

## THESIS / THÈSE

### MASTER EN INGÉNIEUR DE GESTION À FINALITÉ SPÉCIALISÉE EN DATA SCIENCE

#### L'importance du Machine Learning et de la Data Science dans l'Industrie 4.0

BACQ, Mathieu

*Award date:*  
2023

*Awarding institution:*  
Universite de Namur

[Link to publication](#)

#### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



# L'importance du Machine Learning et de la Data Science dans l'Industrie 4.0

**Mathieu BACQ**

**Directeur: Prof. S. FAULKNER**

Mémoire présenté  
en vue de l'obtention du titre de  
Master 120 en ingénieur de gestion, à finalité spécialisée  
en data science

**ANNEE ACADEMIQUE 2022-2023**

# Résumé/Summary

## Résumé

L'industrie 4.0 se développe de plus en plus rapidement pour permettre aux entreprises d'avoir de meilleures performances opérationnelles et financières. La mise en place de l'industrie 4.0 permet de récolter un nombre considérable de données relatives à la production. Il est alors essentiel de trouver des méthodes qui permettent de traiter ces données automatiquement afin d'en tirer de l'information et de la connaissance. Les techniques de machine learning et de data science conviennent exactement pour cette utilisation. Ce mémoire est une revue de littérature questionnant la pertinence des technologies de machine learning et de data science au sein de l'industrie 4.0. Ce travail consiste en l'analyse de sources de littérature scientifique pour déterminer s'il existe un lien entre la mise en application de ces technologies et l'augmentation des performances des entreprises ayant instauré les principes de l'industrie 4.0.

## Summary

Industry 4.0 is developing at an increasing rate to enable companies to have better operational and financial performance. The implementation of Industry 4.0 allows for the collection of a considerable amount of data related to production. It is then essential to find methods to process this data automatically in order to extract information and knowledge. Machine learning and data science techniques are exactly suited for this purpose. This thesis is a literature review questioning the relevance of machine learning and data science technologies within Industry 4.0. This work consists of analyzing scientific literature sources to determine if there is a link between the application of these technologies and the increase in performance of companies that have implemented the principles of Industry 4.0.

# Remerciements

Tout d'abord, je voudrais remercier mon promoteur, le Professeur Faulkner, directeur de ce mémoire, pour tous les bons conseils qu'il a pu me donner et son professionnalisme tout au long de ce travail.

Je tiens aussi à remercier tous les membres de ma famille pour leur soutien inconditionnel. Je remercie tout particulièrement mon frère et ma sœur pour leurs conseils et pour la relecture de ce travail.

Finalement, je remercie les autres étudiants de ma filière qui ont été disponibles pour répondre aux questions que je pouvais me poser pendant la réalisation de ce travail.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Revue de la littérature</b>	<b>7</b>
2.1	Industrie 4.0 . . . . .	7
2.1.1	Internet des objets . . . . .	7
2.1.2	Système cyber-physique . . . . .	8
2.1.3	Big data . . . . .	8
2.1.4	Visualisation des données . . . . .	8
2.1.5	Cycle des données . . . . .	9
2.2	Data science et machine learning . . . . .	9
2.3	Travaux similaires . . . . .	10
<b>3</b>	<b>Méthodologie</b>	<b>11</b>
3.1	Recherche d'articles scientifiques . . . . .	12
3.2	Sélection d'articles scientifiques . . . . .	13
3.3	Analyse d'articles scientifiques . . . . .	14
3.4	Recherche de données chiffrées . . . . .	14
3.5	Recherche d'articles complémentaires . . . . .	14
<b>4</b>	<b>Analyse des données et résultats</b>	<b>15</b>
4.1	Analyse des données . . . . .	15
4.1.1	Articles scientifiques sélectionnés . . . . .	15
4.1.2	Articles scientifiques retenus . . . . .	16
4.1.3	Données chiffrées sélectionnées . . . . .	19
4.2	Résultats . . . . .	19
4.2.1	Mise en pratique du machine learning et de la data science dans l'industrie 4.0 . . . . .	20
4.2.2	Intérêts en termes de performance de passer à l'industrie 4.0 . . . . .	23
4.2.3	Amélioration des performances de l'industrie 4.0 grâce au machine learning et à la data science . . . . .	25
4.2.4	Maturité du machine learning et de la data science dans l'industrie 4.0 . . . . .	28
<b>5</b>	<b>Discussion</b>	<b>31</b>
5.1	Réponses aux questions de recherche . . . . .	32
5.2	Comparaison avec la littérature existante . . . . .	34
5.3	Limitations . . . . .	34
5.4	Recommandations . . . . .	35
<b>6</b>	<b>Conclusion</b>	<b>35</b>
	<b>Références</b>	<b>37</b>
<b>A</b>	<b>Annexes</b>	<b>42</b>

