



Science ouverte

Données ouvertes, logiciel libre

Philippe Gauron, IJCLab





- La recherche se base sur
 - la reproductibilité
 - un processus d'amélioration incrémentale



1. accès ouvert aux publications 

2. données ouvertes 

3. sources logicielles libres 



Science ouverte : quel intérêt ?

transparente



participative



cumulative



favorise le débat
contradictoire



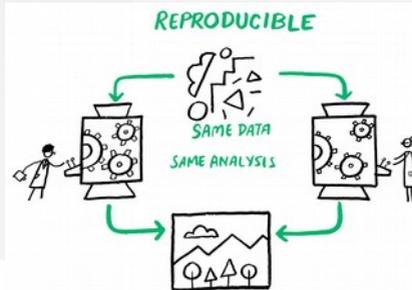
Science ouverte



traçable



reproductible



facilite l'émergence
de connaissances



This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.

This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.



Direct :

- diffuser vos travaux plus vite
- les améliorer
- démarche d'intégrité scientifique

Indirect :

- accès à des idées (collecte, amélioration des protocoles et instruments)
- accès à des données sans nécessiter d'accès aux instruments (accréditations et autorisations, temps)
- évaluation des chercheurs (y compris recrutement)



La science doit être accessible techniquement mais aussi **légalement** et **humainement**

- sensibilisation
- accompagnement
- structures HAL et entrepôt de données, etc.



Scriberia

Oser partager malgré la crainte du jugement

This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.

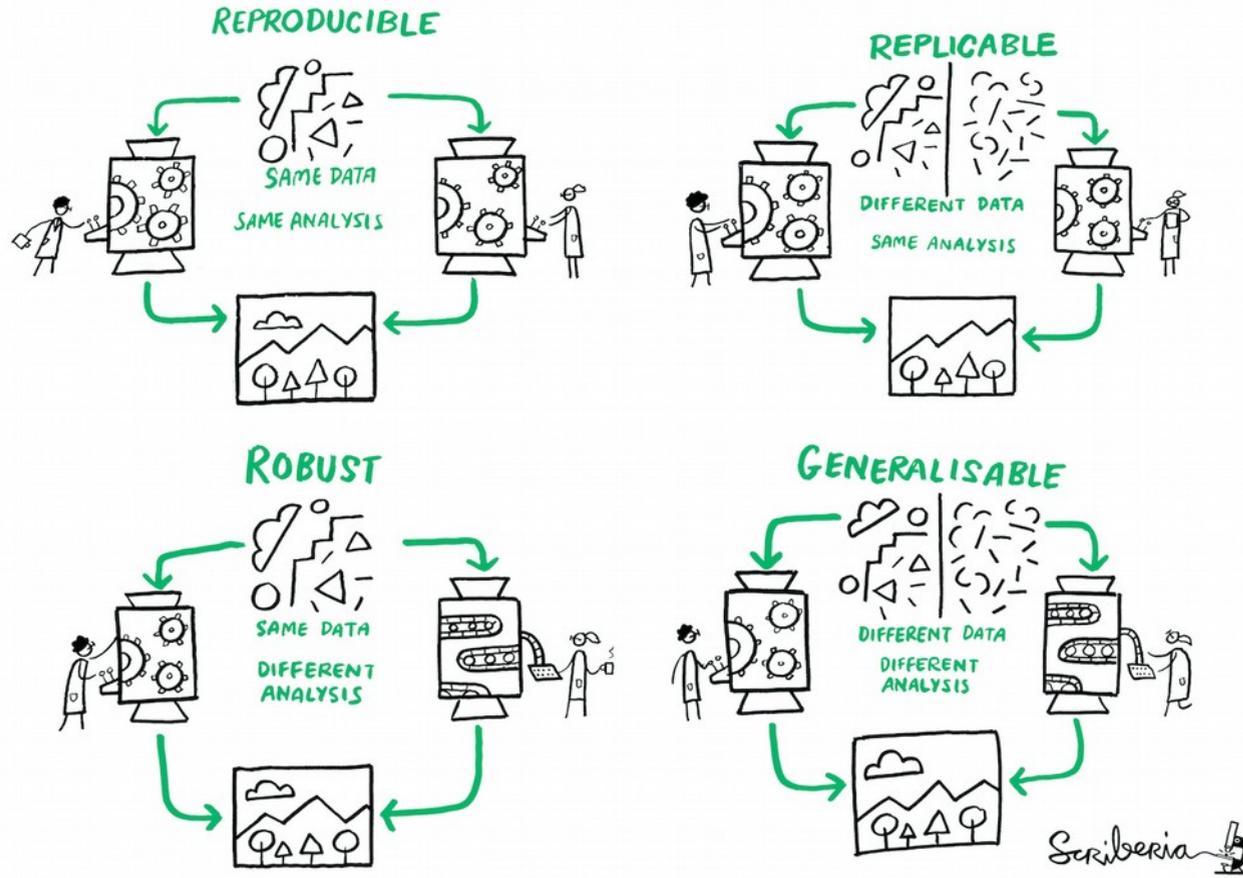


- Plan national pour la Science Ouverte
- Politique des données, des algorithmes et des codes sources de l'ESR
- 2005 : droit de réutilisation des info publiques, données publiques et logiciels
- 12/2015 : réutilisation gratuite **par défaut** ([Loi pour une république numérique](#)), articles financé à +50 % public mis à dispo après embargo de 6 mois en STEM ou 1 an en SHS ([art. 30](#))
- 2018 : [1er plan national pour la science ouverte](#)
- des financeurs demandent un dépôt articles+données en archives ouvertes même si déjà mis à dispo par éditeur **sans embargo** ([Horizon Europe](#) depuis 2021, [ANR](#), etc.)
- 2021 : [2e plan national pour la science ouverte](#)



Données ouvertes, données de recherche

1. Quoi ?
2. Pourquoi ?
3. Comment ?



This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.



