

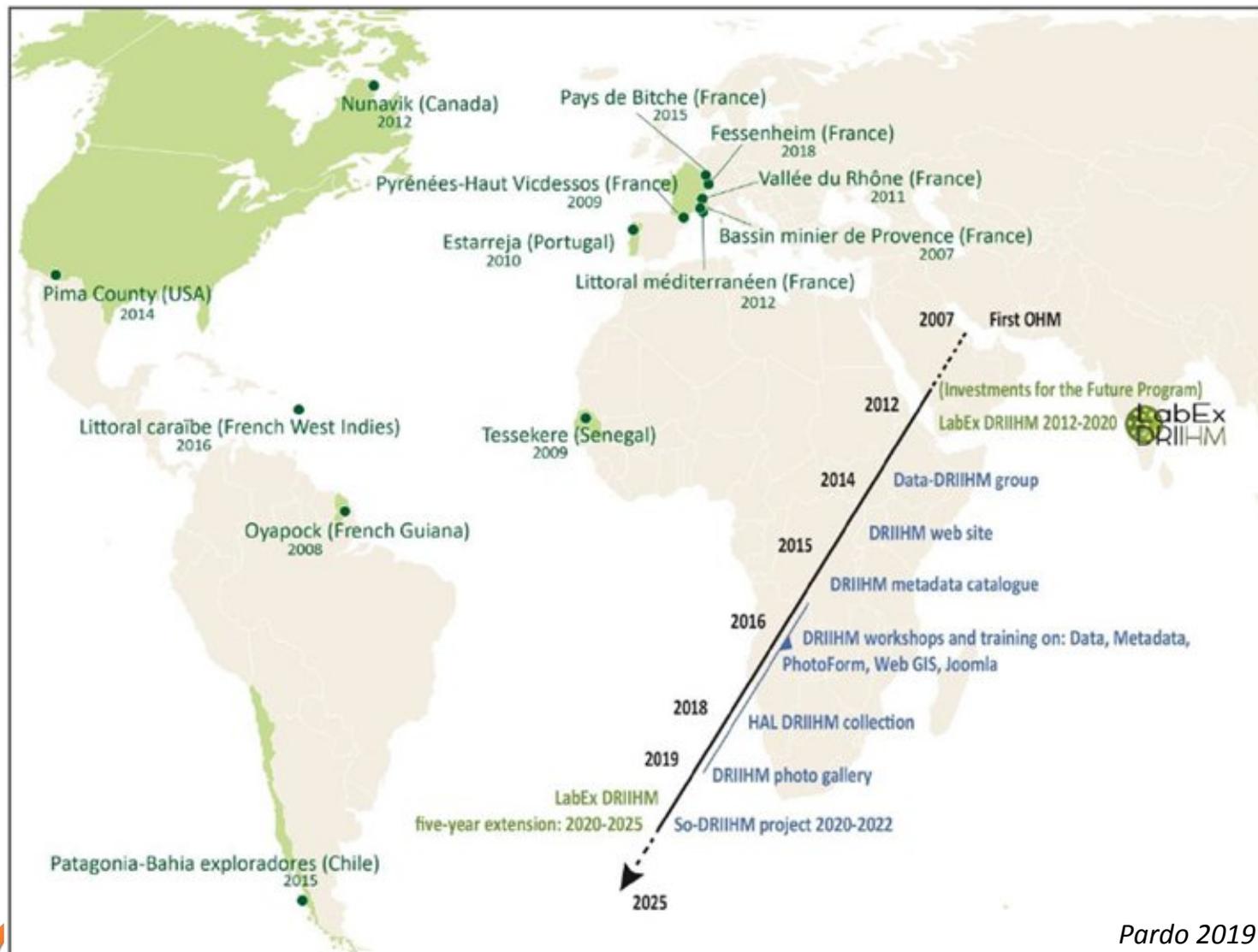
Premiers pas vers le partage et l'ouverture des données des Observatoires Hommes-Milieus : un nécessaire accompagnement pour lever les freins psychologiques et les verrous techniques

Émilie Lerigoleur, Kristell Michel, Corinne Pardo, Marie-Laure Trémélo



1/ CONTEXTE

Le LabEx DRIIHM et ses 13 Observatoires Hommes-Milieus (OHM)



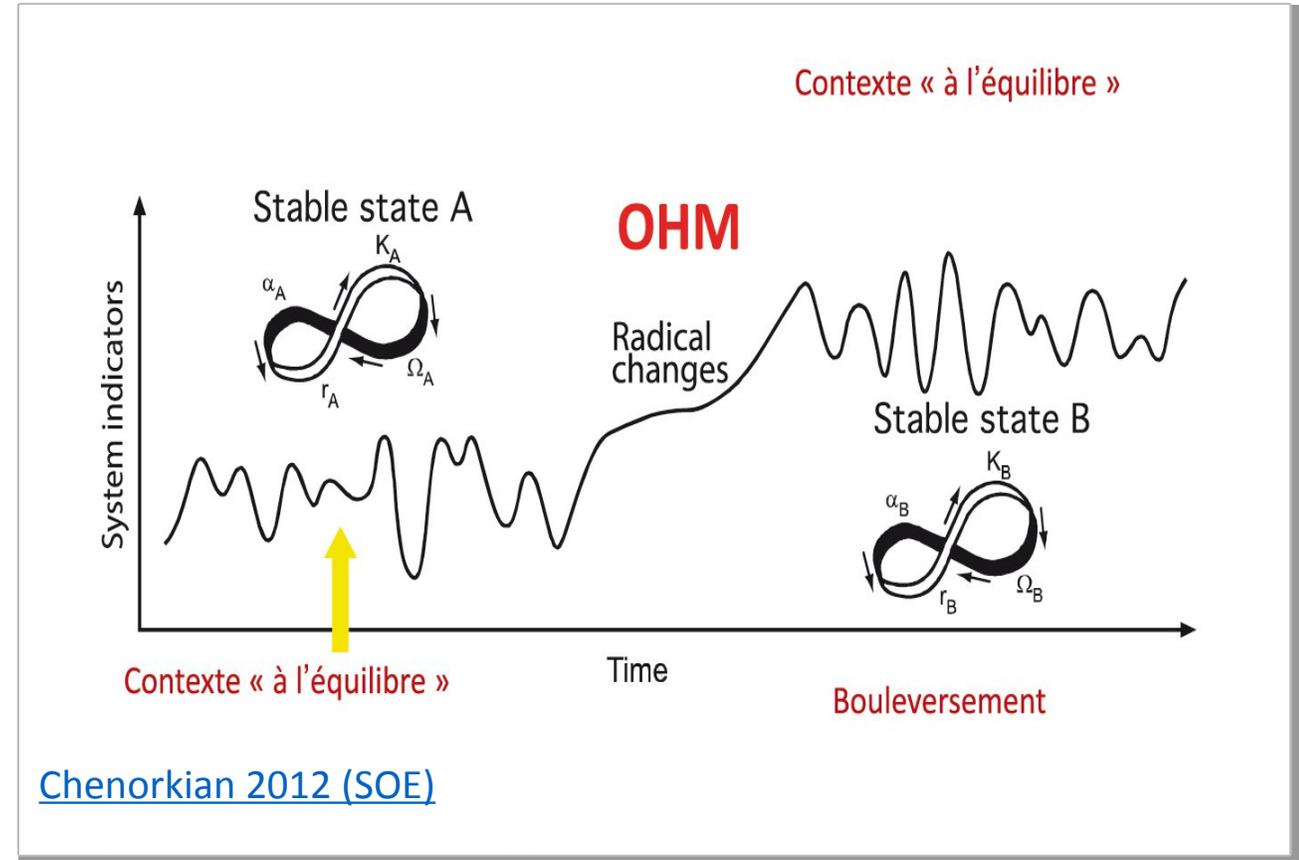
Pardo 2019

- Dispositif CNRS-INEE depuis 2007, financé par le LabEx depuis 2012
 - 13 OHM
 - ~ 1250 pers. / 200 labos / 133 EPST-univ.-grandes écoles / 27 partenaires territoriaux ou associatifs
 - 866 projets annuels/thèses/post-doc
 - 73 champs disciplinaires
 - données très hétérogènes de “la longue traîne”
 - Personnel : 1 Resp. ST + 1 IR gouv. + 1 IE (CDD) informaticien + groupe data-DRIIHM (~ 15 pers. support)
- Une infrastructure de données construite brique par brique géocatalogues, web SIG, photothèque

Fonctionnement d'un observatoire OHM

- 1^{er} OHM créé en 2007, dernier en 2018
- Chaque OHM s'organise autour d'un **objet focal** qu'un **événement fondateur** est venu bouleverser profondément
- Exemples :

Objet focal	Événement fondateur
Bassin minier	arrêt exploitation
Désertification Sahel	Grande Muraille Verte
Fleuve Rhône	crues et Plan Rhône
Vallée pyrénéenne	fermeture usine





