l'accord de consortium s'il existe, et/ou à défaut dans le plan de gestion de données individuel correspondant au projet en question. Doivent être notamment spécifiés : les droits de propriété intellectuelle sur les données, qui en contrôle l'accès, les conditions d'accessibilité, les éventuelles restrictions appliquées, etc.

Afin de s'assurer si les jeux de données nécessitent ou non des formalités spécifiques liées par exemple à l'Accès et au Partage des Avantages ([Protocole de Nagoya](#)), ou à la nature spécifique de certains matériels, le Délégué à la Protection des Données de l'Université Paris 1 peut être sollicité lorsque nécessaire : [dpo@univ-paris1.fr](mailto:dpo@univ-paris1.fr).

Les membres des équipes des projets financés par Sorb'Rising s'engagent à respecter la législation en vigueur tout au long du projet (RGPD, droit à l'image, propriété intellectuelle, etc.) et à accompagner la diffusion et le dépôt de leur jeu de données d'un contrat de

licence. Les licences [Etalab](#) ou [Creative Commun](#) (CC-BY) sont à privilégier, avec un niveau d'ouverture de la licence CC à adapter aux données produites et au niveau d'ouverture souhaité. (Possibilité d'ajouter si nécessaire deux contrats de licence.)

Pour vous aider à choisir une licence, consulter le " [Public license selector](#)", le [site de l'INRAE](#), ou le [site gouvernemental](#) sur les licences de réutilisation qui spécifie la liste des licences pouvant être utilisées par les administrations.

**Pour plus d'informations** sur la propriété intellectuelle des données, se référer aux pages 17-20 du document [« Ouverture des données de la recherche : guide d'analyse du cadre juridique en France »](#).

---

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

Chaque équipe de recherche financée dans le cadre de Sorb'Rising doit s'interroger sur les données qu'elle produit et sur les problèmes d'éthique ou de protection de la confidentialité pouvant éventuellement se poser (données sensibles, données de santé, données personnelles, etc.).

Pour évaluer les situations au cas par cas, chaque responsable de projet peut, préalablement à l'élaboration de son plan de gestion de données individuel, prendre l'attache du [comité d'éthique de Paris 1 ou des référents](#) (déontologie, laïcité, intégrité scientifique, égalité, défense et sécurité), afin de veiller à prendre les mesures nécessaires de protection en terme de récolte de données, stockage, transfert et conditions d'accès. Celles-ci doivent alors être explicitement détaillées dans les PGD individuels des projets financés.

Dans le cas où les jeux de données produits ne soulèvent aucun problème d'éthique, cela doit également être spécifié dans les PGD individuels.

---

## 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 ?

La diffusion des données est une obligation par défaut mais elle s'adapte au principe « aussi ouvert que possible, aussi fermé que nécessaire ». Les modalités et échéances de partage des données de recherche peuvent donc varier en fonction des projets financés par Sorb'Rising, afin de s'adapter au cas particulier et aux spécificités de chacun des sujets. Toutefois, en respect des principes de Science Ouverte promue par l'Université Paris 1, de même que par l'ANR qui finance Sorb'Rising, les responsables des projets financés s'engagent à publier et à partager leurs données aussi rapidement que possible.

A cet égard, les informations suivantes doivent être spécifiées dans les PGD individuels de chacun des projets financés :

- quand et comment les données sont rendues disponibles à la communauté
- les licences ou contrats appliqués aux données partagées
- l'éventuelle protection de certaines données pour des raisons éthiques, de données confidentielles ou personnelles, de propriété intellectuelle ou commerciale, de sécurité, etc.
- s'il est nécessaire de restreindre l'accès à certaines communautés ou d'imposer un accord pour le partage de données
- si un délai de publication (ou embargo) doit ou non être appliqué pour assurer un accès exclusif au producteur de la donnée pendant une certaine durée (auquel cas en expliquer la raison)
- la durée d'archivage des données

#### Instructions en termes de partage :

Il est attendu des responsables des projets financés qu'ils déposent leurs jeux de données sur des entrepôts fiables, tels Nakala (d'Huma-Num), Zenodo, ou éventuellement des entrepôts thématiques de confiance. Outre de garantir le partage des données de recherche avec la communauté scientifique, ces entrepôts permettent en effet d'assurer également une sauvegarde plus sûre des données, des facilités de restauration en cas d'incidents informatiques, mais également de générer un identifiant unique pour les jeux de données (de type DOI).

Pour s'assurer du choix d'un entrepôt FAIR, opter pour l'un de ceux recensés par [Re3data.org](#) ou [Datacite Repository Finder](#), en particulier ceux qui disposent d'une certification de type [Data Seal of Approval \(DSA\)](#) ou [CoreTrustSeal](#).

Recommandation importante : attention à ne pas déposer un même jeu de données dans plusieurs entrepôts, afin d'éviter la dilution des citations et également parce qu'il ne doit y avoir qu'un seul et unique DOI par jeu de données. Les métadonnées et les DOI (des jeux de données) peuvent en revanche être présentés sur plusieurs sites, mais doivent impérativement renvoyer vers le même et unique entrepôt où est déposé le jeu de données.

