



FRANCE-BIOIMAGING



EURO-BIOIMAGING

## France BioImaging (FBI)

# A National Infrastructure for quantitative biological imaging

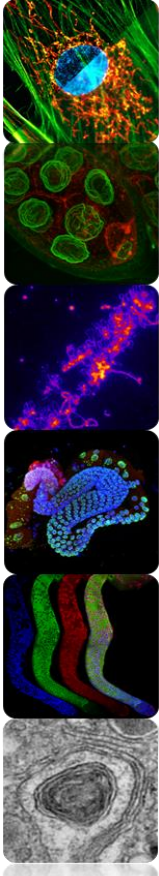
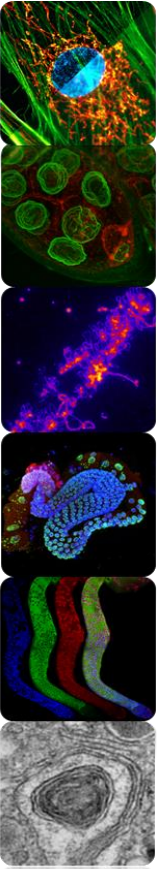
<https://france-bioimaging.org/>

*Edouard Bertrand*

*Scientific Director of the FBI-RI*

*Perrine Paul-Gilloteaux*

*Mission Officer for data management*

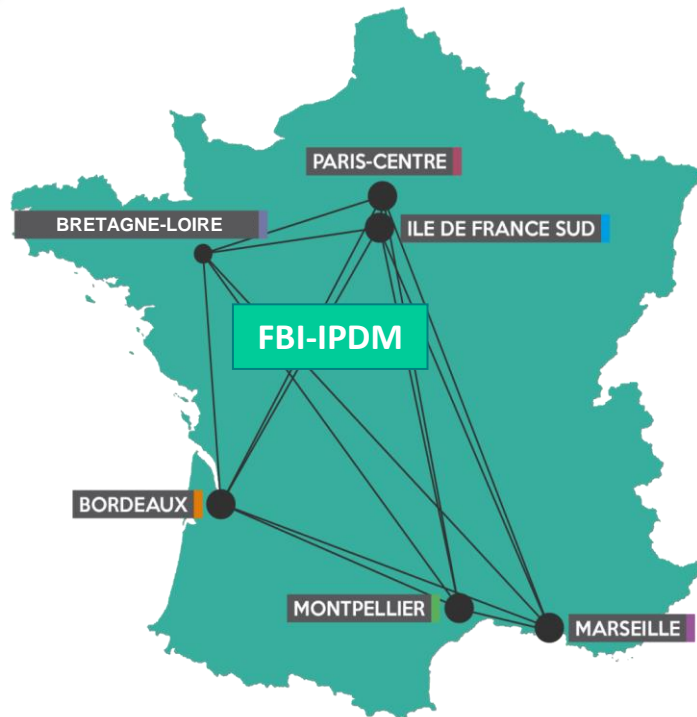


Funded by the Horizon 2020  
Framework Programme of  
the European Union



# FRANCE-BIOIMAGING

## FBI IN A NUTSHELL



- 7 Nodes with complementary expertises
- 18 Imaging Core Facilities
- **25 associated R&D teams**
- ~ 300 staff (engineers & researchers)
- >350 set-ups (45 Super-Res co-financed PIA1)
- 21-25 M€ of budget /year (full cost)

## BORDEAUX NODE

SUPER-RESOLUTION

3D TOMOGRAPHY, CORRELATIVE MICROSCOPY

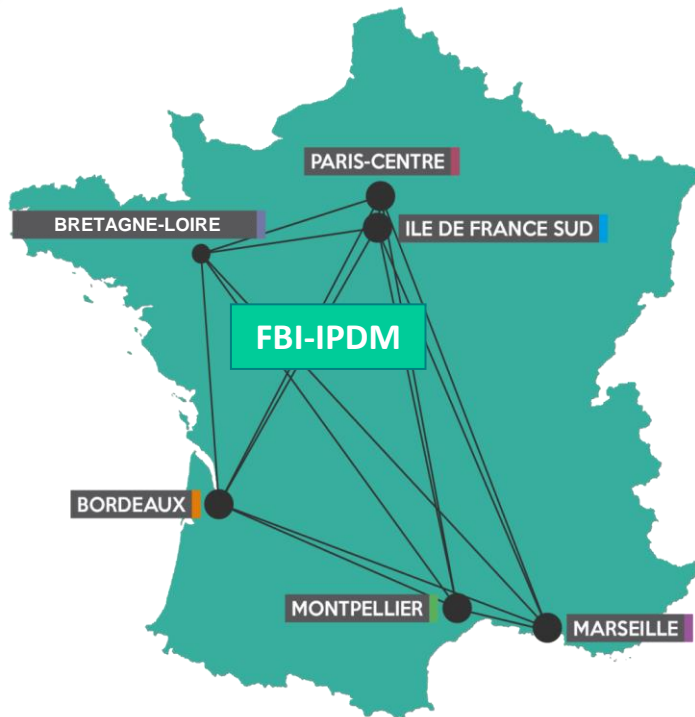
NEUROSCIENCES & PLANT SCIENCE

## MONTPELLIER NODE

SINGLE MOLECULE IMAGING

HIGH-THROUGHPUT MICROSCOPIES

GENOME ORGANISATION AND GENE EXPRESSION



- 7 Nodes with complementary expertises
- 18 Imaging Core Facilities
- **25 associated R&D teams**
- ~ 300 staff (engineers & researchers)
- >350 set-ups (45 Super-Res co-financed PIA1)
- 21-25 M€ of budget /year (full cost)

## Innovation – Access – Training

Bring together  
research and  
innovation

Provide access to  
advanced techs  
and methods

Sustain training,  
education and  
dissemination



NUMBER OF  
USERS (2018)

2020 : >5000

2012 : 1768

---



START UP  
CREATION/SUPPORT

20120: 12

2012: 1

---



# 1

## INNOVATION

- **IMAGE ENTIRE ORGANISMS:**
  - Light Sheet microscopies
  
- **IMAGE AT HIGH RESOLUTION:**
  - Correlative Light / Electron / Atomic
  - Multiscale / Multimodal imaging
  