1920

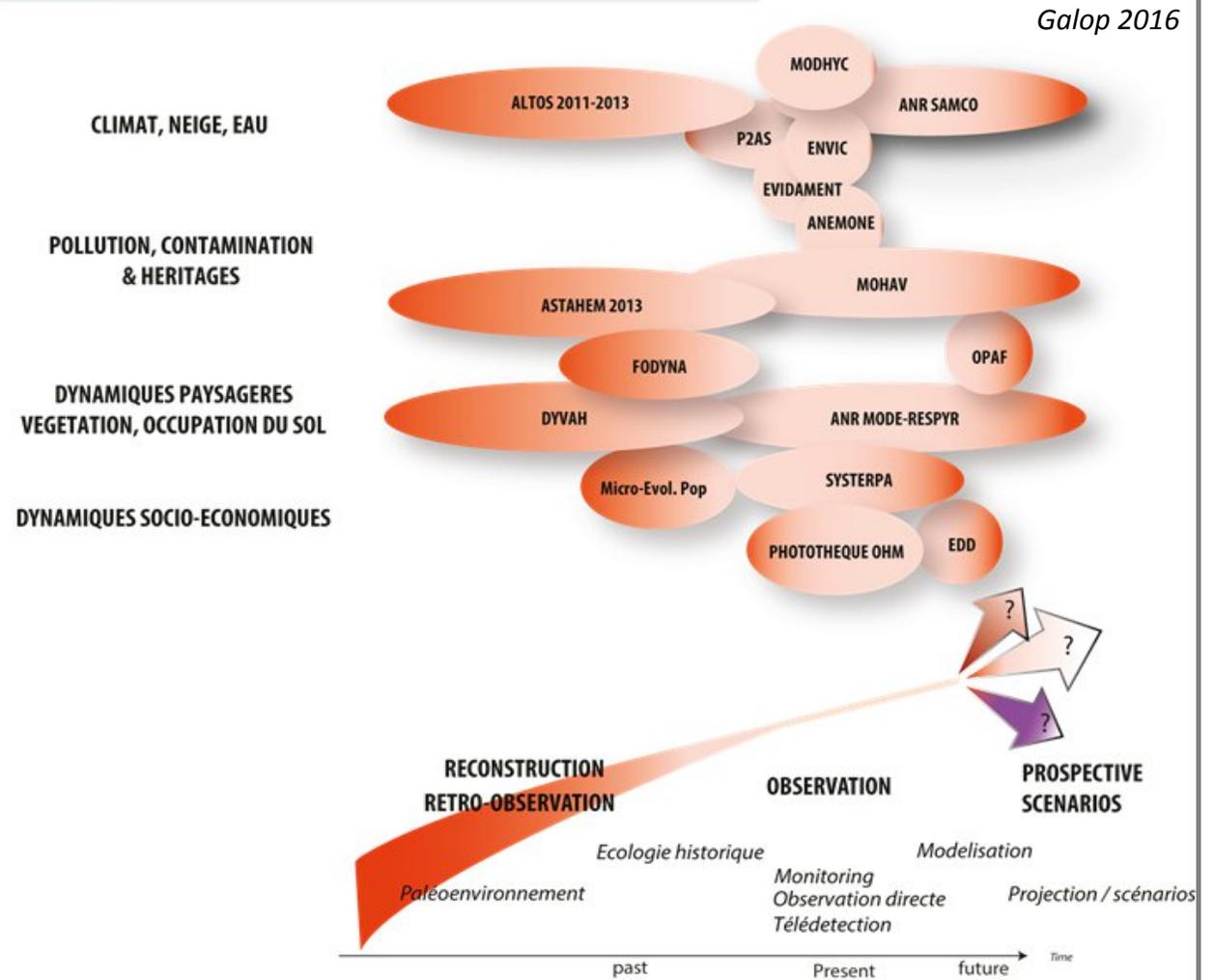
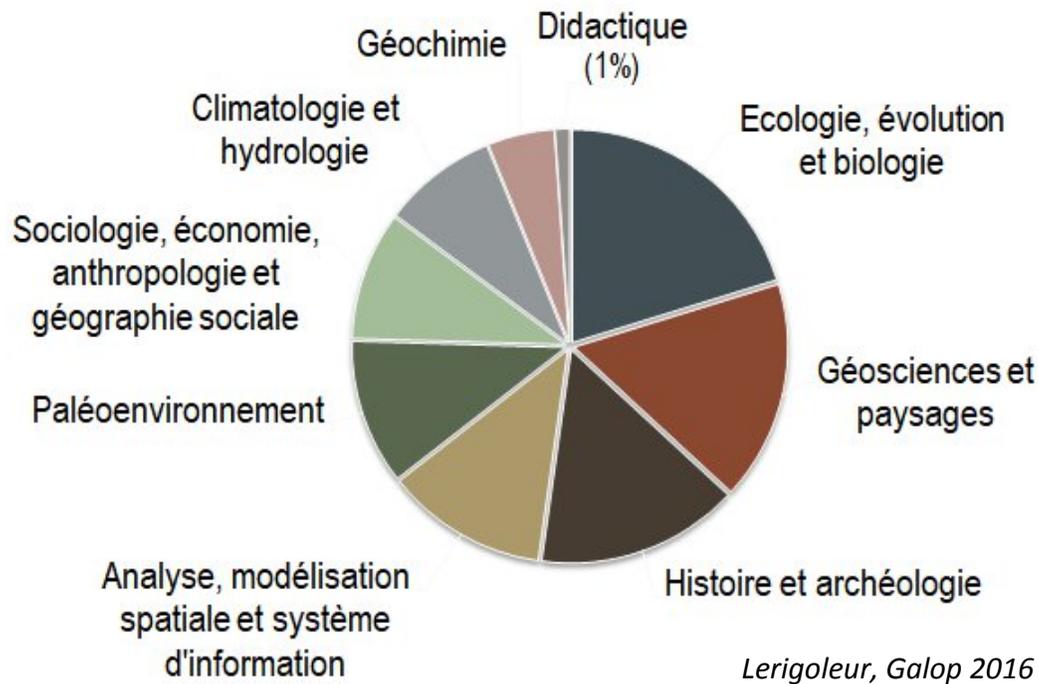


2012

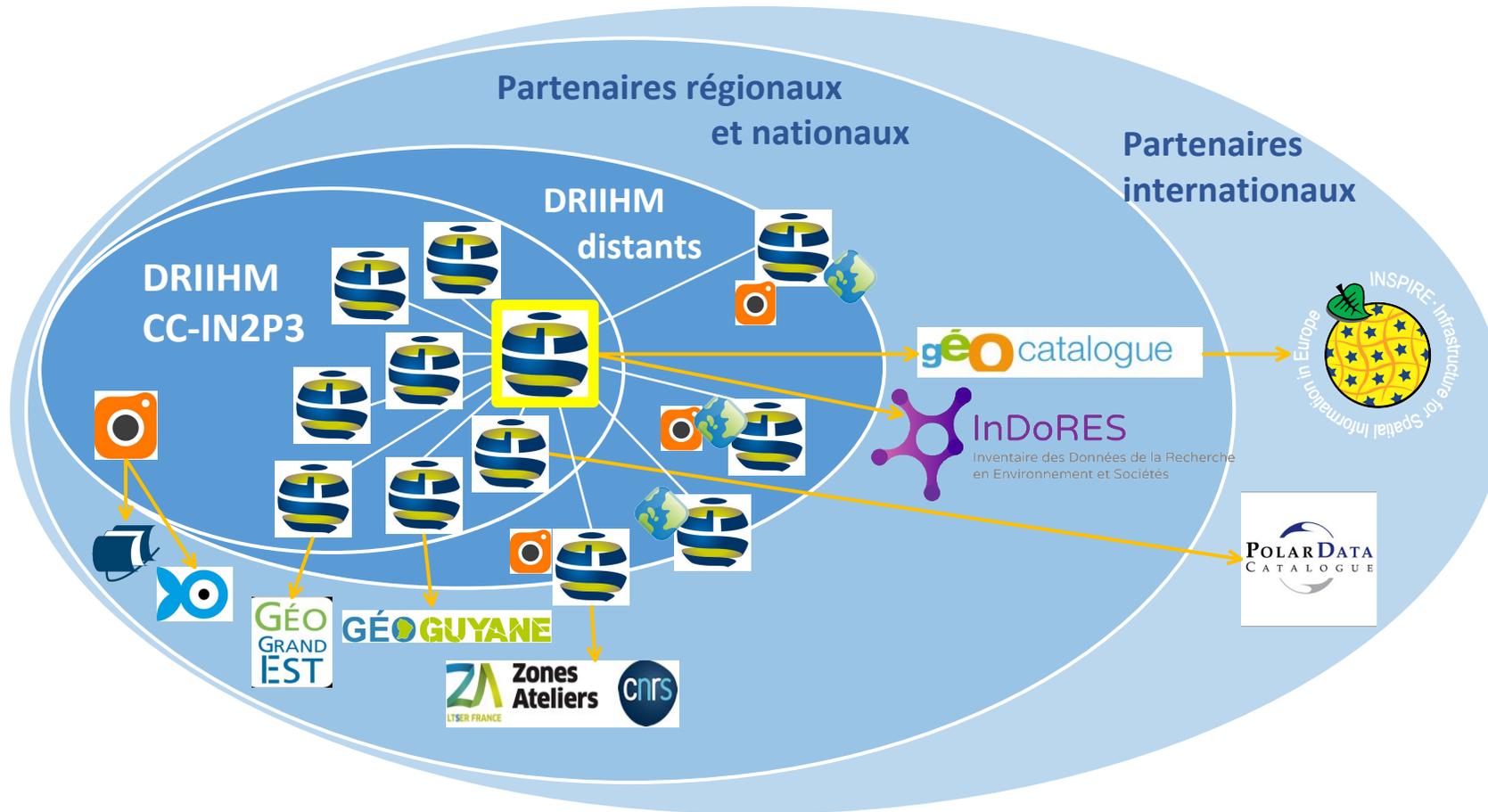
Enjeux de l'interdisciplinarité

Exemple de 18 projets financés pour l'OHM Pyrénées Haut Vicdessos

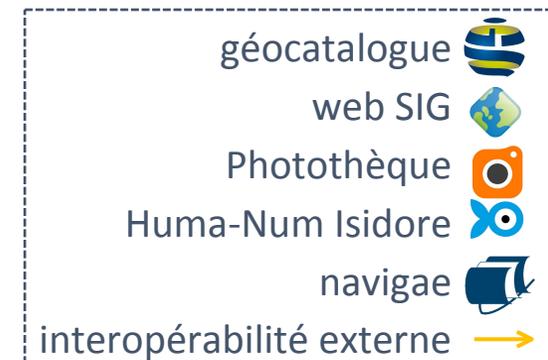
→ nombreuses disciplines et thématiques
→ 3 échelles temporelles



