



Recherche Data Gouv & Dat'AMU

Data submission

Lionel Spinelli – CIML / CB2M - 24/02/2025

Data sharing: a national and international requirement

- 2nd national plan for open science

‘France is committed to ensuring that the results of scientific research are open to everyone - researchers, businesses and citizens - without hindrance, without delay and without payment.

- ANDOR : First “Assises Nationales des Données de le Recherche”, Marseille, November 2024

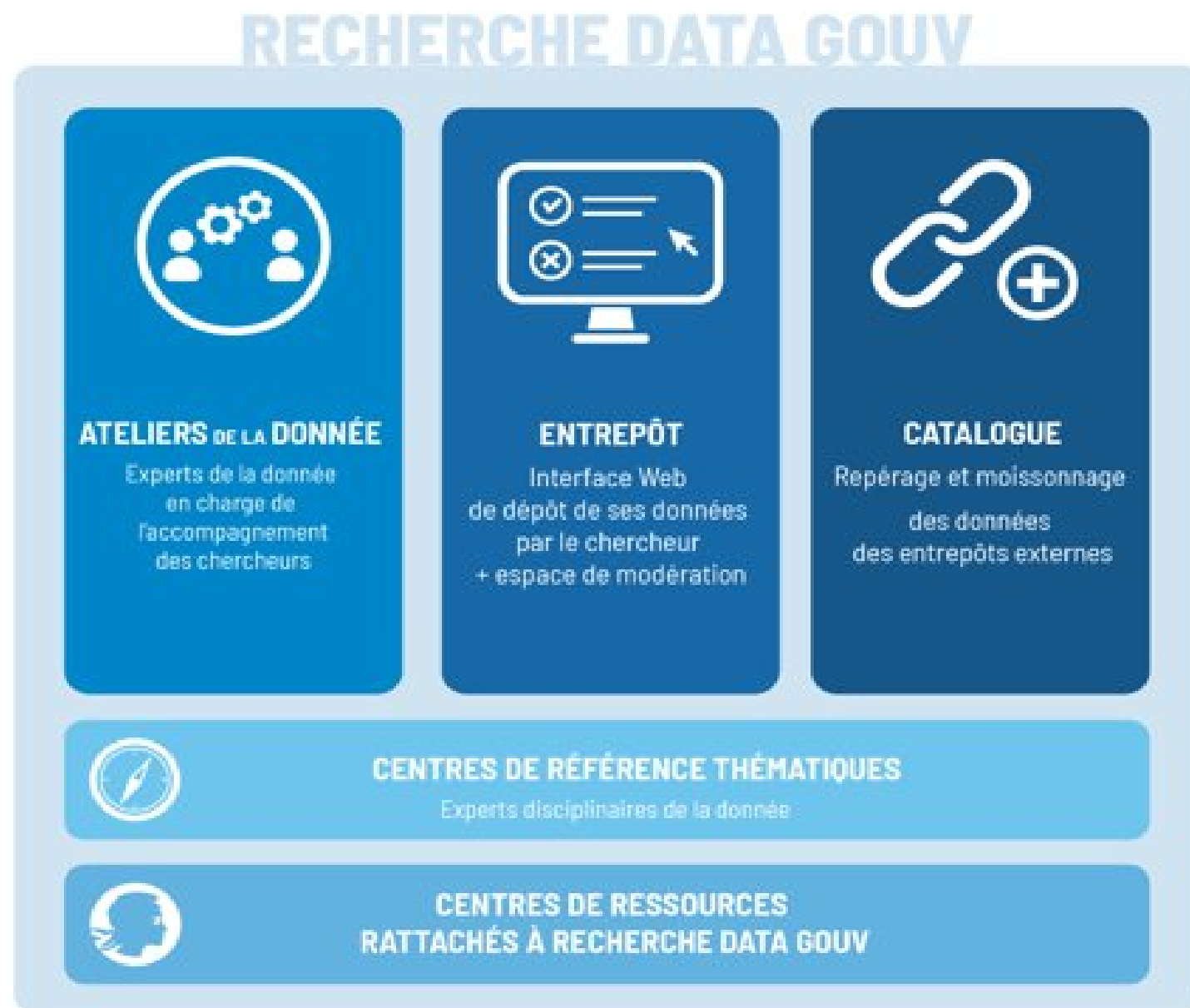
In the presence of the ministry's data and code officer and the directors of several EPSTs (Inserm, CNRS, CEA, INRAE, etc.), it has been confirmed that sharing data :

- *is necessary for more robust and transparent science*
- *is a legal obligation*
- *will be used in the future to evaluate research*

A sovereign solution for depositing, publishing, reporting, discovering, making accessible and reusing research data.

