



UNIVERSITÉ SORBONNE PARIS NORD  
UFR SCIENCES ECONOMIQUES ET GESTION

Mémoire de fin d'année

Sujet : Analyse de l'optimisation des décisions d'investissements en R&D à l'ère de l'IA aux Etats-Unis

Réaliser par :

PHILIPPE ROUMBO, 12300300, étudiant en M1 parcours BIDABI

AMGAD ELABASY, 12101673, étudiant en M1 parcours Conseil en entreprise

Encadré par :

Mme RÉBECCA SIROIS

## Remerciements :

Au terme de la rédaction de ce mémoire, nous souhaitons exprimer notre profonde gratitude envers toutes les personnes qui, par leur disponibilité, leurs compétences ou leur soutien moral, ont rendu ce travail possible.

Nous tenons dans un premier temps à exprimer notre gratitude à notre directrice de mémoire, Mme Rébecca Sirois, Professeure d'étude de marché à l'Université Sorbonne Paris Nord, pour la qualité de son encadrement, attentif et bienveillant. Grâce à ses conseils méthodiques, à ses remarques judicieuses et à son exigence scientifique, nous avons été en mesure de construire avec elle un fil directeur qui nous a permis de bâtir des analyses tout en nous donnant la possibilité d'exercer une démarche critique autonome.

Nous remercions également l'ensemble de l'équipe pédagogique de l'Université Sorbonne Paris Nord et les nombreux intervenants professionnels chargés de notre formation qui nous ont conseillés dans les différentes parties de notre mémoire.

Nous souhaitons également depuis lors faire connaître notre reconnaissance au personnel administratif de l'Université Sorbonne Paris Nord, ainsi qu'aux bibliothécaires, pour leur disponibilité et leur efficacité, qui ont facilité notre accès aux bases de données spécialisées et à des ouvrages rarement consultés.

Nous n'oublions pas nos camarades de promotion qui ont été une source de motivation et d'enrichissement intellectuel lors des échanges parfois agités. Leurs relectures et leurs critiques n'ont pas manqué d'aviver la lumière pour certains de nos passages obscurs.

Enfin, nous tenons aussi à remercier notre famille et nos proches. Leur confiance, leurs encouragements quotidiens, leur patience dans les périodes de concentration intense ont été les conditions qui ont permis d'instaurer un climat souhaitable pour mener ce travail académique jusqu'à son terme.

À chacun d'entre vous, nous vous en sommes reconnaissant.

## Table des matières

Mémoire de fin d'année.....	1
Partie 1 – Introduction.....	6
1.1    Contexte Générale.....	6
1.2    Problématique .....	6
1.3    Enigme mise en lumière dans ce travail de recherche.....	6
1.4    Annonce du plan du mémoire.....	7
1.5    Objectifs .....	7
1)    Quantifier l'effet de l'IA sur le rendement économique et technologique des projets de R&D .....	7
2)    Identifier les secteurs et les types d'entreprises qui ont le plus d'avantage à utiliser l'IA.....	7
3)    Evaluer le niveau de concentration en termes de brevet d'IA aux États-Unis....	7
1.6    Méthodologie .....	7
Partie 2 – Fondements conceptuels et état de l'art.....	8
2.1 Notions clés : Investissement en R&D et financement de l'innovation .....	8
2.1.1 Qu'entendons-nous par Investissements en R&D ?.....	8
2.1.2 Pourquoi optimiser est crucial ? .....	9
2.2 L'IA : Un outil innovant dans l'aide à la décision pour optimiser les investissements ? .....	13
2.2.1 Notions clées : IA, Machine Learning .....	13
2.2.2 Modèles et pratiques existant dans la gestion de portefeuille R&D par IA.....	17
2.2.3 Limites et défis théoriques .....	19
2.3 Politique et écosystème d'IA aux Etats-Unis .....	20
2.3.1 Écosystème d'innovation basé sur la synergie publique-privé.....	20
2.3.2 Stratégies nationales et initiatives fédérales .....	21
2.3.3 Enjeux éthiques et de souveraineté technologique .....	24
Partie 3 – Méthodologie.....	25
3.1 La collecte des données .....	25
3.1.1 Source de données utilisées pour l'analyse quantitative .....	25
3.1.2 Méthodologie de collecte des sources .....	27
3.2 Les variables : variables en lien avec l'Investissement en R&D et l'IA aux Etats-Unis. ....	28

3.2.1 Les variables quantitatives.....	29
3.2.2 Les variables qualitatives .....	31
3.3 Méthodologie d'analyse des variables quantitatives et qualitatives .....	32
3.2.3 Pourquoi avons-nous choisi ces variables ? .....	32
3.3.1 Approche quantitative .....	32
3.3.2 Approche qualitative .....	34
3.3.3 Limites et défis de la méthodologie mixte .....	34
Partie 4 – Résultats & Discussions .....	37
4.1.1Analyses descriptives des données .....	37
4.1.2 Analyse des modèles économétriques .....	55
4.1.3 Gradient du modèle de cycle de time to Market .....	62
4.1.4 Analyse de l'étude de cas.....	63
4.2 Discussions .....	66
4.2.1 Quels secteurs ou types d'entreprises américaines intègrent l'IA dans leur processus de R&D, et comment ? .....	66
4.2.2 Quels sont les impacts mesurables de l'IA sur la performance des investissements en R&D dans le contexte américain ? .....	67
4.2.3 Les recherches associées à l'IA en R&D permet-elle une concentration des brevets qui serait plus privilégié pour les grandes voir très grandes firmes ou, au contraire, serait-elle plus à même d'être généraliser ?.....	67
Partie 5 – Proposition de sujet à traiter .....	68
5.1 L'optimisation des décisions d'investissements en R&D grâce à l'IA : une analyse comparative entre les Etats- Unis, la Chine et l'Europe .....	68
5.2 L'optimisation des décisions d'investissements en R&D grâce à L'IA : Analyse des répercussion politique entre la Chine et Taiwan .....	69
5.3 Quel serait l'impact de l'intelligence artificielle générale (IAG) sur l'environnement politique et écologiques .....	69
Partie 6 –Conclusion.....	70
6. 1 Résumer des résultats obtenus .....	70
6.2 Les apports de notre mémoire .....	71
6.3 Les limites de notre mémoire .....	71
Bibliographie.....	73
Annexes.....	81

Annexe 1 : Information primaire avant nettoyage de la base de données .....	81
Annexe 2 : Information sur le pourcentage de valeurs manquantes de la base de données .....	82
Annexe 3 : Imputation des valeurs manquantes par la méthode des k-NN.....	82
Annexe 4 : Information sur le pourcentage de valeurs manquante après imputation des K-NN.....	84
Annexe 5 : Analyse descriptive des variables de R&D (1ere partie) .....	84
Annexe 6 : Analyse descriptive des variables de R&D (2 <sup>e</sup> partie) .....	85
Annexe 7 : Analyse descriptive des variables de R&D (3 <sup>e</sup> partie) .....	85
Annexe 8: ACF et PACF du model_DIRD .....	86
Annexe 9: Graphique de convergence de la descente de gradient du modèle DIRD .	86
Annexe 10: Comparaison avec les résultats du modèle OLS.....	87
Annexe 11: Courbe de Lorenz de la demande de brevet lié à l'IA .....	88
Annexe 12: Courbe de Lorenz de l'offre de brevet lié à l'IA .....	89
Annexe 13: Représentation graphique de l'évolution de l'investissement en capital risque de 2010-2024 .....	90
Annexe 14 : Code en Langage Python :.....	90
Annexe 15: Lien du Questionnaire en version française.....	90
Index .....	91
<b>TABLEAU DES ILLUSTRATIONS .....</b>	<b>94</b>
<b>IMAGES.....</b>	<b>94</b>
<b>TABLEAUX .....</b>	<b>94</b>

# Partie 1 – Introduction

## 1.1 Contexte Générale

« *Celui qui sera leader dans l'IA, sera le maître du monde* » (Poutine, 2017)

En effet, le discours de Vladimir Poutine résume les enjeux de notre époque. Malgré l'importance croissante des investissements en R&D dans l'économie américaine, leur rendement varie fortement selon les secteurs et les entreprises. L'apparition de l'intelligence artificielle (IA) apparaît comme un levier potentiel pour optimiser ces investissements, mais ses mécanismes d'action et son impact effectif restent encore à clarifier.

## 1.2 Problématique

Cela nous amène donc à nous demander : dans quelles mesures et par quels mécanismes l'intelligence artificielle permet-elle d'optimiser les investissements en R&D aux États-Unis ?

### 1.3 Enigme mise en lumière dans ce travail de recherche

Quels secteurs ou types d'entreprises américaines intègrent l'IA dans leur processus de R&D, et comment ?

Quels sont les impacts mesurables de l'IA sur la performance des investissements en R&D dans le contexte américain ?

Les recherches associées à l'IA en R&D permet-elle une concentration des brevets qui serait plus privilégié pour les grandes voir très grandes firmes ou, au contraire, serait-elle plus à même d'être généraliser ?

#### **1.4 Annonce du plan du mémoire**

Afin de pouvoir apporter les réponses les plus pertinentes possible, nous vous proposons dans un premier temps, avec la Partie 2, d'établir les fondements conceptuels et l'état de l'art : nous y précisons ce que recouvrent l'investissement en R-D et l'IA, tout en décrivant l'écosystème américain qui en résulte.

Ensuite, la Partie 3 exposera notre méthodologie mixte dans lequel y seront détaillés la collecte des données, le choix des variables, ainsi que les protocoles économétriques et d'étude de cas mobilisés pour assurer la robustesse de nos analyses.

Par ailleurs, la Partie 4 présentera les résultats et les discussions. De fait, il y aura des analyses descriptives, modèles économétriques et mise en perspective des effets mesurés de l'IA sur la performance des investissements en R&D.

En outre, la Partie 5 ouvrira le champ des perspectives en proposant trois pistes de recherche futures ; et pour finir, la Partie 6 ne conclue en récapitulant nos apports, en pointant nos limites et en formulant des recommandations à l'intention des chercheurs et des décideurs

#### **1.5 Objectifs**

Toujours dans l'optique de pouvoir répondre au mieux à notre problématique, nous nous sommes fixé les objectifs suivants :

- 1) Quantifier l'effet de l'IA sur le rendement économique et technologique des projets de R&D
- 2) Identifier les secteurs et les types d'entreprises qui ont le plus d'avantage à utiliser l'IA
- 3) Evaluer le niveau de concentration en termes de brevet d'IA aux États-Unis.

#### **1.6 Méthodologie**

Notre mémoire sera réalisé de deux manières différentes. On compte réaliser dans un premier temps une approche quantitative afin de pouvoir faire des analyses statistiques et réaliser des modèles économétriques qui soient solide. Puis d'un autre côté , nous réaliserons une approche qui sera plus qualitative en nous appuyant sur l' étude de cas de Pfizer sans oublier un questionnaire adressé aux praticiens de la R&D et du capital-risque.

Pour le dire autrement, ces deux approches permettront de confronter les mesures statistiques à la réalité de terrain.

## Partie 2 – Fondements conceptuels et état de l’art

### 2.1 Notions clés : Investissement en R&D et financement de l’innovation

Quel que soit le marché, l’investissement en R&D joue un véritable levier stratégique. Dans cette partie nous verrons ce que nous entendons par l’investissement dans la recherche et le développement (R&D), puis nous continuerons sur les raisons pour lequel optimiser est crucial pour ensuite finir sur une présentation des politiques et de l’ écosystème d’IA aux Etats-Unis.

#### 2.1.1 Qu’entendons-nous par Investissements en R&D ?

##### ***Investissement en R&D : définition***

Selon le ministère de l’économie, des finances et de la souveraineté industrielle et numérique<sup>1</sup>, la fonction de recherche et développement représente l’ensemble des processus qui commence de la recherche fondamentale ou d’une invention à une mise en œuvre industrielle.

Autrement dit, il s’agit donc de l’ensemble des étapes permettant de passer du laboratoire de recherche avec le fait d’avoir une idée, à la production industrielle d’un produit concret en usine ou d’un nouveau service dans lequel les avantages concurrentiels obtenue par les agents économiques vont assurer, par extension, le bon développement tant dans les secteurs publics que privé (La fonction Recherche & Développement, s.d.).

##### ***Secteur public VS Secteur privée : définition***

En effet, le secteur public est représenté par l’ensemble de toutes les administrations et organisations qui sont contrôlées ou financées par l’État dont sa mission principale est de réaliser des actions d’intérêt général (INSEE, 2021).

---

<sup>1</sup> [La fonction Recherche & Développement | ministère de l’Économie des Finances et de la Souveraineté industrielle et numérique](#)

Contrairement à ce dernier, le secteur privé rassemble plus les entreprises, les particuliers qui possèdent leur propre capital et les associations qui mettent principalement en place des actions pour générer des bénéfices ou pour réaliser des missions qu'ils se sont eux même fixés tout en étant encadré par la loi et les autorités de régulation (OCDE).

### ***Les Brevets***

Qui plus est, qui dit investissement en recherche et développement, dit aussi dépôt de brevet.

Selon la définition officielle de l'INSEE, le brevet permet de « protéger un produit ou un procédé qui répond à un problème à la condition qu'elle soit conforme aux critères de brevetabilité et de ne pas être expressément exclues de la protection par la loi.

Qui plus est, il peut arriver que certaines inventions ne soient pas brevetables, mais peuvent cependant être conformes à d'autres types de protection, comme le droit d'auteur ou le dépôt de dessins et modèles. » (INSEE, 2021)

#### **2.1.2 Pourquoi optimiser est crucial ?**

##### ***Efficience dans la gestion des ressources***

En effet, lorsqu'il y a investissement dans la recherche et le développement, ça permet d'être plus efficace dans la gestion des ressources car en l'absence d'un bon encadrement, les efforts de recherche peuvent se creuser et se superposer inutilement entraînant donc par la suite une perte directe de temps et d'argent tandis que lorsqu'il est bien géré, c'est un véritable cercle vertueux. (ResearchGate)

Avec de bons outils d'allocation, l'impact technologique de la recherche et développement est aussi maximisé en investissant dans les projets à rentabilité potentielle la plus forte sur le plan scientifique ou commercial.

De fait, de nombreuses études montrent que les rendements de l'innovation sont très déséquilibrés et quelques « super-projets » représentent l'essentiel des retombées (NBER, Septembre 2009).

Par exemple, dans l'industrie pharmaceutique, un dixième des nouveaux médicaments les plus rentables a rapporté plus de la moitié des revenues de leurs catégories, alors que le médicament moyen s'avère très largement moins fructueux ; d'où la nécessité de cibler et de propulser les projets les plus productifs au risque de sacrifier un retour sur investissement en recherche et développement, et par extension le progrès technique et la croissance de l'économie. (Grabowski & J, 2000)

### ***Théorie de Knight***

Traditionnellement, les entreprises ou les instituts de projets de R&D s'appuyaient sur les compétences humaines pour sélectionner et décider les projets qu'ils vont entreprendre ou qu'ils pensent que ce sont des projets prometteurs. Cependant, cette approche n'a pas toujours été si efficace.

Dans le cadre d'un rapport scientifique, intitulé « Prediction of SME<sup>2</sup>s'R&D performances by machine Learning for Project selection » et édité en Juin 2023, les auteurs démontrent, par des observations et des analyses, que l'évaluation de probabilité de succès d'un projet de R&D ex ante par des experts du domaine pose des limites et cette approche a des raisons divers et variées :

Parmi eux, selon la théorie de l'incertitude de Frank H. Knight, à travers de son livre « Risk, Uncertainty, and profit » en 1921, deux principes fondamentaux sont mis en lumière pour mieux comprendre sous quel travers un agent économique investi. (Knight, 2023)

D'un côté, nous avons le risque qui est définie par une situation dans lequel l'agent ne connaît pas l'avenir mais il est certain qu'il existe pour chacun de ces événements une probabilité qui leur est associés et un scénario totalement ou partiellement différent. (Knight, 2023)

Par ailleurs, d'un autre côté, nous avons l'incertitude qui se réfère plus à des événements qui sont imprévisibles, car ils sont uniques ou parce qu'ils n'ont pas de précédent historique.

Pour Knight, c'est justement cette incertitude qu'il qualifie d'incertitude radicale, car cela justifie l'existence d'un super profit pour l'entrepreneur.

À cela s'ajoute le fait qu'un projet de R&D implique divers aspects qui ont besoin chacun d'un expert dans le domaine. (Knight, 2023)

---

<sup>2</sup> SME : Small and Medium Enterprises (en français, Entreprises de petite et moyenne taille)

## *Choix du projet*

En effet, le processus d'un projet de R&D n'est pas structuré, ce qui, en d'autres termes, signifie que la manière dont le projet est choisi n'est pas formalisé ou systématique.

De fait, les critères de sélection sont divers (rentabilité, faisabilité, etc.) et les experts peuvent donner des avis différents. Donc, c'est difficile de trouver un équilibre et d'aboutir à une décision rationnelle. (HyoungSunYoo, YeLimJung, & Seung-PyoJun, 2023)

On peut observer par extension que l'évaluation de probabilité de succès d'un projet de R&D dépend aussi de l'asymétrie d'information<sup>3</sup> qui est plus ou moins conséquente entre le prêteur et l'entrepreneur innovant, car ce dernier détient une meilleure connaissance de ses chances de succès. (Hall & Lerner, 2009)

Comme le pointe Akerlof en 1970<sup>4</sup> au travers de son concept de « marché de citrons », le financement de projets innovants se traduit par une difficulté à trier entre les bons et mauvais risques, ce qui nous amène à demander des primes de rendement élevées pour investir dans la recherche et développement à risques. (Akerlof, 1970)

Il convient également de noter que la théorie économique fait appelle à la notion d'option réelle dans lequel on peut assimiler un projet de recherche et développement à une option que l'entreprise peut décider d'exercer ou d'abandonner à mesure que de nouvelles informations apparaissent. Autrement dit, les décisions d'investissements ou non à la suite des résultats intermédiaires vont définir en grande partie la valeur d'un projet innovant (Princeton U. press, 1994).

## *Le Biais de Jugement*

Au surplus, les biais de jugement déforment la perception sur la décision d'investissement.

Ainsi, les comités de sélection sont souvent averses au risque et pensent que les idées trop novatrices doivent être écartées, et les managers, en général prudents, préfèrent des activités d'investissement plus sûres à celles très risquées, or il est bien connu que ces dernières peuvent avoir un potentiel plus riche. (Singleton, 2021)

---

<sup>3</sup> Pour aller plus loin :

[https://www.nber.org/system/files/working\\_papers/w15325/w15325.pdf#:~:text=In%20the%20innovation%20setting%2C%20the,1970](https://www.nber.org/system/files/working_papers/w15325/w15325.pdf#:~:text=In%20the%20innovation%20setting%2C%20the,1970)

<sup>4</sup> Pour aller plus loin : Akerlof, G. A. (1970). *The Market for "Lemons": Quality Uncertainty and the Market Mechanism*. The Quarterly Journal of Economics, 84(3), 488–500 <https://www.jstor.org/stable/1879431>

Ainsi, le profit n'apparaît pas seulement comme le résultat de la bonne gestion des ressources, mais aussi comme une récompense pour avoir entrepris des décisions dans un contexte où l'avenir lui est totalement imprévisible.

### ***Analyse Moyenne-Variance de Markowitz***

C'est dans ce parallèle la que les travaux d'optimisation de la moyenne-variance de Markowitz (1952) met en lumière comment, grâce à la diversification du portefeuille, un investisseur peut mieux gérer son portefeuille en optimisant son profit et en minimisant les risques qu'il rencontre sachant le rendement qu'il espère obtenir. (Markowitz Mean-Variance Optimization Theory: A Comprehensive Exploration in Portfolio Management, s.d.)

Ainsi, en calculant le rendement moyen du portefeuille de l'agent économique et les covariances entre les actifs financiers, cela permet de répondre au mieux aux besoins des investisseurs. (Markowitz Mean-Variance Optimization Theory: A Comprehensive Exploration in Portfolio Management, s.d.)

### ***Joseph Schumpeter et sa théorie de la Destruction Créatrice***

Joseph Schumpeter s'appuie sur ce processus de recherche et développement pour sortir en 1912 sa théorie sur la destruction créatrice dans lequel il met en exergue le fait que l'innovation va entraîner la disparition de secteurs considérés comme établis par de nouveaux secteurs plus productifs. (La fonction Recherche & Développement, s.d.)

En effet, dans sa théorie, il relate le fait qu'il existe plusieurs formes d'innovations :

Cela peut être représenté par une innovation de produits, une innovation de procédés, une innovation de modes de production, une innovation de débouchés ou encore une innovation en matières premières. (La fonction Recherche & Développement, s.d.)

De fait, au travers de la section suivante, nous verrons comment l'intelligence Artificielle (IA) fonctionne et quels serait son rôle dans le domaine de l'investissement dans la R&D.

## 2.2 L'IA : Un outil innovant dans l'aide à la décision pour optimiser les investissements ?

### 2.2.1 Notions clées : IA, Machine Learning

#### **Définition de l'IA**

Pour le ministère de l'enseignement supérieur et de la recherche, l'intelligence artificielle (IA), n'est pas une innovation récente ; si elle a été créée à la suite des travaux du mathématicien Alan Turing au milieu des années 1950, le parlement européen qui définit l'IA comme l'« outil utilisé par une machine capable de « reproduire des comportements liés aux humains, tels que le raisonnement, la planification et la créativité » déjà dans les années 1970. (Intelligence artificielle (IA) : de quoi parle-t-on ?, 2024)

#### **Le Data Driven Decision Making (DDDM)**

Pour assister les entreprises et les gouvernements à la prise de la décision de sélectionner un projet de R&D, la décision fondée sur les données, ou DDDM, selon IBM, est une méthode permettant d'exploiter les sources de données à son profit afin, au bout du compte, de prendre la décision. Donc, cette méthode est une opportunité de croissance permettant aux entreprises d'optimiser leurs rendements et d'améliorer leurs stratégies par sa modalité itérative. (Mucci, 2024)

#### **L'analyse prédictive**

Qui plus est, les analyses prédictives, aussi connue sous le nom de prédictive Analytics, est une méthode issue des statistiques, de l'extraction de connaissances à partir d'un grand volume de données et du machine Learning afin de pouvoir analyser les faits présents et passés et ainsi, par extension faire des hypothèses prédictives. (Journaldunet, 2023)

#### **Définition du Machine Learning**

Il convient donc parallèlement à ce qui est dit précédemment de noter que pour aider les entreprises et les gouvernements dans la décision de sélection d'un projet de R&D,

l'apprentissage automatique, de son appellation d'origine « Machine Learning », est selon le CNIL est une subdivision de l'IA qui permet aux machines d'être plus performant en apprenant grâce à un ensemble de données d'entraînement et à des modèles mathématiques. (CNIL, s.d.)

On peut ajouter à cela que le modèle de Machine Learning<sup>5</sup> (ML) est composé de plusieurs familles :

- **Le modèle d'apprentissage supervisé (supervised Learning) :** En effet, il est défini selon les travaux de Roy & M. Tobin comme étant un modèle qui « analyse les données et déduit une fonction qui peut être utilisée pour associer de nouvelles observations à leurs cibles prédictives correspondantes. De fait, sachant un ensemble de n observations :  $\{(x_1, y_1), \dots (x_n, y_n)\}$ , avec  $x_i$ , désigne un vecteur prédicteur de la i -ème observation et  $y_i$ , l'étiquette ou la cible de la prédiction.

En d'autres termes, l'algorithme d'apprentissage supervisé ajuste une fonction inférée f sous la forme :  $\hat{y}_i = f(x_i)$ , où  $\hat{y}_i$  est la cible prédictive, qui peut être quantitative ou qualitative. Ainsi, le modèle de machine Learning s'apparente plus à soit des modèles de régressions (les régressions linéaire ou polynomiale) ou des modèles de classifications (les k plus proches voisins (k-NN), la régression logistique, forêts aléatoires) » (Roy & M.Tobin, 2024)

- **L'apprentissage non supervisé (unsupervised Learning) :** En effet, toujours selon le même rapport, « contrairement au modèle de machine Learning supervisé qui dispose d'une entrée et d'une sortie, les modèles d'apprentissage non supervisé apprennent à partir de données non étiquetées. De fait, ce modèle est plus utilisé par le biais du clustering et la réduction de dimensionnalité » (Roy & M.Tobin, 2024)
- **L'apprentissage semi-supervisé (semi-supervised Learning) :** Selon la définition de DataScientest, ce modèle « consiste à partir d'un jeu de données constitué en majorité de données non labellisées et en minorité de données labellisées. Il se situe entre l'apprentissage supervisé et l'apprentissage non supervisé. » (Robert, L'apprentissage semi-supervisé : Tout savoir ce qu'il faut savoir, 2024)
- **L'apprentissage auto-supervisé (self-supervised Learning) :** Effectivement, selon IBM, ce modèle est défini comme « une technique de machine Learning qui utilise

---

<sup>5</sup> Vous pourrez trouver plus d'informations sur la définition des différents types de machine Learning au travers des liens suivants : <https://www.jedha.co/formation-ia/algorithmes-deep-learning> et <https://www.ibm.com/fr-fr/think/topics/machine-learning-types>

l'apprentissage non supervisé pour des tâches qui nécessitent habituellement un modèle d'apprentissage supervisé. Par conséquent, au lieu de s'appuyer sur des jeux de données étiquetés pour les signaux de supervision, les modèles auto-supervisés génèrent des étiquettes implicites à partir de données non structurées. ("Qu'est-ce que l'apprentissage auto-supervisé - IBM") Qui plus est, ce modèle est très utile dans les domaines comme le traitement automatique du langage naturel (NLP) et la vision par ordinateur, qui demandent tous deux une grande quantité de données étiquetées pour entraîner leurs modèles d'IA. » (Bergmann, 2023)

- **L'apprentissage par renforcement (reinforcement Learning) :** En reprenant les travaux de (Roy & M.Tobin, 2024), on peut donc définir ce modèle par rapport aux autres comme un modèle « où l'algorithme est entraîné en récompensant les décisions souhaitées et/ou en sanctionnant les décisions non souhaitées. De fait au travers de ce modèle, les méthodes du Proximal Policy Optimization (PPO), le Soft Actor Critic (SAC) ou encore du Deep Q-Network (DQN) sont les plus utilisés. » (Roy & M.Tobin, 2024),
- **L'apprentissage profond (Deep Learning) :** Enfin, selon FUTURA ce dernier modèle d'apprentissage est défini par le fait « que la machine ait la capacité d'apprendre par elle-même en s'appuyant sur un réseau de neurones artificiels qui s'inspire du cerveau humain. De fait, ces neurones sont composés d'une dizaine voire une centaine de « couches » de neurones qui sont capables eux-mêmes d'interpréter les informations de la couche précédente. » (Futura)

Pour avoir une meilleure idée de ce que cela représente, la Figure 1 présente ci-dessous nous permet de comprendre que ce modèle fonctionne au travers d'une méthode itérative par rapport aux attributs renseignés dans les inputs pour ainsi, après de multiples erreurs et une phase d'apprentissage, répondre au besoin de l'utilisateur.

