

Data sharing : rechercher et partager des données de recherche

Frédérique Flamerie - 2023



Ce contenu est mis à disposition selon les termes de la [Licence Creative Commons Attribution - Partage dans les Mêmes Conditions 3.0 France](#).

Objectifs

- Appréhender concrètement les enjeux liés au partage des données de recherche, du point de vue du producteur de données comme de celui de l'utilisateur de données existantes
- Identifier et savoir utiliser les outils de recherche spécifiques aux données de recherche
- Trouver des données pertinentes pour son thème de recherche
- Evaluer le caractère réutilisable des jeux de données disponibles

Programme

1. Introduction
2. Trouver des données de recherche avec les outils bibliographiques
3. Trouver des données de recherche grâce aux entrepôts de données
4. Trouver des données de recherche pour les réutiliser
5. Politiques de données des éditeurs et des organismes de financement de la recherche



1. Introduction

Terminologie

Enjeux

Modes de partage des données de recherche

Terminologie

→ Définition large OCDE

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 des résultats de recherche.

→ Typologie

- › Données d'observation
- › Données expérimentales
- › Données de références
- › Données de simulation numériques
- › Données dérivées ou compilées

Terminologie

☑ Données d'observation

Capturées en temps réel, souvent uniques, impossibles à reproduire

- Relevés météorologiques
- Images
- Enquêtes sociales
- Fouilles archéologiques

☑ Données de simulation numérique

Générées par des modèles informatiques, souvent reproductibles si le modèle est correctement documenté

- Modèles climatiques
- Modèles économiques

☑ Données expérimentales

Obtenues à partir d'équipements de laboratoires, souvent reproductibles, parfois coûteuses

- Poids biomasse
- Séquences peptide
- Chromatogrammes

☑ Données dérivées ou compilées

Issues du traitement ou de la combinaison de données « brutes », souvent reproductibles, mais coûteuses

- Bases de données compilées
- Fouilles de texte

☑ Données de référence

- Séquences de gènes
- Structures chimiques

Source : Bracco, L., Bouchet-Moneret, F., & Jouneau, T. (2021). Que sont les données de la recherche ?
<https://doi.org/10.5281/zenodo.5495210>

Enjeux

POURQUOI OUVRIR SES DONNÉES ?

POUR VOUS

✓ Visibilité scientifique accrue

Partager ses données et les lier à ses publications augmente le nombre de vos citations : une étude signale jusqu'à 25 % de citations en plus en moyenne.

✓ Traçabilité de vos données

Même si les données ne vous appartiennent pas - elles sont la propriété de l'institution pour laquelle vous travaillez, sauf contrats particuliers avec le privé - les diffuser permet d'y associer votre nom.

✓ Réponse aux exigences de certaines revues

De plus en plus de revues exigent à présent l'accès aux données pour publier un article afin d'en vérifier la méthodologie. Des données publiées accélèrent le processus éditorial.

✓ Augmentation des échanges avec vos pairs

En publiant ou en ouvrant vos données, vous pouvez élargir votre réseau professionnel et nouer de nouvelles collaborations avec d'autres chercheurs.

✓ Préservation de vos données

Le fait de conserver vos données dans un entrepôt n'oblige pas à les partager immédiatement mais vous garantit leur sécurisation.

✓ Réponse aux recommandations institutionnelles

En accord avec le [Plan National pour la Science Ouverte](#), les chercheurs exerçant dans des établissements publics et/ou dont les projets sont financés sur fonds publics sont fortement incités à ouvrir leurs données, sauf exceptions légales.

Exemple : données faisant l'objet de contrats spécifiques public-privé, données du Secret Défense, données soumises au secret professionnel, etc.

POUR LA SOCIÉTÉ

✓ Confiance des citoyens

La transparence de la démarche scientifique augmente la confiance des citoyens dans la recherche : comme elle peut être vérifiée, discutée voire infirmée, elle apparaît ainsi plus fiable.

✓ Utilisation des deniers publics

Les données de la recherche sont considérées comme un bien commun et leur ouverture démontre une bonne utilisation des crédits qui ont financé leur collecte et leur traitement.

✓ Relation entre science et enjeux sociétaux

Plus les chercheurs communiquent et sortent des laboratoires, plus les acteurs politiques et sociaux pourront s'appuyer sur leurs recherches pour faire évoluer la société.

✓ Lutte contre la propagation de fausses découvertes

L'ouverture des données permet l'examen critique, la controverse et la discussion, démarches essentielles à l'identification des pseudo-découvertes.

POUR LA COMMUNAUTÉ SCIENTIFIQUE

✓ Reproductibilité

Elle participe à l'amélioration des connaissances scientifiques en permettant à d'autres de pouvoir reproduire des mesures ou des observations réalisées pendant une expérience.

✓ Lutte contre la déperdition des données

Combien sont oubliées sur des disques durs ? Combien se perdent lors de départs de doctorants ou de chercheurs d'un laboratoire ? Les partager c'est faire en sorte de les sécuriser et leur donner une chance d'être réutilisées. La plupart des licences permettent d'attribuer les jeux de données à leurs auteurs.

✓ Gage de transparence et d'intégrité scientifique

Votre recueil de données, la méthodologie employée pour les traiter et les résultats qui en découlent peuvent être validés par vos pairs et les rendent plus solides.

