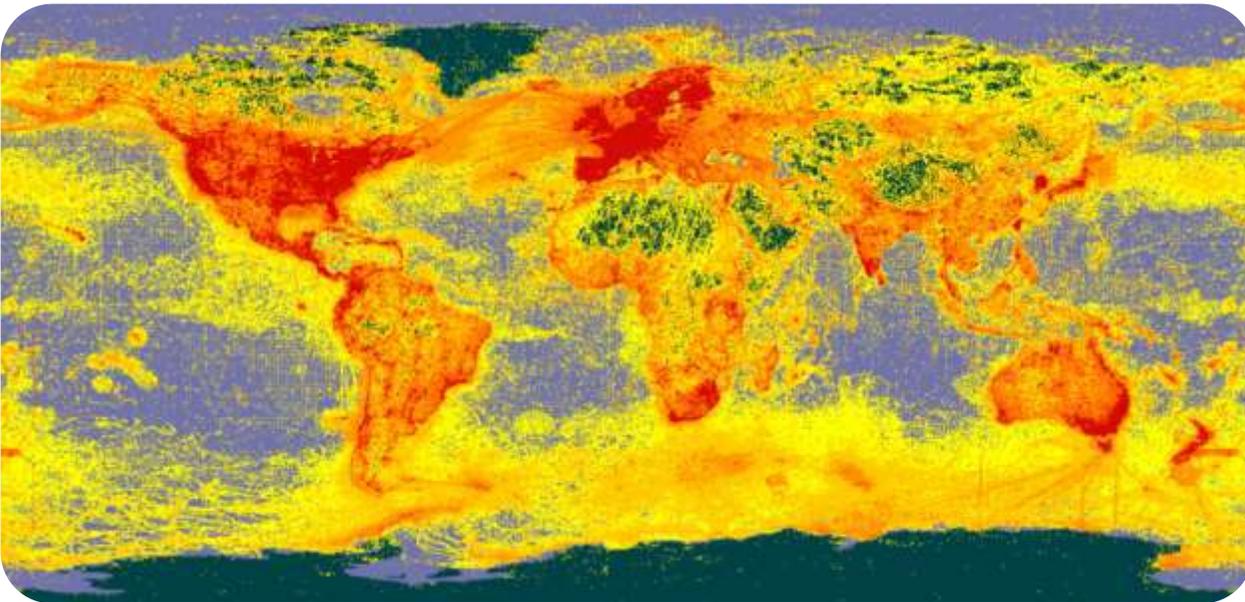


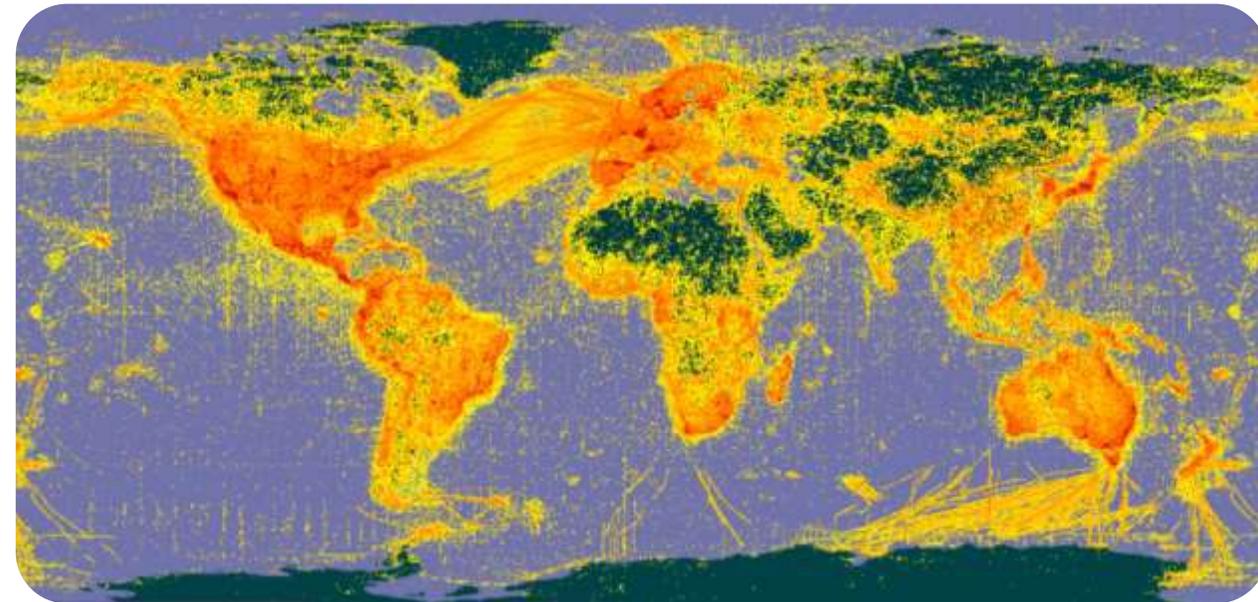
# Enjeux et défis de l'accès aux données des collections naturalistes dans le cadre de l'IR Récolnat

François DUSOULIER (MNHN/DGD-C)





Carte de densité des données issues d'observations naturalistes (GBIF)



Carte de densité des données issues de spécimens en collections (GBIF)



2 328 624 048

Occurrence records



85 890

Datasets



2 054

Publishing institutions



8 869

Peer-reviewed papers using data

=> Connaissance de la biodiversité fondée essentiellement sur des observations naturalistes.

## Raison d'être

- Handicap taxinomique
- Crise de la reproductibilité (Baker, 2016)
- Risque de déconnexion entre science et politiques publiques de conservation
- Difficultés d'accès

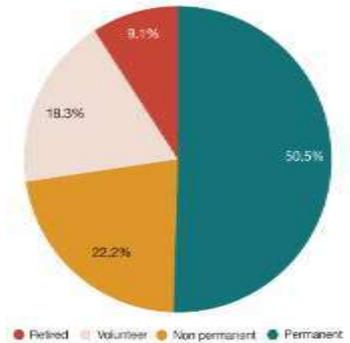
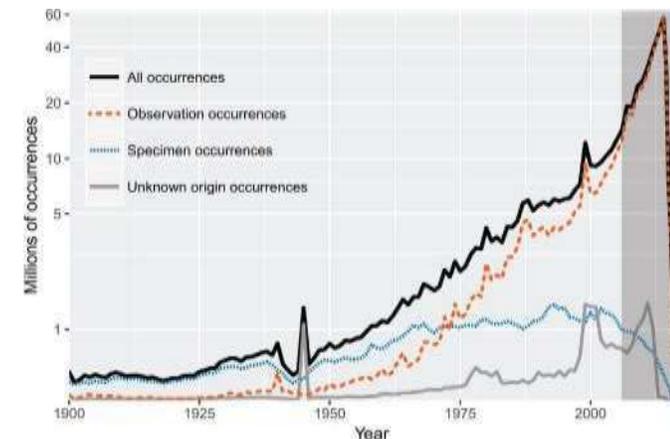


Fig. 11: Employment status of European taxonomists.