Données de la recherche

« 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. »

Source : Organisation de Coopération et de Développement Économiques - OCDE,
<https://web-archive.oecd.org/fr/2018-04-10/137520-38500823.pdf#page=19>

Données ouvertes (Open data)

Open est utilisé comme synonyme de *free* ou *libre* selon les définitions de la [Free Software Foundation](#) et de [Free Cultural Works](#)

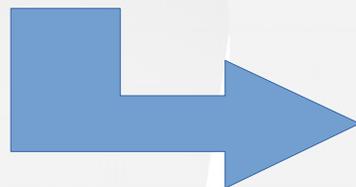
- | | |
|------------------------|----------------|
| 1. Open license/status | 2. Access |
| 3. Machine readability | 4. Open format |

Source : Open Knowledge, Open definition 2.1, <https://opendefinition.org/od/2.1/en/>



permettent de

- comprendre, reproduire, vérifier les résultats
- obtenir des financements
- faciliter les collaborations
- être citée dans une publication



réflexion en amont sur

- les livrables des projets de recherche
- la manière de les organiser.



données ouvertes → FAIR

1. **Facile à trouver** par humains ou automatique, nécessite métadonnées et id pérennes
2. **Accessible** : entrepôt pérenne en accès ouvert, ou sur demande si données sensibles
3. **Interopérable** : téléchargeables, format ouvert, lisible/éditable sur plusieurs logiciels, combinables avec d'autres données
4. **Réutilisable** : doc contextuelle, licence



Besoin d'un entrepôt, mais aussi, pour la pérennité :

- une documentation
- des métadonnées



FAIR n'implique pas données ouvertes

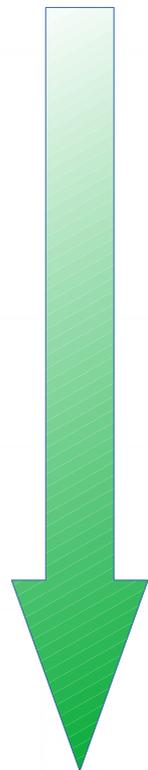
Exemples :

<https://wikidata.org>



<https://opendatabarometer.org/>

Wikidata : Planemad, Public domain, via Wikimedia Commons



1978 : loi CADA (Commission d'accès aux Documents administratifs)

2003 : directive 2003/98/CE (informations du secteur public)

2008 : directive 2007/2/CE INSPIRE (données géographiques)

2013 : [Charte du G8](#) sur les données ouvertes

2015: charte internationale sur les données ouvertes

2016 (France) loi pour une république numérique : par défaut, les données produites par l'administration sont **ouvertes**

2019 : [directive 2019/1024](#) remplaçant celle de 2003



Définition du CRPA (code des relations entre le public et l'administration, art. L312-1-1)* :

- répertoires de données publiques
- base de données (BD) produites/reçues mises à jour régulièrement
- données mises à jour régulièrement présentant un intérêt économique, social, sanitaire ou environnemental

éventuellement avec redevance de réutilisation (art. L324-5 du CRPA)

Décret n°2017-638 recommandant les licences à utiliser, notamment :

- licence ouverte ~CC-BY : réutilisation à toute fin en citant, sans dénaturer données
- ODBL ~CC-BY-SA, inspirée du logiciel libre : réutilisation si partage à l'identique en cas d'enrichissement/modif)
- CC0 ~domaine public : sans mention de la source

* ne s'applique pas aux collectivités territoriales de moins de 3500 habitant·es



Données ouvertes : en France, différents régimes

Les données du domaine public ou données *publiques* sont librement réutilisables par défaut, **MAIS**

- pas de droits d'auteur sur données sauf pour les bases de données :
 - droit d'auteur sur l'architecture de la BD à la développeuse ou développeur
 - droit *sui generis* à l'établissement si investissement substantiel pour création/vérification/présentation de la BD



données personnelles (RGPD) : éviter d'identifier/réidentifier

- Régime spécial :
 - infos environnementales (convention d'Aarhus)
 - infos géographiques (directive européenne Inspire du 14 mars 2007 établissant une infrastructure d'information géographique dans la Communauté européenne, transposée en France par l'[ordonnance n°2010-1232](#))



This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.



Données ouvertes : en France, des exceptions

Quelques exceptions à l'ouverture des données :

- secret d'affaire (clause de confidentialité), en cas de part. Public-Privé (modalités de publication à définir : partenaires d'abord, puis grand public) ;
- données personnelles ;
- données protégées par des droits de propriété intellectuelle ;
- les données soumises au secret défense, ou encore relevant du potentiel scientifique et technique de la nation.

Exemple : en sciences de l'univers, l'investissement est 100% public (parfois international) dans infra lourdes, la publication est donc immédiate,

En 2022 en France, **22 %** des publications mentionnent la **création d'un jeu de données**

(source : [Note Flash SIES n°05](#), février 2024, baromètre science ouverte 2023)



Métadonnées = informations sur les données elles-mêmes

Dans un format standard, idéalement ouvert pour être :

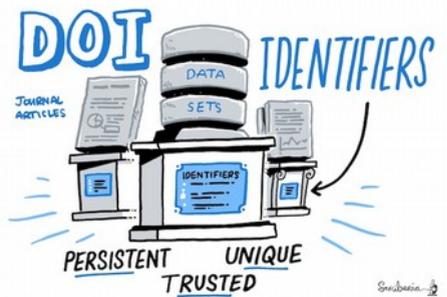
- interrogeable,
- traçable,
- ajoutables automatiquement (collecte)

Exemple : format [Dublin Core](#), rempli quand on utilise un [formulaire HAL](#).



Données ouvertes : identifiants pérennes

PID (persistent identifier) : identifiants des **objets** mais aussi **contributeurs et contributrices**



Exemples :

- **DOI** (Digital Object Identifier) attribué par DataCite (consortium de bibliothèques)
- **ORCID** (Open Researcher and Contributor ID) : persistant, distinctif, évolutif (évolution de carrière), importable, personnalisable (responsabilités, affiliations), mise à jour par tiers sur autorisation, interopérable

Intérêt : lie ressource, localisation web, et ressources associées (article-jeu de donnée), fiabilise ce lien

Facilité la découverte, la réutilisation et la visibilité des publications, auteurs et autrices liées.

These images were created by Scriberia for The Turing Way community and is used under a CC-BY licence.



Données ouvertes : opportunités et points d'attentions

- data paper : article décrivant les données, qui peut être cité (DOI) et valorisé (carrière)
- Prix Science ouverte des données de la recherche



dépôt dans un entretien de données, selon les conditions *a minima* : demande un travail de la mise en forme

This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.



