

Partage et valorisation des données

Questions juridiques
Modalités de partage

Frédéric de Lamotte <https://orcid.org/0000-0003-4234-1172>

Paulette Lieby <https://orcid.org/0000-0002-9289-9652>

Partage et valorisation des données

1. Questions juridiques

Frédéric de Lamotte <https://orcid.org/0000-0003-4234-1172>

Paulette Lieby <https://orcid.org/0000-0002-9289-9652>

avec l'aide de Lionel Maurel <https://orcid.org/0000-0002-8667-5900>

En résumé

- Définition « données de la recherche »
- Présentation du cadre juridique

Les données de la recherche 1.



INSTITUT FRANÇAIS DE BIOINFORMATIQUE



- définition «standard» : « *des enregistrements factuels (chiffres, textes, images et sons), qui sont utilisés comme sources principales pour la recherche scientifique et sont généralement reconnus par la communauté scientifique comme nécessaires pour valider les résultats de la recherche* » (OCDE 2007)
- néanmoins, définitions moins restrictives :
 - données d'observation, données expérimentales, données computationnelles ou de simulation, données dérivées ou compilées, et données de référence (au delà du «factuel», INIST)
 - aussi, les données ne sont pas seulement des données «*nécessaires à la validation des résultats*» :
 - typiquement, en bioinformatique, les données produites sont plus nombreuses que celles strictement nécessaires pour valider un résultat
 - ces données gagnent en valeur précisément si elles peuvent être partagées

Les données de la recherche 2.

- d'ailleurs la dernière mise à jour du document OCDE (2021) sépare bien les données de la recherche et
 - *« autres objets numériques pertinents au regard de la recherche financés sur fonds publics : métadonnées, algorithmes, flux de travail, modèles et logiciels [...] utilisés dans le cadre de travaux de recherche et de développement »*
- pour ensuite toujours référer aux
 - « données de la recherche et autres objets numériques »*
 - dans presque tout le document...
- bref, la définition est large et surtout non limitative
 - voir [Ouverture des données de recherche](#)

Ouverture des données

Exercice

- quel est le statut des données par rapport à leur ouverture ?

voir <http://scrumblr.ca/OpenScience>

Quelles obligations de partage des données ?

- *a priori**, les données issues d'une activité de la recherche sont soumises à une obligation de partage

Les données sont soumises à un principe d'ouverture par défaut et de libre utilisation (Loi République numérique 2016 LPRN)

*«*a priori*» : l'objet d'une partie de ce module est de définir les contours de cette obligation de partage

Les grands principes

Les données de la recherche sont des informations publiques :

- elles sont soumises à un principe d'ouverture par défaut et de libre utilisation (*Loi Lemaire - LPRN 2016*)
- elles sont soumises à un principe de gratuité (*Loi Valter 2015*)
- elles sont protégées contre les risques d'accaparement

Les données de la recherche sont des informations publiques

retour sur la définition

- *Les « documents administratifs » visés par la LPRN sont tous les documents quels que soient leur date, leur lieu de conservation, leur forme et leur support qui sont produits ou reçus, dans le cadre de leur **mission de service public, par l'Etat, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission** (cf. EPICs)*
- qualificatif important : on parle ici des documents **achevés**

Les données de la recherche sont protégées contre les risques d'accaparement

- quels que soient les droits cédés à un éditeur, la partie concernant les données est nulle si elle va contre l'ouverture des données

Les données de la recherche sont soumises à un principe de gratuité

- seule une liste fermée d'administrations peuvent encore valablement fixer des redevances de réutilisation :
IGN, Météo France
- mais une articulation est possible avec le dépôt de brevets et d'autres formes de valorisation

Principe d'ouverture par défaut et de libre utilisation (LPRN)

- la LPRN supprime le principe du régime dérogatoire où chaque établissement devait déterminer les modalités de la réutilisation des données
- il y a des protections possibles pour certains types de données, certains cas
 - il s'agit du **principe de l'exception**, fixé dans la loi et par la commission européenne :

les données sont aussi ouvertes que possible et aussi fermées que nécessaire.

« aussi ouvert que possible ; aussi fermé que nécessaire » : les principes FAIR

- les principes FAIR sont un *ensemble de principes directeurs* visant à rendre les données **trouvables, accessibles, interopérables et réutilisables**
- ces principes fournissent des orientations pour l'ouverture des données scientifiques et sont pertinents pour toutes les parties prenantes de l'écosystème numérique



Exception : la propriété intellectuelle appartenant à des tiers 1.

- le **droit d'auteur**, pour les œuvres de l'esprit
 - les **publications scientifiques** : les chercheurs sont considérés comme des tiers par rapport à l'administration. Ils décident de la manière de diffuser, de partager leur œuvre.
 - l' **image**, qui sous certaines conditions d'originalité, peut aussi être considérée comme une œuvre :
 - la ligne de partage est fixée par l'**originalité**
 - les **logiciels**
 - etc...