=> Utilisation par les sciences environnementales et la société mais potentiel largement sous-exploité

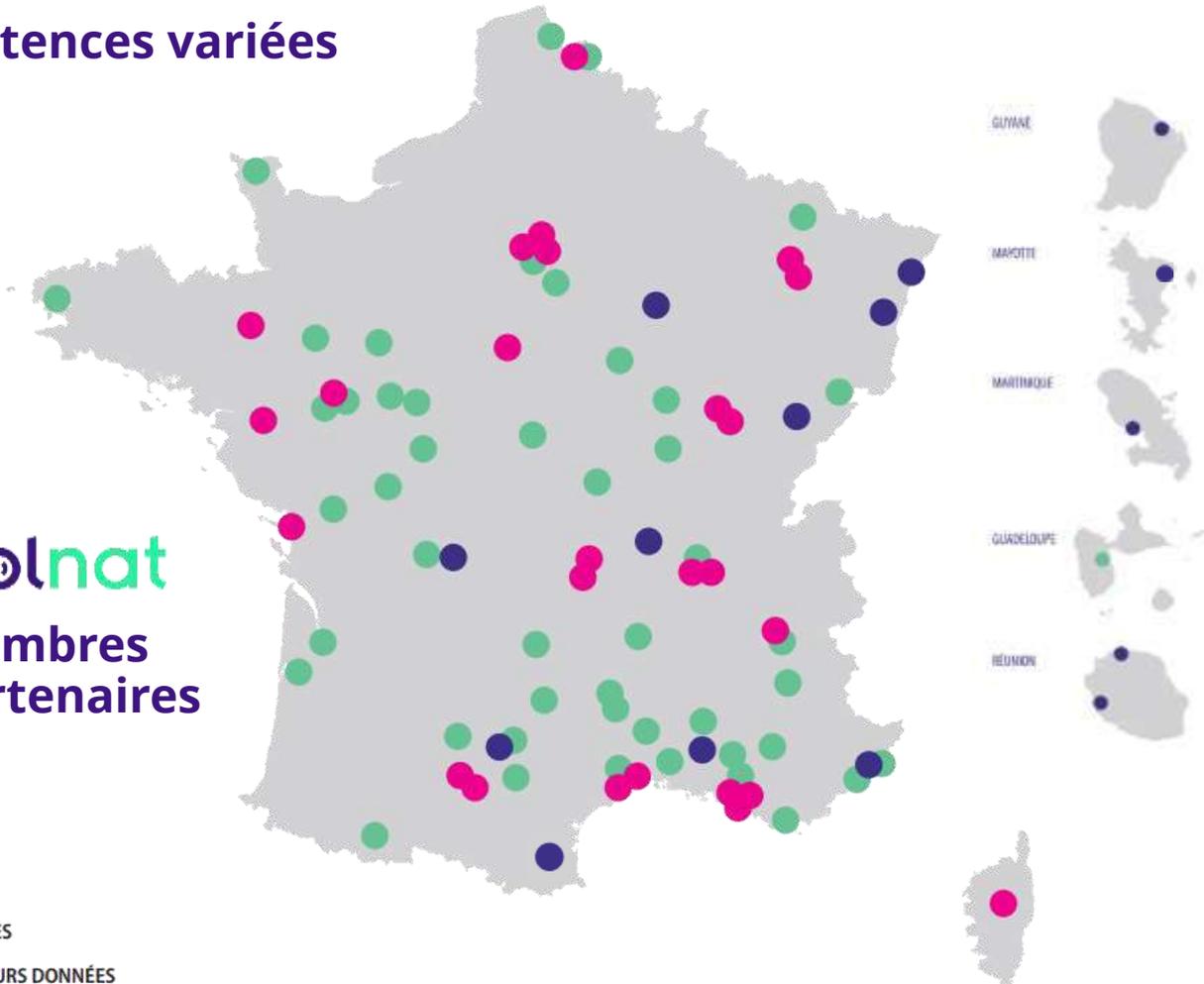


## Une infrastructure distribuée et toute la force d'un réseau de personnes aux compétences variées



 Récolnat  
**92** 26 membres  
66 partenaires

-  MEMBRES
-  PARTENAIRES
-  FOURNISSEURS DONNÉES



## Périmètre des collections naturalistes

Botanique, Mycologie



Zoologie



Ethnobiologie



Tout spécimen, échantillon ou ensemble de collections constituées dans un but scientifique, à l'exception des « formes » domestiquées

Paléontologie



Géologie



Extraterrestre



*Spécimens*  
*Échantillons*  
*Fragments, parties*  
*Traces*  
*Sons*

# Singularité des collections naturalistes

Diversité de matières, de tailles et de supports

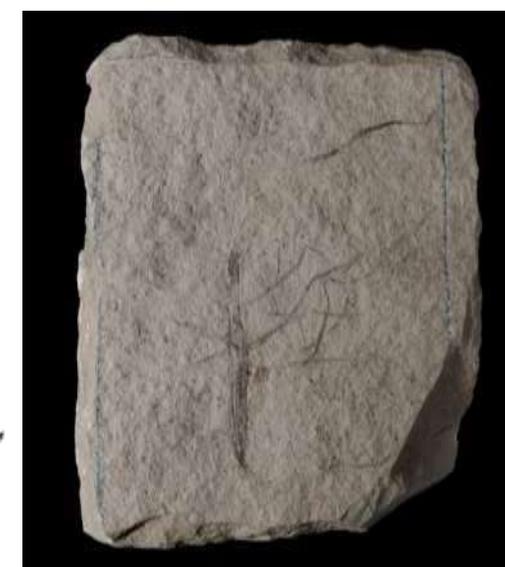
Nombre : fausses « séries » d'individus (ou d'holobiontes)

Profondeur historique (350 ans) et temporelle

Évolutivité des collections avec l'avancée de la science (ressources)

