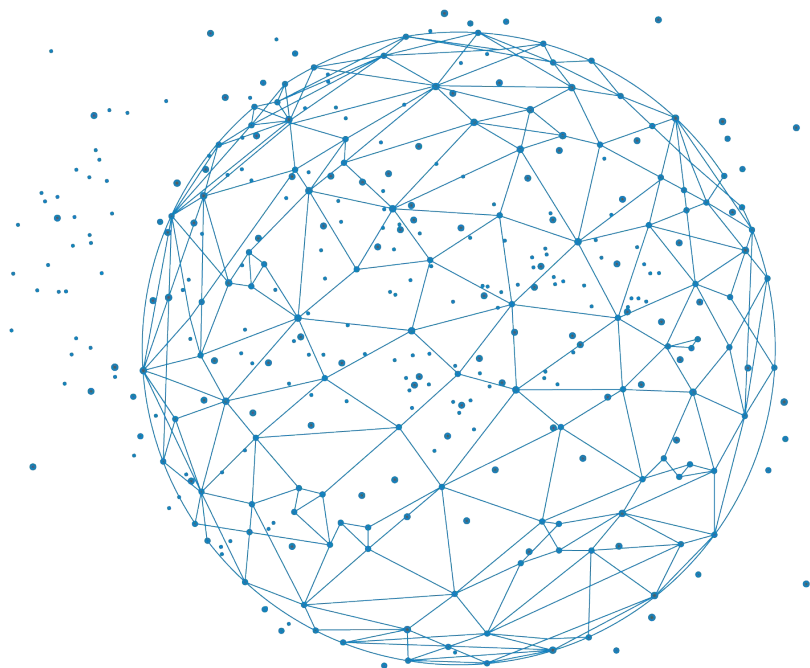


Nouvelle feuille de route nationale des infrastructures de recherche  
8 mars 2022, 9 h 00 – 12 h 00



# Institut Français de Bioinformatique (IFB)

Claudine Médigue & Jacques van Helden



# IFB – Une fédération de ressources en bioinformatique

## Une infra numérique distribuée

- 36 plateformes et équipes
- large couverture géographique
- couverture de tous les domaines d'expertise bioinformatique

22 plateformes-membres  
7 PF contributrices  
8 équipes associées  
>400 experts (~200 FTE)

## Des ressources mutualisées

- Financements d'infrastructure: PIA1 RENABI-IFB (22.8M€), PIA3 MUDIS4LS (16.5M€)
- missions articulées autour de "task forces" mutualisées entre plateformes

## Des projets collectifs

- EMERGEN, ABRomics, PEPR

## Coordination: IFB-core (UAR 3601)

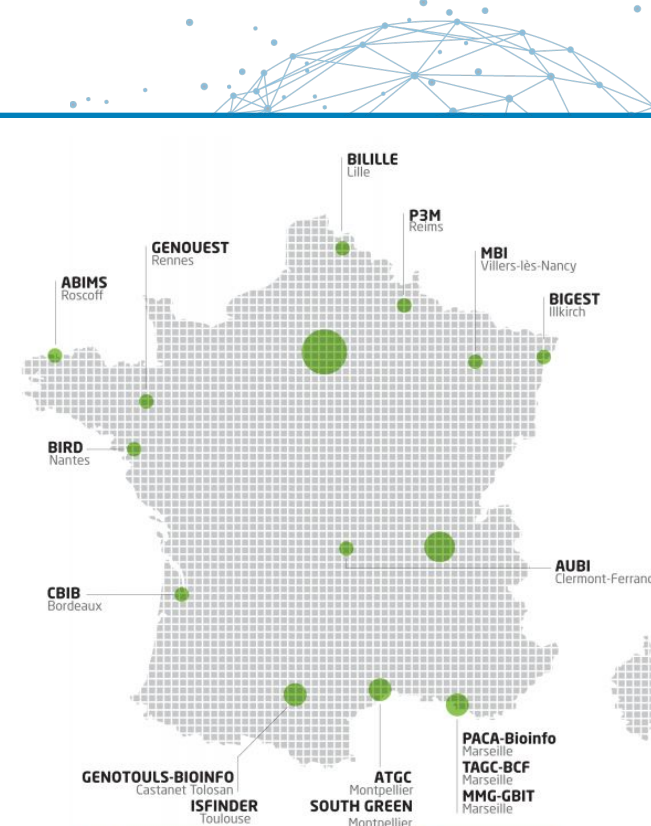
- multi-tutelle (CNRS, Inserm, INRAe, CEA)
- mission : coordination et gestion des moyens mutualisés
- noeud français d'ELIXIR (ELIXIR-FR)