Exception : la propriété intellectuelle appartenant à des tiers 2.

- **les projets partenariaux et les droits des tiers**
 - dans le cas de **recherche partenariale** (acteurs publics et privés), quand une entreprise privée partage des données
 - l'accès, la transmission, etc. sont fixés par des règles, par notamment les accords de consortium et les contrats
- néanmoins, la réutilisation des données est libre si :
 - elles sont issues d'une activité de recherche financée **au moins pour moitié** par des fonds publics

Exception : la protection des données personnelles et de la vie privée

- L'ancienne *Loi Informatique et Liberté* devenue en 2018 le [RGPD](#) ordonne que les données à caractère personnel ne **doivent pas** être soumises à l'ouverture
 - Il est **obligatoire** de recueillir le **consentement** et/ou de **pseudonymiser** ou **anonymiser** les données.
 - note : anonymisation différente de pseudonymisation -- la 1re ne tombe pas sous le RDPD, mais la 2nde oui
- Il existe néanmoins des **dérogations pour la recherche**, notamment pour le partage sécurisé des données qui permet de rester dans le cadre des principes FAIR
- s'applique aux personnes physiques

Exception : les données sensibles, les secrets protégés

- Données sensibles qui ne sont pas des données personnelles
 - données de biodiversité, etc...
- Secret médical, secret des affaires, secret des procédés, secret militaire, ...

sortent de l'ouverture par défaut

Exception : les bases de données

Elles sont susceptibles de recevoir une double protection :

- Un droit particulier, appelé droit "*sui generis*" protège le contenu de la base de données, c'est-à-dire l'ensemble des données qu'elle contient.

Le droit *sui generis* peut protéger les bases de données, même à défaut d'originalité de celles-ci. L'objectif de la protection par le droit *sui generis* est de protéger les investissements réalisés dans le secteur des bases de données et d'empêcher la reprise des bases de données par des concurrents.

- Le droit d'auteur protège la structure de la base de données, si elle est originale.

Un mot sur les licences 1.

- les licences pour encadrer le partage et la réutilisation des données
 - techniquement, suite à la LPRN, une licence n'est pas nécessaire là où aucun droit ne s'exerce (par exemple, si le droit d'auteur ne s'exerce pas)
 - néanmoins, il est (fortement) recommandé d'en utiliser une dans tous les cas
-  **pour clairement afficher les droits afférents**

Un mot sur les licences 2.

- typiquement les licences Creative Commons
 - toutes n'ont pas de jurisprudence spécifique en France mais leur validité n'est pas remise en cause
- LPRN : une liste de licences [Licences - data.gouv.fr](https://data.gouv.fr/licences)
 - etalab : équivalent CC-BY 2.0
 - CC-BY-SA homologuée ; d'autres en cours
 - Note : les licences CC 4.0 sont internationales et ne nécessitent pas/plus de transcriptions nationales

Un mot sur les entrepôts

Le cas des données de la santé :

- le RGPD n'autorise pas le transfert de données hors EU
 - concerne les données sensibles, mais on peut se poser la question des données en général
- les entreprises américaines sont soumises au Cloud Act (Clarifying Lawful Overseas Use of Data Act). Votée en mars 2018, cette loi permet aux États-Unis d'accéder plus facilement aux données stockées sur des serveurs situés hors des États-Unis.
- mais : le cas du Health Data Hub (serveurs Microsoft aux Pays-Bas)
- 14 octobre 2020 : [décision](#) du conseil d'état « *Aucune donnée personnelle ne peut être transférée en dehors de l'Union européenne dans le cadre du contrat conclu avec Microsoft* »
- et pourtant : [Google et Thales s'allient dans le « Cloud de confiance »](#) (d'accord, on ne parle pas de données sensibles, néanmoins...)

Des principes sur lesquels s'appuie la politique publique de science ouverte

- [Deuxième Plan national pour la science ouverte](#)
- [Ouverture des données de la recherche. Guide d'analyse du cadre juridique en France](#)
- [Rapport Bothorel sur la politique publique de la donnée, des algorithmes et des codes sources](#)

Un récapitulatif (non exhaustif) de la législation applicable

- en France
 - Loi Lemaire - LPRN 2016 (*ouverture des données*)
 - Loi Valter 2015 (*gratuité*)
 - RGPD 2018 (*protection des données sensibles*)
 - plus loi Jardé ? encadre expérimentations impliquant la personne humaine
- en Europe
 - [Directive on open data and the re-use of public sector information, Open Data Directive](#), 16 July 2019
 - [Regulation on a framework for the free flow of non-personal data in the EU](#) (EU) 2018/1807 (*free movement of non-personal data*)
 - GDPR 2018
- notons la prééminence de la LPRN