It is composed of :

- A national research data warehouse
- A catalogue of data published on the warehouse itself or on external warehouses.
- A national federation of skills centres:
 - Regional data workshops
 - Thematic reference centres
 - Resource centres



AMU data workshop:

GDsAM -
“Guichet de la donnée
du site d’Aix-Marseille”

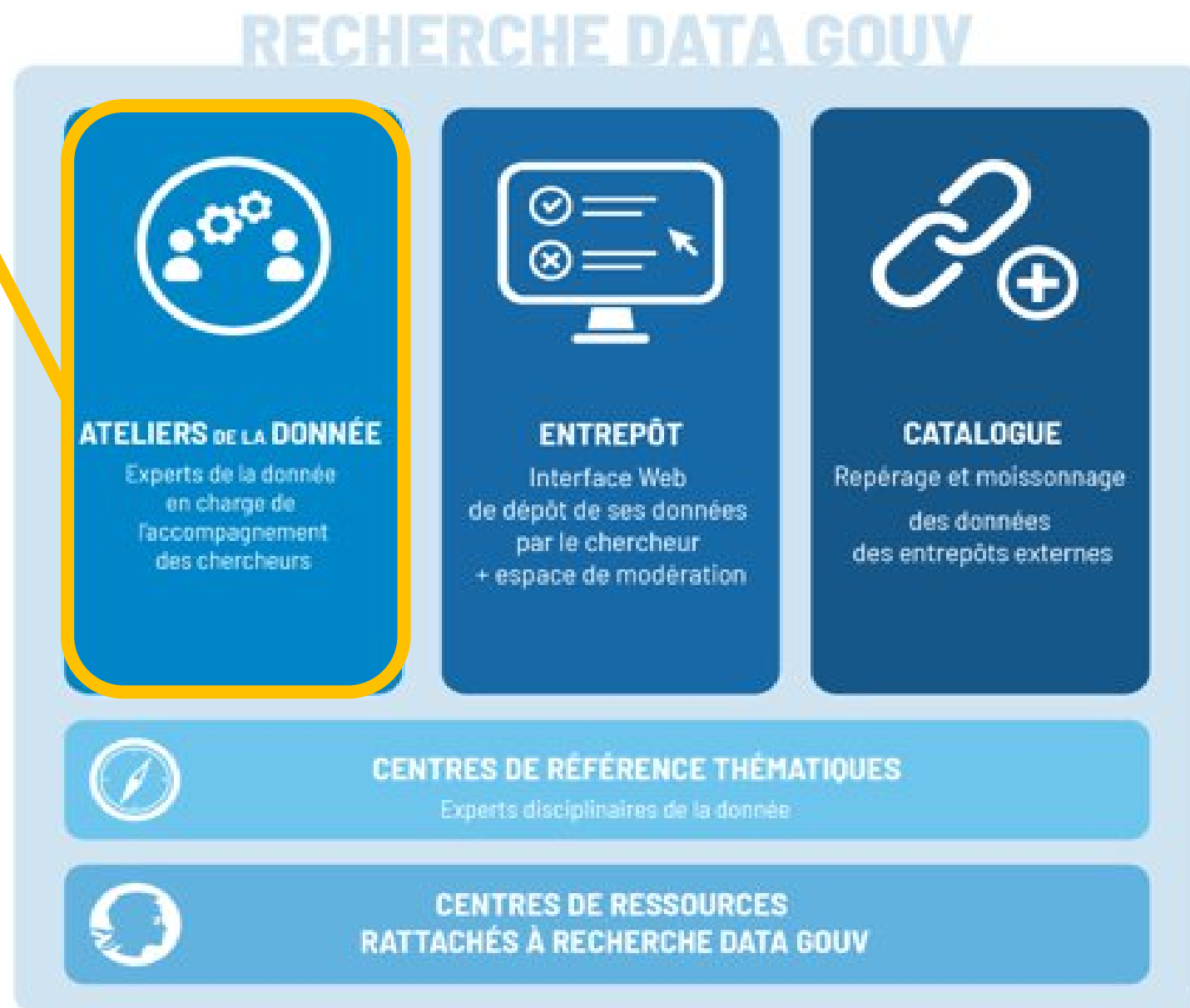


GDsAM is here to help on any aspect of
data management :

<https://gdsam.univ-amu.fr/>

You can use their ticket system to ask all
your questions :

