



Bien gérer ses données pour (peut-être) les partager

Dominique L'Hostis

► To cite this version:

Dominique L'Hostis. Bien gérer ses données pour (peut-être) les partager. "Les datas : quels changements pour la recherche et pour les chercheurs ?", Université Fédérale Toulouse Midi-Pyrénées. FRA., Jan 2016, Rennes, France. pp.44 slides. hal-02799974

HAL Id: hal-02799974

<https://hal.inrae.fr/hal-02799974>

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Bien gérer ses données pour (peut-être) les partager



Matinée Doctorale – URFIST Rennes – 29/01/2016

Dominique L'Hostis - Inra



Plan

1. Contexte général - Big Data vs Small Data
2. Gérer ses données
 - ✓ Rédiger un plan de gestion des données
 - ✓ Bonnes pratiques de gestion
 - Décrire – Documenter
 - Sécuriser le stockage
 - Préserver à long terme
3. Diffuser ses données
 - ✓ Différentes voies
 - ✓ Identifiant, citation, licence
 - ✓ Limite du partage



_01

Contexte général

Big Data vs Small Data

Partager les données : une nécessité, et une problématique mondiale

❖ Des défis scientifiques complexes,

- ✓ Alimentation : Nourrir 9 milliards d'hommes, ...
- ✓ Changement climatique...
- ✓ Agriculture durable, éco-responsable...
- ✓ Santé et nutrition...
- ✓ Biodiversité ...



Numéro spécial – février 2011
<http://www.sciencemag.org/content/331/6018.toc#SpecialIssue>

qui supposent :

- ✓ transdisciplinarité, collaboration des équipes,
- ✓ ouverture et mutualisation des informations, **des données** et des compétences

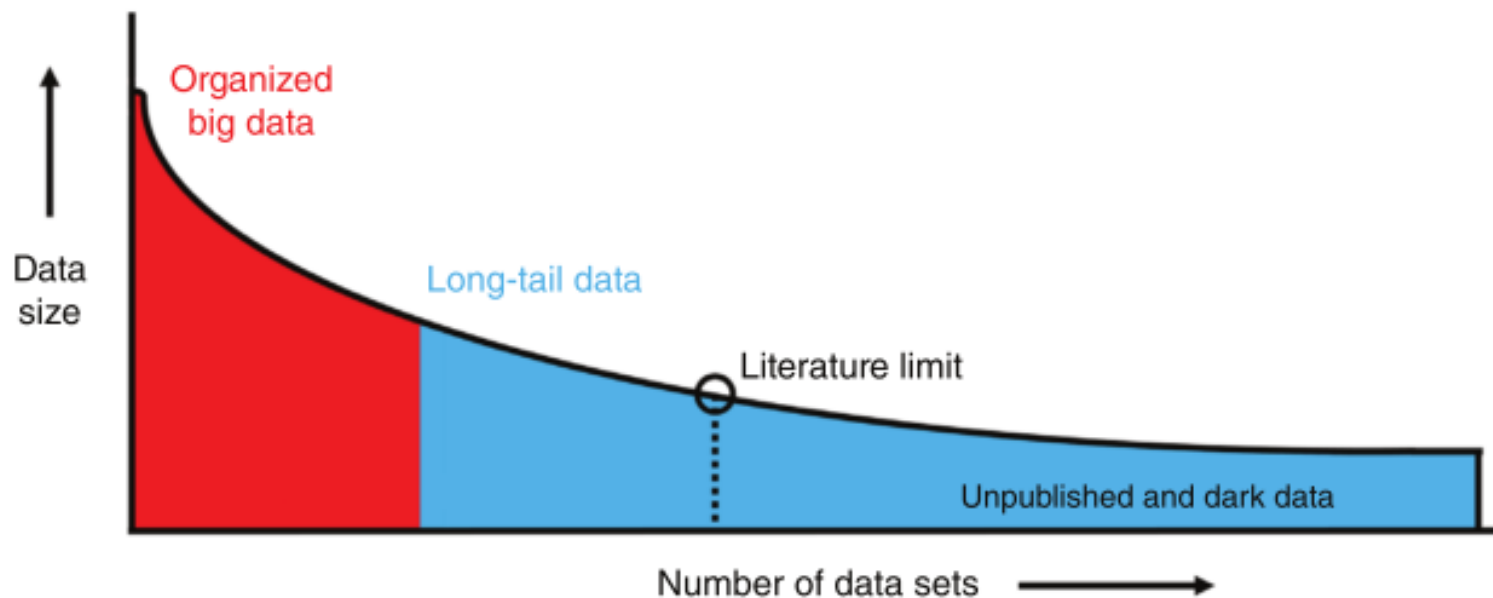
Soutien large de l'ouverture des données : vers l'« Open Science »



- ❖ par la plupart des **organismes de recherche**,
 - De l'Open Access à l'Open Data...
- ❖ par leurs **financeurs**,
 - Union Européenne, NSF, RCUK ...
- ❖ par les **gouvernements**
 - Prise de position du G8...
- ❖ par des **collectifs** transversaux ou spécialisés dans une discipline
 - [Science Europe](#)
 - [Research Data Alliance](#)
 - [Global Open Data for Agriculture and Nutrition \(GODAN\)](#)



« The long tail of data »



(Ferguson, Nielson, Cragin, Bandrowski, & Martone, 2014)

« Open Science needs to be on the ground and in the people, in 10.000s of labs around the world, in 100.000s projects performed by 1.000.000s of researchers » . (Horstmann, 2015)

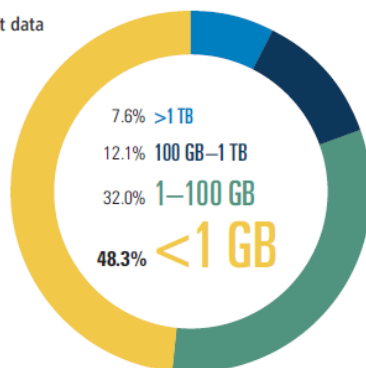
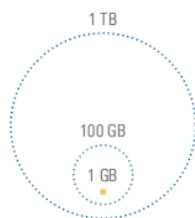
Small Data

❖ Une multitude de datasets

48,3% des répondants travaillent avec des jeux de données < 1GB

**Etude 2011
Science
(1700
réponses)**

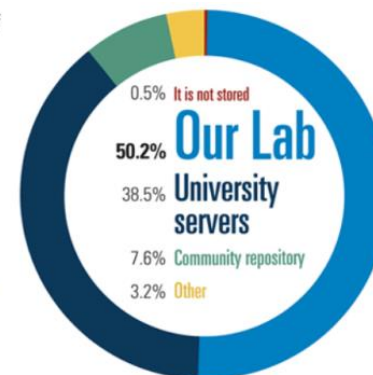
What is the size of the largest data set that you have used or generated in your research?