Logigramme / aide à la décision

- [INRAe](#)

Aides à la décision

- [CIRAD](#)
- [INRAe](#)
- [Pasteur](#)
- [Ponts et Chaussées](#)
- [Ouvrir la science](#)
- un outil : [DAISY](#) (ELIXIR-LU) --pour les données soumises au RGPD

Bibliographie et ressources

- [Aspects juridiques et éthiques – DoRANum](#)
- [Ouverture des données de la recherche : de quoi parle-t-on ?](#)
- [À qui appartiennent les données ?](#) excellent résumé de la présea de Lionel Maurel
- [Les principes FAIR – DoRANum](#)
- [Ouverture des données de la recherche. Guide d'analyse du cadre juridique en France](#) (my first port of call)
- [En route vers la science ouverte : Gérer les données de sa recherche](#)
- [OCDE : Recommandation du Conseil concernant l'accès aux données de la recherche financée sur fonds publics](#)
- [Guide d'application de la loi pour une République numérique](#)

Partage et valorisation des données

2. Modalités de partage

Frédéric de Lamotte <https://orcid.org/0000-0003-4234-1172>

Paulette Lieby <https://orcid.org/0000-0002-9289-9652>

Pendant et après le projet

 **se rappeler les questions du PGD**

- Quel est le potentiel de réutilisation?
- Quels jeux de données sont partagés ?
- Quand ? (durée d'un éventuel embargo)
- Oú ? dans quel entrepôt de données ?
- Comment ?
 - quelles modalités ?
 - quelles licences ?
 - pour quels publics ?

 **ici nous nous préoccupons du partage après la fin du projet**

Quelles obligations de partage ?

« Aussi ouvert que possible, aussi fermé que nécessaire »

Considérer

- les restrictions
 - présence de données sensibles, personnelles, de santé, issues de partenariats avec le privé
- un embargo
 - pour publication, dépôt de brevet, exploitation
- les limites de réutilisation
 - utilisation commerciale, accès sur demande seulement



ceci doit être renseigné et justifié dans le PGD

Obligations de partage

-- spécificités : les bailleurs 1.

- **H2020** – Grant Agreement art. 29.3 :
 - *The beneficiaries **must deposit in a repository***
 - *the data needed to validate the results presented in scientific publications as soon as possible*
 - *other data as specified in the DMP.*
- **H2020** - GA art. 29.2 : *Open access to publications mandatory*
 - obligation de rendre accessible vos articles avant 6 mois (max 6 pour LRPN)
 - publier dans une revue en libre accès
 - choisir une revue qui autorise le dépôt de l'article (ou version acceptée) dans une archive ouverte (voir diapo suivante)

Obligations de partage -- spécificités : les bailleurs 2.

- **ANR**

- demande que toutes les publications consécutives aux projets qu'elle finance, soient déposées en texte intégral dans une archive ouverte, soit directement dans HAL soit par l'intermédiaire d'une archive institutionnelle locale
- suivant en cela l'article 30 de la loi «Pour une République numérique»
- idem pour **H2020**

Quels jeux de données partager ?

- Données ayant un potentiel de réutilisation
 - Nouvelles analyses, nouvelles questions de recherche
 - Méta-analyses, nourrir des modèles
 - Changement d'échelle : spatiale, temporelle, analyses stat.
 - Développement commercial, de services...
 - etc.
- Données utiles
 - jeux de données contrôles, lots témoins
- Données présentant un intérêt pour certains publics
- Données dont le coût d'obtention est élevé

Quels jeux de données partager ?

-- exemples

- données environnementales, climat,
- gestion des territoires
- santé publique
- genre
- favorisant la participation citoyenne
- ...

Potentiel de réutilisation des données



estimer la valeur de ses données

- Données rares ou uniques
 - expérimentation impossible à répéter
 - groupes difficilement accessibles
 - phénomènes rares
- Données à forte valeur scientifique
 - données de référence
 - reproduction difficile ou coûteuse
 - ayant un grand intérêt pour certains publics (ex: société civile, pays du Sud)
- Données ayant une valeur économique
 - perspectives d'application, développement commercial
- Données ayant une valeur environnementale

Identifier les futurs utilisateurs

- Communautés scientifiques
- Enseignants
- Décideurs
- Secteur privé, créateurs de start-up
- ONGs, associations internationales influentes
- Journalistes
- Grand public

Quels jeux de données ne PAS partager ? *(ou du moins se poser la question)*

- Données issues de partenariats privés
- Données sensibles :
 - concernant des espèces protégées ou envahissantes,
 - données cliniques, issues d'expérimentations animales
 - données personnelles (soumises au RGPD)
 - issues de ressources biologiques du Sud (réglementation APA aka Protocole de Nagoya)
- Données stratégiques que vous souhaitez exploiter :
 - identification de marqueurs génétiques, d'arômes
 - création d'une appli, d'une base de données originale
- Jeux de données contenant des données préexistantes
 - produits par d'autres
 - sous licences non ouvertes

Quand partager les données ?