# L'importance du Machine Learning et de la Data Science dans l'Industrie 4.0

Mathieu Bacq

Juin 2023

## 1 Introduction

L'industrie 4.0 est le nom donné à la quatrième révolution industrielle qui est en marche. Dans un monde de plus en plus connecté, où les objets sont aussi connectés à Internet, il est naturel que ces technologies soient intégrées aux processus de fabrication d'un grand nombre d'entreprises. Les machines connectées et les capteurs génèrent alors une grande quantité de données qu'il faut traiter. Étant donné la quantité d'informations, il faut implémenter des méthodes automatiques de traitement de données. Les techniques avancées en data science et en machine learning sont adaptées à ce genre de cas dans lesquels sont générées des données en grande quantité et avec une vitesse élevée. La data science et le machine learning doivent donc permettre de traiter ces données pour en dégager de l'information et de la connaissance. Les entreprises y trouvent un intérêt en améliorant les processus et donc leurs performances et/ou leurs coûts. Il est donc opportun de se demander comment ces technologies peuvent être mises au service de l'industrie 4.0 et quelle est leur utilité.

Depuis la popularisation du terme "industrie 4.0" au début des années 2010 en Allemagne, il est possible de constater à la FIGURE 1.1, qui montre l'intérêt des recherches Google pour le sujet, que l'attrait pour ce domaine n'a cessé de croître jusqu'en 2020. Depuis, l'intérêt a un peu diminué mais il reste assez élevé. De plus, les technologies de machine learning ont aussi de plus en plus d'intérêt, particulièrement dans les entreprises pour essayer de remplacer les tâches pénibles et répétitives par des algorithmes. La data science a aussi beaucoup gagné d'intérêt ces dernières années puisqu'il est de plus en plus facile pour les entreprises de générer ou de se procurer beaucoup de données utiles. L'utilisation de ces technologies est souhaitée dans l'industrie 4.0. Ces domaines sont en croissance depuis quelques années, la littérature est relativement récente sur le sujet, d'où l'intérêt d'en réaliser une revue pour synthétiser ce qui a déjà été fait et aider

à la réalisation de futurs travaux. Au cours de ce mémoire, l'attention sera portée sur l'industrie de la fabrication et de la production de biens. Les autres types d'industries seront brièvement abordés.



**FIGURE 1.1** – Intérêt pour la recherche "industry 4.0" sur Google

Les questions de recherche de ce mémoire sont donc les suivantes : Quelles sont les applications pratiques du machine learning et de la data science dans le monde de l'industrie 4.0? Quels sont les intérêts en termes de performance de mettre en place ces technologies pour des entreprises voulant passer à l'industrie 4.0? Et quelle est la maturité de l'intégration du machine learning et de la data science dans l'industrie 4.0?

Étant donné la quantité limitée de littérature scientifique portant sur ces questions de recherche, il semble important d'y apporter une réponse en parcourant la littérature existante et en la synthétisant pour permettre de trouver des réponses aux questions de recherche de ce mémoire. Ce mémoire étant une revue de littérature intégrative, l'analyse ne sera pas basée sur des données chiffrées mais sur une analyse qualitative de divers articles scientifiques. Ce mémoire peut aussi servir de base pour de futures recherches afin d'approfondir les concepts théoriques sur ce sujet et d'ouvrir la porte à des recherches quantitatives sur l'impact des technologies mentionnées sur la performance au sein des entreprises.

Ce mémoire est composé de plusieurs sections. Tout d'abord, une revue de la littérature décrit les concepts théoriques en ce qui concerne l'industrie 4.0, la data science et le machine learning. La revue de la littérature comporte aussi une section où sont détaillés les travaux similaires qui ont été faits sur ce sujet ou des dérivés. Ensuite, la section méthodologie détaille le processus par lequel les données, dans ce cas, des articles scientifiques et des données chiffrées sont sélectionnés pour leur analyse. Suite à cela, la section analyse des données et résultats détaille quels articles et données ont été collectés et leur représentativité ainsi que les résultats, ici, une synthèse de ces articles scientifiques et de ces données. Après l'analyse, la section discussion présente un résumé des résultats

obtenus pour répondre aux questions de recherche, une comparaison avec la littérature existante, les limitations rencontrées dans la recherche et l'analyse ainsi que des recommandations pour de futurs travaux portant sur le même sujet que ce mémoire. Enfin, la section conclusion synthétise ce mémoire et les réponses aux différentes questions de recherche.

## 2 Revue de la littérature

Cette section est un résumé de la littérature à propos de ce qu'est l'industrie 4.0, de ses composantes et des diverses technologies issues de la data science et du machine learning pouvant être appliquées à ce domaine. Il est principalement question de définir les concepts liés au sujet du présent travail pour déterminer le cadre théorique de ce mémoire.