✓ Gain de temps et d'argent

La réutilisation de données vous permet de vous reposer sur des données déjà discutées et validées par les pairs sans avoir à recommencer la même collecte. Vous pouvez ainsi aller plus loin dans votre recherche en vous appuyant sur des données fiables.

Source : Bracco, L., Bouchet-Moneret, F., & Jouneau, T. (2021). Pourquoi ouvrir ses données ?
<https://doi.org/10.5281/zenodo.5495217>

Enjeux - politique de l'ANR

“Est-ce qu’il existe une obligation à l’ouverture des données ?

Non. En cas de financement le coordinateur ou la coordinatrice s’engage à fournir un plan de gestion de données. Le PGD a pour but de préparer la diffusion potentielle des données mais il ne constitue pas une obligation à l’ouverture. Le principe « aussi ouvert que possible, aussi fermé que nécessaire » est au cœur de la démarche de l’ANR.

Source : [FAQ Plan de Gestion des Données \(PGD\) de l'ANR](#)

A noter que cette politique est commune au réseau des agences de financement françaises (ADEME, ANR, ANRS-MIE, Anses et INCa)

Modes de partage des données de recherche

→ Répondre manuellement à des demandes



Exemple : <https://doi.org/10.1186/s12888-020-03010-3>

→ Annexer à un article des fichiers de *Supplementary data*, publiés sur le site de l'éditeur en même temps que l'article

→ Déposer les données dans un entrepôt de données - ou *data warehouse* ou *data repository*



Exemple : <https://doi.org/10.5255/UKDA-SN-852431>

→ Autre?

Availability of data and materials

Readers and all interested researchers may contact Runsen Chen (Email address: runsen.chen@psych.ox.ac.uk) for details.

UK Data Service
ReShare



Home

Legal

Review procedures

Metrics

Emotional intelligence in social work

Biggart, Laura Paule (2018). *Emotional intelligence in social work*. [Data Collection]. Colchester, Essex: UK Data Archive. [10.5255/UKDA-SN-852431](https://doi.org/10.5255/UKDA-SN-852431)

This interdisciplinary project examines whether emotional intelligence helps professionals in emotionally demanding jobs such as social work to perform well and whether it helps them to manage their anxiety and stress. Emotional intelligence indicates how good someone is at understanding how emotions work, recognising emotions in themselves, recognising emotions in others, managing their own emotions and managing emotions in others. This project aims to assess the influence of Emotional Intelligence (EI) training for social workers in relation to burnout and performance using a randomised control trial. It also aims to produce a psychometric measure of social work performance created from the perspective of social workers, managers, service users and educators. The three year study has four phases: (1) a literature review and focus groups will identify items which tap into behavioural characteristics of social worker performance; (2) will validate a social work performance measure created from phase 1; (3) an emotional intelligence intervention will be designed using elements from existing EI programmes which show effects on outcomes; (4) a randomised control trial will evaluate an emotional intelligence training intervention for social workers and assess its effect on social work performance and stress.

DATA DESCRIPTION (ABSTRACT)

2. Trouver des données de recherche avec les outils bibliographiques

Fonctionnalités intégrées aux bases bibliographiques
Outils de recherche spécialisés



Fonctionnalités intégrées aux bases bibliographiques

- Des outils intégrés aux bases de données bibliographiques permettent d'identifier les jeux de données **liés à des publications**.
 - › [PubMed](#) : attribut *Associated data*, disponible sous la forme d'un filtre à partir d'une liste de résultats ; filtre sur la présence dans l'article d'un lien vers une source de données comme ClinicalTrials.gov, GenBank, Figshare ou Dryad.
 - › [Scopus](#) : encart *Related Research Data*, accessible depuis la notice d'un article en particulier. Cette fonctionnalité est donc très limitée, puisqu'elle ne permet pas de sélectionner un ensemble d'articles ayant des données associées.
- Le **type de publication** peut également être exploité, pour identifier les articles de données, ou *data papers*. Les valeurs associées diffèrent en fonction des bases de données.
 - › Scopus et Web of Science -> choisir la valeur *data paper*
 - › PubMed -> choisir la valeur *dataset*

Outils de recherche spécialisés 1/2

- DataCite Search - <https://search.datacite.org>
 - › Couverture : toutes disciplines, tous types de productions (thèses, preprints, etc.) - [liste des sources](#)
- Mendeley Data, par Elsevier - <https://data.mendeley.com/research-data>
 - › Couverture : toutes disciplines, données, articles ; pour les revues Elsevier : figures, tableaux et fichiers de données supplémentaires
- Google Dataset search - <https://datasetsearch.research.google.com/>
 - › Couverture : tous types de données (pas seulement de recherche), mais seulement des données
- Isidore, par HumaNum - <https://isidore.science/>
 - › Couverture : sciences humaines et sociales, tous types de production scientifique

Outils de recherche spécialisés 2/2

D'autres services indexent des données et peuvent être utiles pour identifier des données existantes ; la liste suivante n'est pas exhaustive.