- **IMAGE AT HIGH-THROUGHPUT :**
  - High-density multiplexing (>10 000 labels);
  
- **IMAGE DATA :**
  - IPDM Node



# 1

## INNOVATION

- **IMAGE ENTIRE ORGANISMS:**
  - Light Sheet microscopies
- **IMAGE AT HIGH RESOLUTION:**
  - Correlative Light / Electron / Atomic
  - Multiscale / Multimodal imaging
- **IMAGE AT HIGH-THROUGHPUT :**
  - High-density multiplexing (>10 000 labels);
- **IMAGE DATA :**
  - IPDM Node

## Image DATA

### Image Data (ESR / EQUIPEX +)

- ❑ **Image analysis :**
  - Machine learning, Deep learning
- ❑ **Image Visualization :**
  - Virtual Reality /Augmented Reality
- ❑ **Data Management :**
  - High Volume (**0.8 Po/year/node ... x7 !**)
  - Data heterogeneity
  - On the fly processing

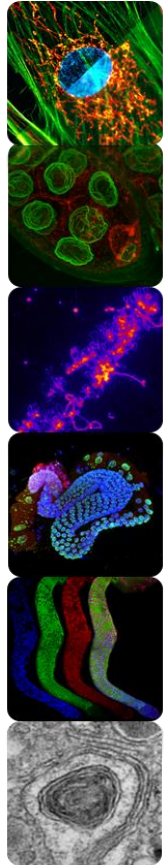


FRANCE-BIOIMAGING

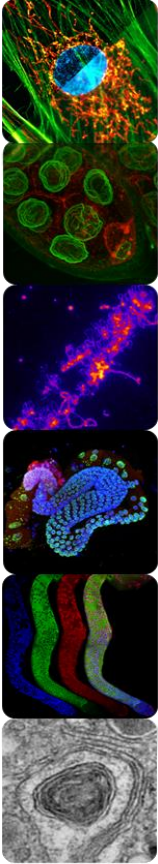


EURO-BIOIMAGING

GLOBAL  
BIOIMAGING  
growing collaboration



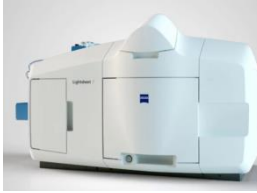
# Challenges et plan data de l'infrastructure France Bio Imaging



SCIENCE OUVERTE 2020, où en sommes-nous ?



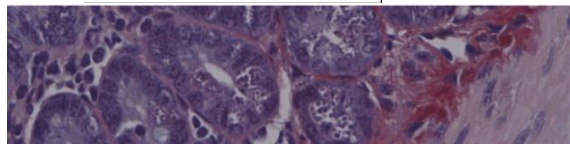
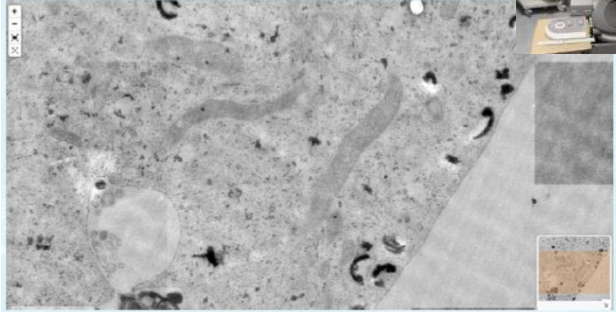
Funded by the Horizon 2020  
Framework Programme of  
the European Union



1040x1288x20x300x2  
colorsx2 bytes =3.214 GB

WF 384x384x50x2  
colorsx2 bytes  
=29.5 MB

10000x10000x1 byte =100 MB



50kx50kx3 bytes=  
7.5 G

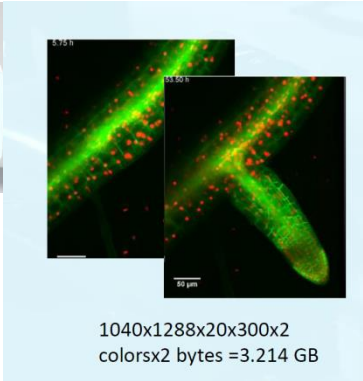
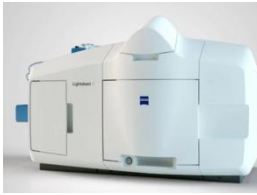


1024x1024x4x96x  
2 byte=0.8G une  
plaque

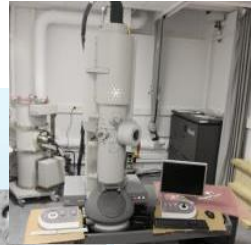
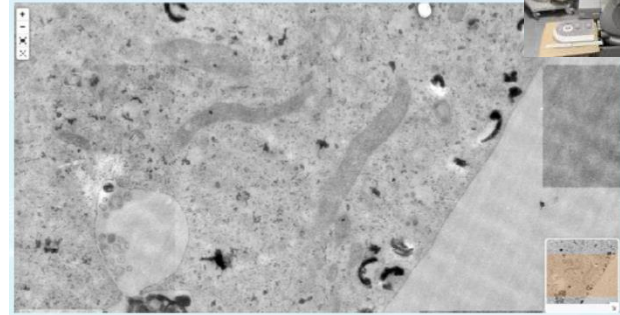


Palm Data  
197\*188\*10000\*2  
bytes= 740 MB





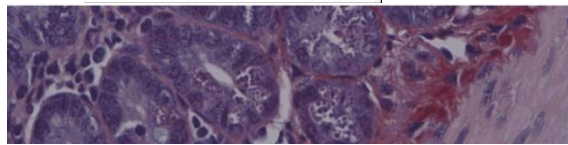
10000x10000x1 byte =100 MB



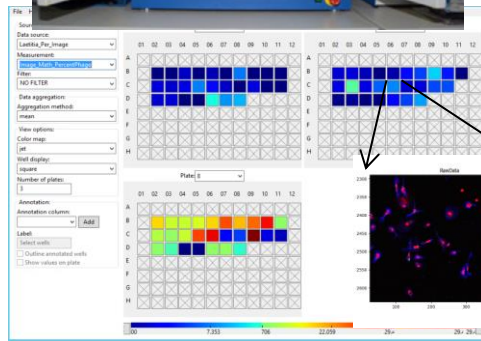
+Données reconstruites, correction ombrage...

