
Bridge Research through Interoperable Data Governance and Environments

Plan de gestion de données créé à l'aide de DMP OPIDoR

Créateurs du PGD : PASCAL AVENTURIER, Luc Decker, Hanka Hensens

Affiliation du créateur principal : IRD - Institut de Recherche pour le Développement

Modèle du PGD : ANR - Modèle de PGD (français)

Dernière modification du PGD : 09/11/2020

Financeur : Agence nationale de la recherche (ANR)

Numéro de subvention : ANR-19-DATA-0013-03

Résumé du projet :

The current organization of national research into joint units and the resulting significant joint digital production requires, at the inter-institute level, the coordination of data policies, which is fundamental for the implementation of Open Science at the national and international levels. Thus, the purpose of the BRIDGE project is to take advantage of a favourable partnership context between IRD, INRA, CIRAD to demonstrate how, by activating policy, cross-community and technical levers, we can promote the implementation of the FAIR principles. Therefore, the BRIDGE project will bring people with various policy, scientific and technical backgrounds and build a practical roadmap towards a concrete implementation of the FAIR principles in each of the partners' institutes. To reach that goal, the BRIDGE project will achieve the following objectives: at the policy level, analyze, compare and harmonize the institutional data governance strategies ; and across-communities, develop practical guidelines for both researchers and data managers. To that aim, the partners will assess existing recommendations and guidelines in order to evaluate how they can be adapted and generalized for different communities. We plan to provide feedback to the developers of the existing recommendations on how applicable they are. At the technical level, the goal is to enhance the current data management and sharing tools of the partners, specifically the data repositories, by increasing their level of compliance with the FAIR principles. The institutional repositories of both partners being based on the same tool, the open source Dataverse, it makes it easier to align strategies and roadmaps. We will organize partners' engagement all along the project for the co-development of the project tools and outputs through a Community of Practice, and ensure its continuation beyond the project's lifetime. Indeed, this Community of Practice will be led by the actors of the data governance in the partners' institutes. The methodology and results envisaged by the project are intended to be generalizable and reusable beyond the three partners and their scientific and technical use cases.

Chercheur Principal : PASCAL AVENTURIER

Identifiant ORCID : 0000-0003-0211-4549

Contact pour les Données : PASCAL AVENTURIER

Produits de recherche :

1. WP1 Gouvernance : Corpus de textes sur la gouvernance des données, Grilles d'analyse, éléments bibliographiques, rapport de synthèse (Collection)
2. WP2 Guidelines : Critères et guide pratiques pour étudier le degré FAIR des entrepôts et application aux cas d'études (Texte)
3. WP3 IT TOOLS : Recommandations pratiques et techniques pour améliorer la FAIRness des entrepôts (Texte)
4. WP3 autocomplétion : Widget d'autocomplétion pour améliorer les métadonnées (Logiciel)

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.

Bridge Research through Interoperable Data Governance and Environments

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

WP1 Gouvernance : Corpus de textes sur la gouvernance des données, Grilles d'analyse, éléments bibliographiques, rapport de synthèse

WP / Tâche / Nom court (max 20 Caract)	Description du jeu de données	Type de collecte
WP1-Gouvernance	Corpus de textes (documents) validés ou en construction sur la gouvernance des données	Manuelle
WP1-Gouvernance	Grille d'analyse des politiques	création
WP1-Gouvernance	Documents de bibliographie	création
WP1-Gouvernance	rapport de synthèse et de conclusion	Création

Les textes seront produits en Open Document Format (ODF) et en PDF

WP / Tâche / Nom court (max 20 Caract)	Description du jeu de données	Type de données
WP1-Gouvernance	Corpus de textes (documents) validés ou en construction sur la gouvernance des données	Texte PDF 100 Mo max
WP1-Gouvernance	Grilles d'analyse originales (texte format tabulaire)	Textes (ODT + PDF) 50 Mo Max
WP1-Gouvernance	Documents de bibliographie	BDD bibliographie zotero 2 00 Mo Max
WP1-Gouvernance	Rapport de synthèse et conclusions	Textes (ODT + PDF) 100 Mo max

WP2 Guidelines : Critères et guide pratiques pour étudier le degré FAIR des entrepôts et application aux cas d'études

Les documents suivants seront rédigés à partir de guides et documents techniques.

WP2- Guide pratique pour les chercheurs	Création
WP2- Evaluation des critères pour étudier la fairisation des entrepôts	Création
WP2- Roadmap de fairisation des cas d'études	Création

WP2 Provide guidelines to the researchers	Guide pratique pour les chercheurs	Texte ODT	200 Mo
WP2	Evaluation des critères pour étudier la fairisation des entrepôts	Texte ODT	200 Mo
WP2	Roadmap de fairisation des cas d'études	Texte ODT	200Mo

WP3 IT TOOLS : Recommandations pratiques et techniques pour améliorer la FAIRness des entrepôts

Les documents suivants seront rédigés à partir de guides et documents techniques. Des tests seront effectués sur les applications "use case" et les entrepôts de données des trois partenaires.