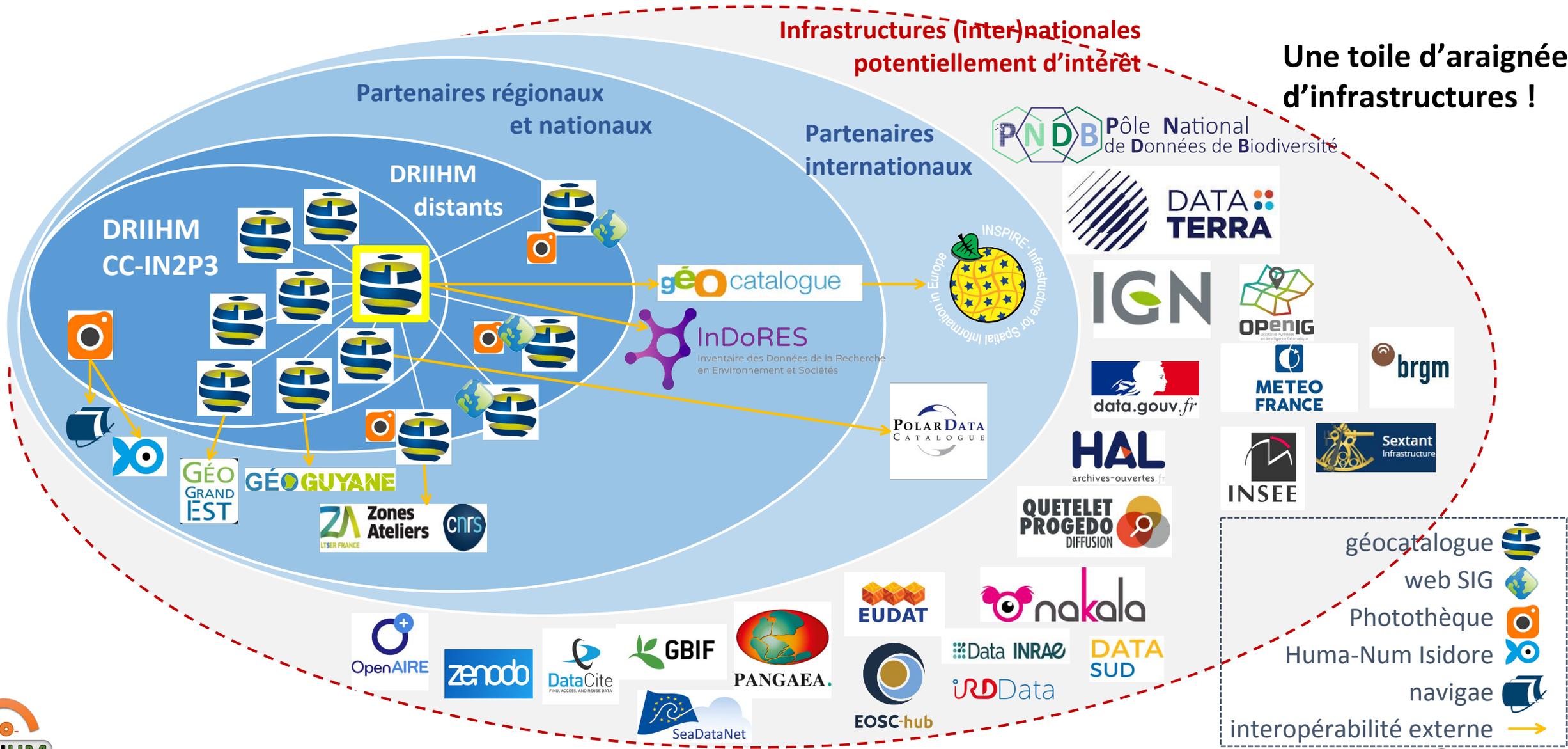
L'infrastructure distribuée existante et ses connexions potentielles



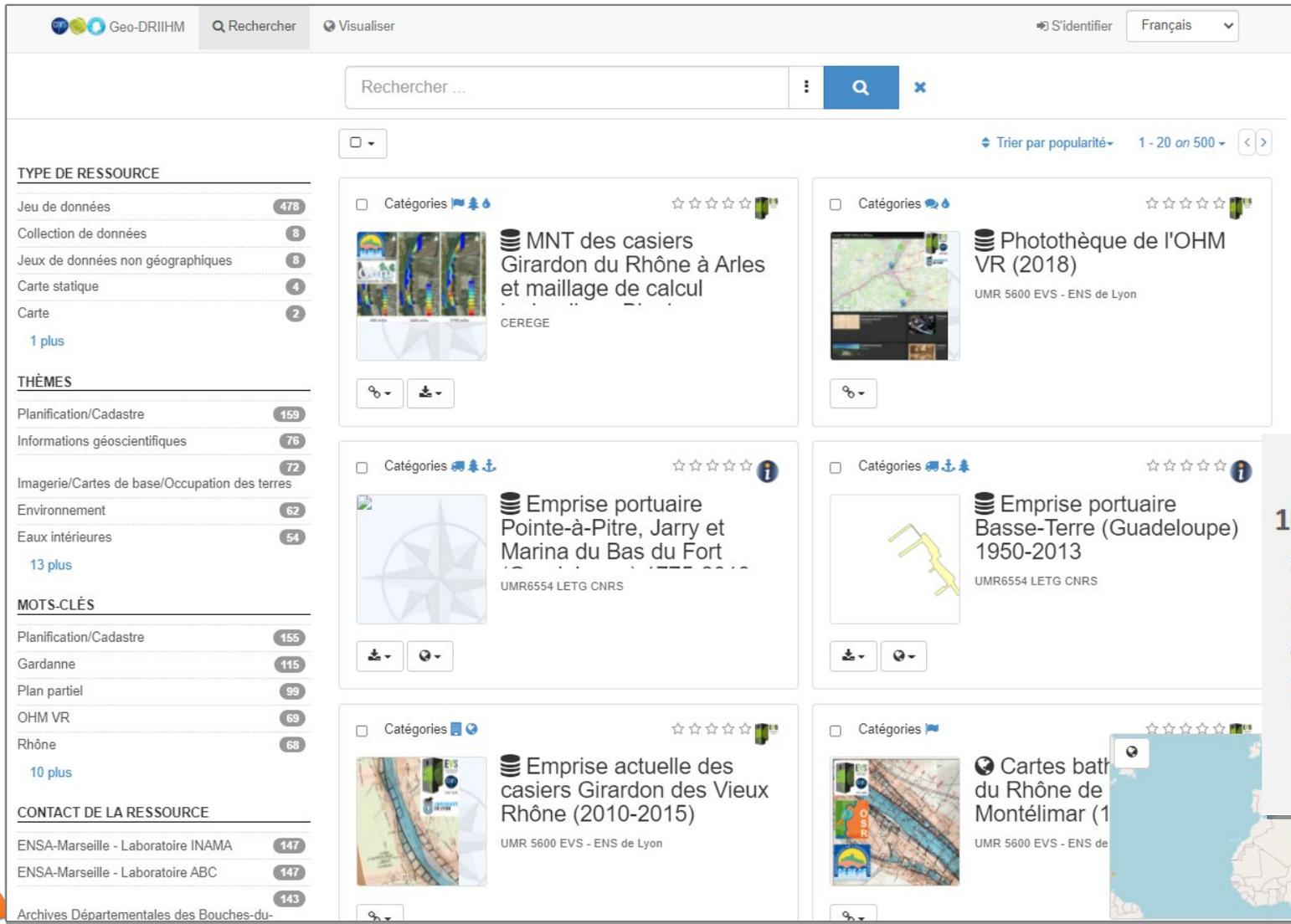
- Des géocatalogues par OHM connectés à un géocatalogue DRIIHM
- Une photothèque mutualisée interopérable
- Quelques connexions externes



