



# EQUIPEMENTS ET INNOVATIONS DE POINTE

## ANALYSE STRATEGIQUE : BILAN ET RECOMMANDATIONS

Septembre 2025

# RAPPORT FINAL SYNTHÉTIQUE DE L'ÉVALUATION

Étude réalisée à la demande du Comité de surveillance des investissements d'avenir et du Secrétariat général pour l'investissement, dans le cadre d'un marché public de l'Agence Nationale de la Recherche, par GRAPHMYTECH, et pilotée par un comité de pilotage partenarial.





# Analyse stratégique des équipements et innovations de pointe : bilan et recommandations pour France 2030



**Rapport final**

11 sept. 2025



# TABLE DES MATIÈRES

<b>11 sept. 2025.....</b>	<b>1</b>
<b>INTRODUCTION.....</b>	<b>4</b>
<b>LES OBJECTIFS DES DEUX ÉTUDES.....</b>	<b>6</b>
1. Objectifs généraux.....	6
2. Périmètre initial défini par l'ANR dans le cadre de France 2030.....	6
3. Questions principales auxquelles le rapport répond.....	7
<b>DONNÉES ET PÉRIMÈTRE FINAL.....</b>	<b>9</b>
1. Nature et provenance des données clients traitées.....	9
Description du processus de sélection des données clients.....	9
2. Données externes.....	11
Processus de collecte des données externes.....	11
Volume final de données exploitées.....	11
<b>MÉTHODOLOGIE.....</b>	<b>12</b>
1. Construction des indicateurs.....	12
L'indicateur d'émergence scientifique.....	14
L'indicateur de frontière scientifique.....	15
2. Construction des référentiels de comparaison.....	16
Recherche d'un référentiel par intelligence artificielle.....	17
Intégration de la classification OpenAlex.....	17
3. Présentation des indicateurs.....	18
Les indicateurs structurants : cohérence et centralité.....	19
L'indicateur d'émergence scientifique.....	20
L'indicateur de frontière scientifique.....	21
L'indicateur de capacité d'absorption.....	22
Les indicateurs d'empreinte ANR dans le cadre de France 2030.....	23
L'empreinte spécifique des projets d'équipement.....	23
Synthèse de la logique de lecture des indicateurs.....	24
4. Limites et précautions d'interprétation.....	25
Les limites liées aux données.....	25
Les limites de la méthode.....	25
<b>PRINCIPAUX RÉSULTATS.....</b>	<b>27</b>
1. Les catégories de domaines.....	27
Cadran 1 : domaines proches de la frontière mais peu émergents.....	29
Cadran 2 : domaines émergents et à proximité de la frontière.....	29
Cadran 3 : domaines faiblement dynamiques et éloignés de la frontière.....	30
Cadran 4 : domaines émergents mais encore éloignés de la frontière.....	30
2. Les équipements par catégories de domaines.....	31
3. Empreinte scientifique.....	33

Empreinte scientifique de l'ANR dans le cadre de France 2030.....	33
Émergence scientifique.....	35
Carte des publications scientifiques par pays.....	37
4. Valorisation.....	38
Empreinte technologique de l'ANR dans le cadre de France 2030.....	39
Capacité d'absorption et identification des acteurs clés.....	41
Carte.....	43
5. Équipements.....	45
Empreinte des dépenses en équipement de l'ANR dans le cadre de France 2030.....	45
Empreinte Avant-Après.....	47
Les dépenses en équipement.....	49
<b>RECOMMANDATIONS.....</b>	<b>52</b>
1. Recommandations pour la collecte des données.....	52
Amélioration des formats de reporting des bénéficiaires.....	52
Normalisation des formats codes des sources.....	52
Intensifier les synergies avec la plateforme HAL.....	52
Création d'un outil de suivi au fil de l'eau.....	53
2. Recommandations opérationnelles et politiques.....	54
Orientations pour l'ANR dans le cadre de France 2030.....	54
Valorisation auprès des acteurs académiques et industriels.....	55
Recommandations par domaine.....	56
<b>CONCLUSION.....</b>	<b>58</b>
<b>ANNEXE.....</b>	<b>60</b>
1. Définitions des indicateurs.....	60
Indicateurs structurants (réseau).....	60
Indicateurs composites (dynamiques & positionnement).....	61
Capacité d'absorption.....	62
Indicateurs d'empreinte (contribution ANR dans le cadre de France 2030).....	62
2. Présentation de l'application.....	64
Statistiques descriptives.....	65
Analyse par domaine.....	68
<b>Références bibliographiques.....</b>	<b>71</b>

# INTRODUCTION

Dans le cadre du programme **France 2030**, l'Agence Nationale de la Recherche (ANR) a été chargée d'évaluer et de qualifier les dynamiques scientifiques et technologiques soutenues par les investissements d'avenir. Ces travaux s'inscrivent dans une démarche stratégique visant à mesurer l'impact des financements publics, à identifier les domaines où la France dispose d'un avantage compétitif, et à repérer les axes nécessitant un renforcement.

Ils prennent également place dans un **contexte international de compétition technologique intense**, marqué par des investissements massifs des États-Unis, de la Chine et de l'Union Européenne dans les infrastructures scientifiques et le soutien aux technologies de rupture. Situer les projets financés par l'ANR dans le cadre de France 2030 y compris les Programmes d'Investissements d'Avenir (PIA)<sup>1</sup> dans ce paysage mondial est donc essentiel pour mesurer leur portée stratégique.

Dans cette perspective, **GraphMyTech** a été mandaté pour la réalisation de deux études :

- « **Qualification et cartographie des technologies de pointe et des innovations de rupture financées par France 2030** », visant à identifier et analyser les technologies émergentes en les positionnant vis-à-vis de la frontière mondiale ;
- « **Bilan des équipements de recherche de taille intermédiaire - cartographie des équipements** », afin de situer le parc français d'équipements financés par les PIA et France 2030 par rapport à l'état de l'art international.

Ces études ont été menées en mobilisant un **cadre théorique et méthodologique rigoureux**, directement issu de la double expertise de GraphMyTech en **économie de l'innovation** et en **analyse de données scientifiques et technologiques**. L'approche adoptée combine l'**intelligence artificielle** (traitement du langage naturel, classification supervisée, détection des proximités sémantiques) et la **théorie des graphes** (mesure de cohérence, détection d'émergence, analyse de réseau), afin de structurer et d'interpréter des volumes massifs de données complexes.

Afin de fournir une **lecture intégrée des résultats**, il a été décidé de combiner ces deux études au sein d'une approche unifiée. Cette démarche présente l'avantage d'offrir une **vision transversale**, croisant à la fois les enjeux liés aux équipements scientifiques et ceux relatifs aux technologies émergentes.

Deux indicateurs centraux ont été retenus : la **frontière scientifique** et l'**émergence scientifique**. Ils s'appuient sur une **analyse structurale des réseaux de connaissances**, qui objective ces notions en étudiant les flux et les combinaisons de savoirs au sein des documents techniques et scientifiques. Dans ce rapport, une définition détaillée et méthodologique sera présentée, mais il convient d'en rappeler les principes :

- La frontière scientifique correspond au **niveau le plus avancé atteint à l'échelle mondiale** dans un domaine donné.

---

<sup>1</sup> dans la suite du rapport, France 2030 fera référence à France 2030 y compris les Programmes d'Investissements d'Avenir

- L'émergence scientifique désigne l'**accélération de la production de connaissances scientifiques dans un domaine** et son **impact croissant sur d'autres**.

Leur mobilisation permet de **cartographier les projets financés**, d'évaluer leur **positionnement par rapport à l'état de l'art international**, et de détecter les **dynamiques d'innovation en cours**. Ce choix répond à un double impératif : mesurer la capacité des projets à rapprocher la France de la frontière scientifique mondiale, et identifier ceux qui contribuent à l'émergence de nouveaux domaines scientifiques et techniques.

Au-delà du bilan factuel, ce rapport vise également à formuler des **recommandations opérationnelles** pour renforcer l'efficacité des futurs dispositifs de soutien. Une attention particulière est portée à la **collecte et à la structuration des données**, élément clé pour assurer la robustesse et la pérennité des analyses. GraphMyTech a veillé tout au long de l'étude à garantir la **neutralité de l'analyse**, la **transparence méthodologique** et la **robustesse scientifique** des résultats, afin d'apporter à l'ANR au Secrétariat Général Pour l'Investissement (SGPI) et au Ministère de l'Enseignement Supérieur et de la Recherche (MESR) un appui fiable et objectif pour la décision publique.

Enfin, le rapport est structuré comme suit :

1. Objectifs des études ;
2. Périmètre et données traitées ;
3. Méthodologie ;
4. Principaux résultats ;
5. Recommandations opérationnelles ;
6. Conclusion.

# LES OBJECTIFS DES DEUX ÉTUDES

## 1. Objectifs généraux

Les deux études confiées à GraphMyTech poursuivent deux objectifs, directement alignés sur les attentes formulées par l'ANR dans le cadre du programme France 2030. La première étude concerne la **qualification et la cartographie des technologies de pointe et des innovations de rupture** financées par France 2030. Il s'agit d'identifier les technologies émergentes effectivement soutenues par les investissements publics, de les positionner vis-à-vis de la frontière scientifique mondiale et de détecter les domaines porteurs d'émergence.

La seconde étude porte sur le **bilan des équipements de recherche de taille intermédiaire - cartographie des équipements**. L'enjeu est de dresser un état des lieux précis des infrastructures financées par les PIA et France 2030, de situer le parc français par rapport à l'état de l'art international et de mesurer leur contribution à la structuration des communautés scientifiques. L'évaluation doit aussi mettre en évidence l'apport de ces investissements en termes de compétitivité, d'attractivité et de synergies créées au sein de l'écosystème national de recherche et d'innovation.

Au-delà de ces deux études, leur combinaison au sein d'un seul et même rapport poursuit un objectif transversal : produire une **vision intégrée des dynamiques françaises**, en croisant l'analyse des infrastructures matérielles et celle des technologies émergentes. En s'appuyant sur les indicateurs de frontière scientifique et d'émergence scientifique, elle vise à cartographier les projets financés, à évaluer leur positionnement par rapport à l'état de l'art mondial et à mettre en lumière les dynamiques d'innovation en cours. Enfin, cette mission s'inscrit dans une perspective opérationnelle : elle ne se limite pas à la production de résultats factuels mais cherche à fournir des éléments stratégiques pour éclairer la décision publique, orienter les futurs financements, renforcer la valorisation des investissements réalisés et améliorer les dispositifs de collecte et de structuration des données, condition indispensable pour assurer la robustesse et la pérennité des analyses.

## 2. Périmètre initial défini par l'ANR dans le cadre de France 2030

Le périmètre des deux études recouvre deux volets complémentaires fixés par l'ANR, le SGPI et le MESR. Le premier concerne la **qualification et la cartographie des technologies de pointe et des innovations de rupture financées par France 2030**. Ce volet vise à identifier les technologies émergentes issues des projets soutenus et à évaluer leur positionnement par rapport aux dynamiques scientifiques et technologiques mondiales. Le second concerne le **bilan des équipements de recherche de taille intermédiaire - cartographie des équipements**. Il s'agit ici d'établir un état des lieux des infrastructures financées par les PIA et France 2030, de les situer par rapport à l'état de l'art international et d'apprécier leur contribution à la structuration des communautés scientifiques. L'articulation de ces deux volets doit permettre de dépasser une lecture en silos et de proposer une compréhension intégrée des investissements réalisés et de leurs effets.

### 3. Questions principales auxquelles le rapport répond

Les deux études conduites par GraphMyTech à la demande de l'ANR, le Secrétariat Général pour l'Investissement et le Ministère de la Recherche et de l'Enseignement supérieur s'inscrivent dans une double ambition stratégique : d'une part, dresser **un état des lieux approfondi des domaines technologiques** soutenus par les financements publics **dans le cadre de France 2030 et des PIA successifs** ; d'autre part, éclairer les décideurs publics sur les effets, **les dynamiques et les perspectives d'impact de ces investissements**, notamment en matière de structuration scientifique, de souveraineté technologique et d'innovation.

[Comment situer la recherche française sur les technologies émergentes et stratégiques dans un contexte international fortement concurrentiel ?](#)

L'analyse permet de positionner les domaines technologiques soutenus par l'ANR dans le cadre de France 2030 dans une cartographie comparative, à partir d'indicateurs croisant production scientifique, innovation technologique et dynamique d'émergence. En confrontant les thématiques couvertes aux grandes tendances de l'innovation mondiale, le rapport met en lumière le positionnement des domaines selon deux axes structurants :

- **La frontière scientifique** désigne ici le niveau de maturité scientifique atteint au niveau international dans un domaine donné. Elle permet d'évaluer dans quelle mesure la recherche française est en phase avec l'état de l'art mondial.
- **La centralité**, mesurée à partir des réseaux de citations, reflète quant à elle la **dynamique d'émergence** d'un domaine : un domaine fortement central occupe une position clé dans l'organisation des connaissances scientifiques et présente des signes de montée en puissance.

Cette approche permet de **cartographier de manière dynamique l'orientation de la recherche française**, d'identifier les opportunités de montée en gamme, et d'anticiper les domaines en transformation, au-delà d'une simple analyse de performance.

[Quels effets concrets les financements de l'ANR dans le cadre de France 2030 produisent-ils sur la production scientifique, l'innovation et la structuration des écosystèmes ?](#)

Le rapport évalue les **retombées des financements** à plusieurs niveaux : production directe des projets, renforcement des laboratoires bénéficiaires, diffusion au sein de l'écosystème national. Il permet **d'appréhender les effets de levier** associés aux investissements publics, en croisant dynamiques temporelles, indicateurs **d'empreinte scientifique et technologique**, et **structuration par domaine, action ou territoire, etc.** Ce regard croisé apporte des éléments tangibles sur l'efficacité des programmes de financement dans la durée.

[Quelle contribution des projets d'équipement à la montée en capacité de la recherche française ?](#)

Les équipements de recherche font l'objet d'une attention à part entière. Ils ont été analysés, non seulement comme **canal de financement**, mais aussi comme **catalyseurs de production scientifique**. Le rapport éclaire leur impact sur les dynamiques de publication, leur distribution territoriale, leur articulation avec les autres actions financées, ainsi que leur positionnement dans les chaînes de valeur technologiques. Il permet ainsi de **mieux comprendre les effets structurants des investissements en équipements sur l'ensemble de l'écosystème de recherche**.

## Comment évaluer le potentiel de transfert et la capacité de valorisation industrielle des résultats financés ?

Une attention particulière est portée à la relation entre la production scientifique et l'innovation technologique. Le rapport identifie, domaine par domaine, les résultats scientifiques **susceptibles d'être valorisés industriellement**, en analysant les publications de brevets déposés dans le cadre des projets ANR dans le cadre de France 2030 et les **capacités d'absorption des entreprises** dans des thématiques voisines. Il met ainsi en avant la capacité d'absorption associée aux différents domaines ainsi que la liste des entreprises les mieux positionnées pour les absorber. Cela fournit une **grille de lecture pertinente pour les politiques de transfert et de partenariat public - privé**.

## Quelle est l'efficacité des dépenses de l'ANR dans le cadre de France 2030 en fonction de leurs caractéristiques et de leur domaine scientifique ?

Le rapport permet d'évaluer **l'effet différencié des financements de l'ANR dans le cadre de France 2030** selon :

- Le **type d'action** (recherche, valorisation, équipement) ;
- Le **profil des projets** (durée, statut, domaine, etc.) ;
- Le **niveau d'investissement consenti** (financement ANR, part équipement, etc.).

En croisant ces paramètres avec les indicateurs d'empreinte scientifique et technologique, le rapport identifie des configurations plus efficaces que d'autres, tant du point de vue de la **valorisation des résultats** que de leur **structuration disciplinaire ou territoriale**.

Cela permet d'éclairer une question centrale des politiques publiques : **comment optimiser les leviers d'investissement selon les objectifs poursuivis ?**

Cette granularité dans l'analyse rend possible une lecture fine du **rendement marginal des dépenses publiques de recherche**, en fonction de leur contexte d'attribution, des domaines concernés, et des effets observables a posteriori.

## Comment assurer une continuité d'analyse et une appropriation des résultats par les services de l'État ?

Le rapport s'accompagne d'un **outil interactif** de visualisation et d'exploration des données, conçu pour faciliter la lecture, la comparaison et la réutilisation des résultats. Cette plateforme permet une **appropriation rapide par les services de l'État**, une mise à jour régulière des analyses, et une **capacité d'exploitation à différents niveaux** de granularité (par domaine, projet, action, territoire, etc.). Elle constitue un levier essentiel pour inscrire le rapport dans une démarche de pilotage, en lien avec les ambitions portées par le MESR et le SGPI.

# DONNÉES ET PÉRIMÈTRE FINAL

## 1. Nature et provenance des données clients traitées

Les données clients fournies concernent **1 277 projets** (projets "eotp"), **83 actions**, **quatre générations de PIA**, **deux types de projets** ('recherche' et 'valorisation'). Les dates de départ de projets s'étalent de **2011 à 2025**.

Ces données sont de deux natures; d'une part **des données structurées** regroupant les informations générales sur chaque projet (**articles, brevets, dépenses, dates, partenaires, pia, région, laboratoires**, etc.); d'autre part **des données non structurées** regroupant **des descriptifs techniques et scientifiques** des projets (**réponses d'appels à projets incluant les annexes techniques, comptes rendus intermédiaires et finaux**, etc.). Au final, **plus de 17 000 fichiers clients** ont été traités pour réaliser ces études.

### Description du processus de sélection des données clients

Une étape de consolidation de l'ensemble de ces données structurées et non structurées a été réalisée pour créer un périmètre homogène avec l'ensemble des projets fournis malgré une hétérogénéité des informations et des échelles d'analyse entre les projets 'recherche' et 'valorisation'.

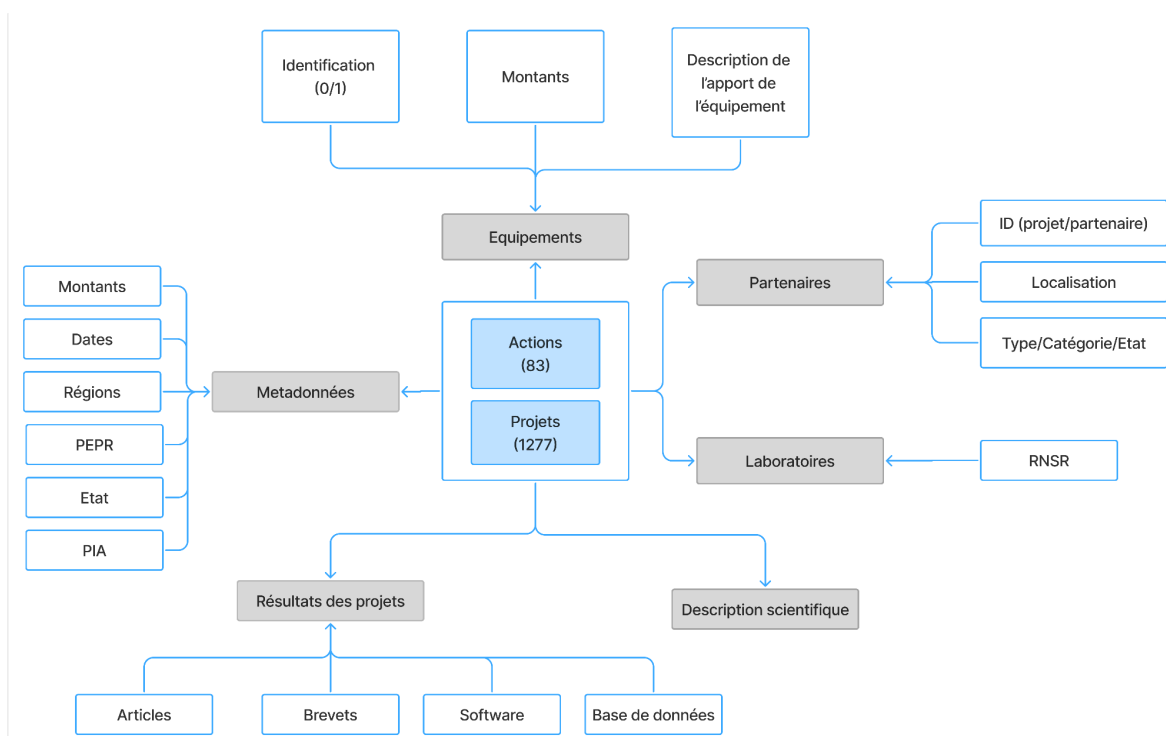


Figure 1: schéma de traitement des données

Le schéma présenté dans la figure 1 synthétise l'organisation du traitement des données mis en œuvre pour cette étude.

Au centre de l'analyse figurent les **projets** et **actions**, qui constituent les unités de base autour desquelles sont rassemblées les informations. Ces projets sont enrichis de **métadonnées structurées** (montants, dates, régions, PEPR, état, PIA), de **résultats de recherche** (articles, brevets, logiciels, bases de données), ainsi que des **descriptions scientifiques** extraites des documents non structurés.

Les **projets d'équipement** font l'objet d'un traitement distinct, avec des **éléments spécifiques** (identification binaire, montants, descriptifs).

Les entités **partenaires** (avec leur type, localisation et statut) et les **laboratoires** (via leurs identifiants **RNSR**) sont également rattachés aux projets, permettant un **croisement systématique entre acteurs, résultats et financements**.

Cette structuration permet d'aligner les données internes, hétérogènes et multi-sources, avec les jeux de données externes collectés (publications, brevets, bases bibliométriques), et de construire ainsi une **analyse multi-niveaux robuste**, à l'échelle des projets, des domaines, et des territoires.

Cette consolidation a mis en évidence :

- L'absence d'informations sur les **laboratoires financés dans le cadre des IRT/ITE**.
- Le petit nombre de **brevets avec des numéros standardisés** qui sont déclarés dans les projets 'recherche'.
- L'absence d'informations complémentaires sur les **résultats des projets (base de données, start-up, logiciels)** a été constatée pour les IRT.
- Pour les IRT, les **données non structurées (descriptions)** et **structurées (articles et brevets)** n'ont pas permis de distinguer les IRT lancés en 2010 de leur **renouvellement à partir de 2023** (23-AIRT-0005, 23-AIRT-0007, 24-AIRT-0001, 24-AIRT-0002). Ces projets ont donc été traités de manière agrégée.
- L'absence de **descriptions scientifiques ou techniques** liées aux projets "**France 2030 - Accélérer les recherches à risque (RRII)**" et aux projets **23-AIRT-0009** et **18-NUDD-0001**. Ils n'ont donc pas pu être pris en compte dans l'analyse.
- Les **données financières** n'ont pas pu être exploitées en raison d'un **manque d'homogénéité** entre les différents types de projets, notamment sur les **postes de dépenses**. En particulier, les **données financières sur les équipements** n'étaient pas disponibles pour les IRT et, pour les autres projets, leurs descriptions étaient **trop hétérogènes pour être exploitées**.
- L'asymétrie de renseignements des articles issus de projets entre les projets "recherche" et "valorisation" est importante. En effet, les publications issues des **projets "valorisation" ne représentent que 2% des articles étudiés**. Ainsi, l'étude d'impact scientifique réalisée dans ces deux études traduit principalement l'effet des projets "recherche".
- Concernant les données issues des **SATT**, **l'absence de renseignements systématiques et normalisés des projets issus de financements de l'ANR dans le cadre de France 2030 par leur identifiant EOTP** et plus généralement de l'ensemble des champs du fichier fourni (brevets, description du projet, numéro siret, ...) a été constatée. Cette hétérogénéité de remplissage à la fois qualitative et quantitative entre les différentes SATT n'a pas permis de prendre en compte ces données dans la présente analyse.

## 2. Données externes

### Processus de collecte des données externes

Pour mener à bien ces études de manière approfondie et obtenir des résultats pertinents, **deux types de données externes** sont collectés.

Premièrement, des données externes sont collectées afin de compléter celles fournies par le client :

- **Collecte des informations relatives aux publications scientifiques** issues des projets, i.e. les **résumés, dates, institutions, pays, citations, classifications OpenAlex** en utilisant les identifiants **DOI et OpenAlex**.
- **Collecte des publications des laboratoires financés par l'ANR dans le cadre de France 2030** en utilisant les **identifiants RNSR** pour requêter la **plateforme HAL**.

**Collecte des informations liées aux brevets** liés aux projets de l'ANR, après un **nettoyage humain des numéros déclarés**.

Cela a mis en évidence un **écart important** entre le nombre de brevets déclarés et les numéros effectivement fournis. Pour toute la période, **4 764 brevets ont été déclarés, 3 661 brevets ont effectivement été retrouvés** grâce à une **recherche hybride** utilisant **numéros fournis et déposants**.

Deuxièmement, des données externes permettant de construire un périmètre de comparables sont collectées :

- **Collecte des publications scientifiques comparables<sup>2</sup>**, à l'échelle mondiale, sur la période **2000 à aujourd'hui**, dans des **revues à comité de lecture**.
- **Collecte des brevets comparables**, en considérant l'ensemble des **brevets déposés dans le monde depuis 2000**,
- **Collecte des informations des entreprises appartenant à l'espace Schengen**. Nous nous concentrons pour cela sur les plus grandes entreprises dans le monde en termes de dépenses de R&D, représentant plus de 90% de ces dépenses dans le monde.

### Volume final de données exploitées

Dans l'ensemble, l'analyse porte sur :

- 1 277 projets financés par l'ANR,
- 131 315 publications,
- 724 familles de brevets,
- 247 domaines OpenAlex,
- 3 800 thématiques de recherche,
- 2 461 laboratoires de recherche impliqués,
- 628 568 publications scientifiques produites par ces laboratoires.

Par ailleurs, **212 projets d'équipement** ont été identifiés grâce à leur **code action (COHO, EQPX, INBS, ESRE, JZAY)**. L'analyse de l'empreinte scientifique s'appuie sur **plus de 12 millions d'articles scientifiques** dans le monde, dont **1,3 million issus d'institutions françaises**. L'étude de

---

<sup>2</sup> Les documents comparables sont extraits par proximité sémantique en utilisant les algorithmes de similarité développés par GraphMyTech

la valorisation technologique s'appuie sur **1,8 million de familles de brevets**, dont **plus de 100 000** déposées à l'**INPI**.