Matérialité qui permet la reproductibilité des expériences

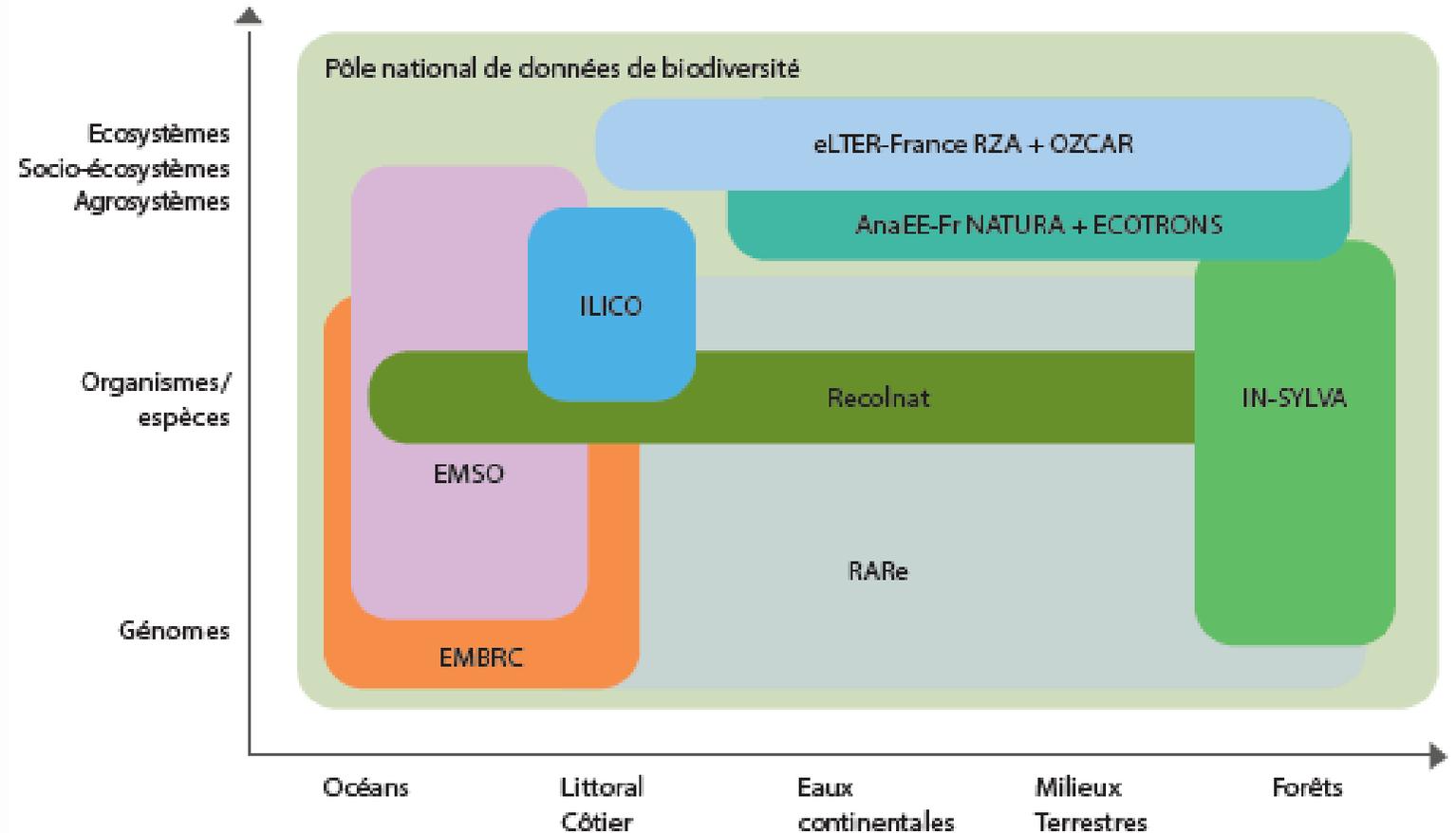
130 millions en France, 1 500 M en Europe,  
6 000 M dans le Monde



## Paysage des IR du domaine STE

**Livre Blanc sur les infrastructures  
 françaises de recherche du domaine  
 des sciences du système Terre  
 et de l'environnement**

Vision stratégique d'AllEnvi 2020-2030



## Livre Blanc sur les infrastructures françaises de recherche du domaine des sciences du système Terre et de l'environnement

Vision stratégique d'AllEnvi 2020-2030



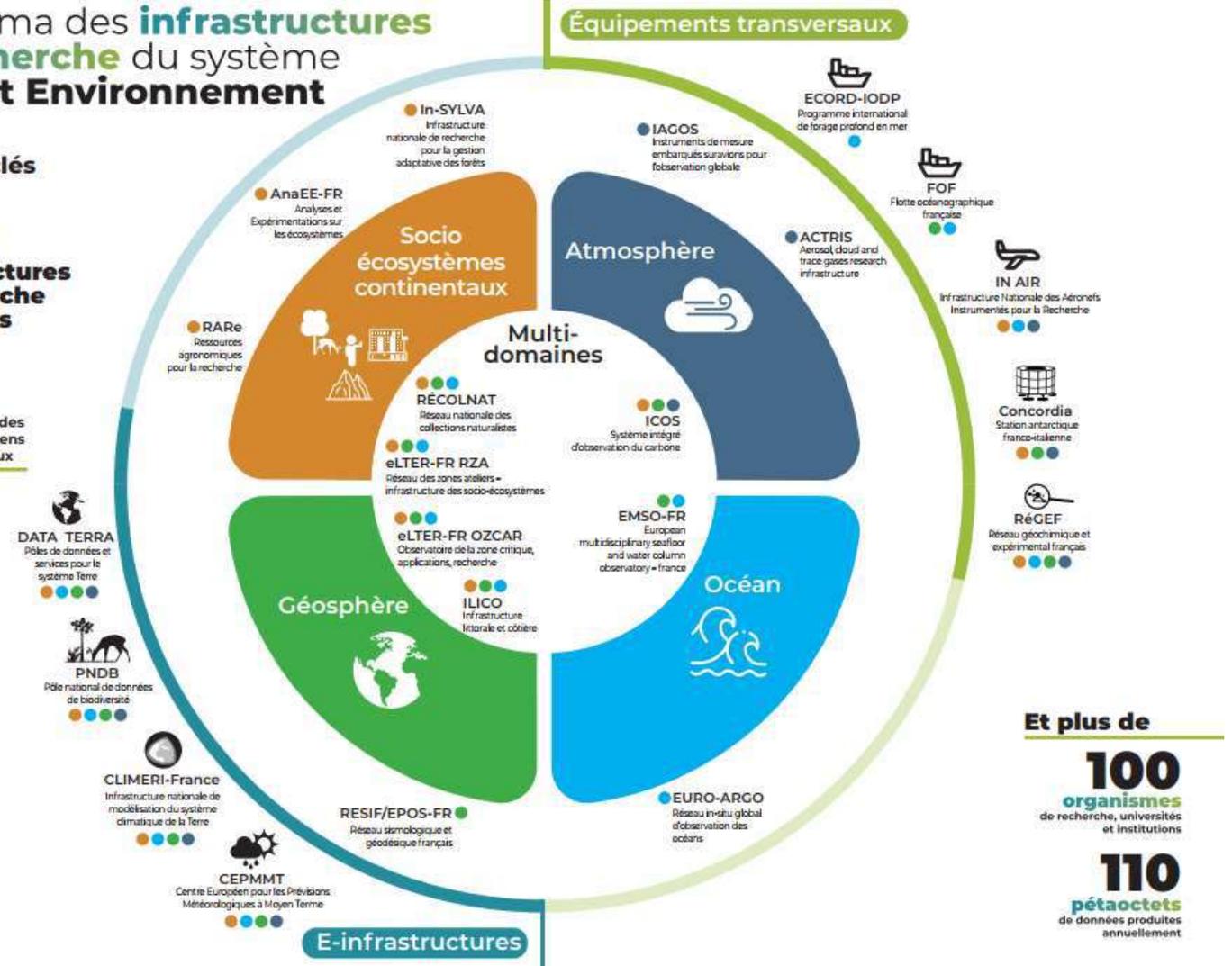
# Paysage des IR du domaine STE

Panorama des **infrastructures de recherche** du système Terre et Environnement

Chiffres-clés

**22** infrastructures de recherche labellisées

**17** intégrées dans des réseaux européens et internationaux



Et plus de

**100** organismes de recherche, universités et institutions

**110** pétaoctets de données produites annuellement

## L'infrastructure Récolnat : des enjeux forts

- Réseau humain réunissant les personnels responsables des collections des universités, du MNHN, des muséums territoriaux et autres établissements de recherche
- Produire, rassembler et donner accès aux collections naturalistes et leurs données
- Offrir un espace de concertation autour des pratiques professionnelles
- Promouvoir les collections naturalistes et l'importance des données scientifiques qu'elles recèlent dans le paysage des IR, en France et en Europe
- S'inscrire dans une démarche FAIR et de science ouverte