+Déconvolution /débruitage

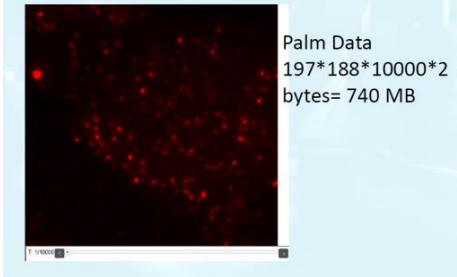
+ Mosaïquage + sections sériées



50kx50kx3 bytes=  
7.5 G



1024x1024x4x96x  
2 byte=0.8G une  
plaque



+Données de localisation et données reconstruites

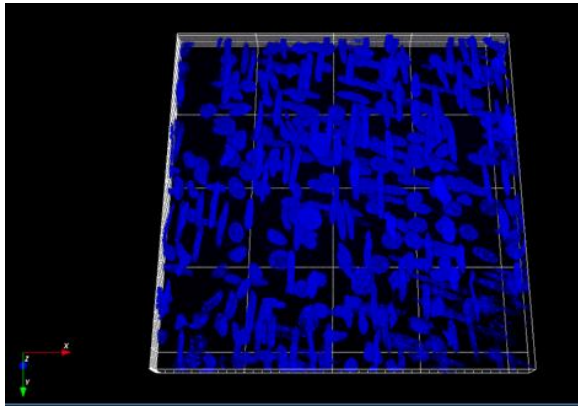
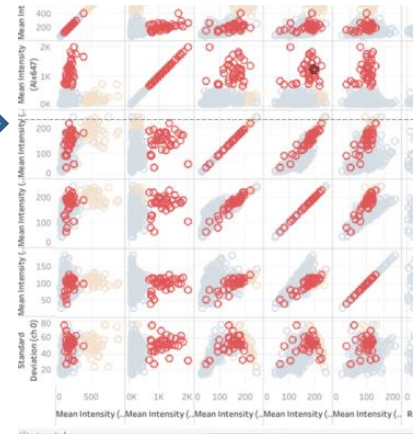
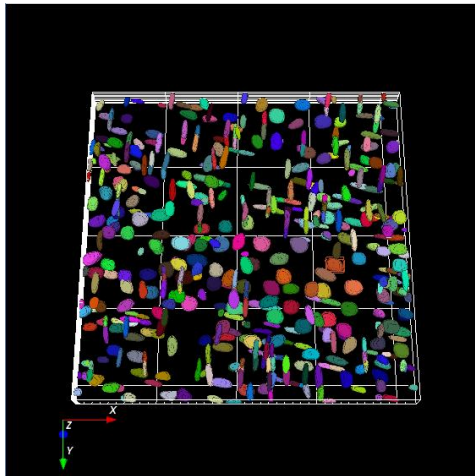
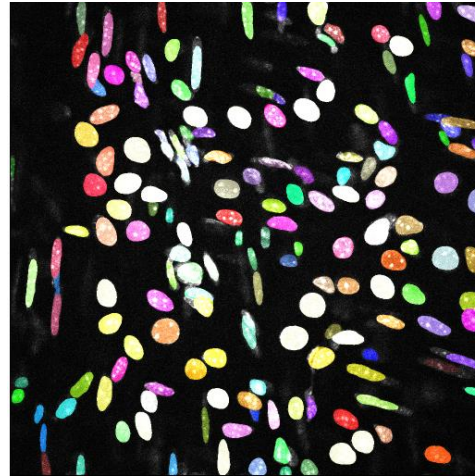
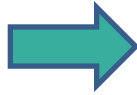