L'approche méthodologique adoptée pour ces études repose sur une articulation entre cadres théoriques issus de l'économie de l'innovation et outils avancés d'analyse de données scientifiques et technologiques. Elle vise à assurer à la fois la robustesse scientifique des résultats et leur pertinence opérationnelle pour l'évaluation des investissements de l'ANR dans le cadre de France 2030.

## 1. Construction des indicateurs

Notre démarche s'appuie sur une conception de l'innovation comme un **processus combinatoire** de connaissances. Dans cette perspective, les avancées technologiques résultent moins de la simple accumulation d'informations que de la **capacité à combiner des savoirs hétérogènes** pour générer de nouvelles solutions. Cette conception s'inscrit dans la lignée des travaux de Weitzman (1998) sur la croissance recombinaire, selon lesquels la valeur économique des idées réside moins dans leur nature intrinsèque que dans leur capacité à se combiner avec d'autres savoirs pour produire de nouvelles solutions. Fleming (2001) a prolongé cette intuition en montrant que l'incertitude recombinaire – c'est-à-dire le risque associé à la combinaison de connaissances hétérogènes – constitue une caractéristique centrale des processus d'innovation technologique, et que la qualité des combinaisons dépend étroitement de la proximité cognitive entre les savoirs mobilisés. L'analyse des réseaux de connaissances, construite à partir des publications scientifiques et des brevets, constitue dès lors un outil central pour objectiver ces dynamiques.

L'originalité de l'approche développée par GraphMyTech réside dans la distinction entre deux niveaux complémentaires :

- Les **briques de connaissances** : il s'agit des unités élémentaires de l'analyse, comprenant les concepts techniques, les publications, les brevets, mais aussi les infrastructures et les acteurs institutionnels. Ces briques sont représentées comme des **nœuds** du graphe.
- L'**architecture de connaissances** : elle correspond à la manière dont ces briques sont reliées, combinées et organisées au sein du graphe. Cette architecture se matérialise par les **arêtes** (cooccurrences, citations, collaborations) et par les sous-structures identifiées (communautés, clusters, sous-graphes).

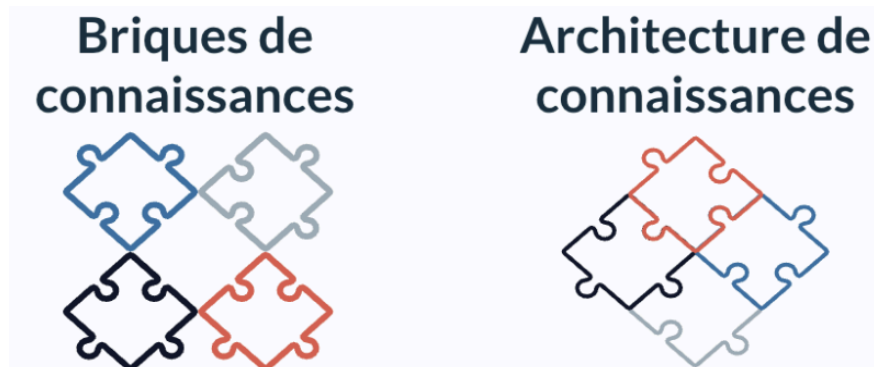


Figure 2: Briques et architecture de connaissances

Cette distinction permet de repérer à la fois les éléments constitutifs d'un domaine (briques) et leur organisation spécifique (architecture) - ce qui est essentiel pour comprendre si un domaine est en train de se structurer, d'émerger ou de se consolider.

Par ailleurs, l'approche de GraphMyTech combine deux dimensions complémentaires d'analyse :

- L'étude des **combinaisons de connaissances**, c'est-à-dire la manière dont les briques s'associent au sein des documents techniques et scientifiques. Elle met en évidence la **structuration interne** des domaines, en révélant leur cohérence, leurs recompositions et les signaux d'émergence.
- L'**analyse des flux de connaissances**, c'est-à-dire la manière dont les connaissances se diffusent via les publications scientifiques et brevets. Elle permet d'étudier les filiations technologiques ainsi que l'**influence d'une technologie ou d'une organisation dans la dynamique de développement des connaissances**.

La combinaison de ces deux perspectives, **combinaisons** et **flux**, offre une vision intégrée de la production de connaissances, articulant à la fois l'**architecture interne des domaines** et leur **ancrage dans l'écosystème global de l'innovation**.

Les indicateurs développés par GraphMyTech croisent ainsi l'analyse des **briques et de l'architecture** avec celle des **combinaisons** et des **flux**. In fine, cette approche permet le calcul de deux indicateurs fondamentaux :

- **La cohérence** : elle correspond au degré de structuration d'un domaine scientifique ou technologique. Dans notre approche, elle est mesurée par l'analyse des combinaisons de connaissances observées au sein des documents. Un domaine cohérent est caractérisé par des rapprochements fréquents et stables entre certains ensembles de savoirs, traduisant une trajectoire technologique dominante et stable. À l'inverse, un faible niveau de cohérence signale soit une fragmentation, soit une phase d'exploration technologique.

Cette conception de la cohérence s'enracine dans deux courants complémentaires de la littérature économique. D'une part, elle prolonge les travaux de Teece, Rumelt, Dosi et Winter (1994) sur la cohérence des firmes, qui montrent qu'**une organisation est cohérente lorsque ses activités reposent sur des ressources et des compétences partageables**, c'est-à-dire sur des savoirs dotés d'une forte relatedness. La notion de relatedness, initialement développée par Teece (1982) pour rendre compte des logiques de diversification des firmes multiproduits, désigne le degré de complémentarité ou de proximité cognitive entre des domaines de connaissance : **deux savoirs sont dits reliés (related) lorsqu'ils mobilisent des actifs, des routines ou des compétences similaires**, permettant ainsi des transferts et des synergies productives.

**Nesta et Saviotti (2006) appliquent ce concept aux connaissances technologique en utilisant les données de brevets** et en exploitant les codes de classification pour identifier les différentes briques technologiques auxquelles ils se rattachent. **Appliquée au niveau scientifique, la relatedness désigne la proximité thématique et méthodologique entre les savoirs mobilisés** dans un corpus de publications : un domaine scientifique cohérent est celui au sein duquel les mêmes combinaisons de concepts et de méthodes reviennent de façon stable et récurrente, traduisant l'existence d'une base cognitive commune.

D'autre part, cette conception de la cohérence comme stabilité des combinaisons de connaissances est directement articulée avec la notion de paradigme technologique développée par Dosi (1982), selon laquelle les trajectoires technologiques se structurent autour d'un ensemble stable de problèmes à résoudre et de règles heuristiques partagées par une communauté. Un domaine scientifique cohérent peut ainsi être interprété comme un domaine dans lequel un paradigme – au sens dosien du terme – s'est consolidé, c'est-à-dire

dans lequel les pratiques cognitives se sont stabilisées autour de combinaisons récurrentes de savoirs. Le glissement de la cohérence technologique vers la cohérence scientifique opère ainsi naturellement : les mêmes mécanismes de stabilisation des combinaisons de connaissances qui caractérisent la structuration d'un paradigme technologique se retrouvent dans la consolidation des domaines scientifiques, rendues observables par la récurrence des associations thématiques dans les corpus de publications.

- **La centralité** : elle traduit la position relative d'un domaine ou d'un acteur au sein des réseaux de citations et de références. La centralité est ici analysée à partir des flux de citations entre publications scientifiques, permettant d'identifier les savoirs, laboratoires ou entreprises occupant une position clé dans la circulation et la diffusion des connaissances. La centralité, au sens retenu ici, s'inscrit dans la tradition des analyses de réseaux sociaux initiée par Freeman (1979), qui a formalisé trois formes de centralité – de degré, d'intermédiarité et de proximité – **pour rendre compte des positions d'influence dans les réseaux**. Dans le contexte des réseaux de citations scientifiques, la centralité d'intermédiarité (betweenness centrality) est particulièrement pertinente : elle mesure la capacité d'un nœud – ici un domaine de connaissance – à se situer sur les chemins les plus courts entre d'autres nœuds, c'est-à-dire à jouer un rôle de pont ou de médiateur dans la circulation des savoirs (Borgatti et al., 2009). Un domaine central dans ce sens est un domaine qui ne se contente pas d'être productif, mais **qui structure les échanges entre d'autres domaines, jouant ainsi un rôle intégrateur dans l'écosystème global de la connaissance**.

### L'indicateur d'émergence scientifique

L'identification des domaines émergents repose sur un indicateur développé spécifiquement par GraphMyTech, fondé sur l'évolution temporelle des positions occupées par les différentes thématiques dans le réseau de connaissances. Cet indicateur n'est pas relatif aux publications issues des projets ANR dans le cadre de France 2030 mais bien **construit sur l'ensemble du domaine à l'échelle mondiale. Il permet donc de mesurer l'émergence relative d'un domaine dans le monde par rapport à l'ensemble des domaines étudiés** (247 domaines sur les 252 domaines que compte la classification OpenALex dans son ensemble, autrement dit il n'y a que 5 domaines au sein desquels on ne trouve aucune publication scientifique issue d'un projet financé par l'ANR dans le cadre de France 2030).

La notion d'émergence scientifique mobilisée ici s'inscrit dans le prolongement des travaux fondateurs de Dosi (1982) sur les paradigmes et les trajectoires technologiques. Dans ce cadre théorique, un paradigme technologique désigne un modèle de résolution de problèmes techniques, fondé sur des principes issus des sciences naturelles, qui oriente et contraint les efforts d'innovation d'une communauté d'acteurs. Une trajectoire technologique correspond alors aux progrès accomplis le long des dimensions définies par un paradigme. L'émergence d'un paradigme technologique se traduit, dans les réseaux de publications et de brevets, par une montée en centralité de certains domaines de connaissance qui deviennent progressivement des références incontournables pour les travaux ultérieurs.

Ce phénomène a été conceptualisé par Nelson et Winter (1982) à travers la notion de régimes technologiques, qui décrivent les conditions épistémiques dans lesquelles s'opère l'innovation. Freeman et Perez (1988) ont étendu cette perspective à l'échelle macroéconomique avec la notion de paradigme techno-économique, en soulignant que certaines technologies génériques deviennent des leviers de transformation de l'ensemble du système productif, se manifestant par une influence croissante et ramifiée dans les réseaux de connaissance. Le glissement du concept d'émergence technologique vers celui d'émergence scientifique repose sur une hypothèse théorique forte : les dynamiques de structuration de la connaissance scientifique obéissent à des logiques similaires à celles de l'innovation technologique.

De même qu'un paradigme technologique se repère à l'augmentation de l'influence d'un ensemble de solutions techniques dans les réseaux d'innovation, un domaine scientifique émergent se caractérise par une croissance rapide de sa centralité dans les réseaux de citations – signe qu'il acquiert un rôle structurant dans la circulation des connaissances. Kuhn (1962), bien qu'il travaillât à l'échelle de la science dans son ensemble, avait déjà souligné que l'émergence d'un paradigme scientifique se manifeste par la cristallisation progressive d'un réseau de références communes, qui finissent par constituer la matrice disciplinaire d'une communauté. C'est précisément ce phénomène que l'indicateur d'émergence développé par GraphMyTech vise à objectiver, en mesurant la variation de centralité d'un domaine dans les réseaux mondiaux de citations scientifiques.

Cet indicateur s'appuie sur la **variation de la centralité au cours du temps**, C'est-à-dire la manière dont un domaine renforce ou perd sa position dans les flux de citations et de références. La centralité désigne donc ici la capacité d'un domaine à se situer au cœur des flux de citations : plus il est cité par d'autres champs, plus il occupe une position centrale dans la circulation des connaissances. Une croissance rapide de la centralité traduit la capacité d'un champ à jouer un rôle de plus en plus structurant dans la circulation des connaissances.

Dans l'approche utilisée ici, la centralité capture la capacité d'un domaine à jouer un rôle d'intermédiation et de circulation des savoirs (centralité d'intermédiation). Sa variation dans le temps révèle des **phénomènes d'influence** dans les réseaux de citations. Lorsque la centralité d'un domaine progresse, il gagne en influence scientifique. C'est cette augmentation de l'influence d'un domaine qui constitue un **signal fort d'émergence**.

Néanmoins, afin de garantir la robustesse de la mesure, cette évolution est contrôlée par le **nombre de publications scientifiques associées au domaine**. Ce contrôle évite que des variations artificielles de centralité, liées à un très faible volume de documents dont l'influence serait très forte, ne soient interprétées à tort comme des signaux d'émergence.

Enfin, il convient de souligner que l'indicateur d'émergence développé par GraphMyTech est un **indicateur relatif**. Il ne s'agit pas de comparer directement les valeurs absolues de centralité entre domaines, mais d'évaluer leur **rang relatif** en fonction de deux dimensions :

- leur **variation de centralité** au cours du temps ;
- leur **volume de publications** observé dans la période considérée.

Un domaine est considéré comme émergent lorsqu'il se situe en **rang élevé sur les deux dimensions simultanément**. Cette approche présente plusieurs avantages :

- elle permet de détecter les domaines dont l'importance relative croît rapidement, indépendamment de leur niveau de départ ;
- elle limite les biais liés à des volumes documentaires trop faibles ;
- elle fournit une **lecture dynamique et comparative**, en identifiant les thématiques qui progressent le plus vite au sein de l'écosystème global.

En combinant ainsi variation de centralité et contrôle par le nombre de publications, l'indicateur d'émergence offre une vision robuste et opérationnelle des dynamiques en cours, permettant de détecter les trajectoires scientifiques et technologiques les plus prometteuses.

## L'indicateur de frontière scientifique

L'évaluation de la proximité d'un domaine avec la **frontière scientifique** repose sur une mesure comparative de sa **cohérence interne par rapport au niveau le plus avancé observé dans**

**l'écosystème mondial. La cohérence interne correspond à la cohérence scientifique des publications issues d'un financement ANR. Tandis que la cohérence mondiale est calculée pour l'ensemble des publications scientifiques appartenant au même domaine à l'échelle mondiale.**

Pour rappel, la cohérence correspond ici à la stabilité et à la répétition de certaines combinaisons de connaissances au sein d'un domaine donné. Elle reflète une consolidation : plus les mêmes regroupements de savoirs sont retrouvés à travers différents documents, plus la cohérence est forte.

L'indicateur est calculé à partir de **l'écart entre deux valeurs** :

- d'une part, le **score de cohérence de l'architecture de connaissances la plus structurée** identifiée au niveau international, qui définit la frontière scientifique dans le domaine considéré ;
- d'autre part, le **score de cohérence calculé pour les publications scientifiques et les productions associées aux projets financés par l'ANR**. Pour éviter les biais liés aux projets mobilisant trop peu de connaissances, l'indicateur n'est calculé que si les projets dans le domaine couvrent au moins 25% des connaissances du domaine.

La cohérence est ici mesurée par l'analyse structurale des réseaux de connaissances, en tenant compte de la densité et de la stabilité des combinaisons observées. Plus l'écart entre les deux valeurs est réduit, plus le domaine étudié se situe près de la frontière scientifique. À l'inverse, un écart important traduit un retard relatif ou une structuration encore incomplète.

Cet indicateur ne fournit donc pas une valeur absolue, mais une **mesure relative de proximité**. Il permet :

- de positionner les projets financés par l'ANR dans le cadre de France 2030 par rapport aux domaines les plus avancés à l'échelle mondiale ;
- d'identifier les champs dans lesquels la France est déjà proche de la frontière scientifique ;
- de repérer, à l'inverse, les domaines où un renforcement des efforts de recherche et d'investissement est nécessaire pour réduire l'écart observé.

En objectivant ainsi la distance à la frontière scientifique, l'indicateur offre un outil robuste pour guider les décisions stratégiques, en mettant en évidence les domaines où les financements publics permettent à la France de consolider son avance, et ceux où ils doivent contribuer à combler des retards.

**Les deux indicateurs d'émergence scientifique et de distance à la frontière sont donc construits à deux échelles distinctes.** L'indicateur **d'émergence à l'échelle mondiale** pour connaître la dynamique intrinsèque du domaine dans le monde et la distance à la frontière technologique est construite **à l'échelle des publications issues des financements ANR dans le cadre de France 2030** pour savoir comment ces dernières se positionnent vis-à-vis de l'état de l'art mondial.

## 2. Construction des référentiels de comparaison

Les indicateurs de cohérence, de centralité, de frontière scientifique et d'émergence reposent sur l'analyse structurale des connaissances.

## Recherche d'un référentiel par intelligence artificielle

Pour leur donner toute leur portée, ils doivent être appliqués à un **corpus homogène et comparable** reliant les productions des projets ANR dans le cadre de France 2030 à celles de la recherche et de l'innovation mondiale. La constitution de **référentiels français et mondial** répond précisément à cette exigence méthodologique.

Un élément central de notre approche réside ainsi dans la construction de ces **référentiels de comparaison**. Leur objectif est de situer les projets financés par l'ANR dans le cadre de France 2030 dans le contexte plus large de la production scientifique et technologique internationale, tout en permettant un positionnement relatif par rapport au paysage français.

Cette étape a reposé sur des techniques d'**analyse du langage naturel (NLP)** et d'**intelligence artificielle** appliquées à de très larges corpus documentaires. Les rapports et livrables fournis par les bénéficiaires ANR dans le cadre de France 2030 ont été traités automatiquement afin d'extraire, filtrer et d'analyser le **contenu sémantique** de chaque document.

Ces contenus ont ensuite été **vectorisés dans un espace multilingue (français-anglais)**, c'est-à-dire transformés en représentations numériques (embeddings) par des modèles d'IA capables de projeter les mots et concepts dans un **espace multidimensionnel**. Dans cet espace, deux documents proches sur le plan thématique se retrouvent à une distance réduite, ce qui permet de mesurer automatiquement leur proximité sémantique. Afin de compléter cette approche par proximité sémantique, un filtre sur les domaines scientifiques et technologiques a été appliqué afin de s'assurer de la pertinence des résultats et de leur alignement avec les thématiques réellement liées aux champs de recherche et d'innovation ciblés.

L'ensemble des documents des projets ANR dans le cadre de France 2030 ainsi vectorisés a été comparé à la **base de données propriétaire développée par GraphMyTech**, qui constitue un référentiel unique de brevets et de publications scientifiques vectorisés. Cette base couvre environ **50% des éditeurs de revues scientifiques** et **90% des brevets mondiaux**, offrant ainsi un cadre exhaustif et homogène. Elle permet une **comparabilité directe entre productions françaises et internationales**, sans biais de langue, et constitue un atout méthodologique distinctif de GraphMyTech.

Sur cette base, deux référentiels ont été construits selon une méthodologie identique, fondée sur la proximité sémantique avec les documents fournis par l'ANR :

- un **référentiel mondial**, constitué de l'ensemble des brevets et publications scientifiques similaires identifiés par rapport aux projets financés par l'ANR dans le cadre de France 2030 ;
- un **référentiel français**, correspondant au sous-ensemble du précédent, filtré pour ne retenir que les brevets et publications d'origine française.

Cette construction parallèle permet de comparer chaque domaine étudié à son **équivalent international**, mais aussi de mesurer la **position relative des projets ANR dans le cadre de France 2030 dans le paysage national**. En consolidant les données à ces deux échelles, l'étude met en évidence à la fois l'**ancrage français** des projets et leur **positionnement international**.

## Intégration de la classification OpenAlex

La constitution des référentiels français et mondial fournit la base documentaire nécessaire pour comparer les projets financés par l'ANR dans le cadre de France 2030 aux trajectoires scientifiques et technologiques globales. Toutefois, pour rendre cette comparaison pleinement opératoire, il est

indispensable de disposer d'une **classification homogène** permettant de regrouper publications, brevets et productions des projets dans des catégories cohérentes. C'est à cette fin que nous avons mobilisé la **classification OpenAlex**.

OpenAlex propose une classification hiérarchique de la production scientifique mondiale. Dans notre approche, cette classification a été utilisée pour **ordonner l'ensemble des résultats des projets ANR**, non seulement les **publications scientifiques**, mais également les **brevets** associés. L'alignement sémantique réalisé en amont grâce à l'IA a permis de rapprocher les brevets des catégories OpenAlex, assurant ainsi une intégration directe de la dimension technologique dans un cadre initialement conçu pour la recherche académique.

Cette étape présente un double avantage. D'une part, elle permet de disposer d'une **grille commune de lecture** couvrant à la fois les dynamiques scientifiques et technologiques, en plaçant publications et brevets dans les mêmes catégories. D'autre part, elle autorise une **consolidation des analyses au niveau des sous-domaines** (subfields), que nous désignons ici sous le terme de **domaines**.

Surtout, l'intégration de la classification OpenAlex permet une analyse **bottom-up** des résultats des projets. Au lieu de préassigner les projets à des thématiques prédéfinies, l'approche consiste à partir directement des **résultats concrets** (publications, brevets, livrables) pour déterminer les **domaines effectifs de contribution**. Cette démarche assure une objectivation maximale et reflète de manière fidèle la réalité des dynamiques scientifiques et technologiques observées.

Bien que les données sources soient nativement structurées par projet, il a été **décidé en comité de pilotage** d'adopter une logique **bottom-up**, c'est-à-dire que l'analyse part des résultats des projets (publications, brevets, livrables) pour **déterminer** à quel domaine ils contribuent. Ainsi, il en résulte qu'un même projet peut contribuer à plusieurs domaines ce qui traduit la nature **multidisciplinaire** de nombreux projets. Ainsi, les analyses présentées ici tiennent compte de cette relation **plusieurs-à-plusieurs** entre projets et domaines, sans imposer d'affectation exclusive.

En travaillant à l'échelle de ces domaines, il devient possible de :

- comparer systématiquement les productions françaises et internationales ;
- mesurer la proximité des projets ANR dans le cadre de France 2030 avec la frontière scientifique mondiale ;
- identifier les signaux d'émergence au sein de champs disciplinaires et technologiques bien définis.

L'intégration de la classification OpenAlex constitue donc une étape clé de la méthodologie. Elle assure la **cohérence sémantique et structurelle** de l'ensemble des analyses, en consolidant toutes les productions dans une même architecture, et en garantissant la robustesse des comparaisons réalisées entre domaines.

### 3. Présentation des indicateurs

Les indicateurs mobilisés dans cette étude ont été conçus pour objectiver les dynamiques scientifiques et technologiques associées aux projets financés par l'ANR dans le cadre de France 2030. Ils visent à répondre à plusieurs questions complémentaires : dans quels domaines les connaissances se structurent-elles ? Quels domaines gagnent en influence dans les réseaux scientifiques mondiaux ? Les projets financés se situent-ils à proximité de la frontière scientifique internationale ? Quelle est l'empreinte relative de l'ANR dans les domaines considérés ? Enfin, quels acteurs économiques apparaissent les mieux positionnés pour absorber ou valoriser les connaissances produites ?

Ces indicateurs reposent sur une approche structurale des connaissances. Ils combinent l'analyse des **combinaisons de connaissances**, mesurées à travers les cooccurrences de topics, et l'analyse des **flux de connaissances**, observés à partir des réseaux de citations scientifiques. Ils permettent ainsi de caractériser à la fois la structuration interne des domaines, leur dynamique d'évolution, leur positionnement relatif par rapport à l'état de l'art mondial et leur capacité à irriguer des acteurs ou des écosystèmes économiques.

Les résultats peuvent être exprimés sous plusieurs formes selon les indicateurs : scores normalisés, rangs, quantiles, variations relatives ou pourcentages. Ces formats n'ont pas tous la même signification. Un score traduit généralement une intensité relative sur une échelle donnée ; un rang ou un quantile permet de situer un domaine par rapport aux autres domaines analysés ; une variation relative mesure une évolution entre deux périodes ; un pourcentage correspond à une part observée dans un ensemble de référence clairement défini. La lecture des indicateurs suppose donc d'identifier systématiquement le dénominateur utilisé : ensemble des publications mondiales du domaine, publications françaises du domaine, brevets du domaine, productions issues des projets ANR, ou encore classement des entreprises dans un périmètre donné.

### Les indicateurs structurants : cohérence et centralité

Deux indicateurs fondamentaux structurent l'ensemble de la démarche : la cohérence et la centralité. Ils constituent les briques de base à partir desquelles sont construits les indicateurs composites d'émergence, de frontière scientifique et d'empreinte.

La **cohérence** mesure le degré de structuration interne d'un domaine scientifique ou technologique. Elle repose sur l'analyse des cooccurrences de connaissances identifiées dans les documents, en particulier les combinaisons de topics issus de la classification OpenAlex. L'objectif est d'évaluer dans quelle mesure certaines associations de connaissances apparaissent de manière stable, récurrente et significative au sein d'un domaine.

Sur le plan méthodologique, l'indicateur compare les cooccurrences observées à des cooccurrences attendues dans un contexte aléatoire. Cette comparaison permet d'identifier les combinaisons de connaissances qui ne relèvent pas simplement du hasard documentaire, mais qui traduisent une organisation cognitive spécifique. Le calcul est réalisé au niveau des combinaisons de topics OpenAlex, à partir d'un référentiel mondial de publications scientifiques construit sur l'ensemble de la période analysée.

La cohérence se lit comme un **score de structuration relative**. Un niveau élevé indique que les connaissances du domaine sont organisées autour de combinaisons récurrentes, ce qui traduit une trajectoire scientifique ou technologique relativement stabilisée. À l'inverse, un faible niveau de cohérence peut signaler une fragmentation, une phase d'exploration, une recombinaison du domaine ou une structuration encore incomplète. Il ne s'agit donc pas d'un pourcentage de publications, mais d'une mesure de la force des liens entre connaissances au sein d'un domaine. Cet indicateur doit toutefois être interprété avec prudence pour les domaines très jeunes, très transverses ou faiblement documentés, qui peuvent être mécaniquement sous-évalués.

La **centralité** mesure, quant à elle, l'influence d'un domaine ou d'un acteur au sein des flux de connaissances. Elle est calculée à partir des réseaux de citations entre publications scientifiques. Alors que la cohérence renseigne sur l'organisation interne d'un domaine, la centralité renseigne sur sa position relative dans l'écosystème scientifique global.

