



# Open Science INRA

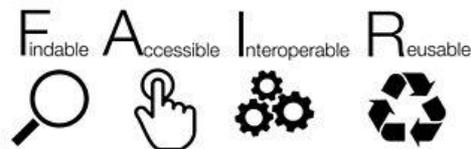
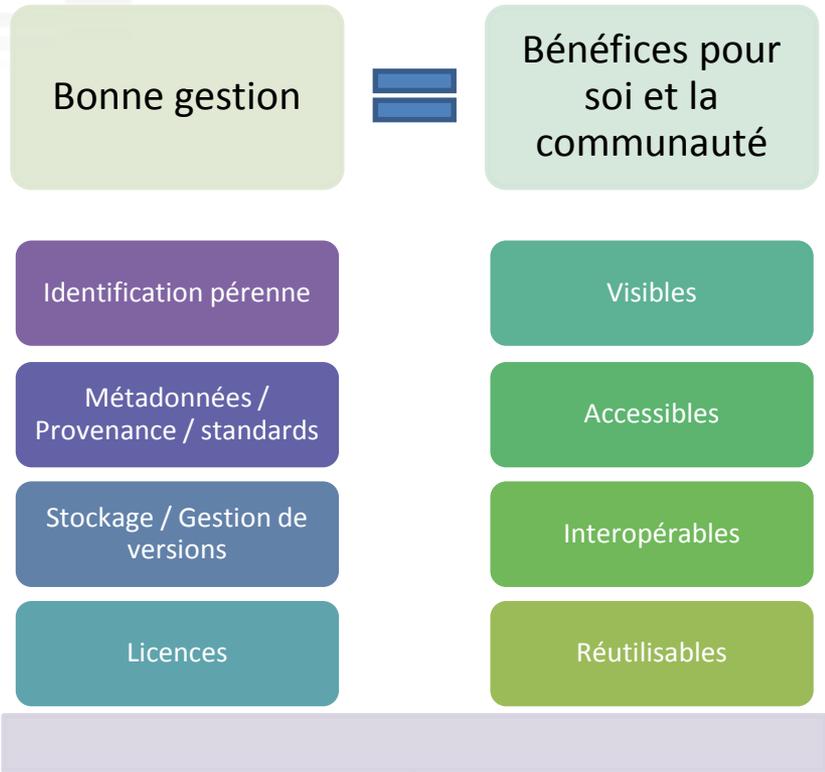
Partager des données FAIR



