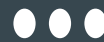




# Présentation issue de discussions au sein d'un groupe de travail du COSO et non validé par le COSO



Cette présentation a été faite à la demande de l'association VPNum



# Open data : stratégie et déploiement

---

**K. Maussang**

*Université de Montpellier*

**kenneth.maussang@umontpellier.fr**



UNIVERSITÉ DE MONTPELLIER

# Comité pour la Science Ouverte (CoSO)

Comité de pilotage de la science ouverte (Copil SO)

Secrétariat permanent (SPSO)

Collèges

Publications

Données de la recherche

Compétences

International

Groupes de travail et projets

269 experts (déc. 2018)

50%

Chercheurs ou  
Enseignants-chercheurs

41%

STM

48%

Universités

55%

femmes

Collège « données »

GT Gouvernance et usages

Hélène Chambefort, INSERM

Juliette Hueber, INHA - CNRS

Claire Lemercier, Science Po - CNRS

Kenneth Maussang, Univ. de Montpellier

Anne Vanet, Univ. de Paris Diderot

<https://www.ouvrirlascience.fr>



UNIVERSITÉ DE MONTPELLIER

# Contexte : LPPR et politique SO du CNRS



## 8. Accélérer le développement de la Science ouverte

- |   |                                                                                                                                                                                             |
|---|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| a | Inciter les opérateurs de recherche à élaborer des feuilles de route pour le développement de la Science ouverte, prenant notamment en compte la problématique des données de la recherche. |
|---|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

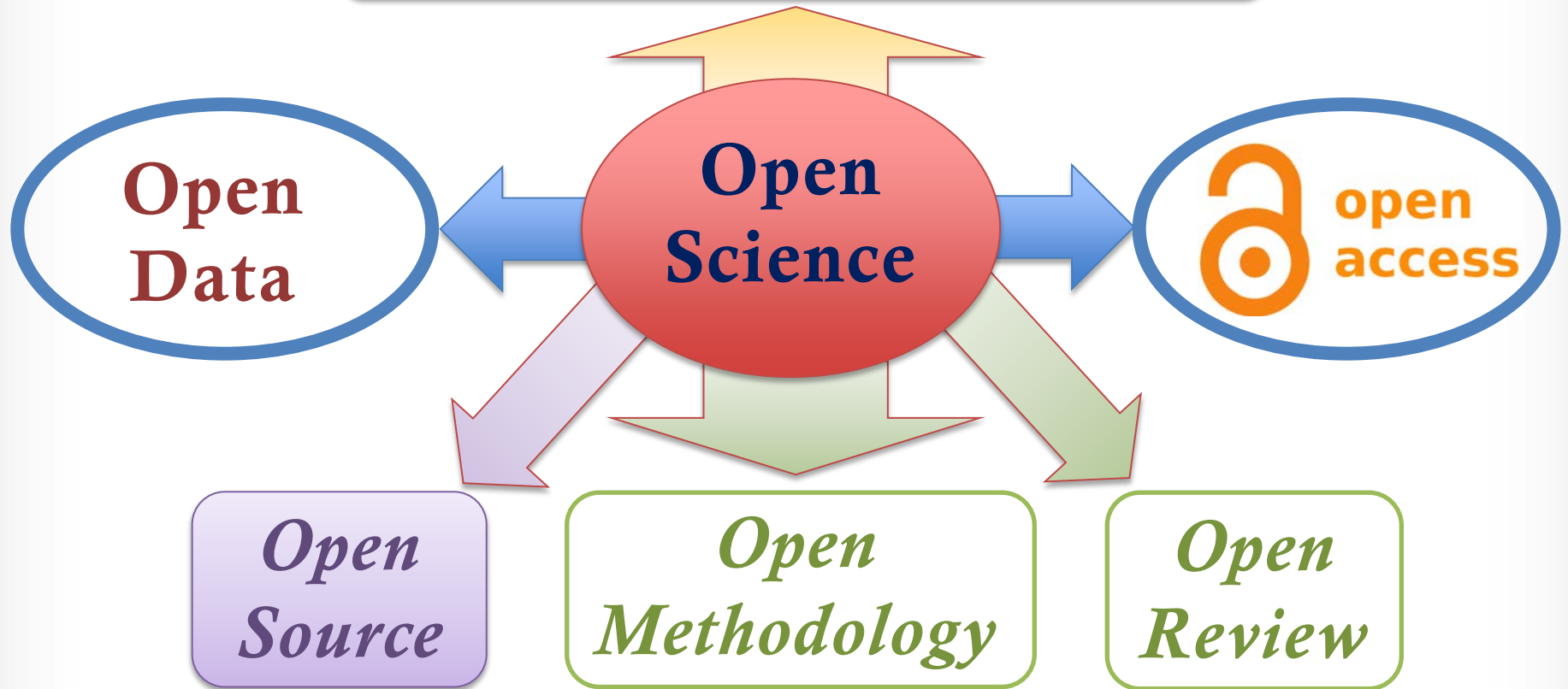


<http://www.cnrs.fr/fr/le-cnrs-se-dote-dune-feuille-de-route-pour-la-science-ouverte>



# What is Open Science ?

*Open Educational ressources*



# *Ouverture des données de la recherche*

**Pourquoi ?**

**Pour qui ?**

**Comment ?**



# *Ouverture des données de la recherche*

**Pourquoi ?**



Motivations des équipes de recherche ?

Opportunités ?