<https://guichet-donnees.univ-amu.fr/>

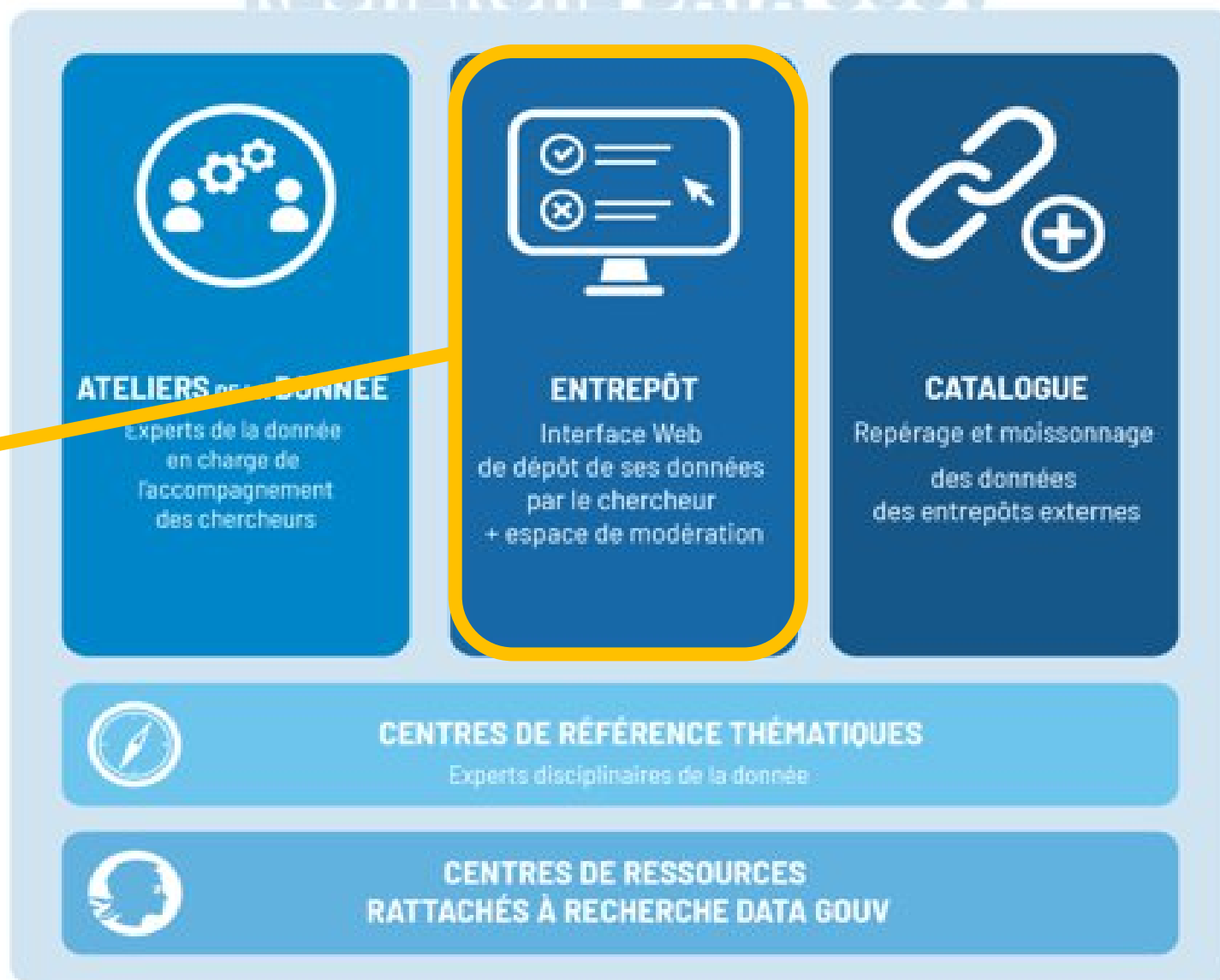




National data warehouse

<https://entrepot.recherche.data.gouv.fr/>

RECHERCHE DATA GOUV





**Thematic reference centre
for bioinformatics:**

IFB
“Institut Français
de Bioinformatique”

<https://www.france-bioinformatique.fr>

RECHERCHE DATA GOUV



ATELIERS DE LA DONNÉE

Experts de la donnée
en charge de
l'accompagnement
des chercheurs



ENTREPÔT

Interface Web
de dépôt de ses données
par le chercheur
+ espace de modération



CATALOGUE

Repérage et moissonnage
des données
des entrepôts externes



CENTRES DE RÉFÉRENCE THÉMATIQUES

Experts disciplinaires de la donnée



CENTRES DE RESSOURCES RATTACHÉS À RECHERCHE DATA GOUV

Why to share ?

- **As a professional** : to meet the requirements of our supervisory bodies
- **As a scientist** : to ensure the transparency and reproducibility of our work
- **As a bioinformatician** : haven't you been frustrated by not being able to have the data from a publication of interest? So why not do it yourself?

Why to share ?

- **As a professional** : to meet the requirements of our supervisory bodies
- **As a scientist** : to ensure the transparency and reproducibility of our work
- **As a bioinformatician** : haven't you been frustrated by not being able to have the data from a publication of interest? So why not do it yourself?

What to share ?