Plus de la moitié n'enregistrent leurs données que dans leur laboratoire

Where do you archive most of the data generated in your lab or for your research?

“Even within a single institution there are no standards for storing data, so each lab, or often each fellow, uses ad hoc approaches.”



<http://www.sciencemag.org/content/331/6018/692.full.pdf>

❖ La richesse dans l'hétérogénéité

“Diversity is likely to remain a dominant feature of research data : diversity of formats, types, vocabularies, and computational requirements, but also of the people and communities that generate and use the data”

https://ec.europa.eu/digital-agenda/sites/digital-agenda/files/data_einfra_h2020_fiches_online_consult.pdf

Bien gérer les données : préalable à l'ouverture et au partage

- ❖ *"L'ouverture des données ne peut constituer un point de départ, elle ne doit être envisagée que comme le résultat nécessaire d'une bonne politique de gestion de données, qui en constitue le préalable indispensable".*

Odile Hologne - cité dans : Gaillard, R., 2014. De l'Open data à l'Open research data : quelle(s) politique(s) pour les données de recherche ?". Université de Lyon, ENSSIB, Lyon. 1-104 p.

<http://www.enssib.fr/bibliotheque-numerique/notices/64131-de-l-open-data-a-l-open-research-data-quelles-politiques-pour-les-donnees-de-recherche>

- ❖ Challenge institutionnel : Développer une **culture du partage**
 - ✓ améliorer la qualité des datasets,
 - ✓ faciliter leur découverte, et leur réutilisation
 - ✓ motiver, informer et former les scientifiques aux bonnes pratiques de gestion et au partage

VIE DE L'INSTITUT

Un nouveau décret statutaire pour l'Inra

Le nouveau décret statutaire de l'Inra est paru au Journal Officiel du 25 novembre 2015 (Décret n° 2015-1517 du 23 novembre 2015). Les missions de l'établissement sont ainsi actualisées, avec deux nouveautés : d'une part, en l'absence de clauses contraíres, l'Institut organise l'accès libre aux données scientifiques et aux publications ; d'autre part, il contribue aux débats qui concernent la place de la science et de la technologie dans la société. Quelques ajustements dans la com-

Rendre visibles, partager ses données...

Disponibilité

Visibilité

Interprétabilité

Réutilisabilité

Citabilité

Curation

Préservation



Producteur/Utilisateur
de données

Décrire et documenter ses données ?

- Notion de plan de gestion
- Quelle documentation?
- Quelles métadonnées?

Gestion

Enjeux du partage

- Pourquoi partager, ou ne pas partager ?

Où et comment diffuser ses données?

- Quelles voies utiliser ?
- Publier dans un Data Paper, quels avantages ? Quels coûts ?

Ouverture

Comment faciliter la réutilisation ?

Partage

- « Trouvabilité », accès pérenne, format durable
- Identifiant - Citation
- Licences

7 critères ~ qualités à atteindre pour faciliter le partage et la réutilisation des données . Rapport ODE (Reilly, Schallier, Schrimpf, Smit, & Wilkinson, 2011)