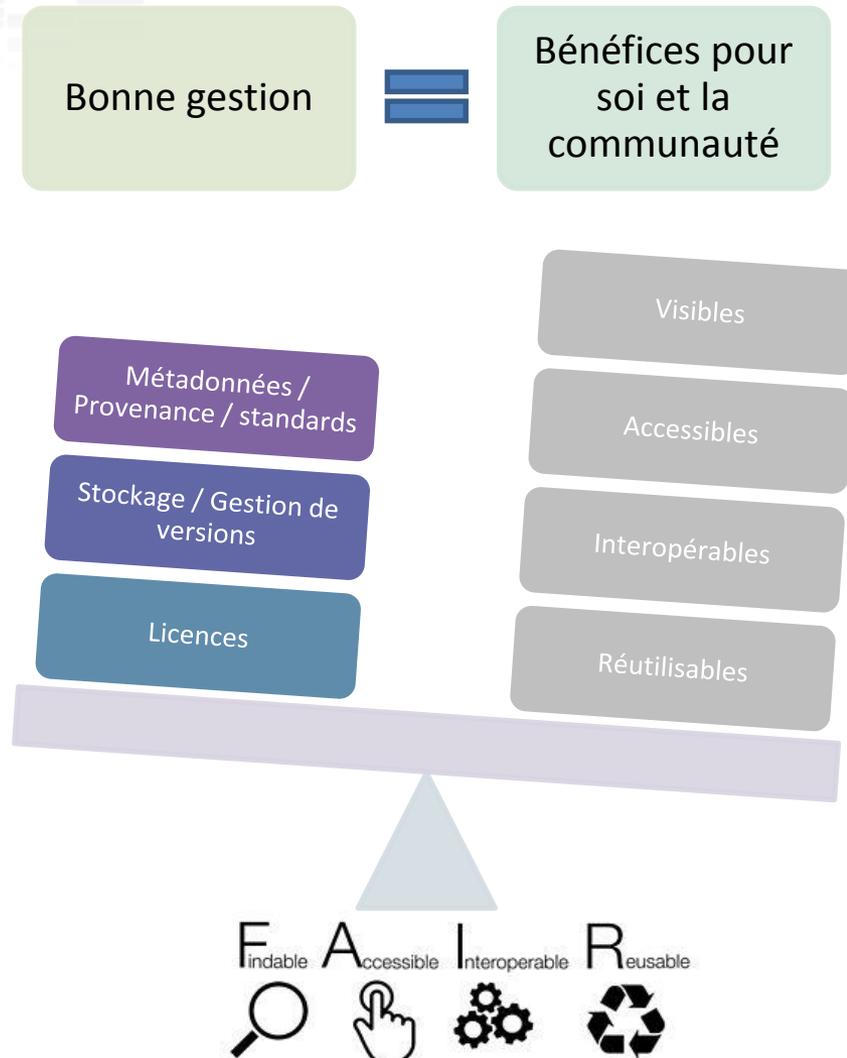
# SOMMAIRE

- ❖ Introduction aux principes FAIR
- ❖ Les données FAIR en pratique

# Les enjeux du partage des données



# Les enjeux du partage des données



# Principes FAIR : Une adoption de plus en plus large



## 1. Growing awareness of FAIRness

Open Science is a growing movement. The European Council adopted Open Science and the reusability of research data as a priority, as did the G7 at their summit in Japan [9]. This provided fertile ground for the rapid uptake of the FAIR Data Principles [25] since their recent publication [3]. The DG RTD (the Directorate General for Research and Innovation) of the European Commission took the lead [6], but in close collaboration with other directorates and the USA-based Big Data to Knowledge (BD2K) of the NIH (National Institutes of Health) [15]. Science Europe has adopted FAIR principles as the basis for sharing administrative data on funding [7]. The G20 went further in the 2016 Hangzhou summit by endorsing the FAIR Principles by name [8]. The Principles have also resonated in many discussions beyond their original scope of research data sharing, in domains as diverse as Archaeology [22], and

---

\*Corresponding author: Barend Mons, Eindhovenweg 20, 2333 ZC Leiden, P.O. Box 9600, 2300 RC Leiden, The Netherlands. Tel.: +31624879779; E-mail: [b.mons@lumc.nl](mailto:b.mons@lumc.nl).

# Qu'est-ce que les principes FAIR

## Findable

- Les (méta)données sont associées à un identifiant unique et pérenne
- Les données sont décrites avec des métadonnées riches
- Les (méta)données sont indexées et peuvent être cherchées

## Accessible

- Les (méta)données sont accessibles par leur identifiant via un protocole ouvert , standard, et permettant un accès contrôlé
- Les métadonnées restent accessibles même si les données ne le sont pas/plus

## Interoperable

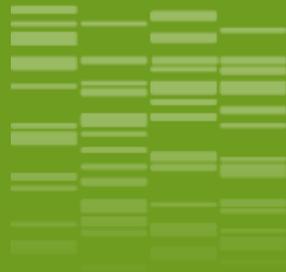
- Les (méta)données utilisent des langages formels et communs pour la représentation des connaissances
- Les (méta)données utilisent des vocabulaires compatibles avec les principes FAIR
- Les (méta)données ont des liens documentés vers d'autres (méta)données

## Reusable

- Les (méta)données sont explicitement associées à une licence
- Les (méta)données sont associées à leur provenance
- Les (méta)données sont conformes aux standards disciplinaires

# Différents types de caractéristiques

- ❖ Certaines caractéristiques sont d'ordre technique
  - ❖ Identifiant pérenne, accès aux données via un protocole libre et ouvert, etc.
- ❖ D'autres relèvent de la gouvernance des données
  - ❖ Accès aux métadonnées même quand les données ne sont plus accessibles, communicabilité des données, choix de la licence
- ❖ Certaines caractéristiques peuvent être implémentées en une fois
  - ❖ Exemples : mettre un identifiant pérenne ou mettre une licence sur un jeu de données
- ❖ D'autres sont implémentées de manière progressive
  - ❖ Richesse des métadonnées, conformité avec les standards