## Chiffres clés

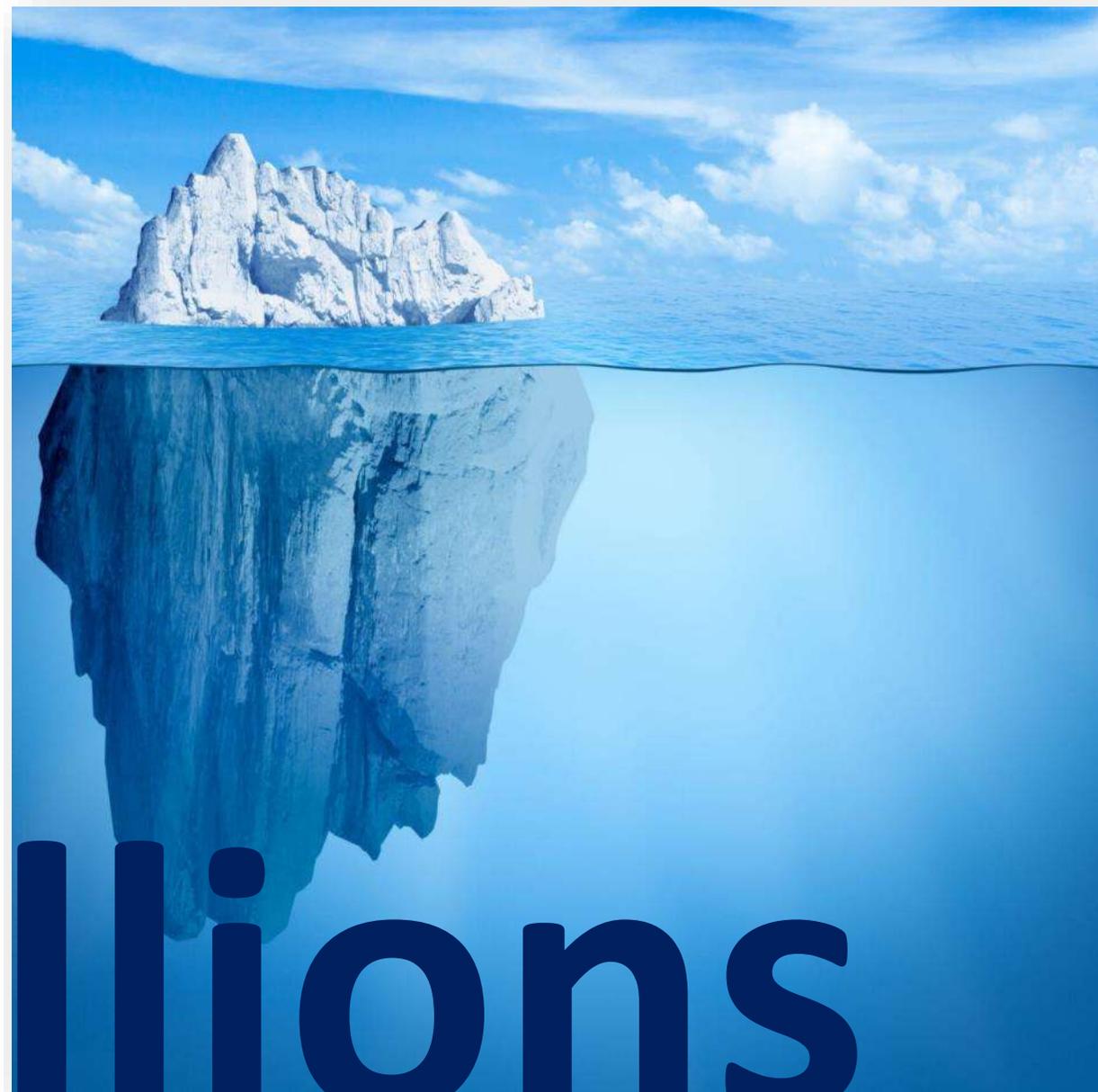
Corpus de **10,9 millions** dont 10 millions d'herbiers !  
Ca. **400 000** types et figurés, **360** collections de **92** institutions  
Ca. **15 000** utilisateurs/an utilisant le portail des collections  
Ca. **1 200** publications/an utilisant l'IR  
**ca. 450** personnes impliquées