### 2.1 Industrie 4.0

L'expression "industrie 4.0" (I4.0) désigne la quatrième révolution industrielle. Après la révolution des machines à vapeur, du travail à la chaîne et de la robotisation, c'est au tour des machines et des opérations de se moderniser à nouveau. "L'industrie 4.0 est une collaboration entre la science, les données, l'information et la technologie avec l'intégration de la science des données et de l'ingénierie pour produire des systèmes industriels intelligents" (Nasution, 2021). L'industrie 4.0 apporte "des améliorations fondamentales des processus industriels impliqués dans la fabrication, l'ingénierie, l'utilisation des matériaux et la gestion de la chaîne d'approvisionnement et du cycle de vie" (Kagermann, Wahlster, Helbig, et al., 2013).

#### 2.1.1 Internet des objets

L'internet des objets est un concept relativement nouveau dans le monde de l'industrie. Une définition simple peut être : l'internet des objets est la mise en réseau d'objets pouvant communiquer entre eux ou avec d'autres entités sur internet. Madakam, Ramaswamy, et Tripathi (2015) proposent néanmoins une définition plus exhaustive, l'internet des objets est : "un réseau ouvert et complet d'objets intelligents capables de s'auto-organiser, de partager des informations, des données et des ressources, de réagir et d'agir face à des situations et des changements dans l'environnement". Une autre définition formulée par Weber (2010) est : "l'internet des objets est une architecture d'information émergente basée sur l'internet qui facilite l'échange de biens et de services dans les réseaux de la chaîne d'approvisionnement". L'internet des objets est un pilier dans le concept général de l'industrie 4.0. Il transforme un environnement physique en un environnement cyber-physique, qui allie et fait coexister éléments réels et numériques. Selon

Aheleroff et al. (2020), ceci donne la possibilité d'avoir de la connectivité en temps réel en échangeant et en stockant une grande quantité de données. Ces données capturées par les objets deviennent une source de valeur pour l'entreprise (Aheleroff et al., 2020).

### **2.1.2 Système cyber-physique**

La modernisation des processus de production comprend la mise en réseau des outils de production pour pouvoir communiquer avec en temps réel et pouvoir en extraire des données avec une grande vitesse. De plus, les technologies de l'internet des objets y sont aussi incorporées. Les objets peuvent aussi communiquer entre eux et aux systèmes d'intelligence qui composent l'écosystème global de l'organisation dans le but d'améliorer l'autonomie des machines, et donc la performance de la production. (Hermann, Pentek, & Otto, 2016). Tout cela est désigné par le terme système cyber-physique (de l'anglais Cyber-Physical System, CPS) qui est une mise en communication entre les différents éléments informatiques d'un écosystème pour qu'ils puissent s'échanger des données et aussi communiquer avec les humains, sous forme de données et d'instructions.

### **2.1.3 Big data**

Le big data est un terme qui désigne des données ayant des caractéristiques spécifiques. Elles sont caractérisées par les 3 V du big data (volume, vitesse, variété). Les données arrivent avec un volume extrêmement élevé, elles arrivent avec une haute vitesse (haute fréquence) et dans un nombre très varié de formats (Sagiroglu & Sinanc, 2013). Ces données servent alors aux entreprises pour de l'analyse, leur quantité permet d'établir des modèles assez précis et leur vitesse permet de travailler en temps quasi réel. Les tendances récentes observées dans le domaine du big data se dirigent vers la proposition d'un degré d'abstraction permettant l'utilisation des plateformes de traitement de données afin de déployer des algorithmes de machine learning sur des flux de données (Gokalp et al., 2016).

### **2.1.4 Visualisation des données**

La visualisation de données est le processus par lequel des grandes bases de données peuvent être résumées via des abstractions sur certains visuels (indicateurs clés de performance, graphes, tableaux de synthèse, etc.). Pouvoir visualiser les données clés dans une entreprise permet d'améliorer le processus de décision, d'améliorer les différentes analyses dans des domaines plus précis et d'améliorer la collaboration (Allen et al., 2021). Selon Allen et al. (2021), donner une abstraction des données importantes permet de "résumer rapidement et intuitivement de grands ensembles de données", là où les décisions doivent avoir comme motivation des données tangibles.