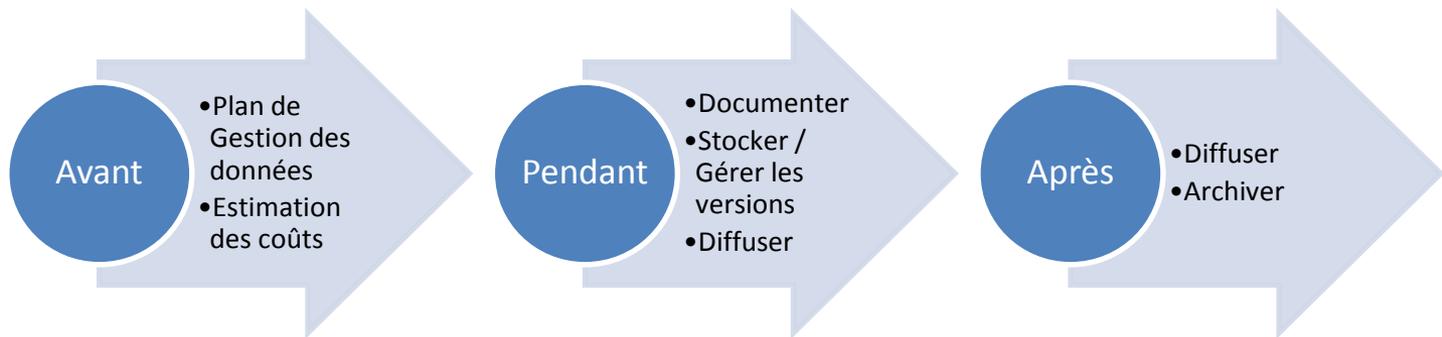


# \_02

## Les données FAIR en pratique

# Données FAIR en pratique :

Y penser avant, pendant et après le projet de recherche





# Le Plan de Gestion des Données (PGD)

- ❖ Document qui décrit la façon dont les données seront obtenues, traitées, organisées, stockées, sécurisées, préservées, partagées, ... au cours et à l'issue d'un projet
- ❖ Créé au démarrage d'un projet de recherche et **mis à jour tout au long du projet**

## Modèle de PGD pour l'Inra, utilisable pour tout type de projet, intégrant les exigences d'H2020 :

1. Informations sur le PGD
2. Informations sur le projet
3. Présentation succincte des données
4. Droits de propriété intellectuelle
5. Confidentialité
6. Partage des données à l'issue du projet
7. Description et organisation des données
8. Stockage et sécurité des données au cours du projet
9. Archivage et conservation des données après la fin du projet

## À l'Inra

- Tutoriels et guides : site <https://datapartage.inra.fr> : Gérer > Plan de gestion
- DMP Opidor (<https://dmp.opidor.fr/>)
- Classes virtuelles pour apprendre à rédiger un PGD : <https://intranet.inra.fr/ist/Se-former-a-l-IST/infodoc-express>
- Aide, conseils, relecture : [digitalist@versailles.inra.fr](mailto:digitalist@versailles.inra.fr)



# Déterminer quelles données seront partagées

- ❖ Guide juridique (<http://www6.inra.fr/datapartage/Partager-Publier/Cadre-juridique>)
- ❖ Outil d'aide à la décision pour la diffusion des données (<http://www.bibliotheque-numerique.fr/DonneesDiffusables.php>)

