

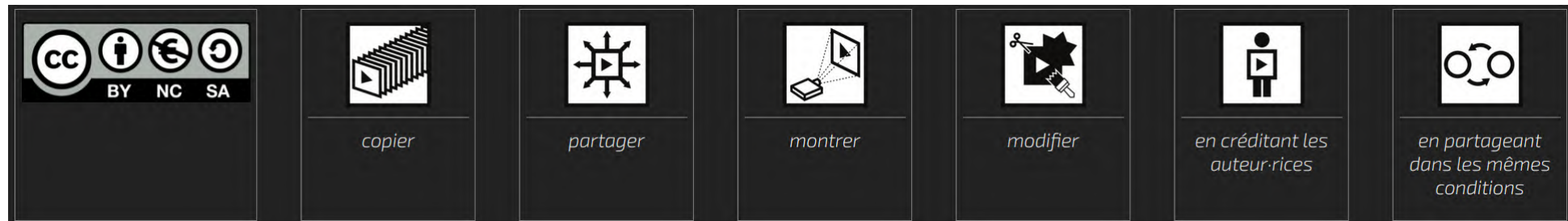
**Données environnementales
des organismes de recherche :
quelles valorisations
économiques ?**



Avertissement préliminaire

Ce document est protégé par la licence de libre diffusion CC-BY-NC-SA :

- [BY] : le document et les éléments qui le composent peuvent être librement utilisés, à la condition de les attribuer à l'auteur. Pour citer ce document ou des éléments qui y figurent : *Analyse stratégique collective « Valorisation économique des données environnementales », AllEnvi Solutions, 2022.*
- [NC] : tant que l'utilisation n'est pas commerciale, la reproduction, la diffusion, la modification sont autorisées ; sinon contacter l'auteur.
- [SA] : le partage du document ou des éléments qui le composent est soumis aux mêmes obligations.



Contributeurs

Cette étude est une Analyse Collective Stratégique (ASCo), qui a été réalisée dans le cadre du Plan d'Investissement d'Avenir attribué au Fonds National de Valorisation.

Le comité scientifique était composé de :

- François Christiaens, AllEnvi Solutions
- Dominique Morin, BRGM

Les membres du groupe de travail [CoVAIEnvi](#) ont participé activement à l'élaboration de cette étude.



Sommaire

AllEnvi Solutions	p. 5
Présentation de l'étude	p. 8
Contexte	p. 11
Réglementations et politiques	p. 17
Enjeux pour la recherche	p. 21
Marchés	p. 36
Acteurs	p. 42
Modèles économiques	p. 62
Valorisations	p. 67

allenviSolutions

● AllEnvi Solutions

Au cœur de la recherche pour l'environnement



AllEnvi Solutions est au cœur d'un réseau de 20 000 chercheurs spécialistes de l'environnement.



Une expertise croisée intelligence économique & environnement : des consultants seniors, de formation scientifique & économique.



À l'interface entre recherche et entreprises, dédiée aux grands enjeux d'innovation durable.

Les grands enjeux d'innovation durable



VILLE DURABLE

Agriculture urbaine
Qualité de l'air # Végétalisation



ECONOMIE CIRCULAIRE

Recyclage des métaux critiques
Réutilisation des eaux usées



BIODIVERSITÉ

Impact des entreprises
Données environnementales



QUALITÉ DE L'AIR

Air intérieur # Air extérieur
Impact de l'exploitation minière



RESSOURCES RENEUVELABLES

Chimie biosourcée # Ressources marines
Biologie de synthèse



USAGES DE L'EAU

Dessalement
Réutilisation des eaux usées



ALIMENTATION

Protéines végétales
Agriculture urbaine



AGRICULTURE

AgriTech # Capteurs connectés
Qualité de l'air



MOBILITÉ

Impact du Tourisme
Circuits courts



CHANGEMENT CLIMATIQUE

Services climatiques



SOL

Dépollution des sols



SERVICES NUMÉRIQUES

Données environnementales



AllEnvi Solutions bénéficie d'un financement du Programme Investissements d'Avenir (PIA) pour ses actions collectives



Introduction

Un objectif commun

Évaluer la nature et l'ampleur de la valorisation économique des données environnementales produites par la recherche publique

L'Analyse Stratégique Collective (ASCo) vise à :

- Contextualiser le cadre réglementaire régissant l'usage des données publiques et les marchés concernés
- Faire connaître les modèles économiques viables pratiqués
- Partager les politiques de valorisation économique existant au sein des organismes
- Contribuer ainsi à la conception de stratégies de valorisation ambitieuses

Ce document agrège les supports présentés lors du webinaire de restitution qui s'est tenu le 7 octobre 2022.



Périmètres de l'ASCo

- ✓ Données
- ✓ Bases de données
- ✓ Logiciels de traitement
- ✓ Logiciels d'analyse

✗ Valorisation **scientifique**
(publication, réutilisation)

✗ Appui aux politiques
publiques

Valorisations économiques :
contrepartie financière

Directe

Indirecte



Contexte

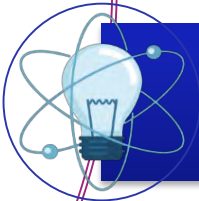
Introduction : vocabulaire



Jeux de données, Big data, smart data, open data, small data... ces concepts passent successivement sous l'éclairage médiatique. Les données y sont souvent comparées au nouveau pétrole, au sens d'une matière première abondante, riche et transformable industriellement, pour donner des produits / services divers, nombreux, utiles et peu chers.



Les métadonnées sont des données servant à définir ou décrire une autre donnée. Ex. : la date à laquelle une donnée a été acquise ; les coordonnées GPS du lieu où une photo a été prise. Les métadonnées précisent l'origine des données, attribuent les données à leurs inventeurs et apportent les bases du traitement que seul l'expert peut fournir.



Les données de la recherche scientifique et académique s'extrait progressivement de leurs support traditionnels tels les articles scientifiques, les rapports, les études et autres publications pour devenir des objets et informations à part entière, qui sont partagés librement sur le Web.

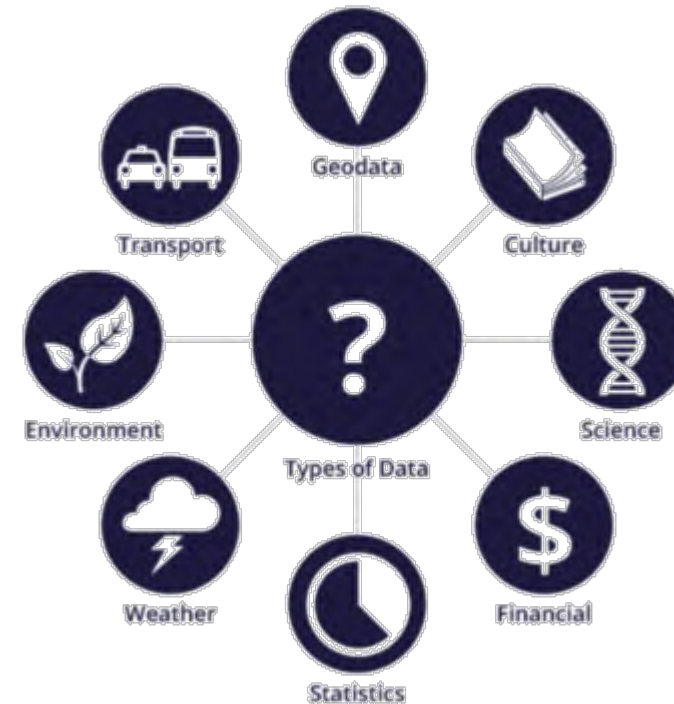


De même, les données publiques sont « ouvertes * » (i. e. accessibles ; sous un format exploitable ; dans un délai raisonnable) pour être mises à disposition de tous les acteurs économiques.

* ouvertes et non « libérées » : des licences sont appliquées

Définition : données ouvertes / open data

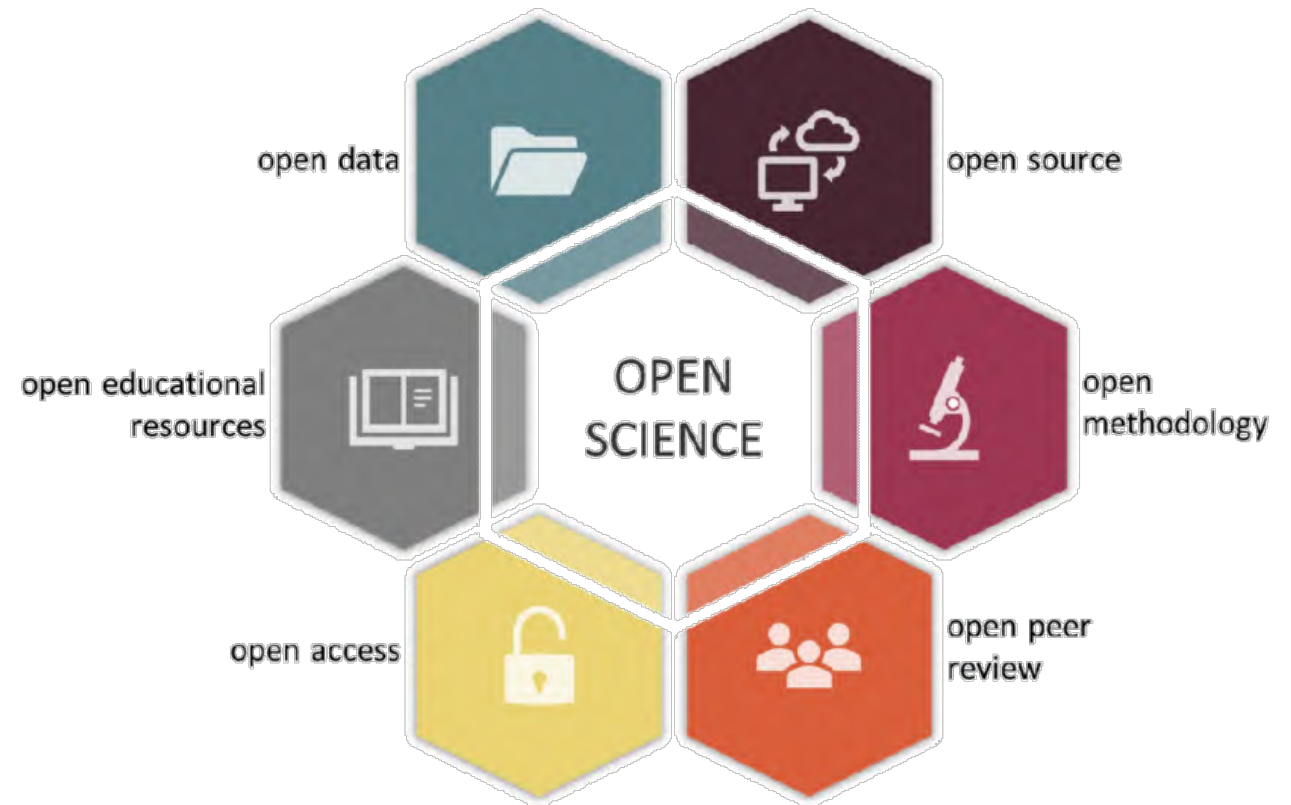
- « Une *donnée ouverte* est une information publique brute, qui a vocation à être librement accessible et réutilisable, [...], sans restriction [...]. »
- Ces jeux de données sont principalement mis à disposition par des organismes publics tels que des gouvernements, ministères, villes, régions, entreprises publiques etc.
- Il peut s'agir du cadastre, de la liste des organismes de formation, du recensement des équipements sportifs ou de l'emplacement des lampadaires... Les thématiques abordées sont larges et très nombreuses.
- L'open data n'est pas un phénomène nouveau et il continue à se développer (i.e. on accède à toujours plus de données), en France et partout dans le monde.



Source : Eduscol

Définition : science ouverte / open science

- Les **données scientifiques**, aussi appelées « données de la recherche » sont « *des enregistrements factuels (chiffres, textes, images et sons), qui sont utilisés comme sources principales pour la recherche scientifique et sont généralement reconnus comme nécessaires pour valider les résultats de recherche.* »
- *Ce terme ne s'applique pas aux éléments suivants : carnets de laboratoire, analyses préliminaires et projets de documents scientifiques, programmes de travaux futurs, examens par les pairs, communications personnelles avec des collègues et objets matériels (par exemple, les échantillons de laboratoire, les souches bactériennes et les animaux de laboratoire tels que les souris). »*
- Le mouvement de l'open science / science ouverte vise à rendre les données de la recherche accessibles à tous.



Les 6 caractéristiques de la science ouverte

Source : OCDE

Données environnementales : définitions

Les données environnementales recouvrent toutes les données qui permettent de caractériser les compartiments de l'environnement suivant les valeurs de paramètres physiques, chimiques ou biologiques dans leurs évolutions temporelles.

Elles sont de natures très variées.

- Pour l'OCDE il s'agit des données air et climat, biodiversité, déchets, eau, forêts, matériaux ou matière et politiques de l'environnement.
- Pour l'INERIS, l'inventaire des bases de données environnementales s'inscrit dans le Plan National Santé et Environnement.
- Pour le CIRAD, il s'agit de toute information relative à :
 - l'état des éléments de l'environnement, notamment l'air, l'atmosphère, l'eau, le sol, les terres, les paysages, les sites naturels, les zones côtières ou marines et la diversité biologique, ainsi que les interactions entre ces éléments ;
 - les décisions, les activités et les facteurs, notamment les substances, l'énergie, le bruit, les rayonnements, les déchets, les émissions, les déversements et autres rejets, susceptibles d'avoir des incidences sur l'état des éléments visés au premier point ;
 - l'état de la santé humaine, la sécurité et les conditions de vie des personnes, les constructions et le patrimoine culturel, dans la mesure où ils sont ou peuvent être altérés par des éléments de l'environnement, des décisions, des activités ou des facteurs mentionnés ci-dessus ;
 - les analyses des coûts et avantages ainsi que les hypothèses économiques utilisées dans le cadre des décisions et activités visées au deuxième point ;
 - les rapports établis par les autorités publiques ou pour leur compte sur l'application des dispositions législatives et réglementaires relatives à l'environnement.

Définition des données environnementales générées par la recherche

Il serait très difficile de définir ce qu'est la « recherche environnementale », autrement dit ce qui en est et ce qui n'en est pas ; toute tentative de définition pourrait passer à côté de nombreuses données tout à fait pertinentes. En conséquence, la présente étude s'intéresse aux **données qui décrivent l'environnement produites dans le cadre de projets de recherche et/ou par les infrastructures de recherche des organismes d'AllEnvi, pour comprendre, prévoir et anticiper.**

Les données produites par les infrastructures de recherche sont soumises aux politiques des organismes en matière d'open data ; ces politiques dépendent de l'organisme.