Dans les conditions de l'article 30 de la Loi «Pour une République numérique», les responsables des projets financés s'engagent par ailleurs à déposer les publications scientifiques (texte intégral) issues de leur projet de recherche, de développement ou d'innovation, dans la collection Sorb'Rising ouverte sur le portail HAL de Paris 1 : LIEN A RAJOUTER.

**Pour plus d'informations** sur les obligations et les limites au partage des données, consulter le document ["Ouverture des données de recherche : cadre juridique»](#).

---

#### 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) ?

Pour chacun des projets de recherche financés par Sorb'Rising, une réflexion doit être conduite par chacune des équipes responsables, afin d'établir un plan de conservation des données et de décider de leur durée d'archivage. Cette décision doit être prise en respect des exigences légales et déontologiques de conservation ou de destruction de certains types de données.

Il est nécessaire de distinguer les données qui sont à conserver sur le long terme (données ayant une valeur juridique, scientifique ou patrimoniale), de celles qui doivent être détruites à l'issue du projet, et de celles qui peuvent être détruites 3 ou 5 ans après la fin du projet (lorsqu'elles ne sont plus nécessaires à la validation des résultats ni pour une éventuelle ré-utilisation).

L'archivage des jeux de données sur le moyen ou long terme doit être assuré au travers d'une plateforme de dépôt de données pérenne et fiable (cf. point précédent), ou d'une plateforme d'archivage électronique comme la plateforme du [CINES](#) (qui héberge notamment les données stockées dans les outils d'[Huma-Num](#) en SHS). Le lieu et la durée d'archivage doivent être clairement explicités dans les versions finales des PGD individuels des projets financés.

**Pour plus d'informations** sur les questions d'archivage de données, consulter le site du [CINES](#), ou une fiche synthétique élaborée par [DORANum](#).

---

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

Chacun des PGD individuels (des projets financés) doit préciser, en fonction du type et de la nature du/des jeu(x) de données produits:

- les conditions d'accès aux données et la démarche à suivre (identifiant / mot de passe unique, gestion d'accès automatique ou par une personne, formulaire de demande d'accès, etc.),
  - si des logiciels sont nécessaires à l'exploitation des données, ou à l'inverse proscrits,
  - dans le cas où un programme ou une interface aurait été développés pour la visualisation et/ou l'analyse des données produites : le lien de téléchargement du code de programmation informatique (voir question 3a).
- 

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

L'attribution d'identifiants pérennes aux jeux de données permet : i) de les localiser, ii) de les référencer de façon fiable et efficace, iii) de comptabiliser les citations et les réutilisations.

Ces identifiants pérennes (de type DOI) sont générés automatiquement par les entrepôts de données de confiance, tel Nakala, Zenodo, etc., sur lesquels il est attendu que les jeux de données soient téléversés (voir point 5a) par les responsables des projets financés.

---

## 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) ?

La gestion des jeux de données produits par les projets soutenus dans le cadre de Sorb'Rising sont sous la responsabilité générale des porteurs de projets. Ceux-ci devront préciser dans les PGD individuels quelle est la liste des différents responsables impliqués à chacune des étapes importantes de production de données, en veillant à désigner une personne responsable, par exemple et de façon non exhaustive, pour:

- la coordination du projet

- la saisie des données
- le contrôle qualité
- le stockage, la sauvegarde et la sécurisation des données
- l'archivage et le partage
- la production des métadonnées
- la mise en œuvre du plan de gestion des données

Au niveau du projet Sorb'Rising dans son ensemble, la Direction des Projets et de la Prospective, en charge de la coordination de Sorb'Rising pour l'Université Paris 1, assure la responsabilité des différentes étapes relatives à la gestion globale, i.e. :

- le stockage, la sauvegarde et le partage du catalogue global des données de recherche produites (incluant les métadonnées de chacun des jeux de données des projets financés)
- la mise en œuvre et la révision périodique du présent PGD cadre
- le suivi de la mise en place des PGD individuels par chacun des responsables de projets financés.

**Contact pour la D2P :**

Stéphanie Chauvet, coordinatrice du projet Sorb'Rising (sorb-rising@univ-paris1.fr)

---

**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) ?**

La gestion de chaque jeu de données est de la responsabilité de chacune des équipes financées par le projet Sorb'Rising, qui doivent donc quantifier le temps de travail nécessaire à la gestion et au partage des données, à leur éventuelle anonymisation, ainsi que les coûts occasionnés en termes de recrutement, de frais de matériel, de stockage, d'archivage pérenne, etc.

Chacun des responsables des jeux de données produits doit veiller à ce que ceux-ci soient "Facile à trouver, Accessible, Interopérable, et Réutilisable" (FAIR).

**Pour en savoir plus** sur le coût de la gestion des données, voir le site d' [OpenAIRE](#). ou de l'[Université d'Utrecht](#).