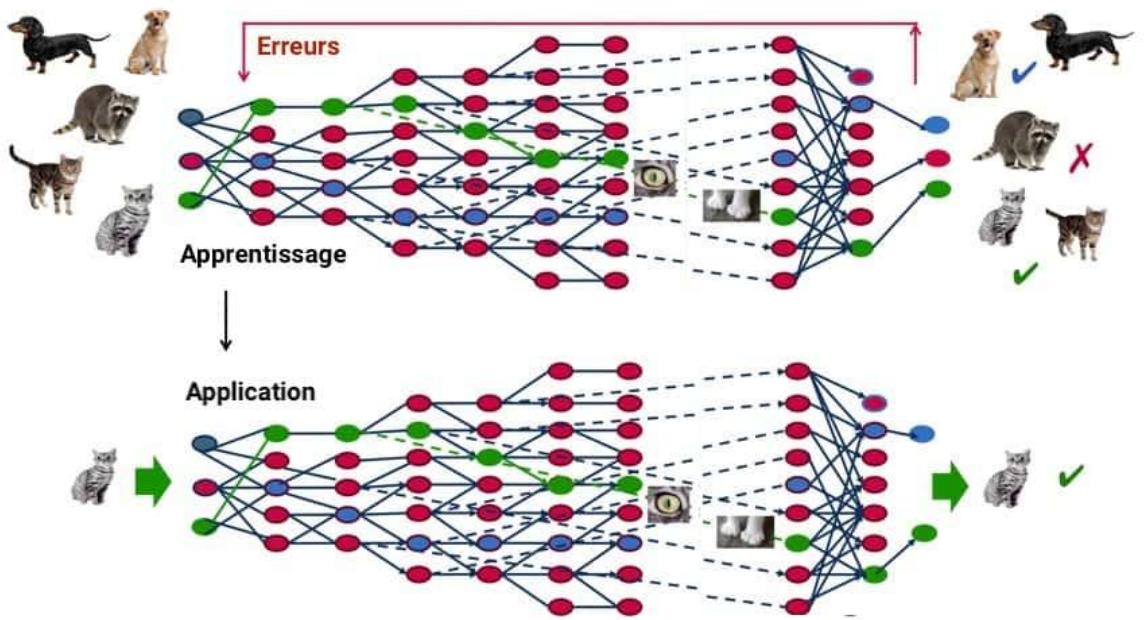
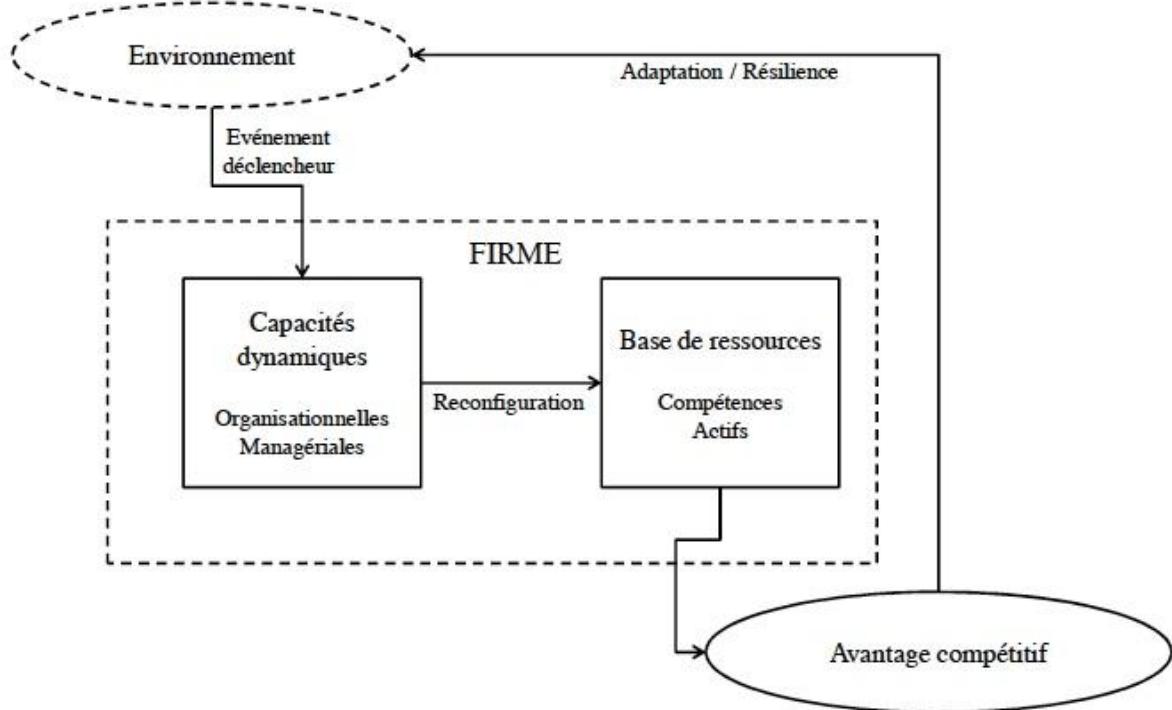


Figure 1: Image du processus d'apprentissage au travers du modèle de Deep Learning (MapR, C.D, Futura)

### ***La théorie des capacités dynamiques***

Sachant ce qui a été dit précédemment, la théorie des capacités dynamiques introduit par D. Teece et G. Pisano en 1994 nous permet de comprendre également à quel point il est important de s'adapter. (Théorie des capacités dynamiques: les compétences pour savoir s'adapter – D. Teece, s.d.)



*Figure 2 : Un modèle des capacités dynamiques (Labrouche, 2014)*

Ainsi, comme l'illustre la Figure 2, une firme qui est réactive face à un événement déclencheur repassera en revue se dont il possède pour pouvoir par extension s'adapter et obtenir un avantage concurrentiel. (Théorie des capacités dynamiques: les compétences pour savoir s'adapter – D. Teece, s.d.)

Qui plus est, l'introduction de l'IA ne fait pas exception car les organisations peuvent adapter et allouer leurs ressources selon des résultats des modèles sortis par le Machine Learning (ML).

## 2.2.2 Modèles et pratiques existant dans la gestion de portefeuille R&D par IA

L'IA joue un rôle crucial dans les processus de sélection de projets de R&D dans les grandes entreprises. Des cabinets de conseil comme EY<sup>6</sup> indiquent que les systèmes d'IA peuvent, non seulement, évaluer les projets de R&D en fonction de la profitabilité, mais aussi de détecter des problèmes techniques, financiers ou humains qui peuvent se retrouver dans ces projets. Donc,

<sup>6</sup> EY, aussi connu sous son ancien nom Ernst & Young, est un cabinet d'audit financier et de conseil. Membre du Big Four, il est le troisième réseau mondial quant au chiffre d'affaires (après Deloitte et PwC) en 2020

l'utilisation de l'IA dans l'évaluation des projets de R&D permet une meilleure allocation de ressources et de prioriser un projet sur l'autre en fonction de sa faisabilité et des contraintes qu'il pourrait poser. Par exemple, dans le secteur pharmaceutique, l'IA est utilisée pour évaluer en temps réel la faisabilité, les taux de réussite clinique, et les coûts de développement des nouveaux médicaments. (EY, 2022).

Plusieurs travaux académiques ont proposé divers modèles pour inclure l'IA dans les processus de sélection prise de décision de projets R&D. Par exemple, le modèle de « Smart R&D Investment Framework » développé par Huang et al en 2020.

Ils ont créé des modèles de Machine Learning et les ont utilisés avec les indicateurs financiers traditionnels, comme la VAN, le taux d'innovation<sup>7</sup> et l'échelle TRL<sup>8</sup> pour prendre des décisions mieux réfléchies et étudiés quant à l'attribution de budget aux projets de R&D.

Actuellement, il existe différents moyens par lequel l'IA peut aider les agents économiques à maximiser leur profit, et parmi eux nous avons :

### ***Les Robots Advisor***

Selon Yomoni, une plateforme française de gestion d'épargne en ligne, les robots Advisor sont des pseudo robot conseiller financier présenté sous forme de lignes de programmes informatique qui vont permettre « d'automatiser ce qui peut l'être (soit parce que la tâche a une faible valeur ajoutée, soit parce qu'elle est plus fiable lorsque réalisée informatiquement), tout en réservant l'intervention humaine aux tâches où elle a le plus de sens. » (Yomoni, s.d.)

De fait, au travers de l'IA, les robots Advisor offres à leurs « clients » des services comme la vérification de documents grâce à la reconnaissance textuelle, elle offre des conseils d'investissement adapté au niveau de risque de l'agent et signale lorsqu'il faut faire des ajustements à la suite d'une variation important du marché. (Yomoni, s.d.)

---

<sup>7</sup> Le taux d'innovation est comme un indicateur de vitesse pour les nouvelles idées et les améliorations. Il nous indique la rapidité et la fréquence avec lesquelles des solutions nouvelles et créatives sont élaborées et utilisées

<sup>8</sup> L'échelle TRL (Technology Readiness Level) est un outil d'évaluation et de mesure de la maturité technologique. Cette échelle va du niveau 1 au niveau 9, c'est-à-dire du stade de conceptualisation à celui du déploiement commercial. (L'échelle TRL : un outil clé pour obtenir des aides à l'innovation !, s.d.)

### ***Le Trading algorithmique (ou algo-trading)***

Selon le Business Review, il s'agit d'« une pratique qui consiste à faire usage de programmes informatique pour exécuter des ordres d'achats et de ventes selon des règles prédéfinies et ainsi offrir la possibilité de rendre le trading plus rapide et plus efficace. » (Business Review Magazine, 2025)

### ***Le Copy Trading AI***

En effet, contrairement à ces derniers, il s'agit d' « une stratégie automatisée qui permet de reproduire en temps réel les transactions de traders expérimentés ou de systèmes piloté par l'IA. » (K, 2025)

#### **2.2.3 Limites et défis théoriques**

Bien que l'IA présente des avantages quant à la sélection de projets de R&D, elle présente également certains défis.

##### ***Historique des projets de R&D***

Étant donné que l'historique des projets de R&D est un des principaux éléments utilisés dans les algorithmes de Machine Learning pour sélectionner les projets de R&D qui ont un potentiel de réussite, cela peut présenter en même temps un inconvénient, c'est ce qu'on appelle le biais algorithmique.

Ce biais algorithmique peut favoriser un projet qui ressemble à un autre projet du passé qui a eu du succès mais ce projet n'est pas forcément innovant. Donc, les algorithmes peuvent surévaluer des projets au détriment des autres qui sont atypiques selon l'algorithme.

Qui plus est, la dépendance sur les données historiques peut mener à une mauvaise identification des bonnes idées innovantes.

Par ailleurs, cela ne constitue pas la seule préoccupation sur l'IA dans le domaine de la R&D, il y en a d'autres comme :

### ***Le manque de transparence dans les modèles prise de décision automatisée***

En effet, les modèles de Machine Learning s'entraîne sur de gros volume de donnée et ils apprennent de ces données pour pouvoir prendre des décisions de manière autonome dans le futur, c'est ce qu'on appelle la prise de décision automatisée, c'est-à-dire sans l'intervention de l'homme.

### ***L'éthique***

Dans une sens générale, « l'éthique de l'intelligence artificielle (EIA) est le domaine qui étudie les valeurs, les principes et les règles qui doivent orienter le développement et l'utilisation de l'IA dans le respect de la dignité humaine, des droits de l'homme, de la démocratie, de l'état de droit et du bien commun. » (Sane, Upmynt, 2025)

De fait, il implique tous les acteurs de la vie économique tant au niveau microéconomique que macroéconomique.

## **2.3 Politique et écosystème d'IA aux Etats-Unis**

### **2.3.1 Écosystème d'innovation basé sur la synergie publique-privé**

Au travers du système publique et privée Américain, selon les études du centre National des statistiques scientifiques et d'ingénierie (NCSES, 2025), le secteur privé occupe une place prépondérante dans l'économie Américaine.

En effet, en 2022 les dépenses nationales en R&D aux Etats Unis s'élavait à 892 milliards de dollars avec près de 600 milliards de dollars qui proviennent du secteur privé et 164 milliards qui a été financé par le gouvernement fédéral. (NCSES, 2025)

De fait, 292 milliards de dollars ont servie à financer le domaine de la recherche, tandis que 600 milliards de dollars ont été consacré pour financer le développement expérimental. (NCSES, 2025)

C'est dire à quel points les grandes entreprises privées Américaines (comme Apple Inc., Microsoft Corporation ou encore Amazon.com, Inc.) sont lucratives pour les Etats-Unis d'Amérique.

En 2024 par exemple, en raison de la forte demande de puces dédiées à l'IA pour palier la hausse de l'utilisation de l'IA dans le secteur public, NVIDIA, une société multinationale américaine spécialisé dans les cartes graphiques, a atteint le top 5 des plus grandes entreprises américaines en termes de capitalisation boursière avec 2.812 Milliards d'euros, et en termes de bénéfices avec 78.92 Milliards d'euros. (NVIDIA, s.d.)

De fait, elle est devenue « un acteur clé de l'intelligence artificielle et de l'informatique accélérée. » (Canonne, 2025).

### 2.3.2 Stratégies nationales et initiatives fédérales

#### *Les modes de financements*

Les modes de financement en recherche et développement, tant dans le secteur public que privé est plus ou moins différent d'une région à l'autre.

En effet, en portant notre regard sur les Etats-Unis, la Small Business Act (SBA), permet à de nombreuses grandes agences fédérales comme le NSF, l'EPA, l'USDA, l'ED ou encore la NASA d'être financées grâce aux programmes comme le Small Business Innovation Research (SBIR) et le Small Business Technology Transfer (STTR) et ainsi permettre de prospérer plus facilement. (About SBIR and STTR)

Qui plus est, les modes de financements dans les projets innovants peuvent en autre se concrétiser à travers le programme Small Business Investment Company (SBIC) qui consiste en « l'investissement des capitaux privés et des capitaux empruntés garantis par la SBA pour réaliser des investissements en capital et/ou en dette à long terme dans des petites entreprises et des start-ups américaines. » (SBA, 2024)

D'autre part, selon les travaux de (Djermane & Benzaim, 2016), aux Etats-Unis la coopération entre les institutions, les incubateurs d'entreprises et le monde universitaire est devenue choses courantes et engendre par conséquent des répercussions positives sur la recherche et le développement.

On peut prendre l'exemple de Y Combinator ou encore techStars qui sont deux incubateurs de start-up les plus prestigieux des Etats-Unis. (Les meilleurs incubateurs de start up aux Etats Unis, 2025)

En effet, grâce à leurs longues années expérience à leurs nombreux mentors et investisseurs, elles ont permis à de nombreuses entreprises comme Airbnb, Dropbox ou encore Reddit d'exceller en sortant de leur programme et d'alimenter à leurs tour le cercle vertueux de recherche et de développement de leurs Pays. (Les meilleurs incubateurs de start up aux Etats Unis, 2025)

On peut par ailleurs prendre l'exemple du cas du réseau Manufacturing USA d'instituts technologiques qui est un réseau national d'instituts de fabrication liés entre eux qui permet ainsi à chaque institut de soit se focaliser sur une technologie unique, ou alors d'accélérer l'ensemble de l'industrie manufacturière américain. (ManufacturingUSA, 2019)

Dès lors, « 14 instituves ont été créés dont huit par le ministère de la défense (DoD), cinq par le ministère de l'énergie (DOE) et un par le ministère du commerce (DOC) » (ManufacturingUSA, 2019).

### ***Les lois***

Effectivement, des lois ont également été mise en place afin de pouvoir faciliter cette dynamique d'innovation.

Si l'on prend exemple du CHIPS and Sciences Act promulguée par le président Biden, un investissement des près de 52,7 milliards de dollars dans la fabrication de semi-conducteurs, la recherche et le développement, et un investissement dans la main d'œuvre américaine a été mise en place afin de pouvoir en fin de compte promouvoir la sécurité nationale et économique. De fait, par extension, en grande partie grâce à un programme d'incitation CHIPS du ministère du commerce ,15 sociétés ont promis d'injecter près de 400 milliards de dollars dans les semi-conducteurs à travers le pays et les États-Unis ont pour objectif de pouvoir être en capacité de produire près de 30 % de l'offre mondiale de puces de pointe d'ici 2032, alors qu'il n'y en avait 0 % lorsque le Président Biden et la Vice-Présidente Harris sont entrés en fonction. (The White House, 2024)

En autre, l'American AI Initiative<sup>9</sup>, lancée par l'ordre exécutif de février 2019 sous l'administration Trump, visait à focaliser les ressources des gouvernements fédéraux au

---

<sup>9</sup> Si vous voulez en savoir plus sur l'American AI Initiative :

<https://www.uspto.gov/sites/default/files/documents/National-Artificial-Intelligence-Initiative-Overview.pdf#:~:text=Leadership%20in%20AI%20Strengthen%20AI,Research%20Infrastructure%20re>

développement de l'IA pour améliorer la sécurité économique, nationale et la qualité de vie des Américains. (Parker, 2022)

L'objectif principal de cette initiative était d'augmenter les investissements dans le domaine de la recherche et du développement de l'IA à travers les différentes agences fédérales, fournir plus d'aide aux institutions académiques, aux entreprises technologiques et de créer, en parallèle, des standards autour de l'IA. (Accelerating America's Leadership in Artificial Intelligence , 2019)

De fait, selon le bureau de la politique scientifique et technologique, « bien que les Américains aient énormément profité d'être les premiers développeurs et leaders internationaux de l'IA, leurs places de leadership n'est pas garantie.

En d'autres termes, il faut qu'ils s'assurent que leurs avancements en matière d'IA restent en constante évolution tout en reflétant leurs valeurs et appliquées pour les bénéfices des citoyens américains ». (Accelerating America's Leadership in Artificial Intelligence , 2019)

Dans cette lutte concurrentielle perpétuelle, la Chine est le principal concurrent des États-Unis dans le domaine de l'IA. Rien ne qu'entre 2000-2015, elle a pu accroître ses investissements dans l'IA de 200% <sup>10</sup>. (Kelly, 2018)

Et ce n'est pas tout, elle a aussi pu sortir plusieurs IA génératives tout aussi performant que ceux des Etats-Unis et à moindre cout.

On parle ainsi de Deepseek par exemple qui est une AI générative multimodal et open-source dotée de capacité de raisonnement avancé et de génération d'image.

Le grand avantage avec cet outil est « qu'elle a été entraîné avec seulement 2 000 puces Nvidia pour un coût total inférieur à 5,58 millions de dollars, soit 10 fois moins que des modèles similaires.

---

<sup>10</sup> Si vous voulez en savoir plus sur le sujet : <https://oversight.house.gov/wp-content/uploads/2018/09/AI-White-Paper-.pdf>

À titre de comparaison, le projet « Star Blueprint » d'OpenAI, mentionné par Donald Trump, a nécessité 50 millions de dollars d'investissement.

Non seulement cet IA surpassé plusieurs IA leaders du marché, mais grâce à sa stratégie de domination par les couts, elle aussi su offrir aux PME et aux startup une voie plus accessible. » (Green, 2025)

Par conséquent, le 21 Janvier 2025, Donald Trump a instauré un nouveau projet, le « projet le plus important de cette époque » dit-il : le projet Stargate. (Franceinfo, 2025)

En effet, selon le JDN il s'agit, « d'un programme d'investissement de 500 milliards de dollars sur quatre ans aux Etats-Unis dans les infrastructures d'IA, conduit par OpenAI, SoftBank, Oracle, et MGX, il est aussi soutenu par Microsoft, NVIDIA et Arm. » (IA : le projet Stargate de Trump prend forme, 2025)

Bien que la majorité des investissements se font en privé, Trump soutient parallèlement par son approche « American first » <sup>11</sup> mettre en place des droits de douane égalitaire car les pays étrangers (France, Chine) importent moins que ce qu'eux même importe comme bien.

De fait, une instauration de droit de douane égalitaire permettrait aux Etats-Unis de trouver non seulement ralentir un tant soit peu les autres dans cette cours à l'IA, réguler la balance commerciale des Etats-Unis mais aussi trouver des fonds pour cette course effrénée de l'IA pour savoir qui atteindra en premier intelligence artificielle générale (IAG ou AGI en anglais), c'est-à-dire une IA qui serait capable d'« exercer la plupart des professions humaines au moins aussi bien qu'un être humain moyen » (Bostrom, 2014).

### 2.3.3 Enjeux éthiques et de souveraineté technologique

En effet, du point de vue de l'éthique de l'IA (EIA) aux Etats-Unis, au cours de la journée du sommet de l'intelligence au Grand Palais à Paris, le vice-président des Etats-Unis J.D. Vance a soutenu que « les États-Unis sont les leaders dans l'IA et leur administration entend qu'ils le

---

<sup>11</sup> Il s'agit d'une approche qui privilégie toujours les intérêts de l'Etat Américain

restent pour être des partenaires de choix pour les autres pays et certainement pour les entreprises à mesure qu'ils étendent leur propre utilisation de l'IA.

De fait, ils sont fermement convaincus que l'IA doit rester libre de tout parti idéologique et que l'IA ne sera pas détournée pour servir d'outil de censure autoritaire. » (Le Parisien, 2025)

## Partie 3 – Méthodologie

Après avoir établit les fondements conceptuels et l'état de l'art de ce qui concerne l'optimisation des décisions de recherche et de développement, ce chapitre servira à vous présenter la méthodologie que nous avons voulu adopter afin de mener à bien notre mémoire.

Au travers de cette partie, nous parlerons des différentes sources de données auxquels nous avons fait appelle, les différentes méthodes d'analyses retenues ainsi que les défis que nous avons rencontré dans cette démarche.

### 3.1 La collecte des données

Les différentes analyses que nous allons mettre en place seront basé sur la constitution d'une base de données solide et d'études de cas tiré de sites certifiés dans le seul et unique but de pouvoir effectuer une analyse mixte de l'optimisation des décisions d'investissements en R&D avec l'IA aux Etats-Unis.

#### 3.1.1 Source de données utilisées pour l'analyse quantitative

## **Sources quantitatives**

Pour mener à bien notre étude quantitative, nous nous sommes penchés sur les sources suivantes :

- **National center for science and Engineering Statistics (NCSES)** qui est définie selon le site officielle comme « la principale agence statistique des Etats-Unis qui a pour mission unique de fournir des données fiables et de qualité à propos de l'entreprise scientifique et technique. » (NCSES, 2025)  
Il s'agit par ailleurs d'une subdivision de la **National science Fondation (NSF)** qui est une organisation scientifique créée en 1950 ayant pour but de « promouvoir les progrès de la science, améliorer le bien être national, la prospérité, assurer la défense nationale et bien d'autres domaines ». (NSF, s.d.)
- **L'Organisation de Coopération et de Développement Economiques (OCDE)** est représentée selon le site officielle comme « une organisation internationale qui s'appuie sur 60 ans d'expérience et de connaissances pour pouvoir promouvoir une croissance économique plus forte, plus verte et plus juste et éléver les niveaux de vie et d'emploi ». (L'OCDE : Des politiques meilleures pour une vie meilleure, s.d.)
- **L'OECD.AI** : Effectivement, pour ce qu'il en est de cette organismes, d'apres le site officiel il s'agit « d'une branche de l'OCDE qui rassemble plus de 250 experts du monde entier qui apporte des réponses politiques sur des sujets émergents, depuis l'évaluation des risques liés à l'IA jusqu'aux outils permettant de mettre en œuvre une IA digne de confiance et de mesurer les capacités nationales en matière de calcul de l'IA. En d'autres mots, cet organisme aide dans la gestion des risques en matière d'IA, elle aide dans la recherche sur l'avenir potentiel de l'IA, elle définit et suit les incidents liés à l'IA (notamment sur le changement climatique) ». (OECD.AI, s.d.)
- **Word Bank Group (WBG)**, aussi connue sous le nom du Groupe de la Banque Mondiale en français (GBM), est dans son cas représenté comme « une organisation internationale créée en 1944 à Washington et constitué actuellement de 189 pays membres qui a pour but de lutter contre la pauvreté et encourager le développement économique du monde entier. En d'autres termes, sans compter son côté humanitaire, elle offre également une expertise technique en fournissant des conseils et des analyses afin d'aider les gouvernements à mieux gérer certaines ressources. » (La Banque mondiale : rôle, missions et impact, s.d.)

- **L'Organisation Mondiale de la Propriété Intellectuelle (OMPI)** est définie selon le site officiel comme « une institution des Nations Unies créée en 1967 à Genève (Suisse) et est constitué actuellement de 197 états membres. De fait, elle a pour objectif de veiller à ce que les idées des innovateurs et des créateurs du monde entier, parviennent en toute sécurité sur le marché et améliorent les conditions de vie de tous et partout. » (À propos de l'OMPI, s.d.)