L'infrastructure distribuée existante et ses connexions potentielles

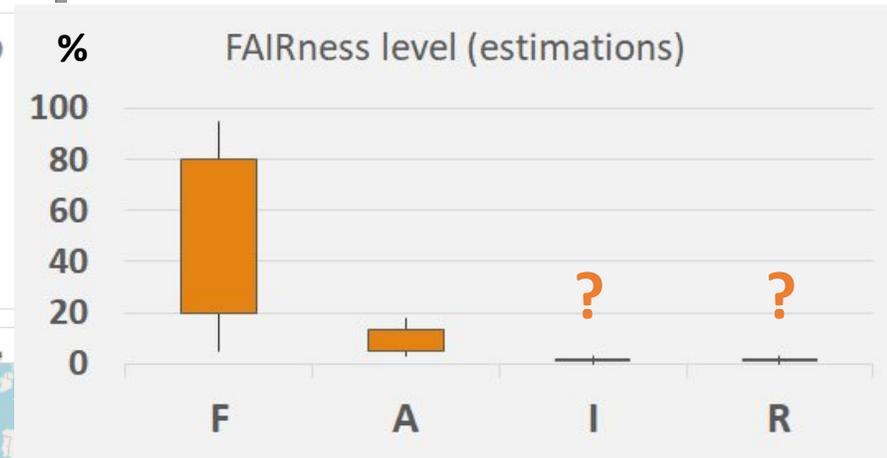


Constat : des outils de catalogage disponibles mais peu utilisés



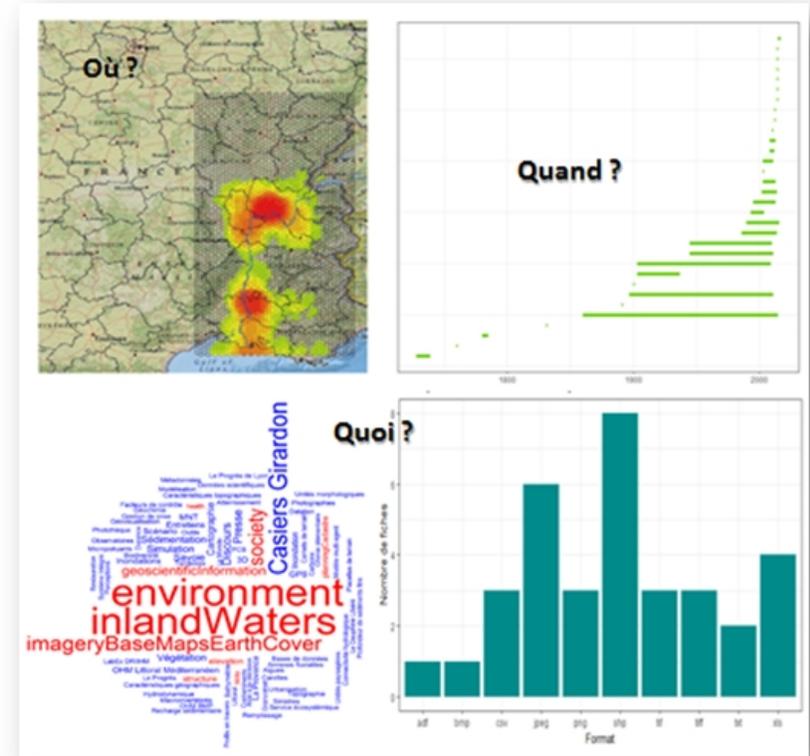
Outil de catalogage existant +
 mais quelques limites -

- Saisie des métadonnées coûteuse et peu ergonomique
- Pas de bénéfice direct dans la démarche
- Difficile d'évaluer l'interopérabilité (I) et la réutilisation (R) des données cataloguées avec ces outils



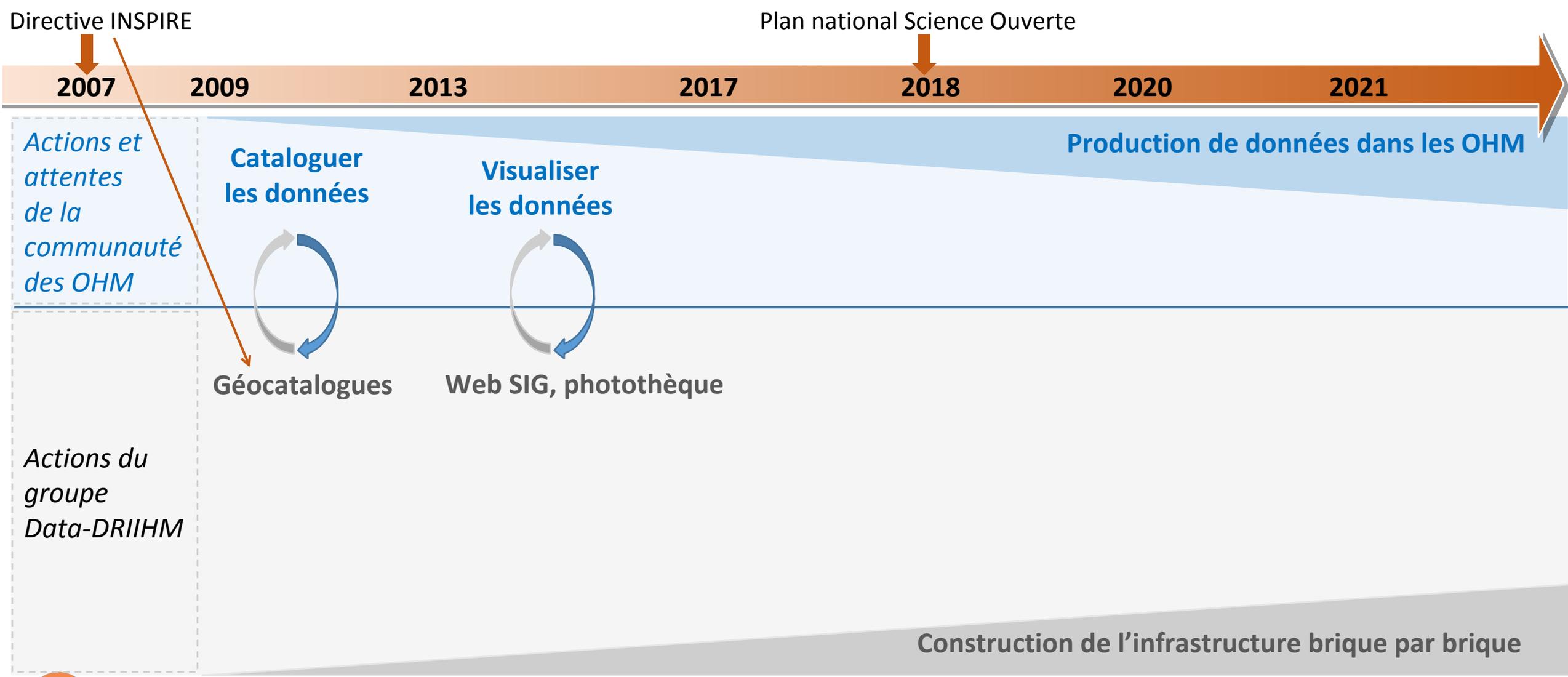
Estimation du degré de FAIRness dans le catalogue [GéoDRIIHM](#)

Constat : des métadonnées disponibles, mais peu de données accessibles. Des usages qui, par ailleurs, se diversifient