- [Web of Science Data Index](#) : sur abonnement - l'université de Bordeaux n'est pas abonnée à cette ressource.
- [B2FIND](#) : composante de l'infrastructure européenne EUDAT ; l'interface permet notamment d'effectuer des recherches par couverture chronologique et géographique.
- [Bielefeld academic search engine](#) : le moteur de recherche de l'université de Bielefeld permet de cibler des recherches sur les données, notamment grâce au critère "Type de document" > "Dataset", "Unknown", etc.

TP1 : trouver des données de recherche avec les outils spécialisés

→ Effectuez les 2 recherches suivantes pour trouver des données pertinentes pour votre domaine de recherche. Prenez note des informations suivantes.

- › Identifiant des données trouvées (DOI et URL si possible)
- › Pour la recherche 1 : sur quel site sont-elles disponibles?
- › Sont-elles indépendantes ou liées à une publication (article, thèse, etc.)?

→ Recherche 1

- › Avec **Isidore** si votre domaine de recherche relève des SHS ou du droit
- › Avec **Mendeley Data** dans les autres cas

→ Recherche 2

- › Avec **recherche.data.gouv.fr**

Intérêt des outils spécialisés?

→ Avantages

- › Ils permettent d'interroger simultanément de nombreuses sources de données.
- › Certains indexent également les fichiers supplémentaires aux articles.

→ Limites

- › Les fonctionnalités de recherche et d'exploitation des résultats sont très limitées.
- › Leur couverture excédant les données de recherche et/ou non spécialisée par domaine scientifique peut être source de beaucoup trop de bruit.

3. Trouver des données de recherche grâce aux entrepôts de données



Entrepôts de données généralistes vs. spécialisés

Trouver un entrepôt de données généraliste

Trouver un entrepôt de données spécialisé

TP2 Entrepôts de données spécialisés vs. généralistes

Choisissez 2 jeux de données dans la liste ci-dessous et affichez-les. Quelles différences entre eux observez-vous (nombre de fichiers mis en ligne, documentation associée, accès aux fichiers, richesse des informations descriptives, etc.) ?

Althaus, C. (2018). *Analyses of Ebola virus disease outbreaks* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.1285677>

Biggart, L. P. (2016). *Emotional intelligence in social work* [Data set]. UK Data Service. <https://doi.org/10.5255/UKDA-SN-852431>

Dashkina, R., Richard, G., & Barbier, M. (2021). *Analyse textuelle des résumés des 44 opérations d'expertise scientifique collective, d'étude et de prospective conduites à l'INRA de 2000 à 2020* [Data set]. Recherche Data Gouv. <https://doi.org/10.15454/QRIZCR>

Fok, M. (2021). *Crédits formel et informel de trésorerie aux producteurs de coton au Bénin* [Data set]. CIRAD Dataverse. <https://doi.org/10.18167/DVN1/NCNHJZ>

Haberle, S., Hopf, F., Tng, D., Johnston, F., & Bowman, D. (2014). *Weekly pollen count data for for the University of Tasmania, Hobart*. [Data set]. ACEAS. <https://doi.org/10.4227/05/5344E9A41A124>

Institut National de la Statistique et des Études Économiques. (2021). *Capacité des communes en hébergement touristique—2021* [Data set]. Progedo. <https://doi.org/10.13144/LIL-1443>

Lebrun, B., & Tribolo, C. (2020). *Optically Stimulated Luminescence data from archaeological site Uzès (Gard, France)* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.4061220>

Critères de choix d'un entrepôt de données

→ Meilleur choix : l'entrepôt de données spécialisé utilisé par votre communauté scientifique

→ Autres critères (liste non exhaustive) :

- › L'entrepôt gère-t-il différents **types d'accès**? Permet-il par exemple un accès restreint ou sous **embargo**?
- › Prend-il en charge un **format de données** particulier?
- › Quelles sont les exigences ou possibilités en termes de **précision et structuration de la description des données déposées**? Devrez-vous ou pourrez-vous fournir des métadonnées riches?
- › Propose-t-il d'autres services particulièrement adaptés à mes besoins? [cf. exemple connexion Zenodo-Github ci-dessous]

Zenodo - connexion GitHub



Search

UploadCommunities

Log inSign up

June 8, 2018

SoftwareOpen Access

Analyses of Ebola virus disease outbreaks

 Christian Althaus

This release contains data and  code files from the following analyses of Ebola virus disease (EVD) outbreaks:

1. Althaus CL. (2014) [Estimating the reproduction number of Ebola virus \(EVOB\) during the 2014 outbreak in West Africa](#). PLoS Curr, 6.
2. Althaus CL, Low N, Musa EO, Shuaib F, Gsteiger S. (2015) [Ebola virus disease outbreak in Nigeria: Transmission dynamics and rapid control](#). Epidemics, 11:80-4.
3. Althaus CL. (2015) [Rapid drop in the reproduction number during the Ebola outbreak in the Democratic Republic of Congo](#). PeerJ, 3:e1418.
4. Althaus CL. (2015) [Ebola superspreading](#)70135-0). Lancet Infect Dis, 15:507-8.
5. Abbate JL, Murall CL, Richner H, Althaus CL. (2016) [Potential impact of sexual transmission on Ebola virus epidemiology: Sierra Leone as a case study](#). PLOS Negl Trop Dis, 10:e0004676.
6. Althaus CL. (2018) [Real-time analysis of the 2018 Ebola outbreak in the Democratic Republic of Congo](#). GitHub.