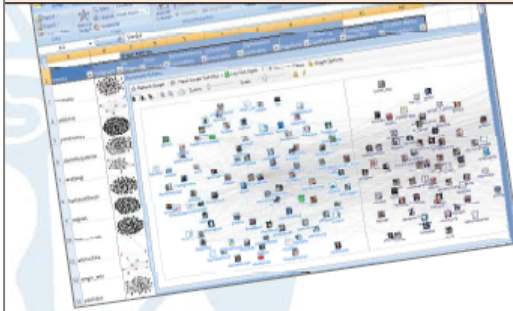


_02

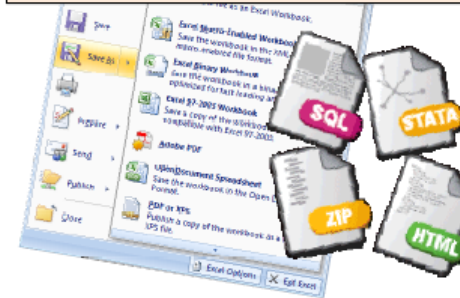
Gérer ses données

Short-term decisions with long-term implications

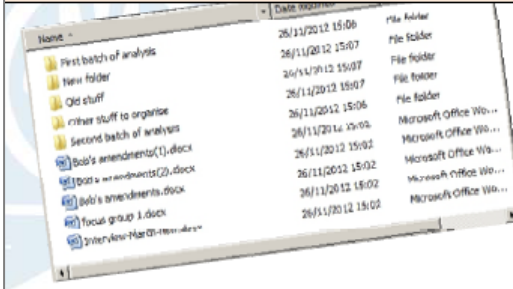
Software products



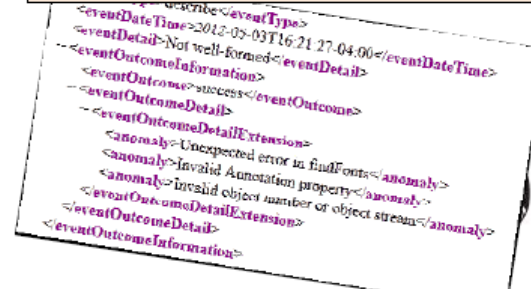
File formats & standards



Data organisation & labelling



Quality Controls



Utiliser des
**formats
durables**

Assurer des
**Contrôles
Qualité**
Fichiers,
Données,
Codes,
Documentation

Utiliser des
Logiciels libres

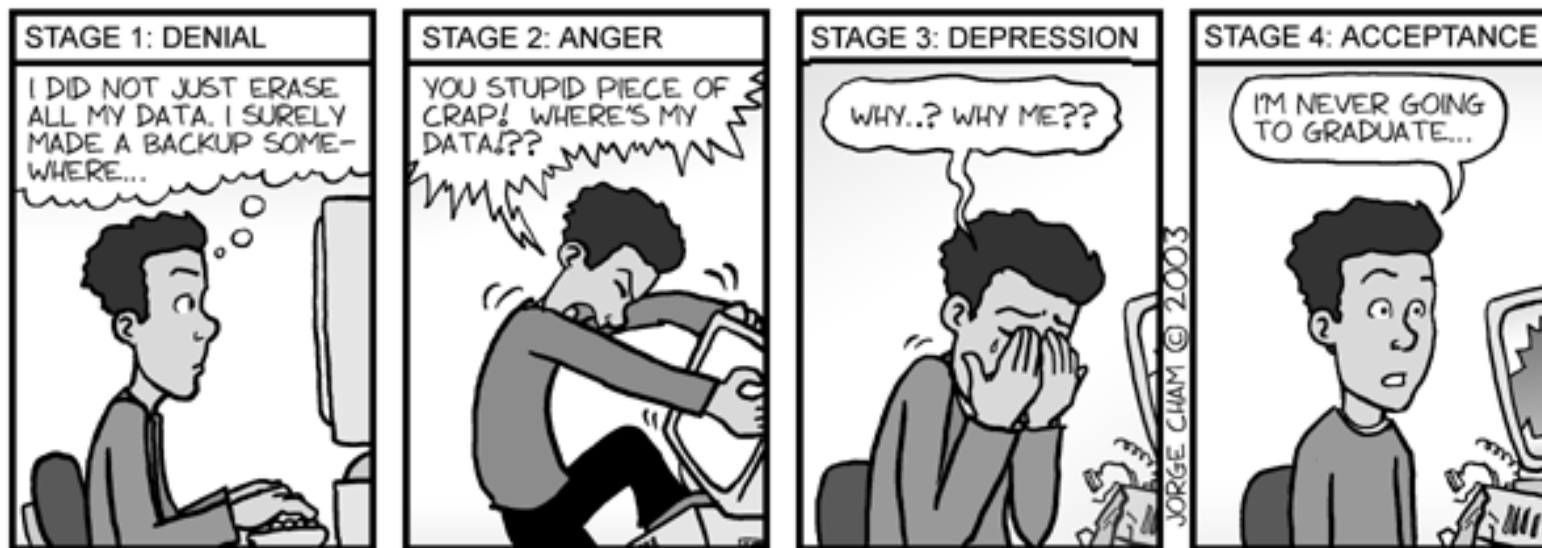
Utiliser un
**rubricage
simple** pour
organiser ses
fichiers
Adopter une
**règle de
nommage** des
fichiers

<http://fr.slideshare.net/GarethKnight/data-management-for-librarians-an-introduction>

Limiter les risques de perte des données !

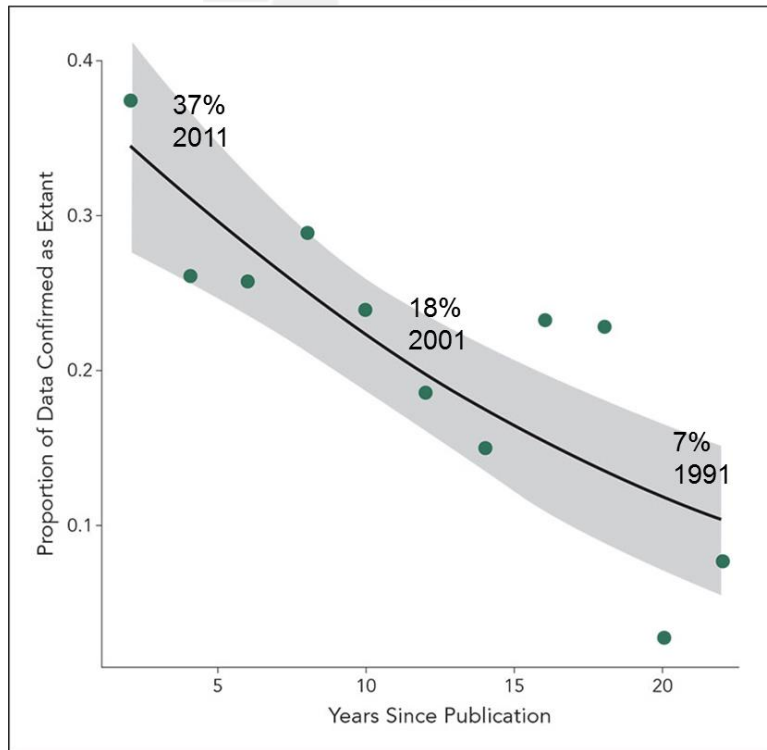
THE FOUR STAGES OF DATA LOSS

DEALING WITH ACCIDENTAL DELETION OF MONTHS OF HARD-EARNED DATA



www.phdcomics.com

20 ans après publication : 80% des données perdues



Etude (réalisée en 2013) portant sur 516 articles publiés entre 1991 et 2011. Domaine : “Morphological data from plants, animals, or other organisms”.

La disponibilité des données chute fortement au cours du temps !

(Vines et al., 2013, 2014)
(Roberge, 2015)

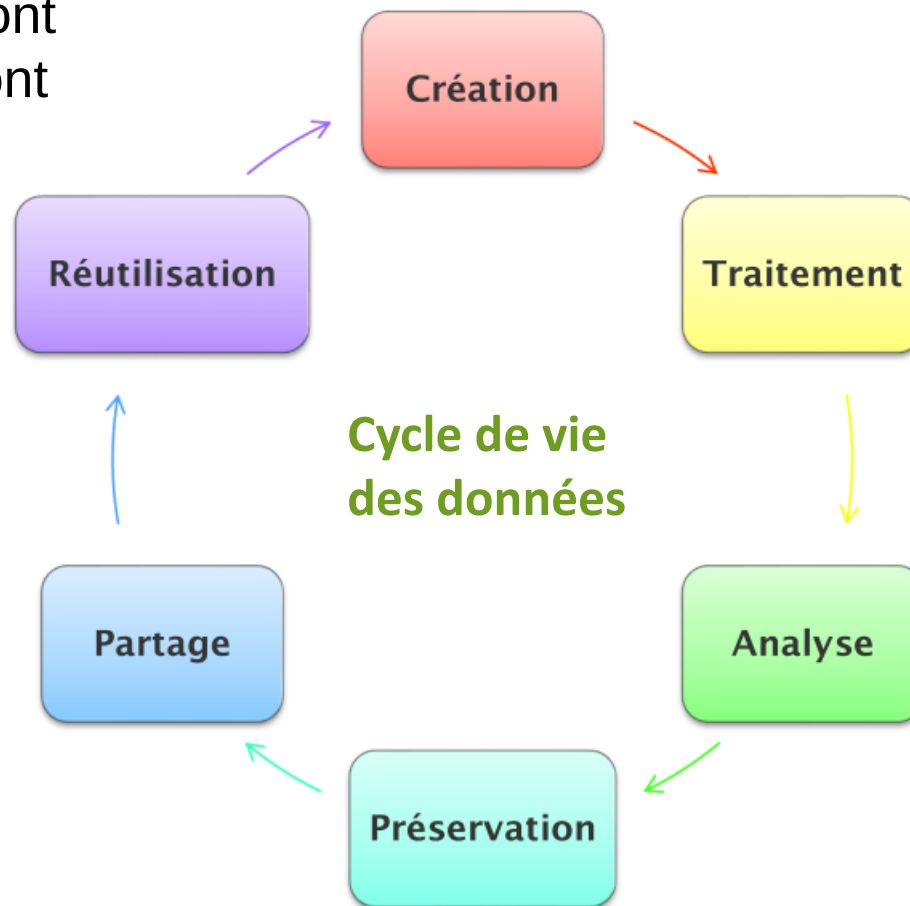
Rédiger un plan de gestion (PGD – DMP)