Données ouvertes : niveaux de classification

5 stars (Tim Berners-Lee) les classe de manière incrémentale :

	données quelconques sur le web, sans licence	
★	données publiées avec licence ouverte	pdf, jpg
★★	données publiées sous forme structurée (tableur plutôt qu'image)	ods, xls, doc
★★★	format ouvert et non propriétaire (csv plutôt que excel)	csv, ods, json
★★★★	URI pour désigner les objets pour permettre les références	xml
★★★★★	lier les données à d'autres données	rdf

Certification possible par <https://certificates.theodi.org/en/>

Objectifs :

- accepter la publication des données
- concourir à la diffusion
- faciliter la réutilisation

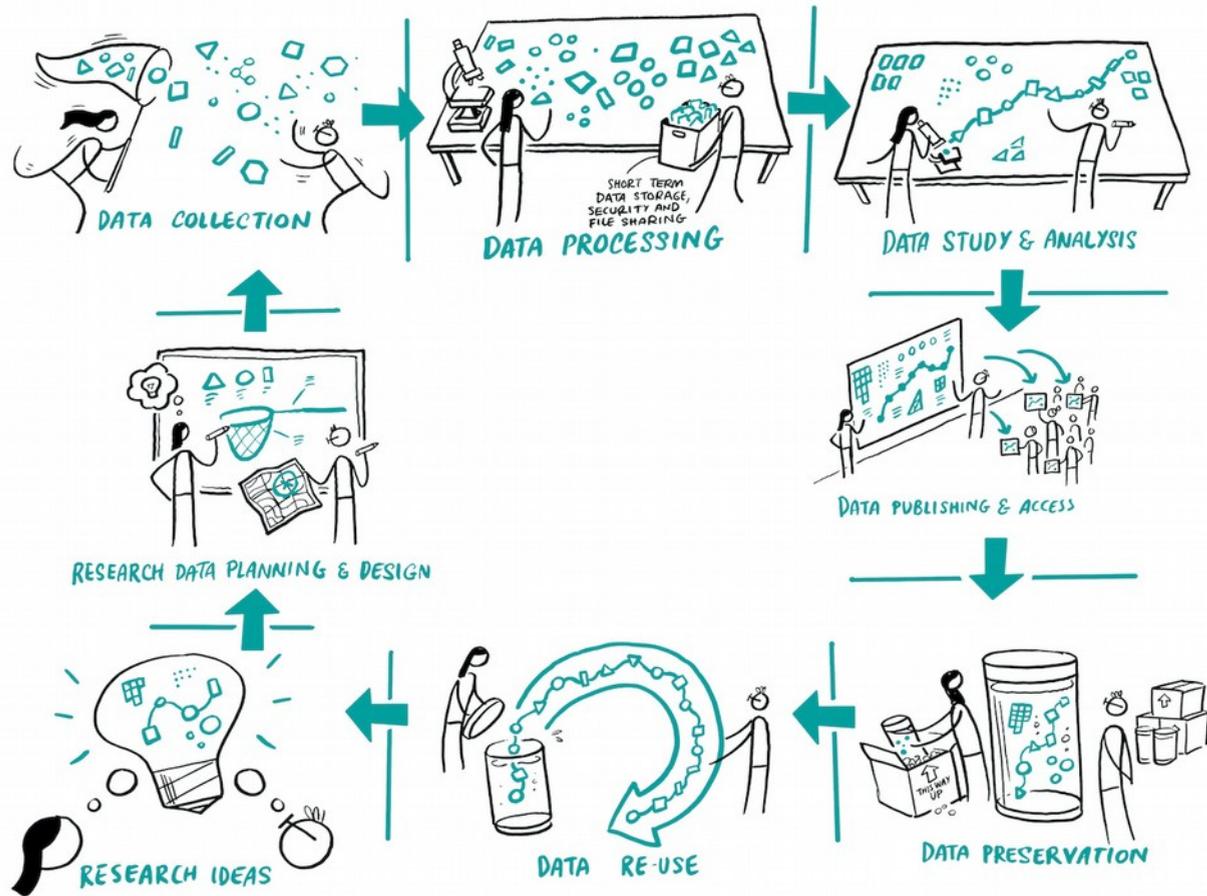


Données ouvertes : des entrepôts à éviter

- absence de modération des dépôts (vérif. métadonnées, contrôle scientifique, pas de mots-clés, de contexte, mal documentées, possibilité d'embargo, filtre de données incomplètes)
- absence d'identifiant pérenne (DOI, etc. pour trouver/citer)
- absence de garantie sur la pérennité de l'infrastructure (au moins 5 ans)
- entrepôts pratiquant la cession de droits (pas de licence ou licence privative)
- politique tarifaire excessive
- localisation des données hors Union européenne pour certaines données (RGPD)
- dépôt restreint par l'affiliation institutionnelle



Données ouvertes : cycle de vie



Scriberia

This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.



Données ouvertes : cycle de vie (1-3)

Principe : quels outils pour inventorier, classer les données, de quoi qlq d'autre aurait besoin comme info pour les exploiter ?

1. Collecte : enregistrement des données et leur sauvegarde (3 copies, sur 2 supports dont 1 à distance : 3-2-1) : format ? nommage ? endroits appropriés ?
2. Traitement et analyse des données :
 - ✓ si données personnelles, prévoir comment conserver/anonymiser (RGPD), quelle accessibilité du logiciel ? identification des versions ?
 - ✓ Analyse : prévoir description objective (mesures, indicateurs, tests)
3. Interprétation & publi : les replacer dans contexte de l'état de l'art et hypothèses *a priori*

Analyse + interprétation → publication



4. Archivage :

- ✓ quoi conserver ? cb de temps ?
- ✓ après le projet : cmt trier (automatiser ?) ? cmt enregistrer ? cmt archiver (verser dans un entrepôt d'archives) ? où (entrepôt de confiance, ex: zenodo) ? support ? accès des personnes autorisées ? quoi supprimer (contraintes légales) ?
- ✓ définir des id pérennes (DOI, Handle system, ARK)

5. Partage et diffusion : quoi ? principes FAIR :

- ✓ articles scientifiques
- ✓ articles décrivant les données
- ✓ répertorier les données dans les bases pluridisciplinaires, moteurs de recherche, catalogues (DataCite, OpenAIRE), entrepôts de données pluridisciplinaires (Zenodo), privés (Figshare, GigaDB, Medelay data), disciplinaires (DRYAD, NCBI, Nakala, Gbif, Seanoe), selon contraintes (certif CoreTrustSeal, financeur, facilité, etc.) et licences



6. Réutilisation :

- ✓ consentement pour l'usage ?
- ✓ citer la source (y compris apprentissage, enseignement, vérification)...
- ✓ doc et accessibilité importantes

7. Pérennisation :

- ✓ prolongement de l'archivage ? suppression partielle ?
- ✓ métadonnées et logiciels nécessaires
- ✓ contraintes légales
- ✓ supprimer données incomplètes ou altérées
- ✓ dépendance format à une techno
- ✓ décodage technique et humain



Entrepôts d'établissement ou thématiques :

- structures très différentes d'une discipline à une autre (différent des articles)
- publics ou privés
- payants ou pas

Identifier les entrepôts thématiques internationaux ou nationaux de confiance (certification [Core Trust Seal](#) possible).