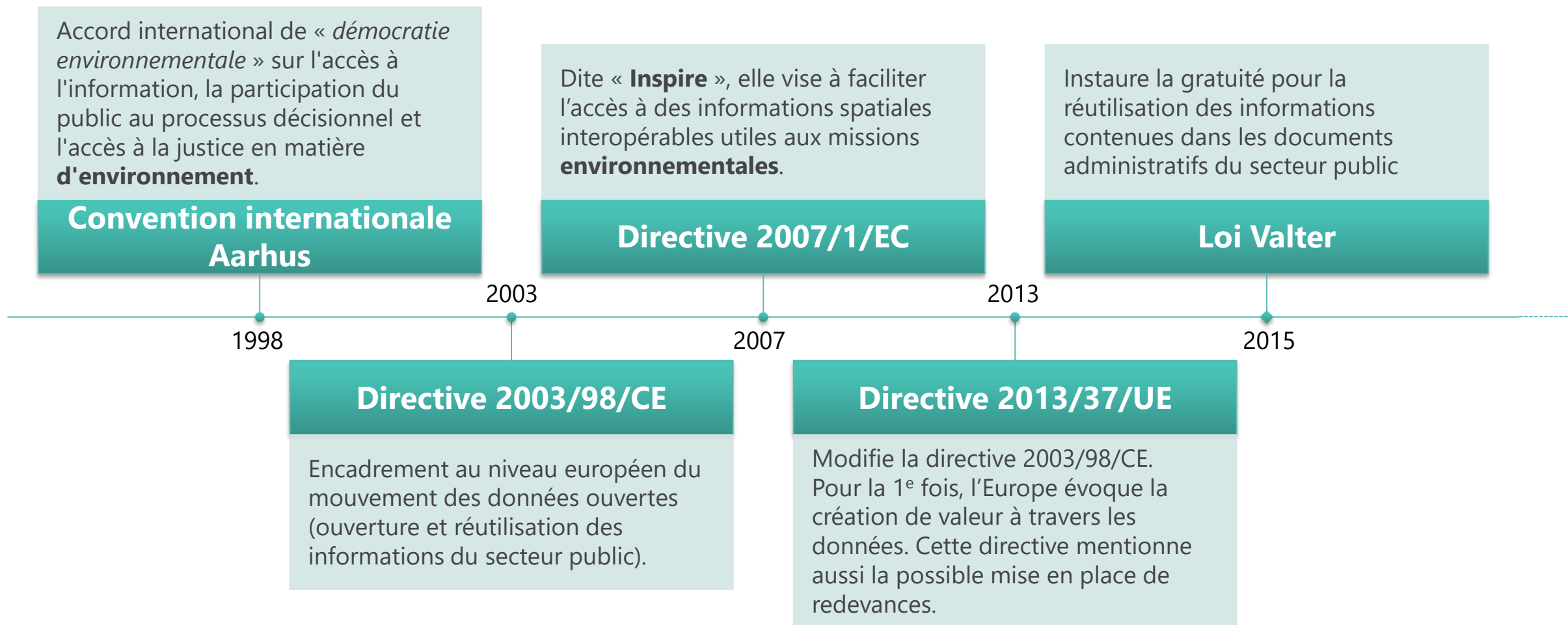
- Les **données environnementales peuvent être synthétiques et réciproquement, les données synthétiques peuvent décrire l'environnement.** Les données synthétiques font l'objet d'un engouement récent et fort, qui a des conséquences pour les données environnementales générées par la recherche. Ces aspects sont abordés dans une [annexe](#) spécifique.
- La norme ISO 14025 quantifie l'impact environnemental des processus de fabrication des produits utilisés en construction. Dans ce cadre, **le terme de données environnementales réfère alors aux [données concernant la durabilité de matériaux](#)** ; cette acception n'est pas celle utilisée dans cette ASCo.



**Réglementations
et politiques**

Un cadre instaure l'ouverture des données environnementales

Des réglementations nouvelles, dont le périmètre est étendu et les applications sont précisées depuis 25 ans



Un cadre instaure l'ouverture des données environnementales

Lois, politiques obligatoires, informations...



Ouverture et gratuité des données sont la règle, assortie d'exceptions utiles pour la valorisation



La « [loi pour une République numérique](#) » vise à donner un accès plus transparent aux administrés aux données les concernant, afin que les citoyens puissent exploiter ces données, pour générer de l'activité économique. Elle modifie marginalement le code de la recherche. Les règles de publication des informations publiques sont précisées dans le code des relations entre le public et l'administration à [l'article L321-1 du chapitre sur l'étendue du droit de réutilisation des informations publiques](#).

« En matière de protection des données entraînant des restrictions à leur ouverture, et sans prétendre à l'exhaustivité, on trouve :

- la protection des données à caractère personnel, en vertu du règlement général européen sur la protection des données (RGPD) et de la loi Informatique et Libertés
- Le secret de la défense, qui protège la sécurité nationale
- Le secret des affaires, qui protège la libre-concurrence
- **[L'article L124-4 du code de l'environnement](#) (par exemple non-divulgarion de l'emplacement d'une biodiversité menacée car convoitée)**
- La souveraineté économique
- La prise en compte des contrats et de leurs clauses. »

L'établissement est concerné par le [dispositif de Protection du Patrimoine Scientifique et Technique de la Nation](#) (PPST) si *la captation induite ou le détournement des savoirs, savoir-faire et technologies développés ou mis en œuvre dans l'établissement peuvent porter préjudice de manière significative à la compétitivité, à celle des partenaires industriels ou à celle du pays.*

* Le PPST est un dispositif français, il ne s'agit pas de la retranscription nationale d'une directive européenne.

Ces deux contraintes, antinomiques, induisent les organismes de recherche à déterminer « au cas par cas » la politique à appliquer, étant donné que les données et logiciels qu'ils produisent sont accessibles par défaut et qu'en même temps dans le cadre de valorisations économiques, ils donnent un avantage à des acteurs privés dans le cadre de collaborations public-privé. En fait la **valorisation économique** des données permet **d'augmenter la compétitivité des partenaires industriels** des organismes de recherche.

Sources : interview D. Margery, Inria ; A. Brunner, Inrae ; Legifrance ; analyse AllEnvi Solutions



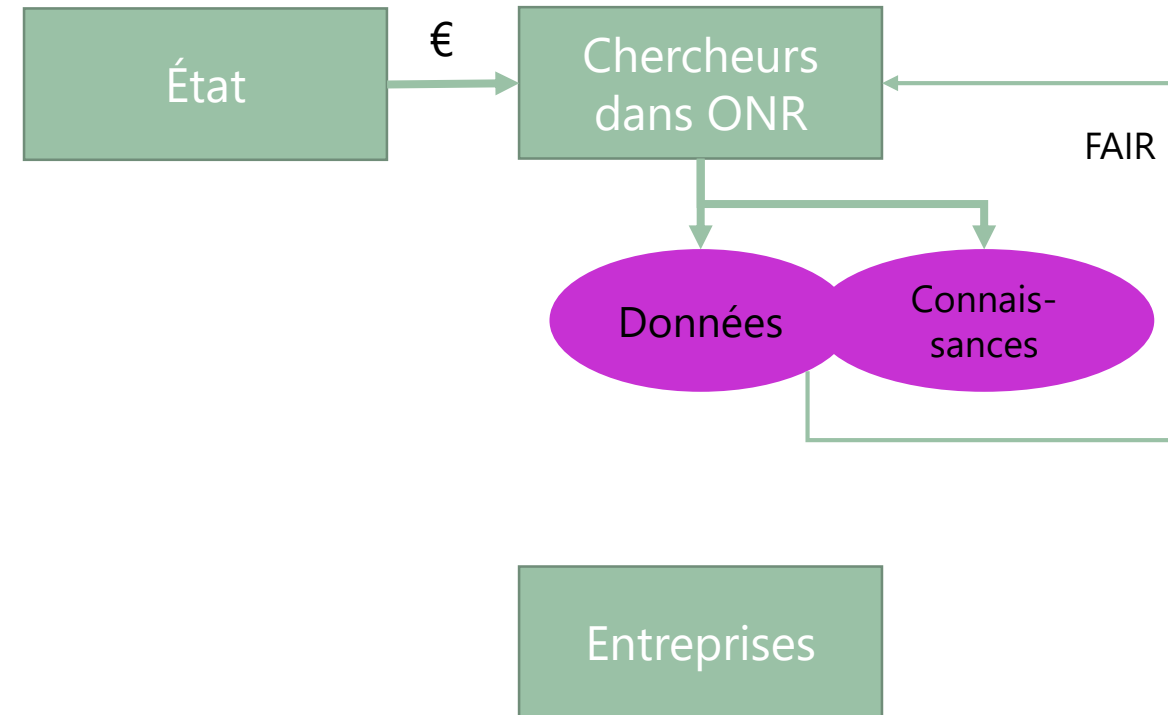
● **Enjeux pour la recherche**

Macro-organisation des acteurs et des flux

Au sein des organismes nationaux de recherche, en suivant [la démarche scientifique](#) les chercheurs produisent des données. Pour cela, ils mobilisent des ressources (équipements, intelligence) auxquelles l'État pourvoit, via le financement et la formation.

À partir de ces données, les chercheurs produisent de la connaissance.

Ces données sont ouvertes, et la politique de la science ouverte impose le processus de « FAIRisation » des données, qui permet aux chercheurs de réutiliser simplement des données produites par d'autres chercheurs.

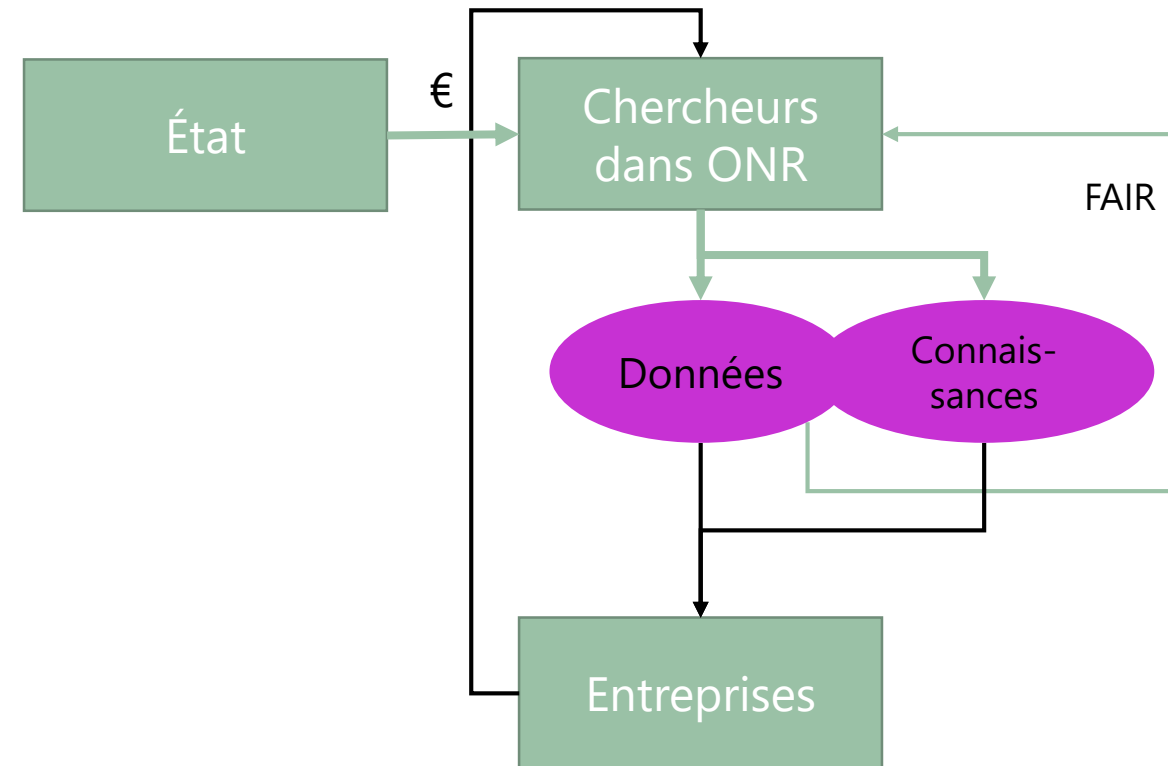


La boucle de la valorisation

L'action des équipes de valorisation au service de la recherche peut être schématisée selon la boucle représentée en noir sur le schéma ci-contre qui consiste, à partir des connaissances, expertises, savoir-faire, technologies et compétences développées par la recherche, à établir des accords avec les entreprises pour qu'elles puissent les utiliser, éventuellement en les adaptant. Le processus de valorisation établit des contrats entre les entreprises et les organismes de recherche, qui induisent

- un flux de projets, de données en retour vers les équipes de recherche
- et un flux financier supplémentaire vers les organismes et les chercheurs.

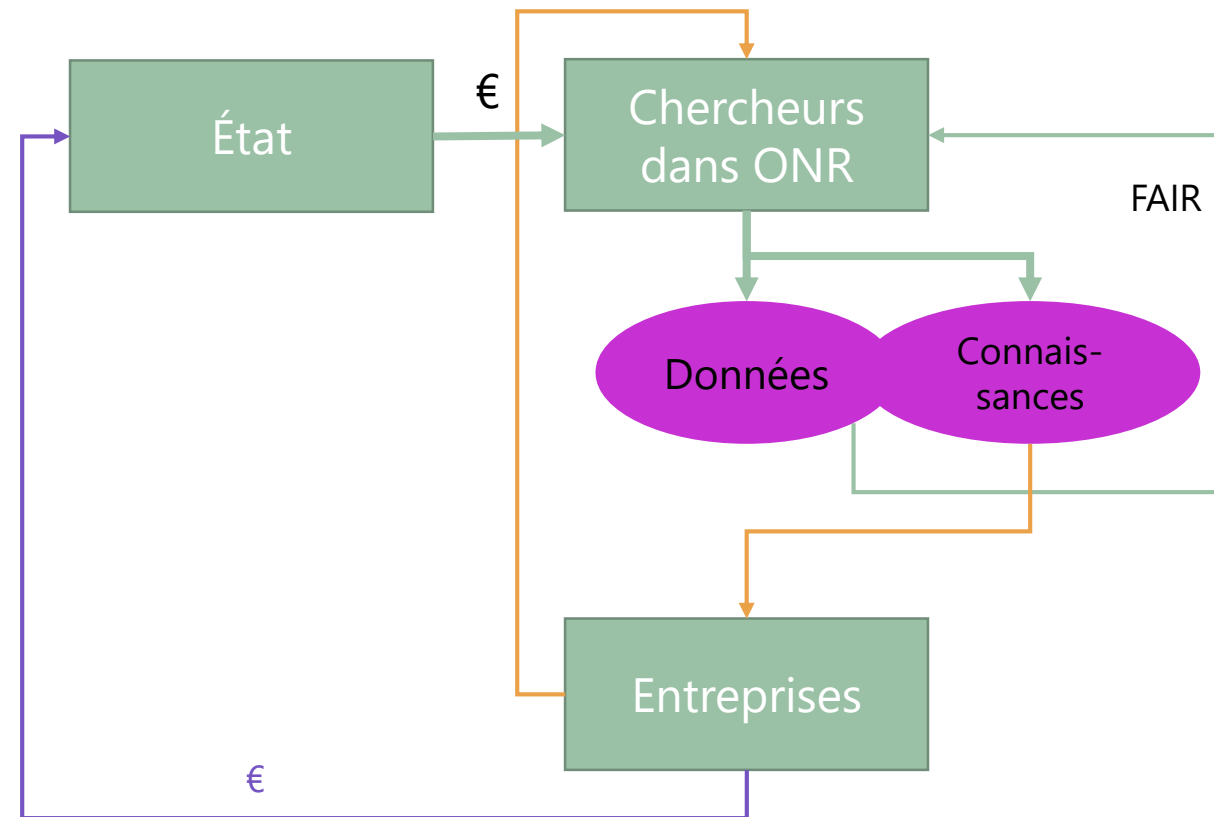
Avant l'ouverture des données de la recherche, celles-ci faisaient partie des ressources mobilisables par les équipes en charge de la valorisation, de façon discrétionnaire.