https://studio.cui.columbia.edu/wp-content/uploads/2013/10/DMPWantedPoster_01.png

Plan de gestion des données

- ❖ Décrit la façon dont les données seront obtenues, traitées, organisées, stockées, sécurisées, préservées, partagées,... au cours et à l'issue d'un projet



- ❖ Aide à la mise en place de bonnes pratiques de gestion à toutes les étapes du cycle de vie des données

Pourquoi rédiger un PGD ? (1)

❖ **Pour se poser les bonnes questions dès le début du projet et...**

- **identifier les risques** liés à la gestion des données, assurer la sécurité et la préservation des données sur le long terme,
- **identifier les responsabilités**, les rôles de chacun dans la gestion des données, **planifier** les ressources et compétences nécessaires à cette gestion,
- **assurer la reproductibilité** de la recherche et permettre à d'autres de **comprendre et d'utiliser** les données (« Planning for sharing »)

"Plan ahead to create
high-quality
shareable research
data"



<https://www.ukdataservice.ac.uk/manage-data/plan>

Pourquoi rédiger un PGD ? (2)

- ❖ Pour répondre aux exigences des financeurs, exemples :

 RESEARCH COUNCILS UK	• 6 des 7 principaux Research Councils
 	• National Science Foundation (NSF) • National Institutes of Health (NIH)
	• H2020 <u>Lignes directrices pour la gestion des données dans Horizon 2020 - version 2.0</u>
 Australian Government Australian Research Council	• Australian Research Council (ARC) • National Health and Medical Research Council (NHMRC)
Canada	• Recommandations plus que des exigences

Comment rédiger un PGD ?

❖ Outils en ligne

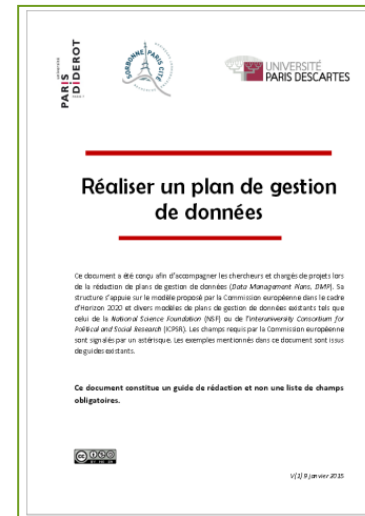
- ✓ [DMPOnline](#) (DCC, UK) - Intègre les directives H2020
- ✓ [DMPTOOL](#) (University of California Curation Center)



❖ Tutoriels et guides



Une série de tutoriels INIST
<http://www.inist.fr/?-Tutoriels-multimedias-H2020->



(Cartier, Moysan, & Reymonet, 2015)

http://www.univ-paris-diderot.fr/DocumentsFC/K/recherche/Realiser_un_DMP_V1.pdf

S'initier au concept et à la gestion des données de la recherche
Public cible : chercheurs, professionnels de l'IST <http://coop-ist.cirad.fr>

**Découvrir des plans de gestion
de données de la recherche
en 4 points**

(Deboin, 2014)
CIRAD

[Plan details](#)[Initial DMP](#)[Mid-term Review DMP](#)[Final review DMP](#)[Share](#)[Export](#)[Signed in as Sylvie Cocaud](#)[View plans](#)[Create plan](#)[About](#)[News](#)[Help](#)

Create a new plan

Please select from the following drop-downs so we can determine what questions and guidance should be displayed in your plan.

If you aren't responding to specific requirements from a funder or an institution, [select here to write a generic DMP](#) based on the most common themes.

If applying for funding, select your research funder.

Otherwise leave blank.

Funder

Funder

Arts & Humanities Research Council

Biotechnology and Biological Sciences Research Council

Cancer Research UK

Economic and Social Research Council

Engineering and Physical Sciences Research Council

European Commission (Horizon 2020)

[Contact us](#) | [Terms of use](#) | [DMPonline previous version](#)

© 2004 - 2014 Digital Curation Centre (DCC)



Bonnes pratiques de gestion ?

- **Décrire / documenter ses données**
- **Sécuriser le stockage**
- **Préserver à long terme**

Documenter ses données



❖ Indispensable pour :

- ✓ Faciliter la **recherche et la localisation** des données
- ✓ Favoriser la **compréhension** et la **reproductibilité**, la **réutilisation** des données

❖ Fournir les informations utiles

- ✓ Sous forme de **métadonnées**, en se référant à des standards (interprétation machine et/ou humaine)
 - généralistes : Dublin Core, DataCite, etc
 - spécifiques à des domaines : Environmental Metadata Language (EML), Darwin Core, INSPIRE, etc.
- ✓ Dans des fichiers séparés (interprétation humaine)
 - Cahiers de laboratoire, fichier « readme », guide d'un questionnaire ou d'une interview, etc.

Métadonnées : données sur les données

❖ Exemples

- ✓ QUI a créé les données,
- ✓ QUE contiennent-elles ?
- ✓ OU sont-elles ?
- ✓ QUAND ont-elles été créées ?
- ✓ COMMENT ont-elles été créées ?
- ✓ POURQUOI ont-elles été créées ?

Entités Dublin Core

Title: _____

Date: _____

Créateur: _____

Droits: _____

Éditeur: _____

Identifiant: _____

Source: _____

Relation: _____

Langue: _____

Mots clés: _____

Portée: _____

Description: _____

Collaborateurs: _____

Documentation, metadata, citation

Match the metadata

Drag the metadata field to its corresponding metadata category.