## Atteindre toutes les communautés

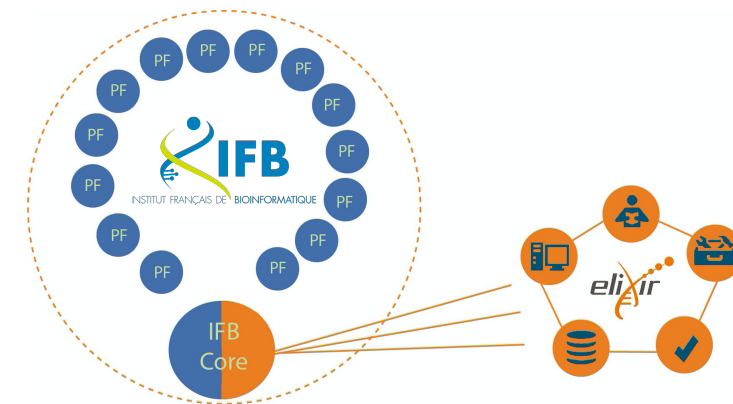
- biologie fondamentale, santé, agriculture, environnement

## Une articulation forte avec l'Europe

- noeud français de l'ESFRI ELIXIR
- point d'entrée pour des projets européens



RÉGION PARISIENNE	RÉGION LYONNAISE
EBIO Orsay Paris	INCA-SLC Lyon
INSTITUT CURIE Paris	PRABI-HCL Lyon
IGR Villejuif	PRABI-AMSB Villeurbanne
MICROSCOPE Evry	PRABI-Lyon-Grenoble Villeurbanne
MIGALE Jouy-en-Josas	PRABI-Lyon-Gerland Lyon
C3BI Paris	
RPBS Paris	
URGI Versailles	
ORPHANET Paris	
ICONICS Paris	
IFB CORE Evry	



# Missions et actions de notre feuille de route 2022-2025

## M1 Infrastructure numérique

- National Network of Computing Resources (**NNCR**)
- Calcul intensif, deux modalités: **cluster + cloud**
- Protection des données (**GDPR**)

11 Po stockage  
21K cœurs

## M2 Outils logiciels

- Développement de logiciels spécialisés
- **Packaging** (conda)
- **Virtualisation** (appliances, containers)
- **Bonnes pratiques** d'ingénierie logicielle
- **Développement collaboratif, ouverture du code**

Plus de **900** outils  
disponibles

## M7 Accompagnement des usagers

- Un point d'entrée unique pour tous les services et toutes les plateformes IFB
- Soutien aux projets dès la conception
- Plans de gestion des données (**PGD**)
- Conception et implémentation de **workflows**
- **Science des données**

400 experts  
>200 FTE

9115 utilisateurs  
310 projets  
collaboratifs



## M3 Bases de connaissances

- Développement de bases de données et de connaissances
- Standards de qualité
- Support technique aux développeurs de BD
- Déploiement de bases de données développées en France

## M4 Formations

- **Ecoles thématiques** (NGS, multi-omiques, phylogénie, biostat, programmation, workflows...)
- Formations **Bring Your Own Data** (BYOD)
- Webinaires, MOOC
- **Nouveau: formations FAIR-data et FAIR-bioinfo**
- Adaptation permanente à l'évolution de la demande

136 formations  
2326 personnes formées  
341 jours de formation

## M6 Innovation & prospective

- Bioinformatique intégrative (multi-omique, imagerie, phénomique, structure...)
- Intégration des données de santé
- Intelligence artificielle pour la biologie-santé