### 2.1.5 Cycle des données

Le cycle de traitement des données qui sortent des capteurs et des autres sources est un cycle en trois phases, comme montré sur la FIGURE 2.1 (Gröger, 2018). Tout d'abord, les données sont produites par les machines des processus de production de l'entreprise. Ensuite vient la première phase d'intégration des données qui consiste à nettoyer et préparer les données pour les étapes suivantes du cycle. Suite à cela, la deuxième phase est l'analyse. Elle consiste à effectuer une analyse de ces données. C'est au cours de cette étape que des méthodes analytiques seront utilisées sur les données pour en dégager de l'information. Enfin, la troisième phase est l'étape d'optimisation, c'est le moment où les informations découvertes seront utilisées pour améliorer le processus de production. Et pour finir, le cycle recommence avec le processus de production amélioré.

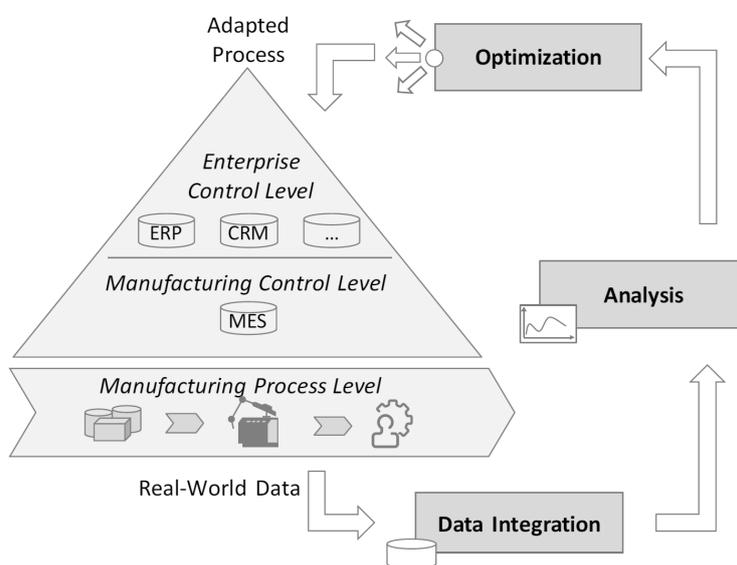


FIGURE 2.1 – Concept de fabrication assistée par les données (Gröger, 2018)

## 2.2 Data science et machine learning

La data science (DS) peut être décrite comme "un ensemble de disciplines nécessaires pour résoudre les défis importants liés aux données" (Song & Zhu, 2016). La data science est aussi "une nouvelle discipline scientifique donnant des stratégies, des techniques et des méthodes pour récupérer des morales et des bribes de connaissances à partir d'index informationnels nouveaux et existants" (Sajid et al., 2021). La data science est aussi décrite comme "la pratique de la collecte de données, de l'analyse des données dans un espace de problèmes, de la déduction de découvertes de modèles significatifs et de la formulation de conclusions, la découverte de modèles de données intéressants et significatifs à l'aide de méthodes d'analyse computationnelle" (Riveret et al., 2019).

Le machine learning (ML) est décrit comme de l'apprentissage automatique par un

ordinateur. Il se compose de deux phases, une phase d'apprentissage (training), où l'algorithme va apprendre des liens entre des données et une phase d'évaluation où l'algorithme va essayer de reproduire ces liens pour prédire une valeur de sortie. L'objectif du machine learning est d'apprendre à produire des sorties en fonction des données d'entrée (Mahesh, 2020).

Pour faire des études sur les données de l'entreprise, il est possible d'utiliser des méthodes qui se basent sur les mathématiques ou sur des analyses statistiques. Au niveau de la data science, ces méthodes sont des statistiques descriptives, des études de corrélation, de la prédiction statistique. Il est aussi possible de faire de la visualisation avancée des données sous plusieurs formes différentes, comme des graphes, cartes ou des indicateurs de performance. Des simulations sont aussi utilisées pour estimer des situations futures. Et du côté du machine learning, ce sont des algorithmes qu'on qualifie d'intelligents (c'est-à-dire qui apprennent des modèles eux-mêmes) qui s'occupent de faire des analyses. Parmi ces algorithmes, il y a les arbres de décision, des algorithmes de regroupements, de la vision par ordinateur ou encore des réseaux de neurones artificiels (Coleman, 2019).

## 2.3 Travaux similaires

Des travaux similaires ont été réalisés sur le même sujet que ce mémoire. Cependant, il est difficile d'en trouver qui parlent de l'impact des technologies comme le machine learning ou la data science sur les performances des entreprises. L'article "Artificial Intelligence, Big Data, and Machine Learning in Industry 4.0" de Lampropoulos (2023) traite du même sujet quant à l'utilisation du machine learning, de la data science, du big data et d'autres technologies. Malheureusement, cet article n'est pas accessible gratuitement sur Internet, il n'a donc pas pu être lu. De plus, le résumé ne précise pas s'il est question de performance quand ils parlent du potentiel des technologies mentionnées, il est marqué que l'article "passe en revue les potentiels que ces technologies pourraient offrir et les mérites qu'elles pourraient produire lorsqu'elles sont appliquées dans le contexte de l'industrie 4.0" (Lampropoulos, 2023).