Hint: You may find it helpful to use the "Full Screen" button in the bottom left-hand corner, and the "Check Answers" button in the bottom right-hand corner, which will display green ticks to highlight the correct answers.

Descriptive

data creator	✓
subject	✓
title	✓
data citation	✓
location	✓
language	✓
project start and end dates	✓
funder	✓

Administrative

rights statement	✓
timestamp of transaction	✓
methodology	✓
file formats	✓
version	✓
sources	✓
Project information	✓
data processing	✓

Structural

database schema	✓
data dictionary	✓
taxonomy	✓
variable list	✓
directory & file listing	✓

Cours MANTRA
University of Edinburgh



http://datalib.edina.ac.uk/mantra/documentation_metadata_citation/

Disciplinary Metadata

While data curators, and increasingly researchers, know that good metadata is key for research data access and re-use, figuring out precisely what metadata to capture and how to capture it is a complex task. Fortunately, many academic disciplines have supported initiatives to formalise the metadata specifications the community deems to be required for data re-use. This page provides links to information about these disciplinary metadata standards, including profiles, tools to implement the standards, and use cases of data repositories currently implementing them.

For those disciplines that have not yet settled on a metadata standard, and for those repositories that work with data across disciplines, the General Research Data section links to information about broader metadata standards that have been adapted to suit the needs of research data.

Please note that a [community-maintained version of this directory](#) has been set up under the auspices of the Research Data Alliance.

Search by Discipline



Biology



Earth Science



General Research Data



Physical Science



Social Science & Humanities

<http://www.dcc.ac.uk/resources/metadata-standards>



Site du DCC - Metadata Tools
Outils qui permettent de capturer et stocker des métadonnées dans le respect des standards

Utiliser des standards de métadonnées

<http://rd-alliance.github.io/metadata-directory/>

Metadata

RDA | Metadata Directory

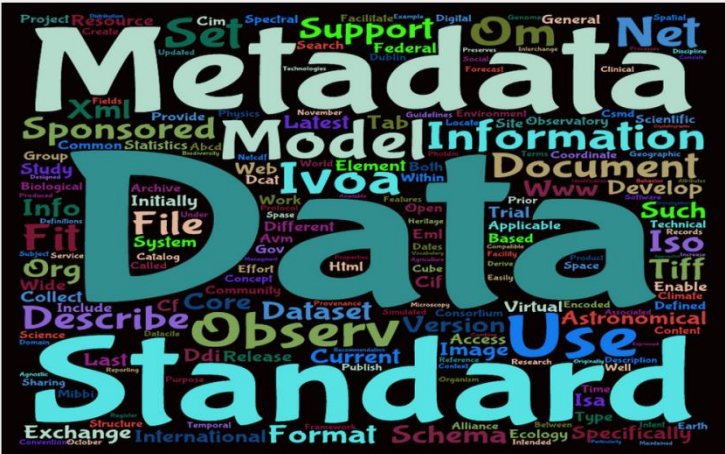
- Getting Started
- View the standards
- View the extensions
- View the tools
- View the use cases
- Browse by subject areas**

- Adding standards
- Adding extensions
- Adding tools
- Adding use cases

github

@twitter

linkedin



Metadata Standards Directory Working Group

The RDA Metadata Standards Directory Working Group is supported by individuals and organizations involved in the development, implementation, and use of metadata for scientific data. The overriding goal is to develop a collaborative, open directory of metadata standards applicable to scientific data can help address infrastructure challenges.



Home

Manage Resources

Administration

About

Basic Metadata: Collection entomologique Barthélémy

You must fill in at least these basic metadata before you can make this resource public.

For each contact you must supply at least a last name, a position, or an organisation. Title and Description are required.

Title

Collection entomologique Barthélémy

Description

Collection composée de lépidoptères français collectés dans la seconde moitié du XXe siècle

Metadata Language

French

Resource Language

French

Subtype

occurrence resource

Resource Contact

First Name

Marie-Françoise

Last Name

Faure

Position

Directrice-adjointe, Responsable des collections

Organisation

Muséum Henri-Lecoq

Address

City

Section

Basic Metadata

Geographic Coverage

Taxonomic Coverages

Temporal Coverages

Other Keywords

Associated Parties

Project Data

Sampling Methods

Citations

Collection Data

External links

Additional Metadata

<http://www.gbif.fr/page/contrib/publication-des-donnees>

Sécuriser le stockage (backup)



❖ Sécurité physique

- ✓ Copies multiples : principe LOCKSS



- ✓ Lieux (copie à distance) et technologies de stockage multiples (supports différents)

❖ Sécurité informatique

- ✓ Protection contre les virus
- ✓ Accès contrôlé
- ✓ Cryptage
- ✓ Etc...

Préserver (à long terme) ses données

- ❖ Triple enjeu : Préservation des contenus, de la lisibilité (être capable d'accéder à, de lire), de l'intelligibilité
- ❖ Différents niveaux de service selon les entrepôts
 - ✓ Stockage simple,
 - ✓ Stockage sécurisé : réplication multi-sites, différents niveaux de protection (selon la nature des données) ...,
 - ✓ Archivage pérenne : réplication + préservation (veille sur les formats et migration en cas d'obsolescence ...)
- ❖ Tout préserver ? rarement possible, ni utile
 - Les données éligibles doivent présenter un intérêt scientifique ou historique à l'avenir. Le coût de la régénération des données doit être suffisamment important pour justifier leur préservation



Consulter la [NERC Data Value Checklist](#)
(Natural Environment Research Council - UK)

Préservation : à qui s'adresser ?

- ❖ Entrepôt institutionnel
- ❖ Entrepôts thématiques
 - ✓ [Pangaea](#),
 - ✓ [Knowledge Network for Biocomplexity](#), etc
- ❖ Entrepôts pluridisciplinaires
 - ✓ Portés par des organisations publiques
([Eudat B2SHARE](#), [Zenodo](#), [3TU.Datacentrum](#), etc.)
 - ✓ Portés par des organisations privées
([Dryad](#), [Figshare](#), etc.)