**L'United States Patent and Trademark Office (USPTO)**, aussi connue sous le nom du Bureau américain des brevets et des marques de commerce en français, est définie selon le site officiel comme « une agence fédérale qui dépend du ministère du commerce et qui est responsable des offres et des demandes de brevet afin de pouvoir soutenir l'innovation et la compétitivité sur le territoire États-Uniens. » (USPTO, s.d.)

### ***Sources qualitatives***

Pour pouvoir effectuer notre étude sur l'optimisation des décisions d'investissement, nous nous sommes servis, pour notre analyse qualitative de :

- **CAIRN.info** qui est une plateforme « qui à favoriser la découverte par le plus grand nombre d'une recherche scientifique francophone de qualité, tout en cultivant l'indépendance et la diversité des acteurs de l'écosystème du savoir. » (Cairn.info, s.d.)
- **JSTOR** qui est également une plateforme de recherche et d'enseignement qui est assez riche en information tant au niveau des livres, des images, d'article de revues ou encore de sources primaires dans plus de 70 disciplines. (About JSTOR, s.d.)
- **Office of Science and Technology Policy (OSTP)** est une des structures attachées au bureau du président des États-Unis. Elle a pour but de conseiller sur les effets de la science et la technologie sur les affaires intérieures et extérieures des États-Unis.<sup>12</sup>

### **3.1.2 Méthodologie de collecte des sources**

#### ***Méthodologie de collecte des sources quantitatives***

Afin de pouvoir mener à bien nos analyses et par extension répondre aux différentes sous-question de notre mémoire, nous avons réalisé nos recherches manuellement et de manière itérative, en utilisant notamment des mots clés en français et en anglais sur les sites officiels mentionnés ci-dessus.

Qui plus est, nous avons limité nos variables sur une période chronologique établit entre 2010 qu'à maximum 2024 afin de pouvoir apprécier les différentes tendances et évolutions qui en découlent.

Sinon, on peut également mentionner le fait que les fichiers sur que nous avons tirés de l'OCDE ont été particulièrement utile pour pouvoir effectuer dans les meilleures conditions nos différents analyses. (Office of Science and Technology Policy (OSTP))

### ***Méthodologie de collecte des sources qualitatives***

Notre collecte de données s'articule sur trois axes principaux :

Premièrement, nous avons décidé de concevoir un questionnaire<sup>13</sup> qui sera envoyé à toutes les personnes travaillant dans le domaine de R&D aux États-Unis via LinkedIn. Ce questionnaire nous permettra de recueillir des données empiriques concrètes sur l'utilisation de l'IA dans les processus de R&D et sur ses impacts perçus.

Deuxièmement, l'analyse documentaire, nous nous appuyons sur des articles scientifiques, académiques et des données publiques afin de renforcer nos analyses avec des données et informations venant du milieu académique. Cette documentation permet de confronter nos résultats empiriques avec les apports théoriques

Troisièmement, nous avons sélectionné des entreprises américaines qui utilisent l'IA dans leurs processus de R&D. Ces études de cas nous permettent d'illustrer concrètement les mécanismes mis en œuvre et les résultats obtenus de la part des entreprises.

### **3.2 Les variables : variables en lien avec l'Investissement en R&D et l'IA aux Etats-Unis.**

---

<sup>13</sup> Vous pourrez accéder au questionnaire en version française via le lien suivant : <https://forms.gle/QYTcLEQyu4Jy8u5j6>

### 3.2.1 Les variables quantitatives

#### ***La Dépense Intérieur de R&D (DIRD)***

Elle est définie selon l'INSEE comme étant un indicateur qui « comprend les moyens financiers (nationaux et étrangers) mobilisés pour l'exécution des travaux de recherche et développement (R&D) sur le territoire national par le secteur des administrations (DIRDA) et par le secteur des entreprises (DIRDE).

Qui plus est, elle comprend aussi les dépenses courantes (masse salariale des personnels de R&D et dépenses de fonctionnement) et les dépenses en capital (achats d'équipements nécessaires à la R&D). » (Dépense intérieure de recherche et développement expérimental / DIRD, 2021)

#### ***Le Produit intérieur brut des Etats-Unis<sup>i</sup>***

Toujours selon l'INSEE, il s'agit de la « richesse créée par tous les agents, privés et publics, sur un territoire national pendant une période donnée.

Par conséquent, elle est le fruit de l'activité de production tant en termes de bien ou de services réalisés par les agents économiques qui vivent à l'intérieur du pays. » (Produit intérieur brut aux prix du marché / PIB, 2021)

#### ***Les offres et les demandes de Brevets<sup>ii</sup>***

Pour reprendre à peu de choses près ce qui a été dit dans la revue de littérature, la demande de brevet permet selon l'INPI de « protéger une innovation technique, c'est-à-dire un produit, un procédé ou une utilisation qui apporte une solution technique nouvelle et inventive à un problème technique donné. » (INPI)

Tandis que l'offre est quant à lui définie par le ministère de l'économie, des finances et de la souveraineté industrielle et numérique comme « l'opération par laquelle le titulaire du brevet

décide de céder à autrui le droit de propriété reconnu sur l'invention brevetée » (Comment transmettre un brevet ?, 2025)

### ***Les offres et les demandes de Brevets d'IA<sup>iii</sup>***

Dans ce cas-là il s'agit de la même logique mais le type de brevet est beaucoup plus axé sur l'IA.

### ***L'investissement en capital risque (ou venture capital (VC))***

Cet indicateur est défini par Bpifrance comme « une prise de participation par un ou des investisseurs, généralement minoritaire(s) et au capital de sociétés non cotées dont l'objectif est de participer financièrement au développement des entreprises innovantes à fort potentiel et d'avoir un retour sur investissement positif lors de la cession de ses titres. » (Capital risque, 2024)

### ***L'investissement en capital risque de l'IA***

Il s'agit de là l'investissement en capital risque mais plus axé dans le secteur de l'IA.

### ***Business enterprise expenditure on R&D (BERD)<sup>iv</sup>***

Effectivement le Business Enterprise expenditure on R&D (BERD) aussi plus communément connue sous le nom de la Dépense Intérieur de Recherche et Développement dans le secteur des Entreprises (DIRDE) est défini selon l'UNESCO comme « la composante des DIRD engagées par les unités du secteur des entreprises au cours d'une période de référence donnée. » (Dépenses intérieures de R-D des entreprises (DIRDE), s.d.)

### ***Le cycle time-to-Market des Brevets<sup>v</sup>***

Selon le JDN il s'agit « d'une expression anglo-saxonne qui correspond à la durée de développement d'une construction d'une offre commerciale ou d'un produit. Autrement dit, c'est un enjeu stratégique majeur pour les entreprises car ça les permet de détenir un avantage concurrentiel par rapport à leurs concurrents. » (JDN, 2019)

### ***Le Crédit Budgétaire public<sup>vi</sup>***

Il est défini comme étant une « autorisation donnée au gouvernement par le Parlement dans la loi de finances » (Mattret, 2006)

*Le Crédit d'impôt recherche (CIR)<sup>vii</sup>*

*Nombre de chercheurs à temps plein<sup>viii</sup>*

*Pourcentage de publications scientifiques impliquant une collaboration internationale*

*Publications scientifiques sur l'IA impliquant une collaboration internationale*

*Nombre total de publications scientifiques*

### 3.2.2 Les variables qualitatives

#### *Etude de cas*

Pour alimenter nos analyses, nous avons prévue de traité les études de cas suivantes :

- **Pfizer** : Selon le site officiel de Pfizer, il s'agit « d'une entreprise pharmaceutique parmi les leaders mondiaux, l'innovation et l'excellence technologique pour faire reculer la maladie, en offrant des médicaments et vaccins qui prolongent et améliorent considérablement la vie du plus grand nombre » (Pfizer)  
De fait, elle nous ait pertinente car elle utilise l'IA pour accélérer la recherche de médicament.

En d'autres termes, cette étude de cas permettra d'appuyer nos hypothèses.

#### *Questionnaire*

Sinon, nous avons également prévue de compléter notre analyse en réalisant un questionnaire dédier aux professionnelles du milieu de l'investissement et de la R&D aux Etats unis.

### 3.3 Méthodologie d'analyse des variables quantitatives et qualitatives

Dans le but de pouvoir correctement analyser notre mémoire qui porte sur l'optimisation des décisions en R&D grâce à l'IA aux Etats-Unis, nous allons nous servir d'une approche méthodologique mixte.

Autrement dit, nous nous servirons d'analyses quantitatives (statistiques descriptives, ACP, modélisation, régressions) et des modèles qualitatifs (questionnaires, études de cas).

Cette approche permettra de dégager des tendances, d'identifier des mécanismes et d'enrichir la compréhension d'un phénomène encore en évolution.

Pour cela, nous avons mobilisé des méthodes quantitatives et qualitatives pour trouver empiriquement des réponses à nos sous-questions.

#### 3.2.3 Pourquoi avons-nous choisi ces variables ?

Nous avons choisi ces variables car elles nous permettront de mettre en lumière l'évolution des décisions d'investissement en recherche et développement sur le sol américain et de trouver des solutions aux questions que nous nous sommes posés.

#### 3.3.1 Approche quantitative

Tout au long de cette analyse quantitative, nous allons établir des opérations statistiques et économétrique qui puissent répondre aux sous-questions du mémoire que nous nous sommes posés, à savoir :

Quels secteurs ou types d'entreprises américaines intègrent l'IA dans leur processus de R&D, et comment ?

Quels sont les impacts mesurables de l'IA sur la performance des investissements en R&D dans le contexte américain ?

Les recherches associées à l'IA en R&D permet-elle une concentration des brevets qui serait plus privilégié pour les grandes voir très grandes firmes ou, au contraire, serait-elle plus à même d'être généraliser ?

De fait nous réaliserons :

### ***Des analyses descriptives***

Dans cette section nous comptons réaliser dans un premier temps des analyses statistiques primaires pour pouvoir poser les bases.

Ensuite, nous continuerons en réalisant des premières analyses de corrélations avec la matrice de corrélation et l'analyse en composante principale (ACP).

Par ailleurs, nous entreprendrons une analyse comparative au travers de la courbe de Lorenz définie par  $L(p_i) = \sum_{j=1}^i x(j) / \sum_{j=1}^n x(j)$  pour visualiser le niveau de concentration des Brevets d'IA et l'indice de Herfindahl-Hirschman (HHI) défini par  $HHI = (\sum s_i^2) * 10\ 000$  qui nous servira plutôt à calculer le niveau de concentration des brevets IA par rapport aux autres secteurs et l'indice de GINI =  $|1 - \sum_{k=1}^{n-1} (X_{k+1} - X_k) * (Y_{k+1} - Y_k)|$  défini pour compléter notre analyse sur le niveau de concentration de ce dernier. (Indice Herfindahl-Hirschman (HHI), s.d.) (Rodriguez, s.d.)

En outre, nous réaliserons également un Gradient et un gradient descendant pour pouvoir d'un côté indiquer la direction dans lequel notre fonction va croître le plus vite et d'un autre côté trouver le point qui va minimiser notre fonction de perte. (Robert, Algorithme de descente de gradient : Qu'est-ce que c'est ?, 2020)

### ***Des modélisations économétriques***

Lorsque cette première partie sera achevée, nous procéderons ensuite à la mise en œuvre et à l'analyse de divers modèles économétrique comme :

-**Le Modèle de Régression Multiple** qui permettra de modéliser trois régressions.

Nous aurons donc en premier lieu les dépenses intérieures en R&D comme variable dépendante par rapport au PIB, les publications scientifiques sur l'IA impliquant une collaboration internationale et le nombre de chercheurs à temps plein.

Ensuite, nous représenterons le cycle de time to Market en tant que variable dépendante par rapport au PIB Américain et les publications scientifiques sur l'IA impliquant une collaboration internationale.

Et enfin nous représenterons également la productivité des brevets comme variable dépendante par rapport aux publications scientifiques sur l'IA impliquant une collaboration internationale, par rapport au nombre de chercheurs à temps plein, par rapport aux dépenses fédérales de R&D et par rapport à la productivité de l'IA.

**-Un Modèle SARIMAX** (Seasonal Autoregressive integrated Moving Average) pour le DIRD sera présenté car elle permettra, selon Data scientest « de modéliser, prévoir des séries temporelles et capturer les motifs qui se répètent à intervalles réguliers afin de pouvoir, par extension améliorer les prévisions ». (Kassel, 2024)

De fait, nous utiliserons donc pour réaliser dans les meilleures conditions nos analyses des visualisations et des modélisations, que nous traiterons grâce aux langages Python essentiellement (avec les packages Numpy, pandas, sklearn, statsmodels, missingno, seaborn et matplotlib), mais nous utiliserons également Excel pour réaliser quelques traitements primaires.

### 3.3.2 Approche qualitative

#### *Etude de cas*

Les études de cas que nous avons trouvé nous permettront d'éclairer plus en approfondis les différentes sous questions que nous avons mise en place.

#### *Questionnaire*

Le questionnaire dédié aux professionnelles du milieu de l'investissement et de la R&D aux Etats unis que nous avons décidé de faire sera partager au travers de LinkedIn, un réseau social où des professionnels peuvent interagir entre eux et proposer des opportunités.

### 3.3.3 Limites et défis de la méthodologie mixte

#### *La disponibilité des données*

En effet, lors de nos recherches divers et variées, certaines variables que nous voulions implémenter dans notre mémoire n'ont malheureusement pas pu l'être car elles étaient soit incomplète, payante ou inaccessible.

#### *Echelle de dimension différente*

Effectivement, en raison des sources de collecte différentes, nos variables n'étaient pas tous représenté sous la même échelle (Ex : les DIRD total par rapport aux PIB Etats-Unien). Pour palier a ce problème, nous avons donc réaliser une indentation par la méthode des k plus proches

voisins <sup>14</sup>(K-NN) à toutes les variables qui ne dépasse pas un taux de 30% de valeurs manquante. Cette méthode consiste donc selon le DataScientest « à trouver dans un premier temps un k voisin , puis calculer la distance soit par le biais de la distance euclidienne =  $\sqrt{\sum(x_i - y_i)^2}$  , ou de la distance de Manhattan =  $\sum|x_i - y_i|$  , ensuite prendre les k voisins les plus proche en fonction de la méthode de calcul de la distance et enfin parmi les k voisins les plus proches voir le nombre point qui appartient à une catégorie et attribuer le nouveau point par rapport à la catégorie qui est la plus présente parmi les k voisins. » (Qu'est ce que l'algorithme KNN ?, s.d.)

Par ailleurs, dans le cadre de ce mémoire, pour plus de sûreté, nous l'avons réalisé avec le package Sklearn.impute de python.

### **Colinéarité**

Il faut dire que certaines de nos variables étaient fortement corrélées (PIB, DIRD et Total R&D). De fait, pour faire par exemple notre matrice de corrélation, nous avons dans un premier temps tourné nos variables quantitatives en format logarithmique et sous forme de taux de croissance pour réduire l'échelle des variables et leurs niveaux de corrélation.

De plus, nous avons également réalisé une analyse des facteurs d'inflation de la variance (VIF) pour tenter de limiter au maximum la multi colinéarité dans nos modèles de régressions linéaire multiple. En effet, cette méthode permet, selon le DataCamp de « quantifier la multi colinéarité d'un modèle.

De fait, cet indicateur est représenté par la formule ci-dessous :

$$VIF_j = \frac{1}{1 - R^2_j}$$

Avec où  $R^2$  est la valeur du coefficient de détermination ajusté obtenue en régressant le jème prédicteur sur les prédicteurs restants.

Le but est que le VIF doit être strictement inférieur à 5 sinon il y aura un sérieux problème de multi colinéarité dans notre modèle. Dans notre cas, nous avons déterminé nos analyses avec

<sup>14</sup> Vous pourrez voir nos traitements par la méthode des k-nn au travers de l'annexe 1 à 3

les variables qui détenaient le VIF les plus faibles, et ainsi permettre par extension de pouvoir réaliser des analyses les plus solide et conforme possible. (Voir annexe) (Singh, 2024)

### ***Collecte de données du Questionnaire incomplet***

Malheureusement bien que nous ayons réalisé notre questionnaire, nous n'avons pas pu récolter assez de données afin de pouvoir mener à bien nos analyses et ainsi, par extension, répondre aux sous-questions que nous nous sommes posées. Nous sommes donc arrivés à la conclusion de mettre de côté notre questionnaire.

### ***Biais d'interprétation***

Effectivement, en raison de nos divers traitements pour s'assurer que notre modèle respecte bien les hypothèses de validité, (sans compter la temporalité de nos rapports études de cas), nos analyses peuvent être biaisées.

## Partie 4 – Résultats & Discussions

Après avoir exposé notre méthodologie sur la façon dont nous allons traiter les différentes variables que nous allons mettre en relation, cette partie sera plus consacrée à l'exposition des résultats que nous avons pu trouver et à l'apport des éléments de réponses aux sous questions de notre problématique.

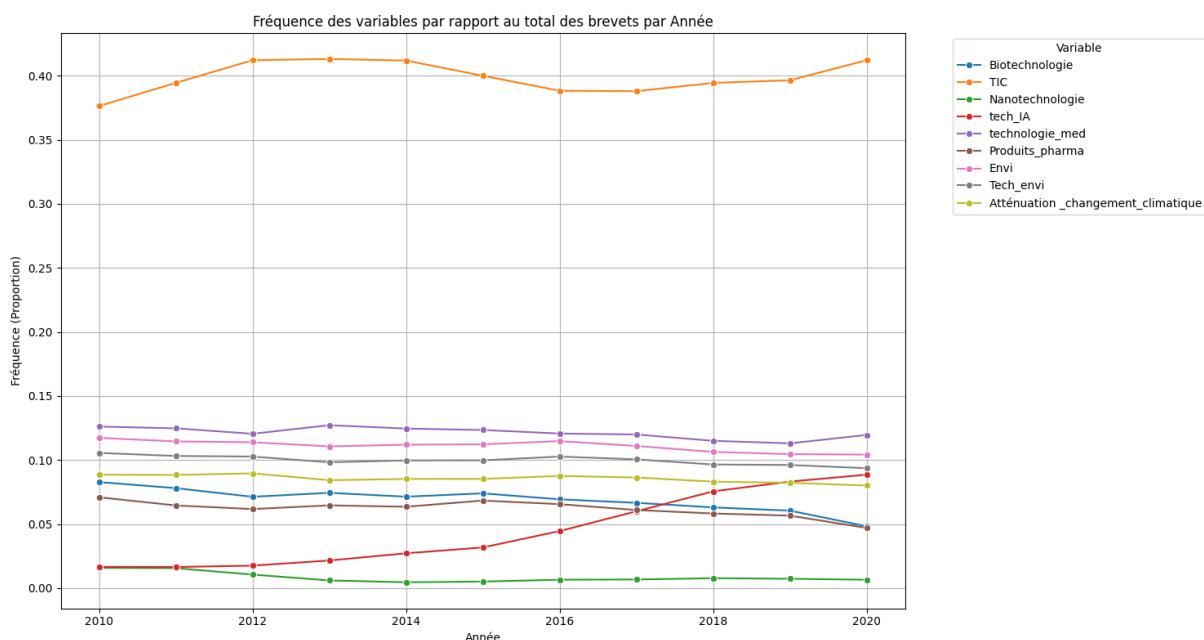
### 4.1 Résultats

#### 4.1.1 Analyses descriptives des données

##### *Analyses statistiques*

###### Répartition des meilleurs secteurs des brevets par années

Figure 3: Représentation graphique de la répartition des meilleurs secteurs du brevet de 2010-2020



Source : Elaboration personnelle tiré des données de l'OCDE

Au travers de la représentation graphique de la répartition des meilleurs secteurs du brevet de 2010-2020, on peut instantanément voir à quel point les brevets de technologie de l'information et de la communication (TIC) surplombe les autres secteurs.

Cependant, bien qu'elle parte toujours sur une bonne tendance évolutive de 2018 à 2020, on peut voir que les brevets dans le domaine de la technologie de l'IA gagnent du terrain et ne cesse de croître.

De fait elle est passé d'une catégorie de brevet qui représentait moins de 5% des Brevets à une catégorie de brevet qui frôle les 10% (soit 5 points de pourcentages en plus).

Qui plus est, au travers toute cette période, on peut apprécier deux catégories différentes.

Une catégorie de brevet qui finit sur une tendance croissante (notamment avec domaine de la technologie de l'IA, les TIC, les technologies du secteur médicale ou encore le domaine environnementale), et une autre catégorie qui présente plus une tendance finale décroissante comme les brevets biotechnologiques ou encore les produits pharmaceutiques) .

En bref, on se retrouve finalement au top 5 des brevets demandés les brevets de type TIC, puis les brevets de dans le domaine de la technologie médicale, suivi de près par ceux dans le domaine environnementale, les produits pharmaceutiques et enfin les brevets associés aux technologie de l'IA.

### ***La productivité de l'IA***

En effet, afin de pouvoir déterminer le niveau de productivité de l'intelligence artificielle , nous avons réalisé la formule suivante :

$$\text{Productivité de l'IA}^{15} = \frac{\text{Demandes de brevet IA} + \text{Offres de brevet IA}}{\text{DIRD total (en million)}}$$

---

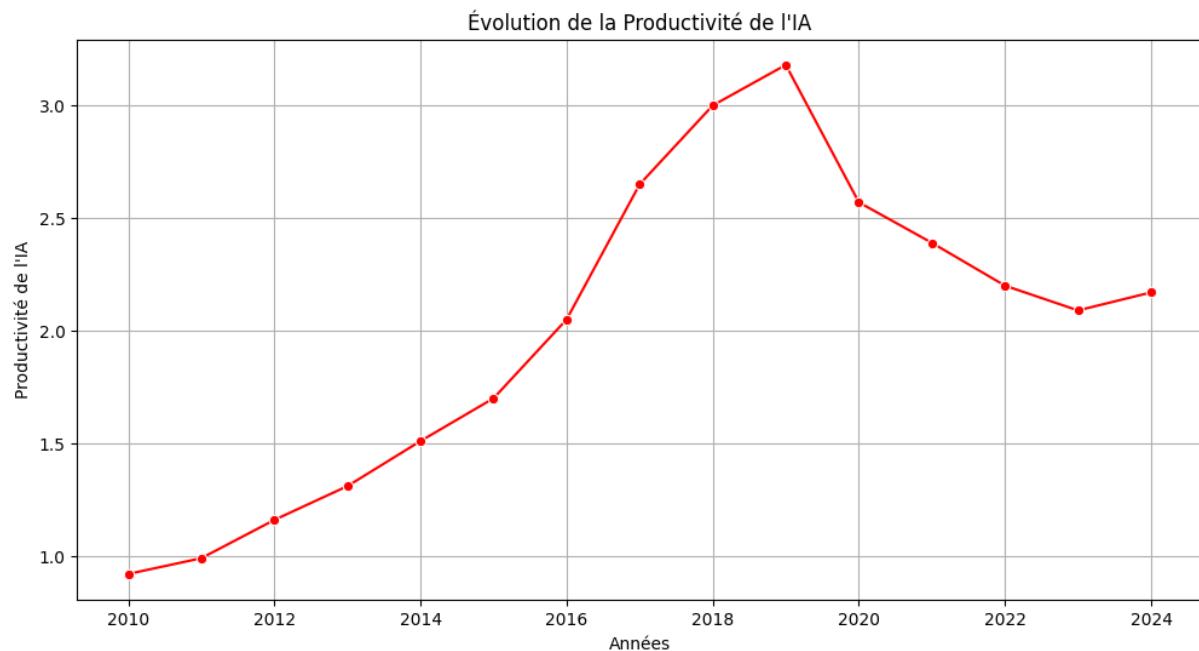
<sup>15</sup> Vous pourrez trouver la formule de cette variable au travers de ce lien : [Intelligence artificielle et productivité : preuves mondiales à partir de données brevetées et bibliométriques d'IA](#)

*Tableau 1: La productivité de L'IA*

	Demande de Brevet	Offre de Brevet	DIRD Total(en million)	Productivité IA (en %)
2010	1966.756958007813	1811.427856445313	408495.0	0.92
2011	2104.708740234375	2133.5830078125	427125.0	0.99
2012	2480.8173828125	2560.498291015625	434424.0	1.16
2013	2970.44873046875	2977.4912109375	455089.0	1.31
2014	3633.637939453125	3564.286376953125	476971.0	1.51
2015	4318.05322265625	4298.70751953125	507372.0	1.7
2016	5439.08837890625	5479.74560546875	533451.0	2.05
2017	7595.18896484375	7379.5302734375	565446.0	2.65
2018	9409.470703125	9131.6064453125	617538.0	3.0
2019	11552.21875	10001.958984375	676995.0	3.18
2020	11436.9189453125	7349.2587890625	729857.0	2.57
2021	10946.8837890625	8675.60888671875	821478.0	2.39
2022	11249.55126953125	8675.60888671875	906953.0	2.2
2023	11249.55126953125	8675.60888671875	955578.0	2.09
2024	11494.56884765625	8675.60888671875	931265.5	2.17

*Source : Elaboration personnelle à partir des données de l'OCDE*

*Figure 4 : Représentation graphique de l'évolution de la productivité de l'IA de 2010-2024*



*Source : Elaboration personnelle à partir des données de l'OCDE*

En effet, au travers de la représentation graphique de l'évolution de la productivité de l'IA de 2010 à 2024, on peut observer trois phases majeures.

On peut donc constater dans un premier temps une tendance haussière assez marqué qui commence en 2010 avec un pourcentage d'environ 0.9% pour atteindre ensuite un pic de plus de 3% en 2019.

Cette première phase pourrait s'expliquer par le fait que l'Amérique à commencer à fortement s'intéressé au potentiel de l'intelligence artificielle (on peut notamment le voir par l'appréciation des demandes et des offres de brevets associé à l'IA par rapport au dépenses internes en recherche et développement américain (voir annexe)).