# *Evolution des pratiques*

## *Injonctions contradictoires*



Science ouverte  
*VS*  
« valorisation »

Science ouverte  
*VS*  
RGPD

Science ouverte  
*VS*  
concurrence accrue entre chercheurs



# *Opportunités*

## **Réutilisation au sein d'une discipline**

- Éviter les redondances (optimisation de l'allocation des moyens) ;
- Théoricien vs expérimentateur ;
- Reproductibilité / intégrité scientifique.

## **Favoriser le dialogue entre les disciplines**

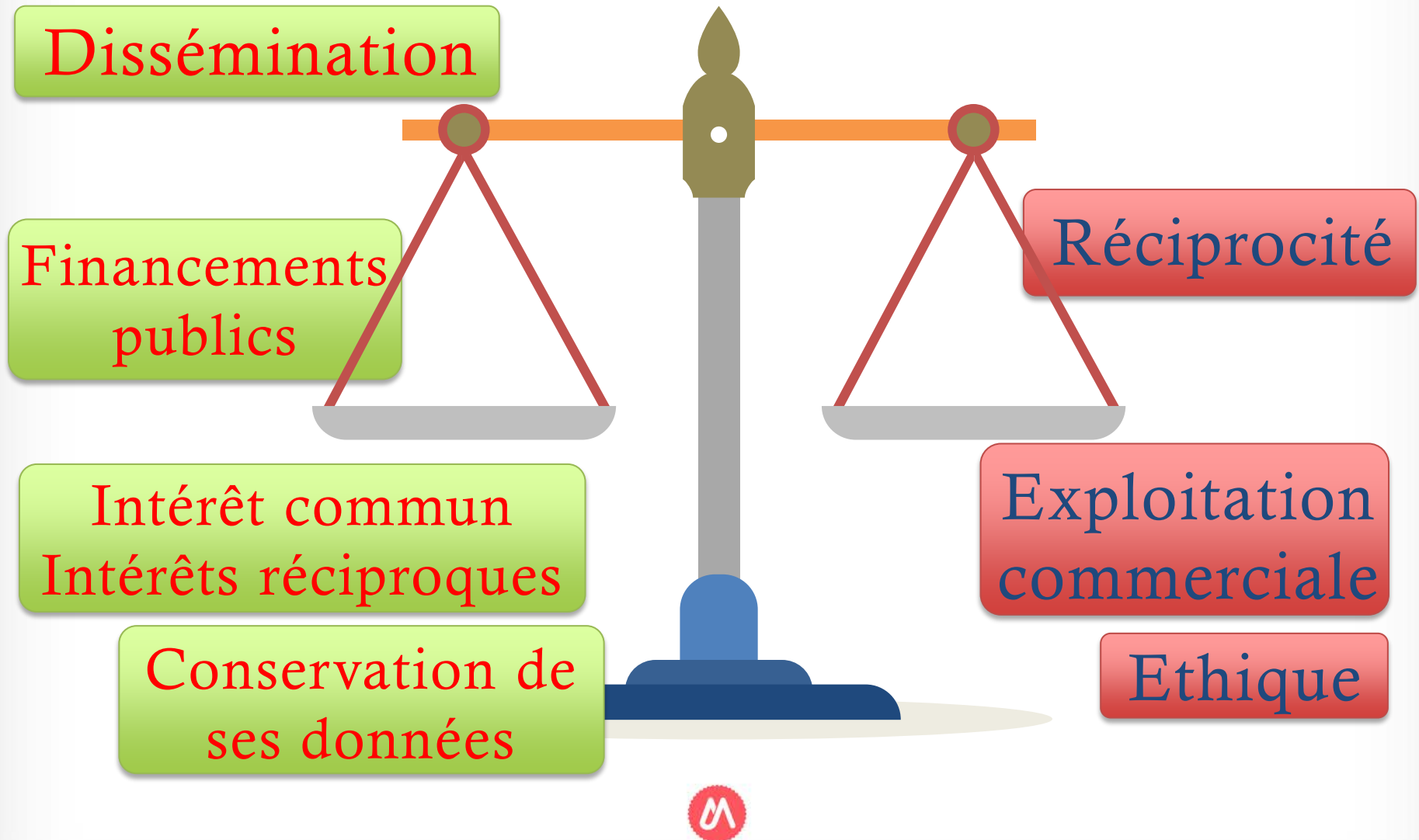
- Émergence de nouvelles disciplines (sociophysique).

## **Réutilisation par les entreprises, les administrations, les associations, le « grand public »**

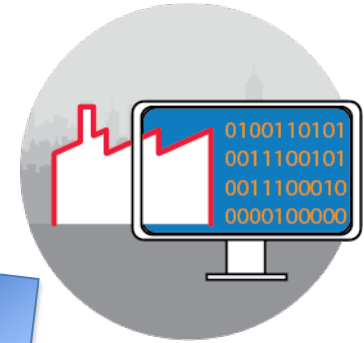
## **Réutilisation par les étudiants et les enseignants**



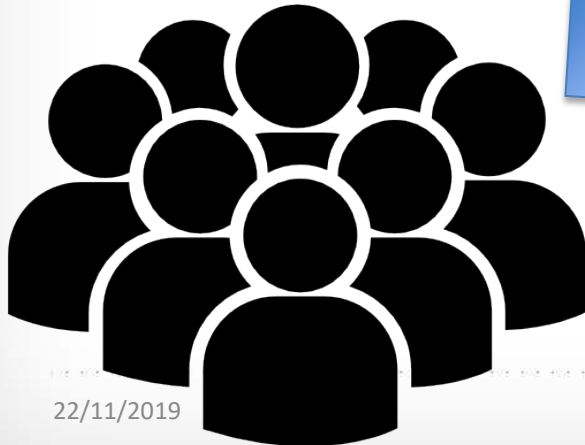
# *Motivations des chercheurs ?*



# *Ouverture des données de la recherche*

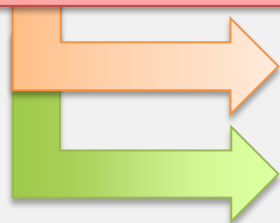


**Pour qui ?**



# Ouverture des données de la recherche

Variétés des publics cibles des données



Quels formats pour quels publics ?