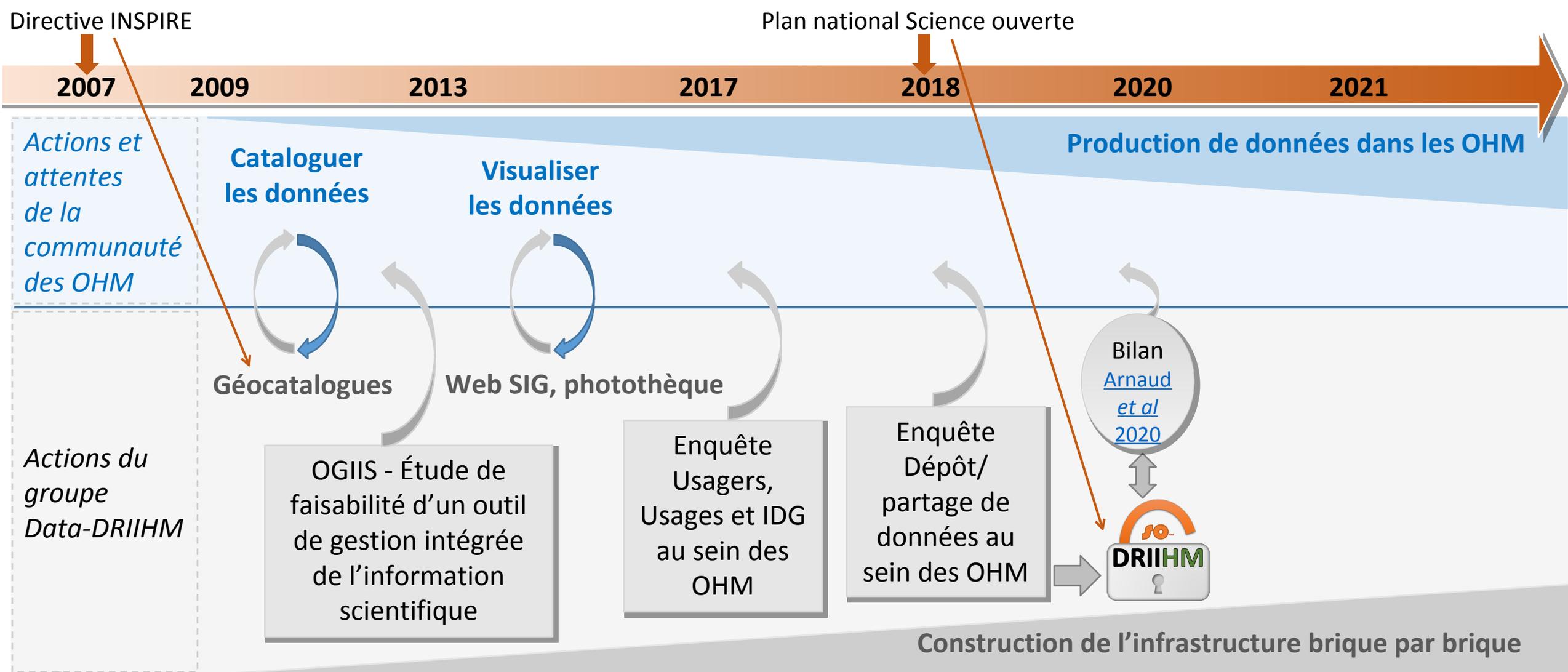


Source : OHM Vallée du Rhône, 2017

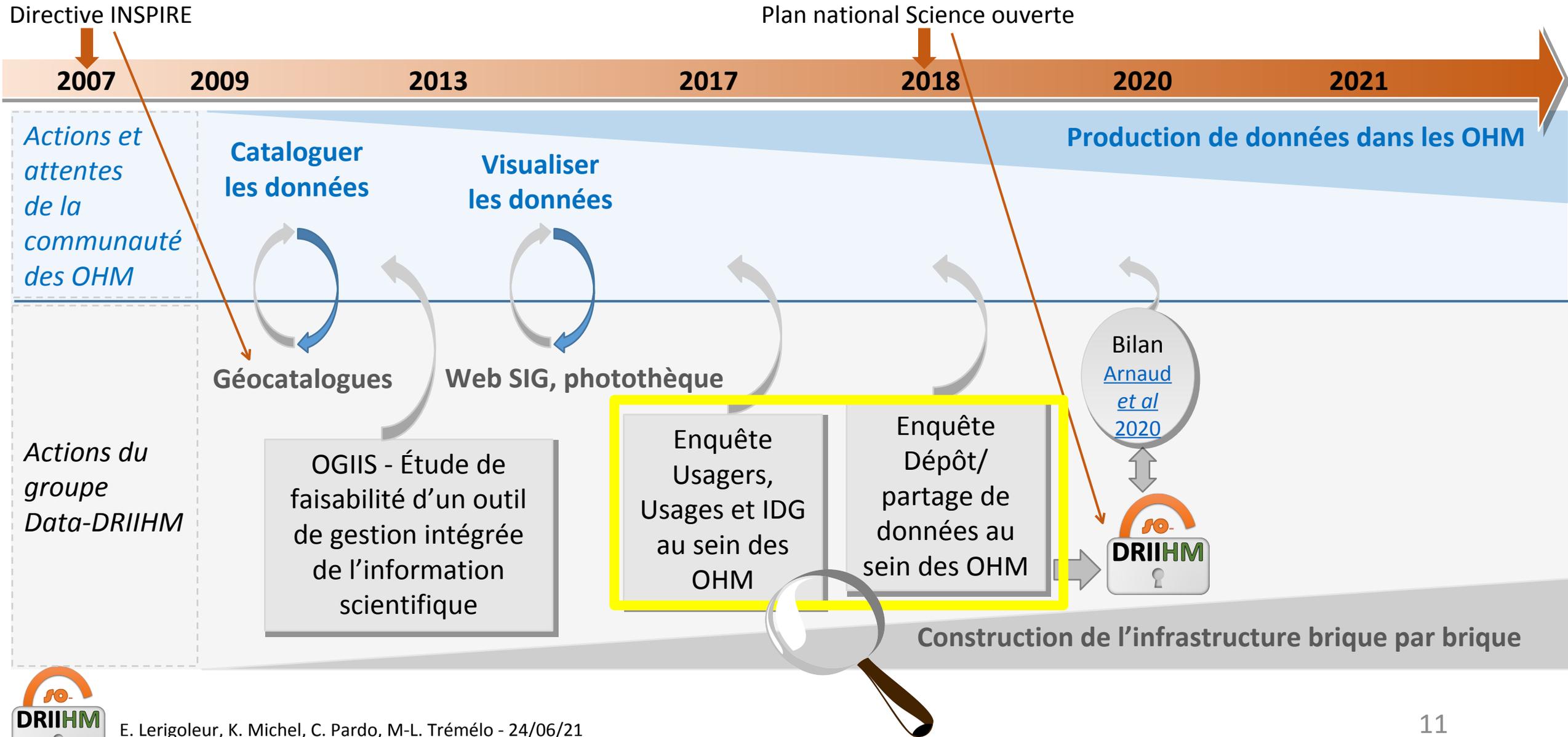
Le groupe Data-DRIIHM en interaction avec la communauté DRIIHM



Le groupe Data-DRIIHM en interaction avec la communauté DRIIHM



Le groupe Data-DRIIHM en interaction avec la communauté DRIIHM



2/ PRINCIPAUX RÉSULTATS DES 2 ENQUÊTES

Enquête 2017

Usagers, Usages et IDG* au sein des Observatoires Hommes-Milieux

Qui sont les usagers ? leurs relations à la données ? leurs attentes ? points de blocage et/ou leviers ?

19 questions (16 QF, 3 QO)

Enquête 2018

Analyse du degré d'acceptabilité de la création d'une plateforme de dépôt et de partage des jeux de données des OHM

Quelles pratiques de dépôt/partage ? intérêt au partage ? limites ? avantages ?

Open access : [10.34847/nkl.3f777nrt](https://doi.org/10.34847/nkl.3f777nrt) 
19 questions (12 QF, 7 QO)

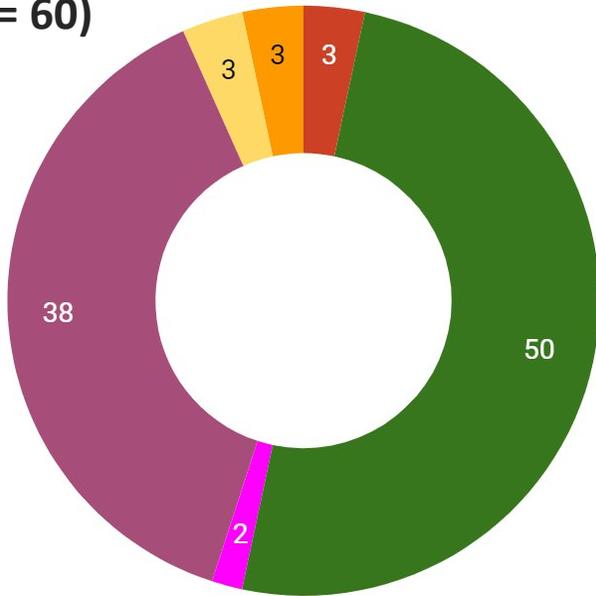
- ✓ Questionnaires en ligne - anonymes
- ✓ Même cible : communauté DRIIHM (producteurs et utilisateurs des données des OHM)

* IDG : Infrastructure de Données Géographiques

Les répondants

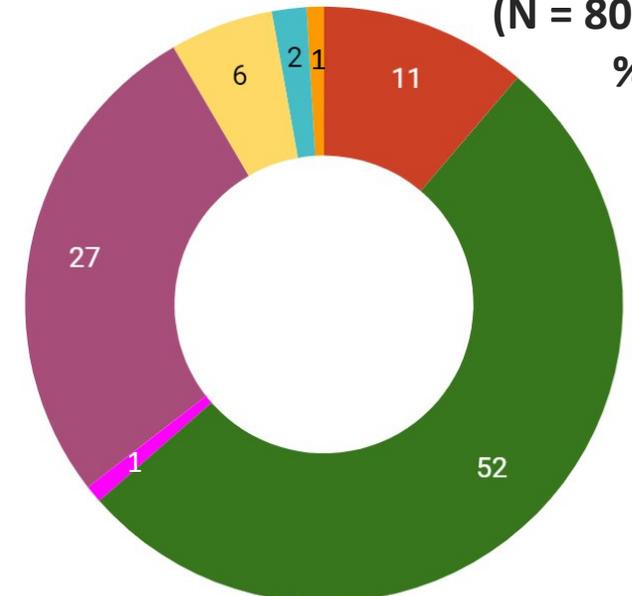
Leurs champs thématiques privilégiés :

2017
(N = 60)
%

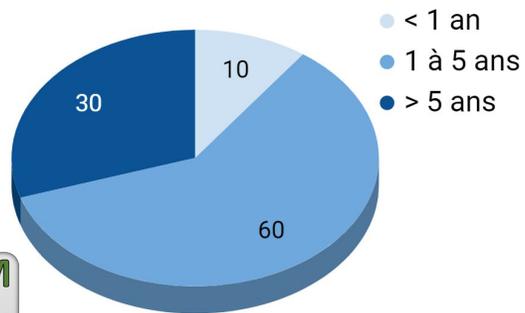


- Sciences de la Terre, de l'Univers et de l'Espace
- Sciences de l'environnement, Sciences Agronomiques et Ecologiques
- Sciences et technologies de l'information et de la communication
- Sciences Humaines et Sociales
- Sciences médicales et sanitaires
- Sciences pour l'ingénieur et technologiques
- Education développement durable