Captation de la valeur ajoutée grâce à la circulation de données

À partir du moment où les données sont ouvertes, les entreprises peuvent s'y intéresser pour ajouter de la valeur à leurs produits ou services, en les agrégeant avec d'autres données, y compris des données dont elles sont propriétaires.

Mais la valeur produite n'est plus captée par les organismes de recherche, elle est récupérée par l'État via l'activité économique supplémentaire générée par les entreprises (taxes, diminution du chômage et donc des aides pour l'emploi, etc.) – si la captation est locale.



Enjeux économiques pour la recherche

Investissement dans les données

- satellites : 1 G€ / an
- autres infrastructures nationales : investissement 1 G€ ; fonctionnement de 200 M€ /an et 1000 ETP.

Pression budgétaire sur les infrastructures de recherche

Réglementation : gratuité d'accès aux données publiques

Évolution des pratiques de recherche : données de plus en plus nombreuses (services d'observation, big data) et outils de traitement de données qui évoluent

L'arrivée de **nouveaux acteurs économiques**, notamment issus des outils de communication et de la collecte et l'exploitation de masse de données

Cost of not having FAIR research data : 3 G€ pour la France. Causes :

- pertes de temps,
- non optimisation des coûts de stockage,
- frais de licence,
- aux problèmes de duplication de la recherche,
- au manque de fertilisation croisée.

Source : Sujets Transversaux Groupe Prospective Octobre 2014 ; analyse AllEnvi Solutions

Forces

Légitimité scientifique / régaliennne



Connaissances expertes des données environnementales

- Nb de chercheurs impliqués sur un même sujet / spécialisation
- Nb de voies explorées
- Diversité et volume des données



Politique appliquée de gouvernance des données

- Conception (BdD, logiciels : « Forge »)
- Entrepôt
- Diffusion - principes de la science ouverte (FAIR)
- + formation sur les plans de gestion | personnel dédié | outils dédiés au service de la donnée



Structuration via les Infrastructures de Recherche, en coordination européenne



Collectif AllEnvi + CoVAIEnvi pour la réflexion et les retours d'expérience

Volonté de s'intégrer dans l'économie de la donnée



Source : interviews CoVAIEnvi ; analyse AllEnvi Solutions

faiblesses

Multiplicité des thématiques

- Données en silos, ou non recensées : non FAIR
- Données sous accord de confidentialité / non propriétaires / dilution des contributions (imbrication dans réseaux)
- Identification des valorisations possibles, depuis le stade des données brutes / au niveau de la stratégie de l'organisme, *a fortiori* quand recherche fondamentale

Dilemme public vs commercial

- Incitation à essayer faible et organisme-dépendant
- Perception négative (ROI + captation) → logiciels à TRL < 4
- Comment faire valoir le retour (notoriété) ?
- Obligation préalable de l'appui aux politiques publiques
- Effectifs en diminution

Diversité des types de données

- Données de format et de nature (images, vidéos, sons, réponses à des questionnaires, cartes etc.) très divers, ce qui rend complexe la mise en place d'ontologies, *a fortiori* interopérables entre les différentes disciplines

Difficulté de FAIRisation

- **Logiciels**
 - La simplification des outils de traitement, étape sous-évaluée
 - Culture de l'open source pas assez répandue dans les organismes

Source : interviews CoVAIIEnvi ; analyse AllEnvi Solutions

Opportunités

Accès plus simple aux

- Outils déployés : acquisition, transmission, stockage, traitement par IA / fusion de données + culture donneurs d'ordres et générationnelle des chercheurs
- Infrastructures : IR Data Terra, EOSC...
- Publics et retours d'informations : communautés d'utilisateurs comptant plus de membres

Réglementation : ouverture des données

- Incitation scientifique : amélioration des sujets de recherche
- Pousse à travailler la valorisation
- Logiciels et expertises sont valorisables

Économie de la donnée en expansion

- Exemples d'acteurs qui donnent de la valeur et diffusent des données
- Augmentation des risques : aléas / exposition / fréquence / vulnérabilité / valeur économique
- Domaines multiples / demande croissante en jumeaux numériques
- Volume croissant de la demande sociétale en services et indicateurs, pour orienter les stratégies et les politiques
- Rythme de conquête des GAFAM moins soutenu concernant l'environnement

Source : interviews CoVAIEnvi ; analyse AllEnvi Solutions

menaces

Complexité technique

- Sauvegarder des données peut être difficile et coûteux (data center privés et onéreux) ; agréger, tracer = difficile
- Pénurie savoir-faire fonctions support : maintenance, interopérabilité, data science

GAFAM / ESN

- Étendue du savoir-faire en matière de collecte et valorisation des données (B2C / B2A, B2B) (météo, qualité de l'air, sols, biodiversité...)
- Efficacité marketing
- Rapidité de déploiement de solutions par les ESN
- Solutions court-circuitant l'expertise scientifique via l'IA
- Captation des bénéfices par pays du Nord

Partenariats public-privé entravé par imprécisions

- Partage des résultats / modèles économiques fondés sur exclusivité d'usage des données
- Délai de négociation des clauses des contrats

Taille des marchés

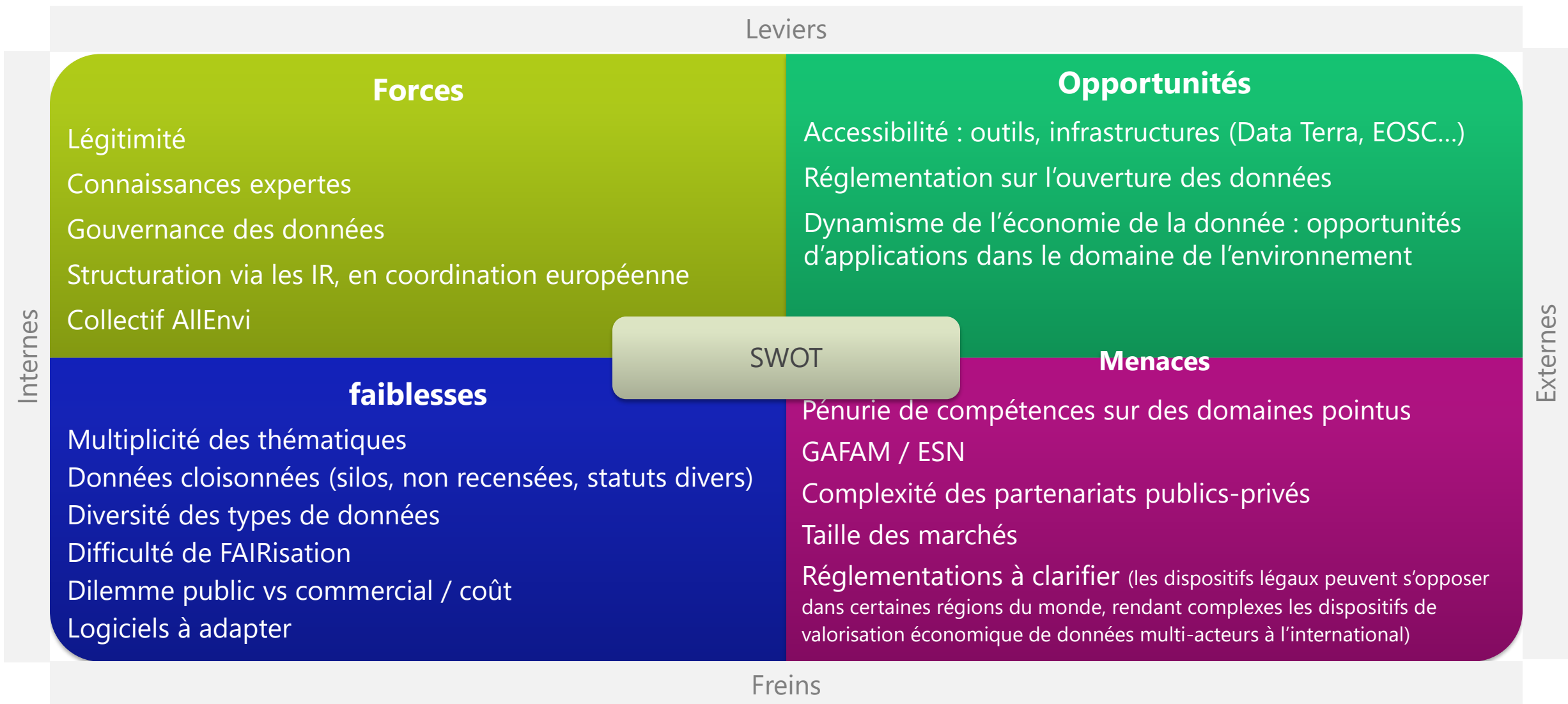
- Méfiance des acteurs par rapport au gratuit → communautés de taille réduite
- Environnement = sujet accessoire pour les industriels, même si très nombreux concernés → marché atomisé

Réglementation

- Politique open-data encore à clarifier (FR)
- Politique nationale parfois plus contraignante qu'europpéenne
- Juridique : pas de propriété sur les données, droit d'auteur peu respecté
- Protocole de Nagoya (accès aux ressources génétiques et le partage juste et équitable des avantages découlant de leur utilisation)

Source : interviews CoVAIIEnvi ; analyse AllEnvi Solutions

SWOT à l'échelle d'AllEnvi



Au niveau des chercheurs, des freins récurrents au partage des données

Depuis 2018, Figshare, Digital Science et Springer Nature mènent une enquête annuelle pour évaluer les freins exprimés par les scientifiques au partage de données. Ces enquêtes sont internationales et le périmètre concerne toutes les données ouvertes issues de la recherche.

Les préoccupations qui arrivent en tête portent sur une mauvaise utilisation des données, la reconnaissance (valorisation scientifique et/ou médiatique), et le manque de connaissances en matière de droits d'auteur et de licences. Ces items ont toujours préoccupé prioritairement les chercheurs, d'après les enquêtes.

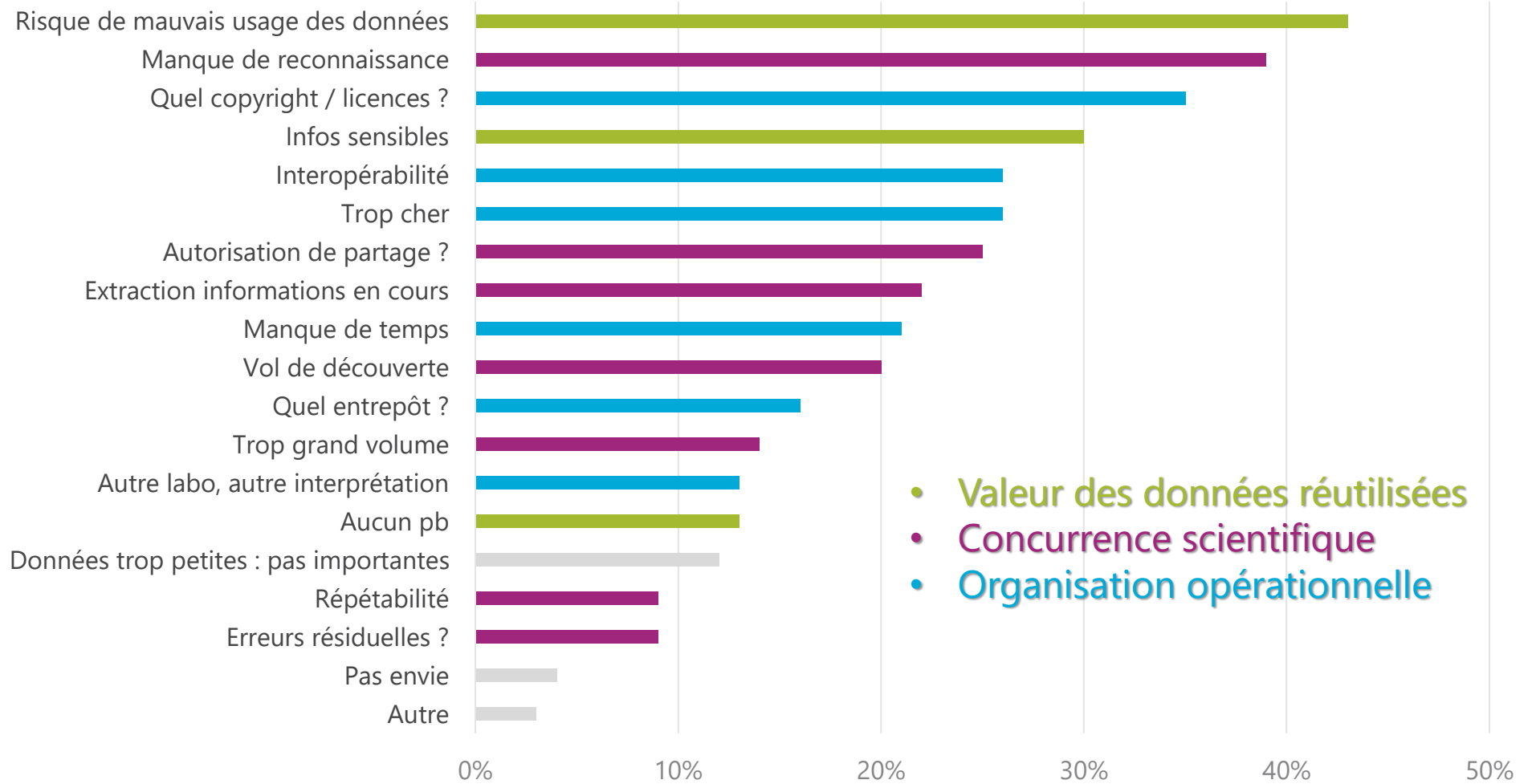


Source : Springer Nature Goup

Priorisation des freins au partage des données pour les chercheurs

Ces freins au partage peuvent être répartis selon 3 catégories :

- Organisation opérationnelle
- Concurrence scientifique
- Valeur des données réutilisées