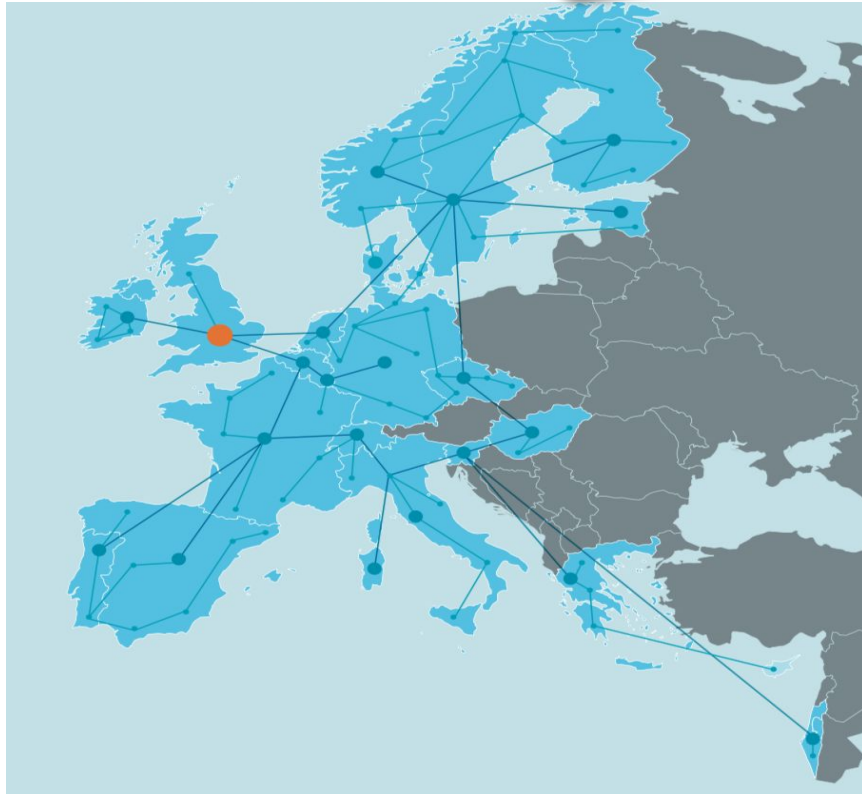
## M5 Science ouverte & Interopérabilité

- Gestion des données
- Interopérabilité, standards, ontologies...
- Courtage des données
- PGD automatisables
- Coopération avec les infra productrices de données

ELIXIR: a distributed infrastructure for Life Science Information

<https://www.elixir-europe.org/>

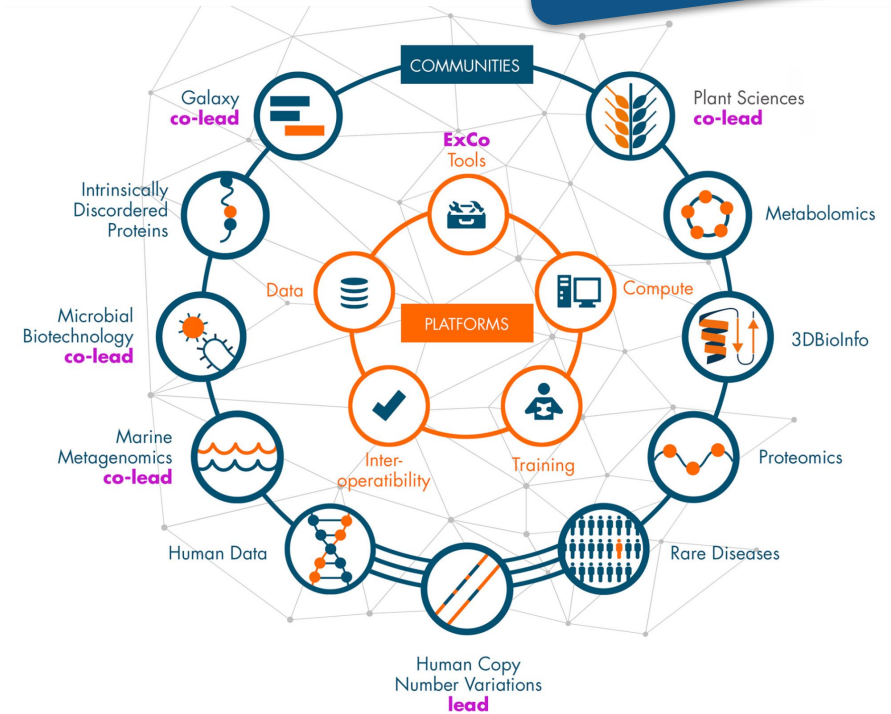
23 noeuds nationaux  
1 hub coordinateur  
200 instituts