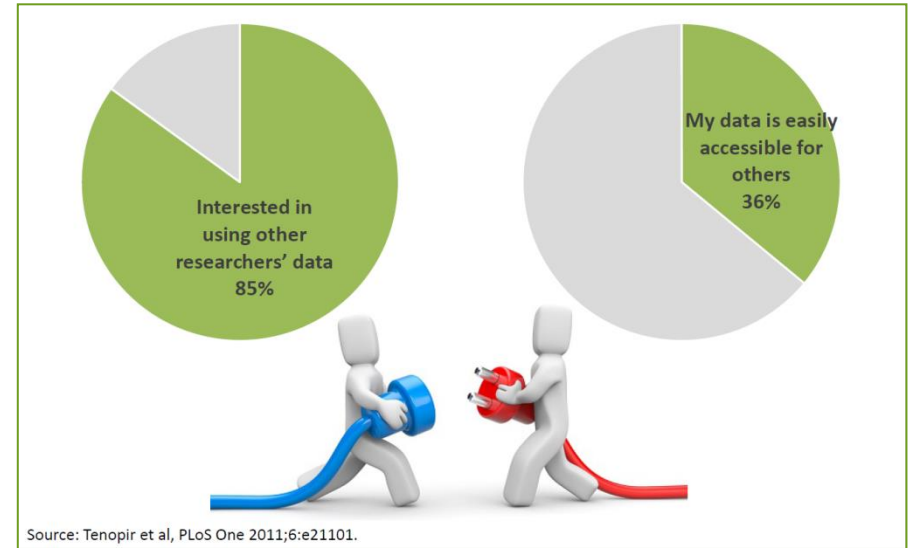
Pour choisir : Utiliser les annuaires d'entrepôts

- ✓ [Re3data](#) Registry of Research Data Repositories
- ✓ [Biosharing](#) Catalogue de Bdds – Sciences de la vie
- ✓ [NIH Data Sharing Repositories](#) ...



Des freins existent ...

❖ Chez les chercheurs



❖ Aux plans organisationnels et techniques

- Des pratiques de gestion des données trop diverses (formats, métadonnées)
- Pas de stratégie de préservation
- Manque d'infrastructures adéquates
- Manque d'incitation (Absence de politiques institutionnelles)
- Manque de reconnaissance dans le processus d'évaluation du chercheur



_03

Diffuser des données



Plusieurs stratégies pour diffuser ses données

- ❖ Publier (=déposer) ses données dans un **entrepôt**
- ❖ Fournir ses données sous la forme de **matériel supplémentaire** à la publication
- ❖ Publier ses données dans un **Data Paper**, publication scientifique spécifique décrivant les données, et publié :
 - soit dans une revue «classique»
 - soit dans un **Data Journal** qui peut fournir ou préconiser des entrepôts de confiance.
- ❖ Publier dans **le web des données (linked data)**

Données partagées vs données publiées ?



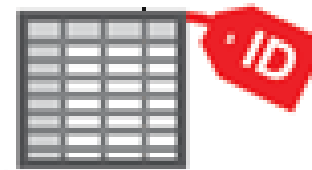
“published” traduit le fait que les données sont :

- **disponibles** au public (car déposées dans un **entrepôt**)
- **citables** (grâce à un **identifiant**),
- **documentées** (métadonnées)
- Et éventuellement **validées** par l'existence d'un processus d'évaluation (**Peer Review**)

F1000Research

(Kratz & Strasser, 2014)
(Kratz & Strasser, 2015)

Identifiants



- ❖ Extrait traduit des « Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020- 2013 »

*« Lorsque cela est possible, **il est recommandé que les contributeurs soient identifiés et les données attribuables, les deux de manière unique, grâce à des identifiants pérennes , non-propriétaires, ouverts et interopérables** (exemple : ORCID pour les identifiants de contributeurs et DataCite pour les identifiants de données)*

- ❖ Exemple d'identifiant pérenne :
 - le **DOI**-Datacite (Digital Object Identifier) : identifiant pérenne et unique permettant de référencer, citer et fournir un lien stable vers un objet numérique et sa citation.

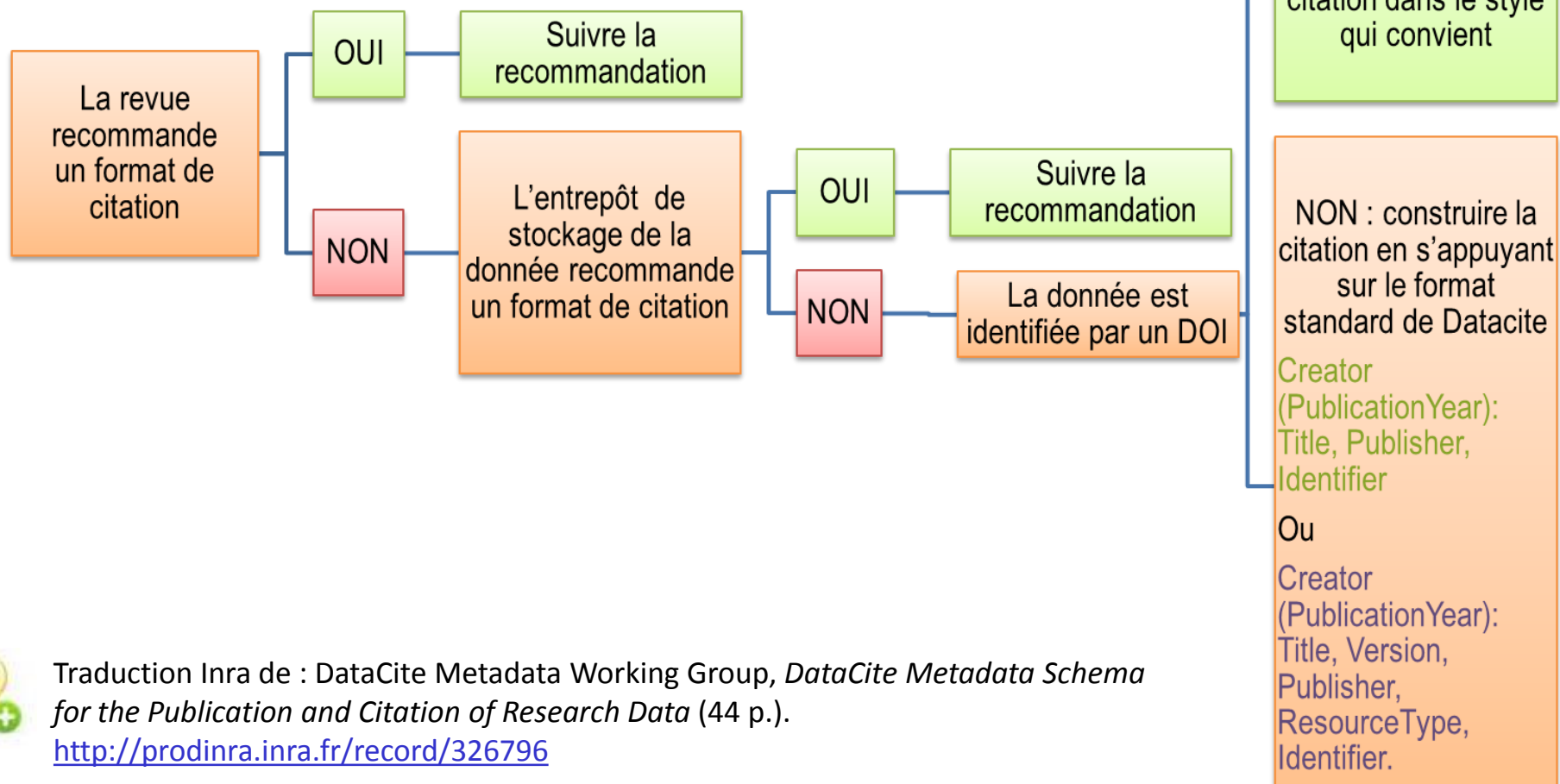
Valeur ajoutée de la liaison données & publication