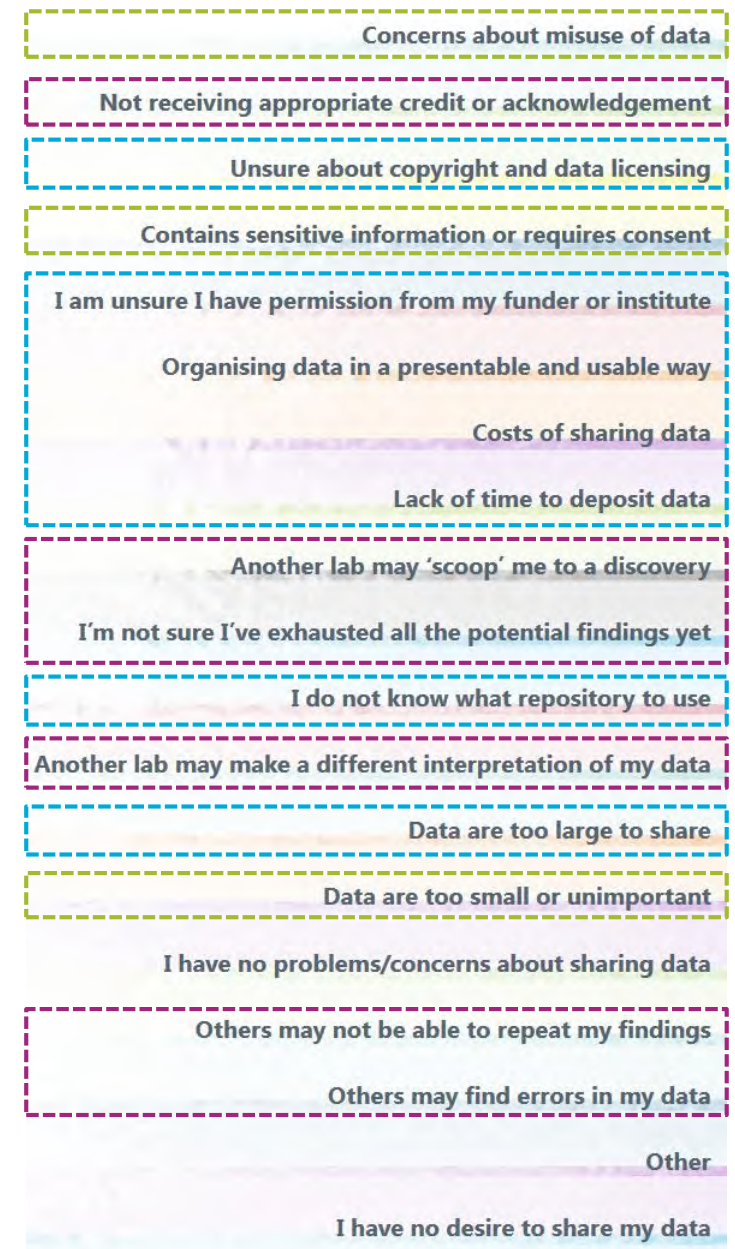


Source : Springer Nature Goup

Les organismes travaillent à lever les freins au partage des données



- La plupart des freins concerne l'organisation opérationnelle. Les organismes de recherche travaillent à réduire les problèmes correspondants, en créant et diffusant des formations pour accompagner les chercheurs dans la démarche, en éditant des guides pratiques...
- Concernant la catégorie « concurrence scientifique », le rapport de la mission Bothorel identifie comme enjeu le plus important le sentiment qu'ont les chercheurs que la production de codes et de données partageables n'apporte rien pour leur carrière. Ce frein apparaît en 2^e position de l'enquête et les problèmes de coût et de temps à passer identifiés dans les enquêtes relèvent de cette catégorie, autant que de la catégorie « organisation pratique du partage ».
- *« Un autre frein est la peur de « mal faire » et, parfois, le manque de compétence. Écrire un code pour ses propres besoins de recherche est sans comparaison avec le fait d'écrire un code pour, potentiellement, les besoins des autres chercheurs. Il y a la crainte d'être jugé sur la qualité de son code et, ainsi de s'exposer. Cela pose la question de la formation des chercheurs. »*



Sources : Springer Nature Goup ; Rapport Bothorel ; analyse AllEnvi Solutions

La mise à disposition des données : des politiques en structuration

Les politiques de mise à disposition des données (dont les données environnementales) sont variables d'un organisme à un autre, d'une plateforme à une autre :

- **maturité de la réflexion** sur la mise à disposition des données plus ou moins avancée
- **ouverture à l'open data** plus ou moins appliquée
- **modèle économique** en cours de réflexion pour la plupart des infrastructures collaboratives
- **enjeux variables** : raison d'être ou simple produit de la recherche, volonté de monétiser ou non
- **valorisation sectorielle** plus ou moins ancienne et intéressante



Cet état de structuration ne permet pas de proposer une offre de recherche lisible.



AllEnvi pourrait-elle jouer le rôle de **vecteur de structuration** ?

Source : interviews chercheurs

L'enjeu de la souveraineté

- Souveraineté : pouvoir supérieur confié aux représentants du peuple pour assurer l'ordre, la continuité, le progrès, la justice, la sécurité et la liberté de la communauté de destin.
- **La maîtrise des données environnementales d'un pays, la connaissance de ses territoires, la préservation des milieux et des populations constituent des enjeux de souveraineté nationale. La disponibilité de ses données massives par une organisation malveillante pourrait constituer des menaces notamment de déstabilisation, dépendance etc.**
- GAFAM et BATX
 - mode de fct = appropriation des données
 - savoir-faire = centralisation de la collecte.
multiplicité des sources = verrou à la collecte des données. **Les entrepôts de données scientifiques, alimentés et maintenus par les organismes de recherche, lèvent un verrou majeur.**
- Réveil des États :
 - lois antitrust aux USA,
 - régulation des marchés dans l'UE,
 - primauté du parti communiste en Chine.
- Respect de la souveraineté économique des producteurs de données → données pas « moissonnées » sans contrepartie directe auprès des producteurs ou indirecte (impôts) auprès des représentants des utilisateurs, par exemple grâce à des réglementations protectrices.
- Producteurs de données souhaitant valoriser économiquement leurs données : sélectionner les acteurs qui respectent des critères de rétribution justes et équitables pour toutes les parties prenantes



● **Marchés et secteurs**

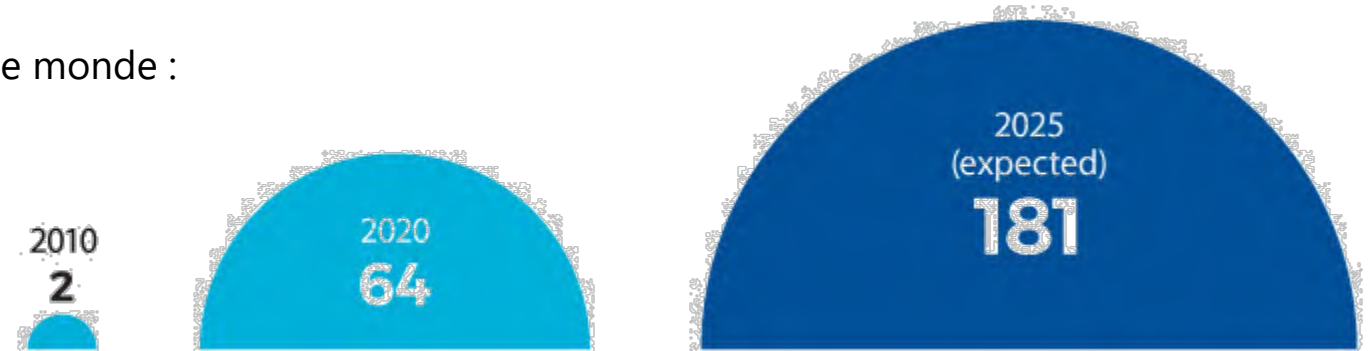
Marchés : indicateurs indirects

Volume de données créées et consommées dans le monde :

En zettaoctets

(1 zettaoctet = 1000 milliards de gigaoctets)

Source : Statista 2021



En 2015, la Commission Européenne publie la « [Stratégie pour un marché unique numérique en Europe](#) » : avec une augmentation de 40 % par an, la croissance des mégadonnées est « 7 fois supérieure à celle du marché de l'informatique ».

Le principe d'**ouverture des données** contribue à la création de valeur économique : 28 G€ pour 2019, selon une étude commandée par la Commission européenne.

En 2020, la valeur du **marché de l'IA** était estimée à 17,3 G\$.

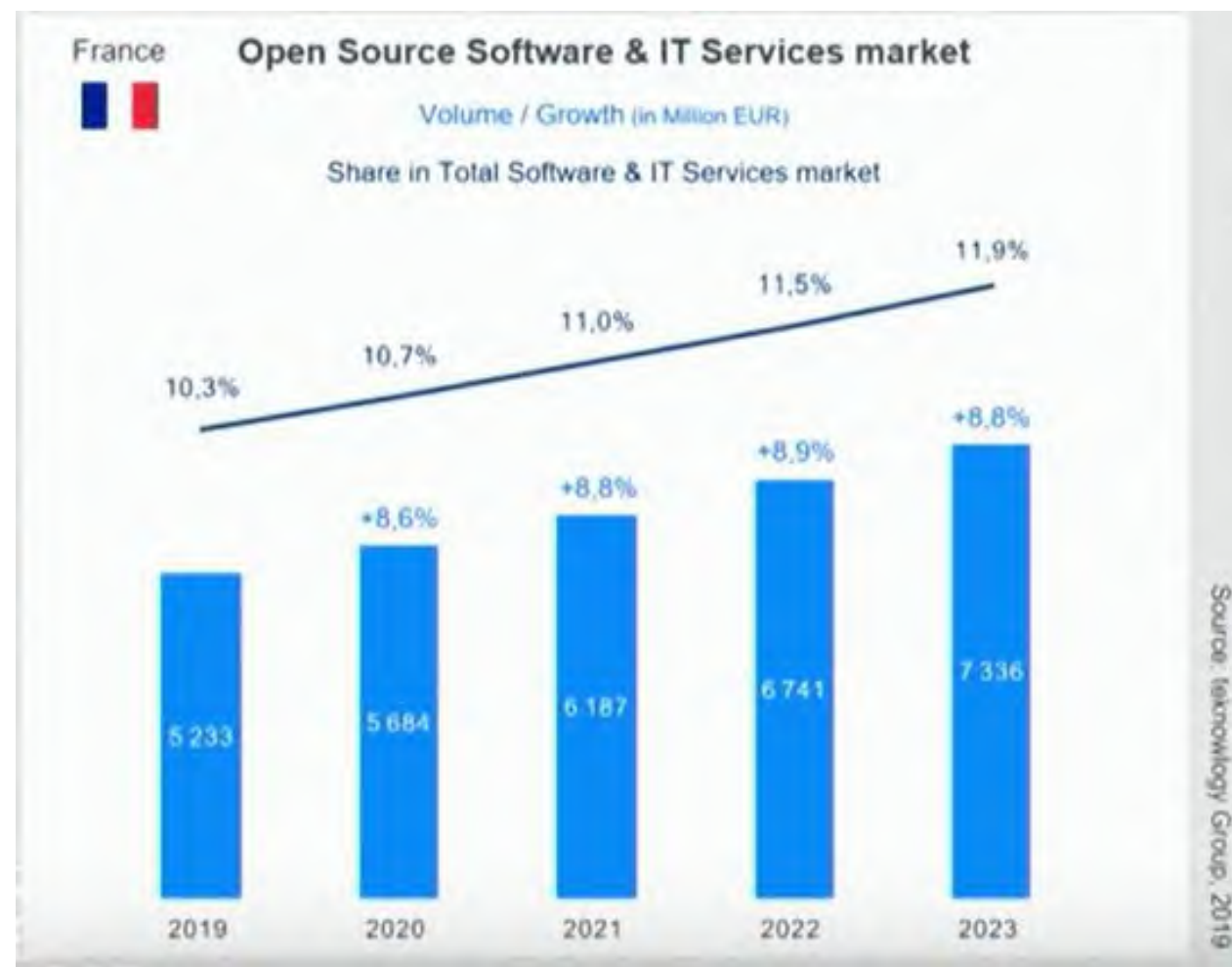
En 2025, l'estimation s'élève à près de 90 G\$. En cinq ans, la valeur du marché va être multipliée par plus de quatre.

Selon un rapport d'IDC publié en mars 2022, les recettes mondiales du marché de l'IA, y compris les logiciels, le matériel et les services, devraient augmenter de 19,6 % (en glissement annuel) en 2022 pour atteindre 432,8 G\$.

Sources : Techniques de l'ingénieur ; gouvernement.fr

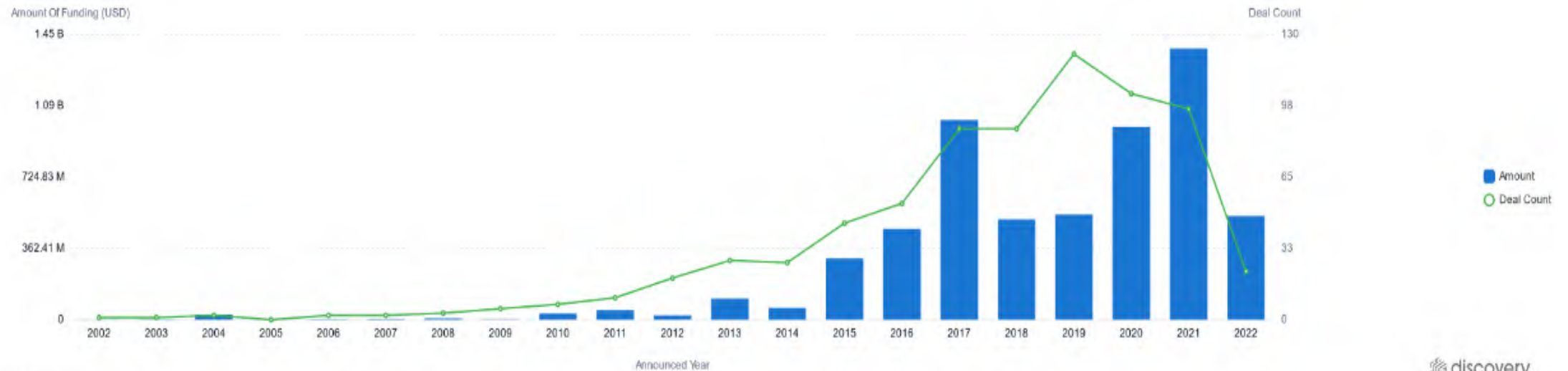
Chiffres sur la place du libre dans l'économie du logiciel

- Un marché en forte croissance : + 8-9 % par an environ. En France, le logiciel libre est un secteur deux fois plus dynamique que le marché IT dans son ensemble.
- Un marché représentant aujourd'hui plus de 10% du marché du logiciel.