2018
(N = 80)
%

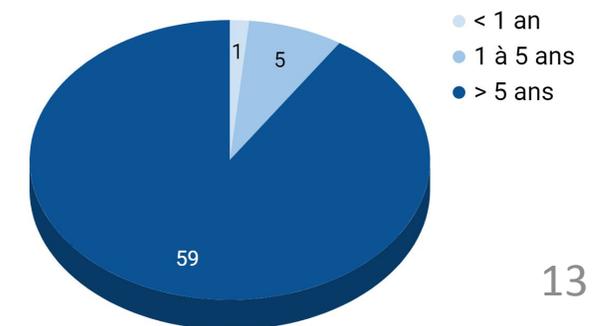


Années d'expérience dans OHM (N = 60) en %



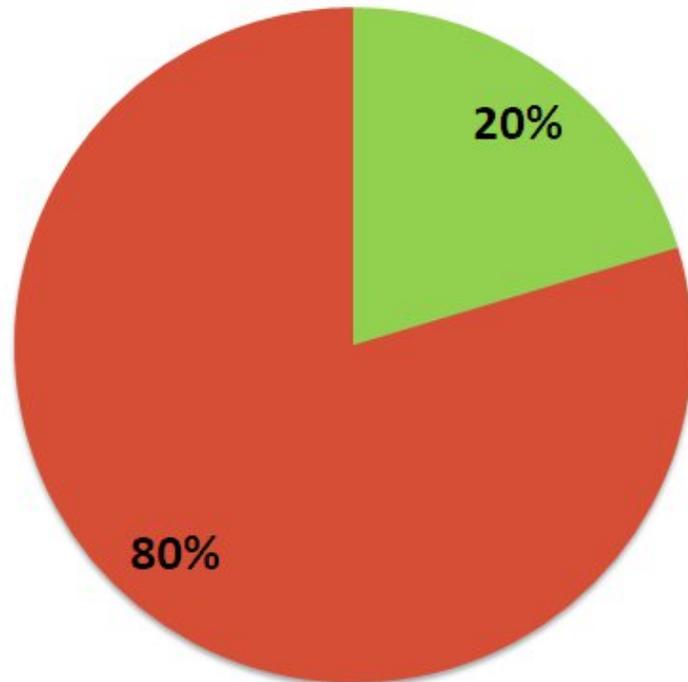
- ✓ Représentants de chaque OHM sauf 3
- ✓ 87% universitaires
- ✓ 86% vit en France métropolitaine
- ✓ 23% ont une pratique régulière de la géomatique 40% peu intensive, 37% ne la pratiquent pas
- ✓ 60% ont plus de 45 ans

Années d'expérience (N = 65) en %

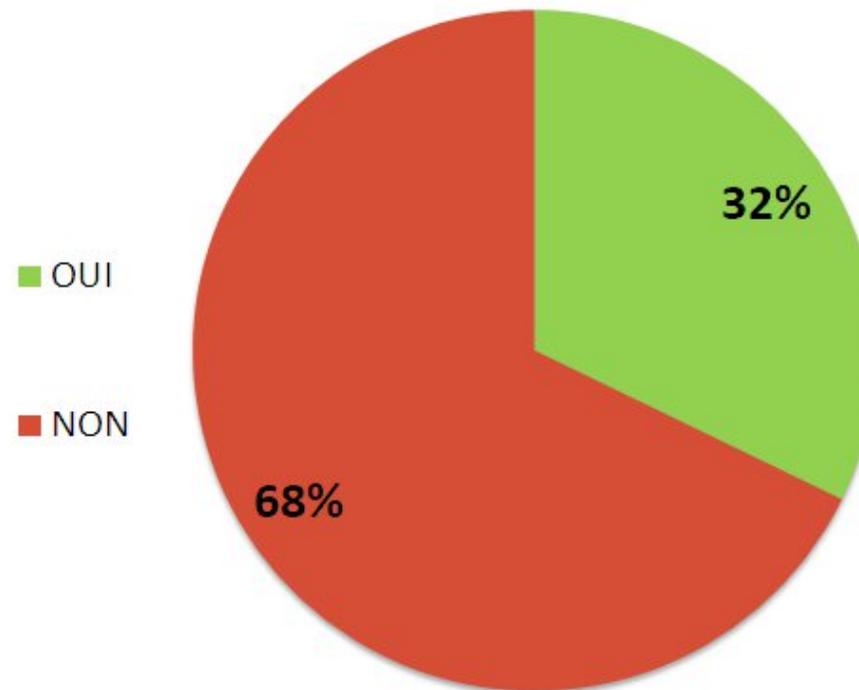


La notion d'IDG non encore intégrée et des IDG sous-utilisées

Connaissez-vous une IDG mise en place par votre OHM ou le LabEx DRIIHM ?



Utilisez-vous, même peu souvent, une IDG dans le cadre de vos activités professionnelles ?



Enquête 2017

(N = 60)

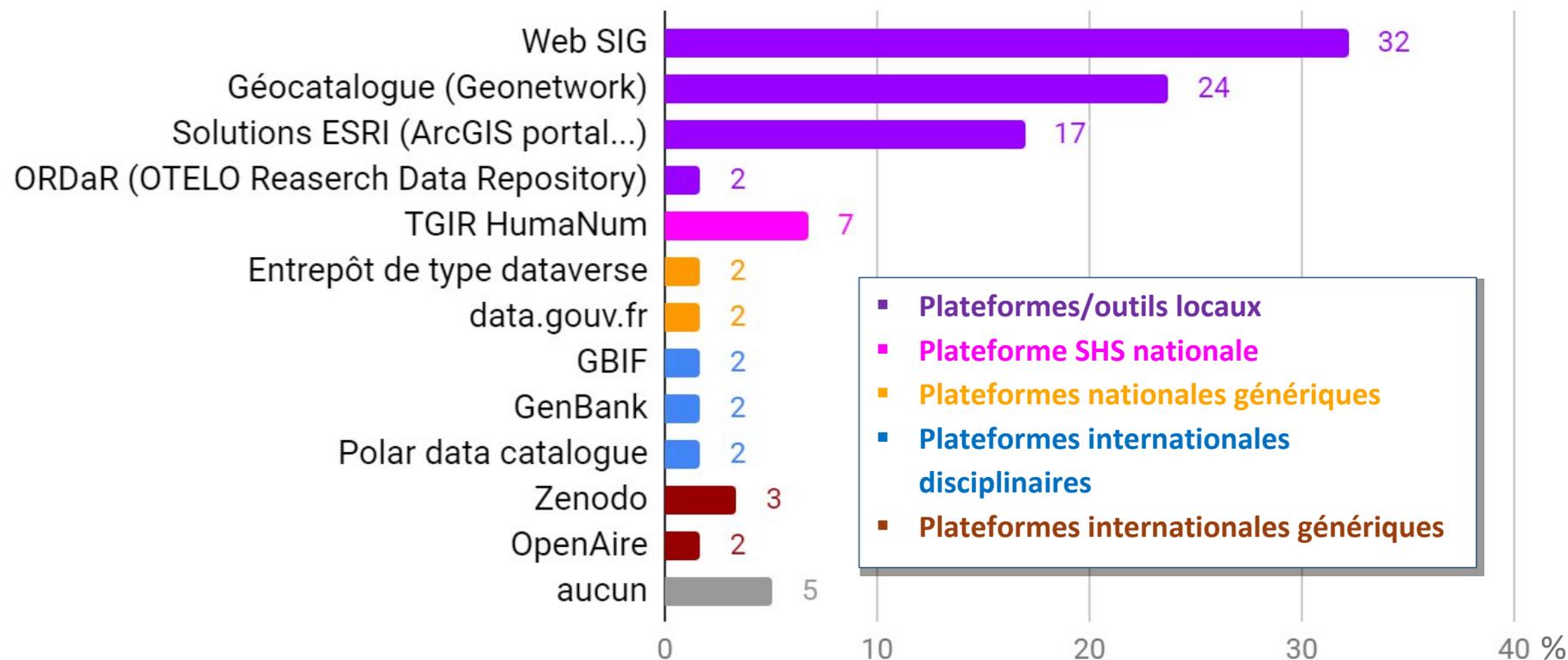
Rappel (donné dans le questionnaire) : une IDG (Infrastructure de Données Géographiques) se compose de "l'ensemble des informations, des systèmes informatiques, des normes et standards, des accords organisationnels, des ressources humaines et des communautés nécessaires pour faciliter et coordonner l'accès et le partage des informations géographiques" (Rajabifard et al. 2002)

Peu de plateformes de dépôt internationales/nationales déjà utilisées par les répondants

[Question à choix multiple + autre(s)]

Enquête 2018

Les plateformes de dépôt déjà utilisées (N = 53)



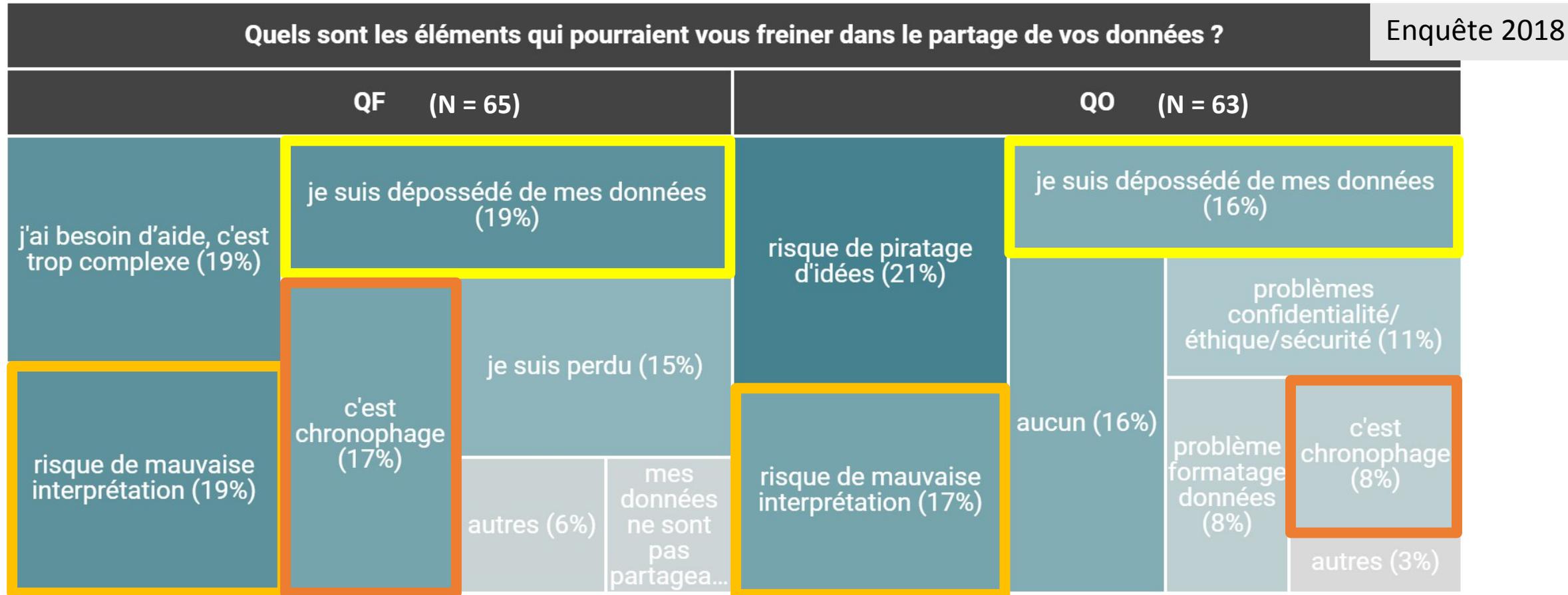
Utilisation essentiellement des outils « locaux » plutôt intra-OHM