- ❖ Permet de retrouver plus facilement les données (pointées par des liens),
- ❖ « Améliore la valeur contextuelle » des données et de l'article associé
 - Un lecteur peut lire le rôle ou la fonction d'une protéine ou d'un gène particulier dans l'article et accéder facilement via un lien vers les données supplémentaires enregistrées dans un entrepôt approuvé (par exemple Protein Data Bank (PDB) , <http://www.rcsb.org/pdb/home/home> ou GenBank <http://www.ncbi.nlm.nih.gov/genbank/>) pour une analyse plus approfondie
- ❖ Facilite l'interprétation des données et donc leur réutilisation en fournissant une explication sur les traitements et les résultats obtenus,
- ❖ Permet à l'auteur d'avoir une reconnaissance immédiate et supplémentaire de son travail.
 - “9-30% increase depending on data publication” (Piwowar H. et al, 2007, 2013)

La citation : élément clé

Comment citer un jeu de données dans un article ?



Traduction Inra de : DataCite Metadata Working Group, *DataCite Metadata Schema for the Publication and Citation of Research Data* (44 p.).

<http://prodinra.inra.fr/record/326796>

Préciser les conditions du partage : définir/choisir une licence?

❖ « *il faut lui (jeu de données) **apposer une licence de diffusion fixant les conditions de son utilisation** : droits d'utilisation et de modification de la donnée, droits de réutilisation commerciale et non commerciale, obligations éventuelles comme la mention de la source des données ou le partage à l'identique. »*

- **vérifiez les exigences de l'entrepôt** dans lequel vous souhaitez déposer votre jeu de données ; **il peut imposer une licence de diffusion particulière** ;
- lorsque les données sont liées à un article scientifique, la licence de diffusion choisie pour les données doit répondre aux exigences de la revue. **Consultez les instructions aux auteurs** ;
- le type de licence préconisé est donc **un critère de sélection important pour choisir un entrepôt ou un éditeur.** »

(Dedieu, 2015)

Les limites du partage ... ?

❖ Limites éthiques

- Protection des données personnelles
- Données sensibles

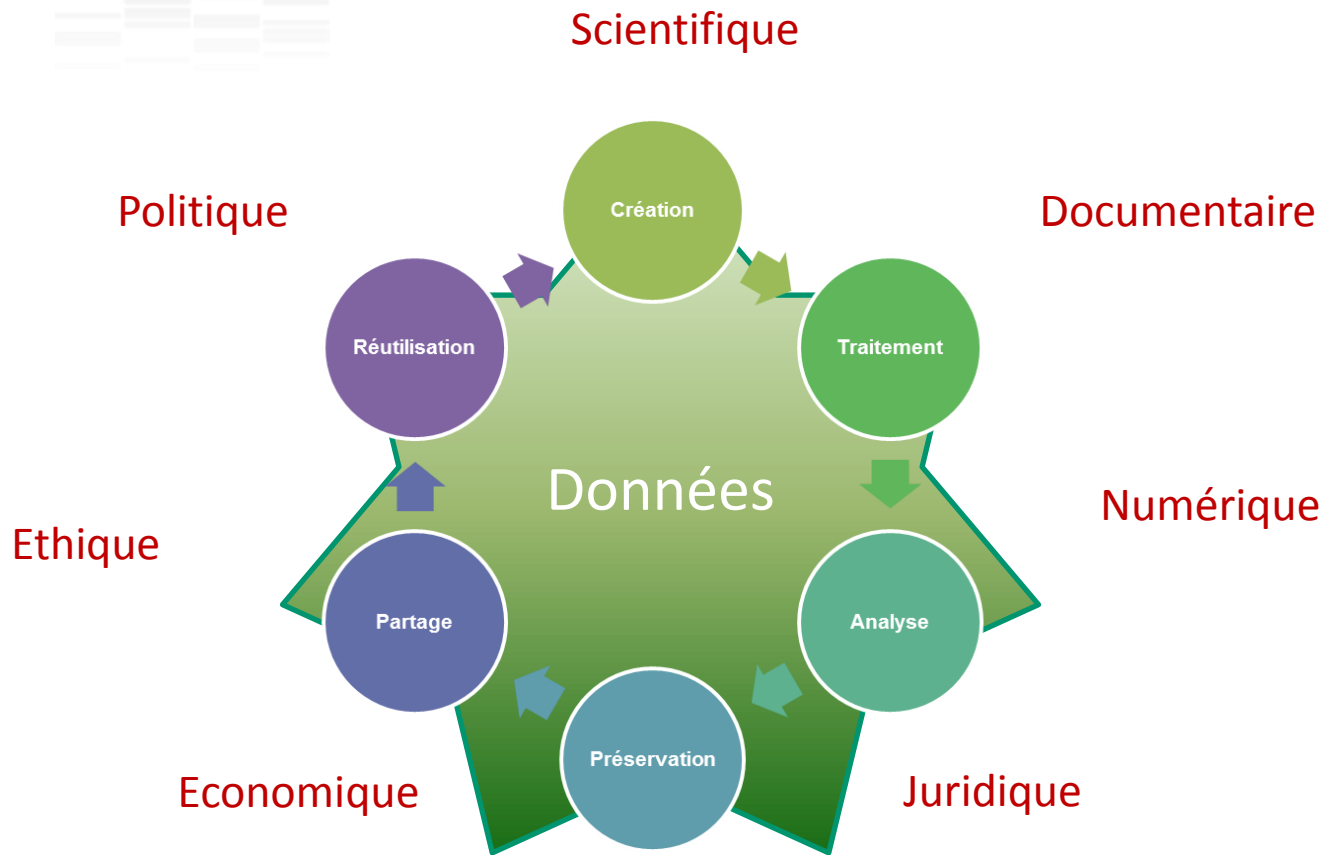
❖ Limites juridiques

- Propriété industrielle
- Propriété intellectuelle
 - Données détenues par un tiers
 - Droit d'auteur sur les données brutes ?
 - Droit des bases de données

Diapo reprise de :

Cosserat, F. – Une introduction à la gestion et au partage des données de la recherche (octobre 2015)

Les données, un objet complexe



Des acteurs et des compétences variées ...