Source : Inrae Transfert

Actions des fonds d'investissements



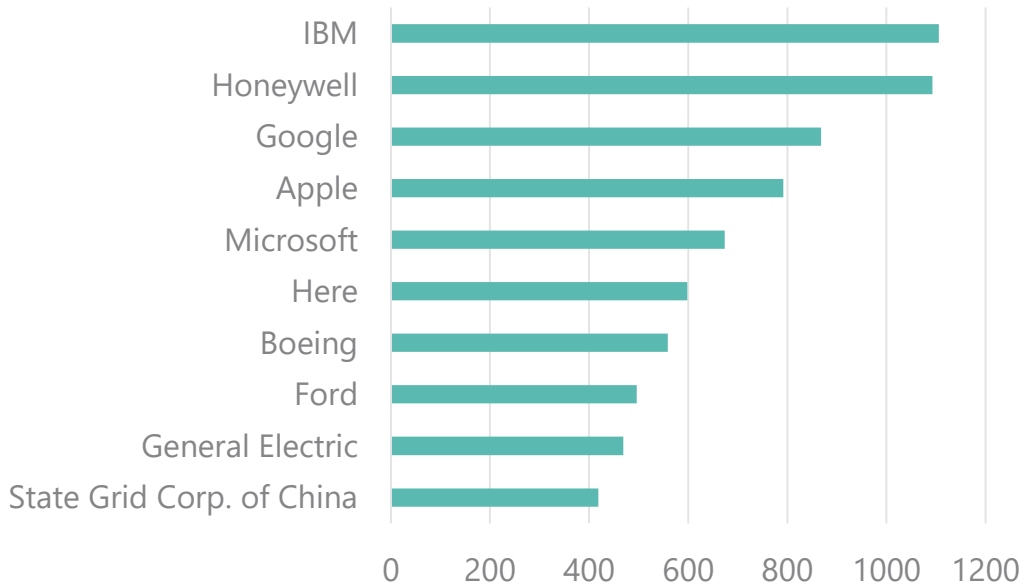
Source: discovery.patsnap.com

discovery
by patsnap

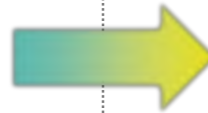
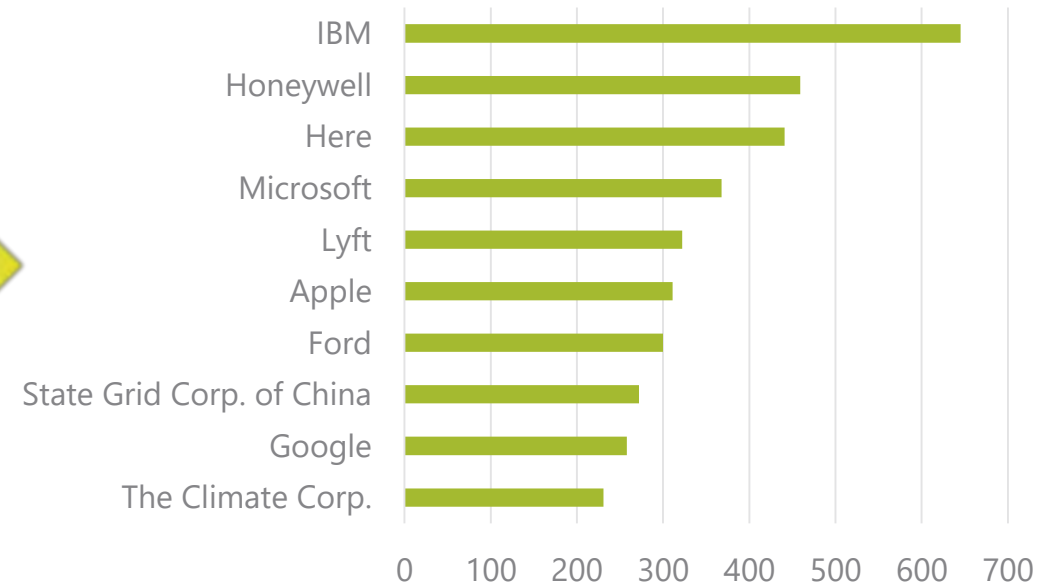
- Les sommes engagées par les fonds d'investissement augmentent, pour atteindre environ 1,4 G\$.
- En 2015, les investissements concernant les données environnementales ont décollé.
- Le nombre de transactions semble diminuer tandis que les sommes ne diminuent pas ; cela peut témoigner d'investissements plus massifs dans des sociétés moins nombreuses.

Brevets concernant les données environnementales

Entreprises déposant le plus de brevets



Entreprises déposant le plus de brevets, depuis 2017



- Les entreprises protègent leurs inventions qui concernent des marchés qu'elles estiment porteurs.
- Les plus gros déposants sont les plus gros utilisateurs de données environnementales.

8 des 10 entreprises qui ont déposé le plus de brevets figurent aussi dans le top 10 de celles qui en ont déposé le plus depuis 5 ans. La fraîcheur de ces dépôts témoigne du dynamisme du sujet.

Source : Patsnap Discovery

De nombreux secteurs économiques exploitent et dépendent des données environnementales

Secteurs économiques exploitant les compartiments du système terre et dépendant des données associées (sous-sol, surfaces continentales, océan, climat, météo, biodiversité) :



- Transport : terrestre, maritime, aérien ; infrastructures



- Eau : approvisionnement, risques d'inondation (pouvoirs publics, collectivités), sécheresses



- Santé-sécurité : qualité de l'air (collectivités, associations...), risques naturels (séismes, éboulements...)



- Assurances



- Appuis aux politiques publiques (bureau d'étude, ingénierie, conseil), exploitants de ressources minérales et végétales (BTP, mines et carrières, forêts...)



- Tourisme : sports d'hiver – enneigement et sécurisation des sites...



- Agriculture (semenciers, coopératives, négociants) : météorologie, nature des sols



- Construction et aménagement (aménageurs, promoteurs immobiliers, notaires) : structure des sols, qualité de l'air...



- Énergie : transformation, transport, obtention

Une étude d'AllEnvi Solutions s'est intéressée aux besoins des acteurs socio-économiques du secteur agricole : startups, PME, grands groupes, instituts techniques, coopératives, agriculteur, organismes de conseil et développement agricole.



● **Acteurs**

Les acteurs de l'open data en France

Les données environnementales de la recherche publique font partie de l'open data.
Les personnes morales de l'open data sont catégorisées ci-dessous.

Collectivités territoriales

- En plus des activités publiques, les activités parapubliques (musées, bibliothèques, conservatoires naturels...) doivent mettre à disposition leurs données

Institutions publiques et administrations

- Les EPA (IGN, Météo France, Cerema et SHOM) en font partie
- Les Organismes Nationaux de Recherche aussi

Entreprises

- Utilisatrices de données ouvertes, y compris les start-ups pure-players, éventuellement créées après des hackathons
- Mais aussi productrices de données

Courtiers (intermédiaires)

- Ils captent les données disponibles, issues de tout type de source. Vocation : guichet unique
- Pas ou peu de valeur ajoutée (mise en forme, croisement, vérification...)

Associations et collectifs

- Par exemple PartiCitaE, qui a pour vocation d'associer la société civile à la recherche.
- Mais aussi l'Ademe, dont [le portail datAgir](#) souhaite « apporter l'information environnementale au plus près des citoyens ».

Source : Interviews ; EGE

Catégories de rôles

Fournisseur

Facilitateur

Agrégateur

Enrichisseur

Développeur

Catégories de rôles

Un même acteur peut jouer plusieurs rôles : une entreprise peut utiliser et **fournir des données**

Fournisseurs

Définition : organisations qui publient leurs données via une interface ouverte pour permettre à d'autres de les utiliser et de les réutiliser.

Il peut s'agir d'organismes du secteur public ou d'entreprises du secteur privé.

Bien que les revenus directs de la publication des données soient nuls, pour le fournisseur les avantages commerciaux de cette transparence peuvent se traduire en des niveaux accrus d'engagement et de loyauté de la part des clients, et à une augmentation des revenus liés aux services ou produits vendus par la même entité, grâce à l'amélioration de sa réputation. L'accès à ces données peut aussi induire des avantages économiques, sociétaux et environnementaux, tels que les avantages liés à une réduction de la congestion grâce à une meilleure coordination des travaux routiers et un plus grand accès du public aux informations sur les travaux, rendu possible par la publication de données sur roadworks.org d'ELGIN.

Facilitateurs

Définition : organisations qui facilitent la fourniture ou l'utilisation de données ouvertes, comme le site web de concours Innocentive, mais qui n'utilisent ou ne réutilisent pas elles-mêmes des données ouvertes.

Ces modèles économiques sont directement générateurs de revenus en offrant des solutions rentables aux entreprises qui n'ont pas nécessairement les fonds nécessaires pour investir dans le développement de plateformes sur mesure ou dans l'analyse de données. Ils stimulent une plus grande production de données ouvertes.

Source : Deloitte

Catégories de rôles

Un même acteur peut jouer plusieurs rôles : une entreprise peut utiliser et **développer des logiciels**

Agrégateurs

Définition : organisations qui collectent et agrègent des données ouvertes et, éventuellement, des données propriétaires, généralement sur un thème sectoriel particulier. Les agrégateurs peuvent proposer des services consistant à trouver des corrélations entre jeux de données, visualiser des relations complexes, voire identifier des facteurs permettant d'optimiser des processus. Par exemple, [namR](#) fournit des renseignements et des services à valeur ajoutée aux professionnels de l'immobilier et aux particuliers, mais aussi aux pouvoirs publics.

Enrichisseurs

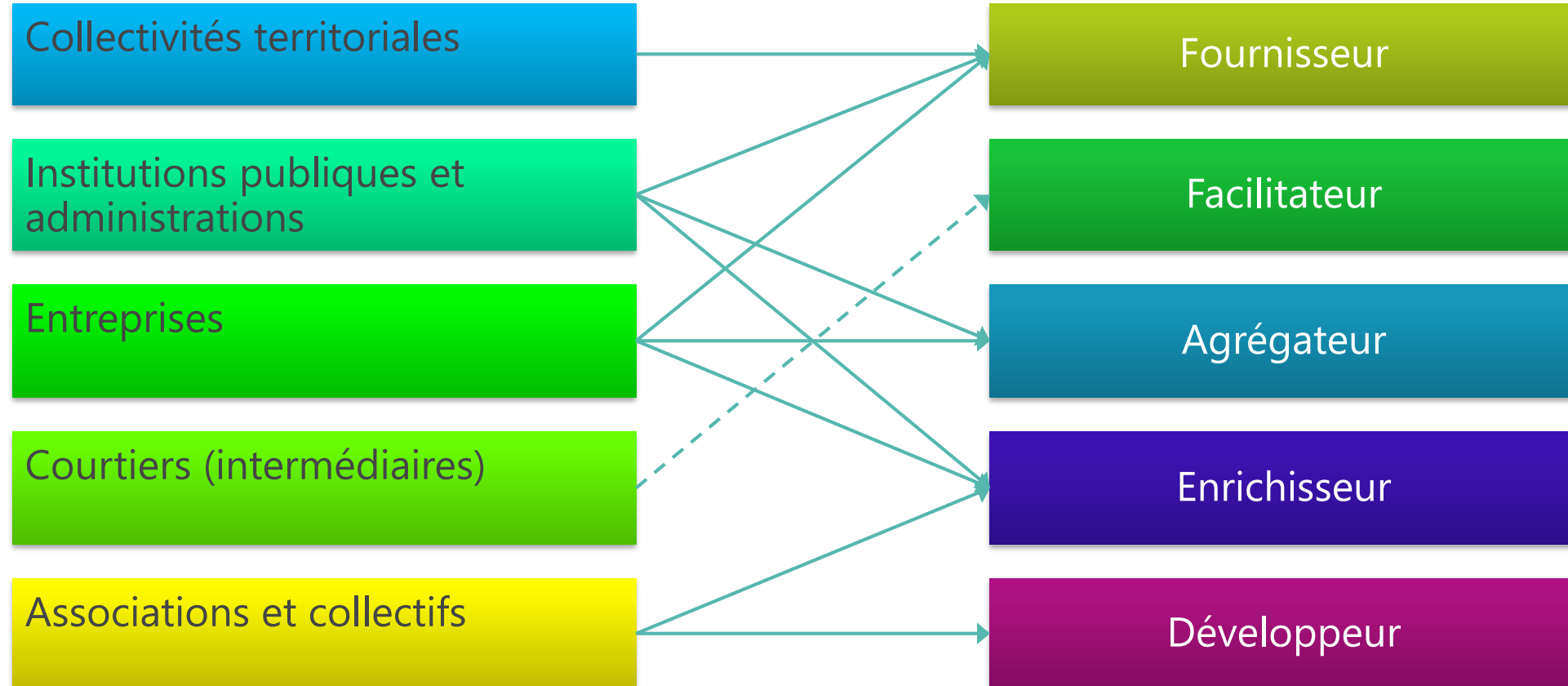
Définition : organisations (généralement de grandes entreprises) qui améliorent leurs produits et services existants grâce à une meilleure connaissance apportée par l'utilisation de données ouvertes. Ces produits et services ne dépendent pas entièrement des données ouvertes. Par exemple, les assureurs et les détaillants ont besoin de mieux comprendre les profils de risque et les données démographiques de leurs clients. Bien que les revenus ne proviennent pas directement des données ouvertes, les entreprises peuvent économiser de l'argent en les utilisant pour rendre leurs activités plus efficaces ou peuvent augmenter les ventes ou les taux de primes pour des produits et services de meilleure qualité.

Source : Deloitte

Développeurs

Définition : organisations et entreprises du numérique qui conçoivent, construisent et vendent des applications destinées à la consommation individuelle. Pure-players du numérique. Ces applications utilisent généralement les données ouvertes fréquemment mises à jour. Citons par exemple les applications de planification des transports personnels qui utilisent des données en temps réel provenant des réseaux de transport terrestres et souterrains, mais aussi météorologiques ou encore de qualité de l'air.

Acteurs et rôles





- **La recherche, productrice de données environnementales**

Où trouver les données environnementales produites par la recherche française ?