Un autre article, qui est un projet de Master intitulé "Impact of Big Data Analytics in Industry 4.0" de Oikonomidi (2020) met en perspective la performance des entreprises mais seulement en ce qui concerne le big data. L'article parle du machine learning et de la data science, mais seulement assez brièvement. "Cette étude vise à fournir l'impact de l'analyse du big data dans les environnements de l'industrie 4.0." (Oikonomidi, 2020). L'article conclut que l'amélioration des performances est un aspect positif de l'industrie 4.0 et qu'en améliorant la chaîne de valeur, il est possible d'avoir des produits de meilleure qualité pour un coût global plus bas.

Au niveau de l'amélioration de la performance pour les entreprises, il est possible

de trouver des articles qui mettent en évidence l'impact positif que l'implémentation de l'industrie 4.0 a sur les entreprises (Szász et al., 2021). Cependant, ils ne parlent pas souvent ou pas assez du rôle du machine learning et de la data science dans cette amélioration et ils ne détaillent pas les techniques de machine learning et de big data relatives à l'amélioration de la performance.

En ce qui concerne la maturité du machine learning et de la data science dans l'industrie 4.0, il n'existe presque pas de littérature qui montre l'avancement et la maturité de ce sujet.

D'une manière générale, les travaux similaires ne font pas réellement le lien entre les technologies qui font l'objet de l'analyse de ce mémoire et le gain de performance pour les entreprises.

Les articles scientifiques qui traitent de l'impact des méthodes d'analyse de données sur l'industrie 4.0 utilisent des méthodes quantitatives, qualitatives et des revues de littérature. L'article de Oikonomidi (2020) est une revue de la littérature systématique de l'impact du big data sur l'industrie 4.0. L'article de Szász et al. (2021) est une revue de la littérature et une recherche quantitative de données déjà présentes sur plusieurs bases de données en ligne qui traite seulement de la performance de l'industrie 4.0 sans parler du machine learning ou de la data science. Les méthodes utilisées dans la littérature sont assez variées, il ne semble pas y avoir de règles particulières pour les travaux se penchant sur ce sujet. Ce travail est une revue de littérature intégrative, c'est-à-dire qu'il n'y a pas de recherche systématique d'articles scientifiques. Il y a, malgré tout, une méthodologie suffisamment précise et qui peut être répliquée pour trouver et analyser différentes sources scientifiques.

### 3 Méthodologie

Dans le cadre de ce mémoire qui est une revue de la littérature, la collecte et l'analyse de données prennent la forme d'une revue de littérature intégrative, c'est-à-dire que la forme de la recherche d'articles scientifiques et la méthodologie de recherche sont relativement libres et ne doivent pas suivre une démarche établie. Pour cela, il a fallu récolter un nombre suffisant d'articles scientifiques pertinents afin de les analyser par la suite. Bien que ce type de revue de littérature ne requière pas de méthodologie spécifique, il est quand même utile d'établir une méthode simple de collecte d'articles scientifiques. De plus, certaines données chiffrées sont aussi collectées pour appuyer les propos de la synthèse de la littérature. Une méthodologie simple a aussi été mise en place pour ces données-ci.

### 3.1 Recherche d'articles scientifiques

La première étape est d'établir une liste de mots-clés à entrer dans des moteurs de recherche d'articles scientifiques pour trouver des articles scientifiques pertinents. Le choix de ces mots-clés est subjectif mais suit une certaine logique. Il y a 5 grands axes autour de la problématique et des questions de recherche : l'industrie 4.0, le machine learning, la data science, la performance dans l'industrie 4.0 et la maturité du machine learning et de la data science dans l'industrie 4.0. Étant donné que le sujet le plus général est l'industrie 4.0, chacun des autres axes est couplé à celui-ci pour effectuer les recherches. Les mots-clés croisés sont alors ceux-ci : industry 4.0, data science industry 4.0, machine learning industry 4.0, industry 4.0 performance et machine learning industry 4.0 maturity. Puis, il faut trouver des variantes de ces mots-clés pour obtenir plus de résultats de recherche. Pour créer ces variantes, il faut ajouter des termes souvent utilisés dans le domaine de la catégorie, par exemple, "Data analytics" rentre dans la catégorie "Data science" Les différentes variantes sont détaillées dans la section suivante, à la TABLE 4.1.