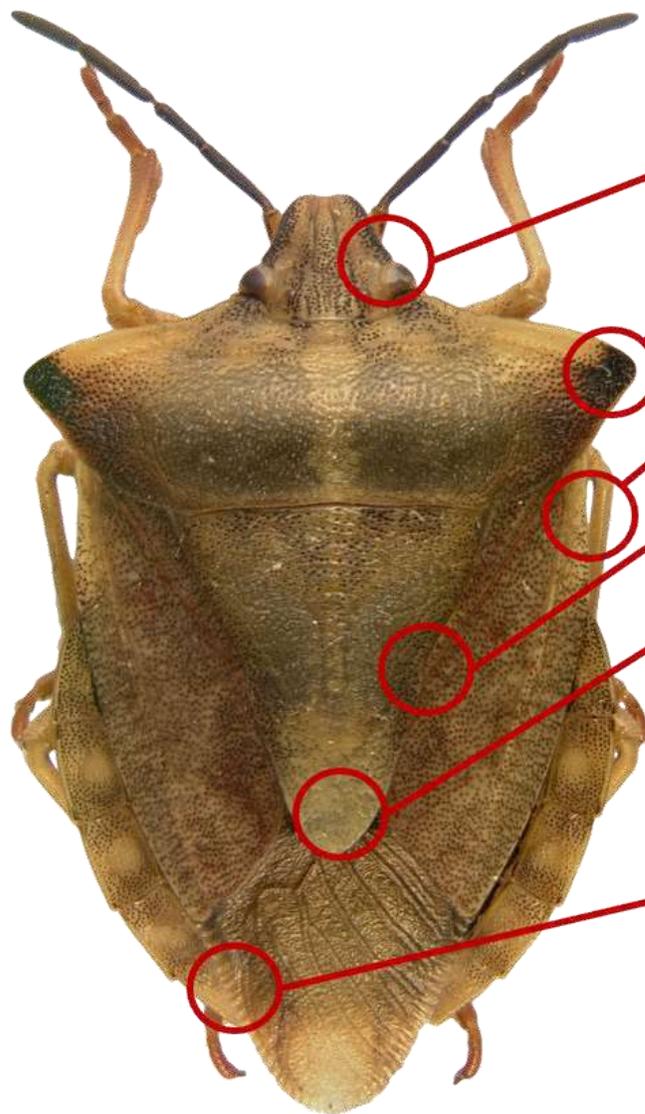
## Les collections

- 10 813 001 spécimens numérisés
- 724 241 taxons (36 %)
- 92 institutions

**130 millions**



## Quelles données dans un spécimen ?



Données génomiques

Données biochimiques

Données morphologiques

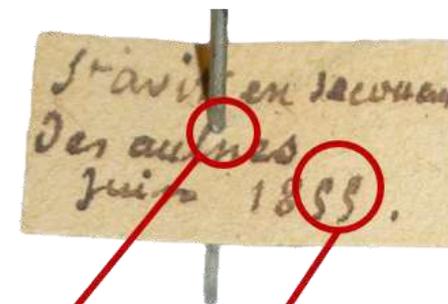
Données écologiques

Données taxinomiques

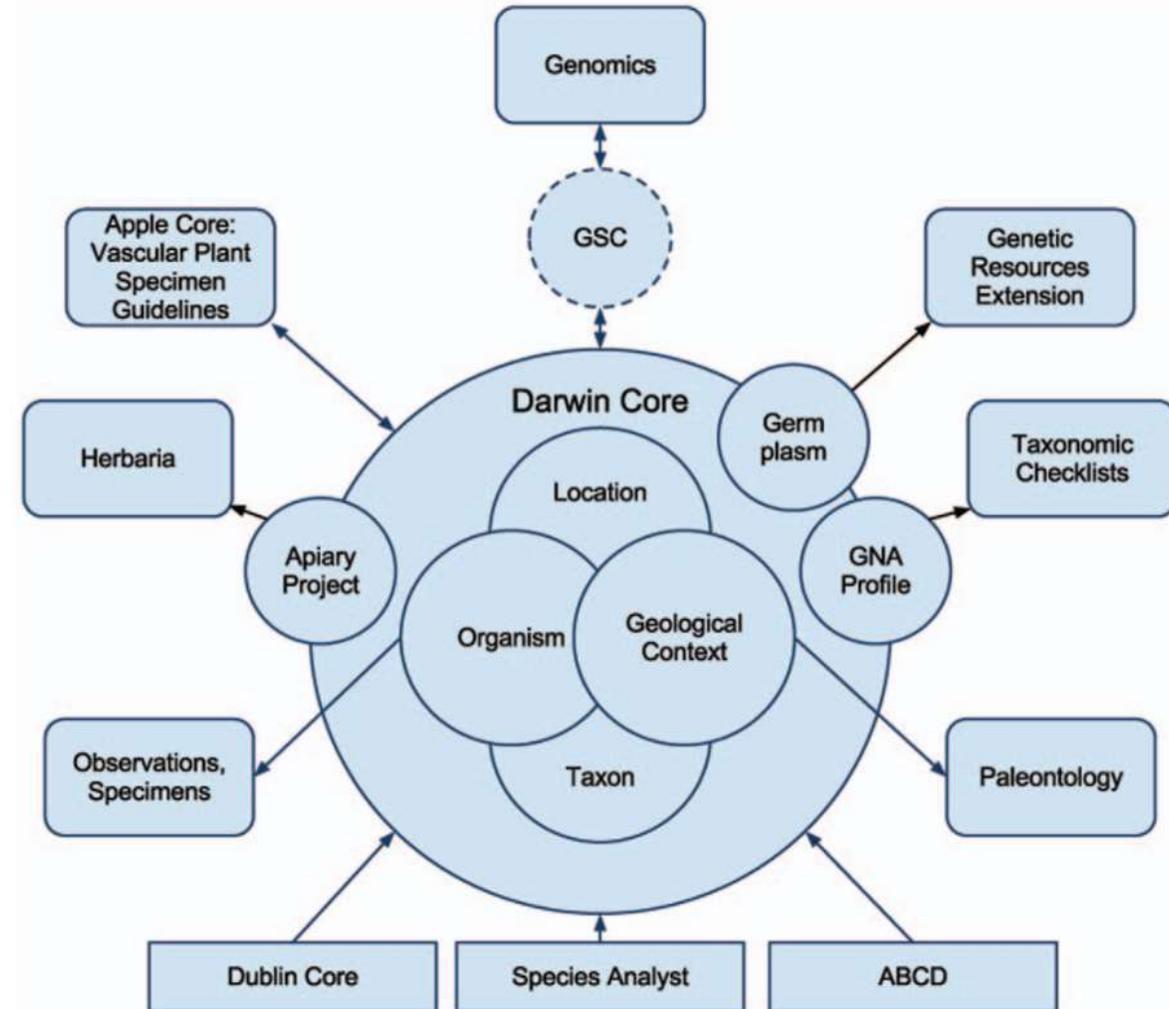
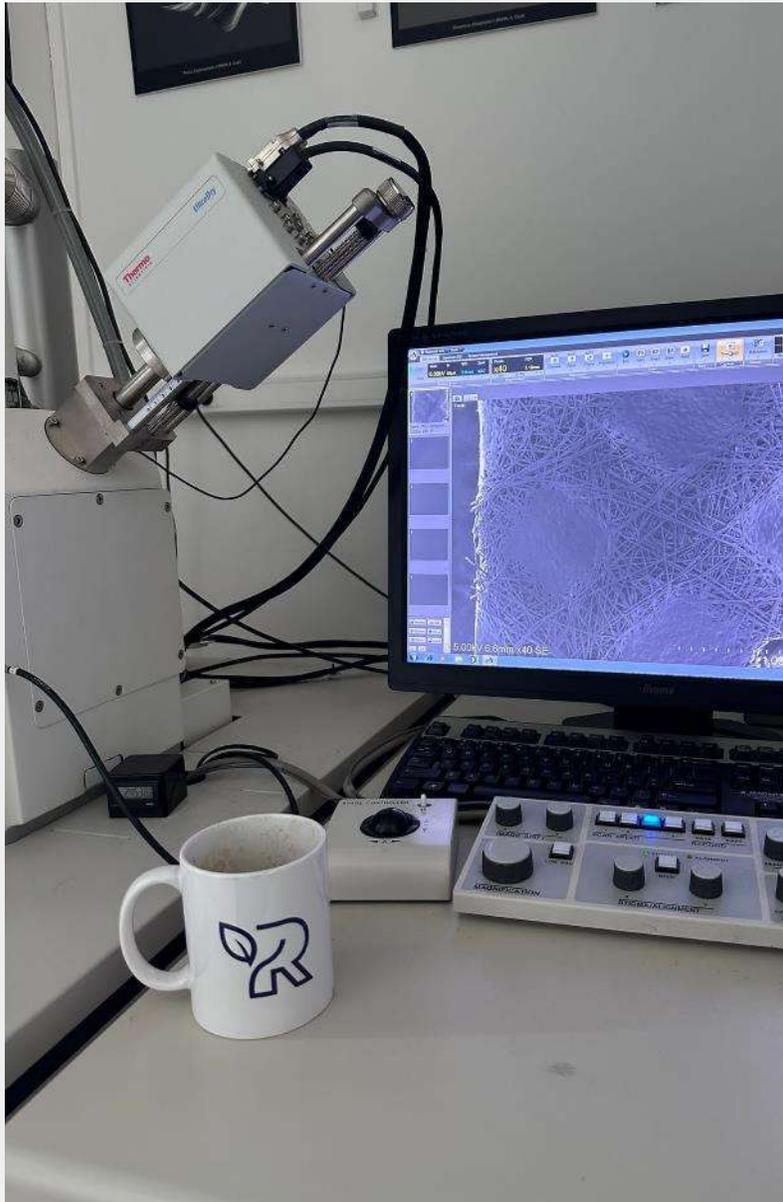
Données géographiques