Preview

 Ebola-v1.0.0.zip

calthaus-Ebola-4888566

DRC (GitHub 2018)

Ebola_outbreak_DRC2018_data.csv1.0 kB

README.Rmd9.8 kB

README.md2.8 kB

README_cache

markdown_github

__packages33 Bytes

fit_ae1c24202d0029d622dc3ee7bdb65569.RData2.7 kB

fit_ae1c24202d0029d622dc3ee7bdb65569.rdb8.6 MB

fit_ae1c24202d0029d622dc3ee7bdb65569.rdx457 Bytes

reproduction_73d668b2a1a28489f76ef1d634598955.RData2.7 kB

reproduction_73d668b2a1a28489f76ef1d634598955.rdb42.9 MB

reproduction_73d668b2a1a28489f76ef1d634598955.rdx251 Bytes

69views

7downloads

See more details...

Available in



Indexed in



Publication date:

June 8, 2018

DOI:

 [10.5281/zenodo.1285677](https://doi.org/10.5281/zenodo.1285677)

Related identifiers:

Supplement to

<https://github.com/calthaus/Ebola/tree/v1.0.0>

License (for files):

 Other (Open)

Exemple :
<https://doi.org/10.5281/zenodo.1285677>

Trouver un entrepôt de données généraliste

Outil d'aide à la décision en ligne UBx : **Trouver un entrepôt de données**

→ Pour identifier l'entrepôt de données généraliste qui répond le mieux aux besoins, à partir de la réponse aux 4 questions suivantes :

- › Allez-vous publier des données dont l'accès doit être restreint?
- › Recherchez-vous un entrepôt avec des options de dépôt gratuit?
- › Souhaitez-vous pouvoir définir librement les conditions d'utilisation de vos données?
- › Souhaitez-vous que votre dépôt de données soit relu avant d'être mis en ligne?

→ Une brève fiche descriptive accompagne chacun des huit entrepôts de données comparés : [4TU.ResearchData](#), [Figshare](#), [Recherche Data Gouv](#), [Dryad](#), [Mendeley Data](#), [Open Science Framework \(OSF\)](#), [B2SHARE](#), [Zenodo](#)

Trouver un entrepôt de données généraliste

Trouver un entrepôt de données

Vous cherchez un entrepôt de données généraliste pour rendre accessibles vos données de recherche ? Cet outil peut vous aider à trouver l'entrepôt qui réponde le mieux à vos besoins. Pour identifier un entrepôt de données spécialisé pour votre domaine de recherche, interrogez l'annuaire **re3data** ou **contactez-nous**. Pour plus d'informations concernant les données de recherche, consultez les pages de notre site web **Les données de recherche**.

Pour toute question, contactez-nous à l'adresse **doc-soutienrecherche@u-bordeaux.fr**

*Cet outil est adapté du **Data Repository Finder** développé par l'université d'Utrecht. Il est en cours de paramétrage et de finalisation, les contenus et les fonctionnalités sont susceptibles d'évoluer.*

1. Allez-vous publier des données dont l'accès doit être restreint ?

☐ Oui ☐ Pas nécessairement

3. Souhaitez-vous pouvoir définir librement les conditions d'utilisation de vos données ?

☐ Oui ☐ Pas nécessairement

2. Recherchez-vous un entrepôt avec des options de dépôt gratuit ?

☐ Oui ☐ Pas nécessairement

4. Souhaitez-vous que votre dépôt de données soit relu avant d'être mis en ligne ?

☐ Oui ☐ Pas nécessairement

Outil d'aide à la décision en ligne

<http://busec2.u-bordeaux.fr/aide-choix-entrepot/>

Limites des entrepôts de données généralistes et avantages des entrepôts spécialisés

Les entrepôts de données généralistes ne permettent généralement pas de bénéficier des avantages suivants des entrepôts spécialisés.

- Prise en compte des spécificités d'un ou plusieurs domaines disciplinaires
- Richesse et précision des informations descriptives
- Interopérabilité : renforcée si l'entrepôt de données requiert la conformité à un format de données
- Réutilisation des données par des applications informatiques (*machine readability*) - cf. point ci-dessus

Trouver un entrepôt de données spécialisé : recommandations - ex SHS

→ Sociétés savantes, éditeurs

- › [Page « Data sharing » sur le site de l'APA](#) : espace sur OSF pour les données liées aux publications APA - « *just one of many available options* »

→ Organismes de financement

- › [ERC - European Research Council](#) : liste commentée d'entrepôts spécialisés par discipline - p. 12 et suiv.
- › [Commission européenne](#) : liste d'entrepôts généralistes et spécialisés par discipline - ces *Data Guidelines* comportent en outre des recommandations pour la préparation des données (notamment tabulaires).

→ Membres d'un consortium : CESSDA - Consortium of European Social Science Data Archives : <https://www.cessda.eu/About/Consortium>

- › Pour la France : PROGEDO - <http://www.progedo.fr/>



Trouver un entrepôt de données spécialisé : annuaire re3data

Registry of Research Data Repositories.