Les moteurs de recherche d'articles scientifiques qui ont été testés sont : WorldWideScience, Science.gov, Refseek, Educational Resources Information Center, ScienceDirect, Google Scholar et ResearchGate. Après des recherches sur chaque site avec les mots-clés "Industry 4.0", "Data science industry 4.0", "Machine learning industry 4.0", "Industry 4.0 performance", "Machine learning industry 4.0 maturity" il est possible de savoir si certains moteurs de recherche sont plus pertinents que d'autres à propos du sujet étudié. La TABLE 3.1 détaille la pertinence des différents moteurs de recherche en fonction des mots-clés testés. La pertinence d'un moteur de recherche est déterminée sur base de l'existence d'un rapport entre les articles présentés sur la première page de recherche pour un mot-clé ou un groupe de mots-clés donné et dans ce cas, les titres sont analysés. Si moins de la moitié des articles présents sur la première page ont un intérêt dans l'analyse de ce mémoire, le moteur de recherche est classifié non pertinent pour le groupe de mots-clés utilisé. Pour définir si un moteur de recherche est pertinent de manière générale, il faut qu'il soit pertinent pour plus de la moitié des différents mots-clés. Il s'avère que les moteurs de recherche WorldWideScience, Science.gov et ERIC.ed.gov ont des résultats relativement peu pertinents pour la majorité des mots-clés. Finalement, pour la recherche d'articles scientifiques, les moteurs de recherche utilisés sont donc : Refseek, ScienceDirect, Google Scholar et ResearchGate, ce sont ceux qui ont donné le plus de résultats pertinents en général.

		Moteurs de recherche						
Mots-clés		WorldWideScience	Science.gov	Refseek	ERIC.ed.gov	ScienceDirect	Google Scholar	ResearchGate
		Industry 4.0					✓	✓
	Data science industry 4.0			✓		✓	✓	✓
	Machine learning industry 4.0	✓		✓		✓	✓	✓
	Industry 4.0 performance	✓		✓		✓	✓	✓
	Machine learning industry 4.0 maturity			✓		✓	✓	✓

**TABLE 3.1** – Pertinence des résultats des moteurs de recherche par groupe de mots-clés

### 3.2 Sélection d’articles scientifiques

Pour qu’un article scientifique soit sélectionné, il faut que l’article pris en compte ait soit un lien avec au moins un des autres sujets du mémoire (machine learning, data science, industrie 4.0, performance de l’industrie 4.0, maturité du ML et DS, etc.), soit un lien direct avec le sujet du mémoire. Si l’article scientifique n’a aucun lien avec un autre sujet ou avec le sujet du mémoire, il n’est pas pris en compte. La TABLE 3.2 présente les critères d’inclusion et d’exclusion des articles scientifiques pour l’analyse. L’existence ou non de ce lien fait l’objet d’un choix subjectif après une lecture du résumé de l’article. Quand un article est sélectionné, il est alors lu en entier. Tant qu’assez d’informations n’ont pas été récoltées à propos d’un sujet, la recherche d’articles scientifiques continue. Il faut noter que tous les articles scientifiques trouvés via les différentes recherches n’ont pas été pris pour chacun des mots-clés, étant donné que certains articles donnés dans les résultats des moteurs de recherche n’apportent pas plus d’informations pour l’analyse.

Critères d’inclusion	Critères d’exclusion
L’article est accessible gratuitement	L’article n’est pas accessible gratuitement
L’article traite 2 sous-thèmes du mémoire ou plus	L’article traite uniquement un des sous-thèmes du mémoire
L’article est rédigé en anglais ou en français	L’article n’est pas rédigé en anglais ou en français

**TABLE 3.2** – Critères d’inclusion et d’exclusion des articles scientifiques pour l’analyse

### **3.3 Analyse d'articles scientifiques**

L'analyse des articles scientifiques consiste en la lecture intégrale de l'article. Ensuite, tous les éléments de l'article qui ont un rapport avec le sujet de ce mémoire sont retenus pour la synthèse de tous les articles lus. Le résultat de l'analyse est une synthèse et une critique de la littérature scientifique à propos du sujet du mémoire. Pour la synthèse, les idées communes de tous les articles sont regroupées, ce qui produit un résumé. Chaque article est mis en perspective au regard des autres articles pour pouvoir les critiquer, noter les différences ou les incohérences.

Pendant la phase d'analyse, il se peut que certains articles scientifiques ne soient finalement pas pertinents. Ces articles sont marqués comme non retenus. Le critère de non-rétention de ces articles est relativement subjectif, mais fait preuve d'un raisonnement qui peut être détaillé comme suit. Si la lecture de l'article n'apporte pas de nouvelle connaissance, il est immédiatement marqué comme non retenu. Si l'article apporte de la nouvelle connaissance à propos du domaine d'intérêt général mais que celle-ci se trouve être hors sujet pour ce mémoire, l'article est marqué comme non retenu. Si l'article apporte de l'information à propos du sujet du mémoire, mais que l'information a déjà été obtenue dans un autre article lu précédemment, l'article est marqué comme non retenu. Cette dernière condition est la plus soumise à la subjectivité : si l'article apporte de la connaissance à propos du sujet du mémoire, mais que l'information n'a que partiellement été lue dans un autre article, il faut alors juger si l'information va apporter de la valeur à l'analyse. Si non, l'article est marqué comme non retenu. Tous les articles dupliqués sont aussi marqués comme non retenus.