## Puis-je diffuser mes données de recherche ?

Veuillez répondre aux questions suivantes :	Ce que nous vous conseillons :
<p>1- Le jeu de données contient-il des données personnelles (qui identifient directement ou indirectement des individus) ?</p> <p><input type="radio"/> Oui</p> <p><input type="radio"/> Oui, mais les formalités CNIL sont respectées, ET les individus concernés ont donné leur accord écrit à la diffusion OU les données ont été anonymisées (non conservation des noms, des identifiants et des pseudonymes, cf <a href="#">CNIL</a> et <a href="#">CNRS</a>).</p> <p><input type="radio"/> Non</p> <p><input checked="" type="radio"/> Vous n'avez pas encore répondu</p> <p>2- Avez-vous utilisé d'autres données que celles que vous avez produites vous-mêmes ?</p> <p><input type="radio"/> Oui</p> <p><input type="radio"/> Oui, mais j'ai bien vérifié la licence d'utilisation de la base de données et j'ai le droit de les extraire et de les réutiliser.</p> <p><input type="radio"/> Non</p> <p><input checked="" type="radio"/> Vous n'avez pas encore répondu</p> <p>3- Est-ce que certaines de ces données, produites dans le cadre de l'établissement public, contiennent des données protégées par le droit d'auteur (textes, images, schémas, ...) ?</p> <p><input type="radio"/> Oui</p> <p><input type="radio"/> Oui, mais j'ai obtenu l'accord écrit de l'auteur.</p> <p><input type="radio"/> Non</p> <p><input checked="" type="radio"/> Vous n'avez pas encore répondu</p> <p>4- Est-ce que ces données ont été produites dans le cadre d'un consortium avec le secteur privé ?</p> <p><input type="radio"/> Oui</p> <p><input type="radio"/> Oui, mais elles peuvent être diffusées librement</p> <p><input type="radio"/> Non</p> <p><input checked="" type="radio"/> Vous n'avez pas encore répondu</p> <p>5- Votre Unité est-elle protégée ou en Zone à Régime Restrictif (ZRR) ?</p> <p><input type="radio"/> Oui</p>	<p>Vous devez encore répondre à 8 questions dans la partie de gauche</p> <p><a href="#">Exporter en PDF</a></p>



# Évaluer les coûts de gestion des données

- ❖ Les coûts de la *FAIRification* des données sont assimilables aux coûts de gestion des données, et de ce fait intégrables dans le financement des projets H2020
- ❖ Attention: ces coûts ne peuvent pas être pris en compte à posteriori des projets H2020
- ❖ Il existe des guides pour évaluer les coûts de gestion des données



## Costing guidance and examples

- Data collection: e.g. salary costs for transcribing interviews
- Storage & back-up: e.g. costs of cloud storage
- Access & security: e.g. costs of third-party data encryption
- Preservation & archiving: e.g. salary costs for converting databases to sustainable formats
- Sharing and reuse: e.g. long-term repository costs

**Data Management Cost Guide** (version June 2017)

Data Management in scientific research comprises more than the just the storage of research data. This 'Data Management Cost Guide' is set up to assist researchers to budget their data management costs. The goal is to have a practical overview in which possible costs per activity in the research process are visible. In general, the sooner certain activities are up taken (such as assigning metadata), the lower the costs will be. A good plan for data management in any case helps prevent extra work and costs. UU also has a [DMP template](#) and provides other data management information on [www.uu.nl/dm](http://www.uu.nl/dm). For those currently applying for funding: Costs for data management are eligible for funding.

DMP phase	ACTIVITY	COMMENTS AND SUGGESTIONS	COSTS
Preparing	Make a Data Management Plan	<ul style="list-style-type: none"> <li>• make a DMP before you start creating data; make decisions about managing your data; consider how you can process, analyse, preserve and share your data</li> <li>• check if there is a department within your organization to support data management planning</li> </ul>	2 hrs to 2 days, depending on the complexity of your project
1. Data Collection	Acquiring External datasets <ul style="list-style-type: none"> <li>• Do you plan to use existing data, and is the data available at a commercial partner?</li> </ul>	<ul style="list-style-type: none"> <li>• your library may be able to help you acquire a license to a crucial database</li> <li>• in research data repositories, data can be available at no or low costs</li> </ul>	Example: A faculty licence on a database for macro-economic analysis: €18.000/yr
1. Data Collection	Formatting and organizing <ul style="list-style-type: none"> <li>• Are your data files, spreadsheets, measurements, interview transcripts, records etc. all in a uniform format or type?</li> <li>• Are files, records and items in the collection clearly named with unique file names and well organized?</li> </ul>	<ul style="list-style-type: none"> <li>• if planned beforehand by developing templates and data entry forms for individual data files (transcripts, spreadsheets, databases) and by constructing clear file structures - low or no additional cost</li> <li>• if needed afterwards - higher cost</li> </ul>	Per project: organize data, format, names can be done by a student, assistant at level 1* salary or data manager at level 2* salary
1. Data Collection	Transcription <ul style="list-style-type: none"> <li>• Will you transcribe qualitative data</li> </ul>	<ul style="list-style-type: none"> <li>• if part of research practice - very low or no additional cost</li> </ul>	Example: Time needed for transcription -