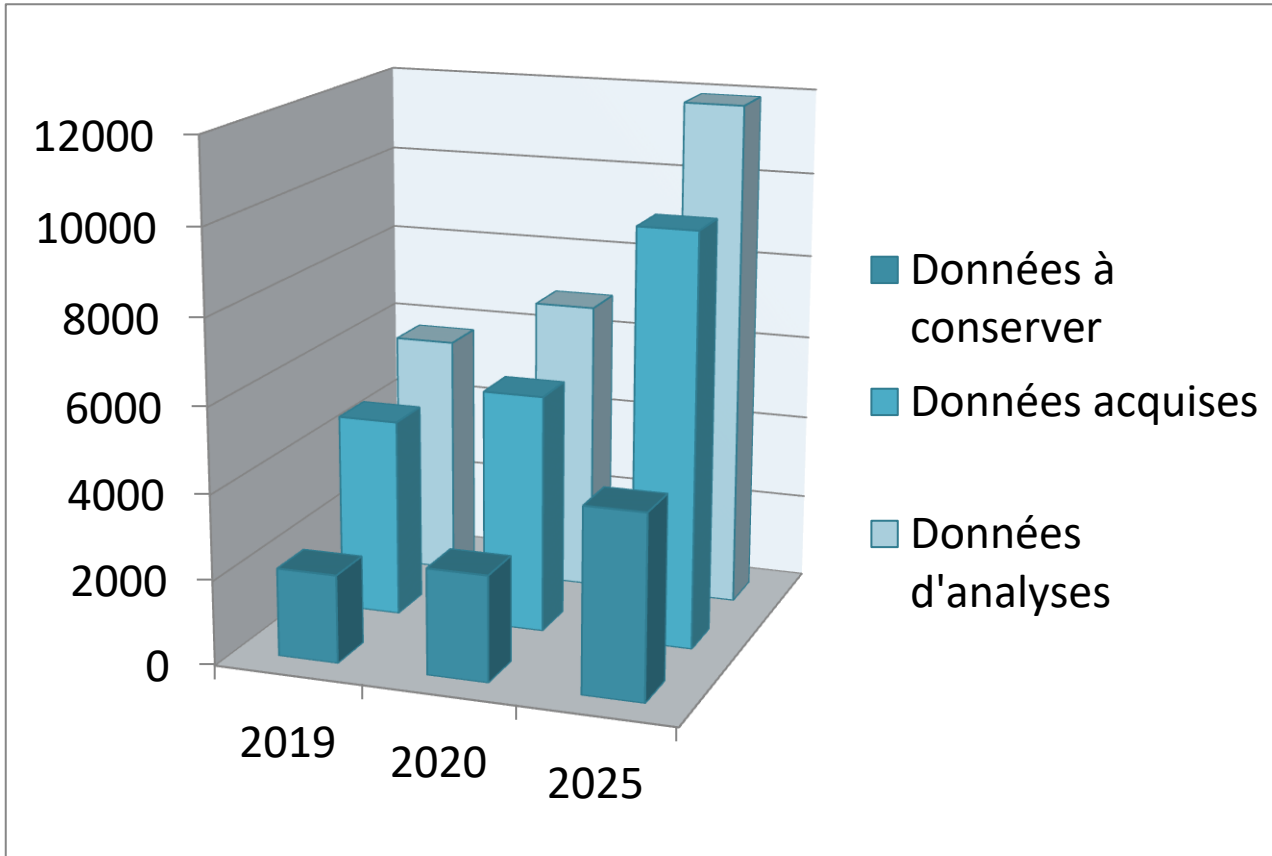
Image acquise  
reconstruite



Traitement et analyse des images:  
annotations spatiales /temporelles et  
mesures par objet d'intérêt



Production de data sur les nœuds France Bioimaging:  
Projection à 5 ans en TeraBytes



### A conserver:

- Data liées à des Publications
- Data réutilisables
- Data couteuses à reproduire
- Data projets de recherche publiques



### Quelles Data?

- **Image acquise (brute) et image reconstruite**
- **Traitement et analyse des images: annotations spatiales /temporelles et mesures par objet d'intérêt**
- **Résultats d'analyses**



### Quelles Data?

- **Image acquise (brute) et image reconstruite**
- **Traitement et analyse des images: annotations spatiales /temporelles et mesures par objet d'intérêt**
- **Résultats d'analyses**

SI réutilisables



## Quelles Data?

- **Image acquise** (brute) et **image reconstruite**
- **Traitement et analyse des images: annotations spatiales** /temporelles et mesures par objet d'intérêt
- **Résultats d'analyses**

SI réutilisables

## Quelle réutilisation?

- Développeurs de nouveaux systèmes d'acquisition ( données de référence)
- **Biologistes:** données de référence ou ré analyse (peu d'exemple aujourd'hui)
- **Les développeurs d'algorithmes:** Partage de data et d'annotations pour tests de nouveaux algorithmes ou entraînement de deep learning pour segmentation ou classification objets ou image.
- Biologistes: quels algo sur quelles données



- Hétérogénéité, volumétrie et visualisation des data à partager
- La Ligue 1 et la science ouverte



Quel rapport entre la ligue 1 et la science ouverte?

-> la ligue 1:

j'en ai entendu parlé

Je ne sais pas quand où comment ca se passe

il y a surement un intérêt mais je n'ai pas encore compris lequel

je n'y passera pas 2h.





Quel rapport entre la ligue 1 et la science ouverte?

-> la ligue 1:

j'en ai entendu parlé

Je ne sais pas quand où comment ca se passe

il y a surement un intérêt mais je n'ai pas encore compris lequel

je n'y passera pas 2h.

-> la science ouverte: la majorité de nos collègues:

en ont entendu parlé

ne savent pas quand où comment ca se passe

il y a surement un intérêt mais il n'ont pas encore compris lequel

ils n'y passeront pas 2h



Quel rapport entre la ligue 1 et la science ouverte?

-> la ligue 1:

j'en ai entendu parlé

Je ne sais pas quand ou comment ca se passe

il y a surement un intérêt mais je n'ai pas encore compris lequel  
je n'y passera pas 2h.

-> la science ouverte: la majorité de nos collègues:

en ont entendu parlé

ne savent pas quand ou comment ca se passe

il y a surement un intérêt mais il n'ont pas encore compris lequel  
ils n'y passeront pas 2h

**Besoin d'acculturation**



### ➤ **Rendre utile et intéressant:**

- premières source de **réutilisation = deep learning et analyse d'image** en général



### ➤ **Rendre utile et intéressant:**

- premières source de **réutilisation** = **deep learning et analyse d'image** en général

### ➤ **Simplifier et faciliter:**

- Fournir des **outils clefs en main et des recettes en lien avec IFB et mésocentres,**
- Limiter **l'effort au minimum utile** e.g metadata minimum ;
- Permettre la **réutilisation du PGD** comme un outil pour faciliter la prise en charge des data d'imagerie (**MaDMP IFB/Opidor**)



### ➤ **Rendre utile et intéressant:**

- premières source de **réutilisation** = **deep learning et analyse d'image** en général

### ➤ **Simplifier et faciliter:**

- Fournir des **outils clefs en main et des recettes en lien avec IFB et mésocentres**,

- Limiter **l'effort au minimum utile** e.g metadata minimum ;

- Permettre la **réutilisation du PGD** comme un outil pour faciliter la prise en charge des data d'imagerie (**MaDMP IFB/Opidor**)

### ➤ **Former**

- créer du contenu **Passeport formation utilisateurs de notre infra et « train the trainers »**,

- contenu à publier sur **DORANUM**



### ➤ **Rendre utile et intéressant:**

- premières source de **réutilisation** = **deep learning et analyse d'image** en général

### ➤ **Simplifier et faciliter:**

- Fournir des **outils clefs en main et des recettes en lien avec IFB et mésocentres**,
- Limiter **l'effort au minimum utile** e.g metadata minimum ;
- Permettre la **réutilisation du PGD** comme un outil pour faciliter la prise en charge des data d'imagerie (**MaDMP IFB/Opidor**)

### ➤ **Former**

- créer du contenu **Passeport formation utilisateurs** de notre infra et « **train the trainers** »,
- contenu à publier sur **DORANUM**