<https://www.re3data.org/>

Il s'agit de la ressource de référence, indexant plus de 2800 entrepôts. On peut rechercher par mot-clé et naviguer dans re3data par :

- [sujet](#),
- [type de contenu](#),
- [pays](#).

The screenshot displays the re3data.org website interface. At the top, there is a navigation bar with the re3data.org logo, a search bar, and links for Browse, Suggest, Resources, and Contact. The DataCite logo is also present. On the left side, a 'Filter' sidebar lists various categories such as Subjects, Content Types, Countries, AID systems, API, Certificates, Data access, Data access restrictions, Database access, Database access restrictions, Database licenses, Data licenses, Data upload, Data upload restrictions, Enhanced publication, Institution responsibility type, Institution type, Keywords, Metadata standards, PID systems, Provider types, Quality management, Repository languages, Software, Syndications, Repository types, and Versioning. The main content area shows search results for 'PubChem'. It includes a search bar, a pagination bar (Previous, 1, 2, 3, 4, 5, 6, 7, ..., 108, Next), and a 'Sort by' dropdown. The results for PubChem are displayed with filters for Subject(s) (Basic Biological and Medical Research, Chemistry, Biology, Life Sciences, Natural Sciences), Content type(s) (Databases, Images, Structured graphics, Scientific and statistical data formats, Raw data, Plain text, Structured text), and Country (United States). A detailed description of PubChem is provided, mentioning its three databases: PubChem BioAssay, PubChem Compound, and PubChem Substance. Below the PubChem results, the 'World Data System' is also listed with its own filters and description.

Trouver un entrepôt de données spécialisé : annuaire re3data

De nombreux filtres permettent ensuite d'affiner les listes de résultats, en fonction par exemple des critères suivants.

- **Certificates** : par quelle certification l'entrepôt est-il qualifié?
- **Accès** : gradient de valeurs de fermé à ouvert - se décompose en 3 types d'accès :
 - › **Database access** : accès à l'entrepôt de données lui-même : sous quelles conditions un utilisateur peut-il accéder à la base de données en général?
 - › **Data access** : accès aux jeux de données déposés dans un entrepôt de données spécifique : sous quelles conditions un utilisateur peut-il accéder à un jeu de données?
 - › **Data upload** : accès à la soumission de données : sous quelles conditions un utilisateur peut-il soumettre des données ?
- **Versioning** : les jeux de données peuvent-ils être versionnés?

Trouver un entrepôt de données spécialisé : service FAIRsharing

[FAIRsharing.org](https://fairsharing.org) répertorie non seulement des entrepôts mais également des standards, des méthodes, des vocabulaires, etc. Pour chaque ressource sont notamment spécifiés les critères suivants.

- › Le **statut** : en développement / opérationnel / incertain / déprécié
- › La **recommandation** : nom de l'éditeur, de la revue, etc. qui recommande

Voir : Sansone, S.-A. et al. (2019). FAIRsharing as a community approach to standards, repositories and policies. *Nature Biotechnology*, 37(4), 358-367. <https://doi.org/10.1038/s41587-019-0080-8>

F. Flamerie - Stage CED Données de recherche - 2023

The screenshot displays the FAIRsharing.org website. At the top, the logo 'FAIRsharing.org' is accompanied by the tagline 'standards, databases, policies'. Below the header, a red button labeled 'Recommended' is visible. The main content area features a 'View as Table' / 'View as Grid' toggle and a 'Sort by' dropdown menu set to 'Best Match'. A section titled 'Recommended Records' contains a table with filters for 'Associated Publication?' (No Publication, Has Publication), 'Claimed?' (No Maintainer, Has Maintainer), and 'Record Status' (Uncertain, Deprecated, In development, Ready). At the bottom, a 'Standard Type' section lists 'Reporting Guideline' (11 items) and 'Terminology Artifact' (7 items).