Données techniques et culturelles

Données d'interactions inter-spécifiques



# Standards internationaux de description



## Données, complétude et qualité

- Données descriptives standardisées des spécimens et échantillons formant les collections naturalistes ;
- Mobilisation des archives compagnes (occurrences spatio-temporelles) ;
- Complétude inégale selon les corpus et leurs modes de production (artisanale vs. industrielle) ;
- Qualité dépendante des « couches hautes » (attributs et fraîcheur du réexamen), du degré d'accrochage et d'interconnexion avec des référentiels, comme d'autres réservoirs de données.



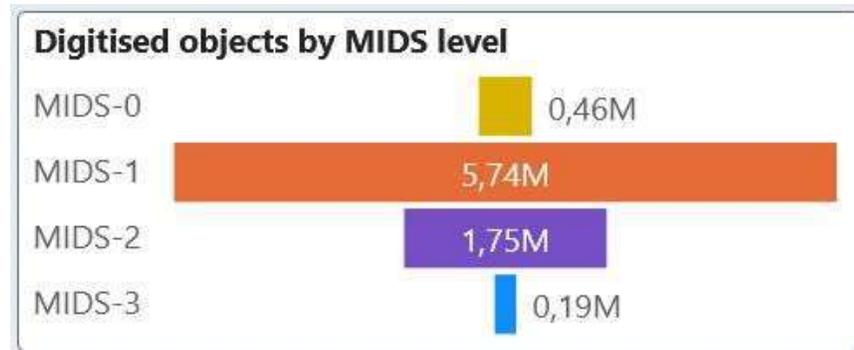


## Données

- 10 812 761 spécimens de botanique, paléontologie et zoologie dont 9 313 573 numérisés en 2D
- 724 134 taxa intégrés
- 80 fournisseurs de données