### ➤ **Communiquer /sensibiliser :**

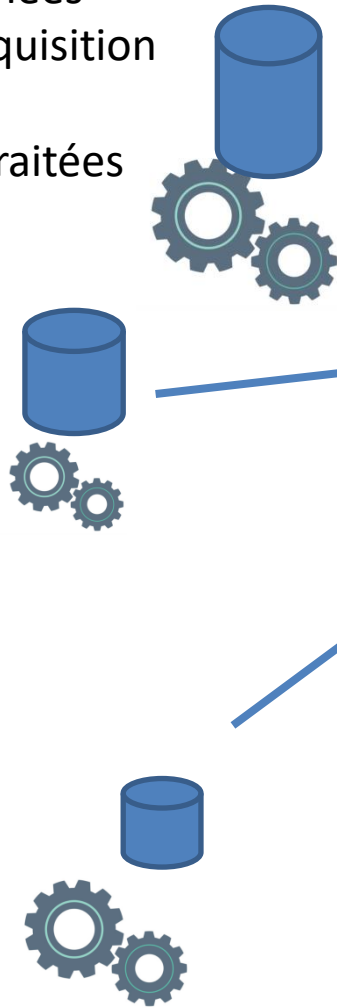
- Impliquer et faire remplir le **PGD FBI** pour chaque plateforme/nœud;
- **Création d'indicateurs** dans nos auto évaluations sur les data et analyses publiés



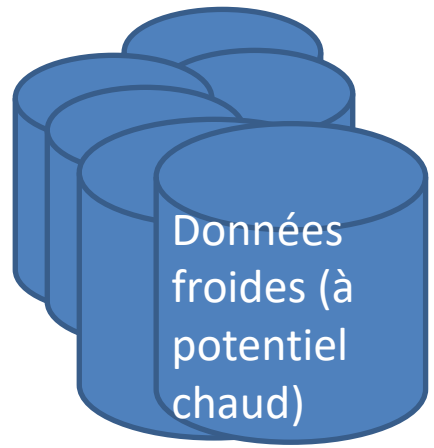
FRANCE-BIOIMAGING

Notre plan de gestion de données

Données  
d'acquisition  
et  
prétraitées



Publication



Données  
froides (à  
potentiel  
chaud)

**ENTREPOTS  
de DONNEES**  
Data publiée  
DOI  
**FAIR**



Mésocentres  
labélisées

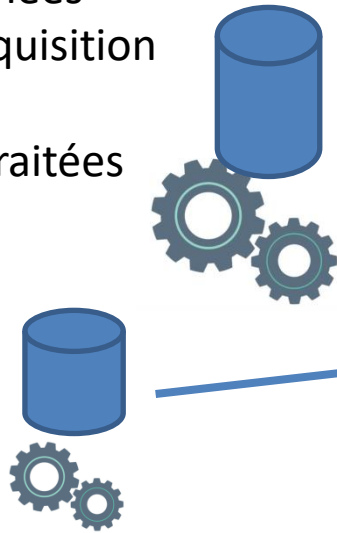




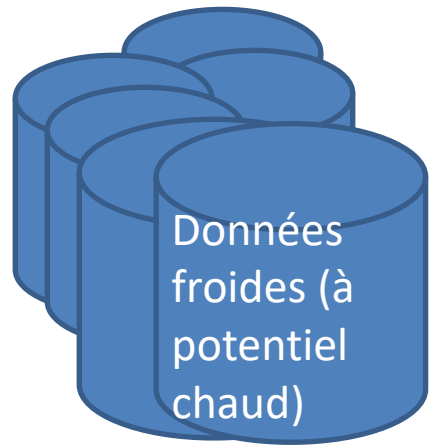
# FRANCE-BIOIMAGING

## Notre plan de gestion de données

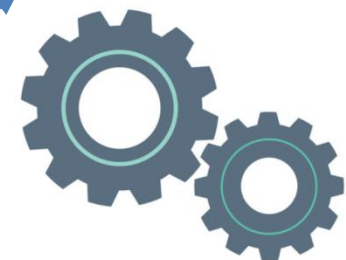
Données d'acquisition et prétraitées



Publication



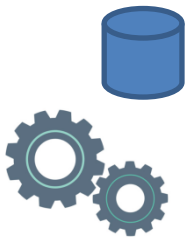
**ENTREPOTS de DONNEES**  
Data publiée  
DOI  
**FAIR**