Harmonisation européenne des services en bioinformatique

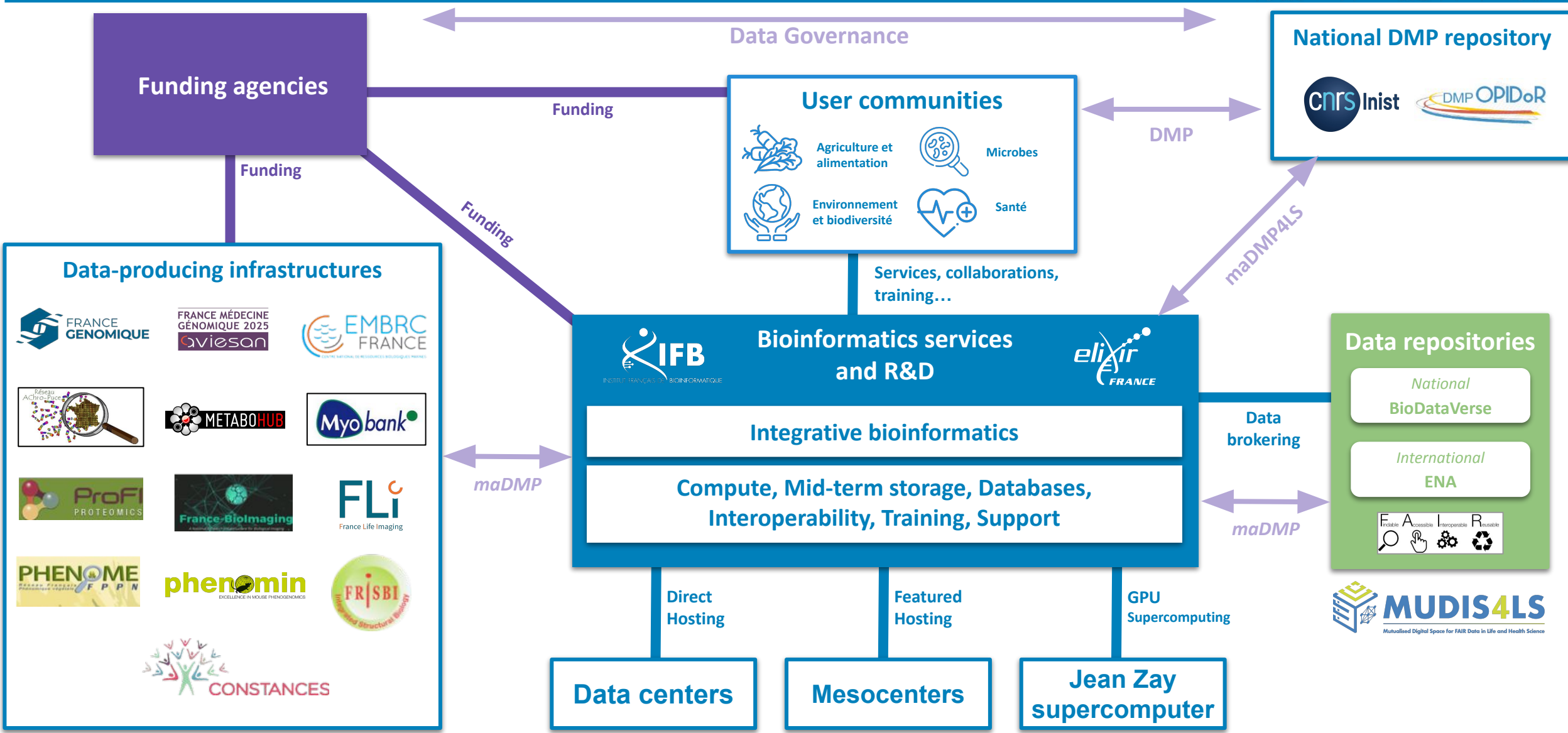
- Activités organisées autour de Commissioned Services
- Un levier pour entrer dans les projets européens (1+MG, B1MG, GID, CONVERGE, ...)