Exemples d'entrepôts publics ou appartenant aux communs numériques :

- [Pangaea](#) (sciences de la terre)
- [Dryad](#) (sciences médicales, payant)
- [Zenodo](#) (prépubli possible, DOI général et DOI version particulière pour gérer les nouvelles versions)



Données ouvertes : outils / accès aux entrepôts

Les entrepôts sont multiples, ils se structurent encore, le bon niveau est à définir : local, national, international ? Besoin de référencer ces entrepôts pour trouver les données.



coût de l'entrepôt des données, financier et environnemental : trier est essentiel !

Existence d'outils numériques pour s'interfacer aux données, comme R Open science :

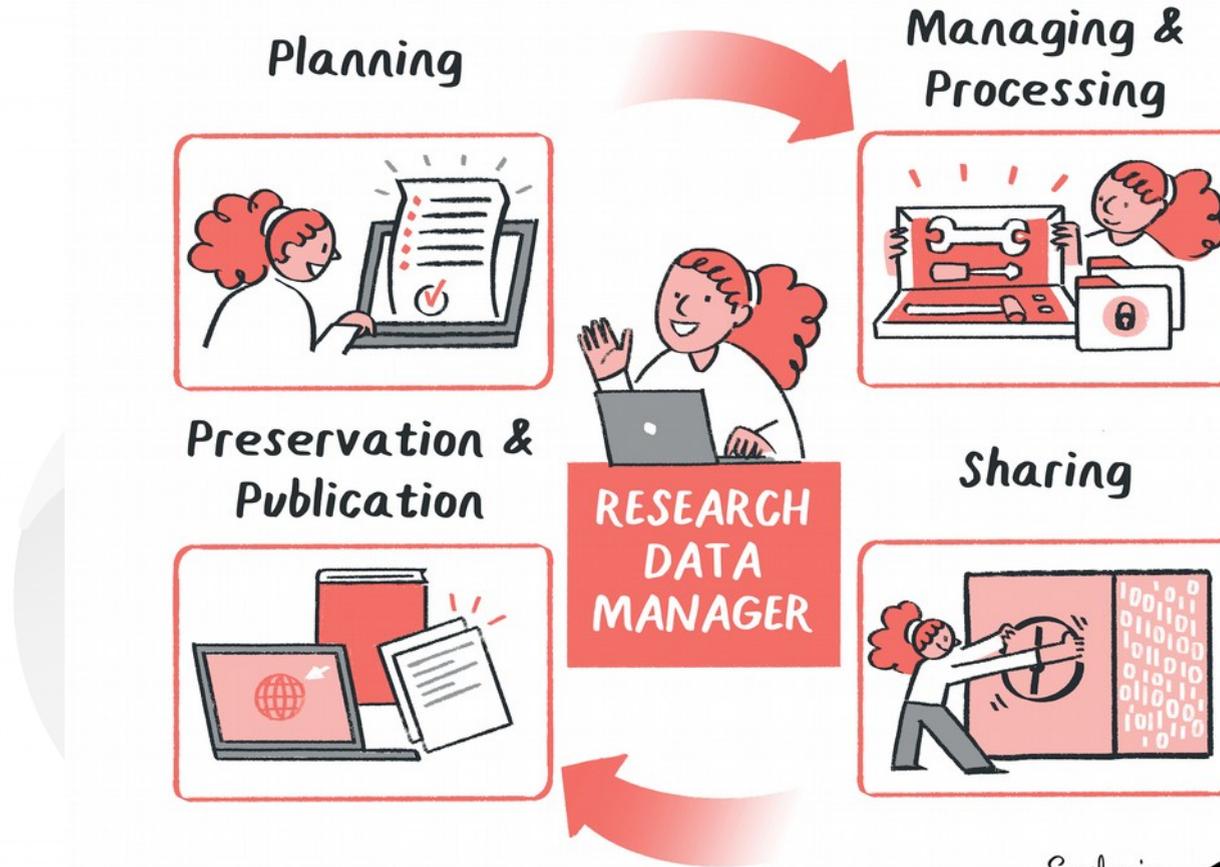
- chercher des données
- visualiser, explorer
- rendre les données interopérable (conversion de format)
- reproductibilité (scripter analyse et rédaction)



This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.



Plan de gestion de données



Scriberia

This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.



Plan de gestion de données (PGD) / Data management plan (DMP)
document **administratif** et **scientifique**, qui synthétise toutes les informations sur les données de recherche d'un projet, au sens large.

Incrémental et évolutif, 1ère version durant le montage du projet

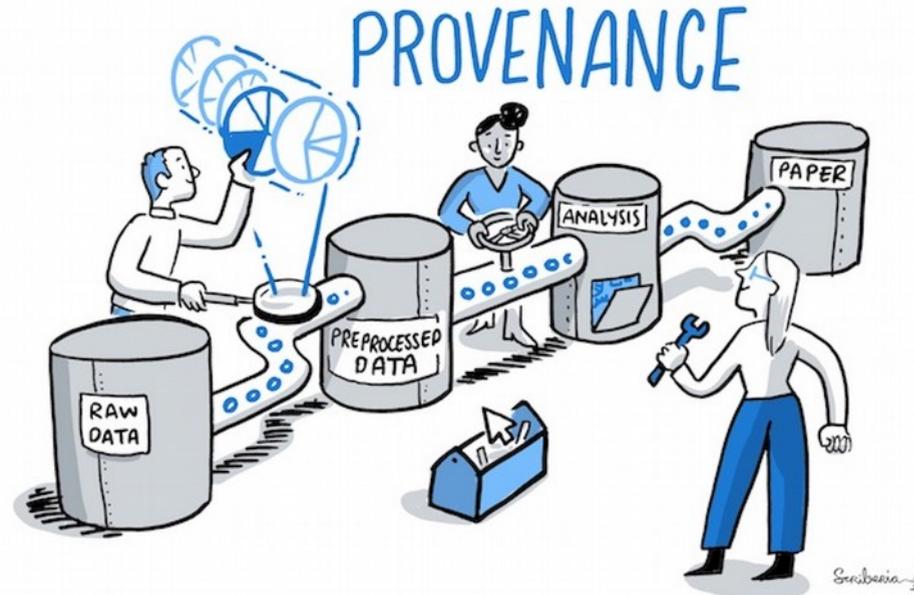
Permet de réfléchir sur les besoins en données, donc anticiper les besoins en infrastructure et en développement, et finalement humains



Demande des financeurs !