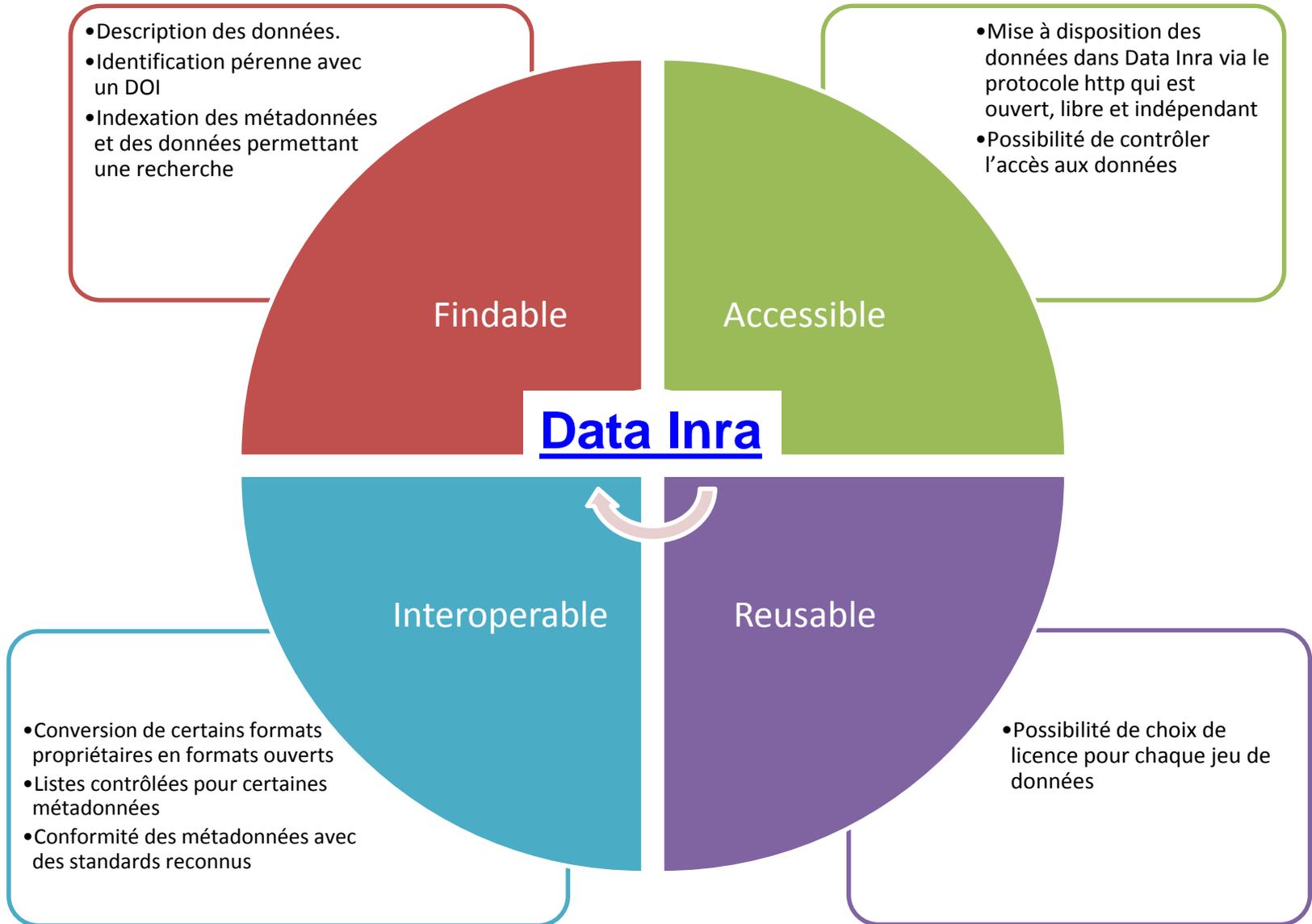
Utrecht University DM Cost Guide: <https://www.uu.nl/en/research/research-data-management/what-we-do/data-management-plan>