L'indicateur repose sur une représentation du réseau de citations sous forme de graphe, dans lequel les domaines, publications ou acteurs sont positionnés en fonction des relations de citation qui les

relient. La centralité est mesurée à partir du poids réel des citations attribuables à un domaine ou à un acteur dans l'ensemble du réseau. Elle constitue ainsi une mesure directe de l'influence scientifique, et non une simple estimation fondée sur le volume de publications.

La centralité se lit comme un **score d'influence dans le réseau de citations**. Une valeur élevée signifie qu'un domaine occupe une position structurante dans la circulation des connaissances : il est fortement relié à d'autres domaines et joue un rôle d'intermédiation dans les flux scientifiques. Là encore, il ne s'agit pas d'un pourcentage de publications, mais d'une mesure de position dans un réseau. Afin de mesurer les évolutions dans le temps, les calculs sont réalisés sur une base annuelle, puis analysés à partir de fenêtres glissantes de trois ans, ce qui permet de lisser les effets de bord liés au découpage temporel.

## L'indicateur d'émergence scientifique

L'indicateur d'**émergence** vise à identifier les domaines dont l'influence scientifique progresse rapidement et atteint un niveau significatif. Il ne mesure donc pas uniquement le volume de publications, mais permet de détecter les domaines qui gagnent en importance relative dans les réseaux mondiaux de connaissances.

Cet indicateur est construit à partir de la variation de centralité observée dans le temps. Une croissance rapide de la centralité traduit la montée en influence d'un domaine dans les réseaux de citations : le domaine devient davantage mobilisé, davantage cité ou davantage connecté à d'autres champs scientifiques. Cette progression constitue un signal d'émergence, dans la mesure où elle indique que le domaine occupe une place croissante dans la circulation internationale des connaissances.

Pour éviter les faux signaux liés à de très petits volumes documentaires, cette variation de centralité est contrôlée par le volume de publications du domaine. L'indicateur combine donc deux dimensions : le rang du domaine en matière de progression de centralité et son rang en matière de volume de publications. La restitution peut être effectuée sous forme de rangs ou de quantiles, afin de faciliter la comparaison entre domaines.

L'émergence se lit donc comme un **indicateur relatif de dynamique**. Un score ou un rang élevé signifie qu'un domaine progresse rapidement dans les réseaux mondiaux de citations tout en reposant sur un volume de publications suffisant. Lorsqu'il est exprimé en rang ou en quantile, il indique la position du domaine par rapport aux autres domaines analysés, et non une part de publications ou une probabilité d'émergence. Par exemple, un domaine situé dans les quantiles les plus élevés d'émergence fait partie des domaines dont la progression relative est la plus forte dans le périmètre étudié.

Un domaine est considéré comme émergent lorsqu'il combine une forte progression d'influence et un volume scientifique suffisant pour rendre cette dynamique robuste. Cette approche permet de repérer les trajectoires scientifiques dont l'importance relative s'accroît rapidement, tout en limitant les biais associés aux domaines trop petits ou trop instables.

L'indicateur reste toutefois sensible à la temporalité retenue. L'émergence n'est pas un processus linéaire : elle peut se traduire par des phases successives d'accélération, de stabilisation ou de ralentissement. Le choix d'une fenêtre de calcul suffisamment large est donc essentiel pour éviter de surestimer une dynamique ponctuelle ou, à l'inverse, de minimiser une trajectoire émergente en cours de consolidation.

## L'indicateur de frontière scientifique

L'indicateur de **frontière scientifique** mesure la proximité des projets financés par l'ANR dans le cadre de France 2030 avec l'état de l'art mondial d'un domaine donné. Il repose sur une comparaison entre la cohérence observée dans les productions associées aux projets ANR et la cohérence maximale observée ou calculable à l'échelle mondiale dans le même domaine.

La logique de cet indicateur est la suivante : plus les productions liées aux projets ANR mobilisent des combinaisons de connaissances proches de celles qui structurent le domaine au niveau mondial, plus elles peuvent être considérées comme proches de la frontière scientifique. À l'inverse, un écart important indique que les projets étudiés se situent dans une architecture de connaissances moins structurée ou plus éloignée des configurations les plus avancées observées internationalement.

L'indicateur est calculé à partir de l'écart entre deux valeurs. La première correspond à la cohérence de référence du domaine à l'échelle mondiale. Cette valeur représente l'architecture de connaissances la plus robuste, obtenue en conservant le minimum de liaisons nécessaires entre les connaissances structurantes du domaine. La seconde correspond à la cohérence observée pour les publications et productions associées aux projets financés par l'ANR dans le cadre de France 2030.

La distance à la frontière est alors mesurée par l'écart entre ces deux valeurs. Pour faciliter la lecture, le résultat est exprimé sous la forme d'un **score de proximité**, défini comme l'inverse de cette distance : plus le score est proche de 1, plus le domaine est proche de la frontière scientifique ; plus il est proche de 0, plus l'écart avec la frontière mondiale est important.

Il s'agit d'un score normalisé de proximité structurale, mesurant la distance des productions ANR par rapport à l'état de l'art mondial. **Ce score est normalisé** entre 0 et 1, où 0 représente l'architecture de connaissance la moins cohérente observée et 1 l'architecture la plus cohérente (définissant la frontière structurale du domaine). Il est crucial de ne pas l'interpréter comme un pourcentage de publications situées à la frontière. Par exemple, un score de 0,80 indique que l'architecture de connaissances observée dans les productions ANR est relativement proche de l'architecture de référence mondiale du domaine, sur l'échelle de proximité retenue, et non que 80% des publications ANR sont à la frontière scientifique.

Cet indicateur doit être interprété comme une mesure relative de maturité structurale, et non comme une mesure absolue de performance scientifique ou technologique. Il renseigne sur la proximité des combinaisons de connaissances mobilisées avec l'architecture scientifique mondiale la plus structurée. Il ne mesure donc ni la qualité intrinsèque des résultats, ni leur valeur économique directe, ni leur niveau de performance technologique.

Plusieurs précautions ont été intégrées au calcul. D'abord, un contrôle sur le volume de documents est appliqué afin d'éviter qu'un nombre très limité de publications, même très cohérentes, ne conduise à positionner artificiellement un domaine au plus près de la frontière. Ensuite, un seuil minimal de couverture des connaissances du domaine est retenu : l'indicateur n'est calculé que lorsque les projets ANR couvrent au moins 25% des connaissances identifiées dans le domaine. Ce seuil garantit que la mesure repose sur une base suffisamment représentative, et non sur quelques contributions isolées susceptibles de donner une image biaisée de la structuration réelle du domaine.

L'indicateur de frontière scientifique complète ainsi l'indicateur d'émergence. L'émergence caractérise la dynamique globale d'un domaine à l'échelle mondiale, tandis que la frontière scientifique permet d'évaluer le positionnement relatif des projets ANR dans ce domaine par rapport à l'état de l'art international.

## L'indicateur de capacité d'absorption

L'indicateur de **capacité d'absorption** vise à identifier les entreprises les mieux positionnées pour internaliser, valoriser ou prolonger les connaissances produites dans un domaine donné. Il répond à une question opérationnelle centrale : quels acteurs économiques disposent déjà d'une base technologique suffisamment proche du domaine étudié pour constituer des partenaires, des bénéficiaires ou des relais potentiels des résultats issus des projets financés dans le cadre de France 2030 ?

L'indicateur repose sur la comparaison entre les connaissances technologiques issues des brevets liés aux projets France 2030 et les connaissances technologiques associées aux portefeuilles de brevets de plus de 2 000 entreprises de référence. Cette comparaison permet d'évaluer la proximité entre un domaine scientifique ou technologique et les compétences déjà observables dans les activités d'innovation des entreprises.

Deux dimensions sont prises en compte. La première est un score de proximité entre l'entreprise et le domaine, fondé sur des proximités sémantiques et technologiques. Il permet d'identifier les entreprises dont les brevets, publications ou signaux publics présentent une proximité forte avec les connaissances du domaine étudié. La seconde est un score de cohérence de l'entreprise dans ce domaine, qui mesure dans quelle mesure les connaissances technologiques associées à l'entreprise sont elles-mêmes structurées et compatibles avec l'architecture de connaissances du domaine. Ces deux dimensions permettent de produire deux types de résultats complémentaires.

Le premier résultat est un classement des entreprises par domaine. Ce classement permet d'identifier les entreprises les mieux positionnées pour intégrer, exploiter ou valoriser les connaissances d'un domaine donné. Il est exprimé sous forme de rang : plus une entreprise est bien classée, plus elle apparaît comme un candidat pertinent pour l'absorption, le transfert, la collaboration ou la valorisation des connaissances du domaine. Ce rang est calculé en combinant le score de proximité technologique et le score de cohérence de l'entreprise.

Le second résultat est un taux de capacité d'absorption du domaine, exprimé en pourcentage. Ce taux indique la part des entreprises de la base de référence qui disposent d'une proximité suffisante avec le domaine pour pouvoir absorber tout ou partie des connaissances associées. Par exemple, un score de 85% signifie que 85% des entreprises de la base de référence présentent un niveau de proximité technologique suffisant pour être considérées comme capables d'absorber, au moins partiellement, les connaissances du domaine concerné. Ce pourcentage ne mesure donc pas la performance d'une entreprise particulière, mais la capacité de diffusion ou de valorisation potentielle du domaine au sein du tissu économique étudié.

L'indicateur peut être calculé à l'échelle française ou européenne, en incluant l'Union européenne ainsi que le Royaume-Uni, la Suisse, la Norvège et l'Islande. Il permet ainsi de repérer à la fois les entreprises les plus directement positionnées sur un domaine et les domaines dont les connaissances présentent le plus fort potentiel d'absorption par les acteurs économiques.

L'interprétation doit néanmoins tenir compte de la disponibilité variable des traces publiques selon les entreprises, les secteurs et les stratégies de publication ou de dépôt. Certains acteurs peuvent être sous-représentés lorsqu'ils publient peu, déposent peu de brevets ou protègent leurs connaissances par d'autres mécanismes, notamment le secret industriel.

## Les indicateurs d'empreinte ANR dans le cadre de France 2030

Les indicateurs d'**empreinte** visent à apprécier la contribution relative des projets financés par l'ANR dans le cadre de France 2030 à la production scientifique et technologique des domaines étudiés. Ils permettent de mesurer le poids des productions associées à ces projets dans un domaine, en les rapportant à des référentiels français ou mondiaux.

L'empreinte scientifique est calculée à partir de la part des publications scientifiques liées aux projets ANR dans l'ensemble des publications d'un domaine. Ces publications peuvent correspondre aux productions directement déclarées dans les rapports transmis à l'ANR par les porteurs de projets. Elles peuvent également inclure, pour les publications scientifiques, les productions appartenant au même domaine et publiées par des laboratoires financés par l'ANR dans le cadre de France 2030. L'empreinte technologique repose, quant à elle, sur les brevets associés aux projets ou aux acteurs concernés.

Ici, le pourcentage se lit bien comme une **part de production dans un ensemble de référence**. Plus précisément :

- lorsque l'empreinte scientifique est calculée sur le référentiel français, elle correspond à la part des publications françaises du domaine qui sont associées aux projets ANR dans le cadre de France 2030 ;
- lorsqu'elle est calculée sur le référentiel mondial, elle correspond à la part des publications mondiales du domaine qui sont associées aux projets ANR dans le cadre de France 2030 ;
- lorsque l'empreinte technologique est calculée sur les brevets, elle correspond à la part des familles de brevets du domaine associées aux projets, laboratoires ou acteurs liés à l'ANR dans le cadre de France 2030, selon le périmètre retenu.

Par exemple, une empreinte scientifique de 7,5% dans un domaine, calculée sur le référentiel français, signifie que 7,5% des publications françaises identifiées dans ce domaine sont associées à des projets financés par l'ANR dans le cadre de France 2030. Si le même indicateur est calculé sur le référentiel mondial, le dénominateur change : il s'agit alors de la part des publications mondiales du domaine attribuable aux productions ANR.

À titre de repère, la part de l'ANR dans le cadre de France 2030 dans les dépenses de la MIREs étant d'environ 5,5%, un résultat supérieur à ce seuil peut être interprété comme une contribution scientifique supérieure à la part budgétaire de référence. Autrement dit, il signale une surperformance relative du domaine ou des projets considérés, au regard du poids financier représenté par les financements concernés.

Cette mesure doit cependant être interprétée avec prudence. Les référentiels français et mondiaux sont construits à partir de l'analyse sémantique des publications et productions liées aux projets ANR. Les publications retenues sont celles qui présentent une forte proximité sémantique avec les documents des projets financés dans le cadre de France 2030. Cette méthode permet de construire un périmètre cohérent et comparable, mais elle peut aussi introduire un biais de sélection, susceptible de survaloriser légèrement la part relative des projets ANR dans l'ensemble des productions françaises ou mondiales.

## L'empreinte spécifique des projets d'équipement

De manière complémentaire, une analyse spécifique a été conduite sur le périmètre des projets ayant mobilisé des dépenses en équipements. L'objectif est d'évaluer plus finement l'effet de ces projets sur la structuration et la visibilité scientifique des domaines concernés.

L'indicateur d'**empreinte équipement** repose sur une analyse différentielle avant/après. Il compare l'évolution de deux dimensions principales : la centralité du domaine et le volume de publications associé. L'enjeu est d'apprécier si la mise en œuvre d'un équipement contribue à renforcer l'influence scientifique du domaine dans les réseaux de connaissances et à intensifier la production scientifique.

Une augmentation de la centralité après la mise en service de l'équipement peut être interprétée comme un gain d'influence scientifique. Une augmentation du volume de publications traduit une capacité accrue de mobilisation de la communauté scientifique autour de l'équipement. La combinaison des deux signale un effet structurant fort, dans la mesure où l'équipement contribue à la fois à accroître la production et à renforcer la position du domaine dans les réseaux de citations.

Les résultats peuvent être exprimés sous forme de variation relative ou de moyenne annualisée. Lorsqu'une variation est exprimée en pourcentage, elle correspond à l'évolution entre la période avant équipement et la période après équipement. Par exemple, une hausse de 20% du volume de publications signifie que la production scientifique moyenne observée après la date pivot est supérieure de 20% à celle observée avant cette date, sur le périmètre considéré. De même, une hausse de centralité indique que le domaine occupe une position plus influente dans les réseaux de citations après la mise en œuvre de l'équipement.

Le calcul est réalisé au niveau des domaines, en sélectionnant les projets explicitement associés à des dépenses d'équipement significatives, telles qu'identifiées dans les rapports d'exécution. Afin de tenir compte du délai entre l'investissement, l'installation effective de l'équipement et la production de résultats scientifiques publiables, la date pivot séparant la période « avant » et la période « après » est fixée à deux ans après la date de démarrage du projet. Les valeurs sont ensuite exprimées en moyenne annuelle afin de garantir la comparabilité des résultats, quelle que soit la profondeur temporelle disponible.

Cet indicateur doit être interprété en tenant compte de la temporalité propre aux équipements scientifiques. Leur effet peut être différé, notamment lorsque la montée en charge opérationnelle, la constitution des équipes utilisatrices ou la publication des résultats nécessitent plusieurs années. L'absence d'effet immédiat ne signifie donc pas nécessairement l'absence d'impact, mais peut traduire un horizon d'observation encore trop court.

## Synthèse de la logique de lecture des indicateurs

L'ensemble des indicateurs constitue un dispositif cohérent d'analyse des dynamiques scientifiques et technologiques. La cohérence et la centralité fournissent les mesures structurantes : la première renseigne sur l'organisation interne des connaissances, la seconde sur leur position dans les flux scientifiques. À partir de ces deux indicateurs fondamentaux, l'émergence mesure les domaines dont l'influence progresse rapidement, tandis que la frontière scientifique positionne les projets financés par l'ANR par rapport à l'architecture mondiale des connaissances.

Les indicateurs d'empreinte permettent ensuite d'apprécier la contribution relative de l'ANR dans le cadre de France 2030 à la production scientifique et technologique des domaines analysés. Enfin, la capacité d'absorption introduit une lecture tournée vers la valorisation, en identifiant les entreprises les plus susceptibles de prolonger, internaliser ou exploiter les connaissances produites.

## 4. Limites et précautions d'interprétation

### Les limites liées aux données

Comme toute démarche d'analyse structurale de données scientifiques et technologiques, la méthodologie adoptée comporte certaines **limites** qu'il convient de prendre en compte pour nuancer l'interprétation des résultats.

Tout d'abord, la **disponibilité et l'exhaustivité des données** peuvent varier sensiblement selon les domaines : certains champs sont largement couverts par les bases internationales et disposent de données de brevets et de publications scientifiques abondantes ; d'autres, en revanche, peuvent être **sous-représentés** ou peu standardisés, ce qui influe sur la robustesse des analyses. À cela s'ajoute le fait que les **stratégies de dépôt de brevets** varient considérablement selon les pays, les secteurs industriels et les cultures d'innovation, ce qui peut **biais**er la **perception comparative** des dynamiques d'un domaine.

Par ailleurs, les **données fournies en entrée** ne sont pas toujours homogènes d'un projet à l'autre : certains projets disposent de documents et rapports riches et structurés, tandis que d'autres présentent des lacunes ou une faible granularité. Cela peut introduire un **biais d'analyse** en faveur des projets les mieux documentés. En particulier, l'empreinte scientifique calculée repose principalement sur les données des projets "recherche" alors que l'empreinte technologique est traduite en grande majorité par celle des projets 'valorisation'.

S'agissant de la **temporalité**, les projets étudiés s'échelonnent entre 2011 et 2025. Deux **biais temporels contradictoires** doivent ici être soulignés :

- d'une part, un **biais favorable aux projets anciens**, pour lesquels le recul nécessaire permet de mieux observer les impacts et les publications associées ;
- d'autre part, une **amélioration progressive de la qualité de collecte des données** pour les projets plus récents, en lien avec l'évolution des procédures de suivi.

Ces effets contraires peuvent avoir un impact sur l'interprétation globale des tendances observées. De plus, les évolutions constatées peuvent également être influencées par des **dynamiques extérieures** aux projets (effets d'opportunité, évolutions institutionnelles, financements parallèles) qu'il convient de **mettre en perspective dans l'analyse**. L'analyse par indicateurs moyens, bien que synthétique, peut aussi **masquer des dynamiques internes hétérogènes**, en particulier dans les domaines multidimensionnels ou en recombinaison.

### Les limites de la méthode

Par ailleurs, plusieurs **limites propres aux méthodes mobilisées** méritent d'être mentionnées :

- Les méthodes d'analyse sémantique et de classification thématique reposent sur des modèles d'**intelligence artificielle**. Bien que entraînés sur de larges corpus, ces modèles peuvent refléter des **biais linguistiques, sectoriels ou disciplinaires difficilement mesurables**, et sous-représenter certains champs ou langues minoritaires.
- Un **décalage structurel** peut exister entre l'innovation effective (notamment en phase de prototypage ou de transfert) et sa **traduction observable** en publications ou en brevets. Certains projets d'innovation avancée peuvent ainsi être **sous-évalués** s'ils ne se traduisent pas immédiatement par une production scientifique ou technologique détectable.
- Les projets **interdisciplinaires**, souvent positionnés à l'interface de plusieurs domaines, peuvent **échapper aux grilles de lecture classiques** : leur contribution peut être diluée dans

des sous-ensembles techniques cloisonnés, ce qui complexifie l'identification de leur **portée réelle**.

- Enfin, certains **effets statistiques** peuvent altérer la mesure : les domaines de très petite taille sont plus sensibles au **bruit** ou aux **variations marginales**, tandis que les domaines très vastes peuvent être affectés par des **effets de seuil** qui atténuent les dynamiques réelles.

Ces limites sont toutefois **atténuées par plusieurs précautions méthodologiques** rigoureusement mises en œuvre à chaque étape de l'analyse.

D'abord, les études reposent sur l'utilisation d'une **base propriétaire internationale et multilingue**, articulée autour d'un moteur sémantique performant et d'un référentiel de classification homogène, permettant d'**aligner les données scientifiques et technologiques à l'échelle mondiale**. Cette approche garantit une couverture large et une comparabilité accrue entre domaines, tout en réduisant les biais induits par des jeux de données incomplets ou hétérogènes.

Ensuite, un **contrôle systématique du volume de publications** a été intégré dans la construction des indicateurs composites, notamment pour l'émergence. Ce choix méthodologique permet de **neutraliser les faux signaux** dans les domaines à faible production, en assurant que les dynamiques identifiées reposent sur des bases empiriques significatives.

De plus, la méthodologie repose sur une **analyse croisée des sources** : publications scientifiques, brevets, rapports et livrables issus des projets financés par l'ANR. Cette pluralité de sources permet de **réconcilier les temporalités différenciées de l'innovation** (court terme pour les publications, moyen terme pour les brevets, long terme pour les effets socio-économiques), tout en **réduisant le risque de surpondération d'un canal de production de connaissance au détriment d'un autre**.

S'agissant des effets temporels, une **fenêtre d'observation étendue** a été retenue afin de lisser les effets de bord, en particulier ceux liés à la montée en charge différée de certains projets (notamment les projets d'équipement ou les projets interdisciplinaires). Des corrections ont également été intégrées pour **comparer les indicateurs entre domaines indépendamment de l'année de démarrage des projets**, notamment via des mesures moyennes annualisées ou des indicateurs différenciés (par exemple, variation de centralité avant/après pour les équipements).

Enfin, la démarche accorde une attention particulière à la **structuration des connaissances** : les indicateurs mobilisés (cohérence, centralité, émergence, frontière scientifique) ont été définis et calculés à partir de **modèles en théorie des graphes** et de **benchmarks technologiques internationaux**, assurant une stabilité et une légitimité méthodologique. Ces indicateurs ont été calibrés de façon à **refléter à la fois les dynamiques internes des projets ANR dans le cadre de France 2030 et leur position dans les réseaux mondiaux de production scientifique**.

La **multidisciplinarité**, source fréquente de bruit statistique dans les analyses classiques, est ici explicitement intégrée : un projet peut être associé à plusieurs domaines, et l'agrégation n'intervient qu'après une lecture par domaine, ce qui permet de préserver la finesse d'analyse. Dans la suite du rapport, cette méthodologie est mobilisée pour :

- **cartographier les dynamiques françaises**, en croisant infrastructures et technologies ;
- **situer les projets par rapport à la frontière scientifique mondiale** ;
- **identifier les signaux d'émergence** susceptibles de structurer de nouvelles trajectoires.

L'analyse des résultats qui suit doit donc être comprise comme l'application opérationnelle des choix méthodologiques présentés ici.

## 1. Les catégories de domaines

Afin de présenter les différents domaines, une matrice d'analyse stratégique a été réalisée. Elle permet de répartir les domaines en quatre cadrans selon deux axes complémentaires : d'une part, l'**émergence** (variation de centralité contrôlée par le volume de publications) ; d'autre part, la **proximité à la frontière scientifique** (score défini comme 1- distance à la frontière).

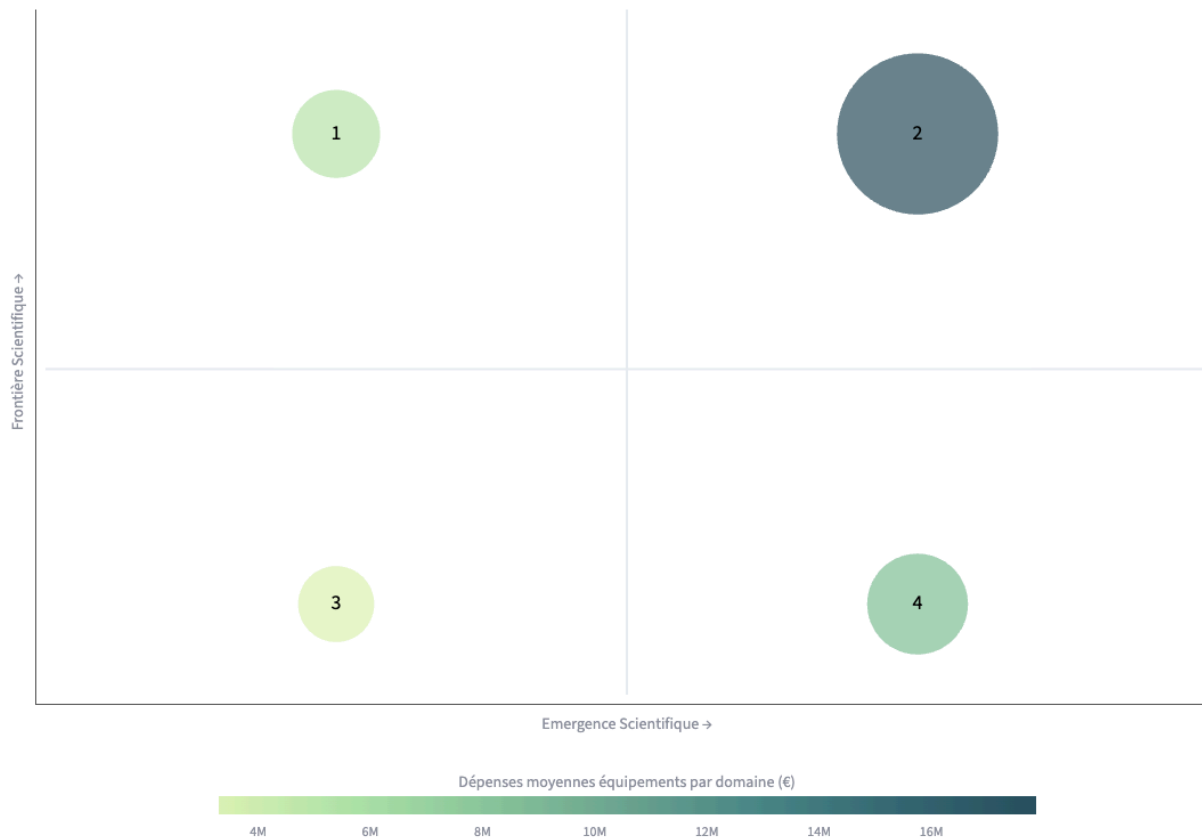


Figure 3: Les quatre catégories (cadrans) de domaine

Lecture: Plus la couleur du rond est foncée, plus les dépenses d'équipement moyennes par domaine sont importantes. Plus la taille du rond est grande, plus les dépenses ANR dans le cadre de France 2030 en moyenne par domaine sont importantes. Plus un domaine se situe en haut, plus il est proche de l'état de l'art mondial ; plus il se situe à droite, plus sa dynamique d'émergence est forte.

