

# Ouverture des données de la recherche

Marin Dacos

Conseiller scientifique pour la science ouverte  
du Directeur général de la recherche et de l'innovation

# UNE MISSION POUR LA SCIENCE OUVERTE

*« Le mouvement de la science ouverte vise à construire un écosystème dans lequel la science sera plus cumulative, plus fortement étayée par des données, plus transparente, plus rapide et d'accès plus universel. »*

# POURQUOI LA SCIENCE OUVERTE ?

Argument de principe : Public money ? Public data.

Mais aussi...



① **IMPACT : CITATIONS ET LECTURES**

② **LE “DERNIER KILOMÈTRE”**

③ **EFFICACITÉ ÉCONOMIQUE**

④ **HEURISTIQUE : CUMULATIVITÉ DE LA SCIENCE VS DUPLICATE EFFORTS**

⑤ **ÉTHIQUE, INTÉGRITÉ, TRANSPARENCE**

Accessoirement : préservation !

# CONTEXTE : AMSTERDAM CALL FOR ACTION ON OPEN SCIENCE - 2016



## *Amsterdam Call for Action on Open Science*

# AMSTERDAM CALL FOR ACTION ON OPEN SCIENCE

« La gestion et le partage des données doit devenir l'approche par défaut pour les recherches financées par le secteur public. Cela impose des **définitions**, des **standards** et des **infrastructures** »

# AMSTERDAM CALL FOR ACTION ON OPEN SCIENCE

## CONCRÈTEMENT :

- Mettre en place des politiques institutionnelles concernant les données de la recherche
- Mettre en place des Data Management Plans (DMPs), faisant partie du processus de recherche, et l'établissant comme condition pour l'obtention de fonds
- Introduire des incitations positives à l'ouverture des données à leur « fairisation » : promouvoir et récompenser les pratiques d'ouverture.
- Tout en définissant l'ouverture comme le principe par défaut, prévoir des dérogations avec une gestion des accès en fonction des contraintes disciplinaires, légales et économiques.
- Actuellement, les chercheurs sont placés devant des injonctions contradictoires face à la science ouverte. Réformer le système d'évaluation pour encourager le partage des données.
- Former les chercheurs aux principes, méthodes et bonnes pratiques de la science ouverte.
- Former des experts des données scientifique et offrir de véritables carrières à ces métiers, qui seront le pont entre l'informatique et la recherche.

# NATIONAL PLAN FOR OPEN SCIENCE (NL)



1. Promoting open access to scientific *publications*.
2. Promoting optimal use and reuse of *research data*.
3. Adapting evaluation and award systems to bring them into line with the objectives of open science (*reward systems*).

# L'EXEMPLE SUISSE

## Open Research Data



Les données de la recherche devraient être libres d'accès pour chacun, tant pour la communauté scientifique que pour le public.

Le FNS soutient ce principe. **Dès octobre 2017**, il introduira de nouvelles mesures dans l'encouragement de projets et étendra progressivement ses principes ORD à d'autres instruments d'encouragement, là où cela s'y prête. Il sera alors obligatoire d'inclure un plan de gestion des données (Data Management Plan ou DMP) avec chaque requête. Le FNS s'attendra en outre à ce que les données produites au cours des travaux de recherche soient archivées dans des bases de données numériques en libre accès, pour autant qu'aucune clause juridique, éthique, concernant le droit d'auteur ou autre ne s'y oppose.

# STRUCTURER : 95 INFRASTRUCTURES NATIONALES DE RECHERCHE

« Les infrastructures de recherche devront à l'avenir être en mesure de **mettre à disposition les données produites, soit immédiatement, soit après une période d'embargo** correspondant aux pratiques internationales du domaine concerné. »



# FAIR DATA PRINCIPLES

Une nouvelle formulation pour des principes très anciens.