### **3.4 Recherche de données chiffrées**

La recherche de données chiffrées est réalisée sur le portail en ligne Statista. Les groupes de mots-clés utilisés pour la recherche sont : "Industry 4.0", "Data science industry 4.0", "Machine learning industry 4.0" et "Industry 4.0 Performance". Les résultats sont ensuite analysés et seules les données pertinentes pour répondre aux questions de recherche sont retenues.

### **3.5 Recherche d'articles complémentaires**

Afin de définir et d'expliquer certains concepts au moment de l'analyse, des articles scientifiques complémentaires seront choisis pour apporter une définition et une explication scientifique aux concepts. Ces articles sont trouvés sur les moteurs de recherche cités à la section concernant la recherche d'articles scientifiques. En revanche, une liste de mots-clés n'est pas établie dans cette situation, étant donné que la recherche de définitions

est trop spécifique et doit être adaptée à chaque concept. Par conséquent, des recherches personnalisées sont faites pour expliquer les concepts qui nécessitent une clarification.

## 4 Analyse des données et résultats

Cette section concerne l'analyse des articles scientifiques récoltés. Il s'agit d'une analyse qualitative basée sur la lecture de ces articles.

La partie qui concerne l'analyse de données présente brièvement les articles scientifiques qui ont été sélectionnés ainsi qu'une description plus poussée des articles qui ont été retenus. Cette partie présente aussi les données récoltées et une brève description de celles-ci.

La partie résultats est la partie centrale de ce mémoire. C'est une synthèse et une critique de la littérature concernant l'utilité, l'impact en termes de performance et la maturité du machine learning et des data sciences dans l'industrie 4.0. Cette partie permet d'avoir la connaissance nécessaire du sujet pour formuler une réponse aux questions de recherche posées dans la section d'introduction.

### 4.1 Analyse des données

Cette sous-section présente les données collectées, c'est-à-dire, les articles scientifiques qui ont été sélectionnés, ceux qui ont été retenus pour l'analyse ainsi que les données chiffrées qui ont été trouvées.

#### 4.1.1 Articles scientifiques sélectionnés

Au cours du processus de sélection d'articles scientifiques, 60 articles ont été sélectionnés au total. Le nombre d'articles sélectionnés par groupe de mots-clés est détaillé à la TABLE 4.1. Dans tous ces articles, il y en a qui n'apportent pas de nouvelles connaissances à propos du sujet du mémoire et il y en a aussi qui sont redondants les uns par rapport aux autres en ce qui concerne leur contenu. Il est maintenant nécessaire de faire un tri grâce à la méthode détaillée à la section 3.3.

Mots-clés par catégorie	Nombre d'articles sélectionnés
<b>Industrie 4.0</b>	<b>7</b>
— Industry 4.0	4
— Industrial internet of things	3
<b>Data science</b>	<b>21</b>
— Data science industry 4.0	5
— Data analytics industry 4.0	4
— Big data industry 4.0	7
— Data mining industry 4.0	2
— Data visualization industry 4.0	3
<b>Machine learning</b>	<b>26</b>
— Machine learning industry 4.0	19
— Artificial neural network industry 4.0	5
— Prediction industry 4.0	2
<b>Industrie 4.0 Performance</b>	<b>10</b>
— Industry 4.0 performance	5
— Industry 4.0 performance increase	4
— Industry 4.0 performance increase firm	1
<b>Maturité du ML et de la DS</b>	<b>4</b>
— Machine learning industry 4.0 maturity	1
— Data science industry 4.0 maturity	1
— Artificial intelligence industry 4.0 maturity	2
<b>Total</b>	<b>68</b>

**TABLE 4.1** – Table récapitulative du nombre d'articles sélectionnés par mot-clé

#### 4.1.2 Articles scientifiques retenus

Après la suppression des doublons et des articles qui n'apportent pas de plus-value à l'analyse, il y a maintenant 31 articles à analyser. Le nombre d'articles retenus par groupe de mots-clés est détaillé à la TABLE 4.2. Le nombre d'articles retenus est inférieur au nombre d'articles sélectionnés. De plus, une partie des articles qui ont été trouvés avec des mots-clés plus globaux ont aussi été classifiés comme non retenus. Cela est dû au fait que l'information trouvée n'était pas assez pertinente pour le sujet de ce travail ou que l'information était trop générique, pas assez précise ou même hors sujet pour répondre aux questions de recherche.