Quels modes de mise à disposition ?

Définir les conditions de réutilisation !

Définir le  
public  
destinataire

Grand public

*structuration  
éditorialisation*

Enseignants

*Jeux de données  
simplifiés*

Entreprises, chercheurs

Rendre des données brutes  
massives utilisables par un tiers

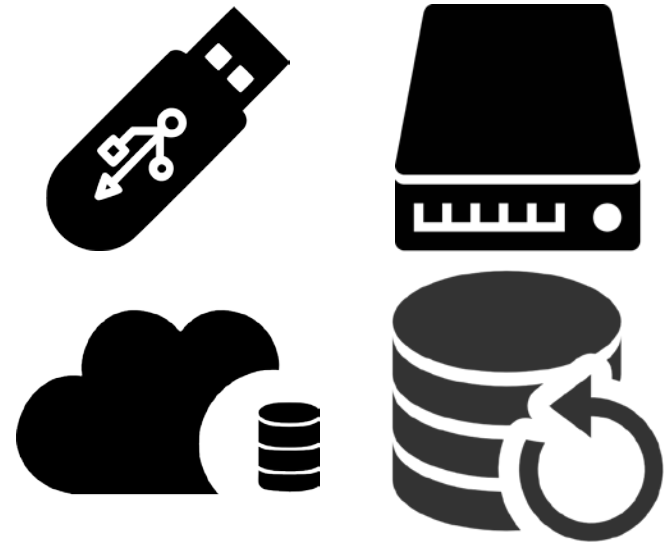
# Ouverture des données de la recherche

Conservation  
des données

*STOCKAGE*

≠

*ARCHIVAGE*



UNIVERSITÉ DE MONTPELLIER

<https://www.cleanpng.com>

# *Formats communs de description des données*

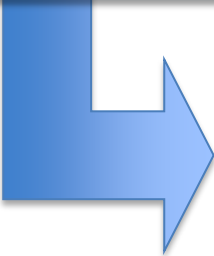
*Degrés de maturités variés suivant les domaines.*

**Instruments partagés** (astrophysique, génétique,...)

→ émergence naturelle de formats communs.

**Données non standardisées** (SHS, Physique,...)

→ données aux formats très liés à la spécificité du sujet d'étude.