Dans cette grille de lecture, les domaines se positionnent comme suit :

- **Cadran 1** (faible émergence / forte proximité à la frontière) : domaines de **maturité élevée**, générateurs de performances à court terme même si la dynamique d'ascension est moindre.

- Orientation** : capitaliser (standardisation, transfert, mise à l'échelle, valorisation socio-économique, indicateurs d'empreinte).
- **Cadran 2** (forte émergence / forte proximité à la frontière) : domaines à la fois dynamiques et avancés, susceptibles de structurer des chaînes de valeur et d'entraîner des effets d'entraînement rapides. **Orientation** : accélérer de manière maîtrisée (renforcement des moyens, sécurisation des compétences, partenariats industriels et internationaux, diffusion des résultats).
  - **Cadran 3** (faible émergence / faible proximité à la frontière) : domaines peu dynamiques et éloignés de la frontière. **Orientation** : reconfiguration (réorientation thématique, rapprochements inter-domaines) ou sortie graduelle.
  - **Cadran 4** (forte émergence / faible proximité à la frontière) : domaines **prometteurs** mais encore distants de l'état de l'art. **Orientation** : arbitrer entre un investissement élevé (équipements, compétences, coopérations ciblées) pour accélérer la montée en gamme, et une temporisation conditionnée à des jalons (preuves d'intégration technologique, validation scientifique) pour vérifier le rattrapage technologique.

Cette représentation propose une lecture intégrée de la **maturité** (proximité à la frontière scientifique) et de la **dynamique** (émergence), évitant de **confondre stabilité et potentiel**. Elle facilite la **priorisation des ressources** sur la base de critères objectifs. Mise en regard des autres indicateurs calculés à l'échelle des domaines, **dépenses moyennes ANR dans le cadre de France 2030 par domaine, nombre moyen de projets, cohérence moyenne, empreinte scientifique moyenne de l'ANR dans le cadre de France 2030**, etc. elle permet d'apprécier, **par grande catégorie de projets, les performances et les moyens mobilisés** dans les différents domaines.

Les seuils sont **relatifs** et définis en fonction de la valeur médiane des 2 indicateurs qui structurent les 2 axes. L'émergence n'étant pas linéaire, une **fenêtre temporelle suffisante** est retenue pour lisser les phases d'accélération et de plateau. Par ailleurs, la **multidisciplinarité** est expressément prise en compte : un projet peut contribuer à plusieurs domaines, et la lecture par domaine précède toute agrégation. Enfin, la matrice n'a pas de portée normative en soi : des considérations **territoriales, sociétales ou de souveraineté** peuvent justifier des décisions distinctes de celles qu'induit la seule position dans la matrice.

La répartition des domaines au sein des quatre cadrans met en lumière une diversité de trajectoires scientifiques et technologiques, tout en révélant certaines constantes structurantes. Sur l'ensemble des projets analysés, les domaines les plus représentés sont ceux des **cadrans les plus émergents** (cadran 2 et cadran 4), qui concentrent à eux seuls **près de 60% des projets**, des publications, et des investissements.

La **centralité moyenne**, indicateur de l'influence dans les réseaux de citations, suit une courbe clairement différenciée selon les cadrans : elle est **près de 15 fois plus élevée dans les domaines du cadran 2 que dans les domaines du cadran 3**, traduisant des écarts significatifs en termes d'ancrage dans les flux de connaissances. Ce différentiel d'influence se retrouve également dans les indicateurs d'empreinte scientifique et technologique ANR dans le cadre de France 2030 (empreinte scientifique ANR dans le cadre de France 2030 dans les domaines du cadran 2 : 7,6% contre 6,6% dans les domaines du cadran 3 en moyenne).

Les **niveaux de cohérence**, qui mesurent la structuration interne des domaines, présentent des valeurs contre-intuitives dans certains cas : les domaines du cadran 3, bien que peu dynamiques, affichent des scores de cohérence élevés, conséquence possible d'une faible recombinaison thématique.

## Cadran 1 : domaines proches de la frontière mais peu émergents

Ces domaines, bien qu'affichant une dynamique d'émergence modérée, sont **solidement positionnés** à proximité de la frontière scientifique. Ils regroupent **759 projets**, pour **30 522 publications** et **111 brevets**, soit en moyenne **40 publications** et **0,15 brevet par projet** ; un rendement correct mais nettement en retrait par rapport aux cadran 2.

Les **dépenses totales ANR dans le cadre de France 2030** atteignent **1,47 milliard d'euros**, soit une moyenne de 1,94 M€ par projet et **26,8 M€ par domaine**. Ces montants plus modérés permettent **un engagement moins significatif des dépenses ANR dans le cadre de France 2030** dans ces domaines plus matures. Ces domaines comptent tout de même **186 projets PEPR**, **67 actions distinctes** et sont répartis sur **14 régions**.

Sur le plan des indicateurs, ces domaines conservent une **empreinte scientifique élevée (7,4%)**, bien qu'inférieure à celle des cadran 2, et une **empreinte technologique** plus contenue (1,8%). Leur centralité (0,016) est également nettement plus faible, mais cohérente avec leur profil de domaines moins dynamique que les domaines du cadran 2.

Les domaines du cadran 1 constituent ainsi des **domaines de stabilité** : ils offrent des résultats consolidés et visibles à court terme, mais avec un potentiel de transformation plus limité. Leur valorisation passe par la standardisation, la diffusion et la capitalisation.

## Cadran 2 : domaines émergents et à proximité de la frontière

Les domaines classés en cadran 2 concentrent la majorité de l'effort scientifique et technologique soutenu par l'ANR. Ils regroupent **1 165 projets**, générant **96 941 publications** et **686 brevets**, soit en moyenne **83 publications** et **0,59 brevet par projet**. Ces chiffres témoignent d'un **niveau d'activité et de valorisation particulièrement élevé**.

La **dépense totale ANR dans le cadre de France 2030** dans ces domaines atteint **6,2 milliards d'euros**, soit 5,3 M€ par projet et **89,8 M€ par domaine** en moyenne, soit des montants nettement supérieurs à ceux des autres cadrans. Ils mobilisent également un **nombre élevé de projets PEPR (444)** et sont associés à **80 actions distinctes**, réparties sur **16 régions**, soulignant leur rôle moteur dans la structuration nationale de l'innovation.

Ces domaines se distinguent aussi par leur **centralité scientifique (0,099)** et leur **empreinte technologique moyen (3,2%)**, confirmant leur influence dans les réseaux de production et de circulation des connaissances. L'**empreinte scientifique moyenne des projets ANR dans le cadre de France 2030** y est aussi la plus élevée (7,6%), traduisant une reconnaissance forte des publications issues des projets soutenus.

En résumé, les domaines du cadran 2 représentent un **noyau stratégique** de domaines déjà structurés, visibles à l'échelle internationale, et susceptibles de produire des effets d'entraînement rapides. Leur pilotage doit viser la consolidation des acquis, la sécurisation des compétences, et l'amplification des retombées.

### Cadran 3 : domaines faiblement dynamiques et éloignés de la frontière

Ce cadran regroupe **660 projets**, pour **16 958 publications** et seulement **37 brevets**, soit une moyenne de **26 publications** et **0,06 brevet par projet**, traduisant une productivité scientifique et technologique nettement plus faible que dans les autres catégories.

Les **dépenses ANR dans le cadre de France 2030** s'élèvent à **1,35 milliard d'euros**, soit une **moyenne de 2,05 M€ par projet** et **30 M€ par domaine**, C'est-à-dire le plus faible niveau d'investissement de toutes les catégories de domaine. On recense toutefois **134 projets PEPR** et **62 actions**, répartis sur **15 régions**, ce qui montre que ces domaines continuent de bénéficier d'un certain soutien institutionnel large mais plus émiété que pour les autres catégories.

Bien que leur **cohérence moyenne** apparaisse élevée (0,321), cela peut s'expliquer par une spécialisation marquée ou un manque de recombinaisons thématiques. Leur **empreintes scientifique (6,6%)** et technologique (1,07%) sont les plus faibles, tandis que leur **centralité (0,0065)** reste marginale en adéquation avec le profil du domaine.

Ces domaines requièrent une **réévaluation stratégique**. Certains peuvent bénéficier d'une reconfiguration (approche interdisciplinaire, rapprochements thématiques), tandis que d'autres pourraient progressivement sortir du périmètre prioritaire de financement.

### Cadran 4 : domaines émergents mais encore éloignés de la frontière

Ce cadran regroupe **826 projets**, générant **31 167 publications** et **159 brevets**, soit une moyenne de **37,7 publications** et **0,19 brevet par projet**. Ces domaines présentent un rendement similaire à celui du cadran 1.

Les **dépenses ANR dans le cadre de France 2030** atteignent **1,93 milliard d'euros**, avec une **moyenne de 2,34 M€ par projet** et **35 M€ par domaine**. Ces montants indiquent un **engagement significatif** mais raisonné, à mi-chemin entre prise de risque et consolidation. Ils impliquent un nombre important de projets **PEPR (234)**, **69 actions** mais sont répartis sur seulement **13 régions**.

Les indicateurs qualitatifs (centralité : 0,0366 ; empreinte technologique : 2,28% ; empreinte scientifique : 7,1%) positionnent ces domaines comme en structuration. **L'empreinte technologique relativement forte montre en revanche l'attrait de l'écosystème d'innovation pour ces domaines** ; leur potentiel est réel, mais leur stabilisation dépendra de l'atteinte de jalons critiques (rapprochement de la frontière scientifique, validation technologique et scientifique, structuration cohérente de communautés).

Ces domaines appellent à une **stratégie différenciée** : dans certains cas, l'investissement plus important se justifiera si l'émergence reste forte et qu'un rapprochement de la frontière scientifique s'opère tandis que dans d'autres, un retrait pourra se justifier si aucun rapprochement de la frontière scientifique ne s'observe).

En conclusion, l'analyse croisée entre **proximité à la frontière scientifique** et **dynamique d'émergence** permet de dépasser une lecture purement sectorielle pour proposer une **typologie opérationnelle** des domaines financés par l'ANR. Cette typologie en cadrans offre une **grille d'évaluation stratégique** permettant d'adapter les modalités de soutien public en fonction du profil réel des domaines.

- Les **projets du cadran 1**, matures mais peu émergents, affichent des rendements scientifiques corrects mais **des performances d'innovation plus modestes pouvant être accentués compte tenu de la maturité des domaines**. Leur soutien doit viser à valoriser les acquis, diffuser les résultats et encourager une exploitation plus systématique, notamment via des partenariats industriels ou territoriaux.
- Les **projets du cadran 2**, à fort impact et fort dynamisme, justifient une **politique de consolidation des acquis**, de **sécurisation des compétences clés**, et de **massification des retombées** scientifiques, technologiques et industrielles. Ils présentent également les **meilleurs rendements** en matière de publications et de brevets rapportés aux montants investis, confirmant leur statut de domaines à effet d'entraînement rapide et à forte efficacité.
- Les **projets du cadran 3**, faiblement dynamiques, présentent les **rendements les plus faibles**, tant sur le plan scientifique que technologique. Néanmoins, certains de ces domaines pourraient regagner en pertinence par des **efforts de reconfiguration** (approches interdisciplinaires, hybridations, mutualisations). Il ne s'agit donc pas de les exclure systématiquement, mais de **repenser les modalités d'intervention**, en tenant compte de leur faible retour sur investissement actuel et de leur position périphérique dans les réseaux de production de connaissances.
- Les **projets du cadran 4**, en phase d'émergence mais encore éloignés de la frontière, combinent des rendements contrastés avec des signaux positifs de structuration. Ils nécessitent une **veille stratégique renforcée**, car certains pourraient évoluer vers le cadran 2, tandis que d'autres justifieraient une réorientation. Dans ces domaines, un soutien plus sélectif et mieux accompagné (structuration des communautés, animation scientifique, outils de valorisation) permettrait de sécuriser les trajectoires.

## 2. Les équipements par catégories de domaines

Les projets d'équipement font l'objet d'un dispositif de financement dédié au sein des actions de l'ANR. Ces projets ont été eux aussi agrégés à l'échelle des domaines puis analysés par cadran. Néanmoins, afin d'analyser spécifiquement les caractéristiques de ces projets en équipement, des indicateurs spécifiques ont été calculés.

- **Cadran 1 :**

Ce cadran regroupe **165 projets d'équipement**, soit environ **16,6 projets par domaine**. Les dépenses totales s'élèvent à **276,6 M€**, pour une moyenne de **5,03 M€ par domaine** et un **coût unitaire de 1,68 M€ par projet**.

Bien que les montants engagés soient plus modérés que dans les domaines du cadran 2, ils permettent de maintenir un bon niveau d'équipement dans des domaines déjà structurés. Le **coût par publication** est estimé à environ **9 000 €**, ce qui est faible. En revanche, le **rendement technologique** apparaît plus limité, avec environ **un coût de 2,49 M€ par brevet**. Ces chiffres confirment le positionnement de ces domaines comme **pilliers stables**, dans lesquels les équipements soutiennent une dynamique de **consolidation** plutôt que de transformation.

- **Cadran 2 :**

Les domaines du cadran 2 concentrent le volume le plus important de projets d'équipement, avec un total de **208 projets**. En moyenne, chaque domaine comprend **42,2 projets d'équipement**, ce qui témoigne d'un niveau d'engagement particulièrement élevé ainsi que de la multidisciplinarité des projets d'équipements. Les **dépenses totales** atteignent **1,23 milliard d'euros**, soit **17,86 M€ par domaine** et environ **5,92 M€ par projet d'équipement**.

Ce niveau d'investissement, à la fois massif et concentré, confirme le caractère stratégique de ces domaines. Rapportées à la production scientifique et technologique, ces dépenses correspondent à environ **12 700 € par publication** et **1,8 M€ par brevet**, traduisant un **bon rendement** des investissements en matière de résultats scientifiques et technologiques.

- **Cadran 3 :**

Avec **150 projets d'équipement**, et en moyenne **12,2 projets d'équipement par domaine**, ce cadran affiche les **niveaux d'engagement les plus faibles**. Les **dépenses totales** atteignent **223,6 M€**, soit **3,29 M€ par domaine** et **1,49 M€ par projet d'équipement**.

Si le **coût par publication**, estimé à environ **13 200 €**, reste comparable à celui des autres cadrans, le **coût par brevet** est particulièrement élevé, dépassant **6 M€**. Ce ratio très défavorable souligne une **efficacité technologique faible** des projets d'équipement pour cette catégorie de domaine.

- **Cadran 4 :**

Les domaines du cadran 4 rassemblent **175 projets d'équipement**, soit en moyenne **22,36 projets par domaine**. Les **dépenses en équipement** atteignent **468,5 M€**, soit **8,52 M€ par domaine** et environ **2,68 M€ par projet**.

Ces investissements traduisent une volonté de **structurer des domaines encore en construction**. Toutefois, le **coût relatif des résultats** est plus élevé : environ **15 000 € par publication** et **2,95 M€ par brevet**. Ce différentiel suggère que les **effets des investissements restent encore à consolider** et que la montée en puissance de ces domaines dépendra de leur capacité à traduire les efforts d'équipement en retombées technologiques tangibles.

En conclusion, les **domaines du cadran 1** bénéficient d'un niveau d'équipement plus modéré mais cohérent avec leur profil. Les dispositifs existants permettent de maintenir une base scientifique solide. Toutefois, ces dépenses en équipement ne permettent pas de contrer le faible rendement technologique de ces domaines.

Les **domaines du cadran 2** concentrent une part importante des efforts d'équipement. Ce choix d'investissement ciblé apparaît pertinent au regard du **bon rendement scientifique et technologique** observé (coût par publication et par brevet), confirmant la légitimité d'une stratégie de consolidation et de montée en puissance. Les équipements y jouent un rôle d'amplification des retombées et de structuration des écosystèmes d'excellence.

Les **domaines du cadran 3** mobilisent des ressources plus modestes et présentent une efficacité plus faible en matière de valorisation technologique.

Enfin, les **domaines du cadran 4** font l'objet d'investissements significatifs en équipements, traduisant un pari stratégique sur leur structuration. Toutefois, les **investissements en équipement ne permettent pas (encore) de significativement accélérer l'empreinte scientifique et technologique**.

### 3. Empreinte scientifique

#### Empreinte scientifique de l'ANR dans le cadre de France 2030

L'indicateur analysé ici correspond à la **part des publications scientifiques françaises dans un domaine donné qui sont issues de projets financés par l'ANR**. Il reflète ainsi le poids relatif de l'ANR dans le cadre de France 2030 dans la production scientifique nationale, et permet d'identifier les domaines dans lesquels son action joue un **rôle structurant**.

Les résultats montrent une **hétérogénéité marquée** des taux d'empreinte selon les disciplines. Si la **moyenne nationale de l'empreinte scientifique** des projets ANR dans le cadre de France 2030 oscille autour de **7,5 %**, certains domaines atteignent des niveaux bien supérieurs, dépassant les **15 %**, voire les **18 %**.

#### Un effet structurant marqué dans certains domaines spécialisés

Le haut du classement est dominé par des domaines où **l'efficacité relative des financements ANR dans le cadre de France 2030 est particulièrement forte**, avec des contributions dépassant les 13% dans la production scientifique nationale.

Ce niveau d'empreinte peut être interprété comme un **indicateur de structuration des communautés de recherche** par les dispositifs ANR dans le cadre de France 2030.

Parmi les domaines les plus concernés, on retrouve notamment :

- **Medicine - Microbiology (2726) (19,4%)** : domaine biomédical stratégique, avec de nombreux projets historiques entre 2011 et 2014, incluant plusieurs LABEX, notamment en Île-de-France.
- **Physics and Astronomy - Acoustics and Ultrasonics (3102) (18,3%)** : domaine technique à forte intensité d'équipement, avec des projets soutenus de manière continue sur plus de 10 ans.
- **Social Sciences - Life-span and Life-course Studies (3319) (16,7%)** : spécialité en sciences humaines où l'ANR dans le cadre de France 2030 a su structurer un pan de la recherche sociale, principalement via deux LABEX : 11-LABX-0066 et 10-LABX-0084.

À l'autre extrémité du spectre, plusieurs domaines présentent des taux d'empreinte **nuls**, suggérant soit une faible visibilité académique, soit des temporalités longues ou des objectifs de recherche non académiques. Exemples :

- Business, Management and Accounting - Industrial Relations (1410)
- Arts and Humanities - General Arts and Humanities (1200)
- Social Sciences - Library and Information Sciences (3309)
- Energy - General Energy (2100)

#### Une forte représentation des sciences de la vie, en particulier la biochimie

Les **sciences de la vie** occupent une place prépondérante dans les domaines à forte empreinte, avec une concentration notable dans la **biochimie et la biologie moléculaire** (Biochemistry, Genetics and Molecular Biology), à travers plusieurs spécialités :

- Biochemistry (1303) : 16,5%
- Structural Biology (1315) : 14,4%

- Molecular Medicine (1313) : 14%
- Genetics (1311) : 13,5%
- Cancer Research (1306) : 13,5%
- Cell Biology (1307) : 12,3%

Cette concentration de spécialités reflète le rôle historique de l'ANR dans le cadre de France 2030 dans le financement de la recherche biomédicale avec de très nombreux projets sur une période de 2011 à 2025. Elle témoigne également d'une forte cohérence intégrant aussi de fortes dépenses en équipement.

#### Une diversité disciplinaire croissante dans les domaines à forte empreinte

Au-delà des sciences de la vie, plusieurs **disciplines variées** affichent également des taux d'empreinte élevés, illustrant le **caractère transversal des appels ANR dans le cadre de France 2030** :

- **Neuroscience - Biological Psychiatry** (2803) : 13,8%, en lien avec la montée des enjeux de santé mentale.
- **Neuroscience - Developmental Neuroscience** (2806) : 11,8%, autour des dynamiques cognitives et développementales.
- **Mathematics - Computational Mathematics** (2605): 13,6%, probablement porté par les thématiques en intelligence artificielle, simulation, ou calcul haute performance.
- **Social Sciences - Cultural Studies** (3316): 12,4%, qui témoigne de la structuration croissante des SHS, notamment via les études culturelles et postcoloniales.

#### Des domaines à plus faible empreinte nécessitant un suivi différencié

Certaines disciplines restent en retrait, malgré une volumétrie de projets parfois significative. Cela peut s'expliquer par :

- Une orientation des projets vers **des finalités non académiques** (valorisation industrielle, transfert technologique...);
- Des **temporalités longues** de publication ou de valorisation ;
- Un **niveau de soutien ANR dans le cadre de France 2030 trop faible** pour produire un effet visible.

#### Enseignements pour le pilotage stratégique

L'empreinte de l'ANR dans le cadre de France 2030 apparaît relativement homogène entre les différents cadrans, mais les **domaines du cadran 2** concentrent en moyenne le **taux d'empreinte le plus élevé (7,6 %)**, suivis des **domaines du cadran 1** (7,4 %), des **domaines du cadran 4** (6,65 %) et des **domaines du cadran 3** (6,51 %). Les **domaines du cadran 2** regroupent également le plus grand nombre de domaines où l'empreinte ANR dans le cadre de France 2030 dépasse 5,5 % (57 domaines).

Ce résultat confirme leur **rôle moteur** dans le paysage scientifique français, mais l'analyse montre aussi que **l'empreinte n'est pas uniquement corrélée au montant des financements ou au nombre de projets**. Certains domaines fortement financés n'affichent pas pour autant une empreinte supérieure à la moyenne. Inversement, certains domaines peu financés peuvent présenter des taux d'empreinte remarquables. **Les écarts entre financement et empreinte scientifique de l'ANR dans le cadre de France 2030 doivent donc être analysés à l'aune des spécificités disciplinaires,**

pour identifier d'éventuels besoins d'ajustement car l'empreinte finale de l'ANR dans le cadre de France 2030 n'est pas strictement dépendant ni du nombre de projets ni du montant dépensé.

## Émergence scientifique

L'indicateur d'**émergence scientifique** reflète le **niveau d'influence et de réorganisation des connaissances au niveau mondial**. Un score élevé signifie que le domaine est au cœur des recombinaisons technologiques actuelles : il attire des contributions croisées, se connecte à d'autres disciplines et devient une zone d'intensification de l'innovation.

### Zoom sur les trois premiers domaines du classement

1. **Mathematics - Applied Mathematics (2604) (0,8943)**

Ce domaine occupe la première place, ce qui reflète son rôle central dans la structuration des réseaux scientifiques mondiaux. Les mathématiques appliquées constituent une base méthodologique indispensable à de nombreux champs : **intelligence artificielle, modélisation multi-physique, optimisation industrielle, traitement de données massives**. Sa place en tête traduit le fait que les avancées méthodologiques se diffusent rapidement dans d'autres disciplines, augmentant ainsi son influence et sa centralité mondiale.

2. **Medicine - Ophthalmology (2731) (0,8923)**

La présence de l'ophtalmologie en seconde position peut surprendre, mais elle traduit une dynamique internationale spécifique : **montée des technologies d'imagerie biomédicale, développement des thérapies géniques et cellulaire appliquées à la vision**. De plus ce domaine était à un niveau très bas en début de période ce qui fait que sa progression sur la période est particulièrement forte.

3. **Mathematics - Modeling and Simulation (2611) (0,8902)**

La modélisation et la simulation se hissent également aux avant-postes. Ce domaine connaît une véritable explosion du nombre de publications au début des années 2020. Sa forte centralité s'explique par la transversalité de ses applications, en particulier avec **l'essor des jumeaux numériques, de la simulation haute performance et de la modélisation prédictive dans les transitions énergétiques et industrielles**.

### Une forte émergence des mathématiques

L'analyse du classement confirme que **les mathématiques figurent parmi les disciplines les plus dynamiques en matière d'émergence scientifique**. Plusieurs domaines apparaissent dans le haut du classement, avec des scores très élevés :

- Mathematics - Applied Mathematics (2604) : 0,8943
- Mathematics - Modeling and Simulation (2611) : 0,8902
- Mathematics - Numerical Analysis (2612) : 0,8862
- Mathematics - Algebra and Number Theory (2602) : 0,8028

Ces résultats soulignent que les mathématiques ne se limitent pas à un rôle théorique, mais qu'elles constituent un **socle transversal d'innovation**. L'analyse numérique et la modélisation/simulation sont directement appliquées à l'ingénierie et à la physique, tandis que l'algèbre et la théorie des nombres trouvent des prolongements dans la cryptographie, la sécurité des systèmes et certaines branches de l'informatique avancée.

L'importance de ce socle se reflète également dans la catégorie Computer Science - Computational Theory and Mathematics (1703) : 0,7947, qui **illustre la convergence entre mathématiques fondamentales et informatique**.

Ces résultats traduisent aussi l'importance des mathématiques dans le développement de l'**intelligence artificielle** : optimisation, calcul numérique, algèbre et théorie des graphes fournissent les outils indispensables à l'apprentissage automatique, aux réseaux de neurones et à la modélisation de systèmes complexes.

### La dynamique des spécialités médicales

L'analyse montre que de nombreuses spécialités médicales figurent dans le haut du classement, avec des scores proches ou supérieurs à 0,75. Aux côtés de l'ophtalmologie (Medicine - Ophthalmology (2731) : 0,8923), de la cardiologie (Medicine - Cardiology and Cardiovascular Medicine (2705) - 0,7988) et de la neurologie (Medicine - Neurology (2728) : 0,7703), d'autres disciplines cliniques affichent également une forte dynamique :

- Medicine - Obstetrics and Gynecology (2729) : 0,7663
- Medicine - Hepatology (2721) : 0,7663
- Medicine - Radiology, Nuclear Medicine and Imaging (2741) : 0,7622
- Medicine - Nephrology (2727) : 0,7541
- Medicine - Infectious Diseases (2725) : 0,7541

Ces résultats traduisent l'essor des **medtech** et l'intégration rapide des innovations technologiques en pratique clinique. L'imagerie médicale (radiologie, médecine nucléaire), portée par l'IA et le deep learning, joue un rôle majeur dans le diagnostic. La néphrologie et l'hépatologie bénéficient des avancées en biologie moléculaire et en dispositifs de suivi. L'obstétrique et la gynécologie voient se développer des solutions numériques et des biomarqueurs pour le suivi personnalisé. Enfin, l'infectiologie, renforcée par la pandémie récente, se trouve au cœur de la recherche sur les vaccins, les antiviraux et la modélisation des épidémies.