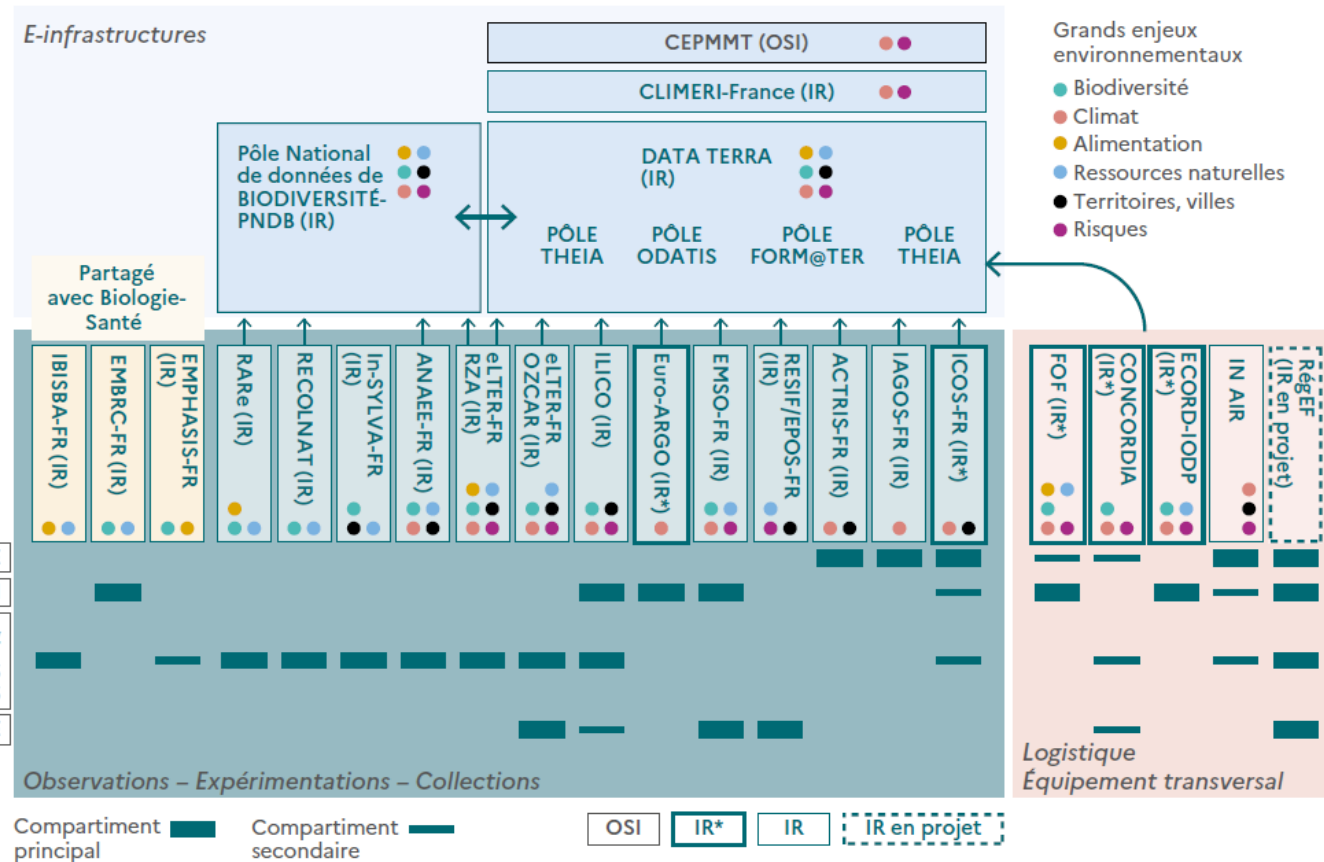


Les projets de recherche qui produisent et mettent à disposition leurs données sont très nombreux et de nouveaux projets démarrent sans cesse. Parmi ces projets, depuis 2020 ceux financés par l'ANR incluent systématiquement un Plan de Gestion des Données, afin de faciliter la réutilisation des données. En 2022, il n'existe pas d'inventaire des projets de recherche générant des données **environnementales**.

Les **infrastructures de recherche** hébergent de très nombreuses données environnementales produites par la recherche, avec leurs métadonnées, en accès libre ; et des données hébergées chez des partenaires, ces données étant librement accessibles ou non. Elles ne produisent pas de données de recherche mais représentent un réceptacle d'intérêt majeur.

Le paysage des infrastructures de recherche qui hébergent des données environnementales se structure. Parmi elles, Data Terra est perçue comme une composante française d'EOSC (voir [annexe sur la science ouverte](#)). Ces infrastructures servent d'abord à la réutilisation et au partage des données à des fins scientifiques, dans la démarche de science ouverte. Ces données peuvent aussi être valorisées pour des usages non scientifiques, et ce paysage aide à repérer quels acteurs produisent quelles données.

Paysage des infrastructures de données environnementales de la recherche française



Les infrastructures de recherche sur l'environnement détiennent des données d'observation, expérimentation et modélisation, destinées d'abord à comprendre le fonctionnement et l'évolution du système terre et de ses différents compartiments. Pour la période 2022-2025, les infrastructures de la recherche publique française du domaine « Environnement et Système Terre » sont organisées selon la carte ci-contre.

Le rapprochement entre le PNDB et Data Terra est recommandé et attendu durant cette période. Cette fusion, issu d'une volonté commune, devra surmonter les difficultés liées à la gouvernance avec un nombre élargi de partenaires (26 à 34) et à des données de natures différentes. Le projet Gaïa Data, défini plus bas, permettra de renforcer les synergies et de mutualiser les offres de services.

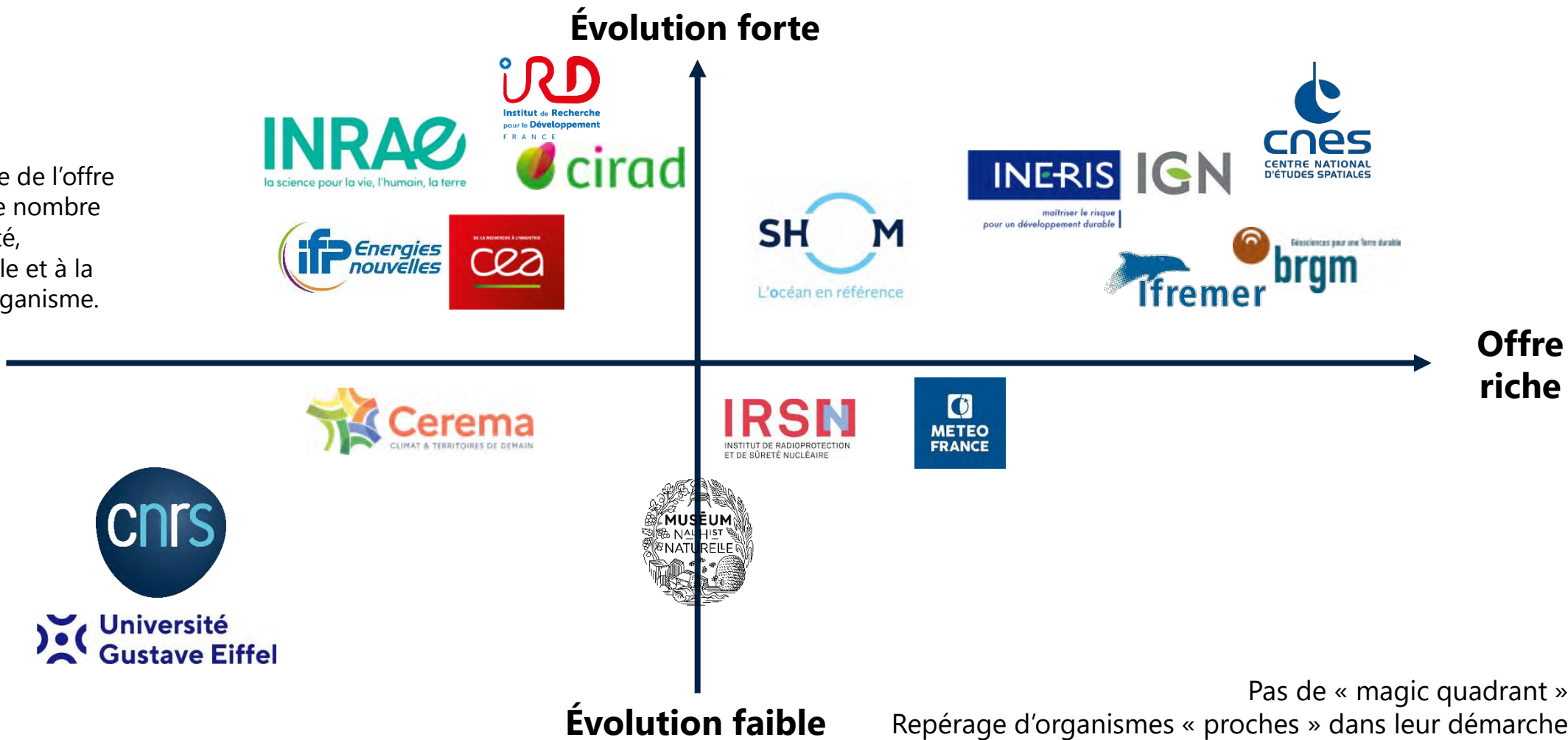
Figure 1 : OSI/IR*/IR du domaine SST & ENV par grands types (observations – expérimentaux – collections, logistiques et e-infrastructures), par grands compartiments du système Terre (atmosphère, océan, hydrosphère–écosystèmes – socio-écosystèmes continentaux – sols et géosphère) et par grands enjeux environnementaux (biodiversité, climat, alimentation, ressources naturelles, territoires-villes, risques). En jaune figurent trois infrastructures partagées avec le domaine Biologie – Santé.

Sources : MESRI ; interviews F. Huynh, R. Moreno

Offre relative des organismes en données environnementales

La notion de richesse de l'offre tente de combiner le nombre d'offres et leur variété, relativement à la taille et à la spécialisation de l'organisme.

Offre pauvre





 **Chaînes de valeur**

La chaîne de valeur des données de la recherche



Quand des tests sont effectués, le processus suit des étapes séquentielles :

- Des mesures, simulations ou des modélisations sont créées, pour tester l'hypothèse : cela génère des données ;
- Les données sont collectées et nettoyées ;
- Le cas échéant, d'autres données sont intégrées, issues de projets de recherche, de science participative ou de partenariats, et elles aussi sont nettoyées ;
- Les données sont traitées et interprétées, en faisant intervenir les expertises scientifiques nécessaires ;
- La valorisation scientifique ultime est la publication.

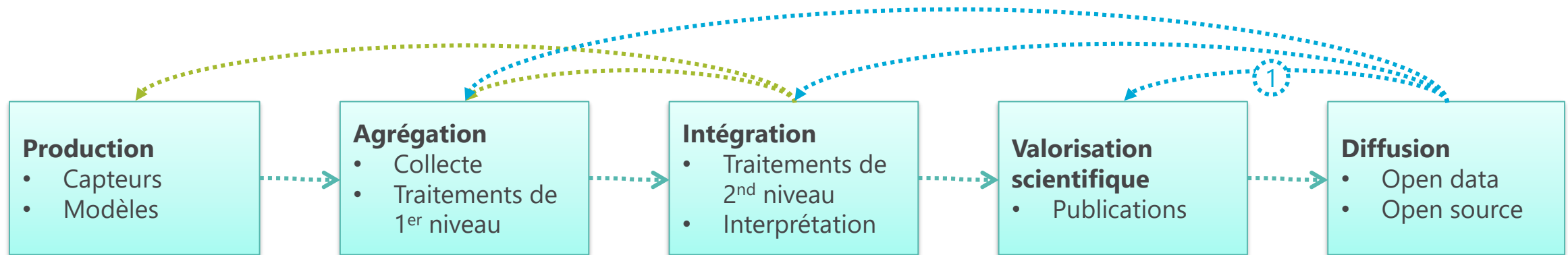
À noter que lorsqu'un projet de recherche implique au moins un partenaire privé, la valorisation économique des données peut faire l'objet d'une répartition exclusive entre les acteurs, et seulement une partie des données est diffusée selon les modalités de l'open data.

Chaque étape ajoute de la valeur aux données, et constitue un maillon de la chaîne de valeur des données de la recherche. Cette chaîne de valeur concerne la recherche.

La valeur d'un jeu de données provient à la fois des métadonnées, qui montrent l'organisation des paramètres d'intérêt pour une thématique donnée, et des données elles-mêmes.

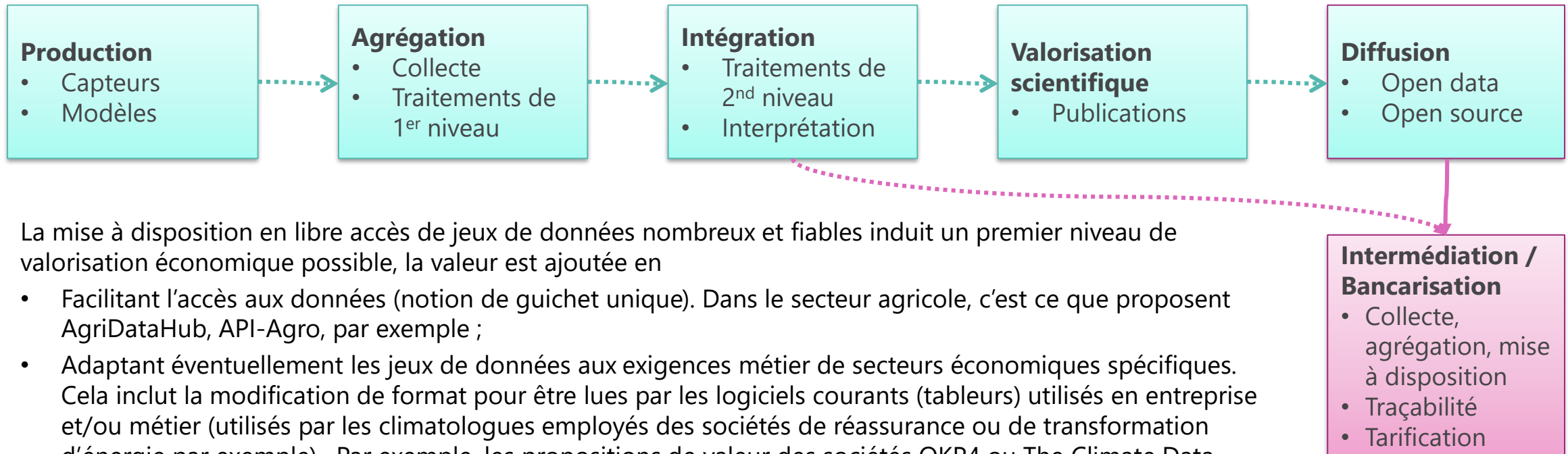
Les données ont une **valeur scientifique, pour le chercheur / le laboratoire qui les génère**. Mais aussi pour les structures qui captent ces données au plus près de leur processus de production, ce qui peut leur permettre de détecter et catégoriser les thèmes émergents de recherche ou les tendances ; par exemple [Octopus](#).

Les données sont produites tout le long de la démarche



- Au cours du processus, l'étape d'intégration peut mettre en évidence la nécessité d'obtenir des données complémentaires ou de raffiner les traitements automatiques sur les données brutes ; les **deux flèches vertes** issues de l'étape d'intégration illustrent ce flux.
- Les données diffusées en open data participent à la valorisation scientifique, via les data papers (**flèche bleue n°1**). Ces données peuvent aussi être agrégées, en tant qu'éléments de réponse de 1^{er} niveau à l'hypothèse testée, ou intégrées (**flèches bleues**).
- Dans le cadre **d'infrastructures de données de recherche**, des données sont produites sur des sujets identifiés d'intérêt majeur (par exemple, la mesure des gaz à effet de serre présents dans l'atmosphère : [ICOS](#)), alimentant en continu des bases de données. Ces données peuvent être utilisées en tant que matière première pour des travaux de recherche, ou bien être agrégées à d'autres jeux de données. Elles sont aussi systématiquement directement mises à disposition de la recherche.