4. Trouver des données de recherche pour les réutiliser

Entrepôts de données et principes FAIR

Focus sur les principes FAIR

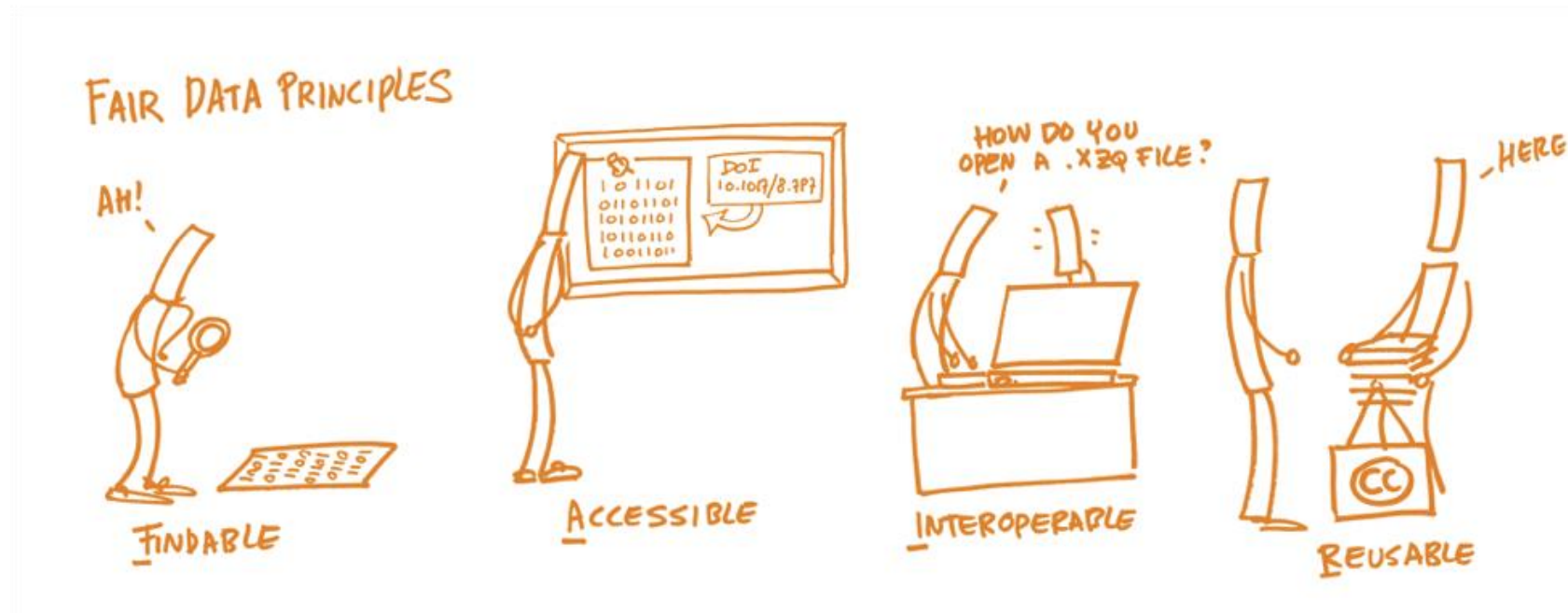


Entrepôts de données et principes FAIR

Principes FAIR -> les données doivent pouvoir être trouvées, comprises et réutilisées.

Cela ne signifie pas qu'elles doivent être nécessairement en accès entièrement libre.

Ces principes s'appliquent également aux métadonnées associées aux données.



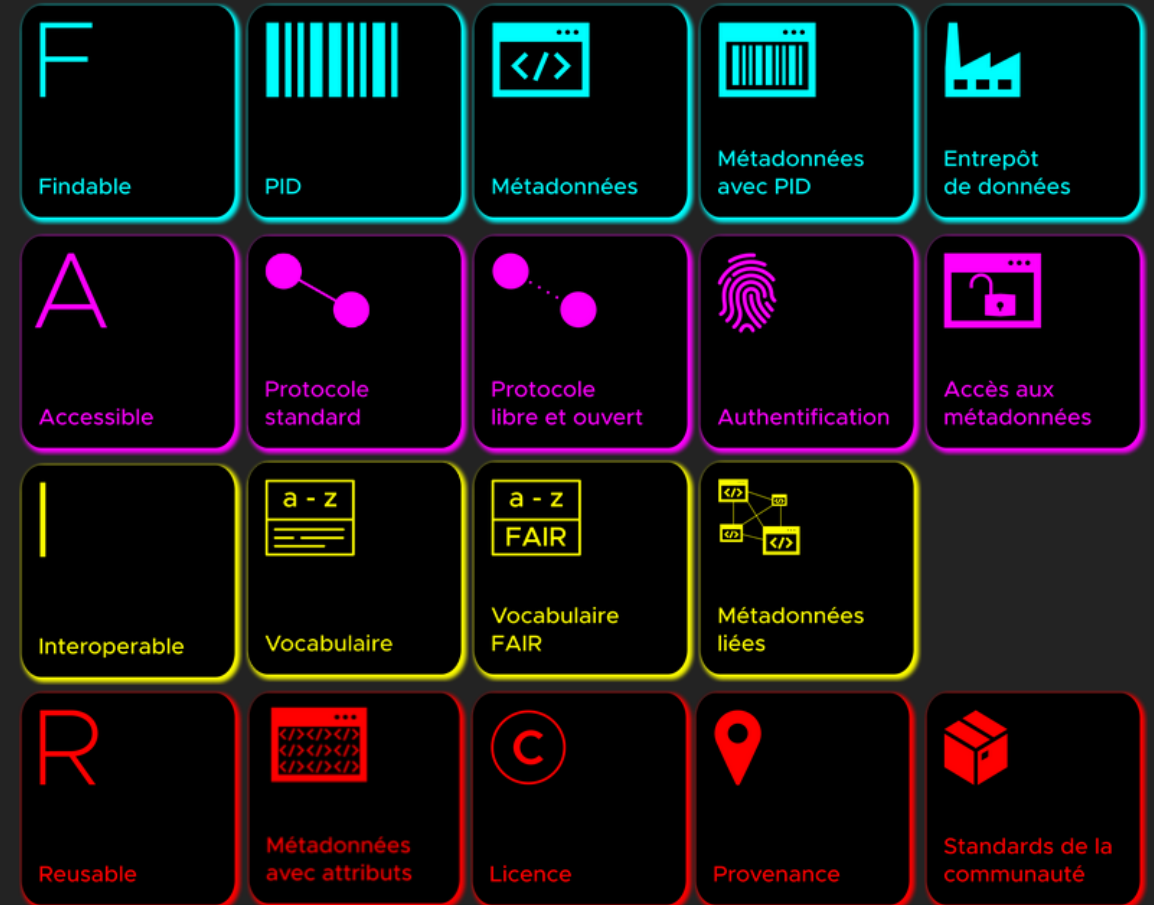
Source image : [The Open Science Training Handbook](#)