A l'inverse, nous observons à la suite de cette première phase une tendance baissière qui va de 2019 avec une productivité d'IA évalué à plus de 3% pour ensuite tomber sur une productivité de l'IA d'environ 2.2% en 2023.

Cela s'expliquerait en partie en raison de la crise conjoncturelle du Covid-19 qui a toucher le monde entier et qui a poussé les Etats-Unis à se confiné. De fait, ce confinement a donc créer

un ralentissement des cycles de recherches et de développement et par extension, de la productivité de l'intelligence artificielle.

Cette récession pourrait également s'expliquer par le fait que les améliorations perçues en début 2010 jusqu'en 2018 étaient des avancé majeurs et qu'au fur et à mesure que ses améliorations se faisaient paraître (couplé à la crise conjoncturelle du covid), le potentiel d'amélioration aurait petit à petit diminué.

Pour donner suite à cette seconde phase majeur, on peut constater une phase reprise dès 2023 faisant ainsi passer la productivité de l'IA d'environ 2.2% à 2.3%. Cette reprise s'expliquerait du fait que les Etats-Unis ont pu sortir de cette crise conjoncturelle et reprendre en main les flux de demandes et d'offre de brevets dans le secteur de l'IA.

### ***La productivité des Brevets***

En effet, la productivité des brevets nous permet de nous rendre compte de l'intensité d'innovation présente sur le sol Américain.

De fait, la productivité des brevets est définie par la formule suivante :

$$\text{Productivité des brevets}^{16} = \frac{\text{Demandes de Brevet} + \text{Offres de brevet}}{\text{DIRD Total (en million)}}$$

---

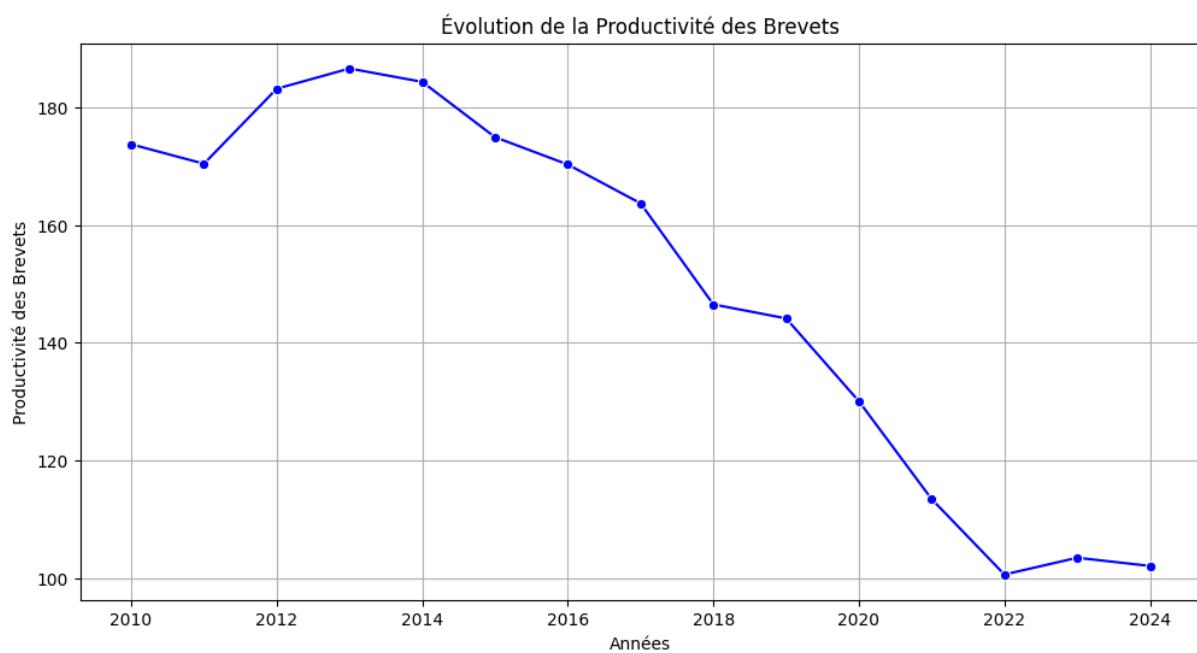
<sup>16</sup> Cette formule a été réalisé au travers du lien suivant : [Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators](#)

Tableau 2: La productivité des brevets

	Demande de Brevet	Offre de Brevet	DIRD Total(en million)	Productivité des Brevets
2010	490226.0	219614.0	408495.0	173.77
2011	503582.0	224505.0	427125.0	170.46
2012	542815.0	253155.0	434424.0	183.22
2013	571612.0	277835.0	455089.0	186.66
2014	578802.0	300677.0	476971.0	184.39
2015	589410.0	298408.0	507372.0	174.98
2016	605571.0	303049.0	533451.0	170.33
2017	606956.0	318828.0	565446.0	163.73
2018	597141.0	307759.0	617538.0	146.53
2019	621453.0	354430.0	676995.0	144.15
2020	597175.0	352066.0	729857.0	130.06
2021	593294.0	338335.0	821478.0	113.41
2022	589155.0	323018.0	906953.0	100.58
2023	642215.0	346152.0	955578.0	103.43
2024	615685.0	334585.0	931265.5	102.04

Source : Elaboration personnelle à partir des données de l'OCDE

Figure 5: Représentation graphique de l'évolution de la productivité des brevets de 2010-2024



*Source : Elaboration personnelle à partir des données de l'OCDE*

Contrairement par rapport au graphique précédent, on peut observer, toujours sur la même période, quatre phases d'évolutions majeures.

De fait, nous pouvons observer dans un premier temps une tendance baissière de un an marqué par un creux à 170 % qui pourrait signifier la fin des retombés de la crise des subprimes qui avait eu lieu en 2008.

Par la suite, nous avons une tendance haussière qui va de 170% à plus de 180%. C'est une phase qui s'expliquerait notamment par des évènements majeurs comme l'adoption de l'American invente Act<sup>17</sup> (AIA) qui avait eu lieu en septembre 2011 pour le droit des brevets ou encore par le fait qu'il y avait eu une forte croissance dans le domaine technologique ( notamment en raison de la concurrence entre Apple , Samsung ou encore Huawei sur la commercialisation des smartphones et des tablettes qui on engendrés une forte croissance des flux de brevet) (Bardy, 2014)

Qui plus est, à la suite de cette seconde phase, nous avons plutôt une tendance décroissante de 2013 à 2022, faisant ainsi passer la productivité des brevets de 180% à un creux sans précédent de 100%.

Cela est la conséquence de multiples facteurs tant conjoncturel que structurelles.

D'une part, elle peut etre du en raison de la guerre commerciale entre les Etats-Unis et la Chine qui a débuté en 2017 qui a été un goulot d'étranglement de l'accès à certains composants électronique et qui, par extension, a occasionné un ralentissement dans la recherche et le développement par soucis de ressources disponible. (Lavignère, 2025)

Par ailleurs, d'une autre part, sans compter la crise du Covid-19, elle peut également etre dû à cause de l'inflation sans précédents des taux de la Fed. (Powell, 2022)

Lorsque cela arrive, la Fed à réagir par une hausse des taux directeur. Par ce constat, les banques augmentent leurs taux d'intérêts et octroient moins de crédits pour les agents économiques. De fait, les ménages consomment et empruntent moins, les entreprises investissement et recrutent moins et les investisseurs, voyant la « mauvaise » santé des entreprises, vont moins etre enclin

---

<sup>17</sup> Pour plus d'informations: [H.R.1249 - 112th Congress \(2011-2012\): Leahy-Smith America Invents Act | Congress.gov | Library of Congress](https://www.congress.gov/112th-congress/bill-status/H.R.1249)

à investir et à acheter des actifs risqués. (Voir annexe pour le graphique de l'évolution du capital risque)

Par extension, l'activité économique est au ralentit, la demande globale diminue et donc la productivité des brevets diminue.

En outre, à la suite de cette phase descendante, on retrouve une tendance plus ou moins stable.

Cela signifie donc que d'un côté les entreprises investissent et embauchent plus car grâce aux ménages qui consomment et empruntent plus, leur profit augmente et d'un autre coté les actionnaires investissent plus car il retrouve confiance dans la capacité de générer des bénéfices dans les entreprises.

Par ailleurs, la valeur de la monnaie est plus ou moins stable et cela permet donc de dégager des financements dans la recherche et le développement, et par extension, la capacité d'innovation grâce aux flux de brevets.

### ***Intensité de la Dépense intérieur en R&D par rapport au PIB***

Dans le but de pouvoir effectuer des analyses sur l'intensité de la dépense intérieur en recherche et développement par rapport à la taille de l'économie nationale, nous nous sommes servi de la formule suivante :

$$\text{Ratio DIRD par rapport au PIB} = \frac{\text{DIRD (en million de dollars courants)}}{\text{PIB (en million de dollars courants)}}$$

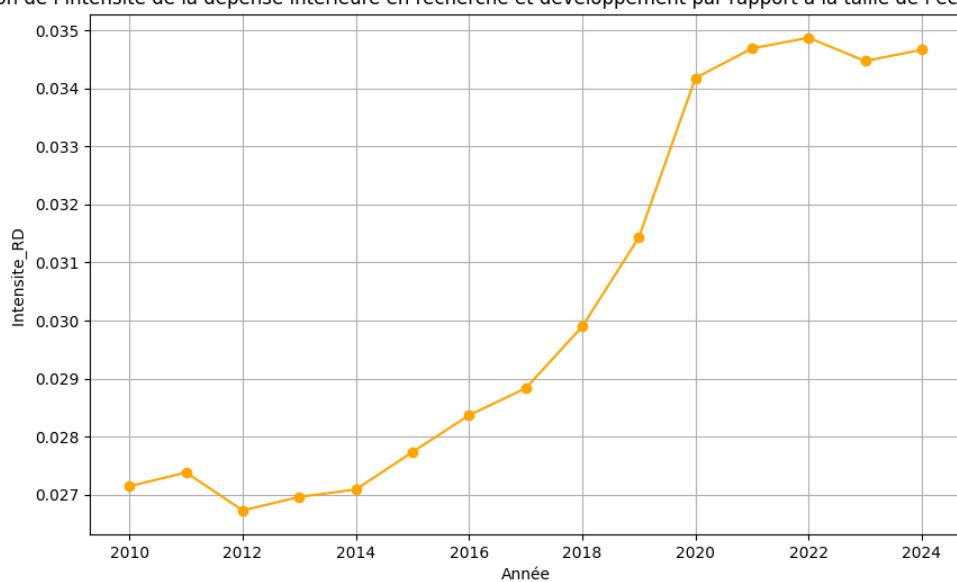
Tableau 3: Intensité de la Dépense intérieur en R&D (DIRD) par rapport au PIB

	DIRD ( en millions de dollars courants)	PIB(en millions de dollars courants)	Ratio DIRD par rapport au PIB (en %)
2010	408495.0	15048971.0	2.714
2011	427125.0	15599732.0	2.738
2012	434424.0	16253970.0	2.673
2013	455089.0	16880683.0	2.696
2014	476971.0	17608138.0	2.709
2015	507372.0	18295019.0	2.773
2016	533451.0	18804913.0	2.837
2017	565446.0	19612102.0	2.883
2018	617538.0	20656516.0	2.99
2019	676995.0	21539982.0	3.143
2020	729857.0	21354105.0	3.418
2021	821478.0	23681171.0	3.469
2022	906953.0	26006893.0	3.487
2023	955578.0	27720709.0	3.447
2024	931265.5	26863801.0	3.467

Source : Elaboration personnelle à partir des données de l'OCDE et de la SNF

Figure 6: Représentation graphique de l'évolution du ratio DIRD-PIB de 2010-2024

Évolution de l'intensité de la dépense intérieure en recherche et développement par rapport à la taille de l'économie nationale



*Source : Elaboration personnelle à partir des données de l'OCDE et de la SNF*

Au travers de ce graphique qui représente l'évolution de l'intensité de la dépense intérieur en recherche et développement par rapport à la taille de l'économie nationale, nous pouvons remarquer quatre grandes phases majeures.

Dans un premier temps, on peut apprécier une première phase plutôt stable marqué par un pic d'environ 2.75% et un creux d'environ 2.68% sur la période de 2010 à 2012.

A la suite de cela, on constate une tendance haussière progressive puis assez marqué pour passer de 2.6% à près de 3.41% (soit un coefficient multiplicateur de près de 1.31<sup>18</sup>)

Nous comprenons donc que cela s'expliquerait en partie du fait que l'Etat s'est assuré de pourvoir renforcer leurs dépenses intérieures pour la recherche et le développement au fil de cette période.

Qui plus est, nous avons ensuite une phase où l'intensité de la dépense intérieure en recherche et développement par rapport à la taille de l'économie nationale ralenti pour ensuite plus ou moins se stabiliser de 2020 à 2024. On a donc une troisième phase qui est établit de 2020 à 2022 puis la dernière de 2022 à 2024.

C'est indubitablement en partie à cause de la crise conjoncturelle du covid-19 et de la guerre commerciale qu'il y a eu un ralentissement de l'intensité de la dépense intérieur en recherche et développement par rapport à la taille de l'économie nationale

### ***Soutien publique total à la R&D des Entreprises***

Afin de pouvoir effectuer des analyses du soutien publique total à la R&D des entreprises, nous nous sommes servis de la formule suivante :

$$\text{Soutien publique à la R&D des entreprises} =$$

$$CIR(\text{en million}) + \text{Crédit budgétaire publique (en million)}$$

---

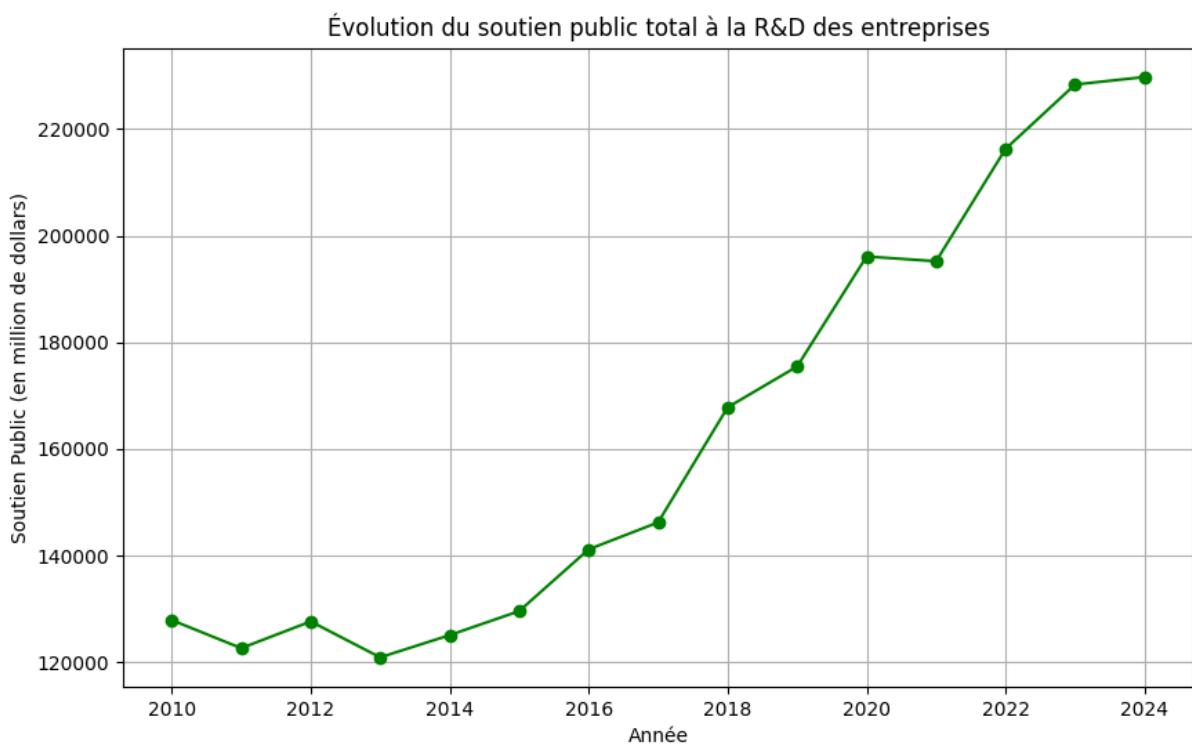
<sup>18</sup>  $3.14\% / 2.6\% \approx 1.3115$

Tableau 4: Soutien public total à la R&D des entreprises

	CIR ( en million )	Crédit Budgétaire publique (en millions)	Soutien public à la R&D des entreprises
2010	8510.7	119382.0	127892.7
2011	9240.5	113386.0	122626.5
2012	10842.6	116796.0	127638.6
2013	11294.4	109608.0	120902.4
2014	12567.6	112502.0	125069.6
2015	14356.8	115220.0	129576.8
2016	15068.1	126093.0	141161.1
2017	18915.8	127306.0	146221.8
2018	23335.5	144459.0	167794.5
2019	25485.3	149971.0	175456.3
2020	26189.7	169901.0	196090.7
2021	33317.6	161885.0	195202.6
2022	29401.45	186876.0	216277.45
2023	29401.45	198941.0	228342.45
2024	25837.5	203919.0	229756.5

Source : Elaboration personnelle à partir des données de l'OCDE et de la SNF

Figure 7 : Représentation graphique de l'évolution du soutien public total à la R&D des entreprises de 2010-2024



*Source : Elaboration personnelle à partir des données de l'OCDE et de la SNF*

Au travers de ce graphique représenté par l'évolution du soutien public total à la R&D des entreprises, nous pouvons observer deux grandes phases majeures.

Nous avons donc une première phase plus ou moins stable et qui oscille sur un minimum de 120 Milliard de dollars et un pic d'environ 130 Milliard de dollars de 2010 à 2013.

Puis, par la suite de cette première phase, nous en avons une seconde marquée par une tendance croissante jusqu'en 2024 allant de 2013 avec 120 Milliard de dollars à plus de 225 Milliard de dollars, soit près 1.87<sup>19</sup> fois la situation initiale.

En effet, cela pourrait s'expliquer par le fait que l'Etat Américain ait mis en place des politiques publiques favorables à la R&D pour les entreprises.

De fait, par extension, cela agit comme un cercle vertueux car l'Etat facilite le développement de l'innovation et cela incite implicitement les entreprises à vouloir en faire plus pour développer leurs parts en recherche et développement et ainsi gagner des parts dans cette lutte de concurrence acharnée sur le marché.

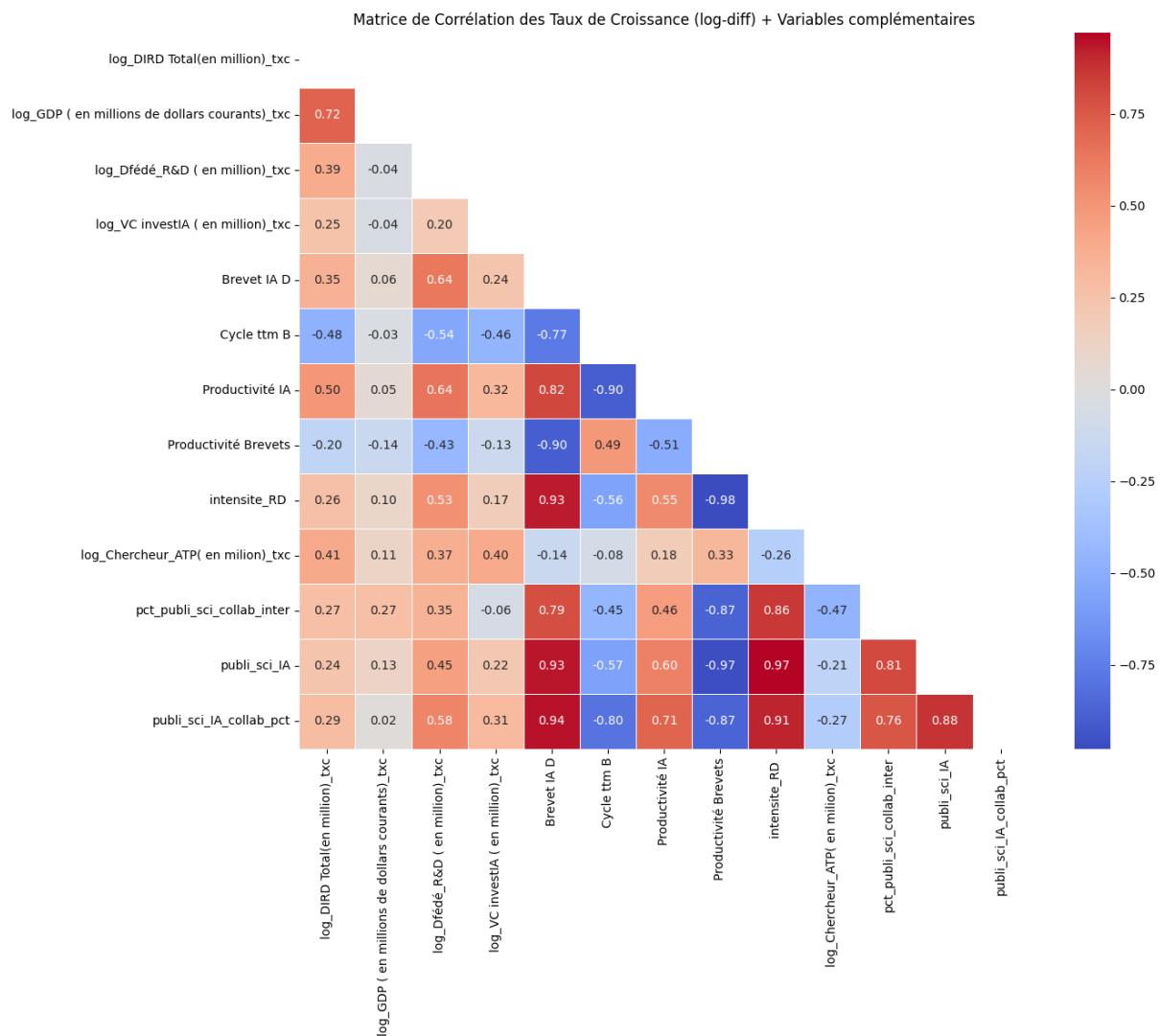
En d'autres termes, les entreprises injectent plus pour pouvoir obtenir plus, et à la fin, l'état fini gagnant dans l'affaire et reprend de nouveau le mécanisme.

---

<sup>19</sup> 120 Milliard/225 Milliard = 1.875

## Matrice de corrélation

Figure 8: MATRICE DE CORRÉLATION



Source : Elaboration personnelle à partir des données de toute les sources de variables quantitatives ( Voir Partie 3)

La matrice de corrélation qui met en exergue l'association entre les variables de notre jeu de données, montre une liaison modérée entre les différentes variables, ainsi on peut par exemple voir que la forme logarithmique des dépenses intérieures en R&D (DIRD total) qui affiche une corrélation négative de -0,48 avec le Cycle de time to Market des brevets.

On comprend donc qu'une augmentation de 1 % des dépenses intérieures de R&D s'accompagne, d'une réduction de 0,48 % du délai de mise sur le marché des brevets.

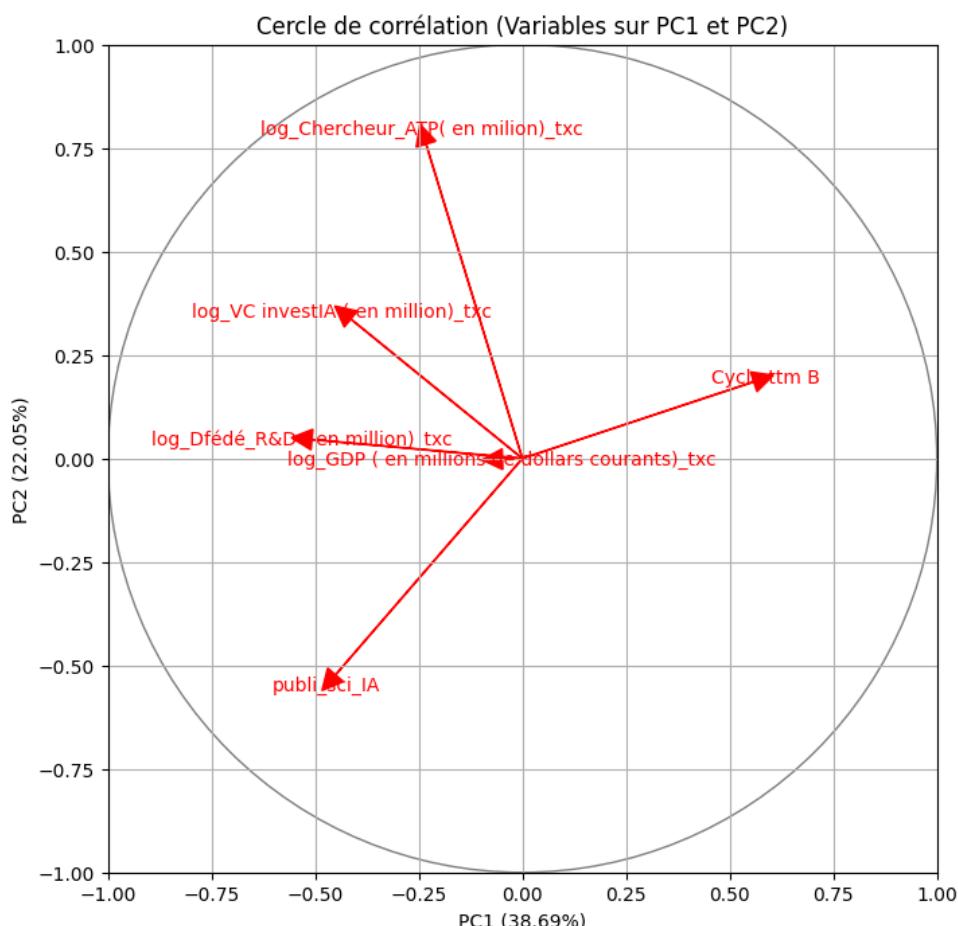
Par ailleurs, cette même variable est positivement corrélée à 0,50 avec la productivité IA, indiquant qu'un effort public de R&D plus soutenu renforce directement l'efficacité de la production IA.