Raw data	: sequencer output, microscopy images, etc
Pre-processed data	: raw data prepared before scientific analysis
Resources	: frozen external data
Analysis code	: scripts/code in programming language
Software environments	: all libraries and software useful for the analysis
Complete results	: all important files resulting from the analysis

Where to share ?

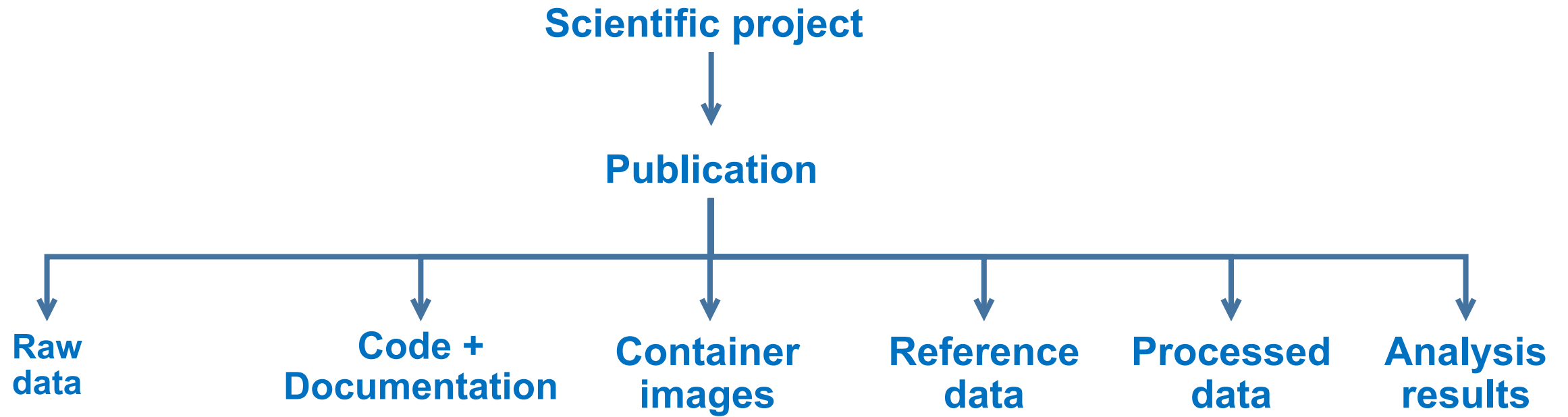
- Sharing must be performed in trusted public related sites selected by specific criteria:

Sélectionner un entrepôt thématique de confiance pour le dépôt de données : méthodologie et analyse de l'offre existante

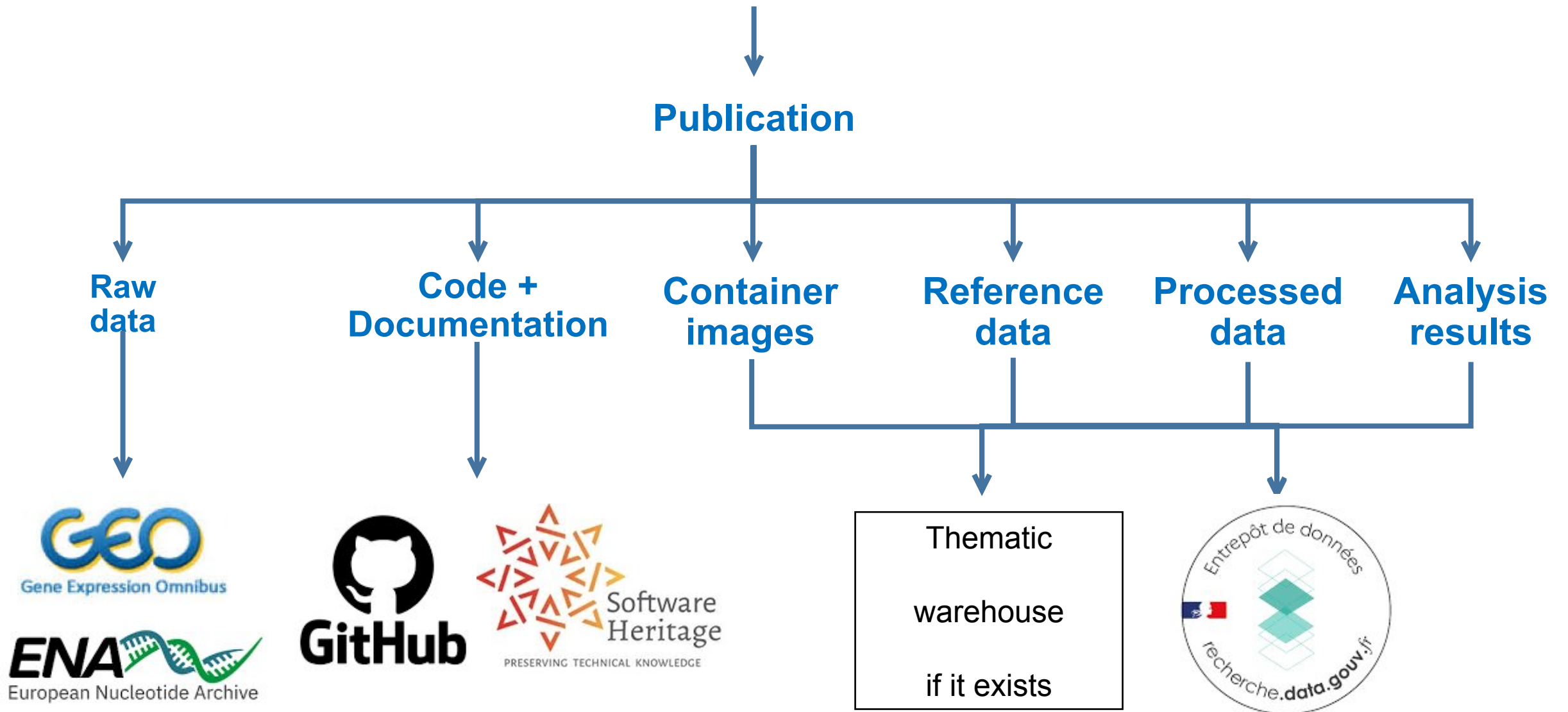
<https://www.ouvrirlascience.fr/selectionner-un-entrepot-thematique-de-confiance-pour-le-depot-de-donnees-methodologie-et-analyse-de-loffre-existante/>