Focus sur les principes FAIR

« [Les principes FAIR](#) » sur le [site DORANuM](#)

Les principes FAIR

Les chercheurs s'appuient sur les connaissances scientifiques antérieures, notamment sur les résultats publiés dans les articles scientifiques. La reproductibilité des résultats, ainsi que leur croisement, ne sont cependant envisageables qu'avec des données originelles et leurs conditions d'obtention. C'est pourquoi la science ouverte vise à faciliter l'accès aux publications scientifiques et aux données de la recherche. Cette facilitation s'accompagne d'un certain nombre de mesures pour rendre les données scientifiques facilement découvrables, accessibles, interopérables et réutilisables. Ce sont les principes FAIR : Findable, Accessible, Interoperable, Reusable.



Entrepôts de données et principes FAIR

DOI et citation :
Findable - Accessible

Documentation :
Reusable

Une documentation plus structurée et standardisée permettrait d'augmenter les scores Interoperable - Reusable

Signage interventions for stair climbing at work: more than 700,000 reasons for caution

Eves, Frank, University of Birmingham, <https://orcid.org/0000-0002-4921-7169>

evesff@bham.ac.uk

Publication date: October 25, 2019

Publisher: Dryad

<https://doi.org/10.5061/dryad.qz612jm91>

Citation

Eves, Frank (2019), Signage interventions for stair climbing at work: more than 700,000 reasons for caution, Dryad, Dataset, <https://doi.org/10.5061/dryad.qz612jm91>

Abstract

Increased stair climbing reduces cardiovascular disease risk. While signage interventions for workplace stair climbing offer a low-cost tool to improve population health, inconsistent effects of intervention occur. Pedestrian movement within the built environment has major effects on stair use, independent of any health initiative. This paper used pooled data from UK and Spanish workplaces to test the effects of signage interventions when pedestrian movement was controlled for in analyses. Automated counters measured stair and elevator usage at the ground floor throughout the working day. Signage interventions employed previously successful campaigns. In the UK, minute-by-minute stair/elevator choices measured effects of momentary pedestrian traffic at the choice-point ($n = 426,605$). In Spain, aggregated pedestrian traffic every 30 min measured effects for 'busyness' of the building ($n = 293,300$). Intervention effects on stair descent (3 of 4 analyses) were more frequent than effects on stair climbing, the behavior with proven health benefits (1 of 4 analyses). Any intervention effects were of small magnitude relative to the influence of pedestrian movement. Failure to control for pedestrian movement compromises any estimate for signage effectiveness. These pooled data provide limited evidence that signage interventions for stair climbing at work will enhance population health.

Methods

Observations from automated counters followed by bootstrapped regressions

Usage Notes

Read me file for UK interventions

Up from ground floor

The coding used to allow analysis of the file with logistic regression is explained below. The dependent variable is stair and escalator choice. If you have any queries, please do not hesitate to contact me directly (evesff@bham.ac.uk).


For ascent, file IJERPHMultiSiteUKUp, the coding is as follows below.

INTSTAGE; 1 =baseline, 2=first intervention, 3=second intervention.

BASE0INT1; the coding allows comparison of the baseline (0) with the first intervention (1).

PH2ASE1; the coding allows comparison of the first intervention (0) with the second (1).

Data Files

 Download dataset

October 25, 2019


IJERPHBarce...iSite.xlsx	276.45 kB
IJERPHMulti...KDown.xlsx	5.92 MB
IJERPHMulti...eUKUp.xlsx	6.48 MB


Related Works


Article

<https://doi.org/10.3390/ijerph16193782>

Metrics

 25 views

 0 downloads

 0 citations

Keywords

population health

signage effectiveness

stair climbing

workplace

License

This work is licensed under a CC0 1.0 Universal (CC0 1.0) Public Domain Dedication license.



Licence
CC:
Reusable

TP3 Evaluer un jeu de données

Option1, évaluer un jeu de données trouvé

Vous pouvez évaluer

- l'un des 2 jeux de données suivants,
- l'un des jeux de données mentionnés en exemple précédemment au cours de la formation,
- ou encore l'un des jeux de données identifiés lors des précédents TP.

Damián, J., Pastor-Barriuso, R., García-López, F., Ruigómez, A., Martínez-Martín, P., & de Pedro-Cuesta, J. (2018). *Facility ownership and mortality among older adults residing in care homes* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.2549444>

Kovacheva, S., & Demireva, N. (2020). *The lived experiences of migration 1996-2017* [Data set]. UK Data Service. <https://doi.org/10/gntpw7>

Une fois sélectionné le jeu de données que vous allez évaluer, considérez les points suivants.

- Les **métadonnées** sont-elles suffisantes?
- La **documentation** (fichier README, etc.) vous paraît-elle suffisante?
- Qu'en est-il de la licence de publication? Les conditions de **réutilisation** sont-elles claires?