Plan de gestion de données : cycle de vie



v1 : dans les 6 premiers mois (programmation)

Mi-projet (bilan d'étape)

définitif (fin de projet)

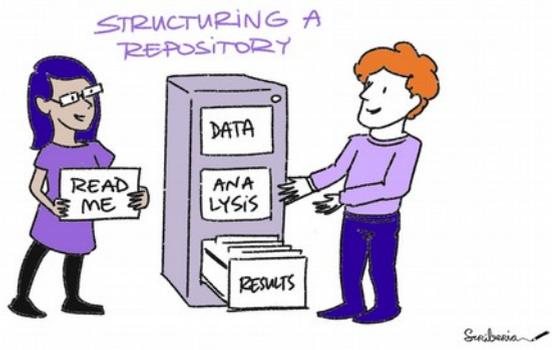
This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.



Plan de gestion de données : données & métadonnées

Données

- données produites/réutilisées
- format, type (numérique, textuels, images, audio, vidéo, composites, etc.)
- id pérennes (PID)
- réutilisation de données existantes, raisons, contraintes
- exemples possibles de réutilisation dans d'autres contextes
- comment sera documentée la provenance ?



Métadonnées

- type, format (standard si possible)
- organisation données : nommage, contrôle de version, structures dossiers
- doc méthodo de collecte, procédure et méthode d'analyse, définition des variables, unités de mesure...
- localisation doc : BD, entête, fichier, livre réf., cahier de labo...

Contrôle qualité et conformité des données

- calibration
- répétition échantillons/mesures, capture standardisée, validation des saisies, revue par les pairs, vocabulaire contrôlé

This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.



Plan de gestion de données : conformité aux règles

RGPD

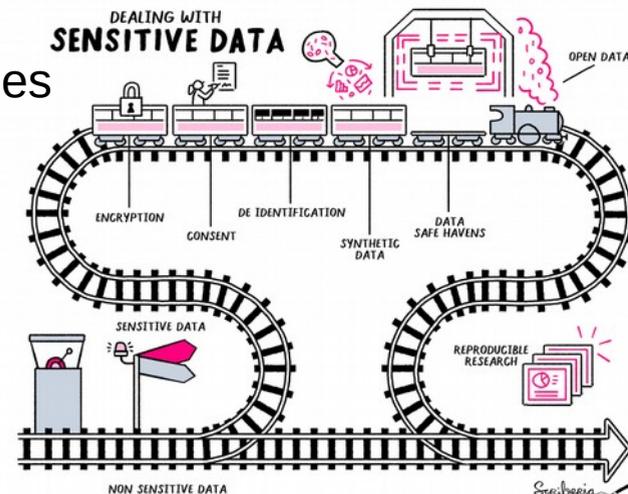
- consentement éclairé pour préservation/partage données personnelles
- Anonymisation ? pseudonymisation (réversible) ?
- chiffrement des données avec entrepôt clé sur tiers de confiance (~pseudonymisation)
- procédure d'accès spécifique pour accès données perso

Propriété intellectuelle

- propriété des données (contrôle d'accès)
- condition d'accès libre ou restrictions, licences d'accès/réutilisation (prévoir cas de multipartenariat)

Éthique

- impact sur entrepôt, transfert, visibilité, utilisation, durée de conservation
- code de conduite inter/nationaux et institutionnel, revue des pratiques nécessaire ?



This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.



Plan de gestion de données : imprévu & analyse

Entrepôt et sauvegarde

- lieu, fréquence sauvegarde (au moins 2 lieux distincts)
- système de stockage robuste, sauvegarde auto
- PRA (plan de reprise d'activité) en cas d'incident
- accès durant processus recherche, contrôle d'accès en particulier si collaboration
- politique institutionnelle de protection des données mise en œuvre



This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.

Traitement et analyse

- logiciels de collecte
- logiciels d'analyse
- version exacte (n° commit [SWHID](#)) pour assurer reproductibilité



Plan de gestion de données : partage des données

Format des données

- description (extension)
- justifier (expertise, préférence format ouvert, standard acceptés par l'entrepôt, usage de la communauté, format du logiciel utilisé)
- privilégier formats standards et ouvert

Volumétrie (espace, quantités fichiers, lignes, colonnes)

Disponibilité :

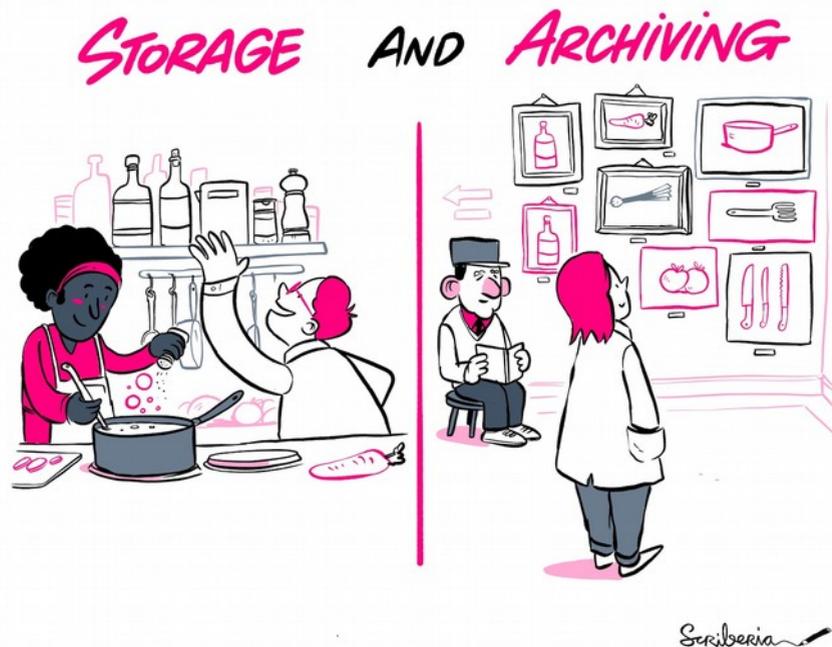
- comment trouver/partager données: type d'entrepôt, indexation, service données sécurisé, traitement direct
- date de dispo ? délais publi prévues ? si utilisation exclusive, raison et durée ? partage différé/limité et raison (PI, brevet...) ?
- raisons et méthode si restrictions, mesures prévues pour dépasser/minimiser restrictions
- utilisations prévisibles des données
- lieu de dépôt, démontrer prise en charge au delà de la durée du projet si pas de dépôt, démontrer vérif. politique des entrepôts/procédures de dépôt (normes métadonnées, coût de mise en œuvre)
- besoin d'outils spécifiques pour accès/utilisation (tenir compte durée de vie des logiciels)
- partage via entrepôt, mise à dispo directe ou autre ?