<https://hal-lara.archives-ouvertes.fr/OUVRIR-LA-SCIENCE/hal-04534321v1>

- Sharing must be performed in priority on trusted thematic depot.
- **Always ask our CRT to be sure : the IFB**
 - <https://www.france-bioinformatique.fr/guichet-conseil-orientation>



Example : Genomics data





How to publish in Recherche Data Gouv warehouse ?

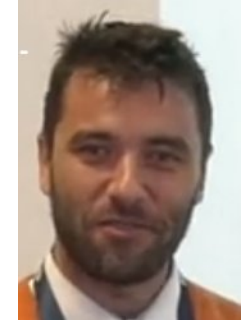
1 . You are not alone



Francine FILOCHE

AMU : Chargée de mission Données de recherche
Guichet unique de la donnée
Service commun de documentation
francine.FILOCHE@univ-amu.fr

- Provide training and start-up assistance
- Help on best practices
- Help on curation and meta-data insertion
- Listen to needs and problems & raise concerns
- Inform on RDG enhancements
- Help on dissemination



Damien SANS

INRAE: Super-administrateur RDG
Direction de la Science Ouverte
damien.sans@inrae.fr

GDsAM is here to help on any aspect of data management :

<https://gdsam.univ-amu.fr/>

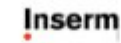
2 . Check for the institutional dataverses of your institute