### Une forte émergence des sciences du vivant

Au-delà des mathématiques, les sciences du vivant occupent également une place majeure dans l'émergence scientifique. Plusieurs spécialités affichent des scores particulièrement élevés :

- Immunology and Microbiology - Parasitology (2405) : 0,8394
- Biochemistry, Genetics and Molecular Biology - Molecular Medicine (1313) : 0,8313
- Biochemistry, Genetics and Molecular Biology - Cancer Research (1306) : 0,7744
- Agricultural and Biological Sciences - Aquatic Science (1104) : 0,7805
- Agricultural and Biological Sciences - Insect Science (1109) : 0,7703

Ces résultats traduisent le dynamisme des **biotechnologies et des sciences médicales appliquées**, avec des avancées fortes dans la compréhension des pathogènes, le développement de thérapies innovantes en médecine moléculaire et en oncologie, ou encore l'exploration de nouvelles ressources biologiques (aquatique, insectes).

### Les domaines faiblement émergents

À l'opposé du classement, plusieurs disciplines se caractérisent par des scores très faibles, souvent inférieurs à 0,25. Parmi elles, on retrouve des champs très divers :

- **Psychologie et neurosciences appliquées** : Psychology - Neuropsychology and Physiological Psychology (3206) : 0,2622 ; Psychology - General Psychology (3200) : 0,2215 ; ou encore Neuroscience - Developmental Neuroscience (2806) : 0,1321.
- **Médecine appliquée** : certaines spécialités cliniques apparaissent en bas du classement, comme Medicine - Anatomy (2702) : 0,2581 ; Medicine - Family Practice (2714) : 0,1606 ou Medicine - Health Informatics (2718) : 0,0894.
- **Sciences des matériaux et chimie** : malgré la présence de biomatériaux dans le haut du classement, d'autres domaines comme Materials Science - General Materials Science (2500) : 0,2602 ; Materials Science - Surfaces, Coatings and Films (2508) : 0,2581 ; ou Chemistry - Physical and Theoretical Chemistry (1606) : 0,1829 ; affichent une dynamique bien plus faible.
- **Sciences sociales et humanités** : Social Sciences - Gender Studies (3318) : 0,1951, Social Sciences - Public Administration (3321) : 0,2073 ; Social Sciences - Safety Research (3311) : 0,1667 ou encore Arts and Humanities - Literature and Literary Theory (1208) : 0,2093 ; illustrent la faible intensité technologique de ces disciplines.
- **Métiers de santé et professions paramédicales** : Health Professions - Occupational Therapy (3609) : 0,2439 ; Health Professions - Radiological and Ultrasound Technology (3614) : 0,2337 ; Health Professions - Emergency Medical Services (3604) : 0,1667 ou Dentistry - General Dentistry (3500) : 0,1016.

Ces résultats confirment que les **sciences humaines, sociales et médicales appliquées** contribuent peu à l'émergence scientifique telle que mesurée par cet indicateur. Leur impact sur l'innovation est probablement plus réduit, moins transversal et plus intra-disciplinaire.

## Carte des publications scientifiques par pays

### Analyse globale

L'analyse agrégée de l'ensemble des domaines met en évidence une forte concentration géographique de la production scientifique mondiale.

Les **États-Unis** dominent largement avec 8 510 682 articles, soit presque deux fois plus que la **Chine**, deuxième avec 4 782 811 articles. Ensemble, ces deux pays représentent un pôle transpacifique incontournable qui structure la dynamique mondiale de recherche.

En **Europe**, plusieurs pays occupent une place de premier plan : le **Royaume-Uni (2 422 700 articles)**, l'**Allemagne (2 422 278)** et la **France (1 680 304)** forment un noyau scientifique solide, complété par l'**Italie (1 409 994)** et l'**Espagne (1 140 188)**. Ces pays, bien que publiant individuellement moins que les États-Unis ou la Chine, ils constituent ensemble un bloc scientifique majeur.

En **Asie**, le **Japon (1 700 204 articles)**, l'**Inde (1 283 186)** et la **Corée du Sud (882 245)** apparaissent comme des pôles de recherche majeurs. Ils confirment la montée en puissance de l'Asie dans la compétition scientifique internationale, aux côtés de la Chine.

D'autres pays de taille plus réduite se distinguent par une **intensité scientifique disproportionnée** : la **Suisse (685 413 articles)**, les **Pays-Bas (823 855)** et la **Suède (565 983)**. Leur forte visibilité témoigne d'écosystèmes de recherche très compétitifs et internationalisés.

Enfin, certains **pays émergents** accèdent au top 20, comme le **Bésil (712 573 articles)** et l'**Iran (398 623 articles)**, illustrant la diffusion progressive de la production scientifique au-delà des pôles traditionnels.

## Analyse sectorielle du leadership scientifique par domaine

L'analyse des pays dominants par domaine met en évidence une concentration très marquée du leadership scientifique mondial.

Les **États-Unis** apparaissent comme leaders dans 193 domaines sur 247 (près de 80%), confirmant la puissance de leur écosystème de recherche et leur capacité à couvrir de manière transversale l'ensemble des champs scientifiques et technologiques.

La **Chine** occupe la deuxième place, en tête dans 38 domaines (environ 15%). Elle s'affirme comme un pôle scientifique en forte croissance, concentrant son leadership sur des niches stratégiques où elle investit massivement.

En comparaison, les autres pays n'apparaissent que de façon ponctuelle : la **Russie** (4 domaines), l'**Italie** (3), le **Royaume-Uni** (2), l'**Inde** (2) et l'**Indonésie** (2). Quelques cas isolés concernent le **Brésil**, l'**Espagne** et la **Corée du Sud**, chacun dominant un seul domaine. Cette distribution illustre une **bipolarisation nette** de la recherche mondiale entre les États-Unis et la Chine.

Au-delà du leadership, l'examen du **top 5 par domaine** met en évidence un autre niveau de compétition scientifique, où plusieurs pays confirment leur rôle structurant :

- Le **Royaume-Uni (203 présences)** et l'**Allemagne (193)** constituent les principaux pôles scientifiques européens, régulièrement positionnés dans le groupe de tête.
- Derrière ce duo, la **France (74)**, le **Japon (59)** et l'**Inde (59)** se distinguent comme puissances scientifiques solides et reconnues à l'échelle internationale.
- Le **Canada (42)** et l'**Italie (39)** maintiennent une présence régulière, témoignant d'écosystèmes de recherche stables et internationalisés.
- L'**Australie (23)**, l'**Indonésie (21)** et le **Brésil (18)** traduisent une montée en puissance régionale, renforçant la visibilité de l'Asie du Sud - Est et de l'Amérique latine.
- Enfin, la **Russie (14)** et l'**Espagne (14)** conservent une visibilité ponctuelle, liée à des spécialisations sectorielles.

Cette répartition dessine une hiérarchie à plusieurs niveaux. On observe d'abord un **duopole constitué des États-Unis et de la Chine**, moteurs principaux de la production scientifique mondiale. Vient ensuite un **noyau dur européen**, regroupant le **Royaume-Uni, l'Allemagne, la France, l'Italie et l'Espagne**, qui consolide la position de l'Europe dans le paysage scientifique international. Enfin, un **cercle élargi de puissances scientifiques régionales**, comprenant le **Japon, l'Inde, le Canada, l'Australie, le Brésil et l'Indonésie**, contribue à la diffusion et à la diversification des dynamiques de recherche à l'échelle mondiale.

## 4. Valorisation

**Pour rappel, l'analyse présentée ici porte sur les dépôts de brevets. La nationalité d'un brevet est définie en fonction de l'office auprès duquel il a été déposé, et non de la nationalité du déposant.** Cette méthode ne reflète donc pas de manière exacte l'origine géographique de l'innovation, mais plutôt sa **destination stratégique** en matière de protection industrielle.

Cela dit, il convient de noter que les entreprises ont généralement tendance à privilégier le dépôt de leurs brevets dans leur pays d'origine, ce qui **limite partiellement ce biais** et permet une lecture raisonnablement représentative des dynamiques nationales. Toutefois, cette limite doit être gardée à l'esprit dans l'interprétation des résultats.

## Empreinte technologique de l'ANR dans le cadre de France 2030

L'indicateur analysé ici mesure, pour chaque domaine, dans quelle mesure les projets financés par ANR dans le cadre de France 2030 se **convertissent en production brevet**. Il renseigne donc le **poids relatif** de l'ANR dans le cadre de France 2030 dans la dynamique de dépôt, et permet d'identifier les domaines où son action exerce un **effet structurant sur la valorisation**.

Les résultats montrent une **forte dispersion des taux d'empreinte** d'un domaine à l'autre. Plusieurs domaines atteignent **des niveaux très élevés, au-delà de 15% et jusqu'à plus de 30%**, signalant une chaîne de valorisation courte et efficace entre recherche financée et dépôts. À l'inverse, de nombreux domaines restent à empreinte faible ou nulle. **159 domaines** sur **247**, soit **64,4%**, ont une empreinte technologique **nulle**. Ceci peut refléter soit des temporalités plus longues de valorisation, soit des résultats non brevetables, ou encore une culture de la valorisation brevet pas encore assez développée dans la communauté scientifique.

### Un effet structurant marqué dans certains domaines spécialisés

Le haut du classement est dominé par des domaines où l'**efficacité relative** des financements ANR dans le cadre de France 2030 est particulièrement forte ; les contributions y dépassent nettement les 10% et atteignent des valeurs exceptionnelles :

- Environmental Science - Water Science and Technology (2312) et Medicine - Oncology (2730) ont une empreinte de 33,33%, mais ceci est fortement nuancé car il n'existe que 3 brevets français sur la période étudiée.
- Physics and Astronomy - Statistical and Nonlinear Physics (3109) domaine pour lequel l'empreinte est de 23,08% : un domaine à la frontière entre le cadran 1 et le cadran 2, pour lequel la rentabilité est assez forte et où les dépenses ANR dans le cadre de France 2030 n'y sont pas parmi les plus élevées.
- Immunology and Microbiology - Microbiology (2404) avec une empreinte de 20,59% : un domaine **du cadran 4** dont les dépenses, y compris d'équipement, sont également relativement faibles, ce qui confère à ce domaine un bon rendement.

Ce top est à fortement nuancer car il concerne des nombres de publications de brevets faibles voire très faibles.

### Un effet structurant marqué dans certains domaines spécialisés

Au-delà de ces 4 domaines à plus de 20%, plusieurs domaines présentent des niveaux élevés et soutenus dans les sciences du vivant et la médecine. Ils concernent aussi au total un plus grand nombre de brevets :

- Biochemistry, Genetics and Molecular Biology - Cancer Research (1306) : 17.65%
- Biochemistry, Genetics and Molecular Biology - Molecular Medicine (1313) : 15.79%
- Biochemistry, Genetics and Molecular Biology - Endocrinology (1310) : 12.5%
- Immunology and Microbiology - Virology (2406) : 10%
- Medicine - Physiology (2737) : 7.46%
- Medicine - Endocrinology, Diabetes and Metabolism (2712) : 7.14%
- Medicine - Dermatology (2708) : 6.67%
- Medicine - Radiology, Nuclear Medicine and Imaging (2741) : 3.07%

Ensemble, ils signalent un continuum de valorisation biomédicale (biomarqueurs, diagnostics, dispositifs, IA d'imagerie, protocoles thérapeutiques) avec un rendement brevet notable.

## Informatique et mathématiques

Les méthodes quantitatives et l'informatique appliquée affichent aussi des empreintes relativement **élevées**:

- Mathematics - Statistics and Probability (2613) : 15%
- Decision Sciences - Statistics, Probability and Uncertainty (1804) : 7.41%
- Computer Science - Software (1712) : 6.06%
- Computer Science - Computer Graphics and Computer - Aided Design (1704) : 5%
- Computer Science - Computational Theory and Mathematics (1703) : 4.85%
- Computer Science - Human - Computer Interaction (1709) : 4.52%
- Computer Science - Computer Vision and Pattern Recognition (1707) : 4.13%
- Computer Science - Artificial Intelligence (1702) : 3.37%

Cela suggère que l'ANR dans le cadre de France 2030 **entraîne la conversion** en actifs protégés dans des segments où la culture du brevet n'est **pas toujours systématique**, notamment sur des **usages logiciels et procédés** (CAO, HCI, vision/inspection), dans des domaines par ailleurs **très dynamiques**.

## Environnement et énergie,

Le bloc environnement et énergie et montre une **empreinte soutenue** de l'ANR dans le cadre de France 2030 avec des nombres de brevets publiés plus élevés :

### Environnement :

- Environmental Science - Health, Toxicology and Mutagenesis (2307) - 9.76%
- Environmental Science - Pollution (2310) : 7.14%
- Environmental Science - Ecological Modeling (2302) : 5.89%
- Environmental Science - Environmental Engineering (2305) : 4.93%

### Énergie :

- Energy - Renewable Energy, Sustainability and the Environment (2105) : 6.56%
- Energy - Energy Engineering and Power Technology (2102) : 2.35%

Ces résultats sont **cohérents** avec la montée des **politiques de transition énergétique** et l'**industrialisation** de technologies bas carbone, qui favorisent la conversion des projets en **actifs brevetables** et leur **transfert** vers les filières.

## Enseignements pour le pilotage stratégique

L'analyse de l'empreinte technologique met en évidence plusieurs leviers d'action pour renforcer la capacité de l'ANR dans le cadre de France 2030 à transformer la recherche financée en résultats valorisables :

### 1. **Consolider les pôles à très haut rendement**

Les domaines affichant des taux de rendement élevés sont rares et concernent parfois des quantités de brevets limitées. Ils justifient un pilotage renforcé, visant à sécuriser et à amplifier la politique de dépôts.

### 2. **Renforcer les dynamiques biomédicales intégrées**

Le continuum observé entre biologie moléculaire, médecine et imagerie invite à développer des dispositifs transversaux combinant recherche fondamentale, infrastructures

technologiques et accompagnement PI. L'objectif est d'accélérer le passage des découvertes scientifiques (biomarqueurs, diagnostics, IA d'imagerie) à des innovations protégées et transférables.

3. **Soutenir la valorisation dans l'informatique et les mathématiques appliquées**  
 Dans des domaines traditionnellement moins orientés vers la Propriété Industrielle (PI) (logiciels, interaction homme-machine, vision), l'ANR dans le cadre de France 2030 joue déjà un rôle différenciateur avec une empreinte relativement forte. Il convient de prolonger cet effet structurant en développant une culture PI affirmée au sein des communautés concernées.
4. **Structurer le lien avec les politiques de transition énergétique et environnementale**  
 L'empreinte notable dans les domaines liés à l'environnement et à l'énergie doit être consolidée par des actions transverses. Celles-ci pourraient associer d'autres disciplines des sciences de l'ingénieur notamment afin de favoriser la montée en maturité technologique et le transfert vers les filières industrielles.
5. **Adapter les modalités de suivi et d'évaluation pour les domaines à faible propension au brevet**  
 Dans les champs où les résultats se traduisent moins par des brevets (par nature disciplinaire ou par temporalité de valorisation), il serait intéressant de compléter l'analyse par d'autres indicateurs : logiciels, normes, bases de données, politiques publiques ou adoption industrielle. Cela permettrait d'éviter de sous-évaluer l'effet réel des financements ANR. Or ces informations sont inexistantes ou trop peu disponibles dans les données fournies pour produire une analyse robuste.

## Capacité d'absorption et identification des acteurs clés

### Capacité d'absorption par pays : une lecture transversale

L'analyse de la répartition géographique des entreprises capables d'absorber les technologies financées par l'ANR dans le cadre de France 2030 met en évidence une structure relativement stable mais différenciée selon les cadrans stratégiques. Pour rappel, cette analyse n'a été réalisée que pour des entreprises domiciliées en Union européenne ainsi qu'au **Royaume-Uni, Suisse, Norvège et Islande. Dans ce périmètre, l'Allemagne et la France forment le noyau central** de cette capacité d'absorption, concentrant ensemble parfois entre 50% et 60% des parts selon les domaines.

Il convient toutefois de **souligner une limite méthodologique importante** : la nationalité des entreprises est ici déterminée à partir de la localisation de leur siège social. Or, cette approche peut **favoriser artificiellement certains pays**, notamment ceux où les sièges sociaux sont implantés pour des raisons fiscales ou administratives, sans que cela reflète nécessairement la localisation effective des activités de R&D ou de production. Ce biais doit être pris en compte dans l'interprétation des résultats, en particulier pour les multinationales dont les centres décisionnels et opérationnels peuvent être géographiquement dispersés.

### Lecture par cadran

Lorsque l'on analyse les choses par cadran **l'Allemagne occupe systématiquement la première place à l'échelle de la catégorie, avec des parts allant de 31% à 40%**, confirmant son rôle de pivot industriel et technologique en Europe. **La France se situe en deuxième position, avec des parts comprises entre 18% et 22%**, témoignant d'une présence forte mais proportionnellement plus limitée.

Autour de ce duo franco-allemand, **les Pays-Bas, le Royaume-Uni et la Suisse apparaissent comme des challengers récurrents, représentant chacun entre 7% et 11%** de la capacité d'absorption dans les domaines du cadran 1, 2 et 4. Leur poids relatif s'accroît dans les domaines

émergents **du cadran 4**, traduisant leur spécialisation dans des niches technologiques en forte dynamique. Enfin, dans le cadran 3, l'Allemagne domine nettement avec 40% des parts, tandis que la France se maintient à 20% et que les Pays-Bas (15%) ainsi que la Suède (5%) témoignent de spécialisations résiduelles.

Cette analyse met ainsi en évidence **une structuration franco-allemande au cœur des dynamiques d'absorption**, complétée par des positions fortes mais plus spécialisées des pays de taille moyenne. Elle souligne également la nécessité, pour la France, de consolider ses positions dans les domaines stratégiques (cadran 2) tout en renforçant sa présence dans les champs émergents (cadran 4), si elle ne veut pas perdre trop de terrain vis-à-vis de l'Allemagne.

#### Analyse du top 5 et top 10 pays par domaine

**L'Allemagne et la France** dominent très largement les classements. **L'Allemagne figure dans le Top 5 de 93% des domaines étudiés, avec une part moyenne de près de 38%, et la France dans 82% des cas, avec une part moyenne proche de 30%**. Ces résultats confirment le rôle moteur de ces deux pays dans la transformation de la recherche en applications industrielles, et leur position de partenaires naturels dans toute stratégie de transfert et de co-développement.

**Pays-Bas, Suisse, Royaume-Uni** affichent une couverture large (plus de la moitié des domaines) et des parts moyennes élevées (environ 20 à 25% lorsqu'ils apparaissent en Top 5). Leur poids, disproportionné au regard de leur taille, témoigne de **la densité de leurs écosystèmes d'innovation et de leur rôle de hubs d'absorption** dans des segments de haute valeur ajoutée.

**Suède, Belgique, Irlande, Autriche, Italie, Danemark, Espagne, Finlande** apparaissent régulièrement dans le Top 10 (entre 20% et 40% des domaines) mais plus rarement dans le Top 5. Leur profil suggère des écosystèmes ouverts et réceptifs, où l'ANR dans le cadre de France 2030 pourrait jouer un rôle d'amplification en renforçant la profondeur des coopérations existantes. Certains d'entre eux, comme **l'Irlande ou le Danemark, affichent des parts particulièrement fortes lorsqu'ils apparaissent en Top 5, traduisant des positions de niche à fort potentiel**.

**Hongrie, Norvège, Luxembourg** ne sont présents que ponctuellement dans le Top 10, mais leur apparition, même limitée, indique l'existence **de niches technologiques ou sectorielles dans lesquelles une stratégie ciblée pourrait s'avérer pertinente**. Ils constituent autant de relais potentiels pour diversifier les partenariats européens.

Cette analyse confirme que la capacité d'absorption des technologies issues des projets ANR dans le cadre de France 2030 repose sur **un noyau franco-allemand, renforcé par quelques hubs européens (Pays-Bas, Royaume-Uni, Suisse)** qui jouent un rôle clé dans la diffusion et la valorisation. Les pays de taille moyenne, souvent spécialisés sur des niches à fort potentiel, offrent des relais complémentaires mais encore insuffisamment exploités.

#### Analyse des top domaines de valorisation

L'enseignement majeur est double : d'une part, il s'agit **de consolider les partenariats structurants avec les pays leaders** pour sécuriser et amplifier l'impact industriel des projets financés ; d'autre part, il convient de développer **une stratégie différenciée avec les pays du second cercle** afin de transformer leurs positions de niche en relais de valorisation plus profonds. Enfin, l'apparition ponctuelle de pays périphériques dans certains domaines invite à explorer de nouveaux axes de coopération ciblés, capables d'ouvrir des débouchés inattendus et de renforcer la souveraineté scientifique et technologique européenne.

Au-delà de la répartition par pays, l'analyse des domaines selon leur niveau de **capacité d'absorption** permet d'affiner la lecture sectorielle et d'identifier les domaines les plus structurants en matière de valorisation industrielle.

Les domaines affichant les **plus fortes capacités d'absorption** sont, sans surprise, concentrés dans les domaines de **l'ingénierie, des matériaux et des technologies numériques appliquées**. Ainsi, le domaine **Engineering – Electrical and Electronic Engineering (2208)** atteint une capacité d'absorption de **38,3%**, confirmant son rôle de domaine technologique structurant au croisement des filières industrielles stratégiques. Il est suivi par **Engineering – Mechanical Engineering (2210) (29,1%)**, **Engineering – Automotive Engineering (2203) (18,8%)** et **Materials Science – Polymers and Plastics (2507) (17,5%)**, qui illustrent la concentration de la capacité d'absorption dans les filières industrielles historiques fortement intégrées.

Le domaine **Computer Science – Computer Vision and Pattern Recognition (1707)** affiche également une capacité d'absorption élevée (**17,3%**), traduisant la dynamique croissante autour des technologies d'intelligence artificielle appliquées à l'image et à la reconnaissance automatique. Cette performance souligne l'attractivité de ce segment pour les entreprises européennes, en particulier dans les secteurs de la défense, de l'automobile, de la santé et de la sécurité.

#### Lecture des principales grandes disciplines

- **Engineering** : moyenne 13.74% (13 domaines)
- **Computer Science** : moyenne 8.07% (10 domaines)
- **Materials Science** : moyenne 8.46% (7 domaines)
- **Energy** : moyenne 7.7% (2 domaines)
- **Physics and Astronomy** : moyenne 5.21% (6 domaines)
- **Chemistry** : moyenne 3.71% (4 domaines)
- **Environmental Science** : moyenne 1.94% (4 domaines)
- **Biochemistry, Genetics and Molecular Biology** : moyenne 1.85% (6 domaines)
- **Immunology and Microbiology** : moyenne 1.25% (3 domaines)
- **Medicine** : moyenne 1.20% (7 domaines)
- **Mathematics** : moyenne 0.90% (3 domaines)

Ces données montrent que la capacité d'absorption des technologies s'incarne dans certains segments technologiques spécifiques, qui concentrent un tissu industriel dense, une demande d'innovation forte, et des opportunités de transfert direct. En cela, ces domaines peuvent être considérés comme des « points d'ancrage » privilégiés pour maximiser l'impact économique des investissements publics dans la recherche. D'autres domaines plus théoriques, ou plus spécialisés ne peuvent en revanche que plus difficilement trouver des débouchés.

**L'intégration systématique de cette lecture sectorielle dans la stratégie de transfert de l'ANR dans le cadre de France 2030 permettrait de mieux cibler les domaines à fort potentiel de valorisation**, de renforcer les interactions avec les filières industrielles concernées.

#### Carte

L'analyse de la répartition mondiale des brevets dans les domaines où l'ANR dans le cadre de France 2030 a financé au moins un projet met en évidence des contrastes très marqués entre blocs géographiques. Elle permet de replacer la production française dans le contexte global et de souligner les écarts de compétitivité technologique.

## Analyse internationale de la production de brevets par domaines

La **Chine** domine massivement la production de brevets, avec plus de **2,4 millions de dépôts**, il faut néanmoins se montrer prudent avec ce chiffre car la qualité des brevets déposés en Chine est parfois très contestable. Quoi qu'il en soit, la Chine est suivie par le **Japon** (1,21M) et les **États-Unis** (889k). Ces trois pays constituent les pôles structurants du système mondial, concentrant à eux seuls l'essentiel des dépôts.

La **Corée du Sud** (406k) et **Taïwan** (181k) complètent ce bloc asiatique, confirmant leur rôle de hubs industriels et de plateformes de dépôt massives, notamment dans les secteurs électroniques et numériques.

Avec **206k brevets**, l'**Allemagne** s'impose comme le premier pôle européen, bien loin devant l'**Espagne** (51,6k) et la **France** (51,5k). Ces deux pays affichent des volumes comparables mais restent très en retrait des leaders asiatiques et nord-américains.

Le **Royaume-Uni** (36,8k) et l'**Autriche** (36,8k) occupent également une place significative, mais sans atteindre les volumes qui caractérisent les pôles dominants.

Des pays comme le **Canada** (85k) et l'**Australie** (83k) apparaissent dans le Top 15, traduisant une capacité de valorisation forte malgré leur poids démographique plus limité.

Le **Brésil** (48,5k) se distingue en Amérique latine, tandis que la **Russie** (65,6k) et le **Mexique** (28,4k) témoignent de positions plus spécialisées mais non négligeables.

## Lecture par grande discipline et technologiques

L'analyse se concentre sur cinq grands domaines scientifiques et technologiques : **Engineering; Computer science; Materials Science; Physics and Astronomy et Chemistry** ; car la production mondiale de brevets y est la plus importante dans tous les pays. Ils représentent ainsi les principaux marchés en matière de valorisation des brevets.