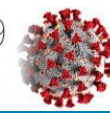
5 plateformes technologiques  
11 communautés thématiques



Food and Nutrition: emerging  
Toxicology: emerging - co-lead

# La bioinformatique au carrefour des données biologie-santé





## M1 Infrastructure numérique

- Déploiement des serveurs: [COVIDici](#), [COVimpact](#), [Covidix19](#), [COVIDSIM](#), [COVIDistress](#)
- Hébergement de calcul intensif (docking structurel)

## M2 Outils logiciels

- Déploiement de logiciel spécialisé
- Génomique comparative des virus: [PIPprofiler](#)
- ELIXIR [COVID-19 Workflows Hub](#)

## M8 User support & expertise

- [Criblage in vitro d'inhibiteurs SARS-CoV-2](#)
- [Analyse phylogénétique de l'origine de SARS-CoV-2](#)
- Développement d'un kit de contrôle de qualité universel

## M3 Bases de connaissances

- Section SARS-CoV-2 de [VirHostNet](#)
- [COVIDScan](#): tomodensitométrie thoracique de patients
- 



## Mutualised task forces



## M4 Formations

- [Home Learning – lessons about COVID-19](#) (ELIXIR biohackathon project)

## M7 Innovation & prospective

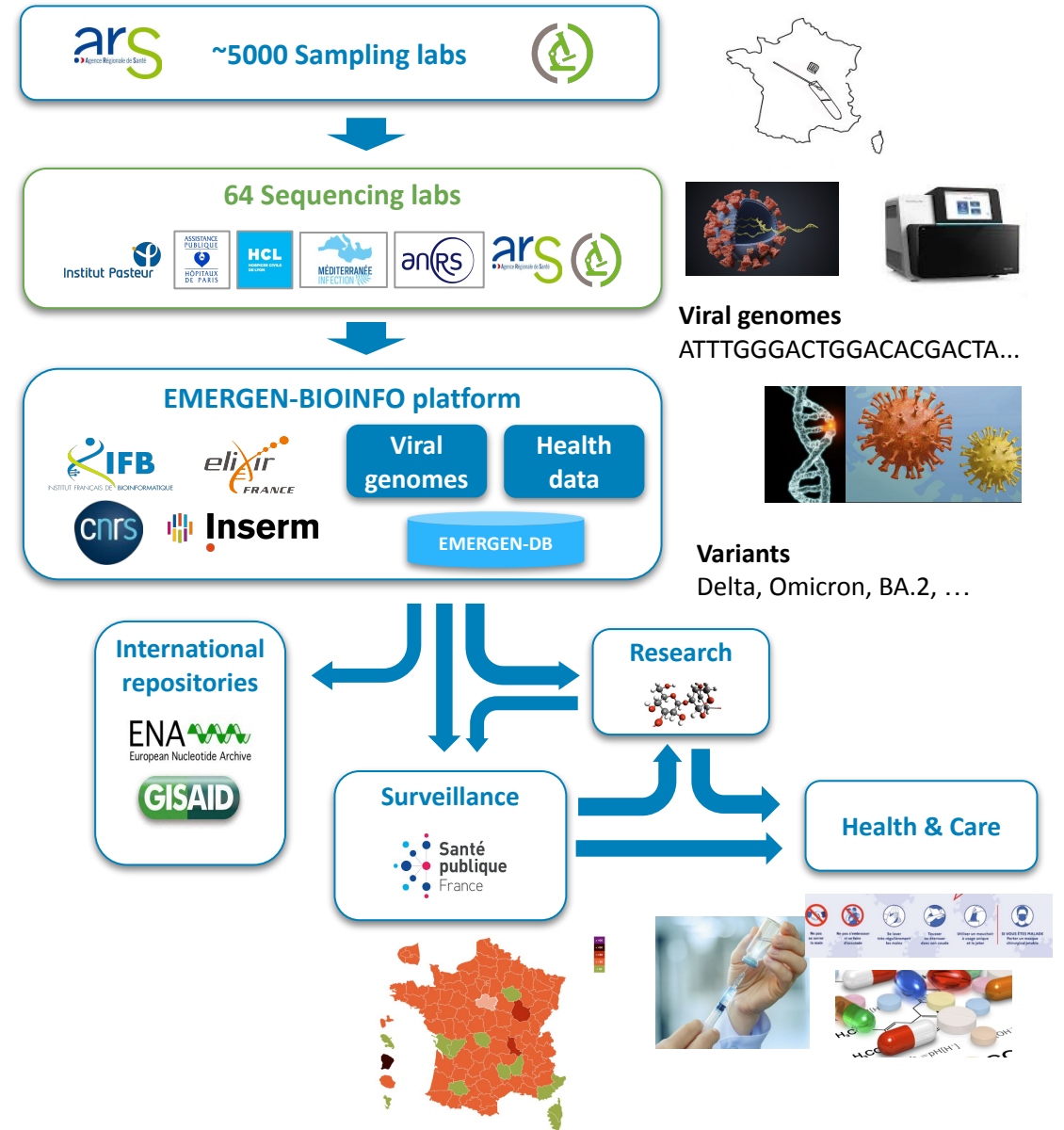
- ... bientôt sur vos écrans

## M5 Open science & M6 Interoperability

- EBI-ELIXIR monthly teleconferences of national node coordinators
- ELIXIR-CONVERGE 5M€ extension for COVID-19