- Dépend
 - du bailleur (H2020, ANR, fondation Gates, ...)
 - de politiques institutionnelles ou nationales
 - de politiques de certains partenaires
 - de la revue de publication
- Après avoir
 - exploité vos données
 - publié vos résultats de recherche
 - mis en forme vos données et métadonnées
 - anonymisé vos données
 - obtenu l'accord de tous vos partenaires

Quand partager ?

- Libre accès aux articles
 - Loi Lemaire (LPRN) : max 6 mois après publication
 - H2020 : idem
 - B&M Gates : immédiatement
- Pour le PGD
 - 6 mois après début du projet: H2020 et ANR
 - Dès début du projet : NIH
- Exemples
 - *Les données seront déposées dans le Dataverse INRAe et seront accessibles après un embargo d'un an (pour publier 2 articles)*
 - *Les données seront disponibles sur demande dès l'année 3 puis seront accessibles sur ENA dès la fin du projet*

Où partager les données ? 1.



Dans un entrepôt de données

→ **visibilité, citabilité, préservation**

→ **optimise les possibilités de réutilisation**

Un entrepôt

- adapté aux données et aux métadonnées
- porté par une communauté scientifique/reconnu dans discipline
- institutionnel
- correspondant aux publics visés
- délivrant un identifiant numérique pérenne et unique
- gratuit (la plupart) ou coût raisonnable

Où partager les données ? 2.

Un entrepôt

- en accord avec le bailleur, l'institution, les partenaires, la revue de publication
 - Exemple : Springer : Where a widely established research community expectation for data archiving in public repositories exists, **submission to a community-endorsed, public repository is mandatory.**

Pourquoi des entrepôts

- Visibilité dans les entrepôts → visibilité dans les moteurs de recherche et sites de data search
 - → Entrepôt qui possède un potentiel de réutilisation de vos données
 - → Entrepôt qui permet le suivi de l'utilisation de vos données
- Sécurité : trois niveaux de certification : CoreTrustSeal, NestorSeal, ISO16363
- Pérennité d'accès (sous condition)

Entrepôts et accès aux données : Quelles licences ?

- privilégiez une licence largement utilisée et compatible avec les autres licences existantes, afin de faciliter la compilation de vos données avec d'autres données mises à disposition sous d'autres licences
- choisissez la licence en tenant compte du potentiel de vos données et des restrictions appliquées selon ce choix
 - moins elle est restrictive: plus grand est le potentiel de réutilisation
 - etalab/CC-BY 2.0 ou CC-BY 4.0 l'une des plus répandues, avec CC-BY-SA 4.0
 - CC-BY-ND (sans modification) et CC-BY-NC (sans usage commercial) sont à éviter
- visiter <https://creativecommons.org/choose/> pour une aide à la sélection
- <https://datapartage.inrae.fr/Partager-Publier/Choisir-une-licence>

Quels entrepôts de données ?

Exercice

- connaissez vous des entrepôts de données dans votre domaine ?
 - pour trouver des données fiables ?
 - pour déposer des données ?

voir <http://scrumblr.ca/entrepots>

considérer scrumblr.ethibox.fr

Les entrepôts de données



- **Institutionnels** Dataverse



- **Europe** Zenodo, B2Share



- **Généralistes** Figshare, Dryad



- **Editeurs** Oxford Univ Press (GigaDB) ; Ubiquity Press (Dataverse)

- [Recherche Data Gouv, the federated national research data platform](#)

- **Thématiques**



- GBIF (Global Biodiversity Information Facility)

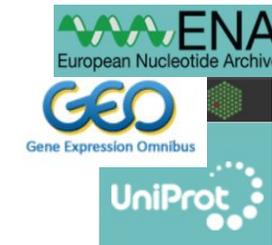
- KNB (Knowledge Network for biocomplexity Initiative)



- Pangaea, SEANOE

- Movebank, WormBase, ViPR, MycoBank, ComBase, FLOW

- GenBank, Barcode of Life Data Systems, UniProt, Intact



- Dataverse, ICPSR, DataFirst, Quetelet, beQuali



Trouver des entrepôts

[re3data](http://re3data.org)

re3data.org

Filter

Subjects ▾
Content Types ▾
Archived data (164)
Audiovisual data (135)
Configuration data (21)
Databases (135)
Images (300)
Networkbased data (27)
Plain text (266)
Raw data (248)
Scientific and statistical data formats (356)
Software applications (96)
Source code (66)
Standard office documents (348)
Structured graphics (173)
Structured text (205)
other (205)
Countries ▾
AID systems ▾
API ▾
Certificates ▾
Data access ▾
closed (50)
embargoed (112)
open (468)
restricted (286)
Data access restrictions ▾
feeRequired (32)
institutional membership (33)
other (136)
registration (176)
Database access ▾
Database access restrictions ▾
Database licenses ▾
Data licenses ▾
Data upload ▾

life science

← Previous **1** 2 3 4 5 6 7 ... 22 Next →

Found 537 result(s)

