

Projet Living Data : Proposition d'héberger un stage de sauvetage de données

Section A : Informations professionnelles

(1) Votre nom : _____

Votre organisation : _____

(2) Qui est le ou la propriétaire initial-e des données et à quelle organisation cette personne est-elle affiliée? Si vous n'êtes pas propriétaire des données, veuillez expliquer pourquoi le ou la propriétaire initial-e des données ne postule pas (par exemple, décès), ainsi que votre intérêt et votre engagement quant aux données.

(3) Expliquez pourquoi vous ou votre organisation ne pouvez pas archiver ces données sans la participation du projet Living Data. Par exemple, vous pouvez manquer de formation spécialisée en gestion de données et en recherche reproductible ou vous pouvez manquer de temps ou de ressources (par exemple, vous pouvez être proche de la retraite ou le propriétaire des données d'origine peut être décédé ou absent).

Section B : Les données

(4) Années au cours desquelles les données ont été collectées : de _____ à _____.

(5) Couverture spatiale de l'ensemble de données (s'il s'agit d'un seul emplacement, indiquez-le; sinon, veuillez expliquer la zone sur laquelle les données ont été collectées).

(6) En termes généraux, quels types de variables ont été collectés dans cet ensemble de données (géolocalisation, température, abondance, diversité, croissance, etc.)?

(7) Quel était le principal objectif de la collecte de cet ensemble de données? Veuillez indiquer dans votre réponse s'il s'agit de données d'échantillonnage ou de données expérimentales.

(8) Pourquoi pensez-vous qu'il est important de préserver ces données pour les générations futures de scientifiques? Les raisons pourraient par exemple inclure : des données historiques appropriées pour établir des bases de référence pour les changements actuels ou futurs; des séries chronologiques utiles à la compréhension des processus

dynamiques en écologie; une expérience classique qui a eu de profondes répercussions sur la théorie; des données sur des espèces ou des écosystèmes de grande importance pour des raisons biologiques, culturelles ou commerciales.

(9) Les données à sauver sont actuellement dans ce format (cochez tout ce qui s'applique) :

- Feuilles de calcul Excel
- Feuilles de calcul dans un autre format (précisez : _____)
- MS Access
- Base de données SQL
- Fiches techniques papier
- Rapports publiés
- Autres : (précisez : _____)

(9a) Est-ce que de la saisie de données est nécessaire? Non Oui

Si oui, indiquez l'importance de la saisie à effectuer :

Section C : Le ou la stagiaire

(10) Le ou la stagiaire doit avoir les compétences suivantes en ce qui a trait aux données (cochez tout ce qui s'applique) :

- Programmation (énumérez les langages : _____)

- Géospatial (par exemple ArcGIS, QGIS, progiciels géospatiaux R, etc.)
- Base de données relationnelle (par exemple, MySQL, etc.)
- Statistiques

(11) Des connaissances dans ces domaines sont nécessaires ou souhaitables pour le stage (cochez tout ce qui s'applique) :

- Entomologie / écologie des insectes
- Ornithologie / écologie aviaire
- Mammalogie / écologie de la faune
- Foresterie / écologie forestière
- Pêche / écologie marine
- Agriculture / agroécologie
- Botanique / écologie végétale
- Limnologie / écologie aquatique
- Écologie des populations
- Écologie des communautés
- Écologie du paysage
- Biologie du changement climatique
- Biologie de l'évolution
- Phylogénétique
- Génétique des populations
- Génomique
- Autres : _____

(12) Quelles sont les contraintes, s'il y en a, liées à l'emplacement physique du stage?

Le ou la stagiaire sera basée dans notre organisation, située à :

Ce stage peut se faire entièrement virtuellement

Autre (décrire) :

(13) Le ou la stagiaire devra maîtriser (cochez tout ce qui s'applique) :

Anglais :

	Lecture	Écriture	Oral
Requis			
Préfééré			

Français :

	Lecture	Écriture	Oral
Requis			
Préfééré			

(14) Veuillez énumérer toutes les autres exigences relatives à la personne stagiaire qui n'ont pas été couvertes.