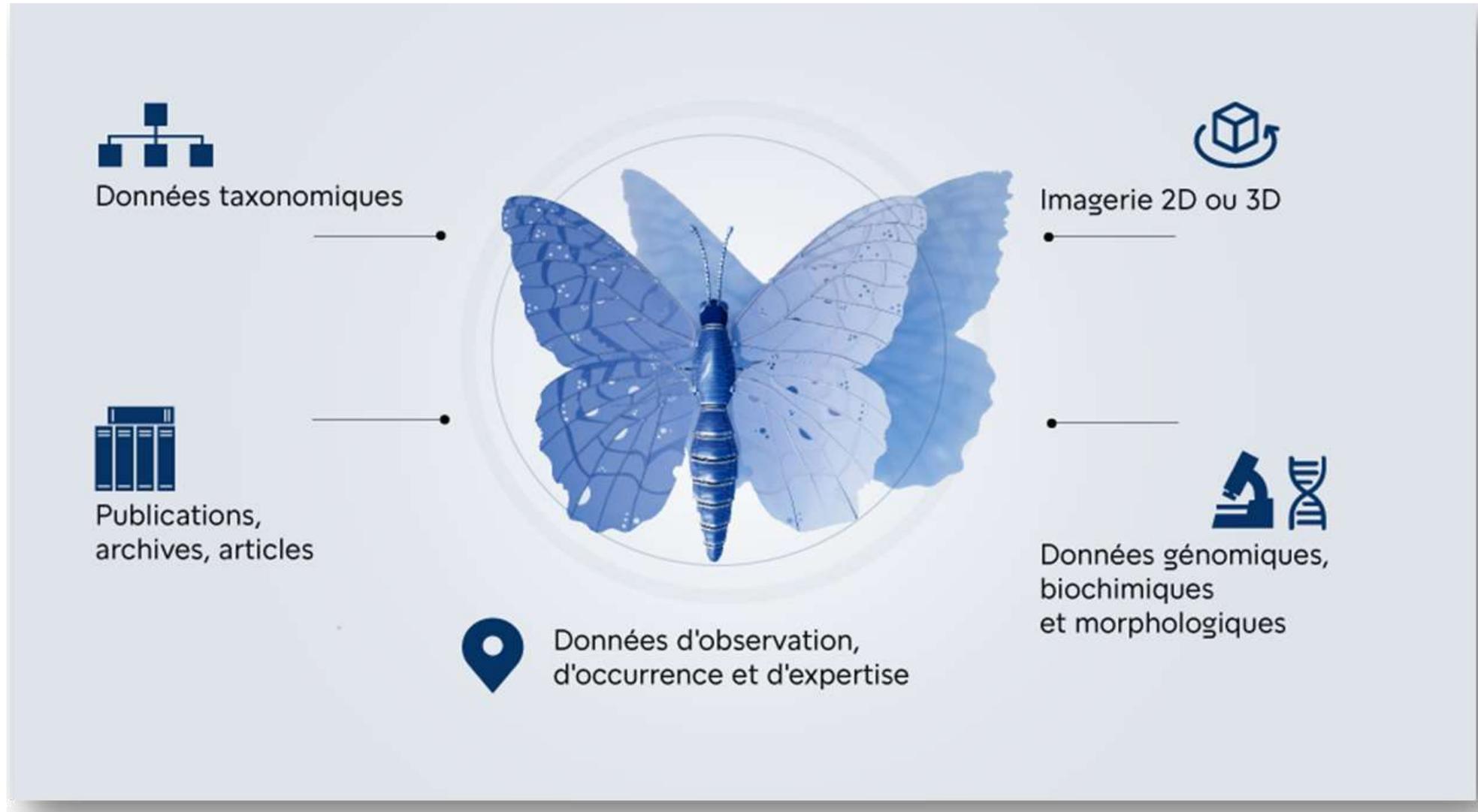
## Données, complétude et qualité

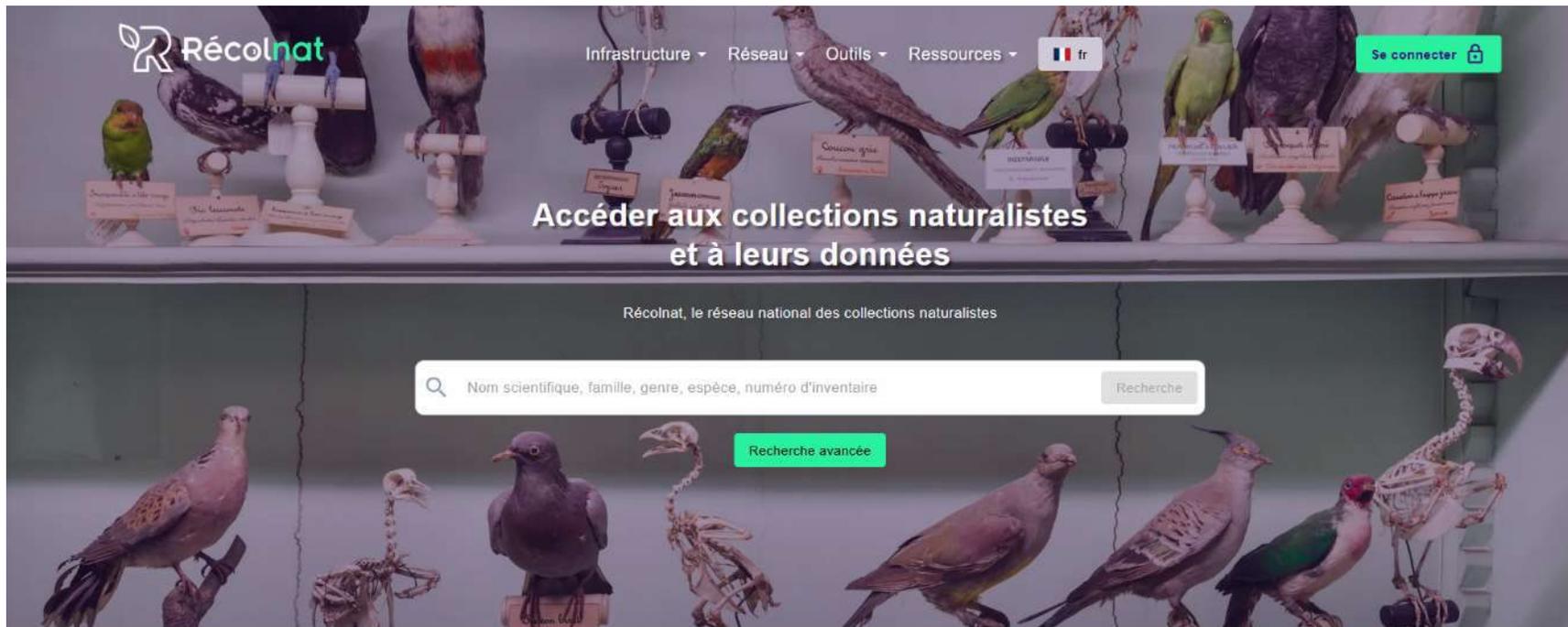


## Minimal Information for digital specimens (MIDS)

- MIDS-0 : métadonnées de base (institutions, code à barres)
- MIDS-1 : informations taxinomiques et géographiques supérieures/macro
- MIDS-2 : verbatim issu de l'étiquette
- MIDS-3 : données d'interprétation et d'attribution, notamment taxinomique et géographique

# Entrée dans le concept de spécimen numérique étendu





recolnat.org  
devient  
recolnat.fr

## Récolnat

### Le réseau national des collections naturalistes

Récolnat est l'infrastructure nationale qui donne accès aux collections d'histoire naturelle (botanique, zoologie, paléontologie, géologie) afin de favoriser leur utilisation dans tous les travaux scientifiques. Pilotée par le Muséum national d'histoire naturelle, elle a pour objectif ambitieux de fonder le catalogue collectif des collections naturalistes de France en coordonnant un réseau de personnes aux compétences et spécialités diversifiées. Récolnat est labellisé infrastructure de recherche (IR) et figure sur la feuille de route nationale du ministère de la Recherche (MESR) depuis 2016.





## Ajouter un spécimen

Domaine: Non défini ▾

Publier

Sauver brouillon

1 Détermination

2 Datation

3 Collecte

4 Spécimen

5 Bibliographie

6 Multimédia

7 Autres

### Détermination

 Doute sur la détermination

Date de la détermination

### Taxinomie

0 référentiels liés 

# Un exemple de complexité : gestion des noms et des attributions

## Multitude de noms pour désigner un taxon

<i>Forficula riparia</i> Pallas, 1773	→	Protonyme
<i>Forficesila riparia</i> (Pallas, 1773)	→	Nouvelle combinaison - aponyme
<b><i>Labidura riparia</i> (Pallas, 1773)</b>	→	<b>Nouvelle combinaison - kyronyme</b>
<i>Forficula pallipes</i> Fabricius, 1775	→	Synonyme subjectif (doxisonyme)
<i>Forficula bilineata</i> Herbst, 1786	→	Synonyme subjectif (doxisonyme)
<i>Forficula gigantea</i> Fabricius, 1787	→	Synonyme subjectif (doxisonyme)
<i>Forficula crenata</i> Olivier, 1791	→	Synonyme subjectif (doxisonyme)
...		...
<i>Elaunon nainitalensis</i> Baijal & Singh, 1954	→	Synonyme subjectif (doxisonyme)
<i>Labiduria riparius</i> (Pallas, 1773)	→	Pb. accord grammatical (apographe)
<i>Labidura riparia</i> (Pallas, 1774)	→	Erreur dans le complexe nominal (année)
...		...