Allez plus loin : calculez le score FAIR (Findable, Accessible, Interoperable, Reusable) de ce jeu de données en répondant au questionnaire suivant, disponible :

- au format statique PDF : [questionnaire FAIR au format PDF](#),
- au format dynamique formulaire web : [Australian Research Data Commons. FAIR self-assessment tool](#)

! Vous pouvez considérer prioritairement les lettres F, A et R

TP3 Evaluer un jeu de données

Option 2, évaluer un de vos jeux de données

Vous préférez travailler sur un de vos jeux de données en cours de préparation?

Utilisez plutôt le questionnaire suivant (en français) : Data Archiving and Networked Services. Outil FAIR-Aware. DoRANum.

<https://doi.org/10.13143/egeo-nw05>

5. Politiques de données des éditeurs et des organismes de financement de la recherche



Politiques des éditeurs

Politiques des organismes de financement de la recherche

Politiques des éditeurs

- L'**accès aux données sous-jacentes** aux articles et la **citation des données** constituent les deux points les plus évidents concernant le partage des données.
- PLOS a été l'un des premiers éditeurs à définir une politique de données en 2014, obligeant notamment au dépôt en accès ouvert des données sous-jacentes aux articles.
- Si on considère plus largement le processus de publication, les données ne sont elles-mêmes qu'une constituante d'une démarche globale d'ouverture et de transparence : voir l'initiative **Transparency and Openness Promotion (TOP)**.

Politiques des éditeurs : TOP

→ [Afficher le tableau sur le site TOP](#)

→ Liste des 16 revues qui atteignent le niveau 3 des TOP Guidelines : [rubrique "Data Reproducibility Policies" sur le site de TOP](#)

TOP Guidelines summary table.

	Not Implemented	Level I	Level II	Level III
Citation Standards	No mention of data citation.	Journal describes citation of data in guidelines to authors with clear rules and examples.	Article provides appropriate citation for data and materials used consistent with journal's author guidelines.	Article is not published until providing appropriate citation for data and materials following journal's author guidelines.
Data Transparency	Journal encourages data sharing, or says nothing.	Article states whether data are available, and, if so, where to access them.	Data must be posted to a trusted repository. Exceptions must be identified at article submission.	Data must be posted to a trusted repository, and reported analyses will be reproduced independently prior to publication.
Analytic Methods (Code) Transparency	Journal encourages code sharing, or says nothing.	Article states whether code is available, and, if so, where to access it.	Code must be posted to a trusted repository. Exceptions must be identified at article submission.	Code must be posted to a trusted repository, and reported analyses will be reproduced independently prior to publication.
Research Materials Transparency	Journal encourages materials sharing, or says nothing.	Article states whether materials are available, and, if so, where to access them.	Materials must be posted to a trusted repository. Exceptions must be identified at article submission.	Materials must be posted to a trusted repository, and reported analyses will be reproduced independently prior to publication.
Design and Analysis Transparency	Journal encourages design and analysis transparency, or says nothing.	Journal articulates design transparency standards.	Journal requires adherence to design transparency standards for review and publication.	Journal requires and enforces adherence to design transparency standards for review and publication.
Study Preregistration	Journal says nothing.	Article states whether preregistration of study exists, and, if so, where to access it.	Article states whether preregistration of study exists, and, if so, allows journal access during peer review for verification.	Journal requires preregistration of studies and provides link and badge in article to meeting requirements.
Analysis Plan Preregistration	Journal says nothing.	Article states whether preregistration of study with analysis plan exists, and, if so, where to access it.	Article states whether preregistration with analysis plan exists, and, if so, allows journal access during peer review for verification.	Journal requires preregistration of studies with analysis plans and provides link and badge in article to meeting requirements.
Replication	Journal discourages submission of replication studies, or says nothing.	Journal encourages submission of replication studies.	Journal encourages submission of replication studies and conducts results blind review.	Journal uses Registered Reports as a submission option for replication studies with peer review prior to observing the study outcomes.

More information available at cos.io/top

TP4 TOP factor

TOP attribue un "**TOP factor**" aux revues, en se fondant sur le tableau présenté. En 2022, plus de 1900 revues ont reçu un "TOP factor".

→ Retrouvez-vous un titre que vous connaissez dans la liste des revues évaluées?

→ Son score vous surprend-il? Pourquoi?

→ [Liste des revues ayant un TOP factor](#)

→ [Présentation du TOP factor 2022 sur le site de TOP](#)

Politiques des organismes de financement de la recherche

Deux aspects constituent des éléments centraux dans la politique des organismes de financement.

- La recherche de la conformité aux **principes FAIR** ; l'approche "*as open as possible, as close as necessary*" de l'Union européenne est ainsi reprise par l'ANR.
- La mise en œuvre de bonnes pratiques de gestion des données, décrites dans un **plan de gestion de données** (DMP - data management plan). [En savoir + sur les DMP sur le site DORANum](#)

La politique de science ouverte de **l'ANR** est détaillée sur [la page "La science ouverte" de son site internet](#)



Merci pour votre attention

Des questions?

Contact : frederique.flamerie-de-lachapelle@u-bordeaux.fr