D'autre part, l'investissement en capital-risque IA ( $\log_{10}$  VC investIA) se distingue par une corrélation négative de -0,54 avec le cycle de time to Market et une corrélation positive de 0,64 avec les publications scientifiques sur l'IA impliquant une collaboration internationale. Autrement dit, un accroissement de 1 % des apports de du capital risques dédiés à l'IA permet de raccourcir de 0,54 % le cycle time-to-Market tout en augmentant de 0,64 % la productivité IA. Ce levier privé présente également une corrélation positive de 0,24 avec les brevets IA, confirmant son rôle majeur dans l'innovation.

En synthèse, ces résultats montrent que les financements publics (DIRD total) et les investissements en capital risque associé à l'IA (VC IA) constituent deux leviers complémentaires pour accélérer la commercialisation des brevets et améliorer la productivité IA. Ils confortent la nécessité, lorsqu'il faut établir une décision , de surveiller et d'ajuster en continu ces deux types d'investissement afin de maximiser l'impact économique et technologique des politiques R&D aux États-Unis.

## Analyse en composante principale (ACP)

Figure 9: Cercle de corrélation (PC1 et PC2)



Source : Elaboration personnelle tiré des données des différentes sources de la partie 3 (à voir)

Sur ce cercle de corrélation d'analyse en composante principale des variables expliquant les décisions en R&D aux Etats-Unis, nous avons deux dimensions qui captent 60,74 % de la variance totale (soit 38,69 % pour la dimension 1 et 22,05 % pour la dimension 2).

De fait, nous pouvons voir que les cycles de time to Market des brevets et les chercheurs à temps plein sont corrélés négativement par rapport aux publications scientifiques sur l'IA impliquant une collaboration internationale, les dépenses fédérales en R&D et les investissements en capital risque associé à l'IA.

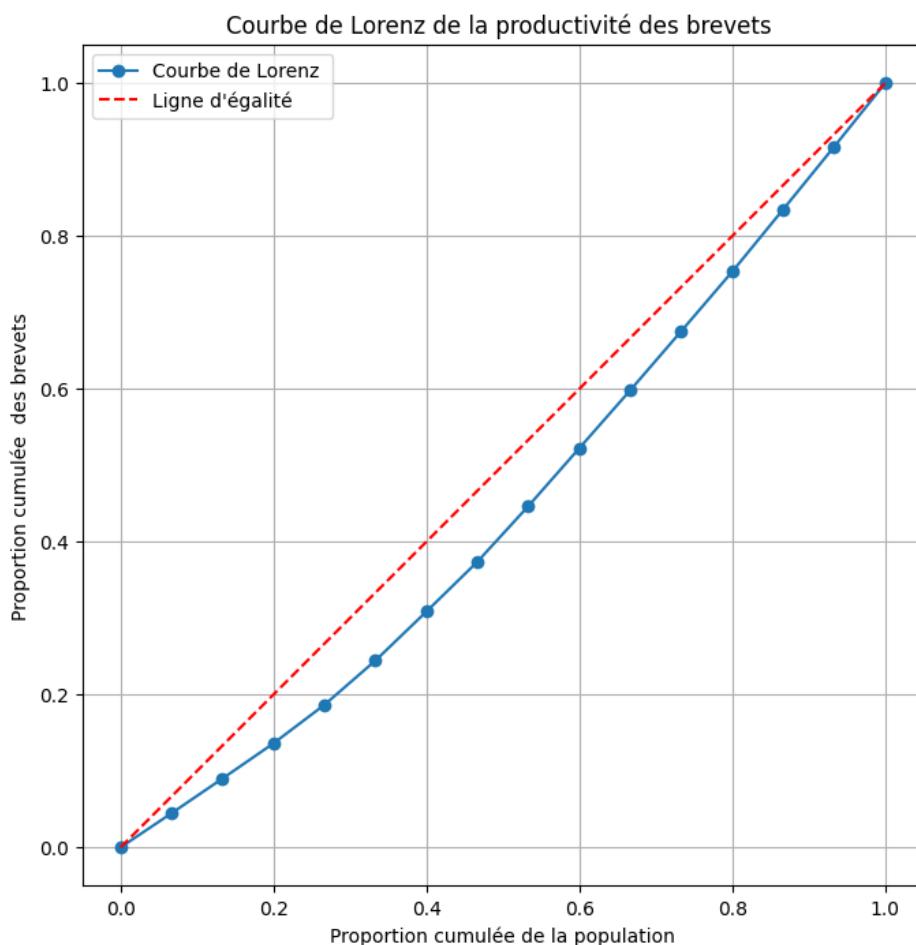
Autrement dit, plus les efforts scientifiques sur les publications en rapport avec l'IA et qui engendre une collaboration internationale, l'investissement en capital risque associé à l'IA et les dépenses fédérales en R&D augmentent, plus le temps de mise sur le marché est réduit, et inversement.

### *Analyses de concentration des brevets*

#### La productivité des brevets

*Courbe de Lorenz de la productivité des brevets*

*Figure 10: Courbe de Lorenz de la productivité des brevets*



Au travers de cette courbe de Lorenz avec en abscisse la proportion cumulée des entreprises et en ordonnée la part de la productivité des brevets et par rapport à la droite d'équipartition , nous pouvons constaté une inégalité de répartition plus ou moins modéré entre les entreprises les plus pauvres et les meilleures entreprises car en effet, 20% des entreprises les plus pauvres ne

captent moins qu'environ 18% de la productivité des brevets, alors que près de 80% des meilleures entreprises profite de près de 87% de cette productivité.

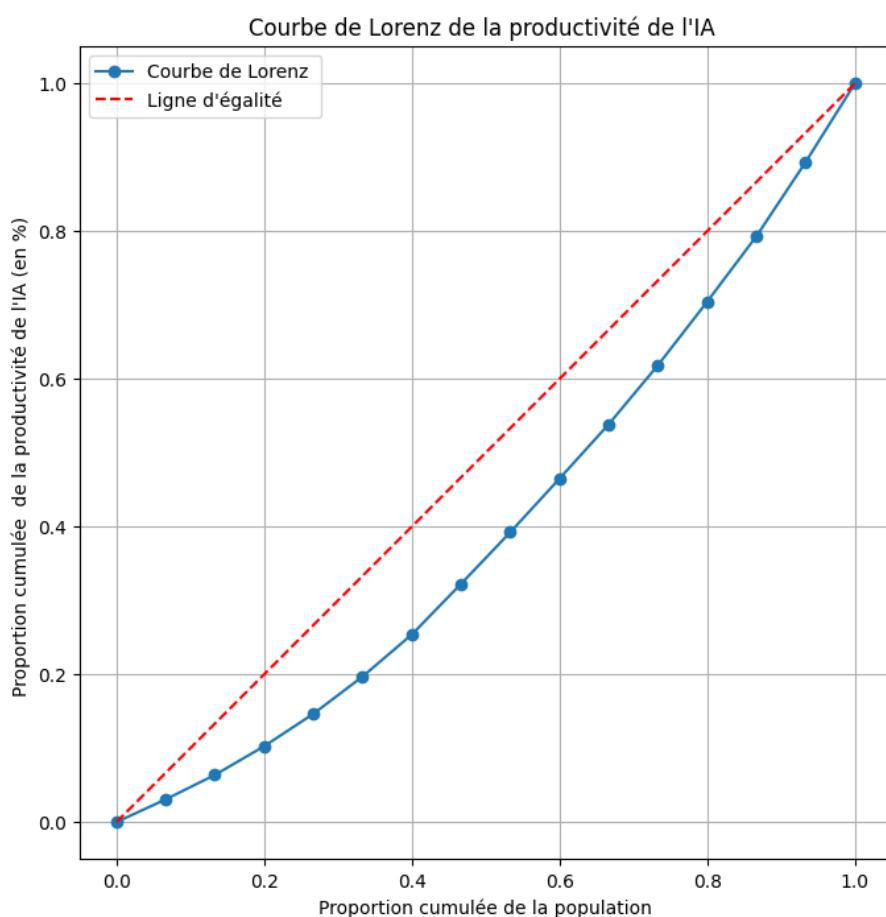
#### *Indice de Gini de la productivité de l'IA*

L'indice de Gini établit à 11.56% nous indique que l'inégalité qui existe entre ces deux groupes est plutôt faible. Nous pouvons donc dire qu'il y a une plutôt bonne répartition des brevets entre les entreprises les plus pauvres et les meilleures entreprises.

#### *La productivité de l'IA*

##### *Courbe de Lorenz de la productivité de l'IA*

Figure 11: Courbe de Lorenz de la productivité de l'IA



Source : *Elaboration personnelle à partir des données de l'OCDE*

Au travers de cette courbe de Lorenz avec en abscisse la proportion cumulée des entreprises et en ordonnée la part de la productivité de l'IA et par rapport à la droite d'équipartition , nous pouvons constaté une inégalité de répartition modéré entre les entreprises les plus pauvres et

les meilleures entreprises car en effet, 20% des entreprises les plus pauvres ne captent qu'environ 15% de la productivité de l'IA , alors que près de 80% des meilleures entreprises jouissent de 85% de cette productivité.

Cet enjeu marqué nous montre à quel point la diffusion de l'innovation avec de l'IA est déséquilibré entre les entreprises.

Par ce constat, ça pourrait s'expliquer en partie du fait les grandes entreprises ont plus de ressources et par conséquent, elles ont plus la capacité d'investir sur des projets pertinents avec un avenir prometteur qui leurs permettront d'obtenir des avantages concurrentiels non négligeable par rapport à leurs concurrents : les projets de R&D liées à l'IA.

Par extension, après avoir investie dans ce secteur, bien qu'ayant un certain niveau de risque et d'incertitudes, si leurs projets de recherche et développement liée à l'IA portent leurs fruits, ils ne peuvent qu'en tirer de meilleure part que ceux qui ont moins de capital et qui sont plus préoccupé par d'autres facteurs externes à cela.

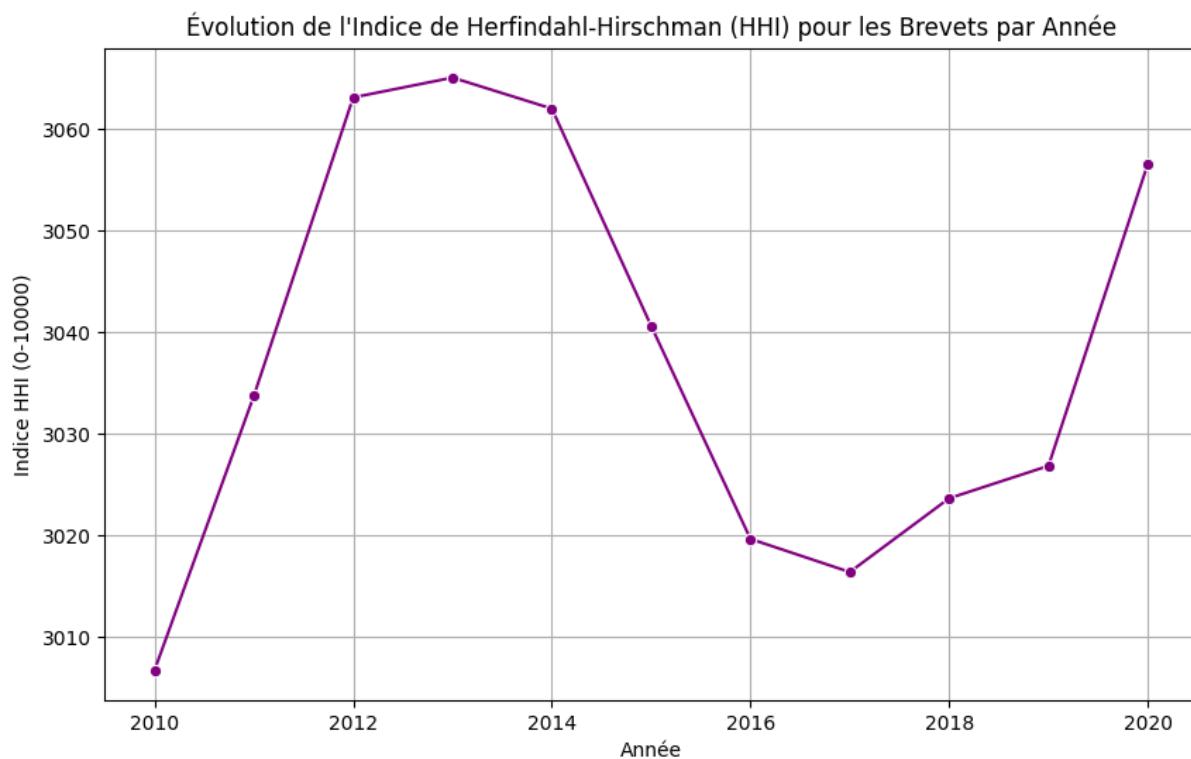
#### *Indice de Gini de la productivité de l'IA*

L'indice de Gini marqué par un indice de 19.69% confirme ce qui avait été dit tout au long de l'analyse sur la courbe de Lorenz car on constate une inégalité modérée qui nous amène donc à dire qu'il y a une concentration partielle de la productivité des brevets qui sont lié avec l'IA.

#### *L'indice de Herfindahl-Hirschman (HHI) des Brevets*

Après avoir appliqué la formule de l'indice de Herfindahl-Hirschman associé aux brevets, voici ce que nous avons pu obtenir :

Figure 12: Représentation graphique du HHI des brevets de 2010-2020



Sources : Elaboration personnelle tirée des données de l'OCDE

Sur cette représentation graphique de l'évolution du HHI pour les brevets par années, nous remarquons dans que le HHI des sont globalement moyennement concentré. Cela validerait en partie la présence de certaines firmes dominant par rapport à d'autres sur le marché concurrentiel sans pour autant avoir le monopole absolu.

#### 4.1.2 Analyse des modèles économétriques

##### Analyse du modèle OLS avec la DIRD

Tableau 5: Modèle OLS avec la DIRD

OLS Regression Results						
Dep. Variable:	log_DIRD_Total_txc	R-squared:	0.676			
Model:	OLS	Adj. R-squared:	0.579			
Method:	Least Squares	F-statistic:	18.05			
Date:	Thu, 12 Jun 2025	Prob (F-statistic):	0.000232			
Time:	11:13:56	Log-Likelihood:	-29.409			
No. Observations:	14	AIC:	66.82			
Df Residuals:	10	BIC:	69.37			
Df Model:	3					
Covariance Type:	HAC					
coef	std err	t	P> t	[0.025	0.975]	
Intercept	-0.5344	1.874	-0.285	0.781	-4.710	3.641
log_GDP_txc	0.6791	0.200	3.396	0.007	0.234	1.125
publi_sci_IA	0.0010	0.001	1.256	0.238	-0.001	0.003
logChercheur_ATP_txc	0.3530	0.161	2.192	0.053	-0.006	0.712
Omnibus:	1.395	Durbin-Watson:	1.605			
Prob(Omnibus):	0.498	Jarque-Bera (JB):	0.891			
Skew:	0.265	Prob(JB):	0.640			
Kurtosis:	1.884	Cond. No.		1.15e+04		

Notes:  
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 1 lags and without small sample correction  
[2] The condition number is large, 1.15e+04. This might indicate that there are strong multicollinearity or other numerical problems.

Source : *Elaboration personnelle à partir des données des sources énoncé dans la partie 3*

Ce modèle de régression linéaire multiple est représenté par les dépenses intérieures en recherche et développement comme variable dépendante et par le PIB Américain, les publications scientifiques sur l'IA impliquant une collaboration internationale et les chercheurs à temps plein comme variables explicatives.

De fait, nous avons donc un modèle sous la forme :

$$\text{Log}(DIRD}_{Total_{txc}}) = \beta_0 + \beta_1 \text{Log}(GDP_{txc}) + \beta_2 \text{publi}_{sci_{IA}} + \beta_3 \text{Log}(Chercheur}_{ATP_{txc}}) + \varepsilon$$

Sur ce, au regard des dépenses intérieur de recherche et développement et par le test de Fisher, l'ensemble du modèle est significatif au seuil de 1% et en regardant le coefficient de détermination ajusté ( $R^2$ ), les variables explicatives sélectionnées pour ce modèle expliquent 57.9% des dépenses internes en recherches et développement ; ce qui signifie que près de 42.1% du modèle n'est pas expliqué.

De plus, fidèlement au test de student, nous remarquons que mise à part les publications scientifiques sur l'IA impliquant une collaboration internationale, nous avons une significativité au seuil de 1% pour le PIB et une significativité au seuil de 10% pour les chercheurs à temps plein.

Qui plus est, nous avons une corrélation positive entre ces deux variables explicatives par rapport à la variable dépendante qui nous amène à dire qu'une augmentation d'1% du PIB va engendrer un accroissement de 0.68 % les dépenses internes en recherches et développement, et un accroissement d'1% de chercheurs à temps plein va être favorable à une augmentation de 0.35 % des dépenses internes en recherches et développement.

En bref, contrairement aux publications scientifiques sur l'IA impliquant une collaboration internationale qui ne semblent pas être un déterminant pour les dépenses internes en recherches et développement, le nombre de chercheurs à temps plein et surtout le PIB le sont ; et cela démontre également par extension que les DIRD sont en partie intimement liés aux cycles économiques.

#### Analyse du modèle SARIMAX avec la DIRD

Tableau 6: Modèle SARIMAX avec la DIRD

```

SARIMAX Results
=====
Dep. Variable: log_DIRD_Total_txc No. Observations: 14
Model: SARIMAX Log Likelihood -33.048
Date: Thu, 12 Jun 2025 AIC 70.095
Time: 10:43:02 BIC 71.225
Sample: 0 HQIC 69.863
- 14
Covariance Type: opg
=====
            coef    std err      z   P>|z|    [0.025    0.975]
-----
log_GDP_txc  1.1609    0.283    4.109    0.000    0.607    1.715
sigma2       9.4535    2.631    3.593    0.000    4.296   14.611
-----
Ljung-Box (L1) (Q): 1.60 Jarque-Bera (JB): 2.99
Prob(Q):          0.21 Prob(JB):        0.22
Heteroskedasticity (H): 0.55 Skew:           1.03
Prob(H) (two-sided): 0.58 Kurtosis:        4.11
=====

```

Warnings:

[1] Covariance matrix calculated using the outer product of gradients (complex-step).

*Source : Elaboration personnelle à partir des données des sources énoncé dans la partie 3*

Ce modèle SARIMAX représenté par les dépenses internes en recherches et développement en fonction du PIB Américain nous permet de comprendre qu'au regard des critères AIC et BIC, notre modèle est significatif.

De plus, en penchant notre regard sur le test de student, le PIB américain est très significatif au seuil de 1%.

De fait nous avons une corrélation positive qui nous indique que lorsque le PIB américain augmente de 1%, nous observons une augmentation de 1.16 points de pourcentages des dépenses intérieurs en recherches et développement.

En bref, au regard du modèle SARIMAX, une bonne condition économique est donc un facteur majeur pour les dépenses en R&D.

## Analyse du modèle OLS par rapport cycle de time-to-Market

Tableau 7: Modèle OLS du cycle time-to-Market

OLS Regression Results						
Dep. Variable:	log_Cycle_ttm_B_txc	R-squared:	0.753			
Model:	OLS	Adj. R-squared:	0.708			
Method:	Least Squares	F-statistic:	30.90			
Date:	Thu, 12 Jun 2025	Prob (F-statistic):	3.06e-05			
Time:	11:23:41	Log-Likelihood:	-30.898			
No. Observations:	14	AIC:	67.80			
Df Residuals:	11	BIC:	69.71			
Df Model:	2					
Covariance Type:	HAC					
coef	std err	t	P> t	[0.025	0.975]	
Intercept	-15.6670	1.960	-7.993	0.000	-19.981	-11.353
log_GDP_txc	0.4014	0.122	3.290	0.007	0.133	0.670
publi_sci_IA	0.0042	0.001	7.838	0.000	0.003	0.005
Omnibus:	3.394	Durbin-Watson:	2.773			
Prob(Omnibus):	0.183	Jarque-Bera (JB):	1.619			
Skew:	-0.824	Prob(JB):	0.445			
Kurtosis:	3.239	Cond. No.	1.10e+04			

Notes:  
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 1 lags and without small sample correction  
[2] The condition number is large, 1.1e+04. This might indicate that there are strong multicollinearity or other numerical problems.

Source : Elaboration personnelle à partir des données des sources énoncé dans la partie 3

Sur ce modèle est représenté les cycles de time to Market comme variable dépendante et le PIB Américain avec les publications scientifiques sur l'IA impliquant une collaboration internationale comme variable explicatives.

Nous avons donc comme expression du modèle :

$$\text{Log}(Cycle_{ttm\_B}) = \beta_0 + \beta_1 \text{Log}(GDP_{txc}) + \beta_2 \text{Log}(publi\_sci\_IA) + \varepsilon$$

En effet, nous pouvons donc constater que notre modèle est globalement significatif au regard du test de Fisher et que les variables explicatives expliquent à 70.8% la période entre le fait d'avoir l'idée d'un produit et sa mise en commercialisation.

De plus, dans la conformité du test de student, on peut apprécier une significativité de 1% pour toutes les variables explicatives de ce modèle.

Par ailleurs, nous avons deux corrélations positives avec d'un coté une augmentation de 1 % du PIB Américain qui engendrera une croissance de 0.4% supplémentaires de la période entre le fait d'avoir l'idée d'un produit et sa mise en commercialisation ; et de l'autre côté, une augmentation d'1% des publications scientifiques sur l'IA impliquant une collaboration

internationale qui engendrera une appréciation de 0.42% de la période entre le fait d'avoir l'idée d'un produit et sa mise en commercialisation.

En bref, le PIB et les publications scientifiques sur l'IA impliquant une collaboration internationale expliquent fortement la période entre le fait d'avoir l'idée d'un produit et sa mise en commercialisation.

#### Analyse du modèle OLS de la productivité des Brevets

Tableau 8: Modèle OLS de la productivité des brevets

OLS Regression Results							
Dep. Variable:	Productivite_Brevets_pct	R-squared:	0.968	Model:	OLS	Adj. R-squared:	0.954
Method:	Least Squares	F-statistic:	370.9	Date:	Thu, 12 Jun 2025	Prob (F-statistic):	5.62e-10
Date:	Thu, 12 Jun 2025	Time:	10:08:08	Log-Likelihood:	-43.978	AIC:	97.96
No. Observations:	14 <th>Df Residuals:</th> <td>9</td> <th>BIC:</th> <td>101.2</td> <th>Df Model:</th> <td>4</td>	Df Residuals:	9	BIC:	101.2	Df Model:	4
Covariance Type:	HAC						
	coef	std err	t	P> t	[0.025	0.975]	
Intercept	236.5445	5.049	46.847	0.000	225.122	247.967	
log_Dfede_RD_txc	-0.6264	0.191	-3.287	0.009	-1.058	-0.195	
publi_sci_IA	-0.0357	0.002	-21.080	0.000	-0.039	-0.032	
Productivite_IA	5.7745	2.680	2.154	0.060	-0.289	11.838	
logChercheur_ATP_txc	1.4217	0.698	2.038	0.072	-0.157	3.000	
Omnibus:	3.112	Durbin-Watson:		2.487			
Prob(Omnibus):	0.211	Jarque-Bera (JB):		1.924			
Skew:	-0.901	Prob(JB):		0.382			
Kurtosis:	2.784	Cond. No.		1.55e+04			

Notes:

[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 1 lags and without small sample correction  
[2] The condition number is large, 1.55e+04. This might indicate that there are strong multicollinearity or other numerical problems.

Source : Elaboration personnelle à partir des données des sources énoncé dans la partie 3

Au travers de ce troisième modèle nous avons plutôt la productivité des brevets en tant que variable dépendante par rapport aux dépenses fédérales en recherches et développements, les publications scientifiques sur l'IA impliquant une collaboration internationale, le nombre de chercheurs à temps plein et la productivité de l'IA en tant que variables explicatives.

Nous retrouvons donc l'expression suivante :

$$Productivite_{Brevets_pct} =$$

$$\beta_0 + \beta_1 \text{Log}(Dfede_(RD_txc)) + \beta_2 \text{publi}_(\text{sci\_IA}) + \beta_3 \text{Productivite\_IA} + \epsilon$$

En effet, nous avons modèle non seulement globalement très significatif mais nous observons également au travers du coefficient de détermination ajusté que les variables explicatives expliquent 95.4% du modèle.