- Dans **Engineering**, la Chine domine nettement avec 1 280 225 brevets, soit 44,98 % du total de la discipline (2 845 972). Le Japon suit à 570 340 (20,04 %), devant les États-Unis à 359 544 (12,63 %) et la Corée du Sud à 181 939 (6,39 %). L'Allemagne complète le quinté de tête avec 105 612 (3,71 %). La France se situe au 9<sup>e</sup> rang avec 25 284 brevets, soit 0,89 % du total.
- En **Computer Science**, la Chine arrive également en tête avec 365 227 brevets (36,55 % d'un total de 999 355), devant les États-Unis à 215 509 (21,56 %) et le Japon à 192 975 (19,31 %). La Corée du Sud occupe la quatrième place à 67 069 (6,71 %), suivie de l'Allemagne à 27 343 (2,74 %). La France se classe 11<sup>e</sup> avec 8 057 brevets (0,81 %).
- Pour **Materials Science**, la Chine mène avec 255 334 brevets (32,54 % d'un total de 784 444), devant le Japon à 180 865 (23,05 %) et les États-Unis à 117 462 (14,98 %). La Corée du Sud affiche 65 299 (8,33 %) et Taïwan 36 717 (4,68 %). La France occupe le 12<sup>e</sup> rang avec 7 147 brevets, soit 0,91 %.
- En **Physics and Astronomy**, la Chine totalise 179 654 brevets (32,58 % d'un total de 551 579), suivie par le Japon à 104 944 (19,02 %) et les États-Unis à 103 460 (18,76 %). La Corée du Sud en compte 49 285 (8,93 %) et Taïwan 26 848 (4,87 %). La France est 9<sup>e</sup> avec 4 365 brevets (0,79 %).
- Enfin, en **Chemistry**, la Chine reste en tête avec 37 260 brevets (19,62 % d'un total de 189 947), le Japon suit à 24 537 (12,92 %) et les États-Unis à 22 056 (11,61 %). Viennent ensuite la Corée du Sud avec 10 185 (5,36 %) et le Canada avec 9 130 (4,81 %). La France se classe 27<sup>e</sup>, avec 1 376 brevets, soit 0,72 % du total.

En comparaison, l'**Allemagne** est présente plus régulièrement, y compris dans les sciences de la vie et l'agriculture, ce qui confirme une meilleure conversion des résultats scientifiques en actifs brevetables dans ces pays.

Cette analyse souligne le **déséquilibre structurel** : la France, bien qu'elle dispose d'une recherche académique reconnue (notamment dans les sciences du vivant), reste en retrait dans la valorisation par brevet. Son effort est concentré dans quelques segments (informatique, ingénierie, environnement), mais peine à rivaliser en volume avec les grands pôles mondiaux.

## 5. Équipements

### Empreinte des dépenses en équipement de l'ANR dans le cadre de France 2030

L'indicateur analysé mesure, pour chaque domaine, la part des publications scientifiques françaises qui provient de projets comportant des équipements financés par l'ANR. Il renseigne donc le rôle structurant de ces investissements sur la capacité de production scientifique nationale.

Au total, les dépenses en équipement ANR dans le cadre de France 2030 contribuent à la production scientifique dans 228 domaines sur les 247 étudiés, avec une empreinte moyenne de 3,26% sur les domaines où l'empreinte est non nulle. Ce chiffre confirme l'importance des investissements en équipements pour soutenir la compétitivité scientifique française, tout en soulignant l'hétérogénéité des résultats selon les disciplines.

#### Domaines à empreinte élevée

- Biochemistry, Genetics and Molecular Biology - Structural Biology (1315) : 12,28%
- Biochemistry, Genetics and Molecular Biology - Biochemistry (1304) : 9,71%
- Neuroscience - Biological Psychiatry (2803) : 6,46%

Ces trois domaines se distinguent par un rôle particulièrement fort des équipements dans la production scientifique française. En effet, la part des publications issus des projets d'équipement de l'ANR dans le cadre de France 2030 y est la plus forte. Ils concentrent des niveaux significatifs de dépenses en équipements, avec respectivement 10,7 M€ pour Biochemistry, Genetics and Molecular Biology - Structural Biology (1315), 22,4 M€ pour Biochemistry, Genetics and Molecular Biology - Biochemistry (1304) et 5,3 M€ pour Neuroscience - Biological Psychiatry (2803).

**Ces montants restent néanmoins très éloignés de ceux observés dans certains domaines concentrant les plus hauts niveaux d'investissement, tels que :**

- Computer Science - Information Systems (1710) : 110,3 M€,
- Biochemistry, Genetics and Molecular Biology - Molecular Biology (1312) : 82,6 M€,
- Earth and Planetary Sciences - Atmospheric Science (1902) : 82,3 M€,
- Environmental Science - Global and Planetary Change (2306) : 76,2 M€.

Ce contraste met en évidence que **l'empreinte scientifique des projets d'équipements n'est pas strictement proportionnelle aux montants engagés**. D'ailleurs aucune corrélation entre les deux n'est constatée sur l'ensemble des données; certains domaines relativement moins dotés en volume financier se révèlent particulièrement performants en termes de contribution à la production nationale.

Ils illustrent ainsi la centralité des **sciences de la vie et de la santé** dans les retombées scientifiques associées aux investissements en infrastructures, et confirment que les plateformes lourdes

(cryo-microscopie, synchrotrons, imagerie cérébrale, dispositifs expérimentaux) constituent des **leviers stratégiques pour maintenir la visibilité internationale de la recherche française.**

#### Domaines à empreinte faible

À l'autre extrémité du spectre, plusieurs domaines affichent une empreinte très faible, voire marginale bien que non nulle :

- Mathematics - Mathematical Physics (2610) : 0,025%
- Mathematics - Applied Mathematics (2604) : 0,059%
- Mathematics - Numerical Analysis (2612) : 0,057%
- Business, Management and Accounting : Marketing (1406) : 0,09%
- Business, Management and Accounting - Management Information Systems (1404) : 0,011%
- Mathematics - Geometry and Topology (2608) : 0,15%

Ces résultats traduisent **apparemment une moindre dépendance aux équipements (disciplines théoriques, sciences de gestion)**. Dans ces domaines, le rôle des équipements est plus périphérique et la production repose davantage sur des approches conceptuelles ou qualitatives.

#### Analyse par grands groupes disciplinaires

##### Biochemistry, Genetics and Molecular Biology

- Structural Biology (1315) : 12,28%
- Biophysics (1304) : 9,72%
- Cell Biology (1307) : 5,56%
- Physiology (1314) : 5,54%
- Aging (1312) : 5,20%

Ce groupe domine nettement. **Les plateformes de pointe (cryo-microscopie, synchrotrons, imagerie structurale) constituent un vecteur central de production scientifique.** Ces disciplines apparaissent comme les plus dépendantes des équipements pour maintenir leur compétitivité.

##### Neuroscience

- Biological Psychiatry (2803) : 6,46%
- Developmental Neuroscience (2806) : 6,27%
- Neurology (2808) : 4,71%
- Cellular and Molecular Neuroscience (2804) : 3,97%

Les neurosciences affichent des scores élevés, traduisant le rôle **des plateformes d'imagerie cérébrale, d'électrophysiologie et d'expérimentation cognitive.** L'ANR dans le cadre de France 2030 contribue ainsi de manière déterminante à la structuration de ce domaine à fort enjeu de santé publique.

##### Agricultural and Biological Sciences

- Plant Science (1110) : 3,59%
- Horticulture (1108) : 3,60%
- Soil Science (1111) : 3,18%
- Insect Science (1109) : 3,15%
- Animal Science and Zoology (1103) : 4,57%

L'empreinte est ici homogène entre 4 et 7%. **Les projets d'équipement ANR dans le cadre de France 2030 soutiennent des infrastructures de terrain (serres, plateformes de phénotypage, stations expérimentales)** essentielles pour la recherche en durabilité, biodiversité et sécurité alimentaire.

### Environmental Science

- Ecology (2303) : 3,34%
- Environmental Chemistry (2304) : 2,49%
- Health, Toxicology and Mutagenesis (2307) : 2,46%
- Pollution (2310) : 2,45%

L'empreinte est homogène et significative, autour de 5 à 7%. Les projets ANR dans le cadre de France 2030 consolident ici des **observatoires et laboratoires analytiques**, renforçant la structuration des écosystèmes de recherche environnementale sur les pollutions, la toxicologie et les risques sanitaires.

### Computer Science, Engineering et Mathematics

- Human - Computer Interaction (1709) : 4,11%
- Computer Vision and Pattern Recognition (1707) : 1,84%
- Information Systems (1710) : 1,36%

Ce groupe montre une forte dispersion. Un segment bénéficie d'un effet notable des équipements mais la très grande majorité des domaines affichent une empreinte inférieure à 2%, ce qui reflète leur **moindre dépendance à des infrastructures matérielles et leur ancrage dans le logiciel, la donnée et la théorie.**

### Lecture transversale et enseignements

- Les sciences de la vie (biologie structurale, neurosciences, immunologie) sont les plus dépendantes des équipements, avec des scores supérieurs à 10% dans plusieurs domaines.
- Les sciences agricoles et environnementales affichent une empreinte intermédiaire (5 à 7%), confirmant que les infrastructures de terrain et d'observation soutiennent une part significative de la production française.
- Les sciences numériques et théoriques montrent des effets plus hétérogènes : un pic élevé en Human - Computer Interaction (1709), mais une majorité de domaines sous 2%, ce qui traduit une dépendance beaucoup plus limitée aux équipements.

En résumé, comme pour **l'empreinte des dépenses de l'ANR dans le cadre de France 2030 en général, le facteur déterminant semble plus lié à la discipline qu'au montant investi.** Les équipements financés par l'ANR dans le cadre de France 2030 structurent massivement les disciplines expérimentales (biologie, neurosciences, environnement), tandis que leur rôle est plus ciblé et hétérogène dans les sciences numériques et théoriques.

### Empreinte Avant-Après

L'analyse des effets « avant / après » met en évidence une dynamique globalement favorable de la production scientifique associée aux projets d'équipement financés par l'ANR.

## Bilan global

Sur les 228 domaines étudiés, 197 présentent un rendement supérieur à 0 % après investissement, contre 31 avec un rendement nul ou négatif. Autrement dit, près de 86 % des domaines connaissent une hausse de leur production scientifique, tandis qu'environ 14 % stagnent ou reculent.

En moyenne, le nombre d'articles publiés par an passe de 70,5 (avant) à 109,2 (après), soit un gain absolu de 38,7 articles/an, correspondant à un rendement moyen d'environ +54,9 % sur la période observée.

## Lecture par grandes familles dans le Top 50

La moyenne des taux de croissance par macro-discipline, calculée sur les 50 domaines connaissant les progressions les plus marquées (Top 50), met en évidence les résultats suivants (le nombre d'entrées est précisé pour limiter les effets de sur-interprétation dans les familles peu représentées) :

- Neurosciences (7) : +104,4 %
- Médecine (23) : +85,5 %
- Biochimie, Génétique et Biologie Moléculaire (8) : +82,8 %
- Sciences Agricoles et Biologiques (7) : +96,8 %

Ces résultats soulignent que les disciplines à forte intensité instrumentale (plateformes cliniques et expérimentales, infrastructures analytiques, dispositifs d'observation) concentrent les rendements les plus élevés. La Neurosciences dépasse en moyenne les +100 %, tandis que les disciplines *Agricultural and Biological Sciences* et *Médecine* enregistrent des gains compris entre +85 % et +97 %, et que *Biochimie, Génétique et Biologie Moléculaire* se situe autour de +83 %.

## Enseignement de pilotage.

La forte proportion de rendements positifs (197/228) ainsi que l'augmentation moyenne de la production annuelle (+38,7 articles/an, soit +54,9 %) confirment l'effet levier des équipements lorsqu'ils sont adossés à des communautés scientifiques déjà structurées.

Il semble encore une fois que l'effet des dépenses soit fortement lié à la discipline dans laquelle la dépense est réalisée. **Des disciplines à forte concentration en instrument occupent les premières places du classement (Top 50), ce qui plaide pour la pérennisation des dépenses dans ces disciplines.** La présence de 31 domaines à rendement nul ou négatif appelle un suivi différencié, notamment par une qualification des causes (délais de mise en service, inadéquation entre équipements et besoins, effets de composition des projets) et, le cas échéant, des ajustements (accompagnement méthodologique, ouverture renforcée, articulation inter-sites).

## Précautions d'interprétation.

Les moyennes par macro-discipline mentionnées ci-dessus portent exclusivement sur les domaines du Top 50, c'est-à-dire ceux présentant les croissances les plus significatives. Elles ne doivent donc pas être extrapolées à l'ensemble des disciplines.

Par ailleurs, l'indicateur de rendement mesure un différentiel « avant / après » en fonction de la date des investissements en équipements, sans que cela implique une causalité directe et exclusive : d'autres facteurs tels que les effets de cohorte, les dynamiques propres à chaque discipline ou le cumul de financements peuvent également contribuer aux hausses observées.

## Les dépenses en équipement

Les projets d'équipements financés par l'ANR dans le cadre de France 2030 représentent un investissement global de **2,2 milliards d'euros**.

### La répartition des dépenses par cadran

La ventilation des **dépenses d'équipement** par cadran stratégique met en évidence des arbitrages différenciés, cohérents avec la structuration des domaines scientifiques et technologiques. **Attention**, un même investissement peut être rattaché à **plusieurs domaines** situés dans **des cadrans différents**. Par construction, les montants présentés **par cadran ne sont donc pas exclusifs** et leur **somme dépasse le total** des dépenses exécutées.

- **Cadran 1 : 277 M€ (13% du total)**  
Les domaines du cadran 1 reçoivent un volume d'équipements plus modéré. Ces montants permettent de maintenir un socle d'infrastructures dans des domaines matures et déjà bien établis. L'objectif de ces dépenses devrait avant tout être de maintenir les équipements pour conserver savoir : faire et connaissances.
- **Cadran 2 : 1,23 Md€ (55% du total)**  
Les domaines classés en cadran 2 concentrent plus de la moitié des dépenses d'équipement. Ce niveau d'investissement traduit la volonté de consolider des champs à la fois dynamiques et proches de la frontière scientifique. L'effet attendu est double : renforcer la compétitivité internationale des communautés concernées et sécuriser les retombées industrielles.
- **Cadran 3 : 224 M€ (10% du total)**  
Même ces domaines, caractérisés par un dynamisme plus faible, continuent de mobiliser près d'un quart de milliard d'euros. Ce choix peut se justifier dans une logique de couverture large de l'ANR, mais pose la question de l'efficacité : dans certains cas, ces dépenses pourraient relever d'une logique de soutien minimal à des communautés fragiles ou de maintien de niches spécialisées.
- **Cadran 4 : 468 M€ (21% du total)**  
Les domaines du cadran 4 bénéficient d'un investissement relativement important, représentant environ un cinquième du total. Cet engagement permet de doter de moyens substantiels des domaines émergents mais encore éloignés de la frontière. Cela devrait leur permettre de franchir des jalons critiques. Toutefois, le rendement scientifique et technologique de ces investissements reste à confirmer, et appelle à un suivi attentif.

### Le rendement des dépenses en équipement

L'analyse des **rendements** des projets d'équipement financés par l'ANR dans le cadre de France 2030 met en évidence des **contrastes marqués entre disciplines**, traduisant des logiques différentes dans l'usage des infrastructures et leur conversion en production scientifique.

En premier lieu, l'examen des domaines les **plus consommateurs de crédits** montre que des investissements massifs ne se traduisent pas toujours par des **rendements élevés**. Certains domaines concentrent des budgets très importants, souvent supérieurs à 70 millions d'euros, pour des **rendements relativement modestes**. C'est le cas de l'informatique, avec un **rendement de 2,36 %** pour Computer Science - Information Systems (1710), malgré un budget de **110,3 millions d'euros**. De même, les sciences de la Terre et de l'environnement présentent des rendements comparables ou inférieurs, avec **environ 2 %** pour Environmental Science - Global and Planetary Change (2306), financé à hauteur de **76,2 millions d'euros**, et **moins de 2 %** pour Earth and Planetary Sciences - Atmospheric Science (1902), financé à **82,3 millions d'euros**.

À l'inverse, certains domaines maintiennent un **rendement significatif** malgré des enveloppes budgétaires importantes. C'est le cas de Biochemistry, Genetics and Molecular Biology - Molecular Biology (1312), qui atteint un **rendement de 6,74 %** pour **82,6 millions d'euros** de dépenses. Cette hétérogénéité témoigne à la fois d'une efficacité différenciée dans la conversion des équipements en résultats scientifiques et de **temporalités variables de production** selon les disciplines.

À l'autre extrémité du spectre plusieurs domaines **modestement financés** affichent des **rendements proportionnellement très élevés**. Arts and Humanities - Visual Arts and Performing Arts (1213) enregistre un **rendement de 3,57 %** pour seulement **84,9 milliers d'euros**, tandis que Social Sciences - Public Administration (3321) atteint **2,99 %** pour un niveau de dépense similaire. Ces résultats soulignent l'empreinte potentielle de **micro-financements bien ciblés**, qui peuvent produire un **effet structurant fort** sur la capacité de publication des communautés concernées. Ils plaident pour une **reconnaissance du rôle stratégique** de ces investissements proportionnellement très rentables.

Les **meilleurs rendements** sont toutefois enregistrés dans certaines disciplines expérimentales, notamment les **sciences de la vie**. Biochemistry, Genetics and Molecular Biology - Structural Biology (1315) affiche un **rendement de 19,13 %** pour **30 millions d'euros** de dépenses. Biophysics (1304) atteint **17,01 %** pour **25 millions d'euros**, et Cell Biology (1307), **8,77 %** pour environ **20 millions d'euros**. Ces **rendements très élevés** illustrent le **rôle central des équipements lourds** tels que la cryo-microscopie, les lignes de synchrotron ou les plateformes biologiques dans la **production scientifique française** et dans sa **visibilité internationale**.

Les **neurosciences** confirment également cette dynamique avec des **rendements situés entre 5 % et 12 %**, traduisant l'efficacité des plateformes d'imagerie cérébrale, d'électrophysiologie ou d'expérimentation cognitive. Ainsi, Neuroscience - Biological Psychiatry (2803) affiche un **rendement de 11,52 %**, et Developmental Neuroscience (2806), **9,08 %**.

Les **sciences agricoles et environnementales** se situent quant à elles dans une position **intermédiaire**. Leur rendement est compris entre **4,5 % et 6,7 %**, avec notamment **6,65 %** en Soil Science (1111), **6,40 %** en Plant Science (1110), et **6,77 %** en Ecology (2303). Ces niveaux suggèrent que, même si les besoins en équipement sont importants (serres expérimentales, infrastructures de terrain, observatoires), leur **conversion en publications** est **moins immédiate** que dans les disciplines biomédicales.

Enfin, les **disciplines numériques et mathématiques** présentent un **profil contrasté**. Certaines niches, comme Human - Computer Interaction (1709), affichent un **rendement de 7,94 %**, proche de celui des sciences expérimentales. À l'inverse, de nombreux domaines présentent des rendements **inférieurs à 3 %**, voire **très faibles**. C'est le cas de Hardware and Architecture (1708), qui enregistre un **rendement de seulement 0,22 %** pour **43,2 millions d'euros** de dépenses, l'un des plus faibles de l'échantillon. Cette disparité s'explique en partie par la **nature des activités** : certains projets informatiques nécessitent des plateformes matérielles complexes, tandis que d'autres relèvent davantage de la modélisation, du développement logiciel ou de l'approche théorique, **moins directement corrélés aux investissements en équipements**.

Au total, l'analyse met en évidence une **double dynamique**. D'un côté, les **sciences de la vie**, les **neurosciences** et plusieurs **disciplines expérimentales** se caractérisent par des **rendements élevés**, qui légitiment pleinement les investissements réalisés. D'un autre côté, certains champs **fortement consommateurs de crédits**, notamment dans le **numérique** et les **sciences de l'environnement**, présentent des rendements **plus faibles**, appelant à un **suivi différencié** et à l'intégration d'indicateurs complémentaires (production de données, innovation logicielle, transfert technologique, production de normes ou effets de long terme).

Cette lecture plaide pour une **stratégie de pilotage différenciée** :

- **Maintenir et consolider** les financements dans les domaines à **fort rendement**, notamment les sciences de la vie et les neurosciences ;
- **Renforcer le ciblage** des investissements modestes dans les disciplines à **haut rendement proportionnel**, y compris certaines sciences humaines et sociales ;
- **Réexaminer les modalités d'évaluation** et d'accompagnement des secteurs à **faible rendement** en publications, en élargissant les critères d'impact **au-delà du seul indicateur bibliométrique**.

Ce cadre d'analyse permet de **renforcer la cohérence** et **l'efficacité** des politiques d'équipement scientifique, en **ajustant l'allocation des ressources aux spécificités de chaque discipline**.

## 1. Recommandations pour la collecte des données

### Amélioration des formats de reporting des bénéficiaires

Il serait pertinent de mettre en place un formulaire en ligne pour recueillir les informations auprès des bénéficiaires, en remplacement des formulaires au format PDF. Une telle approche permettrait de standardiser la collecte des données, tout en garantissant leur exhaustivité. Cela ouvrirait la voie à un traitement automatisé et homogène pour l'ensemble des projets, ce qui faciliterait considérablement l'analyse et la comparaison. Dans cette logique de standardisation, il serait également recommandé de privilégier des questions à choix prédéfinis plutôt que de recourir systématiquement à des zones de texte libre. Cette pratique limiterait les divergences d'interprétation, améliorerait la qualité des données recueillies et simplifierait leur exploitation.

### Normalisation des formats et codes des sources

Nous recommandons de standardiser les formats de codification des différentes sources (brevets, publications, codes laboratoires, codes projets) afin d'assurer une meilleure cohérence et de faciliter leur exploitation. En particulier:

- **Code projets** : généralisation de l'usage des codes EOTP des projets dans les documents à la fois pour les bénéficiaires et les autres services de l'Etat comme les SATT.
- **Numéros RNSR** : collecte systématique avec un format contraint des numéros RNSR des laboratoires participant aux projets. En effet, nous avons observé d'autres numéros que des RNSR (siret notamment) dans ce champ. Un formulaire en ligne vérifiant que le numéro correspond à un format RNSR serait l'idéal.
- **Numéros de brevets** : collecte des numéros de publication de brevets standards, en veillant à intégrer les informations spécifiques qui permettent d'identifier un brevet en précisant bien le type de numéro demandé (numéro de demande, numéro de publication, numéro de priorité, appartenance à une famille, etc). Il est également nécessaire de renseigner les métadonnées essentielles comme l'autorité de dépôt et la date de publication, afin de garantir la traçabilité et de faciliter la recherche dans les bases de données spécialisées.
- **Numéros de publications scientifiques**: collecte systématiques des identifiants universels (DOI<sup>3</sup>) liés aux articles scientifiques publiés dans le cadre d'un projet financé par l'ANR. Cependant, à cela peut se substituer par l'intégration plus grande avec la plateforme HAL présentée dans la partie suivante.

### Intensifier les synergies avec la plateforme HAL

Une intégration entre les projets financés par l'ANR dans le cadre de France 2030 et la plateforme HAL est déjà présente grâce aux informations de financement lorsque l'auteur déclare sa publication sur la plateforme. Cependant, il serait souhaitable d'inciter les auteurs à remplir systématiquement tous les champs requis pour l'ensemble des publications et communications réalisées dans le cadre

<sup>3</sup> <https://www.doi.org>

de ces projets, et plus largement pour l'ensemble de la production scientifique française. Un tel niveau de complétude constituerait un atout majeur pour l'analyse automatisée des données. Il serait également important de renforcer ces correspondances, car des lacunes subsistent, notamment concernant les thèses et les informations sur les institutions. Cette incitation pourrait, par exemple, passer par une intégration de l'API HAL dans les formulaires remplis par les bénéficiaires des financements afin de minimiser leur temps de travail administratif.

## Création d'un outil de suivi au fil de l'eau

La normalisation des données précédemment décrites permettrait de valoriser et d'automatiser le travail réalisé dans ces deux études et de le rendre compatible avec le reste de la Mission Interministérielle Recherche et Enseignement Supérieur (MIREs). Pour en tirer pleinement parti, il serait pertinent de mettre en place un outil de suivi en ligne, accessible à tous les services concernés, offrant :

- **Un traitement homogène de tous les projets** : chaque type de projet serait suivi selon une méthodologie et des formats standardisés. Cela permettrait non seulement de comparer les projets entre eux mais aussi d'identifier rapidement les tendances et les écarts, réduisant les asymétries d'information liées à la multiplicité des formats et des pratiques.
- **Des mises à jour en temps réel et centralisées** : toutes les informations saisies par les bénéficiaires seraient immédiatement disponibles pour l'ensemble des services concernés. Cela limiterait les décalages d'information entre les services, améliorerait la réactivité dans la prise de décision et faciliterait la détection rapide des anomalies ou des données manquantes.
- **Une accessibilité partagée et sécurisée** : l'outil serait accessible à tous les services autorisés, tout en garantissant la sécurité et la confidentialité des données. Cette centralisation simplifierait la collaboration interministérielle et réduirait le besoin de multiplications des fichiers ou des échanges manuels entre services.
- **Une incitation forte pour les chercheurs à renseigner les données** : l'outil générerait automatiquement des rapports et tableaux de bord utilisables par les chercheurs eux-mêmes, donnant un retour concret sur l'usage de leurs données et pouvant faciliter leur travaux en matière d'état de l'art ou de positionnement scientifique. Cette valeur ajoutée encouragerait une saisie complète et correcte des informations, tout en réduisant leur charge administrative.
- **Une automatisation des rapports et analyses** : grâce à la standardisation et à la centralisation des données, l'outil pourrait produire automatiquement des rapports harmonisés et comparables entre projets, laboratoires et types de financements. Cela faciliterait la communication des résultats, la valorisation des projets et la production de statistiques fiables pour l'ensemble de la MIREs.
- **Une exploitation étendue des synergies avec HAL et autres bases** : l'outil pourrait intégrer directement les API des plateformes scientifiques comme HAL pour récupérer automatiquement les publications et communications, garantissant une couverture exhaustive et à jour des résultats scientifiques. Cela permettrait d'automatiser la correspondance entre financements et productions scientifiques et d'améliorer la qualité globale de l'analyse.
- **Une dimension temporelle et de pilotage stratégique** : l'outil pourrait conserver un historique des modifications, publications, résultats et changements administratifs sur toute la durée de vie des projets. Cette capacité permettrait de reconstituer la trajectoire complète d'un financement, d'en comprendre les évolutions et d'appuyer des analyses rétrospectives ou des évaluations ex post.
- **Une interopérabilité avec les standards européens** : l'outil pourrait s'aligner sur les formats de données recommandés par la Commission Européenne (notamment la taxonomie EIC) afin de faciliter les comparaisons internationales et d'anticiper d'éventuelles obligations de reporting à venir.