Pendant

- Documenter
- Stocker / Gérer les versions
- Diffuser

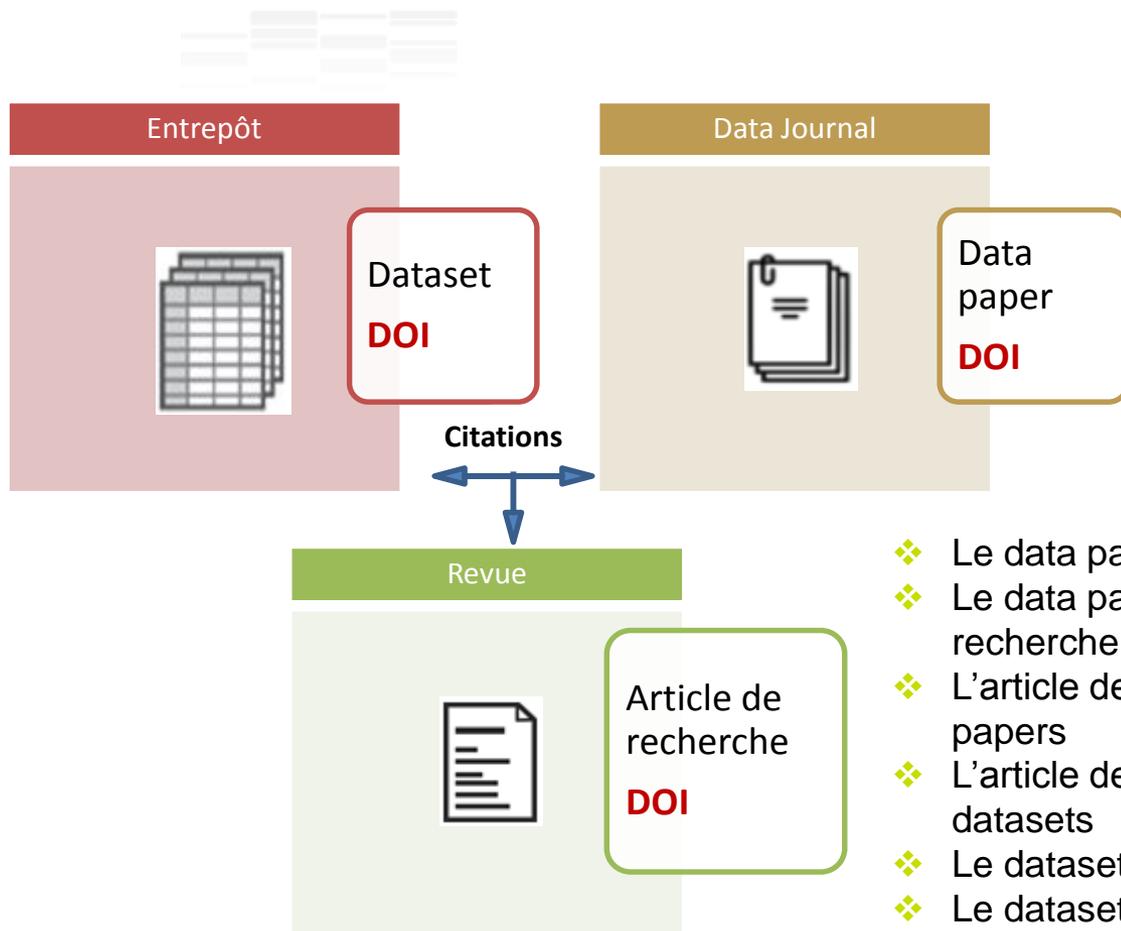
# Data Inra :