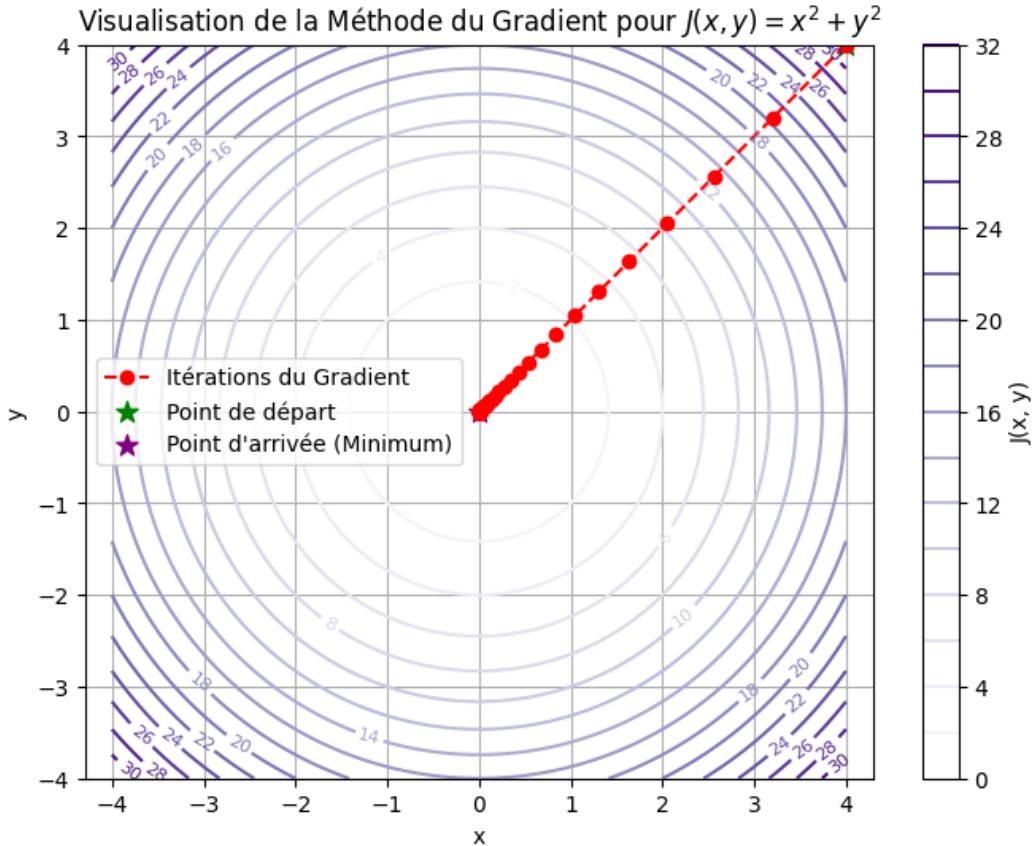
Qui plus est, au regard du test de student, nous observons que les dépenses fédérales pour la R&D et les publications scientifiques sont significatif au seuil de risque de 1% et que la productivité de l'IA et le nombre de chercheurs à temps plein sont plutôt significatif au seuil de 10%.

Par ailleurs, nous avons perspectives de corrélation. Nous avons donc d'un côté une corrélation négative qui nous amène à dire qu'une augmentation d'1% des dépenses fédérales pour la R&D et des publications scientifiques sur l'IA impliquant une collaboration internationale occasionnent une diminution de 0.62% et de 0.35% de la productivité des brevets respectivement ; et d'un autre coté , nous avons une corrélation plutôt positive qui nous conduit au fait qu'un accroissement d'un 1% de la productivité lié à l'IA et le nombre de chercheur à temps plein vont induire une augmentation de 5.77 points de pourcentages et de 1.42 points de pourcentage sur la productivité des brevets.

En bref, il convient donc que les dépenses fédérales pour la R&D et les publications scientifiques sur l'IA impliquant une collaboration internationale ont une influence négative sur la productivité des brevets contrairement au nombre de chercheurs à temps plein et la productivité de l'IA qui soutiennent plus une influence positive mais à un niveau de significativité plus modéré.

#### 4.1.3 Gradient du modèle de cycle de time to Market

Figure 13 : Visualisation de la méthode du gradient du model de cycle time to Market



Source : élaboration personnelle tiré des données de la partie 3

En effet, la visualisation ci-dessus illustre à merveille le fonctionnement de la descente de gradient appliquée à la fonction parfaitement convexe  $J(x, y) = x^2 + y^2$ .

Nous partons d'un point initial situé en (4,4) et matérialisé par l'étoile vert. Dès la première itération, à chaque mise à jour (les points rouges successifs), la valeur de la fonction décroît, les courbes de niveau se resserrent et la norme du gradient se réduit, jusqu'à ce que la trajectoire vienne atteindre le point d'arrivée, c'est-à-dire au minimum global en (0,0), repéré par l'étoile violette. Autrement dit, plus nous avançons, plus les corrections deviennent infinitésimales : le gradient s'éteint graduellement et la convergence est acquise sans aucune oscillation.

En bref, ce schéma confirme qu'un algorithme fondé sur le gradient peut ajuster rapidement et de façon fiable un ensemble de paramètres, dès lors que la surface d'erreur est bien comportée. Dans notre cas, cela signifie qu'à chaque décision, il peut être possible de réestimer, presque

qu'en temps réelle l'effet du PIB ou de l'intensité de publication scientifique sur le cycle time-to-Market des brevets, garantissant ainsi une allocation de ressources R & D toujours alignée sur la trajectoire optimale.

*Figure 14: Descente de gradient sur le modèle de cycle de time to Market*

```
==== Descente de gradient sur model_ttm (approx. OLS) ====
Coefficients (Gradient) : [0.08463413 0.17589951 0.03251331]
Itérations nécessaires : 5000
Norme finale du gradient: 1.319143e-02
Valeur finale J(beta) : 1.335517e-01

Comparaison des coefficients :
const : Gradient = 0.084634 | OLS = 1.410000
log_GDP_txc : Gradient = 0.175900 | OLS = 0.052506
publi_sci_IA : Gradient = 0.032513 | OLS = 0.018547
```

En effet, En effet, notre application de la méthode de descente de gradient sur le modèle

$$\text{Log}(Cycle\_ttm\_B) = \beta_0 + \beta_1 \text{Log}(GDP\_txc) + \beta_2 \text{Log}(publi\_sci\_IA) + \epsilon$$

a permis d'estimer trois coefficients : une constante à environ 0,0846, une élasticité du PIB à 0,1759 et un effet des publications IA à 0,0325. L'algorithme a exécuté la totalité des 5 000 itérations prévues, la norme finale du gradient atteignant une valeur d'environ 0,0132, nous confirme que le processus de convergence est proche de son minimum optimal. De fait, nous avons eu minimum établi à 0,1336.

En bref, l'augmentation du PIB ou du volume des publications scientifiques sur l'IA impliquant une collaboration internationale s'accompagnent d'un allongement du cycle de time-to-Market.

#### 4.1.4 Analyse de l'étude de cas

À l'ère de la révolution industrielle, les entreprises ont commencé à produire en masse, mais à l'ère de l'Intelligence Artificielle, les entreprises ont découvert des mécanismes permettant de travailler plus rapidement et intelligemment. Par exemple, l'IA est capable de prédire, l'efficacité d'un médicament, ses effets secondaires, et gérer la masse de données et documents qui résultent au long du processus du développement produit.

Parmi ces entreprises, Pfizer, une entreprise pharmaceutique américaine, fondée en 1849. En 2013, elle était le leader mondial de son secteur. Dans la suite, nous expliquerons comment intègre-t-elle l'IA dans ses processus de R&D car en effet, ses tentatives sont encore prématurées mais l'entreprise prévoit capitaliser sur l'IA dans les prochaines années.

Un des avantages majeurs de l'IA dans un domaine critique comme la pharmaceutique est la rapidité de mise sur le marché (Go-to-Market) des nouveaux médicaments. Donc, on constate que cela peut évidemment sauver des vies. Mais comment Pfizer fait-elle cela ?

Premièrement, selon l'équipe de Pfizer, l'IA est la base de beaucoup de technologies de pointe dans le domaine de l'Informatique, mais les deux technologies les plus importantes dans ce domaine est la ML (Machine Learning) et NLP (Natural Langage Model)

À chaque stade du développement d'un médicament, il y a des Térabit, voir Pétabit de données qui se génèrent, que ce soit dans des documentations, tableaux ou rapports médicaux. Ces données contiennent généralement des informations ultra-importantes, mais, étant donnée le gigantisme de ces données, les capacités cognitives humaines rencontrent des difficultés à les traiter. C'est là que le ML entre en jeu.

Les logiciels ont des avantages sur l'être humain d'être plus rapide et plus précis. Chez Pfizer, les résultats générés par les machines sont révisés par des experts de la pharmaceutique pour vérifier la fiabilité des résultats et leur précision, car il s'agit d'un domaine critique où l'erreur peut coûter une vie.

Selon le vice-président et le directeur Data du Groupe Pfizer Boris Braylyan, le ML peut désormais prendre en compte les enjeux régulatoires de la production d'un médicament, en identifiant les aspects nécessaires que les régulateurs observent et les anticiper en avance en mettant les réponses dans les soumissions au gouvernement. Cela peut économiser des semaines d'aller-retours avec les régulateurs avant d'entrer sur le marché.

Une des décisions intelligentes d'une entreprise pharmaceutique consiste en choisir quel médicament à ne pas utiliser : il y a des médicaments qui ne sont pas efficaces ou qui peuvent avoir des effets secondaires. Donc, les données peuvent aider à faire une décision rapide sur le potentiel du médicament et permettre une bonne allocation de ressources et d'argent.

#### *Une application concrète de l'IA dans le processus de R&D de Pfizer*

D'après l'équipe de Pfizer, l'IA aide dans la rédaction des documentations, tables, rapports et autres contenus requis pour un nouveau médicament tout au long de son cycle de développement. (IBM and Pfizer to Accelerate Immuno-oncology Research with Watson for Drug Discovery, 2016)

Qui plus est , le cancer est une des causes majeures de décès dans le monde et une des maladies les plus complexes que l'humanité n'a jamais connue et la collaboration entre Watson d'IBM et Pfizer ont permis aux les scientistes de développer des études perspicaces.

De fait, le 1<sup>er</sup> Janvier 2016, Pfizer a annoncé sa collaboration avec IBM pour utiliser son outil « Watson for Drug Discovery » pour aider Pfizer à accélérer la recherche en immuno-oncologie<sup>20</sup>, une approche de traitement du cancer qui consiste à utiliser le système immunitaire du corps pour combattre le cancer.

Par ailleurs, les chercheurs d'oncologie chez Pfizer vont utiliser le « Watson for Drug Discovery » pour analyser de gros volumes dispersés de sources de données (Data), avec licence et disponibles publiquement ainsi que les données disponibles chez Pfizer. Avec ce nouvel outil, les chercheurs chez Pfizer vont pouvoir tester des hypothèses pour générer des compréhensions profondes basées scientifiquement. Cette technologie personnalisée permet aussi de faire des évaluations efficaces et sécurisés.

#### *Qu'est ce que le “Watson for Drug discovery” ?*

C'est le premier outil cognitif basé sur le cloud computing et la machine Learning, le traitement en langage naturel et d'autres technologies de raisonnement cognitif afin d'identifier de nouveaux médicaments, des thérapies combinées pour étude et des stratégies de sélection de patients en immuno-oncologie. (IBM et Pfizer s'allient pour accélérer la recherche immuno-oncologique avec Watson for Drug Discovery, 2016)

L'outil analyse d'immense volume de données, il comprend les questions complexes posés en langage naturel et fournit des réponses fiables sur des bases scientifiques. (IBM and Pfizer to Accelerate Immuno-oncology Research with Watson for Drug Discovery, 2016)

#### *Qu'est ce que l'informatique Cognitif ?*

L'informatique cognitive fait référence aux plateformes technologiques influencées par les sciences cognitives pour simuler le processus de pensée humaine qui englobent l'intelligence

---

<sup>20</sup> L'immunothérapie consiste à modifier le système immunitaire du patient pour pouvoir reconnaître et cibler les cellules du Cancer en utilisant une combinaison de vaccins, immunomodulateurs et des molécules. L'immunothérapie réforme le domaine de l'oncologie.

artificielle et le traitement des signaux. Il peut s'agir de capacités telles que le machine Learning, le raisonnement, le traitement du langage naturel (TLN), la reconnaissance vocale et visuelle, l'interaction homme-machine (IHM), notamment. (Informatique cognitive, s.d.)

## 4.2 Discussions

Après avoir réalisé des analyses diverses et variées, cette partie sera consacrée à répondre aux sous-questions que nous nous sommes posés en introduction.

De fait, nous commencerons par apporter des réponses à la première sous question qui est : Quels secteurs ou types d'entreprises américaines intègrent l'IA dans leur processus de R&D, et comment ?

Ensuite, nous continuerons en expliquant les impacts mesurables de l'IA sur la performance des investissements en R&D dans le contexte américain, et enfin nous apporterons des réponses sur la concentration des brevets IA.

### 4.2.1 Quels secteurs ou types d'entreprises américaines intègrent l'IA dans leur processus de R&D, et comment ?

En effet, au travers de nos divers analyses, nous avons pu constater que En effet, l'étude chronologique de nos jeux de données brevets sur 2010-2020 révèle qu'un noyau restreint de secteurs concentre l'adoption de l'intelligence artificielle. Tout d'abord, les technologies de l'information et de la communication affichent la dynamique la plus marquée : la proportion de dépôts explicitement estampillés « IA » y progresse de moins de 5 % à près de 10 %, signe que software, services cloud et applications de traitement du langage profitent directement des avancées algorithmiques.

Qui plus est, l'industrie pharmaceutique mobilise le raisonnement automatique pour accélérer la découverte de médicaments ; l'exemple phare demeure l'utilisation par Pfizer d'une plate-forme cognitive de type Watson for Drug Discovery, qui permet de cibler en quelques heures des milliers de molécules en immuno-oncologie et, partant, de réduire sensiblement les coûts de sélection des candidats .

Par ailleurs, la santé numérique et la medtech mobilisent la vision par ordinateur ainsi que l'apprentissage profond pour le diagnostic assisté et la médecine personnalisée, si bien que les brevets « technologie médicale » se hissent régulièrement dans le trio de tête des domaines les

plus cités . De leur côté, les technologies environnementales s’emparent à leur tour des modèles prédictifs : gestion d’énergie, optimisation de procédés verts ou prévision climatique, autant de niches où l’empreinte brevetaire ne cesse de progresser sur la décennie.

En bref, ces secteurs se distinguent parce qu’ils réunissent simultanément volumétrie de données, intensité capitalistique et potentiel élevé de retour sur innovation : trois conditions qui permettent à l’IA de démultiplier l’efficacité de la R&D et de constituer, pour les entreprises américaines, un puissant levier de compétitivité technologique.

#### 4.2.2 Quels sont les impacts mesurables de l’IA sur la performance des investissements en R&D dans le contexte américain ?

En effet, le modèle économétrique que nous avons mobilisés est non seulement globalement très significatif mais affiche, au travers de son coefficient de détermination ajusté, une capacité explicative de 57,9% des variations de la dépense intérieure de R&D.

Qui plus est, un accroissement d’1 % du PIB américain induit une hausse de 0,68 % de cette dépense, et une augmentation identique du nombre de chercheurs à temps plein se traduit par +0,35 %, tandis que la productivité IA (brevets IA rapportés à la DIRD) entretient une corrélation positive de +0,50 avec la dépense globale.

Par ailleurs, l’injection de capital-risque dédié à l’IA raccourcit sensiblement le cycle de *time-to-Market*.

Une augmentation d’une unité supplémentaire de l’investissement en capital risque dans le domaine de l’IA réduit en moyenne les délais d’environ 0,54 %, un résultat confirmé par l’analyse en composantes principales où le couple entre le financement IA et les publications scientifiques sur l’IA impliquant une collaboration internationale émerge comme principal moteur de la compression des cycles d’innovation.

En bref, l’IA agit comme un véritable levier : elle augmente le rendement marginal des dollars investis, accélère la valorisation des découvertes et révèle une complémentarité marquée entre soutien public et capitaux privés.

#### 4.2.3 Les recherches associées à l’IA en R&D permet-elle une concentration des brevets qui serait plus privilégié pour les grandes voir très grandes firmes ou, au contraire, serait-elle plus à même d’être généraliser ?

En effet, la lecture conjointe des courbes de Lorenz et des indices d’inégalité laisse apparaître une situation « modérément » concentrée : les 20 % d’entreprises les moins dotées ne captent

qu'environ 15 % de la productivité des brevets IA, contre 85 % pour les 80 % les mieux pourvues.

Qui plus est, l'indice de Gini plafonne à 19,69 % pour l'IA – niveau certes supérieur à celui des brevets totaux (11,56 %) mais encore éloigné d'un oligopole étroit – tandis que le HHI demeure dans la zone « moyennement concentrée » sur l'ensemble de la décennie.

Par ailleurs, l'essor des modèles open-source, les programmes publics SBIR/STTR et la baisse continue des coûts de calcul laissent encore la porte ouverte aux PME innovantes qui souhaitent se positionner.

En bref, même si l'IA renforce indéniablement l'avancé des grands acteurs capables de financer puissance de calcul et talents rares, elle ne verrouille pas totalement l'accès : la diffusion reste partielle mais réelle, et des opportunités subsistent pour les nouveaux entrants chevronnés.

## Partie 5 – Proposition de sujet à traiter

L'analyse que nous avons réalisé tout au long de notre mémoire montre à quel point l'IA influe sur les décisions d'investissements en R&D aux Etats unis. Cependant, cette analyse ne représente qu'une infime partie de ce qu'il en ait de l'optimisation des décisions d'investissement en R&D grâce à l'IA.

De fait, cette section proposera d'explorer d'autres perspectives de ce sujet.

### 5.1 L'optimisation des décisions d'investissements en R&D grâce à l'IA : une analyse comparative entre les Etats- Unis, la Chine et l'Europe

En effet, même si notre modèle états-unien s'avère non seulement globalement très significatif, mais également robuste au regard du coefficient de détermination ajusté, il demeure crucial de l'inscrire dans une logique comparative.

De fait, la Chine et l'Europe mobilisent, chacune à leur manière, des plans d'incitation et de régulation susceptibles de modifier la trajectoire mondiale de la dépense en R-D.

En d'autres termes, confronter ces trois blocs permettrait de mesurer concrètement l'élasticité d'1 % d'investissement IA supplémentaire sur le rendement technologique – valeur qui, rappelons-le, fluctue déjà sensiblement entre secteurs aux États-Unis. En bref, cette analyse

multiniveau offrirait un tableau d'ensemble pour mettre en lumière les convergences que les asymétries structurelles des écosystèmes d'innovation.

### 5.2 L'optimisation des décisions d'investissements en R&D grâce à L'IA : Analyse des répercussion politique entre la Chine et Taiwan

En effet, notre étude a montré que la chaîne de valeur des semi-conducteurs constitue un nœud stratégique dans la course à l'IA.

Qui plus est, la dépendance mondiale à l'égard des fonderies taïwanaises introduit un facteur de vulnérabilité que chaque durcissement des contrôles d'exportation vient accroître.

L'objectif annoncer sera donc plutôt d'évaluer, simultanément, le risque de rupture d'approvisionnement, la capacité d'adaptation des acteurs industriels et l'impact potentiel sur la trajectoire d'apprentissage des modèles IA globaux.

### 5.3 Quel serait l'impact de l'intelligence artificielle générale (IAG) sur l'environnement politique et écologiques

En effet, l'IAG se présente comme la prochaine frontière, capable de reproduire la plupart des tâches humaines. Qui plus est, les infrastructures nécessaires atteignent déjà des ordres de grandeur financiers et énergétiques inédits, à l'image du projet Stargate et de ses centaines de milliards de dollars annoncés.

Mais à quel prix les pays sont-ils prêts à vouloir y arriver et surtout à le faire perdurer sans pour autant infliger d'externalités négatives irréversibles ? Sommes-nous sur le point de causer implicitement notre propre destruction ?

## Partie 6 –Conclusion

En conclusion, à la suite des diverses démarches toute aussi rigoureuses les unes par rapport aux autres pour répondre à notre question centrale qui est : dans quelles mesures et par quels mécanismes l'intelligence artificielle permet-elle d'optimiser les investissements en R&D aux États-Unis, et quels en sont les impacts concrets sur la performance des entreprises et de l'innovation ? Cette partie proposera de faire le point sur les résultats, ensuite sur les apports de notre mémoire, sans compter les limites de notre étude.

### 6. 1 Résumer des résultats obtenus

En effet, notre modèle s'avère non seulement globalement très significatif, mais nous constatons également, au travers du coefficient de détermination ajusté, que les variables explicatives rendent compte de 95,4 % de la variance de la productivité des brevets.

De plus, à la lumière du test de Student, les dépenses fédérales allouées à la R&D et les publications scientifiques sur l'IA sont significatives au seuil de risque de 1 %, tandis que la productivité de l'IA et le nombre de chercheurs à temps plein demeurent significatifs, mais à un seuil plus modéré de 10 % .

Par ailleurs, les perspectives de corrélation font apparaître deux dynamiques. D'un côté, une corrélation négative montre qu'une augmentation d'1 % des dépenses fédérales ou des publications collaboratives internationales se solde, respectivement, par une baisse de 0,62 % et 0,35 % de la productivité des brevets. De l'autre, une corrélation positive indique qu'un accroissement d'1 % de la productivité liée à l'IA et des effectifs chercheurs engendre un gain de 5,77 et 1,42 points de pourcentage sur cette même productivité .

Enfin, l'analyse du gradient et de la descente de gradient appliquée au modèle de cycle time-to-market confirme la robustesse de notre surface d'erreur : itération après itération, la norme du gradient décroît jusqu'à 0,0132, tandis que la fonction-perte converge vers un minimum global de 0,1336, validant ainsi la stabilité des coefficients et la possibilité de ré-estimer ces derniers quasi en temps réel .

En bref, il ressort que les financements publics et la production scientifique exercent un effet contrasté : négatif lorsqu'ils se combinent à une forte collaboration internationale, positif dès lors qu'ils s'accompagnent d'un renforcement du capital humain et d'une montée en puissance de la productivité propre à l'IA .

En d'autres termes, l'intelligence artificielle se positionne comme un levier à double détente dans l'optimisation des investissements américains en R&D.

### 6.2 Les apports de notre mémoire

Ce travail apporte trois contributions majeures. Premièrement, nous avons proposé un cadre méthodologique hybride avec des statistiques descriptives, une analyse en composantes principales, régressions OLS et algorithmes de gradient apte à saisir simultanément la photographie quantitative et la dynamique d'ajustement continu des décisions d'investissement. Deuxièmement, nous introduisons un indicateur original de productivité IA, défini comme le ratio brevets IA / dépense intérieure de R&D ; cet indicateur normalise l'efficacité technologique et permet la comparaison inter-sectorielle.

Troisièmement, nous démontrons la complémentarité opérationnelle entre capitaux publics et capital-risque IA : une augmentation d'1 % du capital risque dédié raccourcit le time-to-market de 0,54 % et accroît la production scientifique IA de 0,64 %, fournissant aux décideurs un signal clair pour calibrer l'allocation des ressources.

### 6.3 Les limites de notre mémoire

En effet, malgré un modèle globalement très significatif, plusieurs contraintes méthodologiques et empiriques viennent modérer la portée de nos conclusions. Premièrement, certaines de nos variables étaient payantes, d'autres lacunaires ou étaient tout simplement absentes ; on a donc été contraints de restreindre l'échantillon temporel et sectoriel, ce cela à , par extension, réduit la finesse de certaines estimations.

Qui plus est, l'hétérogénéité d'échelle entre variables (PIB, DIRD, capital-risque, brevets) a exigé d'intenses opérations de normalisation ; nous avons notamment recouru à l'imputation k-NN pour pallier les vides, puis à des transformations logarithmiques pour stabiliser la variance, démarches qui introduisent un risque de biais de mesure résiduel.

Par ailleurs, si le contrôle du facteur d'inflation de variance (VIF) atténue la multi-colinéarité entre PIB, dépenses publiques et effort privé, il ne le supprime pas complètement; une partie de la colinéarité structurelle subsiste donc dans le jeu de données, et cela peut gonfler l'importance relative de certains coefficients.

De surcroît, notre algorithme de descente de gradient (bien qu'il converge vers un minimum global avec une norme finale de 0,0132 et une fonction-perte de 0,1336) reste tributaire du jeu de paramètres initial et d'une surface d'erreur supposée convexe ; tout changement brusque de régime économique pourrait altérer cette stabilité.

Enfin, le volet qualitatif souffre d'un faible taux de réponse au questionnaire professionnel, limitant la triangulation des résultats et la profondeur des retours managériaux ; certaines perspectives de terrain demeurent ainsi sous-documentées.