WP3-IT Tools	Rapport de FAIRness des entrepôts IRD, CIRAD et INRA	Création
WP3-IT Tools	Recommandations pratiques pour améliorer la FAIRness des entrepôts	Création

WP3-IT Tools	Rapport de FAIRness des entrepôts IRD, CIRAD et INRA	Textes (Latex + ODT + PDF) 100Mo max
WP3-IT Tools	Recommandations pratiques pour améliorer la FAIRness des entrepôts	Textes (Latex + ODT + PDF) 100 Mo max

WP3 autocomplétion : Widget d'autocomplétion pour améliorer les métadonnées

Développement d'un widget d'autocomplétion pour améliorer les métadonnées, en langage *Javascript*.

WP3-IT Tools	Widget d'autocomplétion pour améliorer les métadonnées	Création de code Javascript
--------------	--	-----------------------------

Création du code par le projet pour un volume estimé de moins de 2 Go.

WP3-IT Tools	Widget d'autocomplétion pour améliorer les métadonnées	Code Javascript 10 Mo max
--------------	--	------------------------------

2. Documentation et qualité des données

Les jeux de données seront déposés dans l'entrepôt DataSud. Les métadonnées seront remplies selon le format Dublin Core utilisé par défaut dans cet entrepôt.

Les documents seront accompagnés de fichiers de métadonnées décrivant leur origine, leur mode de production et d'utilisation.

Les développements seront documentés et disposeront d'un fichier readme.txt contenant les instructions d'installation. Le tout sera déposé dans un Github et éventuellement aussi sur DataSuds.

Les rapports scientifiques du projet décriront les livrables, leur processus de collecte et la manière dont il est possible de les réutiliser.

La règle de nommage de fichiers sera la suivante (les informations entre crochets sont facultatives).

BRIGDE_WPX_DATE_[TITLE]_[VERSION]. extension

Les fichiers validés auront la mention de version VF. Les fichiers non validés au minimum par chaque responsable de Work Package.

Les noms des répertoires suivront la règle décrite ci-dessous.

niveau 1 BRIDGE_WPX_[TITLE]

niveau 2 BRIDGE_WPX_[NUMTASK]_[TITLE]

où NUMTASK est le numéro de la tâche décrite dans la description du projet.

Processus de validation collectif des textes produits et du site web . A minima validation par le responsable du Workpackage.

Chaque jeu de données fera l'objet d'un fichier readme.txt ou d'une documentation.

une validation par des juristes sera demandée si la nature des données le justifie.

Les développements informatiques seront précisés dans la prochaine version du DMP

3. Stockage et sauvegarde pendant le processus de recherche

WP1 Gouvernance : Corpus de textes sur la gouvernance des données, Grilles d'analyse, éléments bibliographiques, rapport de synthèse

Alfresco est le système de GED de l'IRD, outil aussi utilisé au Cirad.

WP / Tâche / Nom court (max 20 Caract)	Description du jeu de données	Type de données	Sauvegarde et stockage durant le projet
WP1-Gouvernance	Corpus de textes (documents) validés ou en construction sur la gouvernance des données	Texte PDF 100 Mo max	<i>Alfresco IRD</i>
WP1-Gouvernance	Grilles d'analyse originales (texte format tabulaire)	Textes (ODF + PDF) 50 Mo Max	<i>Alfresco IRD</i>
WP1-Gouvernance	Documents de bibliographie	BDD bibliographie zotero 2 00 Mo Max	<i>Zotero standalone puis Zotero groupe</i>
WP1 Gouvernance	Rapport de synthèse	Textes (ODF + PDF) 100 Mo max	Texte partagé et Alfresco

Le projet ne générera pas à priori de données sensibles. Cependant une vérification sera effectuée avant diffusion finale dans l'entrepôt DataSuds.

Les données sont stockées dans une GED sécurisée. (Alfresco).

WP2 Guidelines : Critères et guide pratiques pour étudier le degré FAIR des entrepôts et application aux cas d'études

Alfresco est le système de GED de l'IRD, outil aussi utilisé au Cirad.

WP2 Provide guidelines to the researchers	Guide pratique pour les chercheurs	Alfresco
WP2	Evaluation des critères pour étudier la fairisation des entrepôts	Alfresco
WP2	Roadmap de fairisation des cas d'études	Alfresco

Le projet ne générera pas à priori de données sensibles. Mais il pourra utiliser et donc protéger des documents classés confidentiels issus des établissements de recherche, qui seront exploités pour créer de nouveaux documents. Cependant une vérification sera effectuée avant diffusion finale dans l'entrepôt Dataverse DataSuds de l'IRD.

Les données sont stockées dans une GED sécurisée (Alfresco).

WP3 IT TOOLS : Recommandations pratiques et techniques pour améliorer la FAIRness des entrepôts

Alfresco est le système de GED de l'IRD, outil aussi utilisé au Cirad.