Bibliographie (1/3)

- ❖ Battisti, M. (2010). Libérons les données ! De quelques aspects juridiques. Retrieved from <http://www.paralipomenes.net/archives/622>
- ❖ Boulton, G. (2015). *Open Data and the Future of Science (ANDS webinar)*. Presented at the Australian National Data Service, 2015/06/17. retrieved from <http://ands.org.au/presentations/2015-06-17-geoffrey-boulton-webinar-slides.pdf>
https://www.youtube.com/watch?v=795i2Hrc_YA.
- ❖ Cartier, A., Moysan, M., & Reymonet, N. (2015). *Réaliser un plan de gestion des données*: Université Paris Descartes, Université Paris Diderot. http://www.univ-paris-diderot.fr/DocumentsFCK/recherche/Realiser_un_DMP_V1.pdf
- ❖ Chaudiron, S., Maignant, C., Schöpfel, J., & Westeel, I. (2015). *Les données de la recherche dans les thèses de doctorat. Livre blanc*
- ❖ Cosserat, F. (2015/10). [Une introduction à la gestion et au partage des données de la recherche]. http://www.u1027.inserm.fr/medias/fichier/florence-cosserat_1444897589645-pdf?INLINE=FALSE
- ❖ Deboin, M. C. (2014). [Découvrir des plans de gestion de données de la recherche]. <http://coop-ist.cirad.fr/gestion-de-l-information/gestion-des-donnees-de-la-recherche/decouvrir-des-plans-de-gestion-de-donnees-de-la-recherche>
- ❖ Dedieu L., Fily M.F. 2015. Rendre publics ses jeux de données scientifiques en 6 points. Montpellier (FRA): CIRAD, 6 p. <http://url.cirad.fr/ist/rendre-publics-ses-donnees>
- ❖ Ferguson, A. R., Nielson, J. L., Cragin, M. H., Bandrowski, A. E., & Martone, M. E. (2014). Big data from small data: data-sharing in the 'long tail' of neuroscience. *Nature Neuroscience*, 17(11), 1442-1447. [10.1038/nn.3838](https://doi.org/10.1038/nn.3838)

Bibliographie (2/3)

- ❖ Gaillard, R. (2014). *De l'Open data à l'Open research data : quelle(s) politique(s) pour les données de recherche ?*. (Conservateur de Bibliothèque), Université de Lyon, ENSSIB, Lyon. Retrieved from <http://www.enssib.fr/bibliotheque-numerique/documents/64131-de-l-open-data-a-l-open-research-data-quelles-politiques-pour-les-donnees-de-recherche.pdf> . Thèse 1-104 p.
- ❖ Horstmann, W. (2015). *Beyond Big Data. The Long Tail of Research*. Presented at the E-IRG, Riga, 2015/06/03. retrieved from <https://speakerdeck.com/subugoe/lettland-3-dot-06-dot-2015>
- ❖ Knight, G. (2013). [Data Management for Librarians : an introduction]. <http://fr.slideshare.net/GarethKnight/data-management-for-librarians-an-introduction>
- ❖ Kratz, J., & Strasser, C. (2014). Data publication consensus and controverses. *F1000Res*, 3, 94. [10.12688/f1000research.3979.2](https://doi.org/10.12688/f1000research.3979.2)
- ❖ Kratz, J., & Strasser, C. (2015). [What does “data publication” mean to researchers?]. <https://zenodo.org/record/13969#.VpOBn1IXLm4>
- ❖ Lannom, L. (2013). *Facilitating new forms of discovery*. Presented at the ARL Fall Forum, 2013/10/11. retrieved from <http://www.arl.org/storage/documents/publications/ff13-lannom.pdf>
- ❖ Piwowar HA, Day RS, Fridsma DB (2007) Sharing Detailed Research Data Is Associated with Increased Citation Rate. *PLoS ONE* 2(3): e308. doi:[10.1371/journal.pone.0000308](https://doi.org/10.1371/journal.pone.0000308)
- ❖ Piwowar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. *PeerJ*, 1, e175. [10.7717/peerj.175](https://doi.org/10.7717/peerj.175)
- ❖ Reilly, S., Schallier, W., Schrimpf, S., Smit, E., & Wilkinson, M. (2011). *Report on integration of data and publications* 1-87. http://www.stm-assoc.org/2011_12_5_ODE_Report_On_Integration_of_Data_and_Publications.pdf
- ❖ Roberge, A. (2015). Les données scientifiques en voie d'extinction Retrieved from <http://cursus.edu/article/24921/les-donnees-scientifiques-voie-extinction/>

Bibliographie (3/3)

- ❖ Vines, T. H., Albert, A. Y., Andrew, R. L., Debarre, F., Bock, D. G., Franklin, M. T., Gilbert, K. J., Moore, J. S., Renaut, S., & Rennison, D. J. (2013). *How Does the Availability of Research Data Change With Time Since Publication? Presented at the Seventh International Congress on Peer Review and Biomedical Publication, Chicago, IL (USA), 2013/09/08-10. Meeting Abstract* retrieved from http://www.peerreviewcongress.org/abstracts_2013.html.
- ❖ Vines, T. H., Albert, A. Y., Andrew, R. L., Debarre, F., Bock, D. G., Franklin, M. T., Gilbert, K. J., Moore, J. S., Renaut, S., & Rennison, D. J. (2014). The availability of research data declines rapidly with article age. *Current Biology*, 24(1), 94-97. [10.1016/j.cub.2013.11.014](https://doi.org/10.1016/j.cub.2013.11.014)



Merci de votre écoute

NB : ce document emprunte de nombreuses diapositives rédigées par E. Dzale - INRA

Contact : dolhostis@nantes.inra.fr