# EMERGEN - une infrastructure répond à l'urgence sanitaire COVID-19

- **Launching : January 2021**
- **Funding**
  - **60M€** Ministry of Health
  - **10M€** Ministry of Research
- **Coordination**
  - Surveillance : Santé Publique France
  - Research ; ANRS | MIE
- **Sample collection : 5 000 labs**
  - **4 400** private labs
  - **~500** hospital facilities
- **Sequencing : 64 labs**
  - **4** National Reference Centers (CNR) associated
  - **46** ANRS | MIE
  - **14** Private labs
- **Data volume (March 3, 2022) :**
  - **329,875** consensus sequences
- **Delays** (average for January samples)
  - Sampling -> Expedition : 5 days
  - Sample delivery → Sequencing : **14 days**
  - Sequencing → EMERGEN-DB submission : 3 days
- **Submission paths**
  - Direct from sequencing lab to GISAID (CNR)
  - Via IFB/ELIXIR-FR (in pre-production)





## COVID-19 Data portal

- Une fédération de portails nationaux
- Tous types de données (séquences, vaccination, épidémio)

## ELIXIR-CONVERGE

- Réseau fondé en 2019, autour du partage des données (WP1–WP6)
- en 2020, la CE le dote de 5 M€ supplémentaires pour 3 WP “cas d’étude” COVID-19 (WP7–WP9)

## Implication de l’IFB dans des projets européens

- [HERA](#): évaluation de la qualité pour les 60 laboratoires séquenceurs d’EMERGEN
- [ISIDORE](#): Integrated Services for Infectious Disease Outbreak Research

## Enjeux sociétaux: accès, protection et ouverture des données

- Dépôts ouverts (INSDC) vs licence protectrice (GISAID)
- Enjeux contradictoires liés à la double finalité surveillance sanitaire / recherche
- A l’échelle nationale
- A l’échelle internationaux (ELIXIR, ECDC, OMS)
- Table-ronde ELIXIR le 30 mars

## Participants

29 research institutes in 22 countries across Europe.

## Work Packages

- WP1: Expert network
- WP2: Training and Capacity Building
- WP3: Common Data Management Toolkit
- WP4: Communications, Industry, International, Impact and Sustainability
- WP5: Demonstrator Projects
- WP6: Project Management and Scientific Coordination

## 2020: 5M€ supplementary funding for COVID-19 “study cases”

- WP7: Federated European Genome-phenome Archives for transnational access of COVID-19 host data
- WP8: ELIXIR-CONVERGE European COVID-19 Data Platform
- WP9: Mobilisation of SARS-CoV-2 variant surveillance data tracking services and tools



## Forces

- Dynamique collective des plateformes IFB
- Proximité des équipes de recherche
- Large spectre de compétences techniques
- Point d'entrée pour appuyer les projets dans tous les domaines thématiques biologie-santé
- Catalyse de projets inter-INBS
- Connecteur à l'international
- Formation
- Capacité à appuyer des politiques publiques de surveillance sanitaire

## Faiblesses

- Limitation des postes statutaires pour affronter la croissance de la demande et des attentes
- Manque de soutien administratif pour projets complexes
- Inadéquation du statut
- Absence (actuelle) de ressource numérique HDS

## Opportunités

- Relever le défi du traitement des données de santé (France: HDS; projets EU 1+MG, B1MG, GDI...)
- Transposer au niveau EU la capacité structurante et l'impact
- IFB centre de référence des données biologie-santé
- Certification organisme de formation
- Interaction avec INBS environnement, biodiversité

## Menaces

- Limitation du vivier des ressources humaines recrutables
- Manque d'attractivité des salaires → fuite des compétences
- Epuisement des ressources humaines du fait de la sur-sollicitation
- Manque de visibilité sur la pérennisation post-2024