**Identifier les disciplines les moins matures.**

**Accompagner le changement de pratique en tenant compte des spécificités disciplinaires.**



# *Nécessité d'un accompagnement juridique.*

## *A qui appartiennent les données ?*

**Qui possède  
les données ?**

**Qui a le droit de  
faire quoi avec  
ces données ?**

**Droits et obligations  
de diffusions ?**

**Conflits  
PI/consortium industriel  
et ouverture ?**



# Guides juridiques et logigrammes établissements

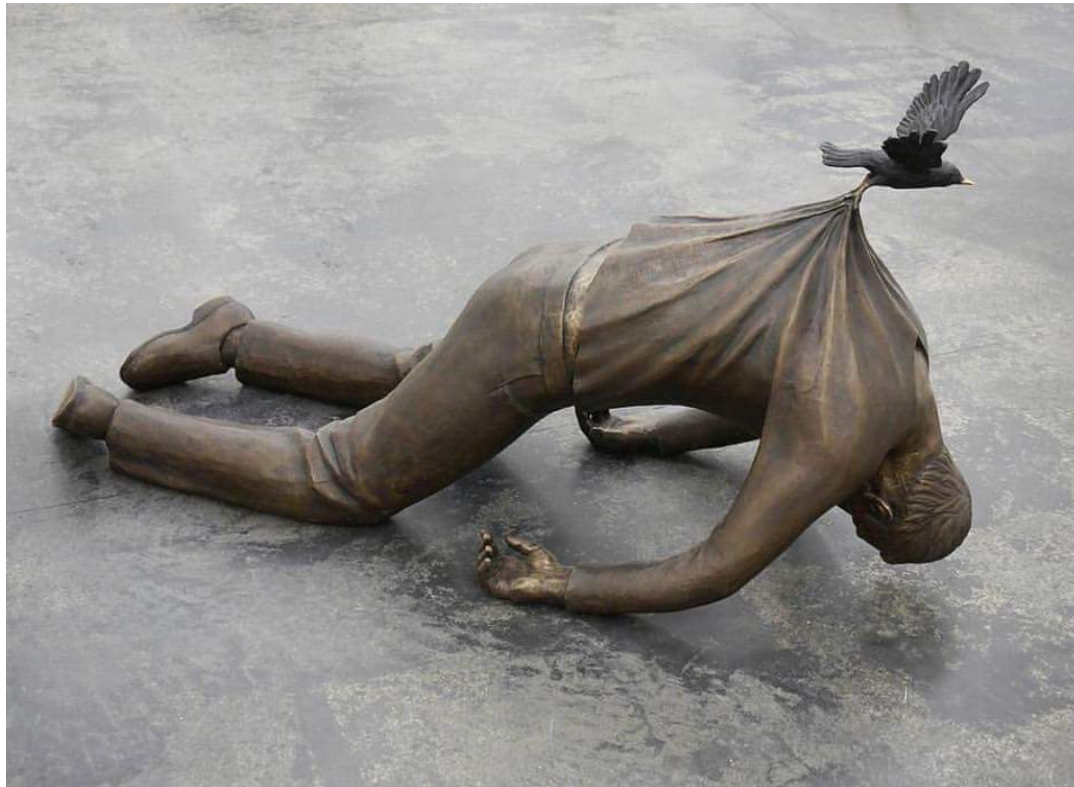
UNIVERSITÉ

Kenneth MAUSSANG – bootcamp VPnum 2019



# *Ouverture des données de la recherche*

**Comment ?**

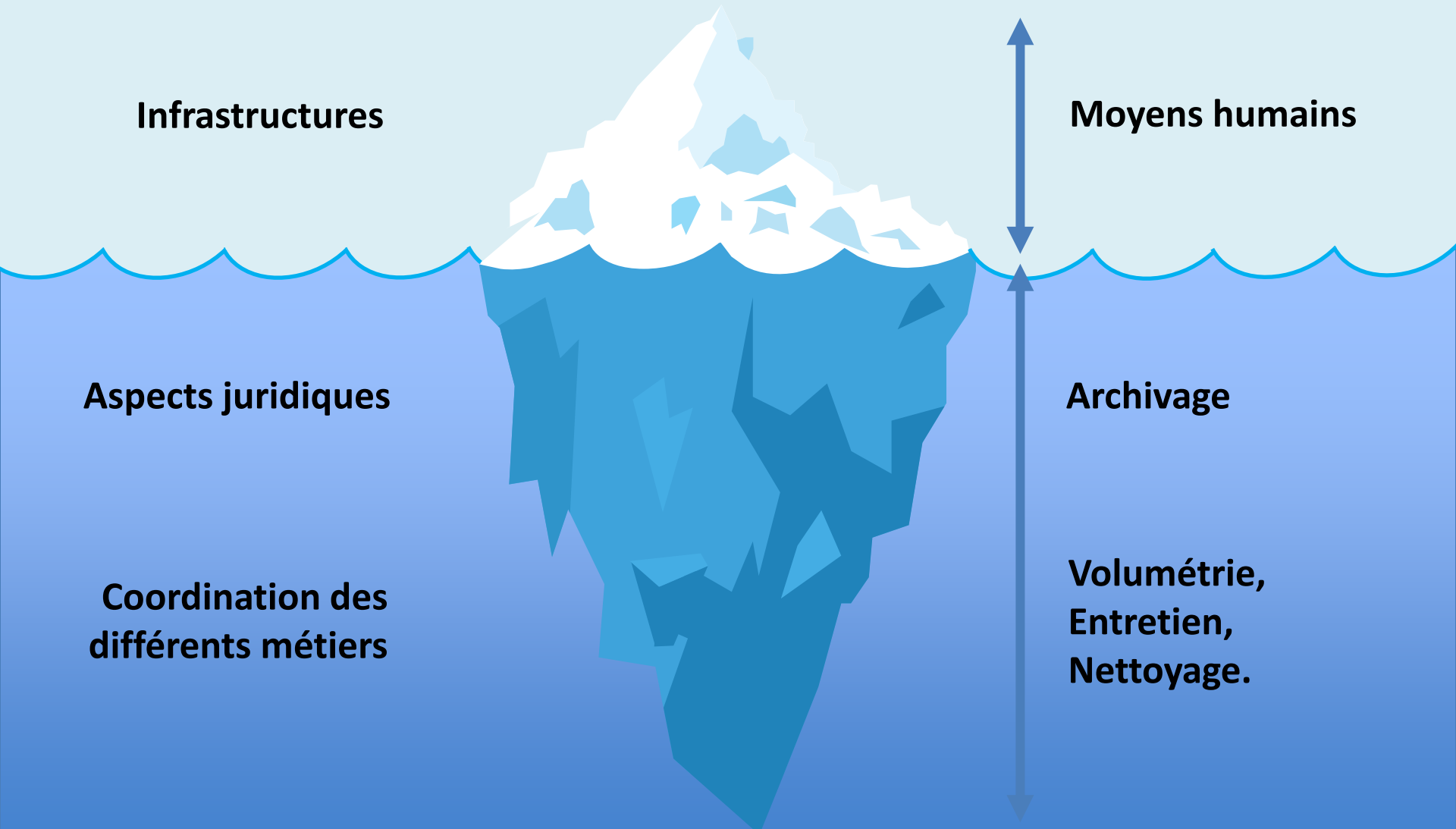


UNIVERSITÉ DE MONTPELLIER

Fredrik Raddum

<http://fredrikraddum.no>

# *Gouvernance ?*

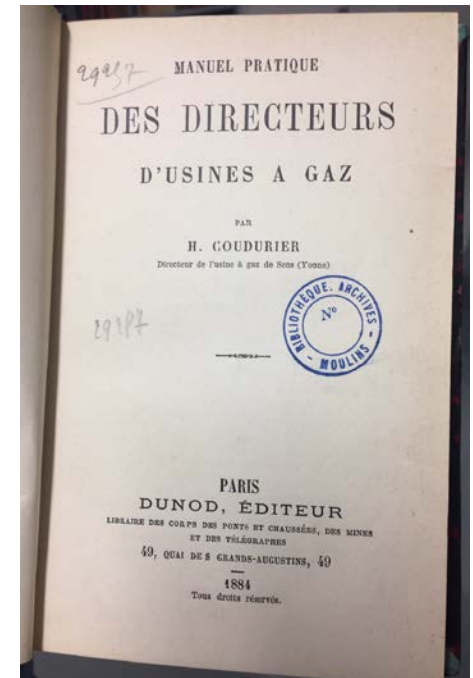