# Bibliographie

(s.d.). Récupéré sur Cairn.info: <https://apropos.cairn.info/fr>

(s.d.). Récupéré sur ResearchGate: [https://www.researchgate.net/figure/The-probability-that-R-D-of-integrative-technologies-will-be-highly-successful-as-a\\_fig3\\_237130837#:~:text=](https://www.researchgate.net/figure/The-probability-that-R-D-of-integrative-technologies-will-be-highly-successful-as-a_fig3_237130837#:~:text=)

(s.d.). Récupéré sur Pfizer: <https://www.pfizer.fr/>

*À propos de l'OMPI.* (s.d.). Récupéré sur OMPI: <https://www.wipo.int/about-wipo/fr/>

*À propos de OCDE.AI.* (s.d.). Récupéré sur OECD.AI: <https://oecd.ai/fr/about/about-gpai>

*About JSTOR.* (s.d.). Récupéré sur JSTOR: <https://about.jstor.org/>

*About SBIR and STTR.* (s.d.). Récupéré sur SBA: <https://www.sbir.gov/about>

*About Us.* (s.d.). Récupéré sur USPTO: <https://www.uspto.gov/about-us>

*Accelerating America's Leadership in Artificial Intelligence*. (2019, Fevrier 11). Récupéré sur The White House: <https://trumpwhitehouse.archives.gov/articles/accelerating-americas-leadership-in-artificial-intelligence/>

Akerlof, G. A. (1970, Aout). *The Market for "Lemons": Quality Uncertainty and the Market Mechanism.* doi:<https://doi.org/10.2307/1879431>

*Algorithme trading : définition & fonctionnement.* (2025, Mai). Récupéré sur Business Review Magazine: <https://business-review.fr/algorithme-trading-definition/#:~:text=Encore%20appel%C3%A9%20trading%20algorithmique%2C%20l%E2%80%99algorithme%20trading%20est%20une,rendre%20le%20trading%20plus%20rapide%20et%20plus%20efficace.>

*Apprentissage automatique.* (s.d.). Récupéré sur CNIL: <https://www.cnil.fr/fr/definition/apprentissage-automatique>

Bardy, C. (2014, Février 18). *2013 : l'année où le smartphone s'est imposé.* Récupéré sur LEMAGIT: <https://www.lemagit.fr/actualites/2240214564/2013-lannee-ou-le-smartphone-s'est-impose>

*Base de données pour l'analyse structurelle (STAN).* (s.d.). Récupéré sur OCDE: <https://www.oecd.org/fr/data/datasets/structural-analysis-database.html>

Bergmann, D. (2023, Décembre 05). *Qu'est-ce que l'apprentissage auto-supervisé ?* Récupéré sur IBM: <https://www.ibm.com/fr-fr/think/topics/self-supervised-learning>

Bostrom, N. (2014). *Superintelligence, Paths, Dangers, Strategies.*

Brevet. (2021, Février 09). Récupéré sur Institut national de la statistique et des études économiques: <https://www.insee.fr/fr/metadonnees/definition/c2054>

Canonne, E. (2025, Janvier 11). *Les "Magnificent Seven" : GAFAM, Nvidia et Tesla, ces géants qui redéfinissent l'économie mondiale.* Récupéré sur PlanèteGrrandeEcole: <https://www.planetegrandesecoles.com/magnificent-seven-gafam-nvidia-tesla>

*Capital risque.* (2024, Mai). Récupéré sur BPIfrance: <https://bpifrance-creation.fr/encyclopedie/financements/recours-a-investisseurs/capital-risque>

*Ce que l'on sait de Stargate, le projet à 500 milliards de dollars de Donald Trump dans l'intelligence artificielle.* (2025, Janvier 22). Récupéré sur franceinfo: [https://www.franceinfo.fr/monde/usa/presidentielle/donald-trump/ce-que-l-on-sait-de-stargate-le-projet-a-500-milliards-de-dollars-de-donald-trump-dans-l-intelligence-artificielle\\_7029521.html](https://www.franceinfo.fr/monde/usa/presidentielle/donald-trump/ce-que-l-on-sait-de-stargate-le-projet-a-500-milliards-de-dollars-de-donald-trump-dans-l-intelligence-artificielle_7029521.html)

*Comment transmettre un brevet ?* (2025, Février 13). Récupéré sur Ministère de l'économie , des finances et de la souveraineté industrielle et du numérique: <https://www.economie.gouv.fr/cedef/fiches-pratiques/comment-transmettre-un-brevet>

*deep Learning : qu'est-ce que c'est ?* (s.d.). Récupéré sur Futura: <https://www.futura-sciences.com/tech/definitions/intelligence-artificielle-deep-learning-17262/>

*DeepMind.* (2018, Avril 05). Récupéré sur Tech Target: <https://www.techtarget.com/whatis/definition/DeepMind>

*DeepMind.* (2018, Avril 05). Récupéré sur techtarget: <https://www.techtarget.com/whatis/definition/DeepMind>

Deepseek. (s.d.). *DeepSeek-R1.* Récupéré sur Git-hub: <https://github.com/deepseek-ai/DeepSeek-R1>

*Definitions of Research and Development: An Annotated Compilation of Official Sources.* (2025, Avril 28). Récupéré sur NCSES: <https://ncses.nsf.gov/pubs/ncses25215>

*Dépense intérieure de recherche et développement expérimental / DIRD.* (2021, Janvier 26).

Récupéré sur INSEE: <https://www.insee.fr/fr/metadonnees/definition/c1226>

*Dépenses intérieures de R-D des entreprises (DIRDE).* (s.d.). Récupéré sur UNESCO:

<https://uis.unesco.org/fr/glossary-term/depenses-interieures-de-r-d-des-entreprises-dirde>

Djermane, R., & Benzaim , S. (2016). *Etats-Unis, l'expérience d'un leader mondial des TIC et NTIC.* Faculté des sciences économiques, commerciales et de gestion , Laboratoire Incubateur d'Entreprises & Développement Local , Khenchela .

*FACT SHEET: Two Years after the CHIPS and Science Act, Biden-Harris Administration Celebrates Historic Achievements in Bringing Semiconductor Supply Chains Home, Creating Jobs, Supporting Innovation, and Protecting National Security.* (2024, Aout 09).

Récupéré sur THE WHITE HOUSE:  
<https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2024/08/09/fact-sheet-two-years-after-the-chips-and-science-act-biden-%E2%81%A0harris-administration-celebrates-historic-achievements-in-bringing-semiconductor-supply-chains-home-creat>

Grabowski, H. G., & J, V. (2000). *The distribution of sales revenues from pharmaceutical innovation.* doi: 10.2165/00019053-200018001-00005

Green, V. (2025, Février 14). *Qu'est-ce que DeepSeek ? Comment cette entreprise devient-elle un acteur disruptif de l'IA, défiant tous ses concurrents ?* Récupéré sur Gitmind:  
<https://gitmind.com/fr/quest-ce-que-deepseek-ai.html>

Hall, B. H., & Lerner, J. (2009). *THE FINANCING OF R&D AND INNOVATION.* NBER , 1050 Massachusetts Avenue. Récupéré sur  
[https://www.nber.org/system/files/working\\_papers/w15325/w15325.pdf#:~:text=In%20the%20innovation%20setting%2C%20the,1970](https://www.nber.org/system/files/working_papers/w15325/w15325.pdf#:~:text=In%20the%20innovation%20setting%2C%20the,1970)

*History.* (s.d.). Récupéré sur US National Science Fondation:  
[https://www.nsf.gov/about/history?utm\\_source=chatgpt.com](https://www.nsf.gov/about/history?utm_source=chatgpt.com)

HyounghSunYoo, YeLimJung, & Seung-PyoJun. (2023, Mai). *Prediction of SMEs' R&D.* Récupéré sur ResearchGate:  
[https://www.researchgate.net/publication/370657138\\_Prediction\\_of\\_SMEs'\\_RD\\_performances\\_by\\_machine\\_learning\\_for\\_project\\_selection](https://www.researchgate.net/publication/370657138_Prediction_of_SMEs'_RD_performances_by_machine_learning_for_project_selection)

*IA : le projet Stargate de Trump prend forme.* (2025, Juin 03). Récupéré sur JDN: <https://www.journaldunet.com/intelligence-artificielle/1542045-ia-le-projet-stargate-de-trump-prend-forme/#confirmation>

*IBM and Pfizer to Accelerate Immuno-oncology Research with Watson for Drug Discovery.* (2016, Decembre 01). Récupéré sur Pfizer: [https://www\(pfizer.com/news/press-release/press-release-detail/ibm\\_and\\_pfizer\\_to\\_accelerate\\_immuno\\_oncology\\_research\\_with\\_watson\\_for\\_drug\\_discovery](https://www(pfizer.com/news/press-release/press-release-detail/ibm_and_pfizer_to_accelerate_immuno_oncology_research_with_watson_for_drug_discovery)

*IBM et Pfizer s'allient pour accélérer la recherche immuno-oncologique avec Watson for Drug Discovery.* (2016, Décembre 1). Récupéré sur IBM: <https://fr.newsroom.ibm.com/2016-12-01-IBM-et-Pfizer-sallient-pour-accelerer-la-recherche-immuno-oncologique-avec-Watson-for-Drug-Discovery>

*Indice Herfindahl-Hirschman (HHI).* (s.d.). Récupéré sur investisor'wiki: <https://investors.wiki/fr/hhi>

*Informatique cognitive.* (s.d.). Récupéré sur Cognizant: <https://www.cognizant.com/fr/fr/glossary/cognitive-computing>

*Institut National de la Statistique et des Etudes Economiques.* (2021, 01 27). Récupéré sur insee.fr: <https://www.insee.fr/fr/metadonnees/definition/c1244>

*Intelligence artificielle (IA) : de quoi parle-t-on ?* (2024, Avril 22). Récupéré sur Ministère chargé de l'enseignement supérieur et de la recherche: <https://www.enseignementsup-recherche.gouv.fr/fr/intelligence-artificielle-de-quoi-parle-t-91190>

*Investment capital.* (2024, Mars 29). Récupéré sur SBA: <https://www.sba.gov/funding-programs/investment-capital>

JDN. (2019, Octobre 31). *Time to market : définition et synonyme.* Récupéré sur Journal du Net ( JDN): <https://www.journaldunet.fr/business/dictionnaire-du-marketing/1208279-time-to-market-definition-et-synonyme/>

Jun, H. S.-P. (2023, Mai). *Prediction of SME'S R&D performances by machine learning for project selection.* Récupéré sur ResearchGate: [researchgate.net](https://www.researchgate.net)

K, A. (2025, Mars 5). *Comment fonctionne le copy trading ? Tout ce que vous devez savoir.* Récupéré sur ALGOBOT: <https://algobot.com/fr/how-does-copy-trading-work->

everything-you-need-to-know#:~:text=Le%20copy%20trading%20est%20une%20strat%C3%A9gie%20automatis%C3%A9e%20qui,vend%20ou%20ajuste%20des%20positions%20en%20temps%20r%C3%A9el.

Kassel, R. (2024, Janvier 08). *Modèle SARIMAX : Qu'est-ce que c'est ? Comment l'appliquer aux séries temporelles ?* Récupéré sur DataScientest: <https://datascientest.com/modele-sarimax-tout-savoir#:~:text=Parmi%20les%20diff%C3%A9rentes%20approches%20disponibles%20%C2%A0mod%C3%A8le%20SARIMAX,dans%20l%E2%80%99analyse%20afin%20d%E2%80%99am%C3%A9liorer%20la%20pr%C3%A9cision%20des%20pr%C3%A9dictions>

Kelly, C. W. (2018). *Rise of the Machines : Artificial Intelligence and its Growing Impact on U.S. Policy.* Récupéré sur <https://oversight.house.gov/wp-content/uploads/2018/09/AI-White-Paper-.pdf>

Knight, F. H. (2023, Janvier 01). *Risk, uncertainty and prophet: The psychological insights of Frank H. Knight.* Récupéré sur CAMBRIDGE UNIVERSITY PRESS: <https://www.cambridge.org/core/journals/judgment-and-decision-making/article/risk-uncertainty-and-prophet-the-psychological-insights-of-frank-h-knight/1B8D5AC369EA841ADAE2DB6675864816>

*L'OCDE : Des politiques meilleures pour une vie meilleure.* (s.d.). Récupéré sur OCDE: <https://www.oecd.org/fr/about.html>

*La Banque mondiale : rôle, missions et impact.* (s.d.). Récupéré sur Novethic: <https://www.novethic.fr/lexique/detail/banque-mondiale.html>

*La fonction Recherche & Développement.* (s.d.). Récupéré sur Ministère de l'Economie, des finances et de la souveraineté industrielle et du numérique: <https://www.economie.gouv.fr/facileco/fonction-recherche-developpement?>

Labrouche. (2014). *Théorie des capacités dynamiques: les compétences pour savoir s'adapter*. Récupéré sur SI & Management: <http://www.sietmanagement.fr/modele-des-capacites-dynamiques-d-teece/>

Lavignère, B. d. (2025, Avril 14). *La guerre commerciale sino-américaine et ses implications pour l'Union européenne*. Récupéré sur Portail de l'IE: <https://www.portail-ie.fr/univers/2025/la-guerre-commerciale/>

*Les cinq types de machine learning à connaître.* (2023, Décembre 20). Récupéré sur IBM: <https://www.ibm.com/fr-fr/think/topics/machine-learning-types>

*Les meilleurs incubateurs de start up aux Etats Unis.* (2025, Mai 01). Récupéré sur FasterCapital: <https://fastercapital.com/fr/contenu/Les-meilleurs-incubateurs-de-start-up-aux-Etats-Unis.html#Les-meilleurs-incubateurs-de-start-up-de-la-Silicon-Valley-aux-tats-Unis>

*Markowitz Mean-Variance Optimization Theory: A Comprehensive Exploration in Portfolio Management.* (s.d.). Récupéré sur AccountEnd: <https://accountend.com/markowitz-mean-variance-optimization-theory-a-comprehensive-exploration-in-portfolio-management/>

Mattret, J.-B. (2006). *Les FINANCES DE L'ETAT catégorie C.* (CNFPT, Éd.) Récupéré sur 123dok: [https://123dok.net/article/d%C3%A9pense-publique-types-cr%C3%A9dits-budget%C3%A9taires.zlg1v74r#:~:text=Le%20cr%C3%A9dit%20budget%C3%A9taire%20est%20une%20autorisation%20donn%C3%A9e%20au,d%C3%A9termin%C3%A9%C3%A9%C3%A9%2C%20un%20montant%20donn%C3%A9e%20et%20une%20dur%C3%A9e](https://123dok.net/article/d%C3%A9pense-publique-types-cr%C3%A9dits-budget%C3%A9taires.zlg1v74r#:~:text=Le%20cr%C3%A9dit%20budget%C3%A9taire%20est%20une%20autorisation%20donn%C3%A9e%20au,d%C3%A9termin%C3%A9%C3%A9%2C%20un%20montant%20donn%C3%A9e%20et%20une%20dur%C3%A9e)

Mucci, T. (2024, Juillet 23). *What is data-driven decision-making?* Récupéré sur IBM: <https://www.ibm.com/think/topics/data-driven-decision-making>

NCSES. (2025, 02 27). Récupéré sur ncses.nsf.gouv: <https://ncses.nsf.gov/pubs/nsf25327>

*NVIDIA.* (s.d.). Récupéré sur <https://companiemarketcap.com/fr/nvidia/benefices/>

*NVIDIA Corporation.* (2025, Janvier 01). Récupéré sur Britannica Money: <https://www.britannica.com/money/NVIDIA-Corporation>

*Office of Science and Technology Policy (OSTP).* (s.d.). Récupéré sur usagov: <https://www.usa.gov/agencies/office-of-science-and-technology-policy>

Parker, D. L. (2022, Juin 29). *National Artificial Intelligence Initiative.* Récupéré sur USPTO: <https://www.uspto.gov/sites/default/files/documents/National-Artificial-Intelligence-Initiative->

Overview.pdf#:~:text=Leadership%20in%20AI%20Strengthen%20AI,Research%20In  
frastructu%20re

Poutin, V. (2017, Septembre 01). *Putin: Leader in artificial intelligence will rule world.* Récupéré sur APnews: <https://apnews.com/article/bb5628f2a7424a10b3e38b07f4eb90d4>

Powell, J. (2022, Juin 16). *Etats-Unis: Face à l'inflation, la Fed procède à la plus forte hausse des taux depuis 1994.* Récupéré sur 20minutes: <https://www.20minutes.fr/economie/3309191-20220616-etats-unis-face-inflation-fed-procede-plus-forte-hausse-taux-depuis-1994>

*Préparation à la preparation d'un contrat de licence d'une demande de brevet ou brevet.* (s.d.). Récupéré sur INPI.

*Produit intérieur brut aux prix du marché / PIB.* (2021, Janvier 28). Récupéré sur INSEE: <https://www.insee.fr/fr/metadonnees/definition/c1365>

*Qu'est ce que l'algorithme KNN ?* (s.d.). Récupéré sur DataScientest: <https://datascientest.com/knn>

*Quels sont les différents algorithmes de Deep Learning ?* (2025, Mai 26). Récupéré sur Jedha: <https://www.jedha.co/formation-ia/algorithmes-deep-learning>

Robert, J. (2020, Juillet 20). *Algorithme de descente de gradient : Qu'est-ce que c'est ?* Récupéré sur DataScientest: <https://datascientest.com/descente-de-gradient>

Robert, J. (2024, Janvier 09). *L'apprentissage semi-supervisé : Tout savoir ce qu'il faut savoir.* Récupéré sur DataScientest: <https://datascientest.com/apprentissage-semi-supervise-tout-savoir>

*Robo-advisor.* (s.d.). Récupéré sur Yomoni: <https://www.yomoni.fr/apprendre/gestion/robo-advisor#Les-robo-advisors>

Rodriguez, A. (s.d.). *Indice de Gini (ou coefficient de Gini).* Récupéré sur <https://statorials.org/indice-de-gini-ou-coefficient-de-gini/>

Roy, A., & M.Tobin, J. (2024, Aout). *Intelligence artificielle et apprentissage automatique : définition des termes et concepts actuels dans la recherche en soins intensifs.* doi:<https://doi.org/10.1016/j.jcrc.2024.154792>

Sane, P. (2021, Février 20). *Les enjeux éthiques de l'IA en 2025*. Récupéré sur Upmynt: [upmynt.com](http://upmynt.com)

Sane, P. (2025, Février 20). *Upmynt*. Récupéré sur Les enjeux éthiques de l'IA en 2025: [upmynt.com](http://upmynt.com)

Singh, V. (2024, Novembre 18). *Facteur d'inflation de la variance (VIF) : Traitement de la multicolinéarité dans l'analyse de régression*. Récupéré sur Datacamp: <https://www.datacamp.com/fr/tutorial/variance-inflation-factor>

Singleton, L. (2021, Décembre 08). *Firms need to tackle biases to drive research and development*. Récupéré sur Imperial College London: <https://www.imperial.ac.uk/news/232474/firms-need-tac8kle-biases-drive-research/#:~:text=reasons%3A>

*Strategic Long-Term Participation by DoD in its Manufacturing USA Institutes*. (2019, Avril 24). Récupéré sur ManufacturingUSA: <https://www.manufacturingusa.com/reports/strategic-long-term-participation-dod-its-manufacturing-usa-institutes>

*Théorie des capacités dynamiques: les compétences pour savoir s'adapter – D. Teece*. (s.d.). Récupéré sur SI & Management: <http://www.sietmanagement.fr/modele-des-capacites-dynamiques-d-teece/>

*Une intelligence artificielle « inclusive » et « éthique » : l'accord de 61 pays, sans les États-Unis et le Royaume-Uni*. (2025, Février 11). Récupéré sur Le PARISIEN: <https://www.leparisien.fr/high-tech/une-intelligence-artificielle-ouverte-inclusive-et-ethique-laccord-de-61-pays-sans-les-etats-unis-11-02-2025-YU57T6UOEVEKLAY2HEW2ZKVMEQ.php>

# Annexes

## Annexe 1 : Information primaire avant nettoyage de la base de données

Années	Type	Valeurs_non_nulles	Valeurs_manquantes
DIRD total ( en millions de dollars courants)	int64	15	0
DIRD secteurs entreprises ( en millions de dollars courants)	float64	14	1
DIRD Secteur public ( en millions de dollars courants)	float64	14	1
DIRD Secteur de l'enseignement supérieur ( en millions de dollars courants)	float64	14	1
DIRD Secteur privé à but non lucratif ( en millions de dollars courants)	float64	14	1
GDP ( en millions de dollars courants)	float64	14	1
Dépenses fédérale de R&D ( en millions de dollars )	float64	15	0
Total BERD	float64	13	2
VC investIA ( en million)	float64	13	2
Demande de Brevet	float64	14	1
Offre de Brevet	float64	14	1
Brevet IA D	float64	12	3
Brevet IA O	float64	11	4
Cycle Time to market des brevets	float64	15	0
Investissement en capital risque VC (million)	float64	15	0
CIR ( en million)	float64	12	3
Credit budgétaire publique ( en million de dollars)	int64	15	0
Chercheur_ATP(en million)	float64	12	3
pct_publi_sci_collab_inter	float64	14	1
publi_sci_IA	float64	14	1
publi_sci_IA_collab_pct	float64	14	1

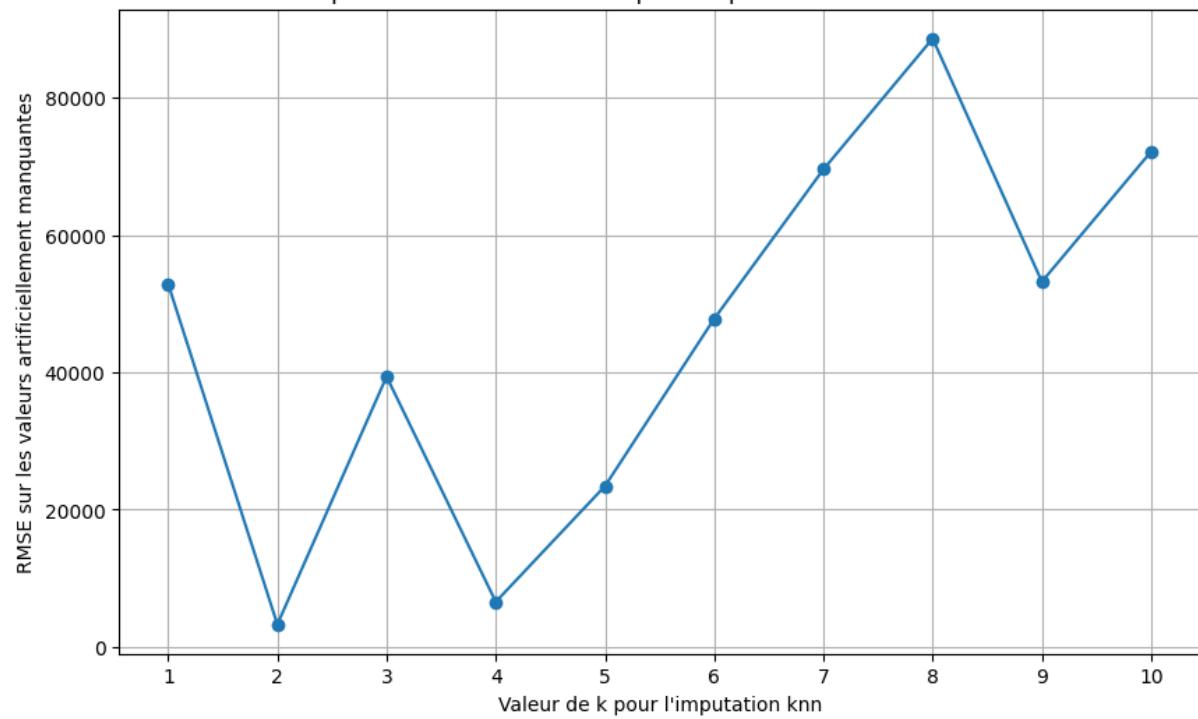
Tableau 9 : Informations primaire avant nettoyage de la base de données

Annexe 2 : Information sur le pourcentage de valeurs manquantes de la base de données

	Valeur_manquante	Pourcentage
Années	0	0.000000
DIRD total ( en millions de dollars courants)	1	6.666667
DIRD secteurs entreprises ( en millions de dollars courants)	1	6.666667
DIRD Secteur public ( en millions de dollars courants)	1	6.666667
DIRD Secteur de l'enseignement supérieur ( en millions de dollars courants)	1	6.666667
DIRD Secteur privé à but non lucratif ( en millions de dollars courants)	1	6.666667
GDP ( en millions de dollars courants)	1	6.666667
Dépenses fédérale de R&D ( en millions de dollars )	0	0.000000
Total BERD	2	13.333333
VC investIA ( en million)	2	13.333333
Demande de Brevet	1	6.666667
Offre de Brevet	1	6.666667
Brevet IA D	3	20.000000
Brevet IA O	4	26.666667
Cycle Time to market des brevets	0	0.000000
Investissement en capital risque VC (million)	0	0.000000
CIR ( en million)	3	20.000000
Credit budgétaire publique ( en million de dollars)	0	0.000000
Chercheur_ATP( en milion)	3	20.000000
pct_publi_sci_collab_inter	1	6.666667
publi_sci_IA	1	6.666667
publi_sci_IA_collab_pct	1	6.666667

Annexe 3 : Imputation des valeurs manquantes par la méthode des k-NN

Imputation des valeurs manquantes par la méthode des K-NN



Source : Elaboration personnelle tiré des données de la partie 3

Annexe 4 : Information sur le pourcentage de valeurs manquante après imputation des K-NN

Σ

	0
Années	0.0
DIRD Total(en million)	0.0
DIRDF(en million)	0.0
DIRDPublic (en Million)	0.0
DIRD Secteur de l'enseignement supérieur ( en millions de dollars courants)	0.0
DIRD SPBNL ( en million )	0.0
GDP ( en millions de dollars courants)	0.0
Dfédé_R&D ( en million )	0.0
Total BERD	0.0
VC investIA ( en million )	0.0
Demande de Brevet	0.0
Offre de Brevet	0.0
Brevet IA D	0.0
Brevet IA O	0.0
Cycle ttm B	0.0
VC(en million)	0.0
CIR ( en million )	0.0
CreditBP (en million)	0.0
Chercheur_ATP( en milion )	0.0
pct_publi_sci_collab_inter	0.0
publi_sci_IA	0.0
publi_sci_IA_collab_pct	0.0

dtype: float64

---

*Source : Elaboration personnelle tiré des données de la partie 3*

Annexe 5 : Analyse descriptive des variables de R&D (1ere partie)

	DIRD Total(en million)	DIRDE(en million)	DIRD Public (en Million)	DIRD Secteur de l'enseignement supérieur ( en millions de dollars courants)	DIRD SPBNL ( en million )	GDP ( en millions de dollars courants)
count	14.000000	14.000000	14.000000		14.000000	14.000000
mean	608340.857143	452469.557143	59952.214286		73010.928571	22908.571429
std	183165.836169	157733.091474	8354.739173		13676.608872	3963.083778
min	408495.000000	278977.000000	52372.000000		58083.000000	17959.000000
25%	460559.500000	327078.000000	53539.000000		61712.750000	18963.500000
50%	549448.500000	403468.500000	54819.500000		69443.000000	23218.500000
75%	716641.500000	543161.500000	66824.500000		80109.500000	26547.000000
max	955578.000000	749390.000000	75823.000000		102044.000000	28320.000000