Mésocentres labélisés

Au niveau régional

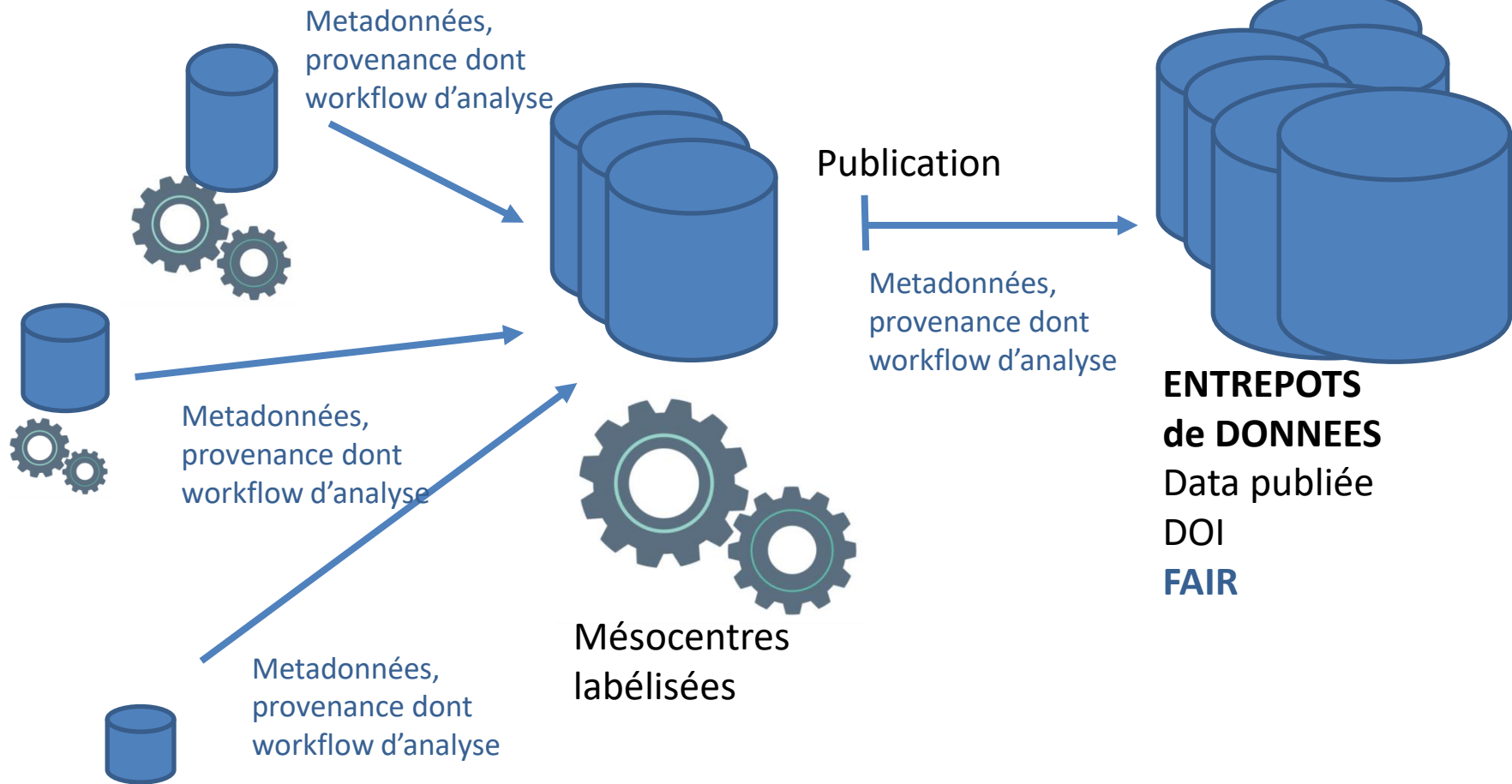
Plateformes des nœuds FBI

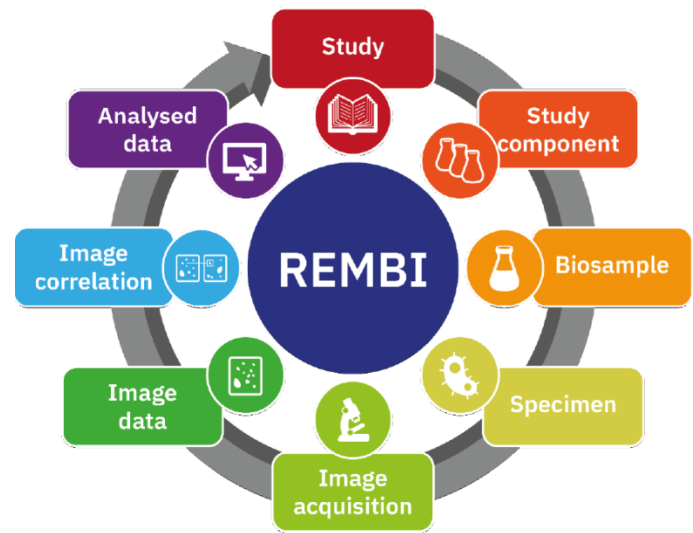
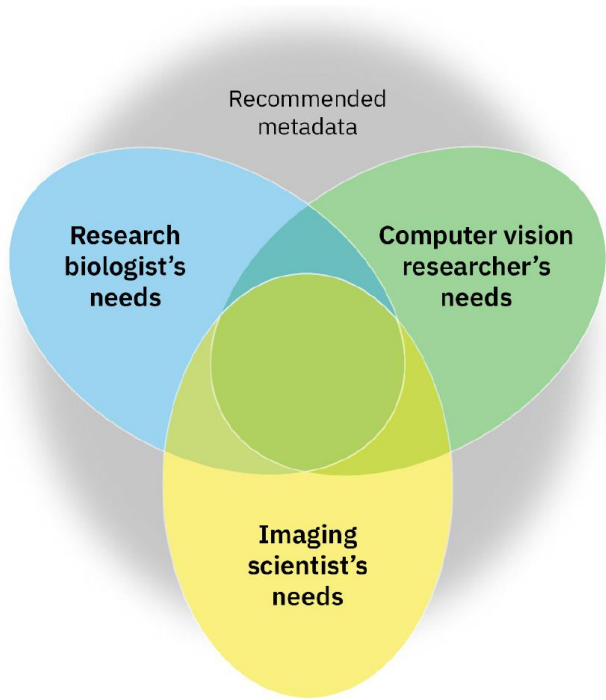


Au niveau national ou européen









Recommended Metadata for Biological Images

Collaboration EUBI (dont FBI), EMBL-EBI, EMBL, Elixir  
Preprint submitted to preprints.org



FRANCE-BIOIMAGING



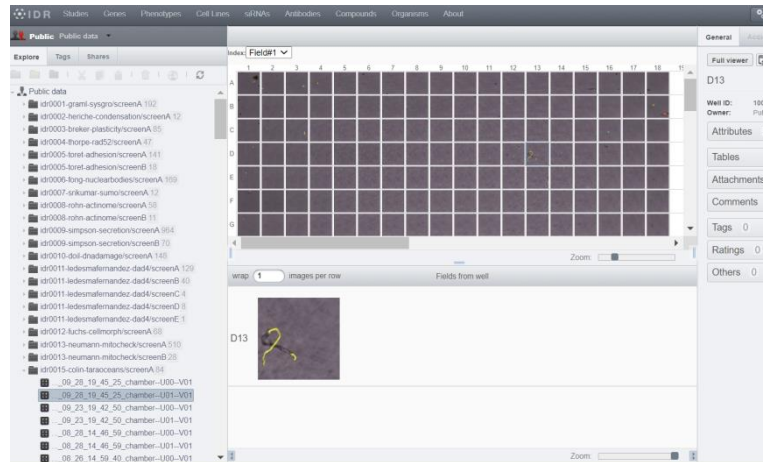
BioImage Archive



HUMAN CELL ATLAS

# Les entrepôts de données image

Scope	Technologies	Biological Samples and Models	Data resource
Molecules to cells	Electron cryo-microscopy (EM)*, Electron cryotomography*, Scanning EM (FIB SEM, SBF SEM, array tomography), Soft X-ray tomography, X-ray microscopy	Macromolecular structure by single-particle methods, high-resolution subcellular organelles and structure.	EMPIAR
Cells	Wide field, TIRF Confocal, Quantitative Phase Microscopy, High content screening	Single molecule (LM), Subcellular and cellular structure architecture, dynamics; Single cell macromolecule localization; Cellular response to small molecule- or genome-wide perturbation.	Cell-IDR
Tissues	Light-Sheet Microscopy, Multiphoton Microscopy, Digital Pathology, Optical Projection Tomography, Micro-CT Mass cytometry	Tissue structure, molecular localization and dynamics, Pathology and histology sections.	Tissue-IDR
All biological images which do not fit into any of the resources above	Magnetic resonance imaging, Ultrasound imaging, Super-resolution microscopy, photoacoustic imaging, light-field microscopy, Brillouin microscopy, Atomic Force Microscopy, other	Human medical data is out of scope.	BioImage Archive