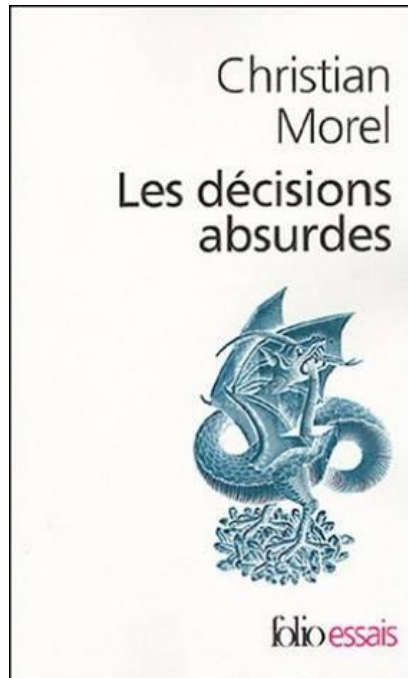
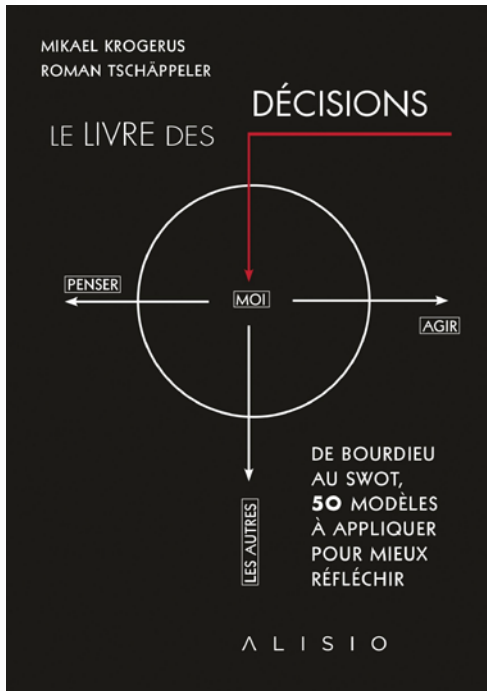


# Gouvernance ?

*Quelle est la bonne échelle d'action ?*

*Quels moyens humains associés ?*

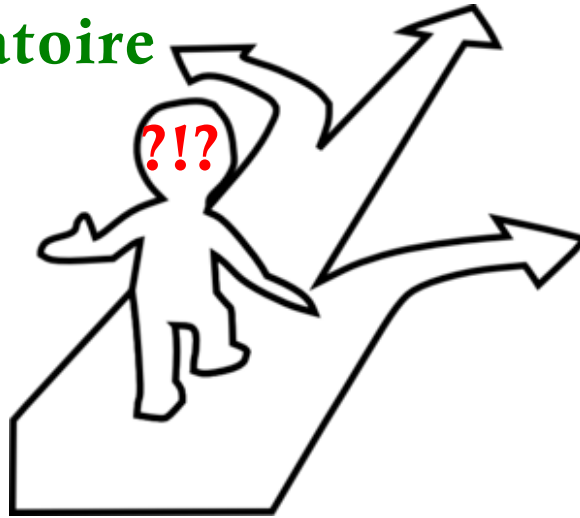
*Quelles infrastructures ?*



# Gouvernance ?

Etablissement

Laboratoire



Structures nationales



# *Gouvernance ?*

## *Différentes échelles*

- *Équipe*
- *Unités / laboratoire*
- *Etablissement*
- *Infrastructures nationales et européennes*

## *Différenciation disciplinaire*

- *Variété disciplinaire du degré de maturité*
- *Variété disciplinaire du type de données*

## *Différenciation en réutilisation*

- *Variété de la mise en forme suivant le public cible de l'ouverture*
- *Cas des partenaires industriels et défense*