CNRS Research Data



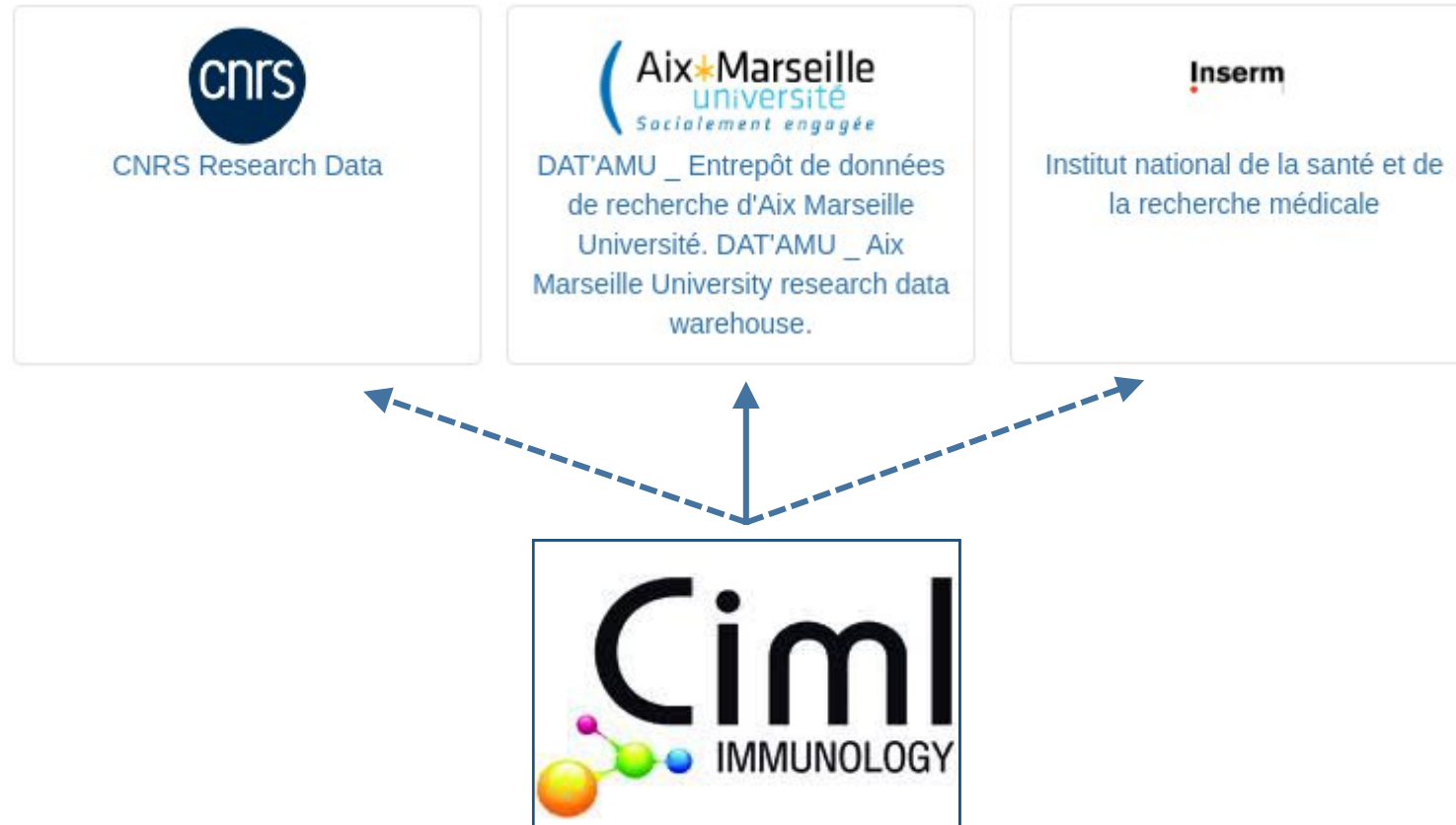
DAT'AMU _ Entrepôt de données
de recherche d'Aix Marseille
Université. DAT'AMU _ Aix