Section D : Le déroulement du stage

(15) Qui, dans votre organisation, sera principalement responsable du mentorat de la ou du stagiaire? Cette personne doit prendre les devants en fournissant les données et en répondant aux questions sur les données et la collecte de données, et en rencontrant le ou la stagiaire toutes les semaines.

(16) Si le ou la stagiaire sera basée dans votre organisation et devra se déplacer pour ce faire, votre organisation est-elle en mesure de contribuer aux frais de voyage et d'hébergement? Le projet Living Data Project ne dispose que de fonds limités pour ces coûts.

(17) En plus de mettre en pratique ses compétences en gestion des données et en archivage, de quels autres avantages potentiels bénéficiera le ou la stagiaire en participant à ce stage? Nous souhaitons particulièrement que nos stagiaires découvrent des lieux de travail en dehors des universités, et qu'ils et elles puissent participer à l'analyse des données, aux publications et à la sensibilisation du public.

Section E : Résultats

(18) Dressez la liste des résultats espérés du stage (productions). Ceux-ci feront partie de l'accord formel avec la personne stagiaire, veuillez donc indiquer si chacun des résultats est requis ou facultatif. Cela pourrait inclure : la création d'un plan de gestion des données pour l'archivage futur des données, la détermination des variables à archiver, la combinaison de plusieurs sources de données, l'exécution d'un contrôle qualité sur les données, la création de métadonnées pour le projet, la publication des données dans le cadre du processus d'archivage (*data paper*). Remarque : tous les projets doivent obligatoirement avoir comme résultat le dépôt des données dans un référentiel public et la rédaction d'une brève description du chercheur ou de la chercheuse et de l'ensemble de données qui sera rendue publique sur le site web de l'ICEE.

(19) Avez-vous un référentiel de données préféré (ou mandaté par l'employeur) pour les données à archiver? Si tel est le cas, veuillez l'indiquer ici (sinon laissez vide, et nous pourrions décider ensemble) :

Nous acceptons tout référentiel de données qui utilise les meilleures pratiques en matière de stockage permanent et d'accès ouvert aux données. Un registre global de ces référentiels peut être trouvé sur r3data.org (veuillez nous contacter si votre référentiel préféré ne figure pas sur cette liste). Les référentiels de données populaires dans les domaines de l'écologie, de l'évolution et des sciences de l'environnement incluent : KNB, GBIF, Environmental Data initiative, Genbank, Pangea, Sequence Read Archive. Les référentiels de données populaires pour les données scientifiques ou académiques en général incluent : Zenodo, Dryad, Figshare et (Scholars Portal) Dataverse. Les organismes gouvernementaux fédéraux sont généralement obligés de s'assurer que les données archivées dans un référentiel peuvent être trouvées par l'entremise du portail du gouvernement ouvert.

(20) Avez-vous des inquiétudes concernant la mise à disposition publique de ces données? Veuillez nous le faire savoir, afin que nous puissions travailler avec vous sur des solutions possibles. Par exemple, les préoccupations concernant une utilisation abusive potentielle des données peuvent parfois être traitées en codant dans les données toute information qui devrait être transmise aux futurs utilisateurs et utilisatrices (par exemple, des changements dans la méthode de collecte ou la fiabilité des données). Les préoccupations concernant les emplacements d'espèces menacées ou ayant une valeur commerciale peuvent être résolues en diminuant la résolution spatiale. Les préoccupations concernant le fait de laisser du temps pour terminer une publication avant que les données ne soient rendues ouvertes pourraient être résolues par la fonction d'embargo du référentiel de Pangaea, ou par la politique d'embargo sans questions d'un an de Dryad et Figshare. Les préoccupations concernant l'utilisation commerciale des données peuvent être résolues par la licence [Creative Commons](#) appropriée (la licence CC NC empêche l'utilisation commerciale).