Plan de gestion de données : long terme

Conservation long terme

- plan de préservation des données (PPD) :
durée d'archivage pérenne
- données à ne pas divulguer/détruire si
raisons contractuelles/légales/réglementaires
- cmt décider quelles données garder ?
données préservées à long terme ?



This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.



Plan de gestion de données : Opidor

Opidor = Optimiser le Partage et l'Interopérabilité des Données de la Recherche

<https://opidor.fr>

- outil CNRS-INIST
- présente des modèles de plan de gestion de données



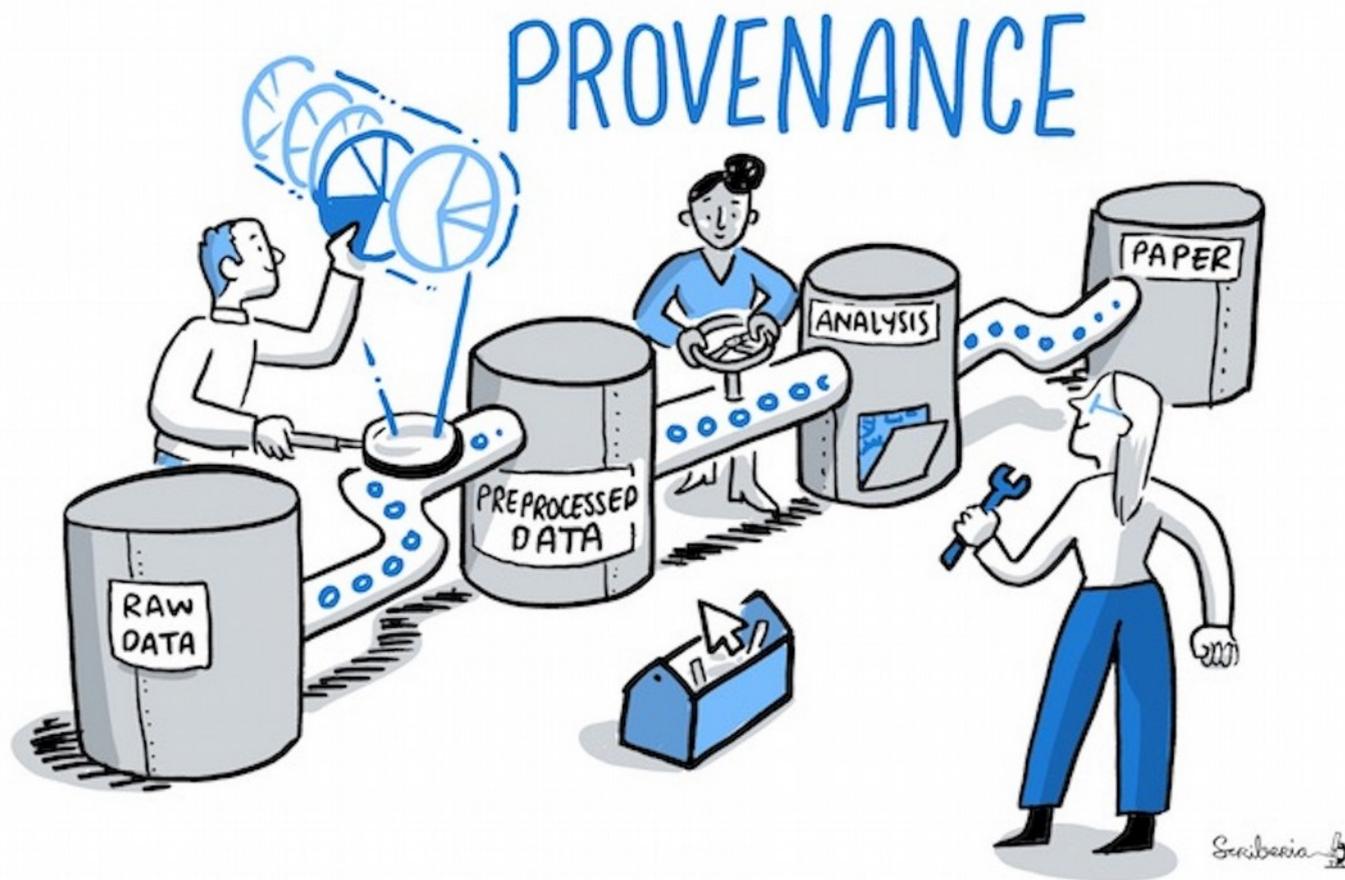
9/4/2024 : nouvelle version !



Questions ouvertes :

- Comment assurer la reproductibilité depuis la collecte des données (diffusion des scripts, documentation)
- Comment assurer la qualité des données ?
- Comment comparer les données ?
- Comment tenir compte de la qualité des données ?

This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.



This image was created by Scriberia for The Turing Way community and is used under a CC-BY licence.



Un **logiciel** est un ensemble des programmes, procédés et règles, et éventuellement de la documentation, relatifs au fonctionnement d'un ensemble de traitement de données (en anglais : software).

Source : [Arrêté du 22 décembre 1981](#) relatif à l'enrichissement du vocabulaire de l'informatique

- Un algorithme décrit les étapes de résolution d'un problème.
- Un (code) source est la mise en œuvre de l'algorithme.
- Un logiciel est processus de traitement, avec algo, code source, doc, exécutable, jeux de tests, licence, etc.

Le logiciel est un résultat de la recherche,
en plus d'être un outil, pas que une donnée.



Logiciels pour la recherche : droit d'auteur et logiciel (1)

Le logiciel *original* est couvert par le droit d'auteur, composé de droits moraux et patrimoniaux (art. [L111-1](#), Code de la propriété intellectuelle). Il dure 70 ans à compter de la mort de l'auteur ou l'autrice de l'œuvre

- droit moral ([L121-1](#) et [L121-2](#) CPI)
 - ✓ perpétuel, inaliénable et imprescriptible
 - ✓ citation, divulgation, repentir (hors logiciel), respect de l'œuvre
- droit patrimonial : ~ employeur si dév. salarié
 - ✓ monopole d'exploitation : reproduction, adaptation, diffusion (licence)

Couvert par le droit d'auteur :

- code source, fichiers objets et exécutables
- tout ce qui a servi à le créer (proto, etc.)
- la doc peut être couverte



Algo, fonctionnalités, cahier des charges non couverts



Couverture : personnes qui ont participé au prorata de leur contribution.
Important de les citer (support, droit d'auteur, documentation)

Droits patrimoniaux et détentrice/propriétaire de **logiciel** :

- structure employeuse si salarié/agent public, y compris si stagiaire de la recherche publique (prime d'intéressement, cf. [décret n° 2023-772](#)),
- il faut donc faire une convention de cession pour les non salariés (et stagiaire de la recherche privée ?)