- **Un module qualitatif intégré** : au-delà des données quantitatives, un espace pourrait permettre la saisie (encadrée) d'éléments qualitatifs sur les retombées des projets (usages industriels, politiques publiques, formations, transferts, etc.), facilitant l'analyse d'impact au-delà des publications, brevets ou autre données quantitatives.

L'ensemble des recommandations présentées (harmonisation des formats, normalisation des codes et métadonnées, intégration avec la plateforme HAL et mise en place d'un outil de suivi en ligne) vise à rendre la collecte et l'exploitation des données plus cohérentes, complètes et automatisables. En structurant les informations de manière standardisée et en facilitant l'accès aux données pour tous les acteurs concernés, ces mesures permettraient non seulement de valoriser les résultats des projets financés par l'ANR, mais aussi de renforcer l'efficacité des analyses et la communication des résultats. La mise en place d'un outil centralisé de suivi constitue un levier essentiel pour transformer les données collectées en informations exploitables, tout en réduisant la charge administrative des chercheurs et en favorisant une meilleure coordination entre les services.

## 2. Recommandations opérationnelles et politiques

### Orientations pour l'ANR dans le cadre de France 2030

L'analyse conduite dans le cadre de cette mission met en lumière l'importance d'une gouvernance de l'action publique en matière de financement de la recherche et de l'innovation. L'architecture de données désormais disponible constitue une base solide pour un pilotage plus fin, plus réactif et plus transparent des financements.

Premièrement, le croisement des deux indicateurs structurants (la **proximité à la frontière scientifique** et la **dynamique d'émergence**) permet d'identifier des domaines à fort potentiel stratégique encore peu soutenus, ou au contraire des domaines bénéficiant d'un financement conséquent mais présentant un faible retour. Cette lecture stratégique facilite une orientation plus ciblée des appels à projets, qu'il s'agisse de renforcer les domaines les plus prometteurs ou de redéfinir les modalités d'investissement dans les secteurs à faible effet de levier.

Deuxièmement, la consolidation des données relatives aux **équipements scientifiques** autorise une lecture transversale et homogène des actions INBS, EQPX, COHO ou ESRE. Cette nouvelle visibilité milite pour la mise en place d'un **pilotage inter-opérationnel des investissements d'infrastructure**, articulé avec les objectifs de France 2030. Elle permettrait d'optimiser l'allocation des ressources, d'accroître les effets d'entraînement sur les projets associés et de renforcer la cohérence stratégique des soutiens accordés. Elle offre également à l'ANR dans le cadre de France 2030 des capacités renouvelées de transparence et de redevabilité vis-à-vis de ses tutelles, du Parlement et des acteurs du système national d'innovation.

Enfin, cette mission confirme l'intérêt d'un **suivi continu et centralisé des données à l'échelle projet**, couplé à une meilleure coordination interministérielle. Elle plaide pour la mise en œuvre d'un **outil pérenne**, adossé à l'infrastructure de l'ANR, permettant une **collecte, une exploitation et une restitution dynamiques des données** (voir propositions techniques dans le rapport). Un tel dispositif constituerait un levier stratégique pour améliorer le pilotage de la recherche française, évaluer l'efficacité des financements publics, et soutenir les ambitions de souveraineté scientifique et technologique.

Les investissements de l'ANR dans le cadre de France 2030 touchent quasiment l'ensemble des sphères de connaissances scientifiques. En effet, l'ensemble des projets financés couvrent 98% de la classification OpenAlex, ce qui illustre l'importante diversité des thématiques soutenues. Cette

couverture très large reflète toutefois une répartition des financements qui, si elle embrasse presque toutes les disciplines, peut aussi diluer l'impact sur certains domaines stratégiques ou émergents.

L'action de l'ANR dans le cadre de France 2030 exerce globalement une empreinte **scientifique positive, avec des taux de production et une influence académique soutenus dans la majorité des domaines**. Toutefois, cette empreinte demeure hétérogène selon les thématiques, les types de projets et la maturité des écosystèmes concernés. Il convient également de souligner que la méthodologie adoptée, centrée sur les données issues des projets financés, peut introduire un **biais de sélection favorable** à l'ANR, malgré les contrôles réalisés par domaine et par thématique.

En matière de **valorisation technologique**, les résultats apparaissent plus contrastés. Les taux d'empreinte sur les brevets restent globalement faibles, ce qui s'explique par la nature des projets soutenus, souvent orientés vers la recherche fondamentale. À cela s'ajoute une **sous-déclaration ou une sous-structuration des données liées à la propriété intellectuelle**, qui limite les capacités d'évaluation exhaustive.

Ce double constat appelle à une **meilleure distinction des finalités** des projets financés. Si les objectifs de production de connaissances et de valorisation technologique ne sont pas nécessairement antinomiques, ils supposent néanmoins des modalités d'accompagnement différenciées. En effet, la publication scientifique peut entrer en tension avec la stratégie de valorisation par le brevet, notamment pour les résultats à fort potentiel applicatif.

Dans cette perspective, l'ANR dans le cadre de France 2030 pourrait renforcer l'**orientation stratégique de ses financements** selon trois leviers complémentaires :

- **Différencier les modalités de suivi et d'évaluation** selon la nature et les objectifs des projets soutenus ;
- **Flécher plus explicitement les moyens permettant d'accompagner la transformation des résultats scientifiques en innovations**, en s'appuyant sur les partenaires industriels ou les dispositifs de soutien à l'innovation aval.

Cette différenciation contribuerait à résorber les tensions entre production de connaissances et valorisation, tout en renforçant la lisibilité, la cohérence et la performance de l'action publique dans le champ de la recherche et de l'innovation.

## Valorisation auprès des acteurs académiques et industriels

Les résultats de ces deux études offrent également des perspectives concrètes de valorisation pour l'écosystème national, au-delà des usages internes à l'ANR dans le cadre de France 2030.

Pour les acteurs académiques, la structuration fine des domaines technologiques, l'analyse des trajectoires de production scientifique, et les métriques de positionnement international constituent des ressources précieuses pour se situer dans leur environnement, ajuster leur stratégie de publication et identifier des opportunités de collaboration, y compris intersectorielles. En facilitant l'accès à des analyses consolidées par domaine ou par région, cette mission ouvre la voie à un dialogue renforcé entre chercheurs, établissements et agences de financement.

Pour les acteurs industriels, les résultats agrégés et comparatifs de cette cartographie offrent une grille de lecture structurée des technologies de rupture soutenues par la puissance publique. Ils peuvent ainsi mieux identifier les domaines en émergence, les zones de faible densité technologique à combler, ou encore les champs porteurs de partenariats public-privé. L'analyse de la capacité d'absorption par les entreprises, croisée aux dynamiques d'empreinte scientifique, permet par ailleurs

de renforcer l'alignement entre les priorités de l'ANR dans le cadre de France 2030 et les besoins du tissu productif national.

De manière plus large, la mise à disposition d'un applicatif interactif constitue une innovation majeure en matière de valorisation. Il pourra servir d'interface commune entre financeur, bénéficiaires, évaluateurs et partenaires. Son usage dans le cadre des dialogues stratégiques, des revues d'orientation de programme ou des exercices de prospective collective permettra d'enrichir les processus décisionnels à tous les niveaux.

## Recommandations par domaine

L'analyse croisée des dynamiques scientifiques, des investissements réalisés et des indicateurs de performance permet de formuler plusieurs recommandations spécifiques selon les grands domaines disciplinaires, dans une logique de pilotage différencié. Ces recommandations visent à renforcer l'efficacité des investissements, à optimiser les effets leviers des équipements, et à orienter les choix futurs de financement en fonction des dynamiques observées.

### Sciences de la vie : renforcer les plateformes à forte intensité technologique

Les domaines relevant de la biologie moléculaire, de la génétique, des neurosciences et de la médecine présentent des rendements particulièrement élevés, tant en termes de production scientifique que de centralité internationale. La structuration autour de plateformes lourdes (cryo-microscopie, imagerie biomédicale, plateformes expérimentales) explique en grande partie cette performance. Les données confirment l'importance de ces équipements pour maintenir l'attractivité et la compétitivité de la recherche française.

**Recommandation** : Pérenniser et mutualiser les plateformes à forte intensité instrumentale, en accompagnant les investissements initiaux par des dispositifs de maintenance, de gouvernance des données et de soutien technique aux équipes utilisatrices. Une attention particulière doit être portée aux disciplines comme *Structural Biology*, *Molecular Medicine* ou *Biological Psychiatry*, qui présentent à la fois un fort rendement et une centralité élevée.

### Sciences de l'ingénierie et des matériaux : cibler les filières à forte capacité d'absorption

Les domaines de l'ingénierie (mécanique, électronique, automobile) et des matériaux (polymères, surfaces) concentrent les plus fortes capacités d'absorption, traduisant une forte intégration dans les filières industrielles stratégiques. Ces disciplines présentent une valorisation technologique élevée, notamment par le biais des brevets, et un potentiel de transfert important.

**Recommandation** : Renforcer les dispositifs d'articulation entre recherche académique et industrie, en soutenant les projets à forte proximité sémantique entre publications et brevets, et en priorisant les partenariats sur les segments les plus structurants (ex. : *Electrical and Electronic Engineering*, *Mechanical Engineering*, *Polymers and Plastics*). Les indicateurs d'absorption doivent guider les appels à projets pour maximiser l'effet levier en matière d'innovation industrielle.

### Mathématiques et informatique : piloter selon les temporalités de valorisation

Les mathématiques appliquées, la simulation numérique et l'intelligence artificielle apparaissent comme des domaines à très forte centralité, moteur de la structuration des réseaux scientifiques internationaux. Néanmoins, la valorisation directe via les brevets y demeure relativement faible, reflétant des temporalités plus longues ou des formes de diffusion non propriétaires.

**Recommandation** : Maintenir un soutien fort aux domaines méthodologiques à haute centralité (ex. : *Modeling and Simulation*, *Computational Mathematics*, *AI and Computer Vision*), tout en développant des instruments de suivi adaptés à leurs logiques de valorisation (open source, logiciels, transferts implicites). Le rendement des investissements doit être évalué sur des horizons longs et en tenant compte de l'effet d'entraînement transversal.

#### Sciences de l'environnement et de la planète : mieux calibrer les investissements

Certains domaines des sciences de l'environnement présentent un déséquilibre entre l'ampleur des investissements et les rendements observés. Par exemple, *Global and Planetary Change* et *Atmospheric Science* ont mobilisé des budgets significatifs ( $\geq 75$  M€) pour un rendement moyen inférieur à 2 %, suggérant une montée en charge plus lente ou une inadéquation partielle entre équipements financés et besoins scientifiques.

**Recommandation** : Procéder à une analyse approfondie des temporalités de production scientifique dans ces domaines, afin d'identifier les effets de composition, les éventuelles redondances ou les écarts entre besoin et équipement. Des mécanismes de suivi différencié et de réallocation partielle des crédits pourraient permettre d'ajuster les futurs investissements aux trajectoires réelles de structuration des communautés.

#### Sciences humaines et sociales : valoriser les micro-investissements à forte empreinte

Les SHS présentent des niveaux d'équipement modestes, mais certains domaines affichent des rendements élevés rapportés aux montants engagés. Ces résultats confirment l'effet structurant de micro-financements bien ciblés. Néanmoins leur empreinte est globalement bien plus restreinte surtout quand on s'attarde sur la valorisation des brevets ou l'empreinte en équipements. Ceci est fortement lié à la nature même des disciplines concernées.

**Recommandation** : adapter l'analyse et l'évaluation des SHS en tenant compte de leurs spécificités propres. Les indicateurs utilisés pour les disciplines expérimentales ou technologiques (brevets, intensité d'équipement) ne reflètent pas toujours l'empreinte réelle de ces sciences. Il conviendrait d'intégrer des métriques complémentaires (structuration des réseaux de recherche, production de données ou de ressources numériques) afin d'appréhender de manière plus juste l'efficacité des investissements dans les SHS et d'en optimiser le pilotage futur.

## CONCLUSION

Le présent rapport propose une lecture intégrée des effets de l'ANR dans le cadre de France 2030 en croisant deux volets complémentaires : **la qualification des technologies de pointe et des dynamiques d'émergence, et le bilan des équipements de recherche de taille intermédiaire**. Il s'appuie sur un cadre méthodologique robuste combinant analyse sémantique (IA) et théorie des graphes, avec deux indicateurs centraux (proximité à la frontière scientifique et émergence) définis et mobilisés de manière homogène à l'échelle des domaines, en référence à des corpus France/Monde construits spécifiquement pour ces deux études. Cette approche unifiée **a permis d'objectiver le positionnement des projets, d'en mesurer l'apport relatif au regard de l'état de l'art international, et d'identifier les trajectoires de structuration des connaissances**.

Sur le plan des **objectifs et du périmètre**, le rapport répond aux questions de mesure de l'empreinte scientifique de l'ANR, d'identification des domaines porteurs d'émergence, et d'évaluation du rôle structurant des équipements, avec une déclinaison systématique des résultats au niveau des domaines et des comparaisons France/Monde. **L'architecture du rapport a été conçue pour relier les objectifs des deux études, la nature et la provenance des données, la méthode, les résultats principaux et les recommandations opérationnelles, assurant ainsi une continuité analytique et décisionnelle**.

La **méthodologie** distingue précisément les « briques » de connaissances et leur « architecture » (combinaisons et flux), afin d'apprécier à la fois la cohérence interne des domaines et leur centralité dans les réseaux de citations. L'indicateur d'émergence repose sur le rang de variation de centralité (fenêtres temporelles glissantes, contrôle par volumes), et l'indicateur de frontière mesure la distance de cohérence à une référence mondiale (architecture minimale la plus robuste). **Ces choix évitent de confondre volumes et qualité, croissance et impact**, et offrent des critères comparables entre domaines, tout en explicitant les limites (sensibilité des petits corpus, non-linéarité des trajectoires, caractère théorique de la frontière).

Les **résultats** confirment une diversité de trajectoires, synthétisée par la matrice « émergence × proximité à la frontière » : cadran 1 (proches de la frontière mais moins émergentes), cadran 2 (dynamiques et proches de la frontière), cadran 3 (faibles sur les deux axes) et cadran 4 (émergentes mais encore éloignées de la frontière). La répartition observée montre une concentration des projets, publications et investissements dans les cadrans les plus émergents (cadran 2 et cadran 4), ce qui reflète un mix d'excellence consolidée et de paris stratégiques sur des domaines en rattrapage. **Cette typologie n'a pas de portée normative : elle éclaire la priorisation mais laisse place aux arbitrages de souveraineté, de territoire ou d'opportunité socio-économique**.

Sur l'**empreinte scientifique** de l'ANR dans le cadre de France 2030, l'indicateur « part ANR dans le cadre de France 2030 » (exprimée en pourcentage de la production d'un domaine) **permet d'identifier des surperformances relatives au regard d'un repère budgétaire (les dépenses ANR dans le cadre de France 2030 représentent 5,5 % de la MIREs)**, tout en rappelant la précaution inhérente à la construction sémantique des référentiels (biais de sélection possible). La cartographie par pays et l'évolution temporelle (fenêtres glissantes) complètent cette mesure par des éléments de positionnement et de trajectoire.

Concernant les **équipements**, l'analyse différentielle « avant/après » met en évidence une dynamique globalement favorable : sur 228 domaines, 197 affichent un rendement positif, avec un passage moyen de 70,5 à 109,2 articles/an (+ 38,7 ;  $\approx$  + 54,9 %). **Les effets sont particulièrement marqués dans les disciplines à forte intensité instrumentale (biologie structurale, neurosciences, infrastructures analytiques), tandis que les sciences numériques et théoriques montrent des empreintes plus hétérogènes**, souvent < 2 % hors segments spécifiques (p. ex. *Human-Computer Interaction*). Les sciences agricoles et environnementales présentent des effets intermédiaires (5–7 %), cohérents avec des temporalités d'observation plus longues. Ces constats confortent l'idée que la *discipline*, plus que l'ampleur des montants, détermine le levier des équipements.

En **valorisation**, l'analyse articule la proximité sémantique entre publications et brevets, l'identification des familles technologiques et la capacité d'absorption des entreprises européennes. Cette grille de lecture offre aux acteurs industriels une vision des domaines en émergence, des « faiblesses » technologique à combler, et des terrains propices aux partenariats public-privé ; elle fournit parallèlement aux acteurs académiques des repères pour se situer, ajuster leurs stratégies de publication et renforcer les coopérations. **L'ouverture via l'application interactive fournie en complément constitue, de ce point de vue, un vecteur de dialogue et d'appropriation partagée des résultats.**

Les **recommandations** proposées déclinent ces constats : (i) ajuster les cadres d'évaluation aux profils disciplinaires (indicateurs alternatifs pour le numérique et les SHS, reconnaissance des effets immatériels), (ii) financer le **cycle de vie** des plateformes (au-delà de l'acquisition : maintenance, personnels, gouvernance des données), (iii) mettre en place un **suivi différencié** (jalons, fenêtres temporelles adaptées), (iv) généraliser la **mutualisation** inter-équipes et inter-sites pour maximiser l'usage marginal des infrastructures, et (v) **articuler systématiquement** les projets à des dispositifs de transfert afin d'accélérer la conversion des résultats en innovations. **Ces orientations répondent directement aux tensions identifiées entre production de connaissances et valorisation** et visent à renforcer l'efficacité de l'action publique.

Enfin, le rapport souligne des **limites** assumées et documentées (qualité hétérogène des données sources, biais possibles liés à la construction sémantique des référentiels, sensibilité des signaux dans les petits périmètres, non-linéarité des trajectoires d'émergence) qui appellent à **poursuivre l'effort de structuration et de normalisation des données et à consolider les outils de suivi « au fil de l'eau »**. La mise en œuvre des pistes opérationnelles proposées (collecte, formats, interfaces, applicatif interactif) doit permettre d'améliorer la robustesse, la pérennité et l'utilité décisionnelle de ces analyses pour les futurs millésimes.

Globalement, les résultats confirment l'existence d'un **effet structurant** des investissements de l'ANR dans le cadre de France 2030 sur les domaines expérimentaux et instrumentés, valident la pertinence d'une lecture par **cadres stratégiques** pour orienter l'allocation des moyens, et mettent en évidence la nécessité d'un **pilotage différencié** selon les familles disciplinaires et les finalités recherchées. En adossant la décision publique à des indicateurs transparents et comparables, et en dotant l'écosystème d'un outil de visualisation commun, l'ANR dans le cadre de France 2030 se dote d'une base solide pour **prioriser, suivre et ajuster ses investissements au service de la souveraineté scientifique et de la compétitivité technologique.**

## 1. Définitions des indicateurs

### Indicateurs structurants (réseau)

#### Cohérence (par domaine)

**Objet.** Cet indicateur mesure la structuration interne d'un domaine en évaluant la stabilité des combinaisons de connaissances qui le composent.

**Méthode.** Il repose sur une analyse de graphe des **cooccurrences de connaissances** identifiées dans les documents. En analysant l'écart entre les cooccurrences **observées** et les cooccurrences **théoriques** dans un contexte aléatoire, l'indicateur mesure la cohérence de chaque combinaison de connaissances et permet ainsi **de comparer les niveaux de cohérence** dans le temps, **entre domaines** ou au sein de n'importe quel sous-ensemble (entreprises, **pays**, groupes de documents, etc.).

**Périmètre & niveau.** Le calcul s'effectue au niveau des **combinaisons de Topics** de la classification **OpenAlex**, avec un référentiel de publications scientifiques construit à l'échelle du monde et sur toute la période analysée.

**Sortie & lecture.** Le résultat est fourni sous forme de score ou de rang : un niveau élevé indique un domaine **consolidé, stable et mature**, tandis qu'un niveau faible signale une **fragmentation**, une **exploration** ou une **recomposition** technologique.

**Limites / précautions.** L'indicateur peut sous-évaluer des domaines très jeunes ou très transverses et reste sensible au périmètre de concepts retenu. Par ailleurs, un domaine au sein duquel le **nombre d'observations** est très réduit peut également être sous-évalué.

#### Centralité (par domaine/acteur)

**Objet.** Cet indicateur quantifie l'influence d'un domaine ou d'un acteur au sein des **flux de connaissances**, C'est-à-dire au sein des réseaux de citations (articles).

**Méthode.** Il est construit à partir d'une **représentation de référence en graphe**, sous forme **matricielle**, reprenant l'ensemble des flux de citations au sein des publications scientifiques. La centralité d'une partie (domaine ou acteur) est **mesurée à partir du poids réel des citations** attribuables à cette partie dans l'ensemble du graphe. Il en résulte une **mesure directe** (et non une estimation) de l'influence de cette partie dans le réseau des publications. Cette influence peut ensuite être analysée à différentes échelles (nationale, régionale, mondiale).

**Périmètre & niveau.** Le calcul est produit au niveau de chaque domaine (**subfield OpenAlex**), avec un référentiel de publications scientifiques construit à l'échelle du monde et sur toute la période analysée.

**Sortie & lecture.** Le résultat est livré en **score** : une valeur élevée traduit un **rôle d'influence et d'intermédiation** (centralité d'intermédiarité) entre différents domaines au sein du réseau de connaissances.

**Limites / précautions.** La mesure dépend de la fenêtre temporelle sélectionnée. Ici, afin de mesurer des évolutions, **une agrégation annuelle a été réalisée**, puis **une mesure des variations sur des fenêtres de trois années glissantes** a permis de **lisser les effets de bord** liés à ce découpage.

## Indicateurs composites (dynamiques & positionnement)

### Émergence (par domaine)

**Objet.** L'indicateur vise à détecter les domaines **dont le développement s'accélère et atteint un niveau élevé**.

**Méthode.** Il s'agit d'une mesure **relative** fondée sur le **rang de la variation de centralité** observée dans une fenêtre temporelle glissante, **contrôlée par le rang du volume de publications** afin de limiter les faux signaux sur faibles volumes ; la restitution s'effectue en rangs ou en quantiles.

**Périmètre & niveau.** Le calcul s'applique au niveau de **chaque domaine** analysé, sur des séries temporelles.

**Sortie & lecture.** Un **rang composite élevé** signale un domaine émergent, combinant **progression rapide d'influence et volume suffisant**.

**Limites / précautions.** Le bruit peut persister dans les plus petits domaines ou ceux dont les frontières sont très **floues** et **mouvantes**. Par ailleurs, l'émergence n'étant pas un processus linéaire, elle peut se composer d'une succession de phases d'**accélération**, de **stagnation** ou de **régression** ; il est donc important de prévoir une fenêtre de calcul suffisamment large pour englober l'intégralité d'un cycle. Il reste toutefois possible que ces fluctuations de rythme **minimisent** ou **accentuent** le score d'émergence.

### Frontière scientifique (par domaine)

**Objet.** L'indicateur mesure la **proximité d'un domaine avec l'état de l'art mondial**.

**Méthode.** Il évalue l'écart de **cohérence** entre, d'une part, la valeur **maximale** observée à l'échelle mondiale et, d'autre part, la cohérence calculée pour les projets ANR dans le cadre de France 2030 dans le même domaine. La valeur maximale correspond à la **cohérence la plus forte calculable en conservant le minimum de liaisons de connaissances**, ce qui permet d'identifier l'**architecture de connaissances la plus robuste**. La **distance à la frontière scientifique** est alors la différence entre cette valeur de référence et la cohérence observée pour l'ANR.

**Périmètre & niveau.** L'analyse est conduite au niveau du **domaine**, avec un **référentiel technologique mondial** construit sur toute la période de l'analyse.

**Sortie & lecture.** Le résultat est exprimé sous la forme d'un **score défini comme 1 moins la distance à la frontière** : plus le score est proche de **1**, plus le domaine est **proche de la frontière scientifique** ; un score proche de **0** indique un **éloignement** important.

**Limites / précautions.** La **frontière** ainsi définie est **théorique** : elle reflète une **maturité structurale** (cohérence de l'architecture des connaissances) **sans renseigner en absolu la performance technologique**. Par ailleurs, un **contrôle sur le nombre de documents** est appliqué : un volume

trop faible, même extrêmement cohérent, ne peut positionner à lui seul un domaine au plus près de la frontière. De même, afin d'éviter les biais associés à des productions reposant sur un nombre trop limité de connaissances mobilisées, un seuil minimal est appliqué. Concrètement, l'indicateur n'est calculé que lorsque les projets financés par ANR dans le cadre de France 2030 couvrent une part significative des connaissances du domaine, fixée ici à 25%. Ce choix vise à garantir que la mesure repose sur une base suffisamment représentative et non sur quelques contributions isolées, qui pourraient donner une image faussée de la structuration réelle du domaine.

## Capacité d'absorption

### Capacité d'absorption (entreprise par domaine)

**Objet.** L'indicateur identifie les **entreprises** les mieux positionnées pour **internaliser** les connaissances d'un domaine, en vue de **transferts** ou de **partenariats**.

**Méthode.** Il combine un **score de proximité** entreprise/domaine (fondé sur des proximités sémantiques et technologiques) et un **rang de cohérence** de l'entreprise dans le domaine, agrégés en un **classement synthétique**.

**Périmètre & niveau.** Le calcul s'effectue au niveau de l'**entreprise** par domaine, en **France** et/ou au niveau **européen** (Union européenne ainsi que **Royaume-Uni, Suisse, Norvège et Islande**).