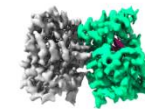


## pheromone receptor Rgg3 in complex with SHP3

**Publication:** Structure-function studies of Rgg binding to pheromones and target promoters reveal a model of transcription factor interplay  
 Capodagli GC, Tyler KM, Kaelber JT, Petrou VI, Federle MJ, Neiditch MB  
*Proc. Natl. Acad. Sci. U.S.A.* 117 24494-24502 (2020)  
 PMID: 32907945  
 DOI: 10.1073/pnas.2008427117

**Related PDB entry:** 7jjo  
**Related EMBD entry:** EMD-22341  
**Deposited:** 2020-07-22  
**Released:** 2020-11-13  
**Last modified:** 2020-11-13  
**Dataset size:** 3.0 TB  
**Dataset DOI:** 10.6019/EMPIAR-10535  
**Experimental metadata:** [Download xml](#)

**Contains:**  
 micrographs



### Image set

- 1. Unaligned multi-frame micrographs of Rgg3-SHP3 complex

**Category:** micrographs - multiframe  
**Image format:** MRC  
**No. of images or tilt series:** 1464  
**Frames per image:** 40  
**Image size:** (3838, 3710)  
**Pixel type:** 32 BIT FLOAT  
**Pixel spacing:** (1.038 Å, 1.038 Å)  
**Details:** The dataset contains unaligned gain-corrected micrograph movies in mrc format. Cryo-EM data were collected at the Rutgers New Jersey

Statistics

Re-use case study

EMPIAR in the news

Contact us

EMPIAR@PDB

### EMPIAR citations

Topaz-Denoise: general deep denoising models for cryoEM and cryoET.  
 Bepler T, Kelley K, Noble AJ, Berger B. (2020)

High-resolution cryo-EM structure of urase from the pathogen *Yersinia enterocolitica*.

Righetto RD, Anton L, Adaxio R, Jakob RP, Zivanov J, Mahi MA, Ringler P, Schwede T, Maier T, Stahlberg H. (2020)

Semantic segmentation of HeLa cells: An objective comparison between one traditional algorithm and four deep-learning architectures.

Karabag C, Jones ML, Peddie CJ, Weston AE, Collinson LM, Reyes-Maqueo G. (2020)

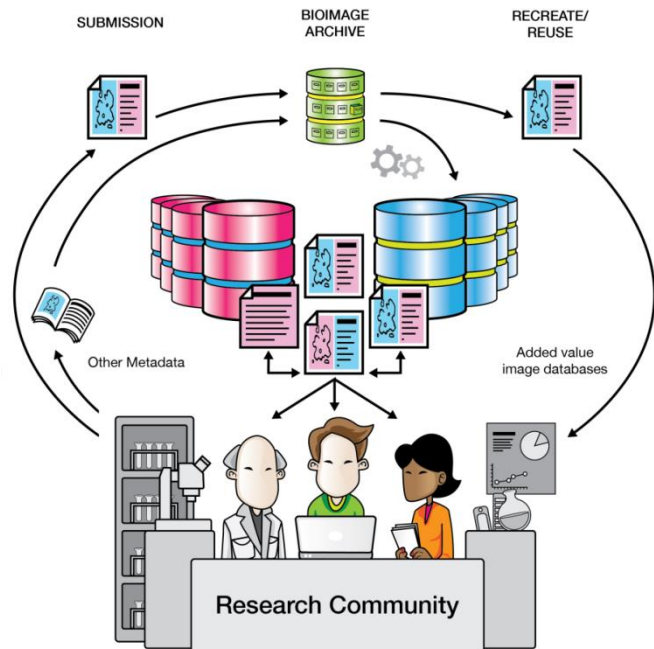


FRANCE-BIOIMAGING

# Les entrepôts de données image

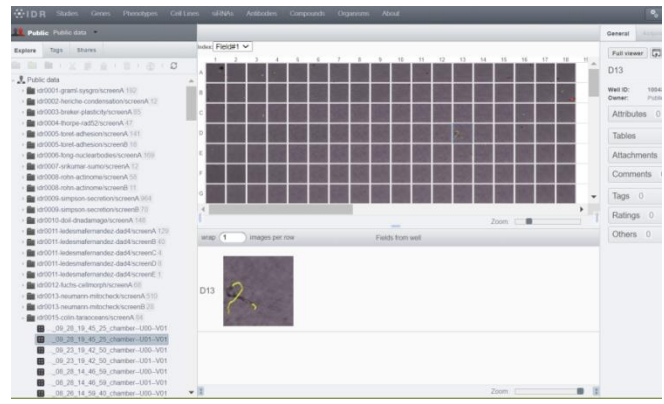


Scope	Technologies	Biological Samples and Models	Data resource
Molecules to cells	Electron cryo-microscopy (EM)*, Electron cryotomography*, Scanning EM (FIB SEM, SBF SEM, array tomography), Soft X-ray tomography, X-ray microscopy	Macromolecular structure by single-particle methods, high-resolution subcellular organelles and structure.	EMPIAR
Cells	Wide field, TIRF Confocal, Quantitative Phase Microscopy, High content screening	Single molecule (LM), Subcellular and cellular structure architecture, dynamics; Single cell macromolecule localization; Cellular response to small molecule- or genome-wide perturbation.	Cell-IDR
Tissues	Light-Sheet Microscopy, Multiphoton Microscopy, Digital Pathology, Optical Projection Tomography, Micro-CT Mass cytometry	Tissue structure, molecular localization and dynamics, Pathology and histology sections.	Tissue-IDR
All biological images which do not fit into any of the resources above	Magnetic resonance imaging, Ultrasound imaging, Super-resolution microscopy, photoacoustic imaging, light-field microscopy, Brillouin microscopy, Atomic Force Microscopy, other	Human medical data is out of scope.	BioImage Archive