F  
indable



A  
ccessible



I  
nteroperable



R  
eusable



# FAIR DATA PRINCIPLES

## Box 2 | The FAIR Guiding Principles

### To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

### To be Accessible:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
  - A1.1 the protocol is open, free, and universally implementable
  - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

### To be Interoperable:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

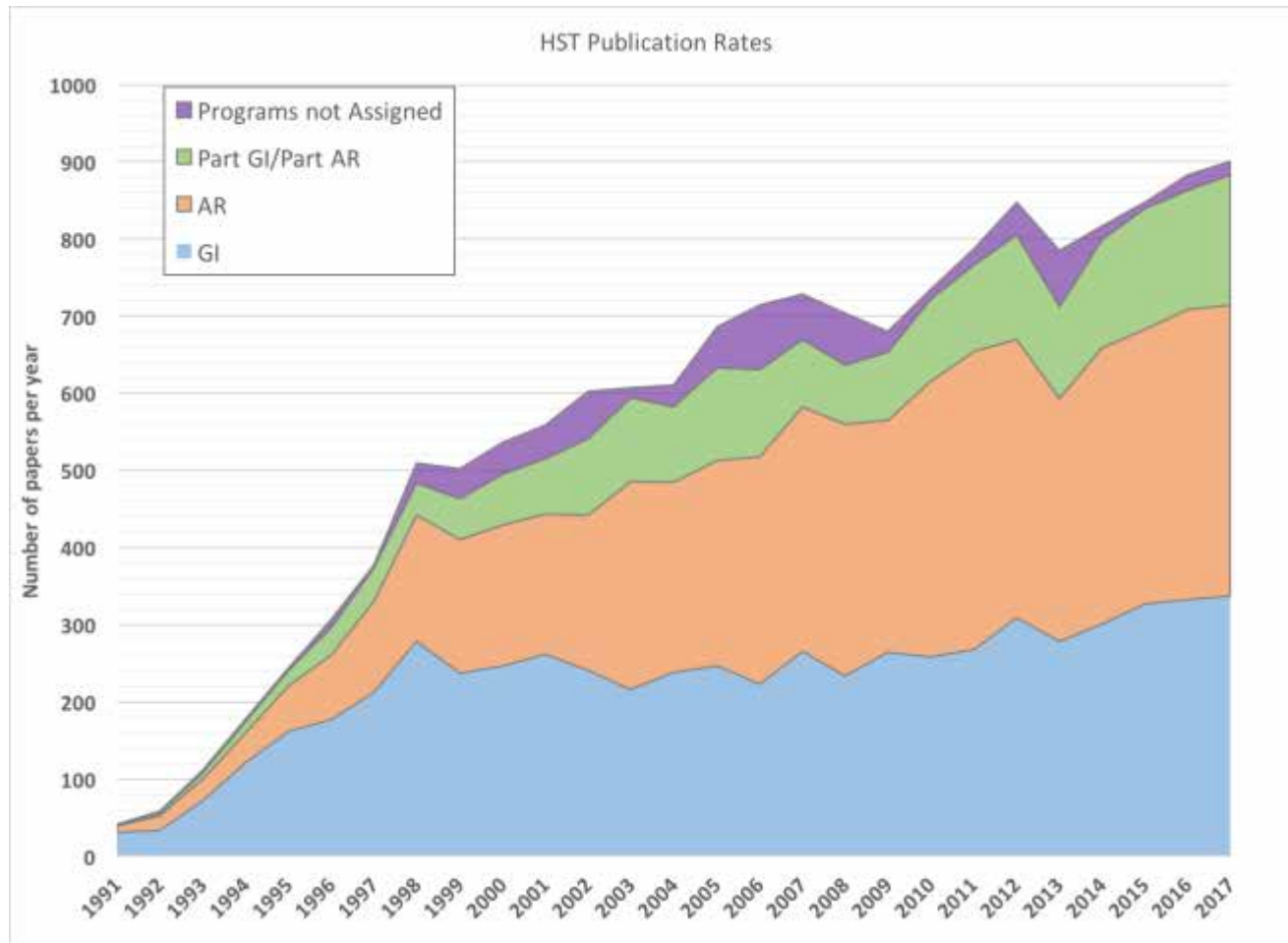
### To be Reusable:

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
  - R1.1. (meta)data are released with a clear and accessible data usage license
  - R1.2. (meta)data are associated with detailed provenance
  - R1.3. (meta)data meet domain-relevant community standards

# LOI POUR UNE RÉPUBLIQUE NUMÉRIQUE D'OCTOBRE 2016

- Un nouveau devoir pour les acteurs publics
- L'article 6 définit le principe d'ouverture par défaut pour les données administratives : données construites dans le cadre des missions de service public, auxquelles les données de recherche appartiennent. *Il crée donc une obligation.*

# CAS DE L'ASTRONOMIE : LES DONNÉES « ARCHIVES » PRODUISENT PLUS DE PUBLICATIONS QUE LA COLLECTE INITIALE.



# QUESTIONS SOULEVÉES PAR L'OUVERTURE DES DONNÉES

## 1. C'est compliqué

C'est la raison pour laquelle les infrastructures de données et les plans de gestion des données existent (DMP).