Plusieurs étapes de la chaîne de valeur permettent d'engager un processus de valorisation économique



La mise à disposition en libre accès de jeux de données nombreux et fiables induit un premier niveau de valorisation économique possible, la valeur est ajoutée en

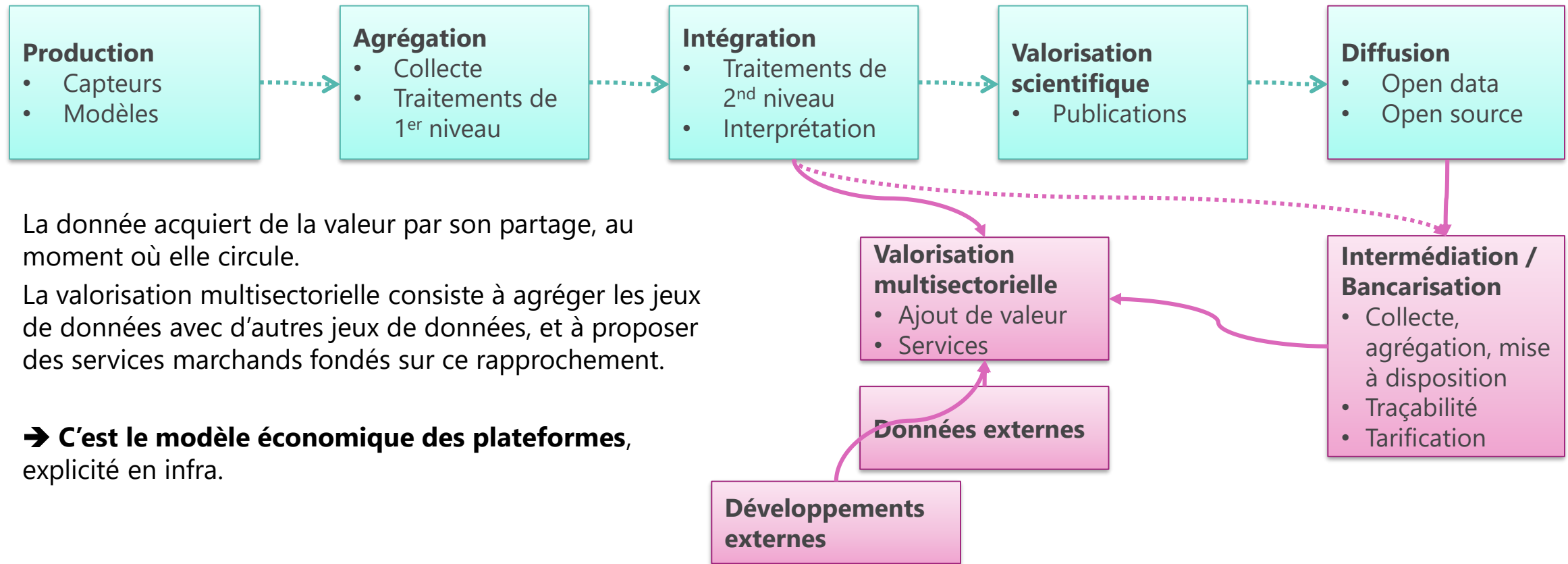
- Facilitant l'accès aux données (notion de guichet unique). Dans le secteur agricole, c'est ce que proposent AgriDataHub, API-Agro, par exemple ;
- Adaptant éventuellement les jeux de données aux exigences métier de secteurs économiques spécifiques. Cela inclut la modification de format pour être lues par les logiciels courants (tableurs) utilisés en entreprise et/ou métier (utilisés par les climatologues employés des sociétés de réassurance ou de transformation d'énergie par exemple). Par exemple, les propositions de valeur des sociétés OKP4 ou The Climate Data Factory suivent cette logique ;
- Garantissant une traçabilité, gage d'expertise associée à la génération des données.



Modulo l'existence d'accords entre laboratoires et plateformes, les plateformes publient des jeux de données (flèche pourpre pointillée) avant publication, si la publication de ces jeux de données n'est pas prévue.

Le savoir-faire mis à disposition par **Données BRUTES** couvre toutes les étapes de la chaîne de valeur : il consiste à récupérer les données, les nettoyer, les enrichir, les centraliser, les analyser, les comprendre et les analyser. *Work With Data* collecte des données à partir de sources fiables et les structure dans un modèle unique. *Work With Data* les transforme en ressources que chacun peut utiliser grâce à une interface claire facilement exploitable.

Plusieurs étapes de la chaîne de valeur permettent d'engager un processus de valorisation économique



La donnée acquiert de la valeur par son partage, au moment où elle circule.

La valorisation multisectorielle consiste à agréger les jeux de données avec d'autres jeux de données, et à proposer des services marchands fondés sur ce rapprochement.

→ **C'est le modèle économique des plateformes,** explicité en infra.

- Par exemple, Kayrros intègre des images satellite, des données diverses et des informations de marché pour fournir
- des outils d'aide à la décision aux acteurs des secteurs de l'énergie, des ressources naturelles et de l'industrie,
 - des informations aux gouvernements, aux entreprises et aux fonds d'investissement pour les aider à réduire les émissions environnementales et gérer les risques liés au climat et à la transition énergétique.



Plateformes : description

Définition

- Ossature numérique qui permet de connecter collaborateurs, clients, partenaires, fournisseurs, objets... et aux uns et aux autres de partager leurs données.

Rôles

- Sur une plateforme, chacun peut être tour à tour producteur ou consommateur de données ou de services les exploitant.

Périmètre

- Une plateforme peut être interne à l'entreprise, privée, c'est-à-dire étendue aux fournisseurs, clients et partenaires, ou publique, fédérant autour d'elle tout un écosystème.

Compétition

- Plus la plateforme est vaste, possède de membres et collecte de données, plus ses algorithmes sont riches et puissants, et plus elle creuse l'écart avec ses concurrentes. Elles mènent donc une course effrénée à la taille critique d'où ne devrait émerger qu'une poignée de leaders.
- Les plateformes peuvent se rémunérer par abonnement, à la transaction ou en fonction du chiffre d'affaires induit.
- La pluridisciplinarité, la confiance et la technologie sont les clés de leur succès.

Source : Atos

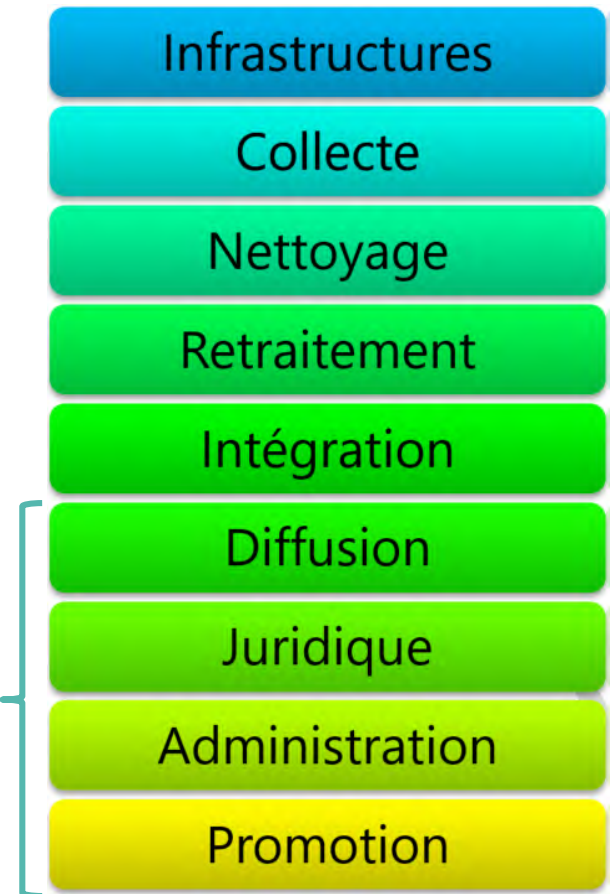
Positionnement des plateformes

La donnée en soi n'a pas grand intérêt, à part pour celui qui la collecte. C'est la mise en relation avec d'autres données qui peut révéler une nature tout à fait différente et une information inattendue, créant une nouvelle utilité et donc une valeur. Le modèle économique choisi doit concourir dans sa logique à cette création de valeur et se structure à partir des plateformes.

La plateforme peut être publique ou privée, l'accès aux données peut être libre ou restreint. La réexploitation des données n'oppose pas la gratuité à la licence.

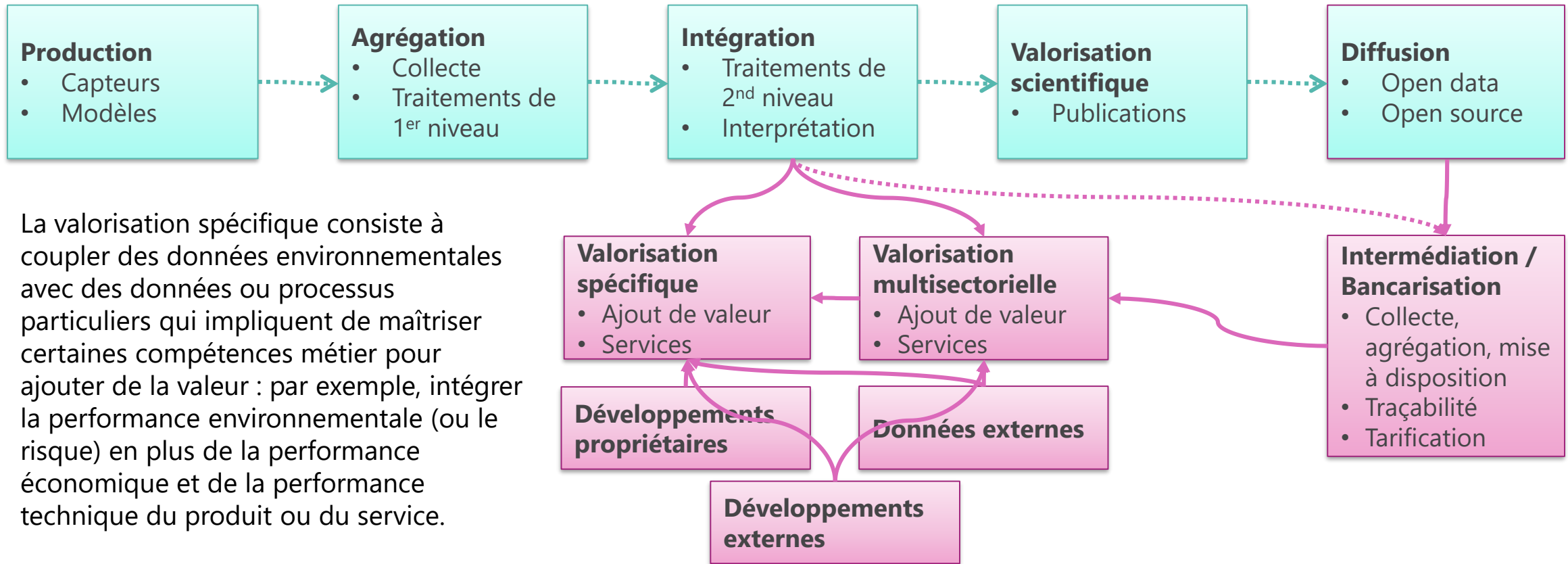
L'État ou un opérateur public participent au financement des premiers maillons : l'open data est assimilée à un bien commun.

Le nouveau paradigme des données ouvertes a permis l'émergence de nouveaux modèles commerciaux : les plateformes numériques. Ces dernières se sont positionnées avec succès en tant que parties prenantes uniques et inévitables entre l'offre et la demande. Elles incluent systématiquement les maillons à partir de la diffusion, dans la chaîne de coûts ci-contre. Leur modèle de développement est fondé sur les effets de réseau, ce qui signifie que la valeur de leur offre de service augmente de manière exponentielle avec la croissance du nombre d'utilisateurs. Les plateformes numériques ont largement contribué à favoriser l'innovation et le développement économique dans de nombreux pays.



Sources : Open data et business models ; [Quatre plateformes pour un SI](#). ZDNet, 2019-03-03

Plusieurs étapes de la chaîne de valeur permettent d'engager un processus de valorisation économique



La valorisation spécifique consiste à coupler des données environnementales avec des données ou processus particuliers qui impliquent de maîtriser certaines compétences métier pour ajouter de la valeur : par exemple, intégrer la performance environnementale (ou le risque) en plus de la performance économique et de la performance technique du produit ou du service.

Par exemple, Septeo Proptech agrège les données environnementales concernant le sous-sol (argiles, carrières, pollutions des sols...), les risques d'inondation, l'ensoleillement... mises à disposition par le BRGM et d'autres organismes de recherche, les croise avec des informations juridiques, administratives... pour construire et proposer des services aux notaires, agences immobilières ou aménageurs.

Données environnementales : un gradient de criticité pour les entreprises

Essentielles



Accessoires

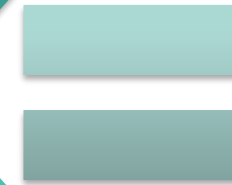


Les « communs numériques », une source de revenus ?

Préconisation : outils numériques développés par les administrations publiques = communs numériques par défaut



Secrétariat général à l'investissement public : intérêt pour les communs numériques



Déclarer les données et codes sources produits par la recherche en tant que communs numériques serait financé au niveau européen et apporterait « un levier essentiel pour la souveraineté européenne »,

La création du guichet unique de financement et d'une fondation européenne semblent acquises



Le rapport sur les communs numériques : un levier essentiel pour la **souveraineté** européenne - Min. des Affaires Étrangères et Min. de l'Économie



● **Modèles économiques**

Modèle économique : défini pour les entreprises

Un modèle économique explique comment l'entité gagne de l'argent avec son offre, quels sont les différents flux de revenus et en quoi sa manière de gagner de l'argent est cohérente avec son offre.



- La valorisation des organismes de recherche ne développe pas de modèle économique au sens strict, mais vise à augmenter la valeur produite pour les organismes en améliorant les interactions avec les entreprises.
- Les entreprises élaborent des modèles économiques qui intègrent assez souvent plusieurs flux de revenus.

Source : LanceTonIdee.com ; analyse AllEnvi Solutions

4 ressources pour créer un avantage concurrentiel

Ressources pour les entreprises :

- **ressources matérielles** spécifiques (puissance de calcul, stockage etc.),
- **compétences**,
- **données**,
- **algorithmes**.