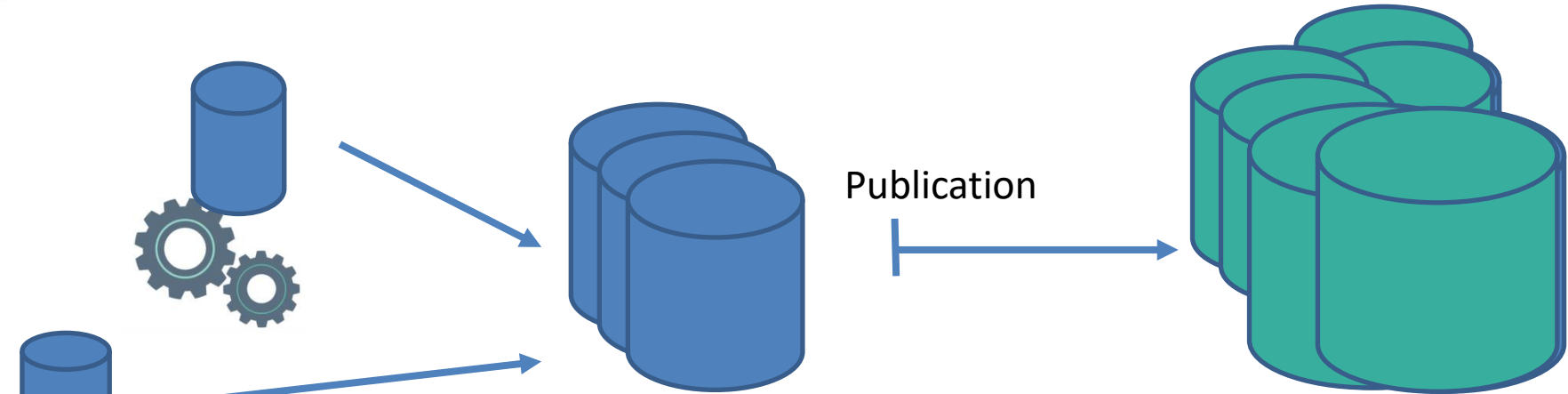
par rapport à Zenodo par exemple:

- Interface d'accès ++ par exemple: calcul sur entrepôts sans téléchargements des data.
- Structuration metadata spécifique au domaine
- Personnels CURATEURS dédiés en assistance aux chercheurs.
- Pas de limite de taille





Notre plan de gestion de données



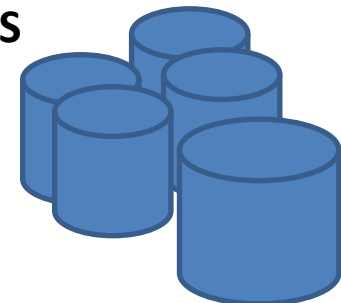
**ENTREPOTS de DONNEES Nationaux**

Création d'un entrepôt national : data broker, alignement sur les entrepôts européens , miroir ou complément? A construire avec partenaires européens

Mésocentres labélisées

Portail web d'accès unique de consultation et d'analyse (infrastructure distribuée)

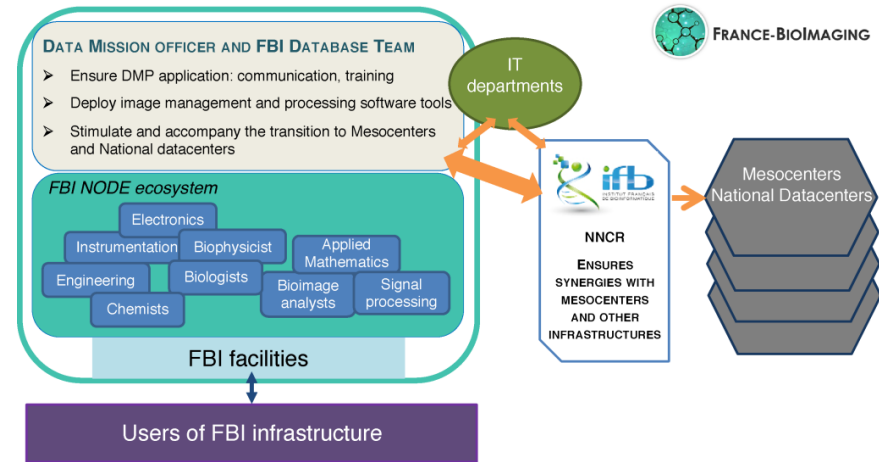
**ENTREPOTS de DONNEES**





D'ici 2024:

- Plan de gestion de données structure productrice de donnée imagerie (collaboration FBI, IFB, EMBRC, INIST OPIDOR)
- Déployer les outils nécessaires sur l'ensemble de l'infrastructure en collaboration avec l'IFB et les mésocentres labélisés, et former les plateformes, équipes R&D et utilisateurs
- Archive publique des données de bio imagerie de l'infrastructure et utilisation pour développement d'algorithme de machine learning (utile pour les biologistes, utiles pour les chercheurs en analyse d'image)



**Aligné avec les recommandations EOSC-life (partenaire EuroBioimaging-ERIC, implication européennes des chargés de missions FBI data)**





FRANCE-BIOIMAGING



THANK YOU FOR YOUR ATTENTION !!

- Jean Salamero FBI
- Pauliette Liebly IFB
- Emmanuel Faure FBI
- Jean-Fançois Dufayard IFB
- Olivier Colin IFB
- Faisal Beckouche EMBRC
- Frederic De Lamotte Inrae
- Fred Brau univ-nice
- Projet openlink
- Bio Image Analyst FBI
- Jean-Christophe Olivo-Marin FBI
- Charles Kervrann FBI

Equipe Opidor inist (MC Jacquemot)



Funded by the Horizon 2020 Framework Programme of the European Union