Marseille University research data
warehouse.



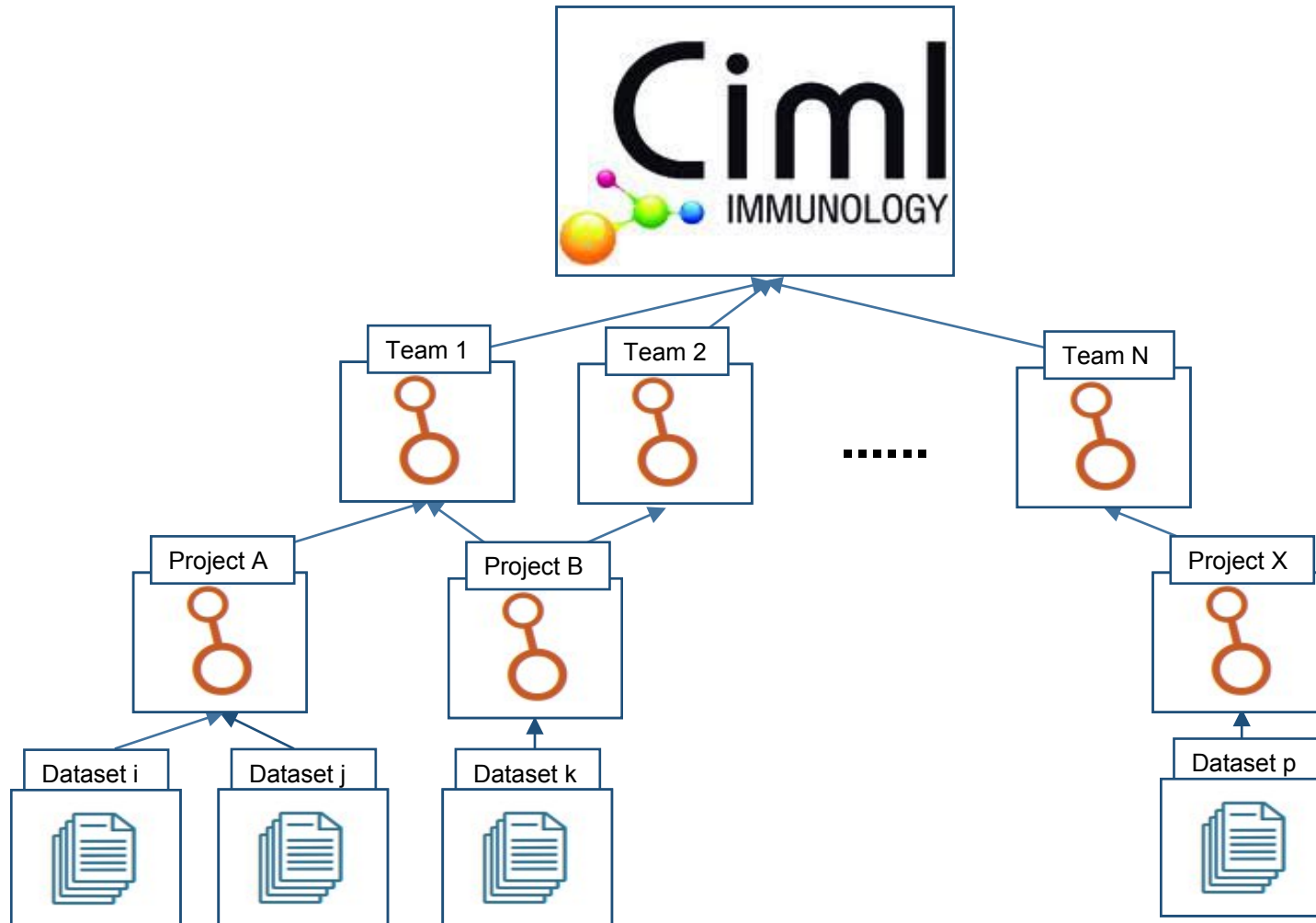
Institut national de la santé et de
la recherche médicale