## Un service pour partager des données FAIR



# Diffuser :

dépôt des données dans un entrepôt / Data paper / article de recherche 3 possibilités complémentaires qui augmentent la visibilité des résultats de la recherche



- ❖ Le data paper cite un ou plusieurs datasets
- ❖ Le data paper cite zéro ou plusieurs articles de recherche
- ❖ L'article de recherche cite zéro ou plusieurs data papers
- ❖ L'article de recherche cite zéro ou plusieurs datasets
- ❖ Le dataset cite zéro ou plusieurs data papers
- ❖ Le dataset cite zéro ou plusieurs articles de recherche



## En résumé

- ❖ Faire un PGD en amont du projet de recherche
- ❖ Inclure une évaluation des coûts de gestion des données dans le PGD
- ❖ Utiliser des formats ouverts
- ❖ Utiliser autant que possible des standards de sa discipline
- ❖ Déposer les données dans un entrepôt de données compatible avec les principes FAIR + idéalement écrire un data paper
- ❖ Ne pas oublier de mettre une licence sur les données
- ❖ Citations croisées données / data papers / articles de recherche

# Portail d'information Datapartage



<http://datapartage.inra.fr>

- actualités,
- informations,
- accès direct aux services...

[digitalist@versailles.inra.fr](mailto:digitalist@versailles.inra.fr) pour toute information complémentaire

The screenshot shows the INRA Datapartage portal. At the top left is the INRA logo with the text 'SCIENCE & IMPACT'. Below it is a navigation menu with the following items: 'Gérer', 'Partager / Publier', 'Réutiliser', 'Technologies', and 'Documents de référence'. The main header features the text 'gestion et partage des données scientifiques' over a world map background with various scientific icons. Below the header is a 'Services, outils et bonnes pratiques recommandés par l'INRA' section. This section includes a 'Zoom sur...' sub-section with a featured article titled 'Note choix licence logicielle' and a 'Boîte à outils' section with a list of links: 'Questions/Réponses sur les données de la recherche', 'Rédiger un plan de gestion', 'Obtenir un DOI', 'Arbre de décision, statut des données', 'Déposer dans Zenodo Inra', 'Choisir un entrepôt', 'Publier un Data Paper', 'Citer des données', and 'Publier un vocabulaire/une ontologie'. There is also an 'On vous signale' section with links to 'Tutoriels, supports de formation, glossaires', 'A lire, à voir', and 'S'abonner aux actualités'. At the bottom, there are two columns of 'Actualités' (News) with titles like 'Textes officiels sur la mise à disposition des données de la recherche' and 'Délais légaux pour la diffusion des données de la recherche', each with a 'Lire la suite' link.

**INRA**  
SCIENCE & IMPACT

**Gérer**

**Partager / Publier**

**Réutiliser**

**Technologies**

**Documents de référence**

gestion et partage  
des données scientifiques

Accueil

Services, outils et bonnes pratiques recommandés par l'INRA

**Zoom sur...**

**Note choix licence logicielle**

Note sur le choix des licences logicielles suite à la parution du décret 2017-638 du 27 avril 2017 relatif aux licences de réutilisation à titre...

Lire la suite

**Boîte à outils**

- > Questions/Réponses sur les données de la recherche
- > Rédiger un plan de gestion
- > Obtenir un DOI
- > Arbre de décision, statut des données
- > Déposer dans Zenodo Inra
- > Choisir un entrepôt
- > Publier un Data Paper
- > Citer des données
- > Publier un vocabulaire/une ontologie

**On vous signale**

- > Tutoriels, supports de formation, glossaires
- > A lire, à voir
- > S'abonner aux actualités

**Actualités**

Toutes les actualités

**Textes officiels sur la mise à disposition des données de la recherche**

Où puis-je trouver les textes officiels sur la mise à disposition des données de la recherche publique ?

Lire la suite

**Délais légaux pour la diffusion des données de la recherche**

Quels sont les délais légaux pour la diffusion des données de la recherche dans le cadre de la politique open data de l'INRA ? Y-a-t-il un...

Lire la suite

**INRA**  
SCIENCE & IMPACT

# Merci!