# *Définir ce qui doit être conservé*

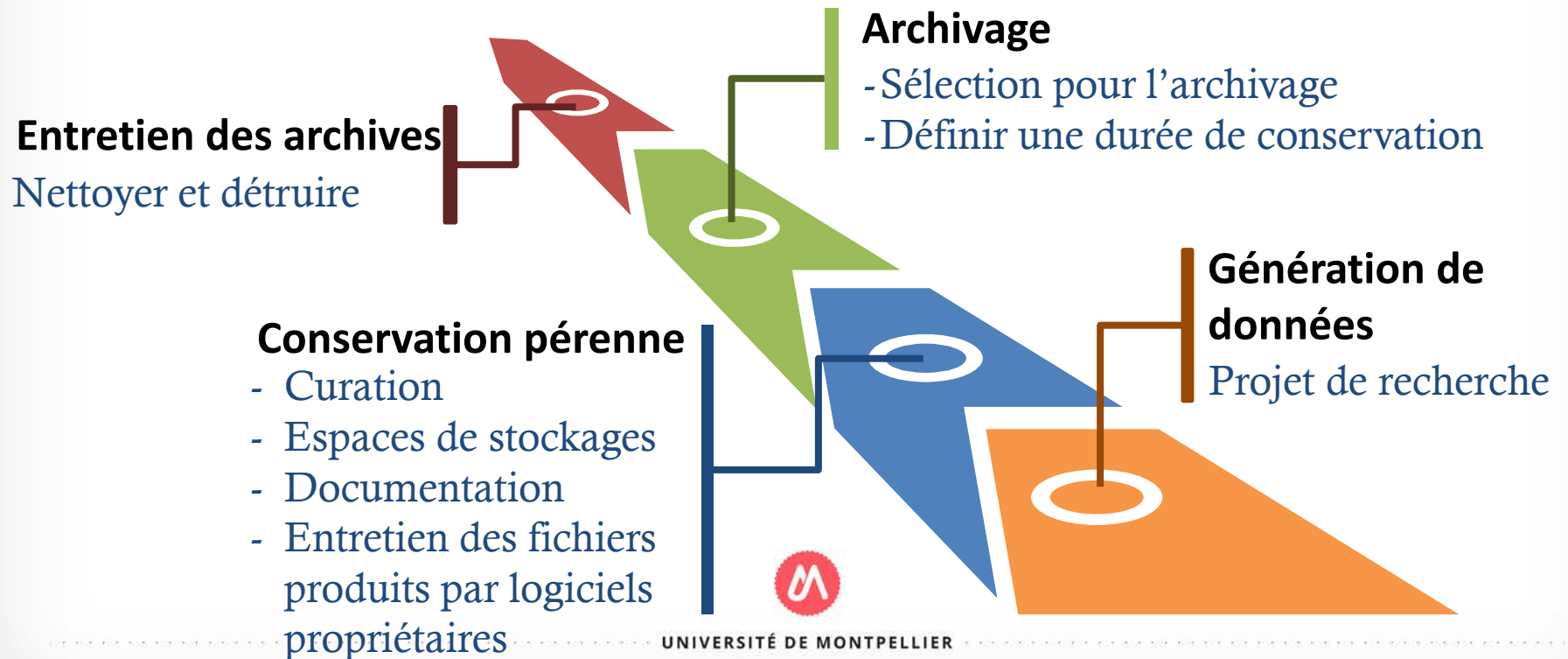


CC BY 2.5 Denmark license [www.digitalbevaring.dk](http://www.digitalbevaring.dk).  
Illustration by Jørgen Stamp.

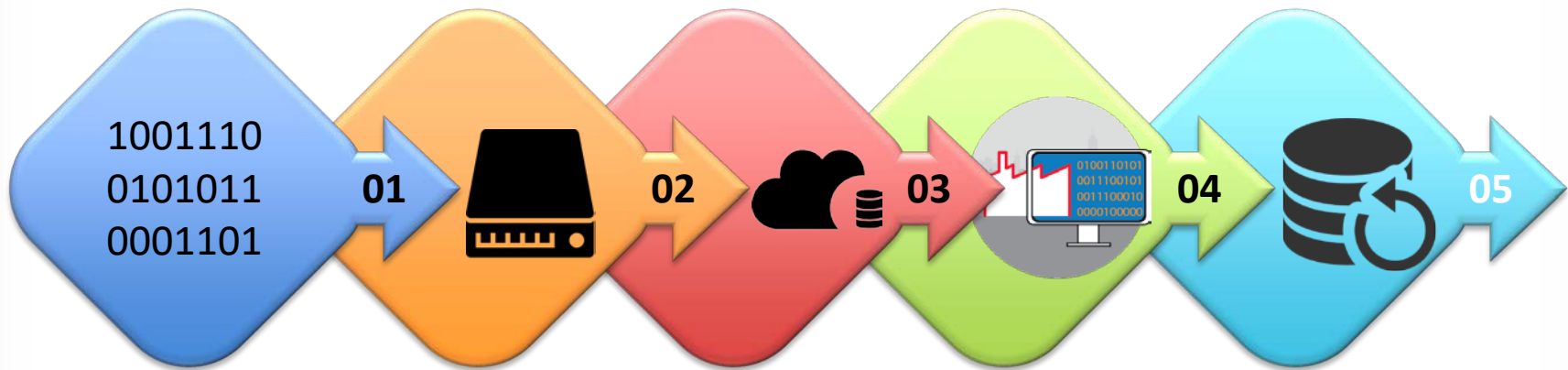


# *Science ouverte implique des choix*

- **Sélection pour l'ouverture** (définir ce qu'il est prioritaire d'ouvrir) ;
- **Conservation avec accès limité pour un temps** ;
- **Sélection pour l'archivage** ;
- **Accès immédiat mais non archivé** ;
- **Durée de conservation ; gestion des archives.**



# Définir ce qui doit être conservé



## Production de données

Plan de gestion de données

## Conservation

- Curation
- Sélection
- Documentation
- **Volume de données produites annuellement ?**

## Ouvrir ?

- Définir ce qui doit être ouvert ou non
- Juridique
- Protection (PI)
- Mise en forme suivant le public destinataire
- Embargo ?