<https://entrepot.recherche.data.gouv.fr>

3 . Ask for your institute dataverse creation and link it to the institutional dataverse

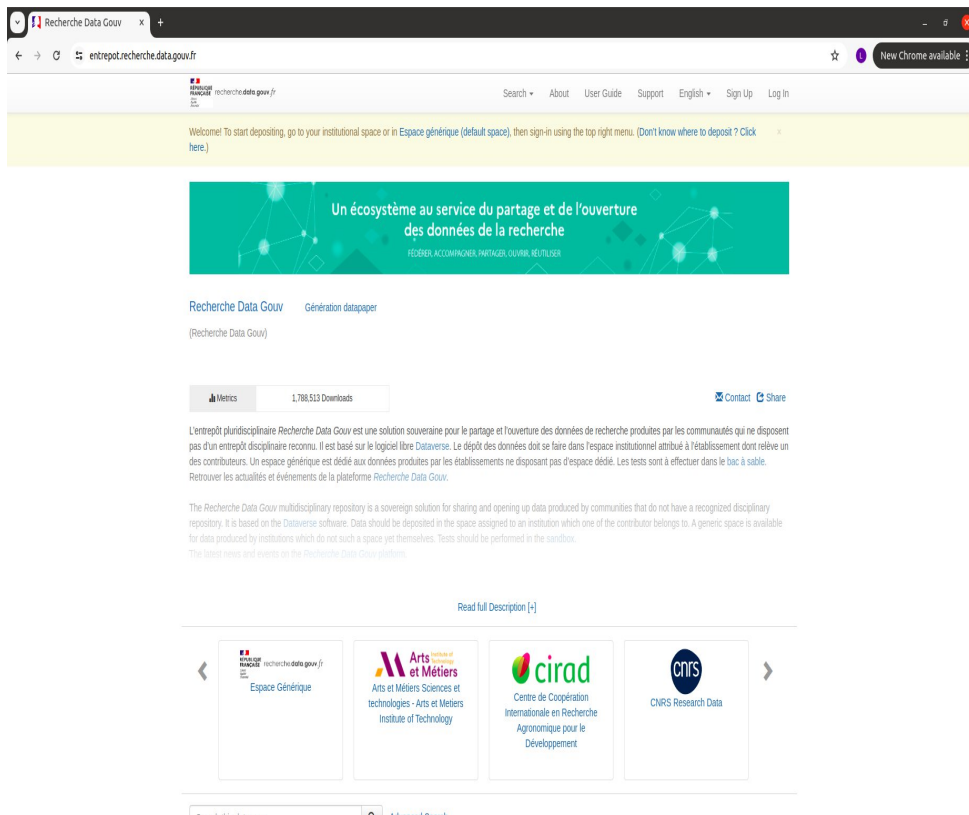


4 . Decide the dataverses/datasets structure in your institute dataverse

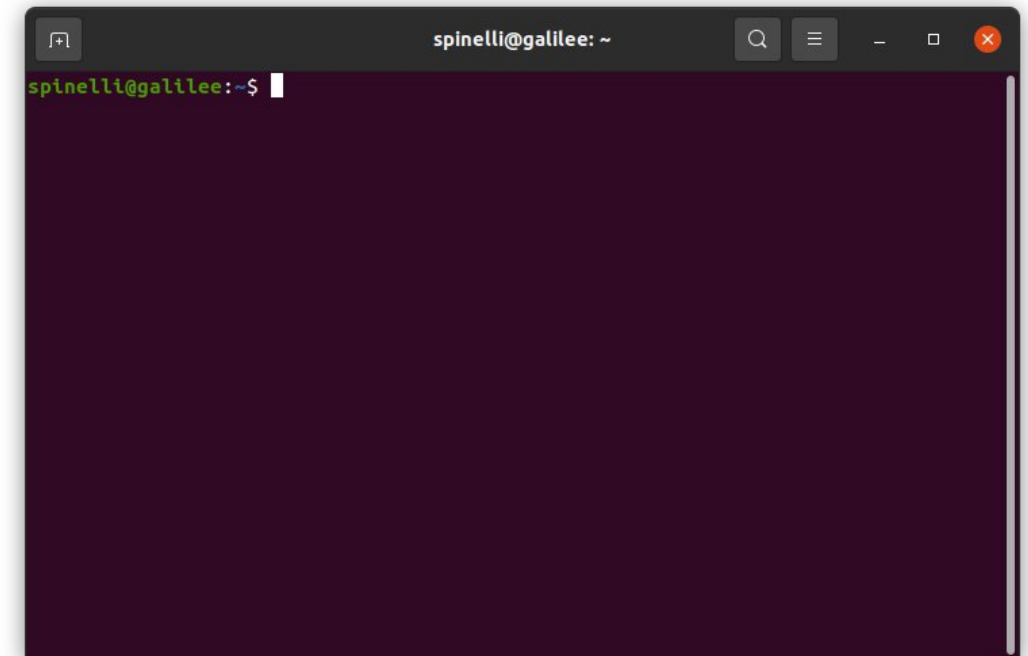


5 . Upload your data to the dataset

With the web interface



With the CLI tool : DVUploader



6 . Fill the meta-data of the dataset

entrepot.recherche.data.gouv.fr/dataset.xhtml?persistentId=doi:10.57745/4EFY2Z&version=DRAFT

RECHERCHE DATA GOUV

Search About User Guide Support English Lionel SPINELLI

> BMMLab_Prostate-CD45Pos-PTENKO >

Edit Dataset Metadata – Add more metadata about this dataset to help others easily find it.

Host Dataverse BMMLab_Prostate-CD45Pos-PTENKO

Save Changes Cancel

Metadata

Please enter metadata in the **English** language, or change the metadata language choice in the parent dataverse before proceeding.

*Asterisks indicate required fields

Common Metadata

Title * VH00228_144_AAAGNCJHV_181_AAAGNF7HV_AllMerges

Subtitle

Alternative Title

Link to data Enter full URL, starting with https://

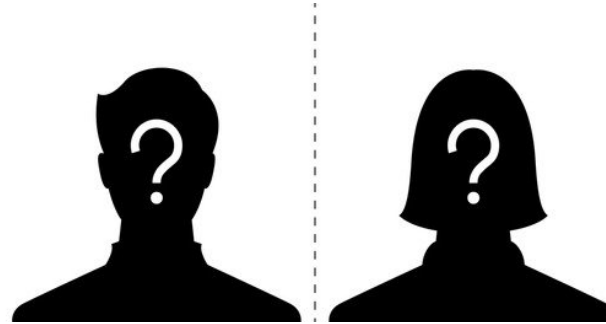
Other Identifier

Agency **Identifier** +

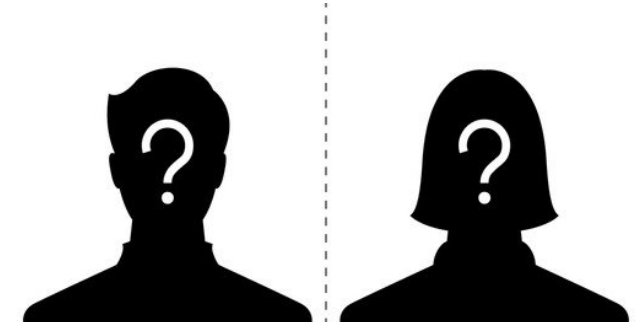
7 . Curate the dataset with the referee



Francine FILOCHE
AMU : Chargée de mission Données de recherche
Guichet unique de la donnée
Service commun de documentation
francine.FILOCHE@univ-amu.fr



???
Inserm



???
CNRS

8 . Publish your data

On 29th of April,

**curathon day organized by Recherche Data Gouv
on the Luminy campus: whole day with lunch offered**

**An excellent moment to prepare and submit your first
datasets**

**Please contact Francine Filoche for more details and
prepare your participation**

francine.FILOCHE@univ-amu.fr

Let's look at an example