Pendant la durée du projet, les documents produits seront stockés dans l'outil *Alfresco* qui fait partie de l'Intranet de l'IRD, ainsi que sur le site Overleaf pour permettre une écriture formatée des documents techniques (et qui permet d'avoir un suivi de versionning via GitHub).

WP3-IT Tools	Rapport de FAIRness des entrepôts IRD, CIRAD et INRA	Texte (Latex et ODT) 100 Mo max	Overleaf GitHub
WP3-IT Tools	Recommandations pratiques pour améliorer la FAIRness des entrepôts	Texte (Latex et ODT) 100 Mo max	Overleaf GitHub

Le projet ne générera pas à priori de données sensibles. Mais il pourra utiliser et donc protéger des documents classés confidentiels issus des établissements de recherche, qui seront exploités pour créer de nouveaux documents. Cependant une vérification sera effectuée avant diffusion finale dans l'entrepôt DataVerse DataSuds. Les données sont stockées dans un système de Gestion Electronique de Document sécurisé (outil *Alfresco*), ainsi que sur le site Overleaf pour permettre une écriture formatée des documents techniques (et qui permet d'avoir un suivi de versioning via GitHub).

WP3 autocomplétion : Widget d'autocomplétion pour améliorer les métadonnées

Alfresco est le système de GED de l'IRD, outil aussi utilisé au Cirad.

Les documents produits seront stockés dans l'outil *Alfresco* de l'IRD pendant la durée du projet. Les codes seront entreposés dans *GitHub* et documentés dans *Overleaf*.

WP3-IT Tools	Widget d'autocomplétion pour améliorer les métadonnées	Code Javascript 10 Mo	GitHub
--------------	--	--------------------------	--------

Le projet ne générera pas à priori de données sensibles. Cependant une vérification sera effectuée avant diffusion finale dans l'entrepôt DataVerse DataSuds.

Les données sont stockées dans une GED sécurisée (*Alfresco*), ainsi que sur le site Overleaf pour permettre une écriture formatée des documents techniques (et qui permet d'avoir un suivi de versioning via GitHub).

4. Exigences légales et éthiques, codes de conduite

Il n'y aura pas de traitement de données à caractère personnel dans le contexte du projet Bridge. Cependant une vérification sera effectuée avant diffusion finale des données dans l'entrepôt DataVerse DataSuds.

Les droits de propriétés intellectuelles sont partagés par les trois partenaires. Toutes les productions du projet seront en accès ouvert (*open access, open source*). Par exemple les codes produits seront sous licence CeCill-C compatible GPLv3, et les textes sous CC-BY compatible Etalab.

Ne s'applique pas à ce projet.

Mais une analyse et éventuellement une mise à jour du DMP aura lieu pour vérifier ces questions.

5. Partage des données et conservation à long terme

Lorsqu'elles seront finalisées et au plus tard à la fin du projet, les données seront diffusées sous licence CC-BY dans l'entrepôt de données DataVerse DataSuds, qui garantit une pérennisation des données sur le long terme. Un embargo ne se justifie pas. De même les codes seront distribués sous licence CeCill-C compatible GPLv3.

Les données seront déposées dans l'entrepôt de données DataVerse DataSuds. Les codes informatiques dans *GitHub* et dans Software Heritage ou éventuellement dans DataVerse DataSuds. Les documents de type analyse ou guidelines seront déposés dans les archives ouvertes de chaque partenaire, leur code source (documentation technique) archivée dans DataVerse DataSuds et sur Overleaf.

Il n'y a pas de méthode spécifique. Les vocabulaires qui seront exploités par le code sont propres à chaque organisme et décrits dans un langage standardisé (SKOS). Tout utilisateur pourra utiliser ses propres vocabulaires. Les développements seront codés dans des langages très utilisés (*JavaScript*) et indépendants d'outils propriétaires.

Des identifiants *DOI* seront attribués par l'entrepôt DataVerse DataSuds, ou directement par l'IRD si nécessaire pour des documents spécifiques.

Pour le code informatique : identifiant attribué par *Software Heritage et lien GitHub*.

6. Responsabilités et ressources en matière de gestion des données

Chaque responsable de workpackage listé ci-dessous est responsable de la gestion des données durant le projet et de la qualité des données finales.

Luc Decker - IRD Administrateur de l'entrepôt DataVerse DataSuds assurera le dépôt et la gestion des données (et métadonnées) dans l'entrepôt DataVerse DataSuds.

François Sabot IRD responsable du WP3 et Luc Decker s'assureront du dépôt des composants logiciels dans Github. Les données seront examinées et validées par un membre désigné de chaque institution partenaire avant diffusion finale.

Le PGD sera mis à jour régulièrement : deux versions mises à jour seront fournies, à mi-projet et en fin de projet.

Pascal Aventurier, responsable scientifique du projet, assurera le suivi du plan de gestion des données.

Il n'y a pas de budget spécifique alloué pour la "fairisation" des données produites par le projet.

Tous les participants au projet Bridge sont des spécialistes des données FAIR qui auront à coeur de produire des données faciles à trouver, accessibles, interopérables et réutilisables.