Oeuvres orphelins (pas d'auteur ou autrice)

Oeuvres du domaine public (décès depuis plus de 70 ans)



Par défaut, un logiciel n'accorde aucun droit, même s'il est disponible (usage, copie, diffusion, modification, art. [L335-2](#) du CPI).

Une licence est donc nécessaire pour permettre son utilisation.

La licence est définie par défaut par le titulaire des droits patrimoniaux par défaut. Tout ce qui n'est pas autorisé est interdit par défaut.



- si tout le monde n'est pas salarié/agent public
 - si usage de images/sons/vidéos/textes etc., vérifier les droits
 - si usage d'outils/logiciels pour développer, vérifier limites des licences
- En cas de doute, contacter votre service valorisation.



Prévenir son employeur si logiciel assez mature (permet rétribution si valorisation).

Une preuve de date de création peut être nécessaire (souvent **APP**) → service valorisation

Brevet : logiciel non brevetable en UE **mais** brevet sur dispositif matériel
Nommage du logiciel : droit des marques (y compris nom de domaine)
IHM protégée si originale



Licence logicielle = contrat entre titulaire droits patrimoniaux et usager·e

- **Licences libres** créées par FSF (1985) avec 4 libertés :

0. utiliser

1. étudier et modifier

2. redistribuer une copie

3. redistribuer une version modifiée

- Licences dites privatives ou propriétaires, accordent un droit d'utilisation pour un type d'usage
- Plusieurs licences peuvent co-exister (ex : académique et entreprise), plusieurs composants peuvent être sous licences différentes.



Logiciel libre et logiciel open source : 2 définitions différentes

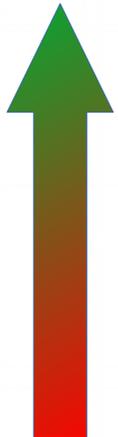
- Free software foundation
- Open source initiative

On trouve parfois le mélange *FLOSS* (Free libre open source software)
L'incompréhension et la confusion sont fréquentes.

Licences libres :

- à réciprocité/copyleft/gauche d'auteur (distribution dans les mêmes conditions y compris œuvre dérivée) : GPLv3, AGPLv3, Cecill-2.1
- à réciprocité faible/amointrie (bibli combinable/utilisable par logiciel propriétaire): LGPL (GPL amoindrie), Cecill-B, Mozilla MPL, EUPL
- permissives (intégration possible dans logiciel propriétaire): MIT, Apache, BSD, Cecill-B

Codification normalisée **SPDX**





« Ces conditions ne peuvent apporter de restrictions à la réutilisation que pour des motifs d'intérêt général et de façon proportionnée. »

Source : [Art. L323-2](#) du Code des relations entre le public et l'administration

Ouverture des logiciels développés dans le secteur public dès 2016 ([Loi pour une république numérique](#)).

Devenu une des 3 composantes du plan pour la science ouverte 2021.

« aussi ouvert que possible, aussi fermé que nécessaire »



Logiciels pour la recherche : quelle licence choisir ?

1. contacter le service de valorisation de votre employeur
2. choisir une licence **avant** de débiter : dépend des collaborations du projet, du niveau d'autonomie, de l'objectif recherché à long terme...
 - × éviter les licences nationales (Cecill) pour faciliter la diffusion
 - ✓ choisir parmi les **licences recommandées** par la DINUM
 - ✓ EUPL compatible droit de chaque pays UE, GPL, CC-BY, etc. (**matrice de compatibilité**)
 - ✓ tenir compte des licences des dépendances (**Fossology**)
3. signaler dans HAL (description) et entrepôt de données de confiance si un code source est associé à un jeu de données (ex.: <https://recherche.data.gouv.fr>)
4. utiliser une forge logicielle (rédaction, partage, maintenance collaboratives), permet gestion des problèmes, tâches, wiki, doc, codes sources, tests, intégration continue, reproductibilité (ex.: <https://gitlab.in2p3.fr>)



Logiciels pour la recherche : opportunités

- **Citation** dans vos publis des versions des logiciels utilisés avec Software heritage (**SWHID**) :
 - ✓ fichier
 - ✓ contenu d'un dossier pour un commit précis
 - ✓ commit précis
 - ✓ version précise d'un logiciel
 - ✓ dépôt dans un commit précis
- Prix **Science ouverte du logiciel libre de la recherche**
- En 2022 en France, **19 %** des publications mentionnent le **partage de sources logicielles** et **40 %** mentionnent **l'usage de logiciel** (source : Note Flash SIES n°05, février 2024, baromètre science ouverte 2023)



Logiciels pour la recherche : bonnes pratiques

- Utilisez un gestionnaire de version (Mercurial, git, bazar, etc.) couplé à une forge, idéalement de l'ESR, locale, nationale ou thématique (<https://gitlab.in2p3.fr>, <https://gitlab.huma-num.fr/>, etc.), archivées par [Software Heritage](#) (ou manuellement dans [HAL](#))
- Ajoutez à votre dépôt les métadonnées pointant sur chaque « objet » concerné (ORCID, DOI, idHAL, etc.)
 - ✓ une licence (LICENCE),
 - ✓ une description (README),
 - ✓ les auteurs et autrices (AUTHORS)
 - ✓ *les instructions d'installation (INSTALL)*
 - ✓ *les instructions pour participer (CONTRIBUTING)*
 - ✓ *mettez à jour la liste des évolutions à chaque version (CHANGELOG)*
- Organisez le support pour faciliter l'adoption, l'identification et la correction des erreurs, ouvrez la possibilité de proposer des correction/fonctionnalités
- Communiquez sur vos montées de version, démo, etc. citez votre logiciel dans vos articles ([biblatex-software](#))