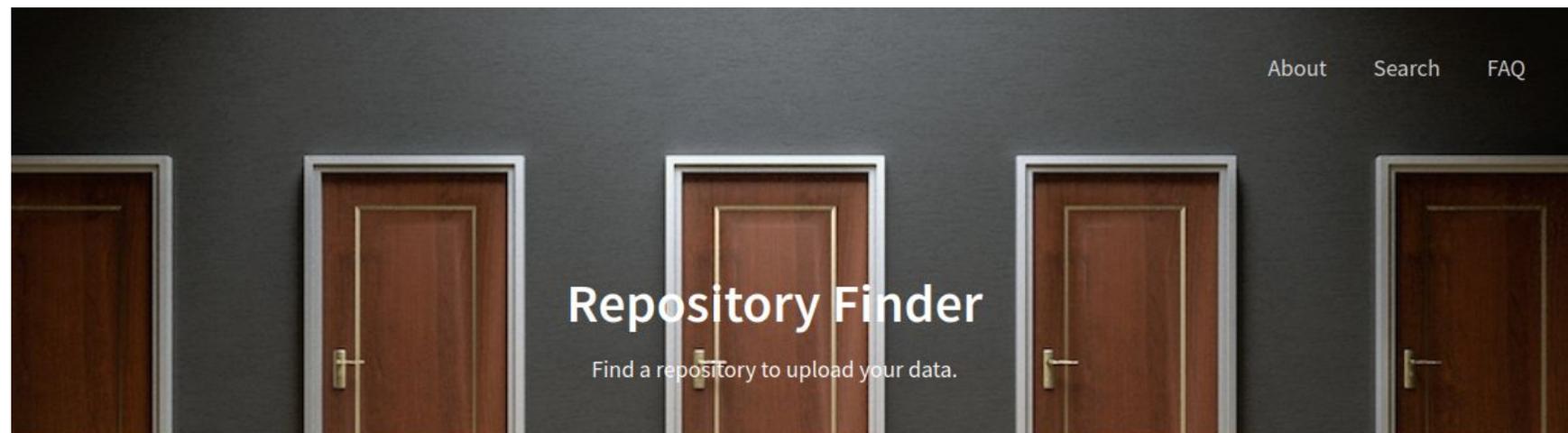
Life Science Database Archive
LSDB Archive
Subject(s) Basic Biological and Medical Research Biology Life Sciences
Content type(s) Databases Images Raw data Plain text Structured text
Country Japan
The Life Science Database Archive maintains and stores the datasets generated by life scientists in Japan in a large number of datasets by metadata (description of datasets) in a unified format, and to access and download the datasets with ease in public and private institutions, and thereby supports further contribution of each research to life science.

SciLifeLab Data Repository
SciLifeLab figshare
Subject(s) Life Sciences Natural Sciences
Content type(s) Scientific and statistical data formats Standard office documents
Country Sweden
This repository accepts data from life science researchers and service units in Sweden. The repository is operated by SciLifeLab in Sweden. This repository replaces NBIS DOI repository: <http://doi.org/10.17616/R3CW52>

Brunel figshare
Brunel research data repository & registry
Subject(s) Humanities and Social Sciences Engineering Sciences Geosciences

Trouver des entrepôts

DataCite Repository Selector



Repository Finder, a pilot project of the [Enabling FAIR Data Project](#) led by the American Geophysical Union (AGU) in partnership with DataCite and the Earth, space and environment sciences community, can help you find an appropriate repository to deposit your research data. The tool is hosted by DataCite and queries the re3data registry of research data repositories.

As part of the [FAIRsFAIR project](#), which aims to supply practical solutions for the use of the FAIR data principles throughout the research data life cycle, the Repository Finder is extended to query for repositories relevant to FAIRsFAIR Project.

Search re3data for a repository to upload your data

Exercice : trouver l'entrepôt adapté à ses données

En fonction de vos jeux de données à partager

- **Recherchez un entrepôt, à partir de :**

- Re3data (<https://www.re3data.org/>)
- Revues scientifiques dans vos domaines
- Votre communauté scientifique
- Sites de recherche de données

DataCite, WoS (Data Citation Index), Google data search,...