Restreindre l'accès à au moins 1 ressource → **avantage concurrentiel**

- **Ressources matérielles** : accès banalisé
 - **Compétences** : pénurie, forte concurrence → ROI du recrutement
- sécuriser et restreindre l'accès aux **données** et/ou aux **algorithmes**.
- Pour les entreprises, le coût perçu du partage de ces actifs est principalement attribué à une perte d'avantage sur les concurrents.

Un éventail de possibilités entre accès ouvert et fermé

Les modalités d'accès aux données sont multiples et façonnent les relations partenariales.

Donner accès aux données, et/ou aux logiciels, n'est pas un choix binaire. Quelle que soit la structure, il est possible et nécessaire de décider quoi partager, avec qui et comment. Pour les données ouvertes, l'ouverture peut être graduée :

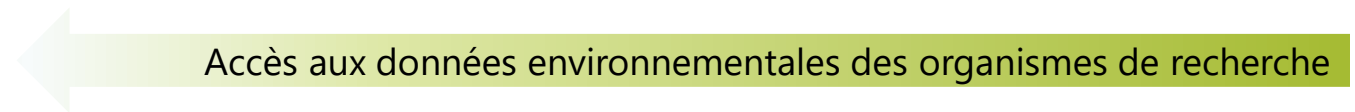


Accès	Interne	Nominatif	Groupes d'utilisateurs	Public	Libre
Contractualisation	Contrat de travail + politique interne	Accord spécifique	Authentification	Licence utilisateur	Licence ouverte
Exemple de données	Chiffre d'affaires	Permis de conduire	Recherche médicale	Fil Twitter	Horaires des bus

Fermé

Partagé

Ouvert



La réussite entrepreneuriale modifie son pilier économique : de l'hypercompétition à l'écosystème

Années 2000

Années 2010

Concurrence

Hypercompétition

Écosystème

Conséquences pour la valorisation

Utilisateur de données environnementales + membre d'un écosystème = contribution des données
« diluée »

→ plateforme « AllEnvi », pour fédérer un écosystème ?



● **Valorisations**

Moyens de protection

- via l'Agence de Protection des Programmes,
- « software heritage » (entrepôt du code + publication HAL qui décrit = équivalent métadonnées),
- preuve d'antériorité,
- data papers ou data software,
- licence GNU ou autre

Modalités de valorisation – vers des solutions types

Évaluer, au cas par cas, quelles données relèvent des informations publiques et quelles données n'en relèvent pas.
Valorisation **des** données → valorisation **avec** des données.

- Valoriser économiquement la valeur ajoutée obtenue grâce à **l'interprétation et au traitement** des données.
- Critères : l'origine du financement + Contrats d'Objectifs et de Performances
 - Données ou codes sources créés grâce au financement public = accès libre (ressources génériques)
 - Données ou codes sources créés et financés par acteurs privés = valorisables (ressources spécifiques)
 - Fonctionne mieux quand consortiums établis avant le projet ; gouvernance

Échanges de données, les entreprises partageant les données qu'elles obtiennent avec leurs partenaires de recherche, ce qui alimente les projets de recherche avec des données nombreuses, réelles et représentatives.

Ainsi, les données obtenues par les capteurs piézométriques du BRGM sur le niveau des nappes phréatiques permettent des prévisions de sécheresse ou d'inondations, prévisions actuellement vendues à Veolia et à des collectivités territoriales.

Cependant, si l'ouverture des données a été rendue obligatoire par la législation pour les entreprises publiques et la plupart des collectivités, il n'en va pas de même pour les applications : Uber, Waze ou Citymapper n'ouvrent pas leurs données mais pourtant se servent des données publiques existantes pour fonder leur modèle économique.

Sources : EGE ; BRGM

Le modèle de plateforme, solution avantageuse ?

Bénéfices d'un modèle économique fondé sur une plateforme publique

vecteur de modernisation de l'action publique :

- facilitation des échanges de données de recherche
- mutualisations techniques possibles (économies d'échelles)
- incitation à l'amélioration de la qualité du service, à la suite des demandes des réutilisateurs ou du fait d'éventuels mécanismes de retour financier

développement d'un écosystème :

- meilleure connaissance des usagers et des usages
- maîtrise de la donnée
- régulation des usages par les licences
- plateforme mobilisable par l'ensemble des services publics
- potentiels effets de traction du fait d'un modèle « freemium »
- gisement de « travail gratuit »

source de financement complémentaire (redevance, voire taxe)

contrepartie financière aux comportements susceptibles d'engorger l'infrastructure technique

frein à l'établissement de positions dominantes

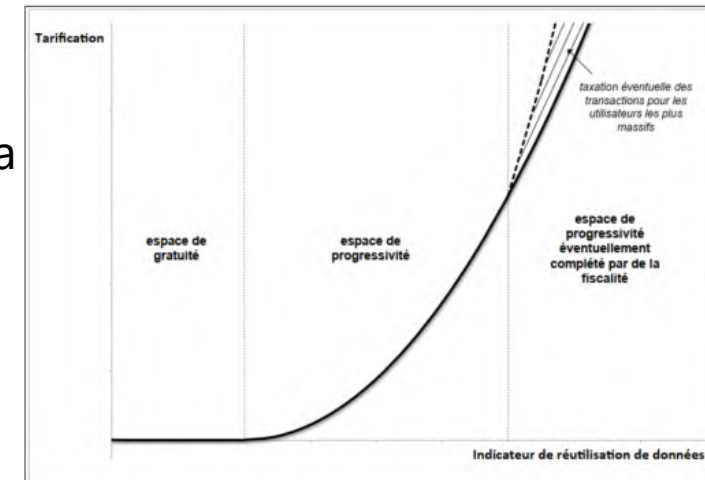
rémunération pour services rendus conforme à l'esprit et à la lettre de la loi organique relative aux lois de finances

Modèles de revenus pour le(s) plateforme(s) de données

Financements possibles :

- **Coûts de production** et de collecte des informations publiques : budget de l'État (dépenses permanentes du service public)
- **Coûts marginaux** inhérents à la diffusion des informations sur les plateformes : contribution prélevée sur la valeur que cette plateforme procure aux utilisateurs
 - **Redevance**, si elle ne constitue ni un frein à la réutilisation ni une barrière à l'entrée des réutilisateurs.
 - Modèles combinant **gratuité et tarification progressive**, selon le type de réutilisateurs et en fonction de critères correctement définis au regard de la valeur ajoutée par la plateforme, à privilégier.
 - Autres modes de couverture des coûts marginaux : contributions volontaires ou financements coopératifs : à expérimenter

Les organismes de recherche, éventuellement via l'action des infrastructures de recherche consacrées aux données environnementales, sont invités à construire un écosystème reposant sur une plateforme.



Source : rapport Trojette ; analyse AllEnvi Solutions



 **AllEnvi :
quelles valorisations ?**

Projets en lien avec la valorisation économique des données hébergées par Data Terra

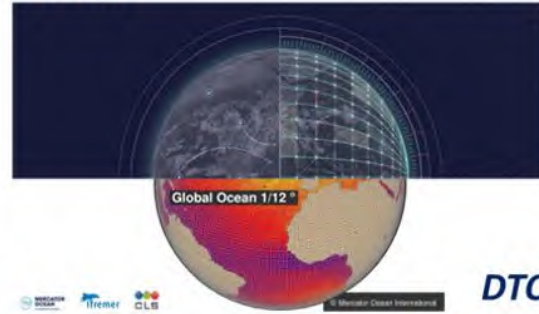
- Le [projet Gaïa Data](#) vise à permettre et faciliter la **mise en œuvre de cas d'usage transversaux**, c'est-à-dire impliquant des moyens informatiques et des données hétérogènes. Il s'agit de proposer un accès transparent et continu à l'ensemble des données multi-sources du système Terre et des services associés.
- Gaïa Data contient une batterie de services et d'outils permettant la découverte et l'utilisation des données, ainsi que la mise au point de nouveaux produits et traitements et leur exploitation systématique. Les 2 IR Data Terra et Climeri opéreront les services développés dans le cadre du projet Gaïa Data.
- La réponse aux grands enjeux **scientifiques** transversaux sera favorisée par le [projet Gaïa Data](#) (IR Data Terra / 4 pôles + IR PNDB + IR Climeri). Le budget attribué est de 16,2 M€. Le coût total du projet est de 62 M€ sur 8 ans. Il a débuté en juin 2021.
- [Destination Earth](#), alias DestinE, est une initiative de la Commission Européenne qui vise à créer une simulation numérique de la Terre, avec l'ambition d'aboutir à un jumeau numérique. Elle sera utilisée pour mieux comprendre les effets du changement climatique et des catastrophes environnementales et pour permettre aux décideurs politiques de réagir plus efficacement à ces problèmes. Une ambition réaliste de Gaïa Data est de contribuer à des briques qui pourront être intégrées dans des jumeaux numériques d'écosystèmes.
- La **combinaison de données spécifiques à certains besoins** est assurée par les bureaux d'études en charge de traiter le cas d'usage.



fondée sur la plateforme de Dawex dans le cadre du Plan de Relance, cette plateforme d'intermédiation débute son développement. Le plan France Relance la finance de 2021 à 2023. En 2022, elle propose majoritairement les données spatiales optiques (ni radar, thermique ou d'autres type), pas les données *in situ*. Ses liens avec Gaïa Data ne sont pas encore établis, le CNES y procédera ultérieurement.

Le schéma page suivante propose une articulation des données environnementales spatiales par rapport aux différents projets.

DestinE



DTO



DAWEX
Space Data Market Place



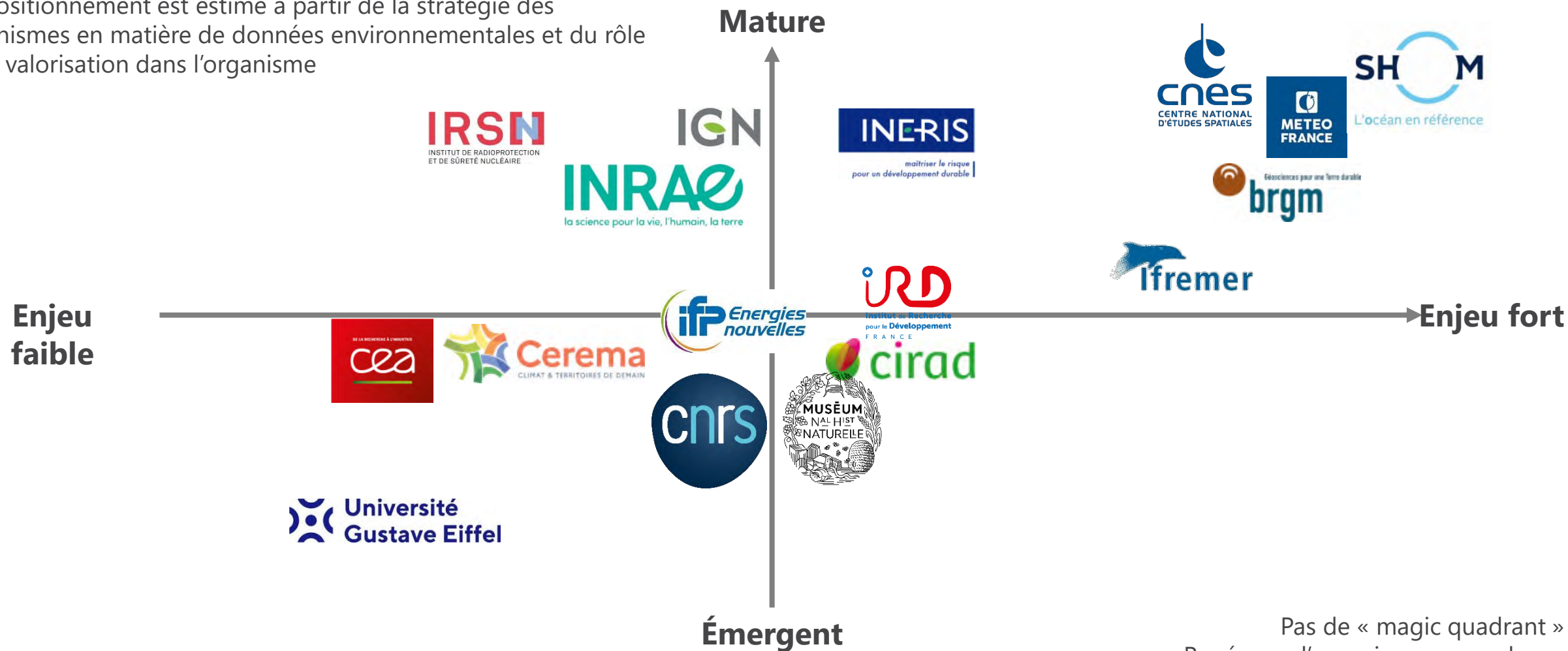
WEKEO

AIRBUS
DINAMIS
DATA TERRA
Theia
ForM@Ter
OBATIS
ERIS

© CNES Document Cnes soumis à la licence créative commons by-nc-sa 4.0

Positionnement relatif des organismes sur la valorisation des données environnementales

Ce positionnement est estimé à partir de la stratégie des organismes en matière de données environnementales et du rôle de la valorisation dans l'organisme



Pas de « magic quadrant »
Repérage d'organismes « proches »
dans leur démarche

Voies envisagées par les organismes

	Création d'activités spécifiques de valorisation liées à la diffusion de données	Plateformes de diffusion : mutualisation ?	Co-innovation avec des opérateurs du numérique	Maintien du système existant de valorisation
CEA		Data Terra + partage métadonnées		
Cirad		Oui, avec peu partenaires	Oui, voie principale (Orange)	
CNES	Oui + création continuum science-éco	Data Terra au moins	Oui, industriels hors ESN demandeurs	Oui, mais moyens déjà tous déployés
Ifremer	Oui, mais sans modèle économique	Oui ; d'abord pilotes	Oui, pilotes	Non
CNRS (INSU)	Oui	Oui ; clef = conviction chercheurs, pas ROI	Oui	Besoin d'une rupture, pas d'amélioration incrémentale
IGN	Oui	Pourquoi pas	Oui	Oui mais système à améliorer
Ineris	Non, pas économique	Si centralisation des demandes d'expertises	Non	Oui + ajout de services fondés sur l'IA (extraction, croisement)
Inrae	<ul style="list-style-type: none"> Données : libres et gratuites, pas de valorisation économique. Développement en cours de nouveaux modèles de valorisation des logiciels. 	Chaque communauté continue d'utiliser les plateformes du domaine	Oui + codévelopper des logiciels avec des privés pour répondre à leurs besoins propres => logiciels propriétaires	non, en lien avec la création de nouvelles activités. Logiciels libres par défaut.
IRD	Pourquoi pas	Mandat accordé	Non, services trop petits	
MNHN	Oui, essais en cours		Non	Oui
UGE	Un peu	OK de principe ; apport UGE = ?	Oui	Évolution en cours
Bilan	Oui	Mitigé	Plutôt oui	Au moins



Contact : François Christiaens

•

Tél. : 06 79 97 70 50

Email : francois.christiaens@inrae.fr

•

28, rue du Docteur Finlay

75015 Paris