## 2. C'est cher

C'est la raison pour laquelle les frais de structuration et d'ouverture sont éligibles dans les appels à projets.

## 3. C'est risqué pour les données sensibles

Données personnelles sans consentement. Données présentant des risques pour la sécurité publique ou la sécurité de l'établissement. Secret de la défense nationale, secret professionnel, secret des affaires, secrets industriels et commerciaux. Données présentant des risques pour la protection du potentiel scientifique et technique de la nation.

Dans tous ces cas, il ne faut pas ouvrir les données, mais on peut les fairiser.

**Aussi ouvert que possible, aussi fermé que nécessaire.**

# QUE FAIRE EN ÉCOLE DOCTORALE ?

## ■ Former aux compétences de la science ouverte.

- Objectifs à définir, en prenant toujours en compte la dimension disciplinaire : les écoles doctorales sont le bon niveau de granularité. Comité pour la science ouverte peut être mobilisé pour y réfléchir.
- Publications : accès ouvert, bonnes pratiques, paysage, bibliométrie, droit, éthique...
- Données : formation aux fondamentaux disciplinaires, paysages, outils, bonnes pratiques, études de cas, DMP, droit, éthique...

## ■ Encourager à rédaction d'un DMP pour la thèse.

- Sera souvent articulé avec le DMP de labo ou d'établissement

## ■ Encourager à publier *sous forme structurée* les données citées dans de la thèse.

## ■ Encourager à publier les données de la thèse dans le respect de la législation, des principes FAIR et d'un éventuel embargo définit par la communauté scientifique.

## ■ Recommander le dépôt des données dans des entrepôts certifiés.

# CRÉATION DU COMITÉ POUR LA SCIENCE OUVERTE : APPEL À MANIFESTATION D'INTÉRÊT

- Publications
- Données de la recherche (dont code source)
- **Compétences**
- International

Les groupes du Comité devront répondre à la question :

*« Que faut-il faire pour que la recherche soit plus ouverte, plus durable, plus mutualisée, plus transparente et d'accès plus universel ? »*

# DMP OPIDOR

POUR FACILITER LA RÉDACTION DE PLANS DE GESTION DE DONNÉES



# DORANUM

DONNÉES DE LA RECHERCHE : APPRENTISSAGE NUMÉRIQUE À LA GESTION ET AU PARTAGE

**DORA Num** ACCUEIL ACTUALITÉS THÉMATIQUES RESSOURCES À IMPRIMER LE PROJET CONTACT

## La minute publier un data paper

INRA Contenu développé en partenariat avec l'Inra

Comment publier un data paper ? Cette vidéo schématise toutes les étapes pour y parvenir.

\* Facultatif si données intégrées dans l'article

### La minute "Publier un Data paper"

- 1 Choisir une revue  
Data Journal, Peer-review
- 2 Choisir un entrepôt & déposer les données  
données, entrepôt, DOI, Recommandé
- 3 Rédiger le data paper & la cover letter  
DATA PAPER, Cover letter, Instructions auteurs, Modèles...
- 4 Lier le data paper aux données déposées  
DATA PAPER, entrepôt, données, DOI
- 5 Soumettre le data paper  
Originalité des données, Cover letter, DATA PAPER
- 6 Attribuer un identifiant au data paper  
DOI, Data paper citable, Liens réciproques

Script (PDF)