**Sortie & lecture.** Le résultat se présente sous forme de **rang** et de **score global** pour l'ensemble du domaine, **exprimé en pourcentage** ; une position élevée d'une entreprise dans le classement désigne un **bon candidat** pour l'absorption ou la collaboration dans le domaine considéré.

**Limites / précautions.** La **qualité** et le **volume** des traces publiques varient selon les entreprises, et des **biais sectoriels** peuvent subsister.

## Indicateurs d'empreinte (contribution ANR dans le cadre de France 2030)

### Empreinte ANR dans le cadre de France 2030 (par domaine)

**Objet.** L'indicateur apprécie la **contribution relative** de l'ANR dans le cadre de France 2030 à la **production scientifique** d'un domaine.

**Méthode.** Il calcule la part des publications (article scientifique pour l'empreinte scientifique et brevet pour l'empreinte technologique) **liées aux projets ANR**. Ces publications sont soit directement celles **déclarées** dans les rapports remis à l'ANR par les porteurs de projets, soit, pour les publications scientifiques, **toutes celles appartenant** au même domaine et **publiées par un laboratoire financé par l'ANR**. Ce **nombre de publications** est ensuite rapporté au **total du domaine**, dans le référentiel **Monde** ou **France**, et sur **toute la période** analysée.

**Périmètre & niveau.** Le calcul s'effectue au niveau du **domaine**, avec une déclinaison **France** et **Monde**.

**Sortie & lecture.** Le résultat est exprimé en **pourcentage** de la production du domaine attribuable à ANR dans le cadre de France 2030 : une valeur élevée traduit un **poids scientifique** plus important. À titre de repère, la part de l'ANR **dans le cadre de France 2030** dans les dépenses de la **MIRES** étant d'environ **5,5%**, un résultat **supérieur à ce seuil** signale une contribution scientifique **au-delà**

de la **part budgétaire**, autrement dit, une **surperformance relative** du domaine (ou du projet) considéré. Les projets ayant une empreinte **supérieure** ont ainsi pu être identifiés.

**Limites / précautions.** Les référentiels de publications (monde et France) sont construits à partir de l'**analyse sémantique** des publications liées à l'ANR dans le cadre de France 2030 : ce sont les publications présentant une **proximité sémantique forte** avec les documents des projets ANR dans le cadre de France 2030 qui sont retenues. Il existe donc un **biais de sélection**, susceptible de **survaloriser légèrement** la part des projets ANR dans le cadre de France 2030 dans l'ensemble des publications françaises ou mondiales.

De manière complémentaire à l'ensemble des indicateurs précédents, une analyse a été conduite sur le périmètre spécifique des projets ayant mobilisé des **dépenses en équipements**. Cette analyse vise à mesurer plus finement l'empreinte de ce type de projet, en particulier sur la structuration des connaissances scientifiques dans les domaines concernés.

#### Empreinte équipement (par domaine)

**Objet.** Cet indicateur vise à évaluer l'effet des projets d'équipement sur la structuration et la visibilité scientifique d'un domaine, en mesurant les dynamiques associées à la mise en service de ces équipements.

**Méthode.** L'indicateur repose sur une analyse différentielle, comparant la **centralité** du domaine et le **volume de publications** avant et après la mise en œuvre du projet d'équipement. Il s'agit d'apprécier dans quelle mesure l'équipement contribue à une montée en influence du domaine dans les réseaux de connaissances (publications, citations), et à une intensification de la production scientifique. Le calcul s'appuie sur des fenêtres temporelles glissantes et des règles de contrôle de la taille des corpus pour assurer la robustesse de la comparaison.

**Périmètre & niveau.** Le calcul s'effectue au niveau du domaine, en sélectionnant les projets ANR dans le cadre de France 2030 explicitement associés à des dépenses d'équipement significatives, telles que recensées dans les rapports d'exécution.

**Sortie & lecture.** Le résultat est exprimé sous forme de variation relative :

- une **augmentation de centralité** post-équipement traduit un gain d'influence scientifique ;
- une **augmentation de la production** traduit une capacité de mobilisation autour de l'équipement ;
- La combinaison **des deux** indique un effet structurant fort du projet.

**Limites / précautions.** L'interprétation de l'effet équipement nécessite de tenir compte du délai entre l'investissement, l'installation de l'équipement et la montée en charge effective des activités scientifiques associées. Afin de garantir la comparabilité des résultats, quelle que soit la profondeur temporelle disponible, les valeurs sont exprimées en moyenne annuelle. En outre, pour intégrer le temps nécessaire à la production et à la publication des résultats, la date pivot séparant la période « avant » et la période « après » équipement est fixée à deux années après la date de démarrage du projet.

## 2. Présentation de l'application

L'application interactive développée en appui au présent rapport permet à l'ANR dans le cadre de France 2030 d'explorer l'ensemble des résultats de manière fluide et comparative.

### Statistiques descriptives



Sélecteur de page  
 ● Statistiques descriptives | ● Analyse par domaine

#### Analyses globales

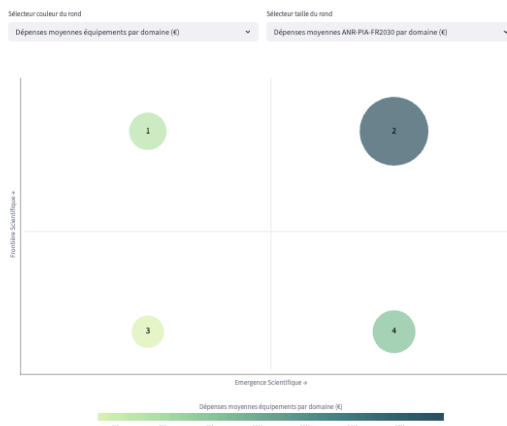
##### Périmètre

Projets : 1275 Actions : 83 Domaines : 247 Topics : 3800 Articles : 131315 Brevets : 724

Catégorie  
 PIA

Catégorie	Articles	Brevets	Dépenses ANR-PIA-FR2030 (K)	Dépenses Équipements (K)	Projets	Projets avec équipements	Partenaires	Basis de données	Logiciels
1	113,965	696	5,183,955,523	1,252,962,743	376	132	3,998	35	82
2	4,207	0	577,896,873	183,000,000	54	5	531	0	0
3	6,992	0	1,671,301,403	620,103,311	194	54	0	0	0
4	1,758	24	2,436,883,624	0	484	0	3,152	378	324
FR30	3,053	4	1,815,211,739	146,998,175	159	21	907	81	89

##### Les principaux indicateurs



##### Indicateurs par cadran

	1	2	3	4	Total*
Centralité moyenne par domaine	0.082	0.095	0.066	0.096	0.044
Cohésion moyenne par domaine	0.1275	0.1439	0.1210	0.1866	0.2200
Dépenses ANR-PIA-FR2030 (K)	41,472,258,857	48,187,455,787	41,355,306,040	41,939,228,230	150,955,258,913
Dépenses moyennes ANR-PIA-FR2030 par domaine (K)	436,762,343	499,837,263	426,593,683	435,096,004	444,953,117
Dépenses moyennes Équipements par domaine (K)	45,808,940	417,952,327	43,288,715	48,517,311	48,910,295
Dépenses équipements (K)	4276,951,709	41,292,155,122	4221,632,640	4461,452,128	42,209,832,089
Empreinte scientifique moyen de l'ANR-PIA-FR2030 par domaine (%)	7.43	7.08	6.31	6.85	7.05
Empreinte technologique moyen de l'ANR-PIA-FR2030 par domaine (%)	1.82	3.36	1.28	2.28	2.09
Nombre d'actions uniques	67	89	62	69	83
Nombre d'articles uniques	30522	98941	18958	31167	109108
Nombre de brevets uniques	111	686	37	159	724
Nombre de domaines	55	69	68	55	247
Nombre de domaines avec une empreinte technologique > 5.5%	4	12	5	10	31
Nombre de domaines avec une empreinte scientifique > 5.5%	39	57	43	37	176
Nombre de projets	799	1365	666	826	1275
Nombre de projets PEPH uniques	186	444	134	234	529
Nombre de projets Équipements uniques	183	208	156	173	312
Nombre de régions uniques	14	16	15	13	17
Nombre moyen d'actions par domaine	11.50	23.78	19.87	14.35	14.42
Nombre moyen d'articles par domaine	353.62	1,432.48	143.28	387.75	604.69
Nombre moyen de brevets par domaine	2.02	15.26	0.57	3.16	3.66
Nombre moyen de projets Équipements par domaine	18.40	42.20	12.33	21.36	23.83
Nombre moyen de projets PEPH par domaine	4.48	16.29	2.29	6.14	7.51
Nombre moyen de projets par domaine	75.67	181.79	98.88	93.18	104.02
Nombre moyen de régions par domaine	8.78	13.97	8.87	19.38	9.83

\* Rappel : les projets étant distribués au sein des domaines en fonction de leurs résultats (articles, brevets), ils peuvent contribuer à plusieurs cadrans.

##### Valeurs des indicateurs

###### Indicateurs par domaine

Sélectionnez l'indicateur  
 Centralité

Titre	Centralité
Engineering - Electrical and Electronic Engineering (2595)	0.4039
Physics and Astronomy - Astronomy and Astrophysics (1348)	0.3687
Physics and Astronomy - Nuclear and High Energy Physics (3395)	0.3438
Biochemistry, Genetics and Molecular Biology - Molecular Biology (1312)	0.3394
Physics and Astronomy - Atomic and Molecular Physics, and Optics (3107)	0.3320
Materials Science - Materials Chemistry (5555)	0.3133
Computer Science - Artificial Intelligence (1792)	0.2523
Mathematics - Applied Mathematics (2684)	0.2213
Mathematics - Geometry and Topology (2096)	0.2189
Medicine - Epidemiology (2712)	0.2133

###### Indicateurs par projet

Sélectionnez l'indicateur  
 Dépenses ANR-PIA-FR2030 (K)

IDPP	Dépenses ANR-PIA-FR2030 (K)
22-NEXG-0001	408,000,000
19-ART-0005	228,220,844
19-ART-0001	180,137,793
19-ART-0002	166,842,875
19-ART-0003	130,696,540
19-ART-0008	126,584,563
19-ART-0007	121,730,594
16-EGPH-0001	107,200,000
22-BIOC-0001	100,000,000
23-BIOC-0001	99,141,420

La première page, intitulée "Statistiques descriptives", offre une vue d'ensemble des projets, actions, domaines et résultats associés. Elle est structurée en quatre sections principales.

### Périmètre

Cette section permet de **restreindre dynamiquement l'analyse à un sous-ensemble de projets** selon les critères suivants :

- Type d'action ANR
- Présence d'équipements (avec ou sans)
- État du projet (actif ou inactif)
- PIA concerné (PIA 1, 2, 3, 4 ou France 2030)
- Type de projet (recherche ou valorisation)
- Région administrative

Une fois ces filtres appliqués, un **tableau de synthèse** s'affiche, présentant :

- Articles
- Brevets
- Dépenses ANR dans le cadre de France 2030 (€)
- Dépenses équipements (€)
- Projets (nombre)
- Projets avec équipements (nombre)
- Partenaires (nombre)
- Bases de données (nombre)
- Logiciels (nombre)

Cela permet à l'utilisateur de valider le périmètre d'analyse avant d'explorer les autres sections.

### Les Principaux indicateurs

Une visualisation synthétique permet de **cartographier les domaines** dans une matrice **d'analyse stratégique** en croisant deux indicateurs structurants :

- Proximité à la frontière scientifique (cohérence)
- Dynamique d'émergence (centralité)

Chaque domaine est représenté par un **rond positionné dans un des quatre cadrans**.

Deux **sélecteurs** permettent à l'utilisateur de choisir quels indicateurs utiliser pour :

- **La taille** du rond (ex. : dépenses, nombre de projets...)
- **La couleur** du rond (ex. : empreinte scientifique ou technologique, ratio publications/brevets ...)

Cette flexibilité permet d'identifier visuellement les domaines les plus stratégiques ou les plus sous-performants selon les critères choisis.

De plus, l'utilisateur peut visualiser pour l'ensemble des cadrans plus de 25 indicateurs disponibles pour visualiser la structure et les effets des financements selon le périmètre retenu. Ces indicateurs couvrent notamment :

- La **production scientifique et technologique** (nombre d'articles, brevets, projets...)
- L'**empreinte** moyen de l'ANR dans le cadre de France 2030 par domaine (%)
- Les **financements** (dépenses ANR dans le cadre de France 2030 totales ou par équipement)
- Les **dynamiques structurelles** (centralité, cohérence, nombre de domaines)
- Les **ratios moyens par domaine**

Les valeurs sont présentées sous forme de graphiques et de chiffres clés afin d'identifier les grands ordres de grandeur.

## Valeurs des indicateurs

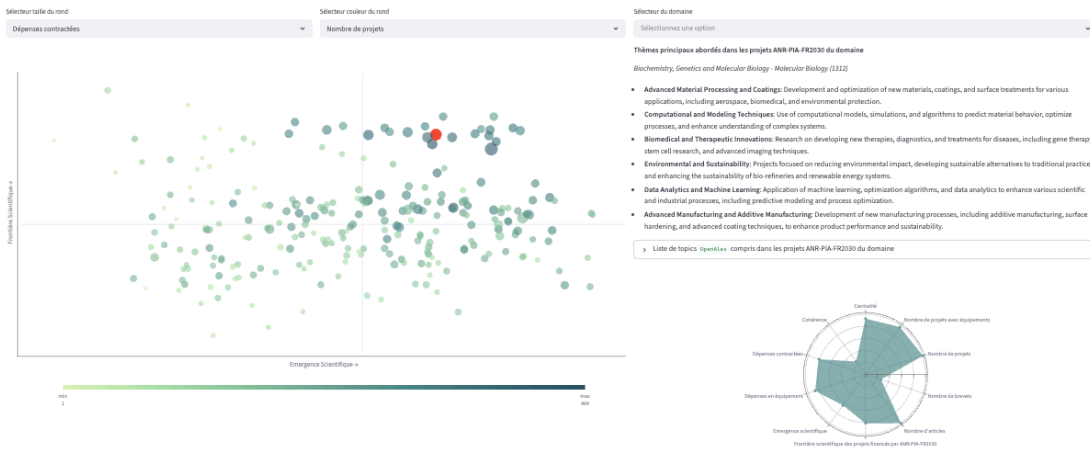
En complément de la vue agrégée, cette section permet de visualiser le classement de tous **les domaines et projets** selon tous les indicateurs pertinents au niveau des domaines ou des projets.

# Analyse par domaine

Sélecteur de page

Statistiques descriptives | Analyse par domaine

## Analyse par domaine financés par ANR-PIA-FR2030



## Empreinte Scientifique

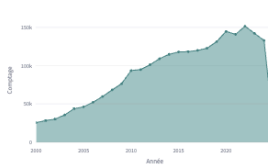
Biochemistry, Genetics and Molecular Biology - Molecular Biology (1312)

Empreinte ANR-PIA-FR2030 14,70%

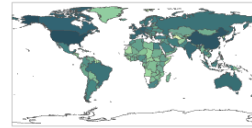
Empreinte ANR-PIA-FR2030



Empreinte dans le temps



Carte du nombre d'articles par pays



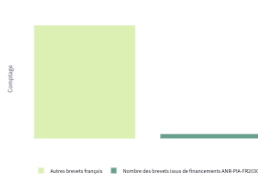
## Valorisation technologique

Biochemistry, Genetics and Molecular Biology - Molecular Biology (1312)

Empreinte ANR-PIA-FR2030 3,86%

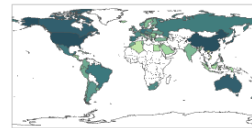
Capacité d'absorption du domaine 5,43%

Empreinte ANR-PIA-FR2030 brevets



Entreprise	Classement	Pays
INFINEON TECHNOLOGIES	1	DEU
ROBERT BOSCH	2	DEU
XTRON	3	BEL
LFB	4	FIN
NOVZEMES	5	DNK
DSM	6	NLD
TE CONNECTIVITY	7	CHE
VALEO	8	FIN
CONTINENTAL	9	DEU
ZF	10	DEU

Carte du nombre de brevets par pays

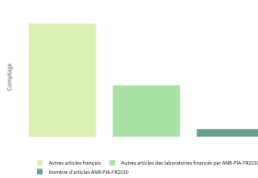


## Equipment

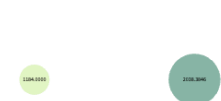
Biochemistry, Genetics and Molecular Biology - Molecular Biology (1312)

Nombre de projets avec Équipements 100

Empreinte ANR-PIA-FR2030 12,93%



Empreinte dans le temps 72,16%



Dépenses par action



Sélecteur du projet

10-COHO-0002

Descriptif des équipements:

Le logiciel EDMUS est un outil de gestion de données pour le suivi de la schémose en plaques. Comparé aux équipements précédents, le logiciel EDMUS offre une couverture et des fonctionnalités améliorées pour répondre aux exigences des neurologues français. Grâce au logiciel, un suivi longitudinal généralisé et une association systématique des données cliniques avec des prélèvements biologiques et la neuro-imagerie sont désormais possibles, améliorant la qualité et l'efficacité des soins.

## Projets

Biochemistry, Genetics and Molecular Biology - Molecular Biology (1312)

Nombre de projets 454

idno_projet	Date de début	Date de fin	Résumé	Région	Nombre d'articles	PIA	Dépenses ANR-PIA-FR2030	Action	Etat	PEPR
10-LABR-0008	2011-04-13	2020-12-31	The LabEx CORM, focuses on coral reefs, emphasizing socio-economic governance, innovation	Ile-de-France	1093	1	6,164,811	LABX	Closed	<input type="checkbox"/>
11-INRS-0013	2012-01-01	2020-12-31	The French Bioinformatics Institute (IFI) is a national infrastructure aimed at providing bioinformatics	Ile-de-France	1183	1	22,800,000	INRS	Actif	<input type="checkbox"/>
11-LABR-0029	2011-01-03	2020-12-31	SynOrg LABEx, a French cluster in molecular chemistry, aims to address global health challenges	Normandie	1107	1	11,405,571	LABX	Actif	<input type="checkbox"/>
10-LABR-0002	2011-04-13	2020-12-31	The document describes the LabEx 'Integrative Biology of Emerging Infectious Diseases (IBED)	Ile-de-France	2086	1	44,594,395	LABX	Actif	<input type="checkbox"/>
10-LABR-0033	2011-04-13	2020-12-31	ERMITE is a multidisciplinary laboratory focused on combating cancer, cardiovascular disease	Ile-de-France	1437	1	19,300,000	LABX	Closed	<input type="checkbox"/>
10-LABR-0054	2011-04-13	2020-12-31	The LabEx MemeLife project aims to achieve its objectives through effective governance at the	Ile-de-France	841	1	10,409,848	LABX	Closed	<input type="checkbox"/>
10-LABR-0036	2011-04-13	2019-12-31	The document outlines the selection and funding of 11 LabEx Laboratories of Excellence (LoE)	Grand Est	247	1	9,300,000	LABX	Closed	<input type="checkbox"/>
10-INRS-0005	2011-01-05	2024-12-31	FRSBI (French Infrastructure for Integrated Structural Biology) is a national distributed infra	Grand Est	1814	1	35,800,000	INRS	Closed	<input type="checkbox"/>
10-INRS-0009	2011-01-12	2020-12-31	This document outlines the proposal for the 'France Genomics' infrastructure, aimed at inf	Ile-de-France	1426	1	45,387,574	INRS	Actif	<input type="checkbox"/>
11-LABR-0071	2012-01-03	2022-03-31	The Exploring Identity Initiative aims to foster interdisciplinary research and training to advan	Ile-de-France	474	1	13,177,862	LABX	Closed	<input type="checkbox"/>

La seconde page de l'application interactive permet d'explorer en détail les données agrégées et comparatives à l'échelle de chaque domaine scientifique et technologique couvert par les projets financés par l'ANR. Elle est structurée en cinq sections principales permettant une lecture à la fois quantitative, qualitative et temporelle des résultats par domaine :

### Sélecteurs d'exploration comparative

En haut de page, l'utilisateur peut :

- Sélectionner un **domaine scientifique** via un menu déroulant ou en cliquant directement sur un point de la représentation par cadran
- Personnaliser la **taille des ronds** (ex. : dépenses contractées, nombre de projets).
- Choisir la **couleur des ronds** (ex. : nombre d'articles, nombre de brevets, etc.).

Le domaine sélectionné est positionné automatiquement dans la **matrice stratégique des cadrans** (cadran 1, 2, 3 et 4).

Le reste de la page est automatiquement mis à jour en fonction du domaine sélectionné dans la matrice stratégique.

En complément de la matrice stratégique, un **graphe radar** situé en haut de la page permet de **visualiser simultanément plusieurs indicateurs clés** pour le domaine sélectionné. Chaque axe représente un indicateur distinct (Centralité, Nombre de projets avec équipements, Nombre de projets, Nombre de brevets, Nombre d'articles, Frontière scientifique des projets financés par l'ANR, Émergence scientifique, Dépenses en équipement, Dépenses contractées, Cohérence)

Cette représentation permet de comparer d'un coup d'œil le **profil global du domaine**, en identifiant ses forces relatives (valeurs proches du bord extérieur) et ses marges de progression (valeurs proches du centre). Elle offre une lecture synthétique du positionnement du domaine au regard des principaux critères d'analyse retenus dans l'étude.

Enfin, cette première partie contient pour chaque domaine scientifique sélectionné, une description avec :

- Une **description obtenue par IA** du contenu et des principaux thèmes traités dans les projets financés par l'ANR.
- Une liste déroulante permet d'explorer les **topics précis** identifiés dans les projets ANR dans le cadre de France 2030 du domaine. Ces topics sont directement issus de la taxonomie OpenAlex<sup>4</sup>.

### Empreinte scientifique

Cette section comporte :

- Une représentation de "**empreinte de l'ANR dans le cadre de France 2030**" sur la production scientifique avec trois **niveaux de comptage** : nombre total d'articles scientifiques français, nombre d'articles des laboratoires financés par l'ANR, et nombre d'articles directement issus des projets ANR.
- Un **graphique de tendance temporelle** montrant l'évolution d'un indicateur de 2000 à 2025, du nombre d'articles scientifiques.

---

<sup>4</sup>[https://docs.google.com/spreadsheets/d/1v-MAq64x4YjhO7RWcB-yrKV5D\\_2vOOsxl4u6GBKEXY8/edit?gid=983250122#gid=983250122](https://docs.google.com/spreadsheets/d/1v-MAq64x4YjhO7RWcB-yrKV5D_2vOOsxl4u6GBKEXY8/edit?gid=983250122#gid=983250122)

- Une carte de la production scientifique dans le monde avec le nombre d'articles scientifiques du domaine dont la similarité sémantique avec les publications issues des projets ANR dans le cadre de France 2030 est supérieure au seuil défini par GraphMyTech.

### Valorisation technologique

Cette section présente :

- Une représentation de "**empreinte de l'ANR dans le cadre de France 2030**" sur la production de brevets avec deux **niveaux de comptage** : nombre total de brevets français, nombre de brevets issus des financements ANR.
- Un indicateur spécifique : la **capacité d'absorption du domaine**, qui traduit le pourcentage d'entreprises capables d'absorber la technologie et le classement des entreprises en fonction de cette capacité d'absorption.
- Une carte de la production technologique dans le monde avec le nombre de brevets du domaine dont la similarité sémantique avec les publications issues des projets ANR dans le cadre de France 2030 est supérieure au seuil défini par GraphMyTech.

### Équipements

- Une visualisation de l'**empreinte scientifique des projets avec équipements**, selon les mêmes indicateurs que l'empreinte scientifique mais limitée aux projets d'équipement.
- Un **graphique comparatif "Avant / Après"** indiquant l'évolution du nombre moyen d'articles publiés par an.
- Une représentation des **dépenses d'équipements** selon l'action ANR dans le cadre de France 2030 (EQPX, INBS, COHO, etc.).
- Un accès au **descriptif du projet** sélectionné (via un sélecteur), avec résumé des objectifs et spécificités de l'équipement.

### Projets

Enfin, un rappel du **nombre total de projets** dans le domaine sélectionné est affiché et la liste des projets avec les principales informations à cette échelle : projet EOTP; Date de début ; Date de fin ; Résumé ; Région ; Nombre d'articles ; PIA ; Dépenses de l'ANR dans le cadre de France 2030 ; Action ; État ; PEPR.

## Références bibliographiques

- Borgatti, S.P., Mehra, A., Brass, D.J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, 323(5916), 892–895.
- Dosi, G. (1982). Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. *Research Policy*, 11(3), 147–162.
- Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science*, 47(1), 117–132.
- Freeman, C., & Perez, C. (1988). Structural crises of adjustment: Business cycles and investment behaviour. In G. Dosi, C. Freeman, R. Nelson, G. Silverberg, & L. Soete (Eds.), *Technical Change and Economic Theory* (pp. 38–66). Pinter.
- Freeman, L.C. (1979). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215–239.
- Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, 41(5), 685–697.
- Kuhn, T.S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.
- Nelson, R.R., & Winter, S.G. (1982). *An Evolutionary Theory of Economic Change*. Harvard University Press.
- Nesta, L. Saviotti, P-P. (2006). Firm knowledge and market value in biotechnology, *Industrial and Corporate Change*, Volume 15, Issue 4, August, 625–652
- Singh, A., Triulzi, G., & Magee, C.L. (2020). Technological improvement rate estimates for all technologies: Use of patent data and an extended domain description. *SSRN Electronic Journal*.
- Teece, D.J. (1982). Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior & Organization*, 3(1), 39–63.
- Teece, D.J., Rumelt, R., Dosi, G., & Winter, S. (1994). Understanding corporate coherence: Theory and evidence. *Journal of Economic Behavior & Organization*, 23(1), 1–30.
- Triulzi, G., Alstott, J., & Magee, C.L. (2017). Estimating technology performance improvement rates by mining patent data. *SSRN Electronic Journal*.
- Weitzman, M.L. (1998). Recombinant growth. *Quarterly Journal of Economics*, 113(2), 331–360.