Bibliographie : données ouvertes

- DINUM ; Plan d'action G8 sur l'ouverture des données publiques ; <https://www.numerique.gouv.fr/actualites/plan-daction-g8-sur-louverture-des-donnees-publiques/>
- décembre 2017 ; MESRI ; Ouverture des données de recherche Guide d'analyse du cadre juridique en France ; Licence ouverte 2 ; https://www.ouvrirlascience.fr/wp-content/uploads/2018/11/Guide_Juridique_V2.pdf
- 23 janvier 2023 ; CNRS-MITI, Groupe de travail « Atelier Données » Mission pour les Initiatives Transverses Interdisciplinaires ; Guide de Bonnes Pratiques sur la gestion des données de la Recherche ; CC-BY 4 ; <https://hal.archives-ouvertes.fr/hal-03152732>
- novembre 2023 ; Ouvrir la science ; Sélectionner un entrepôt thématique de confiance pour la diffusion des données de recherche : note méthodologique Collège Données de la recherche ; CC-BY-ND 4 ; <https://www.ouvrirlascience.fr/selectionner-un-entrepot-thematique-de-confiance-pour-la-diffusion-des-donnees-de-recherche-note-methodologique/>
- MESR-Collectif Ouvrir la science ; Gérer et partager les données de recherche pour les rendre réutilisables ; https://www.ouvrirlascience.fr/wp-content/uploads/2023/10/Livret_pour_impression.pdf
- 22/02/2024 Données de la recherche :
- données ouvertes en France (Wikipédia) : https://fr.wikipedia.org/wiki/Donn%C3%A9es_ouvertes_en_France
- CNRS, INIST ; site Opidor (outils pour les données ouvertes de l'ESR) : <https://opidor.frbadges> 5 stars open data ; Linked Open Data star badges ; <https://web.archive.org/web/20171129210613/http://lab.linkeddata.deri.ie/2010/lod-badges/>
- Dataactivist ; Formation « Comprendre l'open data » ; https://dataactivist.coop/oracle/comprendre_open_data/comprendreod.html#44
- Dublin Core ; format de métadonnées ; <https://www.dublincore.org/>
- Loi pour une république numérique, Article L312-1-1 du CRPA qui définit les exceptions à l'ouverture des données produites par l'administration ; https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000033205512
- choix de licence CC : <https://creativecommons.org/choose/>
- Choo-choo - choose your license : https://framapiaf.org/@Bristow_69/101376108789300072
- CIRAD ; arbre de décisions CC-BY ; <https://coop-ist.cirad.fr/actualites/avez-vous-le-droit-ou-l-obligation-de-diffuser-vos-donnees-parcourez-l-arbre-de-decision-du-cirad>
- open data 5 star ; classification de Tim Berners-Lee des données ouvertes ; <https://5stardata.info/fr/>
- Open data institute ; certification Open Data ; <https://certificates.theodi.org/en/>
- Linked Data Basics for techies ; https://web.archive.southampton.ac.uk/openorg.ecs.soton.ac.uk/wiki/Linked_Data_Basics_for_Techies.html
- 2023 ; Science ouverte, données de recherche, collection Passeport pour la science ouverte ; <https://www.ouvrirlascience.fr/science-ouverte-donnees-de-la-recherche/>
- 2023 ; Définir et reconnaître les activités de gestion et diffusion des données (étude), ministère de l'Enseignement supérieur et de la Recherche ; ???
- 2022 ; Partager les données liées aux publications scientifiques, guide pratique pour les chercheurs ; <https://www.ouvrirlascience.fr/partager-les-donnees-liees-aux-publications-scientifiques-guide-pour-les-chercheurs/>
- janvier 2023 ; Guide de bonnes pratiques sur la gestion des données de la recherche ; <https://mi-gt-donnees.pages.math.unistra.fr/guide/>
- 2023 ; Guide d'application de la loi pour une République numérique pour les données de recherche ; <https://www.ouvrirlascience.fr/guide-dapplication-de-la-loi-pour-une-republique-numerique-pour-les-donnees-de-la-recherche/>



Bibliographie : logiciels libres

- août 2022 ; MESR-Collectif pour la science ouverte ; Passeport pour la science ouverte - Codes et logiciels ; CC-BY-SA 4 ; <https://www.ouvrirelascience.fr/passeport-pour-la-science-ouverte-guide-pratique-a-lusage-des-doctorants/>
- septembre 2023 ; MESR, DGRI ; État des lieux de la production et de la valorisation des logiciels issus de la recherche publique française ; <https://www.enseignementsup-recherche.gouv.fr/sites/default/files/2023-12/-tat-des-lieux-de-la-production-et-de-la-valorisation-des-logiciels-issus-de-la-recherche-publique-fran-aise-30402.pdf>
- Article L323-2 du Code des relations entre le public et l'administration https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000033219073
- Guide juridique logiciels libres - Guide juridique pour la publication des codes sources de l'administration : <https://code.gouv.fr/guides/juridique/>
- MESR-Collectif Ouvrir la science ; Algorithmes, codes sources et logiciels libres ; https://www.ouvrirelascience.fr/wp-content/uploads/2023/10/Livret_pour_impression.pdf
- Rapport sur les Forges de l'ESR - Définition, usages, limitations rencontrées et analyse des besoins ; <https://www.ouvrirelascience.fr/forges-de-lesr-definition-usages-limitations-rencontrees-et-analyse-des-besoins/>
- Guide pour archiver et référencer des codes sources dans Software Heritage ; <https://www.softwareheritage.org/howto-archive-and-reference-your-code/>
- Présentation synthétique de la citation des logiciels. ; <https://www.softwareheritage.org/2020/05/26/citing-software-with-style/?lang=fr>
- Liste de licences commentées du Système exploitation GNU (soutenu par la Free Software Fondation) ; <https://www.gnu.org/licenses/license-list.html>
- Sandra Brandstätter - Ada et Zangemann ; CC-BY-SA 4 ; <https://cfeditions.com/ada/>
- 2011-09 Teresa Gomez-Diaz - Les aspects juridiques dans la production de logiciel ; http://igm.univ-mlv.fr/~teresa/logicielsLIGM/documents/JourneesLL/2011septJDEVPI_Lic.pdf
- DevLog ; Je code : quels sont mes droits ? Quelles sont mes obligations ? CC-BY-SA 4 ; <https://hal.science/hal-02399517>
- 2020 ; A policy and legal Open Science framework: a proposal ; Teresa Gomez-Diaz, Tomas Recio ; CC-BY 4 ; <https://hal.archives-ouvertes.fr/hal-02962399>
- 2023 ; Les logiciels de la recherche et leurs licences : trois visions sur un objet ; Teresa Gomez-Diaz ; CC-BY-SA-NC-ND 4 ; <https://hal.science/hal-02434287v2>
-

748 Membres

233

Chercheurs & Enseignants Chercheurs



Ingénieurs & Techniciens

344

150

Titulaires de HDR



171

Doctorants & Postdoctorants



50

Bourses de Recherche Européennes & Internationales

150

Bourses de Recherche Nationales & Locales

600/an

Articles dans des revues internationales



7 Pôles Scientifiques



1 Pôle Ingénierie



4 Plateformes de Recherche



6 Plateaux Techniques



50 000 m² de locaux