**Télécharger le dossier ZIP**

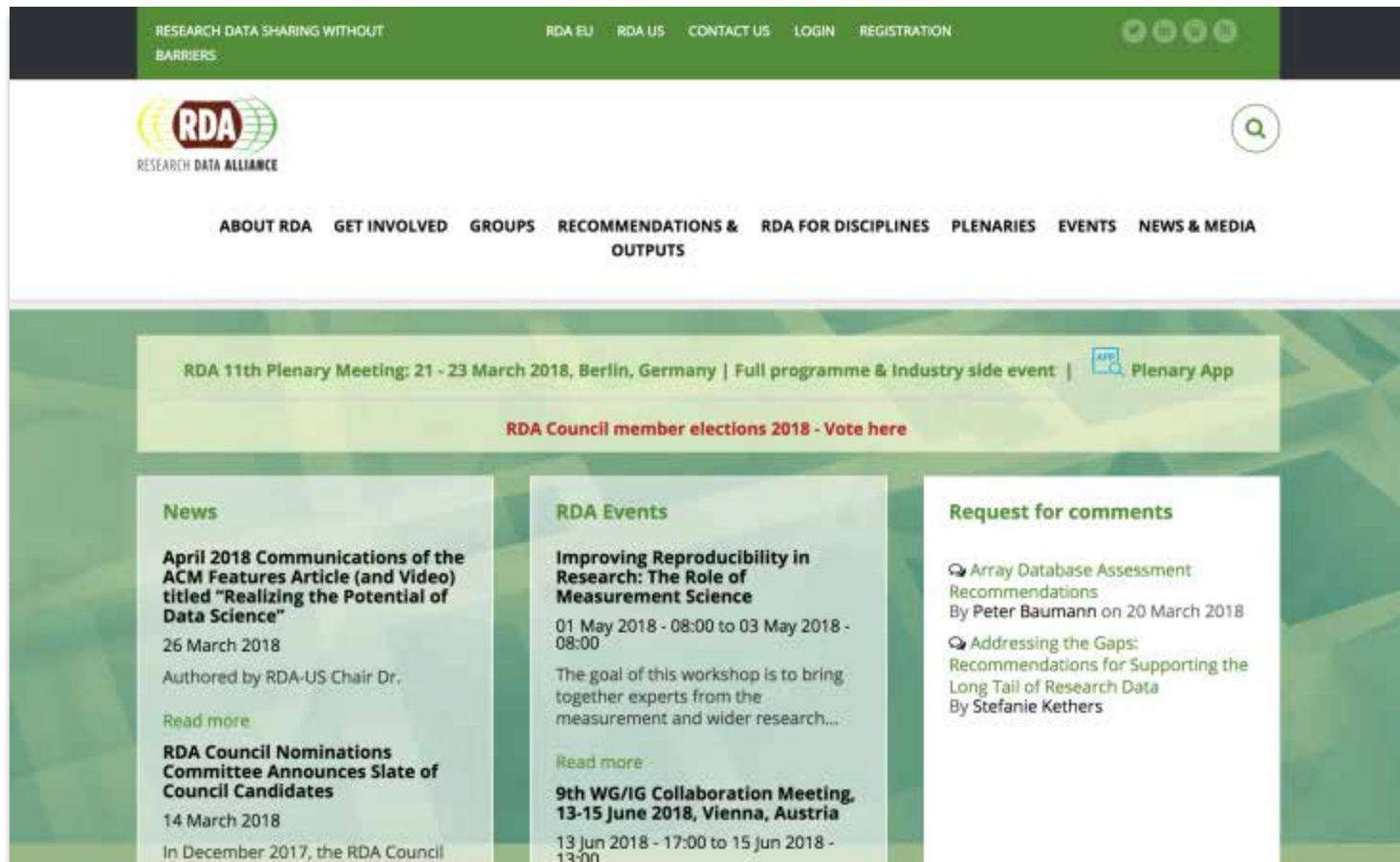
Ce dossier comprend :

- Les informations sur la ressource
- La vidéo (MP4)
- Le script (PDF et TXT)

<http://doranum.fr/>

# RESEARCH DATA ALLIANCE

"RESEARCH DATA SHARING WITHOUT BARRIERS"



<https://www.rd-alliance.org/>



LIBERTÉ • ÉGALITÉ • FRATERNITÉ  
RÉPUBLIQUE FRANÇAISE

MINISTÈRE  
DE L'ENSEIGNEMENT SUPÉRIEUR,  
DE LA RECHERCHE  
ET DE L'INNOVATION

## 20

# Merci !

Marin.dacos@recherche.gouv.fr

Twitter : @marindacos