Identifiez 1 entrepôt adapté

+ les licences proposées

+ modalités de dépôt (standard métadonnées, format, taille fichiers...)

R23data : Trouver les entrepôts pertinents

Repository details

TRY

General Institutions Terms Standards

Name of repository	TRY
Additional name(s)	Plant Trait Database
Repository URL	https://www.try-db.org/TryWeb/Home.php
Subject(s)	Plant Sciences Biology Life Sciences
Description	This database is a global archive and describes plant traits from throughout the globe. TRY is a network of vegetation scientists headed by DIVERSITAS, IGBP, iDiv, the Max Planck Institute for Biogeochemistry and an international Advisory Board. About half of the data are geo-referenced, providing a global coverage of more than 8000 measurement sites.
Contact	https://www.try-db.org/TryWeb/Contact.php
Content type(s)	Standard office documents Structured graphics images Archived data
Keyword(s)	botany plant morphology plant anatomy plant phenology plant physiology characteristics of plants ecosystem ecology biology biogeography terrestrial vegetation biodiversity
Repository size	7 million trait records; 140.000 plant taxa
Repository type(s)	disciplinary
Mission statement for designated community	https://www.try-db.org/TryWeb/About.php
Research data repository language(s)	eng
Data and/or service provider	dataProvider

TRY

General Institutions **Terms** Standards

Policies (1)

Policy Name	TRY intellectual property guidelines
URL	https://www.try-db.org/TryWeb/TRY_Intellectual_Property_Guidelines.pdf

Database access

Type of access to research data repository	open
--------------------------------------------	------

Data access (2)

Type of access to data	open
------------------------	------

Type of access to data	restricted
------------------------	------------

Data access restriction type(s)	other
---------------------------------	-------

Data licenses (3)

DataLicense	other
-------------	-------

URL	https://www.try-db.org/TryWeb/TRY_Intellectual_Property_Guidelines.pdf
-----	-------------------------------------------------------------------------------------------------------------------------------------------------------------

DataLicense	other
-------------	-------

URL	https://www.try-db.org/TryWeb/TRY_Policy_Free_Access_Data.pdf
-----	-------------------------------------------------------------------------------------------------------------------------------------------

DataLicense	other
-------------	-------

URL	https://www.try-db.org/TryWeb/TRY_Data_Release_Notes.pdf
-----	---------------------------------------------------------------------------------------------------------------------------------

Data upload (1)

Type of data upload	restricted
---------------------	------------

Data upload restriction type(s)	other
---------------------------------	-------

Data upload license (1)

Data upload license name	Intellectual Property Guidelines for the TRY Initiative
--------------------------	---------------------------------------------------------

Data upload license URL	https://www.try-db.org/TryWeb/TRY_Intellectual_Property_Guidelines.pdf
-------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------

Entrepôts recommandés par les revues



Research Data Policies

- Data policy types
- Data availability statements
- Data policy FAQs
- Journal policies & services
- Recommended repositories list
- Research Data Helpdesk
- Research Data Support

Recommended Repositories

In general, data should be submitted to discipline-specific, community-recognised repository where possible, or to [generalist repositories](#) if no suitable community resource is available.

This list is derived from *Scientific Data's* recommended repository list. The list is also [available in figshare](#) for free reuse, with attribution, by others. *Scientific Data* is an open access data journal published by Springer Nature.

Authors should consult journal information for authors in case of more specific repository recommendations. Some Springer Nature journals, such as *Human Genome Variation*, maintain their own research data repositories.

If an author's preferred repository is not listed below we encourage repository managers to investigate [listing their repositories with re3data.org](#) and/or [FAIRsharing.org](#) in the first instance and, for those that meet the criteria, applying for listing with *Scientific Data*.

For more information browse our [repositories FAQs](#).

Find the best repository for your data

[Biological sciences](#)

[Chemistry and Chemical biology](#)

[Earth, Environmental and Space sciences](#)

[Generalist repositories](#)

[Health sciences](#)

[Materials science](#)

[Physics](#)

[Social science](#)

[Other repositories](#)

Entrepôts recommandés par les revues



Earth, Environmental & Oceanographic Data

Data Repository	How articles and data are linked (learn more)	More Information
Atmospheric Radiation Measurement (ARM) Data Archive	Authors should Include data DOI's In their manuscript, following these citation Instructions >	ARM Data homepage > Data citation Instructions >
BGS GeoScenic	Authors should specify BGS GeoScenic numbers, e.g. <i>GeoScenic: P603281</i> .	BGS GeoScenic homepage >
Carbon Dioxide Information Analysis Center (CDIAC)	Authors should Include data DOI's In their manuscript.	CDIAC homepage > Data submission Instructions > Example article >
Natural Environment Research Council (NERC), including BADC, BODC, EIDC and NGDC.	Authors should Include data DOI's In their manuscript	NERC Data Centres > Example article >
System for Earth Sample Registration (SESAR), registry for International Geo Sample Numbers (IGSN)	Authors should specify IGSN numbers, e.g. <i>IGSN: HRV003M16</i> .	SESAR homepage > Registering samples > Example article >