Source : Elaboration personnelle tiré des données de la partie 3

### Annexe 6 : Analyse descriptive des variables de R&D (2<sup>e</sup> partie)

	Dfédé_R&D ( en million)	Total BERD	VC investIA ( en million)	Demande de Brevet	Offre de Brevet	Brevet IA D	Brevet IA O
count	15.000000	13.000000	13.000000	14.000000	14.000000	12.000000	11.000000
mean	153094.222867	429629.307692	43721.694839	580671.928571	301273.642857	6154.516042	5153.463124
std	29663.736436	137989.406471	34627.743331	42294.503570	43900.517531	3819.660891	2902.071262
min	118273.784000	278977.000000	3581.508444	490226.000000	219614.000000	1966.756958	1811.427856
25%	132178.651500	322528.000000	17123.106662	573409.500000	282978.250000	2848.040894	2768.994751
50%	140670.243000	390336.000000	38796.072598	591352.000000	305404.000000	4878.570801	4298.707520
75%	182816.785500	508420.000000	62633.574678	603472.000000	334505.750000	9793.823975	7364.394531
max	200928.462000	711044.000000	117347.342261	642215.000000	354430.000000	11552.218750	10001.958984

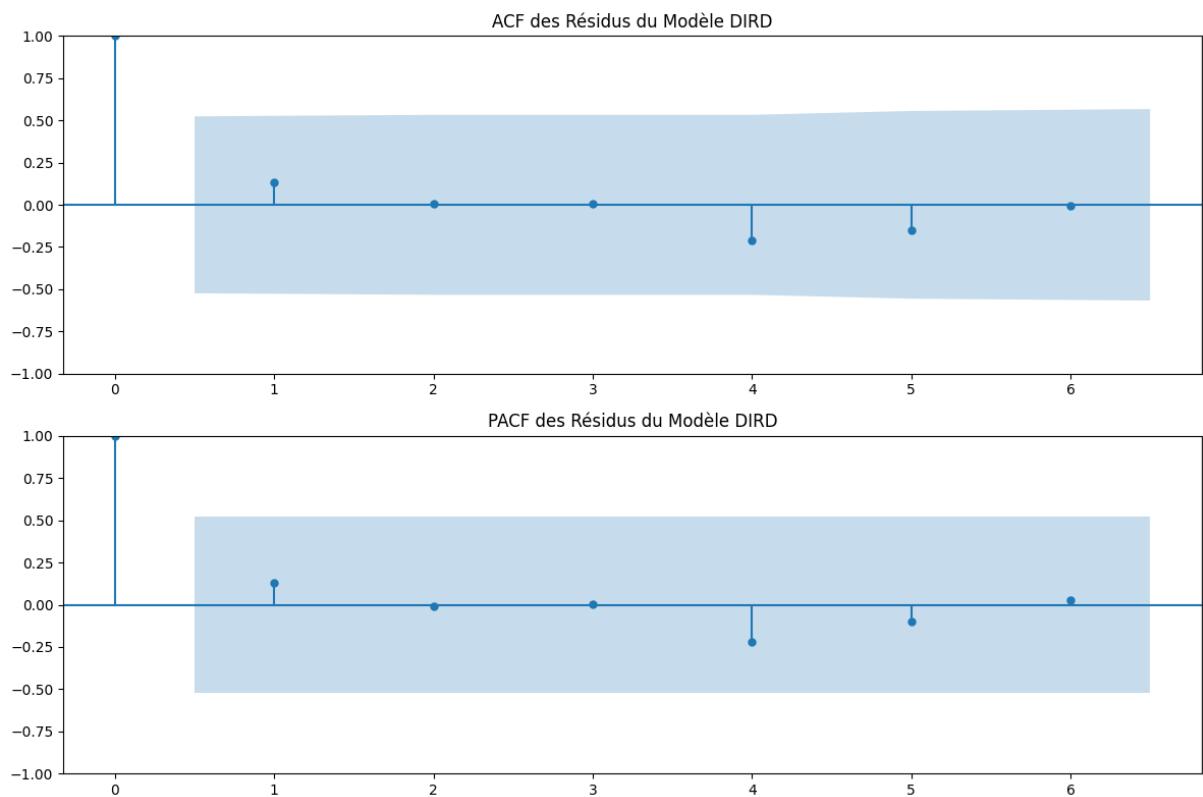
Source : Elaboration personnelle tiré des données de la partie 3

### Annexe 7 : Analyse descriptive des variables de R&D (3<sup>e</sup> partie)

	VC(en million)	CIR ( en million)	CreditBP (en million)	Chercheur_ATP( en milion)	pct_publi_sci_collab_inter	publi_sci_IA	publi_sci_IA_collab_pct
count	15.000000	12.000000	15.000000	12.000000	14.000000	14.000000	14.000000
mean	100961.409326	17427.050000	143749.666667	1330.794264	26.221061	2716.075041	21.015463
std	66654.779813	7962.673009	33121.748200	166.136483	2.598763	833.592193	2.780672
min	30481.141083	8510.700000	109608.000000	1127.894732	22.721334	1779.974426	15.821487
25%	50306.605852	11181.450000	116008.000000	1206.373975	24.430061	2035.533535	19.132295
50%	71057.580677	14712.450000	127306.000000	1279.481049	25.759141	2517.784488	21.052767
75%	129659.042997	23872.950000	165893.000000	1449.129222	27.224406	3364.964650	23.149839
max	266686.174847	33317.600000	203919.000000	1664.938655	32.418245	4234.041958	24.814020

Source : Elaboration personnelle tiré des données de la partie 3

### Annexe 8: ACF et PACF du model\_DIRD

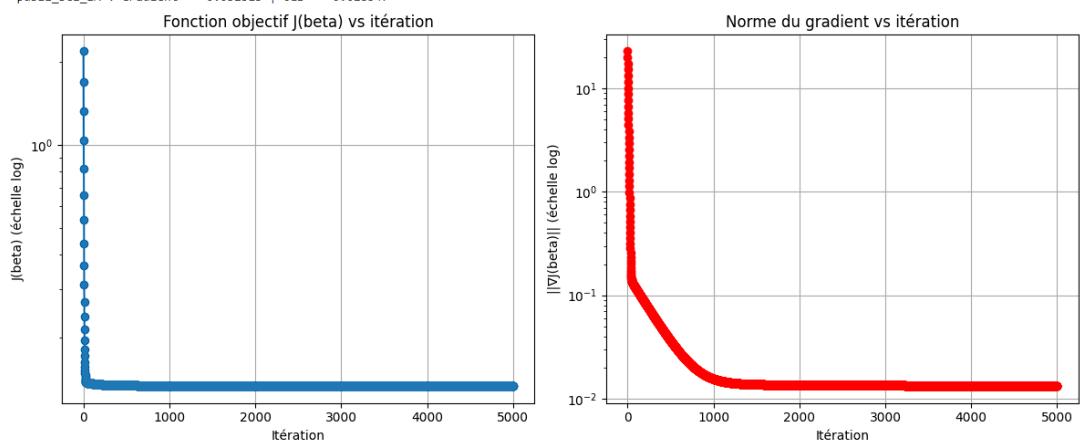


Source : Elaboration personnelle tiré des données de la partie 3

### Annexe 9: Graphique de convergence de la descente de gradient du modèle DIRD

```
--- Descente de gradient sur model_ttm (approx. OLS) ---
Coefficients (Gradient) : [0.08463413 0.17589951 0.03251331]
Itérations nécessaires : 5000
Norme finale du gradient: 1.319143e-02
Valeur finale J(beta) : 1.335517e-01
```

```
Comparaison des coefficients :
const : Gradient = 0.084634 | OLS = 1.410000
log_GDP_txr : Gradient = 0.175900 | OLS = 0.052506
publi_sci_IA : Gradient = 0.032513 | OLS = 0.018547
```



Ces deux graphiques confirment la convergence de l'algorithme de gradient vers un minimum local, en accord avec les résultats OLS.

*Source : Elaboration personnelle tiré des données de la partie 3*

Annexe 10: Comparaison avec les résultats du modèle OLS

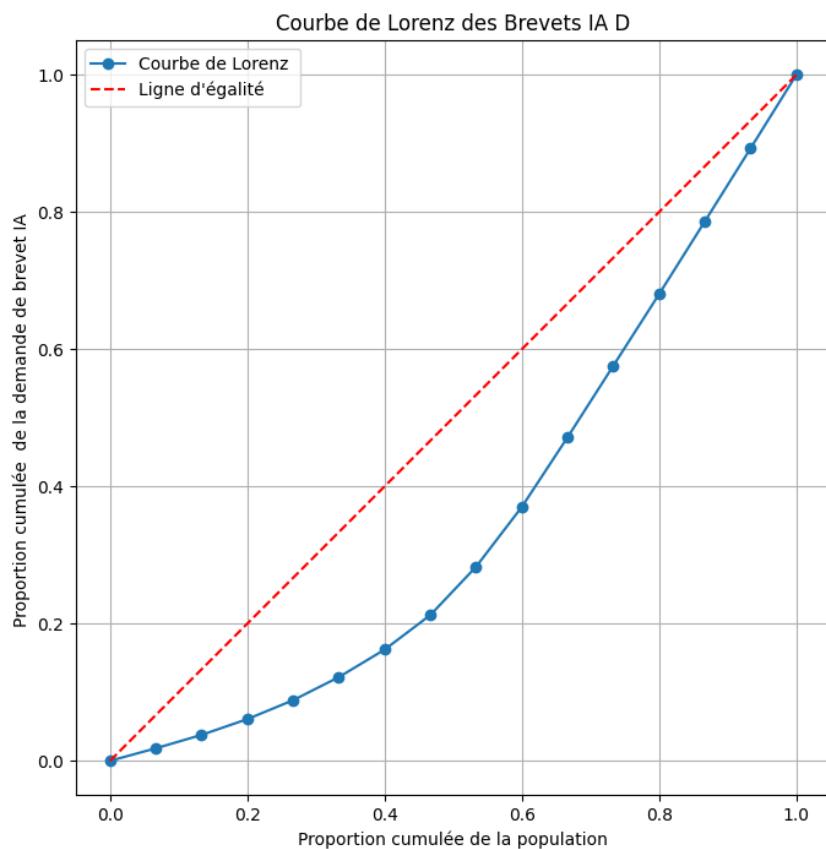
```
Comparison with OLS results:
OLS Regression Results
=====
Dep. Variable: log_DIRD_Total_txc R-squared:      0.676
Model:          OLS   Adj. R-squared:    0.579
Method:         Least Squares F-statistic:     18.05
Date:           Thu, 12 Jun 2025 Prob (F-statistic): 0.000232
Time:           11:04:44 Log-Likelihood:   -29.409
No. Observations: 14 AIC:             66.82
Df Residuals:    10 BIC:            69.37
Df Model:        3
Covariance Type: HAC
=====
      coef    std err      t      P>|t|      [0.025      0.975]
-----
Intercept     -0.5344    1.874   -0.285    0.781     -4.710     3.641
log_GDP_txc    0.6791    0.200    3.396    0.007      0.234     1.125
publi_sci_IA    0.0010    0.001    1.256    0.238     -0.001     0.003
logChercheur_ATP_txc  0.3530    0.161    2.192    0.053     -0.006     0.712
-----
Omnibus:            1.395 Durbin-Watson:       1.605
Prob(Omnibus):      0.498 Jarque-Bera (JB):  0.891
Skew:               0.265 Prob(JB):        0.640
Kurtosis:            1.884 Cond. No.  1.15e+04
=====

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 1 lags and without small sample correction
[2] The condition number is large, 1.15e+04. This might indicate that there are
strong multicollinearity or other numerical problems.

Note: Gradient Descent results may differ slightly from the OLS results
due to convergence properties, learning rate, iterations, and tolerance settings.
OLS provides the exact solution analytically (for standard OLS).
```

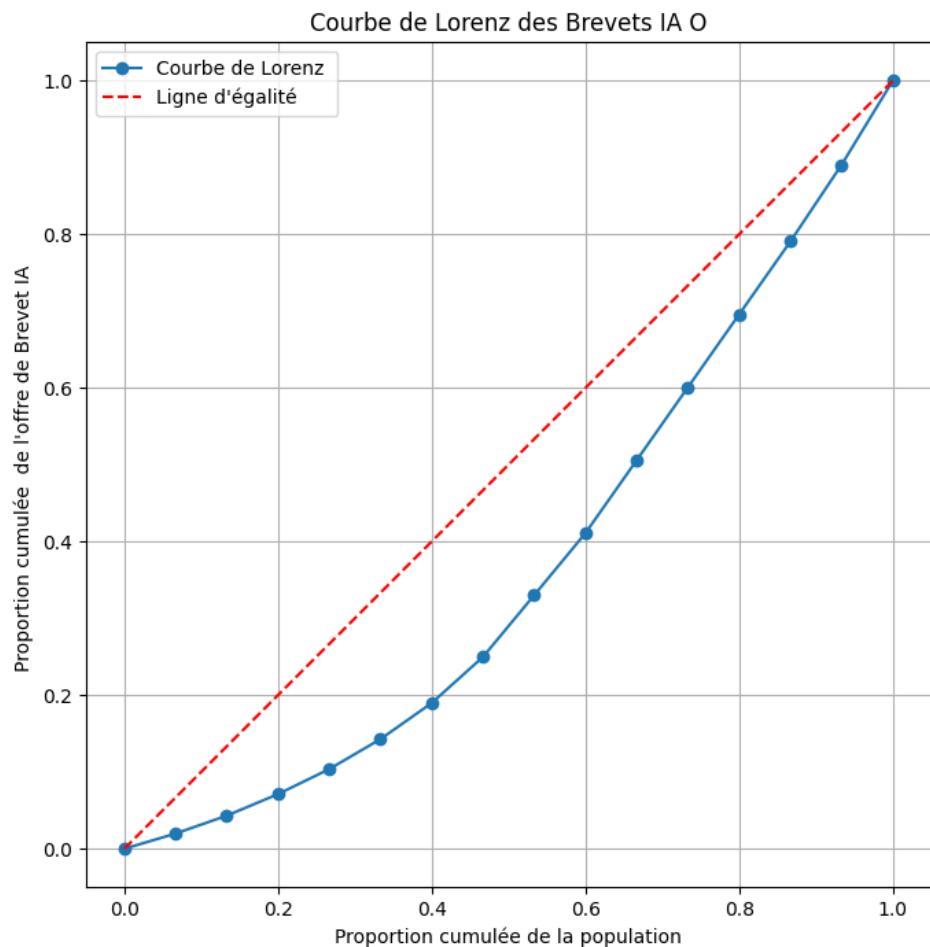
*Source : Elaboration personnelle tiré des données de la partie 3*

### Annexe 11: Courbe de Lorenz de la demande de brevet lié à l'IA



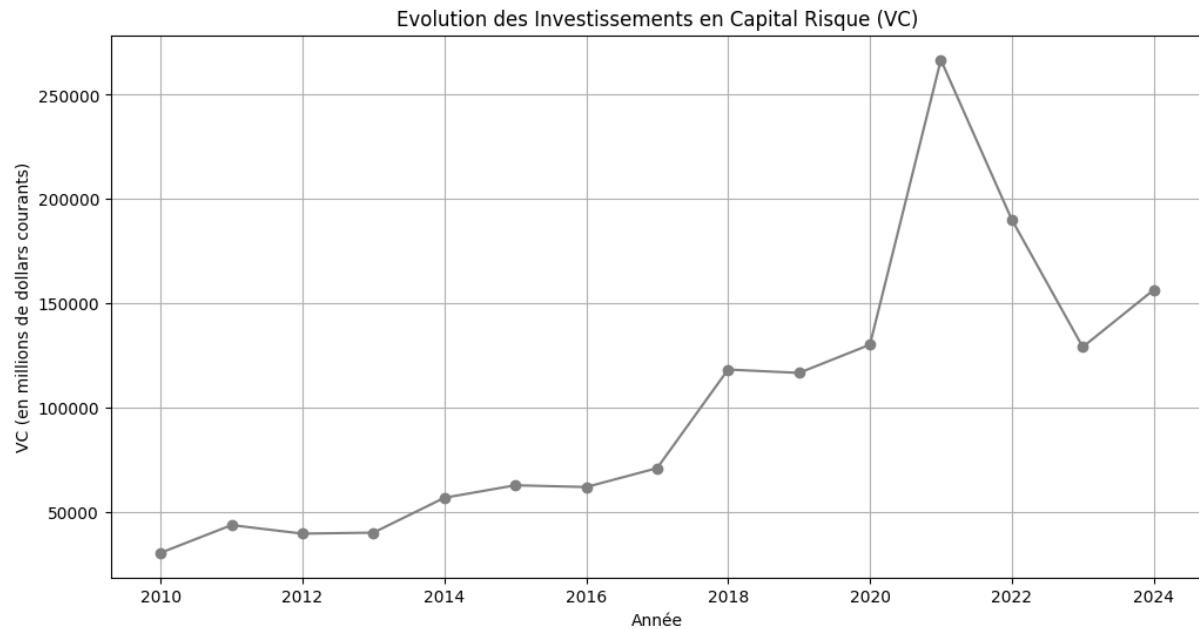
Source : *Elaboration personnelle à partir des données de l'OCDE*

#### Annexe 12: Courbe de Lorenz de l'offre de brevet lié à l'IA



Source : *Elaboration personnelle à partir des données de l'OCDE*

Annexe 13: Représentation graphique de l'évolution de l'investissement en capital risque de 2010-2024



*Source : Elaboration personnelle à partir des données de l'OCDE*

Annexe 14 : Code en Langage Python<sup>21</sup>:

Annexe 15: Lien du Questionnaire en version française

<https://forms.gle/QYTcLEQyu4Jy8u5j6>

---

<sup>21</sup> Vous pourrez trouver tous les codes que nous avons utilisé via le lien de partage suivant :  
[https://colab.research.google.com/drive/1irTyA2-\\_m8-ExP6Oibc0-bQaP8wcTBkS?usp=sharing](https://colab.research.google.com/drive/1irTyA2-_m8-ExP6Oibc0-bQaP8wcTBkS?usp=sharing)

# Index

## A

ACP

Analyse en composantes principales..... 31, 32

## C

CNIL..... 13

## D

DDDM

Data Driven Decision Making..... 13

*DIRDE*

Dépense Intérieur de Recherche et Développement dans le secteur des Entreprises..... 28, 29

DQN

Deep Q-Network..... 15

## E

ED..... 20

EIA

Ethique de l'intelligence artificielle ..... 19, 23

EPA..... 20

United States Environmental Protection Agency ( ou L'Agence de protection de l'environnement des États-Unis en Français )..... 20

## H

HHI

Indice de Herfindahl-Hirschman ..... 32

## I

IAG

intelligence artificielle générale ..... 23, 67

IBM .....	13
<b>J</b>	
JDN	
Journal du net .....	23, 29
<b>K</b>	
k-NN	
Les k plus proches voisins.....	14
<b>N</b>	
NASA	
National Aeronautics and Space Administration.....	20
NSF.....	20, 25
<b>O</b>	
<b>OCDE</b>	
Organisation de Coopération et de Développement Economiques .....	8, 25, 27
<b>OMPI</b>	
Organisation Mondiale de la Propriété Intellectuelle.....	26
<b>OSTP</b>	
Office of Science and Technology Policy .....	26
<b>P</b>	
PPO	
Proximal Policy Optimization.....	15
<b>R</b>	
<b><i>R&amp;D</i></b>	
<i>Recherche et développement</i> 1, 6, 7, 8, 10, 11, 12, 13, 16, 17, 18, 19, 24, 27, 28, 29, 31, 32, 33, 34, 64, 65, 66, 67, 68	
<b>S</b>	
SBIR	
Small Business Innovation Research .....	20

SME.....	10
<b>STTR</b>	
Small Business Technology Transfer .....	20
<b>U</b>	
<b>USDA</b>	
United States Department of Agriculture ( ou Département de l'Agriculture des États-Unis en français) .....	20
<b>USPTO</b>	
United States Patent and Trademark Office .....	26
<b>V</b>	
<b>VAN</b>	
Valeur actuelle nette	
c'est un indicateur financier qui peut être utilisé pour apprécier la rentabilité d'un investissement .....	17
<b>W</b>	
<b>WBG</b>	
Word Bank Group .....	25

# TABLEAU DES ILLUSTRATIONS

## IMAGES

<i>Figure 1: Image du processus d'apprentissage au travers du modèle de Deep Learning (MapR, C.D, Futura)</i>	16
<i>Figure 2 : Un modèle des capacités dynamiques (Labrouche, 2014)</i>	17
<i>Figure 3: Représentation graphique de la répartition des meilleurs secteurs du brevet de 2010-2020</i>	37
<i>Figure 4 : Représentation graphique de l'évolution de la productivité de l'IA de 2010-2024</i>	40
<i>Figure 5: Représentation graphique de l'évolution de la productivité des brevets de 2010-2024</i>	42
<i>Figure 6: Représentation graphique de l'évolution du ratio DIRD-PIB de 2010-2024</i>	45
<i>Figure 7 : Représentation graphique de l'évolution du soutien public total à la R&amp;D des entreprises de 2010-2024</i>	47
<i>Figure 8: MATRICE DE CORRÉLATION</i>	49
<i>Figure 9: Cercle de corrélation (PC1 et PC2)</i>	51
<i>Figure 10: Courbe de Lorenz de la productivité des brevets</i>	52
<i>Figure 11: Courbe de Lorenz de la productivité de l'IA</i>	53
<i>Figure 12: Représentation graphique du HHI des brevets de 2010-2020</i>	55
<i>Figure 13 : Visualisation de la méthode du gradient du model de cycle time to Market</i>	62
<i>Figure 14: Descente de gradient sur le modèle de cycle de time to Market</i>	63

## TABLEAUX

<i>Tableau 1: La productivité de L'IA</i>	39
<i>Tableau 2: La productivité des brevets</i>	42
<i>Tableau 3: Intensité de la Dépense intérieur en R&amp;D (DIRD) par rapport au PIB</i>	45
<i>Tableau 4: Soutien public total à la R&amp;D des entreprises</i>	47
<i>Tableau 5: Modèle OLS avec la DIRD</i>	55

<i>Tableau 6: Modèle SARIMAX avec la DIRD</i>	57
<i>Tableau 7: Modèle OLS du cycle time-to-Market</i>	59
<i>Tableau 8: Modèle OLS de la productivité des brevets</i>	60
<i>Tableau 9 : Informations primaire avant nettoyage de la base de données</i>	81

---

<sup>i</sup> Vous pourrez trouver les données du PIB Etats-Unien via ce lien-là :

<https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?end=2023&locations=US&start=2010>

<sup>ii</sup> Vous pourrez trouver les données des offres et des de mande de brevets via ces liens là :

[https://www.uspto.gov/web/offices/ac/ido/oeip/taf/h\\_counts.htm](https://www.uspto.gov/web/offices/ac/ido/oeip/taf/h_counts.htm)

Et : <https://www.uspto.gov/sites/default/files/documents/USPTOFY21PAR.pdf>

<sup>iii</sup> Vous pourrez trouvé la base de donnée des Offres et des demandes de brevets IA ( sans compter les autres types de brevets via ce lien la : [OECD Data Explorer • Patents by technology](#)

<sup>iv</sup> Vous pourrez trouvé la base de donnée des différents types de DIRD via ce lien la :

[https://data-](https://data-explorer.oecd.org/)

[https://data-explorer.oecd.org/vis?tm=R%26D%20expenditure&snb=108&isAvailabilityDisabled=false&pg=0&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_RDS\\_GERD%40DF\\_GERD\\_TORD&df\[ag\]=OECD.STI.STP&df\[vs\]=1.0&dq=.A..\\_T...\\_T...XDC.&pd=2015%2C&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?tm=R%26D%20expenditure&snb=108&isAvailabilityDisabled=false&pg=0&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_RDS_GERD%40DF_GERD_TORD&df[ag]=OECD.STI.STP&df[vs]=1.0&dq=.A.._T..._T...XDC.&pd=2015%2C&to[TIME_PERIOD]=false)

<sup>v</sup> Vous pourrez trouver la base de données du cycle time to market des brevets via ces liens-là : <https://www.uspto.gov/sites/default/files/documents/USPTOFY2014PAR.pdf>

Et <https://www.uspto.gov/dashboard/patents/pendency.html>

<https://www.jct.gov/getattachment/ec4fb616-771b-4708-8d16-f774d5158469/x-23-20.pdf>

<sup>vi</sup> Vous pourrez trouvé la base de donnée du crédit budgétaire publique via ce lien là :

[https://data-](https://data-explorer.oecd.org/)

[https://data-explorer.oecd.org/vis?lc=fr&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_RDS\\_GOV%40DF\\_GBARD\\_NABS07&df\[ag\]=OECD.STI.STP&dq=USA.A..\\_T...\\_T...XDC.&pd=2010%2C&to\[TIME\\_PERIOD\]=false&vw=tb](https://data-explorer.oecd.org/vis?lc=fr&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_RDS_GOV%40DF_GBARD_NABS07&df[ag]=OECD.STI.STP&dq=USA.A.._T..._T...XDC.&pd=2010%2C&to[TIME_PERIOD]=false&vw=tb)

<sup>vii</sup> Vous pourrez trouver la base de données du CIR via ces liens là :

<https://www.jct.gov/getattachment/ec4fb616-771b-4708-8d16-f774d5158469/x-23-20.pdf>

---

Puis : [https://data-explorer.oecd.org/vis?df\[ds\]=DisseminateFinalDMZ&df\[id\]=DSD\\_RDTAX%40DF\\_RDTAX&df\[ag\]=OECD.STI.STP&dq=USA.A..XDC..&pd=2010%2C&to\[TIME\\_PERIOD\]=false&vw=tb](https://data-explorer.oecd.org/vis?df[ds]=DisseminateFinalDMZ&df[id]=DSD_RDTAX%40DF_RDTAX&df[ag]=OECD.STI.STP&dq=USA.A..XDC..&pd=2010%2C&to[TIME_PERIOD]=false&vw=tb)

viii Vous pourrez trouver la base de données des variables restantes via ce lien là :  
[https://stip.oecd.org/stats/SB-StatTrends.html?i=G\\_XGDP&v=3&t=1998,2021&s=CHN,EU27\\_2021,JPN,KOR,OECD,USA,GBR](https://stip.oecd.org/stats/SB-StatTrends.html?i=G_XGDP&v=3&t=1998,2021&s=CHN,EU27_2021,JPN,KOR,OECD,USA,GBR)