## Archiver

- Sélection
- **Durée de vie ?**

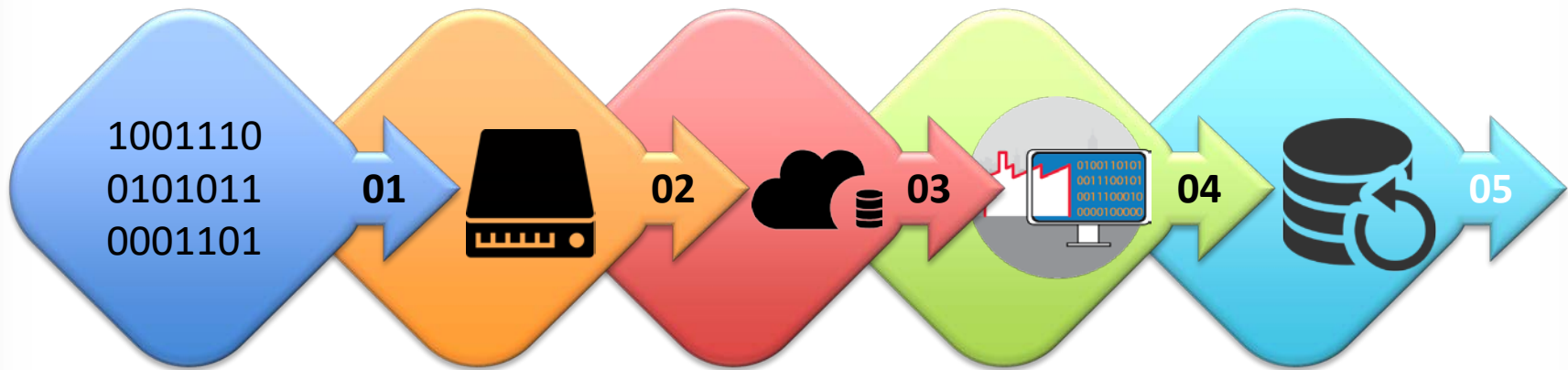
## Entretien

- **Destruction programmée de données**
- Traitement approprié des données conservées (formats propriétaires)

***Réflexion sur la stratégie d'archivage à mener dès le début (à l'élaboration du PGD)***



# Définir ce qui doit être conservé



## Production de données

Plan de gestion de données

## Conservation

- Curation
- Sélection
- Documentation
- **Volume de données produites annuellement ?**

## Ouvrir ?

- Définir ce qui doit être ouvert ou non
- Juridique
- Protection (PI)
- Mise en forme suivant le public destinataire
- Embargo ?

## Archiver

- Sélection
- **Durée de vie ?**

## Entretien

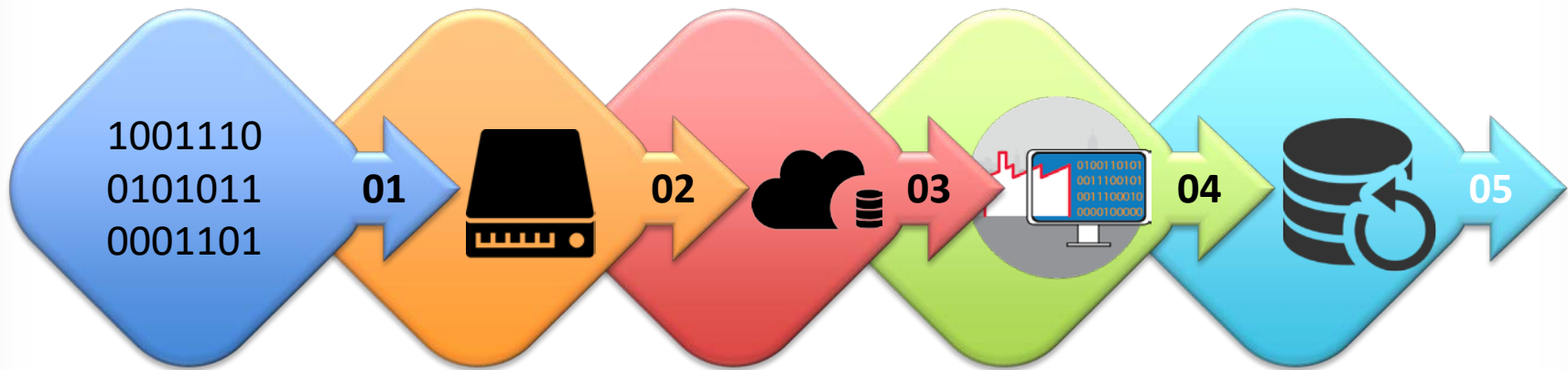
- **Destruction programmée de données**
- Traitement approprié des données conservées (formats propriétaires)

*Prévoir une durée de vie par type de données*

*Autoriser un moment de révision avant destruction*



# Définir ce qui doit être conservé



## Production de données

Plan de gestion de données

## Conservation

- Curation
- Sélection
- Documentation
- **Volume de données produites annuellement ?**

## Ouvrir ?

- Définir ce qui doit être ouvert ou non
- Juridique
- Protection (PI)
- Mise en forme suivant le public destinataire
- Embargo ?

## Archiver

- Sélection
- **Durée de vie ?**

## Entretien

- **Destruction programmée de données**
- Traitement approprié des données conservées (formats propriétaires)

### Cas des données massives :

*conservation d'échantillons (aléatoires ou non), de manière similaire aux données administratives*