Engineering & Materials Standards

Data Repository	How articles and data are linked	More Information
ASTM (formerly, American Society for Testing and Materials)	Authors should specify ASTM Identifiers, e.g. <i>ASTM: G63</i> .	ASTM standards homepage > Example article >

Genes & Gene Expression

Data Repository	How articles and data are linked	More Information
Allele Frequency Net Database (AFND)	Authors should specify AFND accession numbers, e.g. <i>AFND: AFND001243</i>	AFND homepage > Submitting data >
ArrayExpress	Authors should specify ArrayExpress accession numbers, e.g. <i>ArrayExpress: E-MEXP-3783</i> .	ArrayExpress homepage > Submitting data > Example article >
GenBank	Authors should specify GenBank accession numbers, e.g. <i>GenBank: BA123456</i> . ScienceDirect displays and visualizes supporting Information using Information from and linking to the repository.	GenBank homepage > Submitting data >
Gene Expression Omnibus (GEO)	Authors should specify GEO accession numbers, e.g. <i>GEO: GSE27196</i> ; <i>GEO: GPL5366</i> ; <i>GEO: GSM9853</i> . ScienceDirect displays supporting Information using Information from and linking to the repository.	GEO homepage > Submitting data > Example article >
Genome Sequence Archive	Authors should specify GSA Identifiers, e.g. <i>GSA: GSA000134</i>	GSA homepage >

Interdisciplinary

Data Repository	How articles and data are linked	More Information
4TU.ResearchData	Authors should Include data DOI's In their manuscript.	4TU.ResearchData homepage > Uploading data > Citing data >
Australian National Data Service	Authors should Include data DOI's In their manuscript.	ANDS homepage > Example article > Registering data > Citing data >
Dataverse	Some Journals have a dedicated Dataverse repository set up for authors to upload their data that belongs with the article. Authors should Include the dataset DOI In the article, e.g. <i>doi:10.7910/DVN/26868</i>	Harvard Dataverse page > Example of the Journal Dataverse repository > Example article >
Mendeley Data	Mendeley Data banners will be shown on ScienceDirect when the repository has data for the article	Mendeley Data homepage > Example article >

Life Sciences

Data Repository	How articles and data are linked	More Information
CryptoDB	Authors should specify CryptoDB accession numbers, e.g. <i>CryptoDB: cgd2_220</i>	CryptoDB homepage > Submitting data >
EMBL-EBI OLS Molecular Interaction Ontology (MI)	Authors should specify EMBL-EBI OLS accession numbers, e.g. <i>EMBL-EBI MI: 0218</i> .	EMBL-EBI OLS homepage > Example article >
FungiDB	Authors should specify FungiDB accession numbers, e.g. <i>FungiDB: NCU06658</i>	FungiDB homepage > Submitting data >
Mycobank	Authors should specify MycoBank accession numbers, e.g. <i>Mycobank: 476</i> .	Mycobank homepage > Submitting data >
NCBI Taxonomy	Authors should specify NCBI Taxonomy accession numbers, e.g. <i>NCBI Taxonomy: 48184</i> .	NCBI Taxonomy homepage > Example article >
PlasmoDB	Authors should specify PlasmoDB accession numbers, e.g. <i>PlasmoDB: PF3D7_0417200</i>	PlasmoDB homepage > Submitting data >
ToxoDB	Authors should specify ToxoDB accession numbers, e.g. <i>ToxoDB: TGME49_239250</i>	ToxoDB homepage > Submitting data >
TriTrypDB	Authors should specify TriTrypDB accession numbers, e.g. <i>TriTrypDB: Tb927.11.3120</i>	TriTrypDB homepage > Submitting data >

Materials Science

Data Repository	How articles and data are linked	More Information
Institut Laue-Langevin (ILL)	Authors should specify Data DOI's In their manuscript.	ILL homepage > ILL data policy >

Model Organisms

Data Repository	How articles and data are linked	More Information
-----------------	----------------------------------	------------------



Comprendre la politique de publication Sherpa Romeo



INSTITUT FRANÇAIS DE BIOINFORMATIQUE



Jisc Digital Resources > Open Access

Sherpa Romeo

About Search Statistics Help Support

Search

Enter a journal title or issn, or a publisher name below:

Journal Title or ISSN

Publisher Name

Open access services

Sherpa Services

Ma

Comprendre la politique de publication Sherpa Romeo



Jisc Digital Resources > Open Access

Sherpa Romeo

About Search Statistics Help Support Us Contact Admin

Genomics, Proteomics and Bioinformatics

Publication Information

Title	Genomics, Proteomics and Bioinformatics [English]
ISSNs	Print: 1672-0229
URL	https://www.journals.elsevier.com/genomics-proteomics-and-bioinformatics/
Publishers	Elsevier [Commercial Publisher]
DOAJ Listing	https://doaj.org/toc/1672-0229
Requires APC	Yes [Data provided by DOAJ]

Publisher Policy

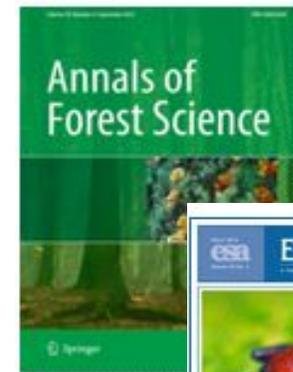
Open Access pathways permitted by this journal's policy are listed below by article version. Click on a pathway for a more detailed view.

Published Version [pathway a]	None CC BY PMC
	Any Repository, Journal Website, +1
OA Publishing	This pathway includes Open Access publishing
Embargo	No Embargo
Licence	CC BY
Publisher Deposit	PubMed Central
Location	Any Repository Named Repository (PubMed Central) Journal Website
Conditions	Published source must be acknowledged Must link to publisher version with DOI

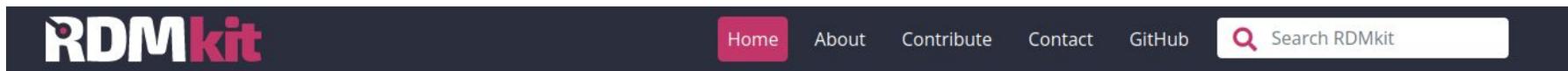


Valoriser les données

- Publier un **datapaper**
- Publier le **PGD**
- Publier un **article de recherche**
- Rédiger une brève pour un **magazine** spécialisé
- Contribuer à un **blog**,



L'outil à ne PAS oublier : RDMkit



- Data life cycle
- Your role
- Your domain
- Your tasks
- Tool assembly
- All tools and resources
- All training resources

Are you working with data in the Life Sciences? Do you feel overwhelmed when you think about Research Data Management?

The ELIXIR Research Data Management Kit (RDMkit) is an online guide containing good data management practices applicable to research projects from the beginning to the end. Developed and managed by people who work every day with life science data, the RDMkit has guidelines, information, and pointers to help you with problems throughout the data's life cycle. RDMkit supports FAIR data — Findable, Accessible, Interoperable and Reusable — by-design, from the first steps of data management planning to the final steps of depositing data in public archives.

The RDMkit organises information into the six sections displayed below, which are interconnected but can be browsed independently.

Data life cycle

Start here to get an overview of research data management. Click on a section of the diagram below to get an introduction to that stage of the data management life cycle.



RDMkit : conserver

Data life cycle ^



Your role ▾

Your domain ▾

Your tasks ▾

Tool assembly ▾

All tools and resources

All training resources

Preserving [Edit me](#)

- [What is data preserving?](#)
- [Why is data preserving important?](#)
- [What should be considered for preserving data?](#)
- [More information](#)

What is data preserving?

Data preservation consists of a series of activities necessary to ensure safety, integrity and accessibility of data for as long as necessary, even decades. Data preservation is indeed more than just data storage and backup, since data can be stored and backed up without being preserved. Data preservation prevents data from becoming unavailable and unusable over time through appropriate activities, such as:

- Ensure data safety and integrity.
- Change the file format (format migration) and update software to make sure that they do not become outdated or obsolete.
- Change hardware and other storage media (such as paper, magnetic tape, etc) to avoid degradation.

RDMkit : partager



Your role ▾

Your domain ▾

Your tasks ▾

Tool assembly ▾

All tools and resources

All training resources

Sharing [Edit me](#)

- [What is data sharing?](#)
- [Why is data sharing important?](#)
- [What should be considered for data sharing?](#)
- [More information](#)

What is data sharing?

Sharing data means making your data known to other people.

You can share your data with collaboration partners in the context of a collaborative research project, or you can publish your data to share it with the global research community and society at large.

It's important to know that data sharing doesn't mean open data or public data. You can choose to share your data with restricted access or even closed access. Moreover, sharing or publishing data is different from publishing a paper or a manuscript in a journal. Here we focused on data (i.e. raw observations and measurements, analysis workflows, code, etc), not on papers or articles.

Data sharing can be done at any time during the research data life cycle but, at the latest, data should be made available at the time of publication of articles that use the data to make scientific conclusions.

Merci !



Commentaires ? Questions ?

Remerciements :

- toute l'équipe de formation @ IFB
- ELIXIR-CONVERGE