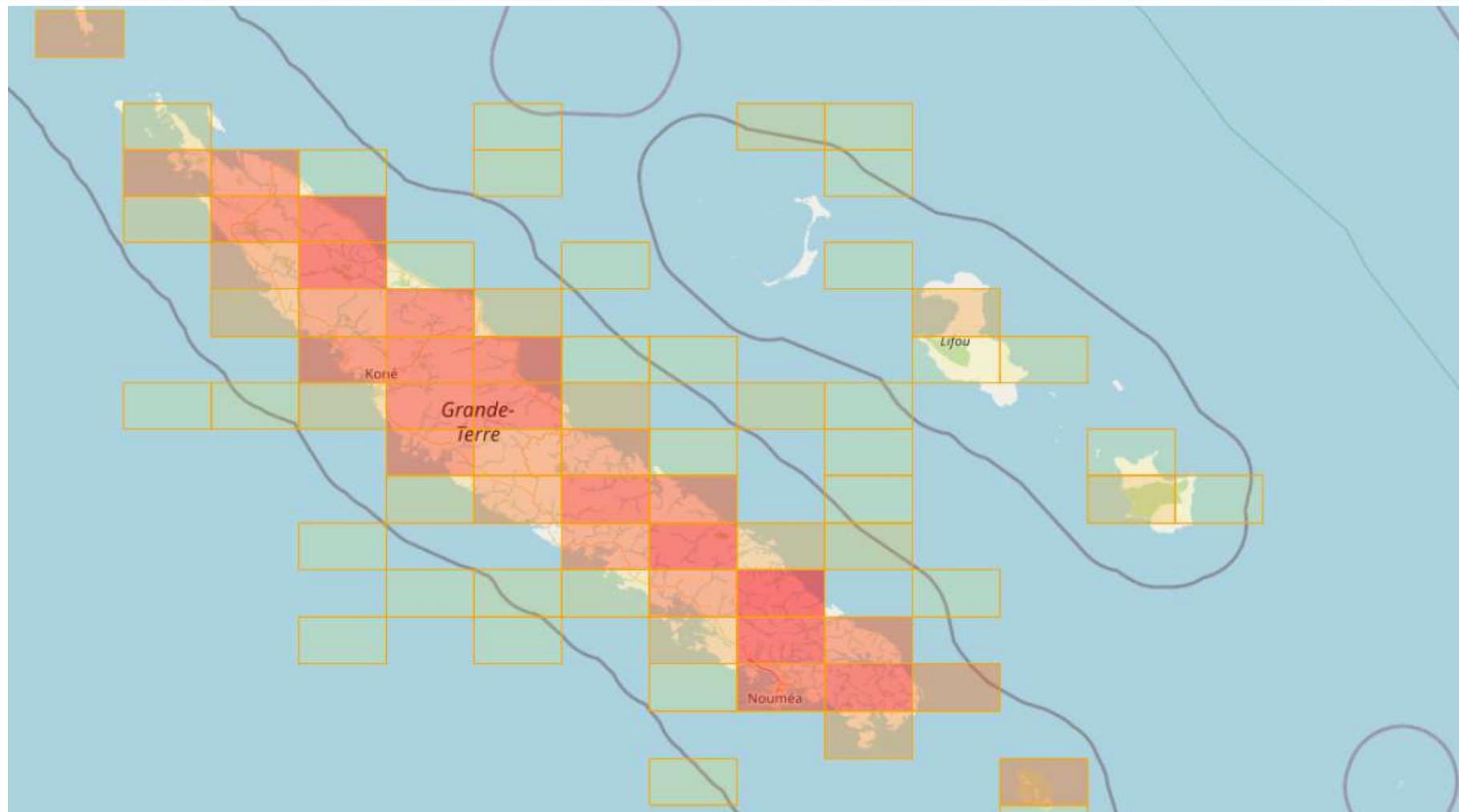
## Historique des attributions taxinomiques

***Parus caeruleus* Linnaeus, 1758**  
déterminé par BUISSON Alexandre en 1856

***Parus major* Linnaeus, 1758**  
déterminé par POULAIN Régis en 1898

***Cyanistes caeruleus* (Linnaeus, 1758)**  
déterminé par BLAVET Georgette le 13 mai 2021

## Recherche



[Recherche avancée \(2\)](#) [Carte](#)

collecte	
nie	
nie osaki, N.	
nie	
nie	

# Spécificités de la réglementation liées aux collections naturalistes

Statut juridique de l'objet

Préparation

Attribution taxinomique

Provenance du spécimen

Date de collecte

Attachement au fonds

Origine du spécimen

Exposition au public



## Récolnat porte le point nodal français de DiSSCo

170+  
institutions



1,5 milliard de  
spécimens



23 pays



Une collection européenne  
ouverte

**DiSSCO**  
Distributed System of Scientific Collections



Un catalogue de données  
exploitables constitué par les  
institutions membres de DiSSCo  
et des services associés



Disponibilité des données



Préparation scientifique

Préparation organisationnelle



Préparation financière



Préparation technique



## Perspectives européennes



Distributed System of Scientific Collections

**DiSSCover** Home All Specimens About Take a tour Login / Sign-up

Specimens

Quality/completeness indicators:  
MIDS Level:

Search filters:

- Topic Discipline:
- Taxonomy:
- Country:
- Organisation:
- Source System:
- Type status:
- Living or preserved:
- License:
- Has Media:

DOI	Specimen Name	Physical Specimen ID	Topic Discipline	Country of Origin	Date Collected
10.3535/ZTG-RDM-BNE	<i>Flaccisagitta hexaptera</i> (d'Orbigny, 1836)	<a href="https://data.biodiversitydata.nl/naturalis/specimen/ZMA.VER.20087">https://data.biodiversitydata.nl/naturalis/specimen/ZMA.VER.20087</a>	Zoology		5-VI-1970
10.3535/CWJ-S8C-T2R	<i>Heurste diversicolor</i> (O.F. Müller, 1778)	<a href="https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.808.43">https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.808.43</a>	Zoology	Netherlands	1835-05-20
10.3535/MBZ-VDN-EA2	<i>Labidocera tentaculata</i> (Horst, 1917)	<a href="https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.5281">https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.5281</a>	Zoology	Indonesia	1900-02-22
10.3535/QQP-C89-RKK	<i>Arenicola marina</i> (Linnaeus, 1758)	<a href="https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.92">https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.92</a>	Zoology	Netherlands	1915-07-04
10.3535/MSO-CQL-FHH	<i>Hydroides elegans</i> (Hatswell, 1883) (nomen protectum)	<a href="https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.5016">https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.5016</a>	Zoology	Netherlands Antilles	1907-10-14
10.3535/HGC-HVS-E75	<i>Heurste diversicolor</i> (O.F. Müller, 1778)	<a href="https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.808.41">https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.808.41</a>	Zoology	Netherlands	1934-06-13
10.3535/D83-9HG-RKN	<i>Aikmaea romijni</i> Horst, 1919	<a href="https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.5">https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.5</a>	Zoology	Netherlands	1924-10-07
10.3535/SE3-AWD-9WT	<i>Scolecops (Scoleopsis) squaricata</i> (O.F. Müller, 1806)	<a href="https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.988">https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.988</a>	Zoology	Netherlands	1915-07-03
10.3535/STY-3P9-D46	<i>Pseudovermilia</i> Bush, 1907	<a href="https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.4771">https://data.biodiversitydata.nl/naturalis/specimen/ZMA.V.POL.4771</a>	Zoology	Indonesia	1984-09-26

Specimen found: 50251

1 2 ... 399 >



## Nos 4 challenges du moment

1. Mener à bien nos chantiers en cours : import de données en masse, description des échantillons géologiques, requêtage spatial fin des données et intégration des modèles 3D ;
2. Réduire la *dette* de description des collections naturalistes et augmenter la qualité du corpus partagé dans une perspective FAIR et de science ouverte ;
3. Passer à l'échelle et accompagner la conduite du changement, notamment en vue de la mise en œuvre du concept de spécimen étendu ;
4. Continuer l'engagement des adhérents vers une nouvelle structuration pour 2026-2030 et organiser notre participation/articulation avec l'organisation européenne DiSSCo.





Mais n'oubliez pas que les disques durs sont plus périssables que les herbiers.

**Merci pour votre attention**

[francois.dusoulier@mnhn.fr](mailto:francois.dusoulier@mnhn.fr)