# *Définir ce qui doit être conservé*

**Définir des lignes directrices par discipline au niveau national ?**

*Elaborer des guides de la science ouverte*

- ✓ *Motivations*
- ✓ *Principes généraux*
- ✓ *Cadre juridique*

**Adaptés à chaque discipline !**

- ✓ *Que peut-on ouvrir ?*
- ✓ *À qui ? Quand ?*
- ✓ *Pour quelle durée ?*

***Rôle des alliances (Athena, Aviesan, Allistene, ...) ?***



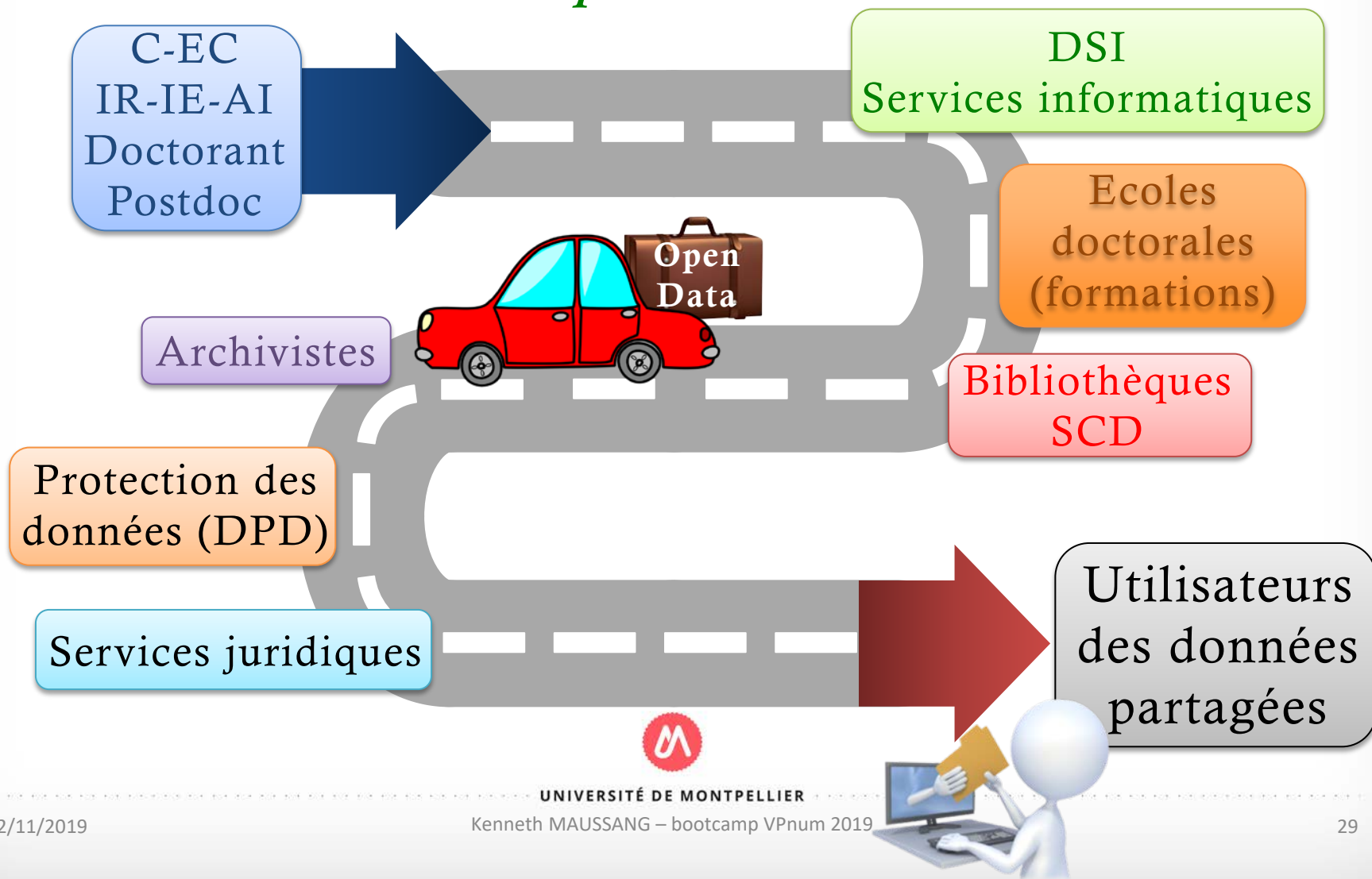
*La science ouverte a besoin de moyens*

**Au confluent de plusieurs métiers aux  
compétences complémentaires**



# *La science ouverte a besoin de moyens*

*Au confluent de plusieurs métiers aux compétences complémentaires*



# *Et le chercheur dans tout ça ???*



ELZO DURT POUR "LE MONDE"

[https://www.lemonde.fr/sciences/article/2014/02/17/le-burn-out-des-labos\\_4368153\\_1650684.html](https://www.lemonde.fr/sciences/article/2014/02/17/le-burn-out-des-labos_4368153_1650684.html)



UNIVERSITÉ DE MONTPELLIER

## ACADEMIA

PRO

YOU CAN WORK WHENEVER  
YOU WANT! EVERY DAY  
IS A SATURDAY!



JORGE CHAM © 2017

CON

YOU WORK ON SATURDAYS.



WWW.PHDCOMICS.COM

# *Et le chercheur dans tout ça ???*

## Current Status

### Peer Review

Peer Review

| Reviewer   | Request Status                        | Review Status                                                      |
|------------|---------------------------------------|--------------------------------------------------------------------|
| Reviewer 1 | Waiting, Due (22 janv. 2019)          |                                                                    |
| Reviewer 2 | Agree (19 janv. 2019)                 | Waiting, Due (02 févr. 2019)<br>Last Reminder Date (03 févr. 2019) |
| Reviewer 3 | Not Area of Expertise (19 janv. 2019) |                                                                    |
| Reviewer 4 | Too Busy (20 janv. 2019)              |                                                                    |
| Reviewer 5 | Too Busy (27 janv. 2019)              |                                                                    |
| Reviewer 6 | Too Busy (28 janv. 2019)              |                                                                    |
| Reviewer 7 | Waiting, Due (31 janv. 2019)          |                                                                    |



# *Reconnaissance dans les carrières*

**La science ouverte ne progressera pas sans reconnaissance du travail correspondant dans les évaluations des chercheurs.**

- ✓ Suivi de formation en matière de gestion et d'ouvertures de données ;
- ✓ Temps consacré à rendre ses données pérennes et partageables ;
- ✓ Efforts d'ouverture effectif des données ;
- ✓ Efforts de prise de connaissance des pratiques d'autres disciplines.





**[kenneth.maussang@umontpellier.fr](mailto:kenneth.maussang@umontpellier.fr)**

UNIVERSITÉ DE MONTPELLIER