Principales raisons :

- complexes
- peu ergonomiques
- peu connues
- pas le temps !

« Rapport coût/bénéfice : temps passé important pour voir des données jamais utilisées par d'autres car trop spécifiques à une recherche donnée »

Les freins : risque de mésusage, perte de temps, ne pas savoir comment faire ou avoir peur de mal le faire, données non partageables...

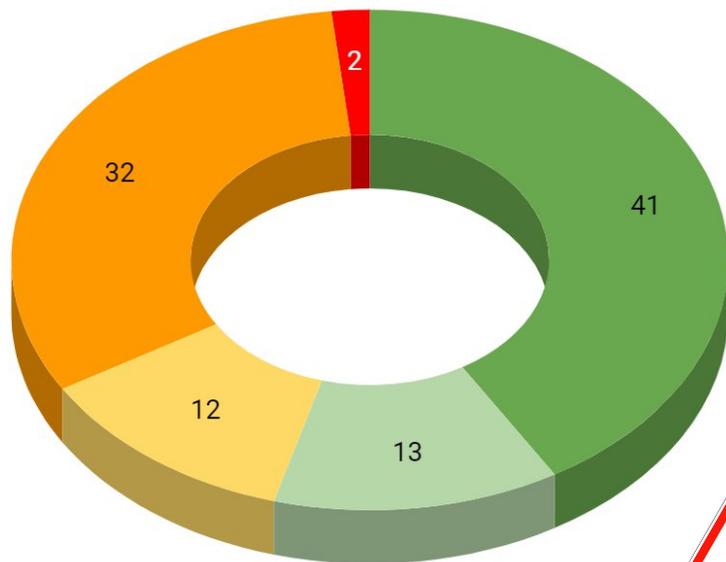


« Diffusion trop précoce, risque de dévoiler des idées mais tout dépend du périmètre de la diffusion »

De bonnes intentions quant au partage... à consolider

[Question fermée]

2017
(N = 60)
%



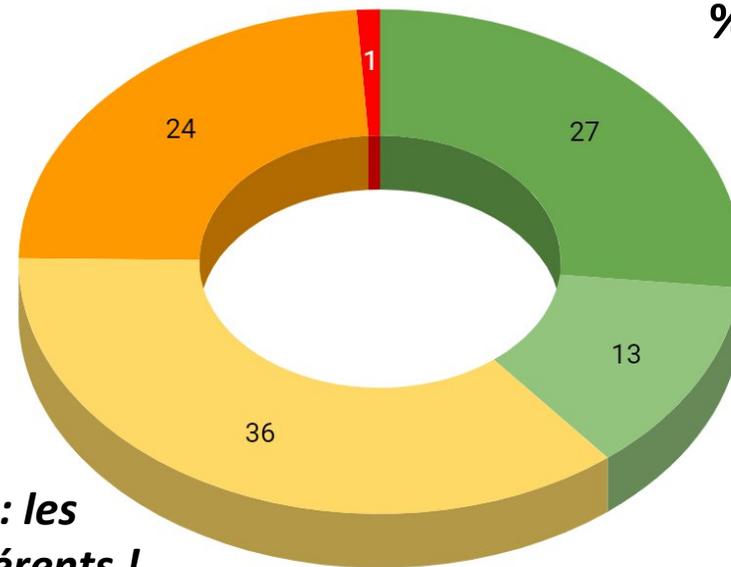
Quel accès à vos données ?

- libre accès
- accès restreint
- accès sur demande
- non accessibles avant publication
- pas d'accès



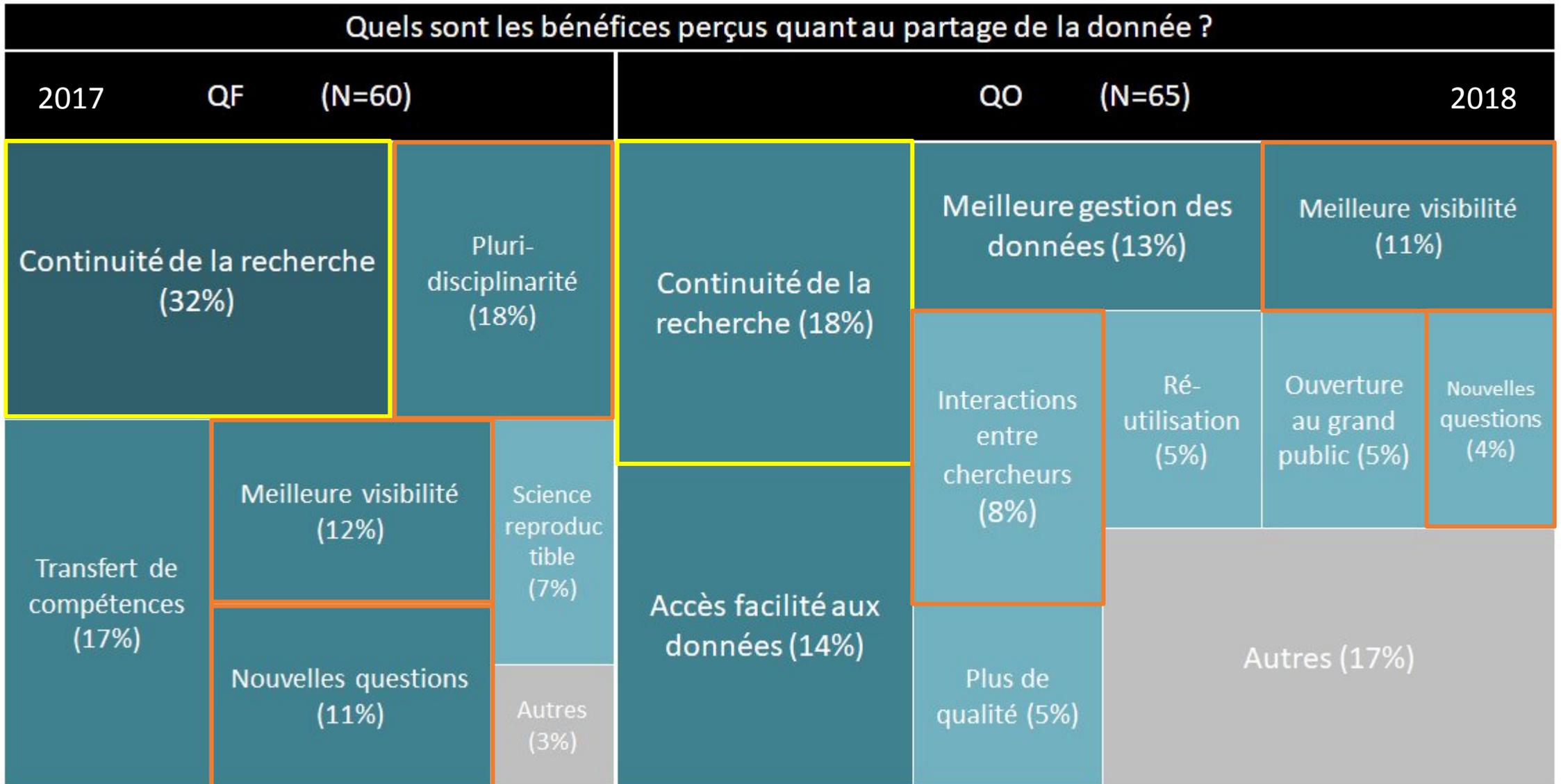
Pas de comparaison possible : les répondants peuvent être différents !

2018
(N = 80)
%



*« données publiées accessibles à tou.te.s ;
données en cours de valorisation: accessibles sur demande pour favoriser les partenariats et l'émergence de questionnement;
de par leurs natures, certaines données ne peuvent être rendues publiques (...) mais les métadonnées peuvent l'être »*

Les bénéfices perçus du partage de la donnée



Les leviers d'action pour une IDG opérante

FT : frein technique
 FP : frein psychologique

Enquête 2017

Quels sont les leviers d'action pour une IDG opérante ?				
QFCM (N=60)				
Formats de diffusion des résultats (20) FT	Formation et information (23) FT	Pérennité de l'infrastructure (27) FT		Spécifications des règles de diffusion (29) FP
Information détaillée (25) FP	Automatisation (18) FT	Aspects juridiques et réglementaires (19) FT	Ressources financières (10) FT+FP	Adhésion et appropriation (27) FT+FP
Ressources humaines (20) FT+FP		Formes d'organisation (10) FT+FP		

Actualisation des leviers d'action

- **Lever les verrous techniques**

- ✓ utiliser un vocabulaire simple, non technique
- ✓ améliorer l'ergonomie des plateformes
- ✓ *simplifier les tâches*
- ✓ guider pas à pas, expliquer, rassurer
- ✓ « humaniser les interfaces »
- ✓ *fournir des fonctionnalités utiles*
- ✓ *fournir des garanties : sécurité/pérennité/contrôle diffusion*
- ✓ *informer sur l'existence des plateformes et former à leur utilisation !*

- **Lever les freins psychologiques**

- ✓ accepter que toute donnée ne soit pas « partageable »
- ✓ *former aux « bonnes pratiques » de partage de données, aux principes « FAIR »*
- ✓ *informer sur les questions de propriété/éthique/confidentialité*
- ✓ *rassurer sur le maintien du contrôle de la diffusion de ses données (licence/droits d'accès)*
- ✓ *s'assurer d'une politique de diffusion des données transparente/claire*
- ✓ questionner régulièrement la communauté pour s'assurer de l'appropriation des outils

en bleu : freins et verrous identifiés lors des enquêtes

3/ CONCLUSION-PERSPECTIVES

- **1er bilan des enquêtes : un pas vers le partage et l'ouverture des données « oui, mais »**
 - Une part importante de scientifiques motivés par le partage de leurs données, mais nécessaire accompagnement
 - Des données d'observations pluridisciplinaires d'intérêt au partage, mais aussi des données « non partageables »
 - Des limites dans l'interprétation des données d'enquêtes :
 - biais proportion (~10%) et représentativité des répondants
 - choix des modes de représentation, regroupements...
 - non recoupement entre disciplines et motivations/freins au partage => effet de la discipline sur les retours, analyses par sous-groupes

➔ Hypothèse : volonté de partage/ouverture = $f(\text{discipline, type de données, organisation réseau disciplinaire})$

Étude sociologique dans contexte particulier
pour question scientifique spécifique
= **données « peu partageables »**

facilité de partage

Collecte de données météo. sur
plusieurs stations en France
= **données quanti. corrigées ouvertes**

3/ CONCLUSION-PERSPECTIVES

- **1er bilan des enquêtes : un pas vers le partage et l'ouverture des données « oui, mais »**
 - Une part importante de scientifiques motivés par le partage de leurs données, mais nécessaire accompagnement
 - Des données d'observations pluridisciplinaires d'intérêt au partage, mais aussi des données « non partageables »
 - Des limites dans l'interprétation des données d'enquêtes :
 - biais proportion (~10%) et représentativité des répondants
 - choix des modes de représentation, regroupements...
 - non recoupement entre disciplines et motivations/freins au partage => effet de la discipline sur les retours, analyses par sous-groupes

➔ Hypothèse : volonté de partage/ouverture = $f(\text{discipline, type de données, organisation réseau disciplinaire})$

Étude sociologique dans contexte particulier
pour question scientifique spécifique
= **données « peu partageables »**

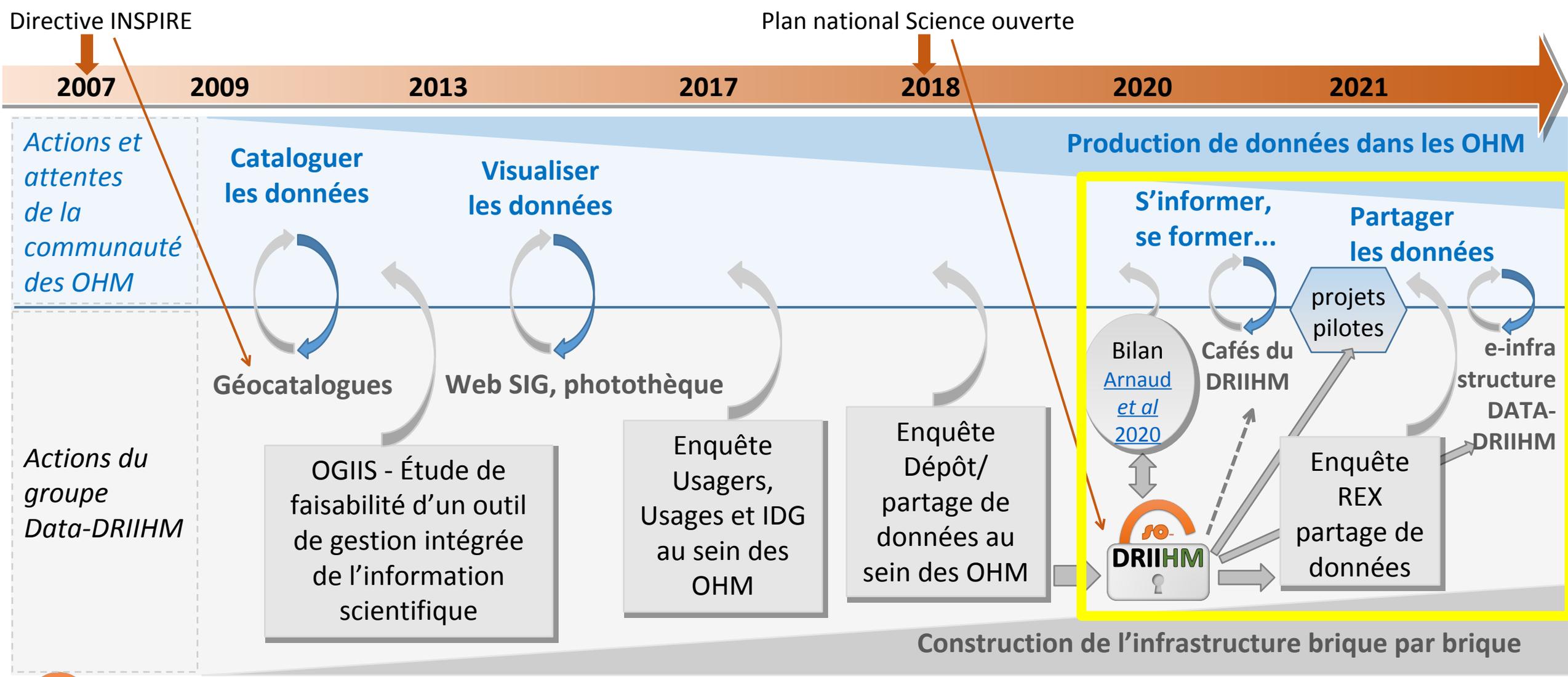
facilité de partage

Collecte de données météo. sur
plusieurs stations en France
= **données quanti. corrigées ouvertes**

- LabEx DRIIHM = pluri-/interdisciplinarité + international + RH techn. distribuées faibles (~1.2%) + € limités
- Data-DRIIHM = relais feuilles de route « science ouverte » + relais besoins communauté des OHM

➔ **ANR SO-DRIIHM (2020-2023)** : Impulser la science ouverte au sein du Dispositif de Recherche Interdisciplinaire sur les Interactions Hommes-Milieus (DRIIHM) : co-design d'une e-infrastructure intégrant les principes FAIR

Le groupe Data-DRIIHM en interaction avec la communauté DRIIHM



Merci !



<http://www.driihm.fr/les-ohms>



emilie.lerigoleur@univ-tlse2.fr

ORCID: [0000-0002-0864-659X](https://orcid.org/0000-0002-0864-659X)

[CV HAL](#)

kristell.michel@ens-lyon.fr

[CV](#)

corinne.pardo@univ-amu.fr

ORCID: [0000-0001-5273-7117](https://orcid.org/0000-0001-5273-7117)

[CV HAL](#)

marie-laure.tremelo@univ-amu.fr

ORCID: [0000-0002-4802-7179](https://orcid.org/0000-0002-4802-7179)

[CV HAL](#)

ANNEXE - Les récentes actions du groupe Data-DRIIHM

ANR SO-DRIIHM (2020-2023)



1. **Promotion de la science ouverte** :
montrer les bénéfices du partage et de l'ouverture des données à la communauté scientifique DRIIHM
2. **Co-construction d'une e-infrastructure ergonomique et interopérable**,
intégrant les outils existants et accompagnant les scientifiques dans la « FAIRisation » de leurs données
= **portail aiguilleur du « FAIR »**

- Webinaires « Cafés du DRIIHM »
 - 24/03/21 : [RGPD et Plan de Gestion de Données](#)
 - prochain en octobre sur Accès et Partage des Avantages issus de la biodiversité (APA)
- SO-DRIIHM
 - communications lors de 3 séminaires des OHM
 - accompagnement dans projets pilotes
 - 21/06/21 : lancement nouvelle enquête « Partagez votre expérience de partage de données ! »
 - 06/09/21 : session plénière SO-DRIIHM ouverte à tous lors du [séminaire DRIIHM](#) -> **inscrivez-vous !**
 - démarrage de la co-conception de l'e-infrastructure avec de futurs ateliers de test réguliers



Annnonce - Communication journées S.O. MATE-SHS - **23 septembre 2021**
 « L'e-infrastructure DATA-DRIIHM, un aiguilleur de la Science Ouverte pour renforcer l'interdisciplinarité et le partage des données des OHM »