Il est aussi intéressant de voir en quelle année les articles retenus ont été publiés. Les données sont présentées sur la FIGURE 4.1. Il n'y a aucun article retenu qui a été publié avant 2016. Il est possible de voir qu'il y a une forte montée en popularité de

ce sujet entre 2017 et 2019. Il y a malgré tout une baisse d'intérêt pour le sujet entre 2019 et 2021. Ces observations indiquent que le sujet est encore relativement récent. Cependant, ces statistiques ne sont pas forcément représentatives de la réalité au niveau de l'intérêt pour ce sujet étant donné que cela n'est pas une recherche systématique et que l'échantillon ne comporte pas un grand nombre d'articles, il ne peut donc pas être suffisamment représentatif.

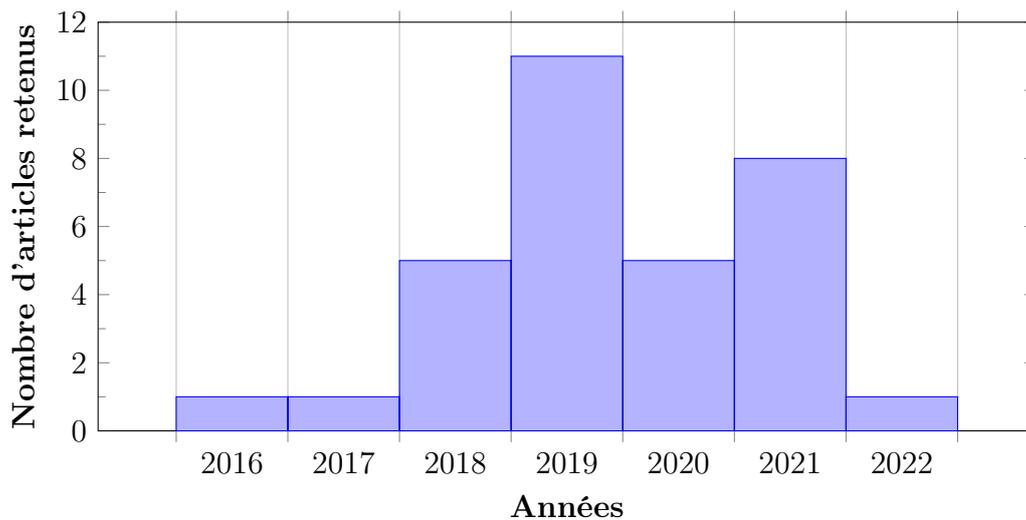
Il est aussi d'usage de vérifier dans quels journaux scientifiques les différents articles ont été publiés. La liste des journaux ainsi que le nombre d'articles venant de chaque journal sont détaillés à la TABLE 4.3. Ce sont majoritairement des journaux et des collections d'articles de conférence portant sur l'ingénierie, la production et les technologies.

<b>Mots-clés par catégorie</b>	<b>Nombre d'articles retenus</b>
<b>Industrie 4.0</b>	<b>1</b>
— Industry 4.0	1
— Industrial internet of things	0
<b>Data science</b>	<b>7</b>
— Data science industry 4.0	2
— Data analytics industry 4.0	1
— Big data industry 4.0	1
— Data mining industry 4.0	0
— Data visualization industry 4.0	3
<b>Machine learning</b>	<b>15</b>
— Machine learning industry 4.0	13
— Artificial neural network industry 4.0	1
— Prediction industry 4.0	1
<b>Industrie 4.0 Performance</b>	<b>7</b>
— Industry 4.0 performance	4
— Industry 4.0 performance increase	2
— Industry 4.0 performance increase firm	1
<b>Maturité du ML et de la DS</b>	<b>1</b>
— Machine learning industry 4.0 maturity	0
— Data science industry 4.0 maturity	0
— Artificial intelligence industry 4.0 maturity	1
<b>Total</b>	<b>31</b>

**TABLE 4.2** – Table récapitulative du nombre d'articles retenus par mot-clé

Nom du journal scientifique	Nombre d'articles
<b>Journaux scientifiques :</b>	<b>25</b>
Journal of Manufacturing Technology Management	2
Sensors	2
Annals of Operations Research	1
Computers in Industry	1
Data & Knowledge Engineering	1
Enterprise Information Systems	1
Expert Systems with Applications	1
Frontiers of Mechanical Engineering	1
IEEE Internet of Things Journal	1
IFAC-PapersOnLine	1
Information Fusion	1
International Journal of Information Management	1
International Journal of Operations & Production Management	1
International Journal of Production Economics	1
International Journal of Production Research	1
Journal of Intelligent Manufacturing	1
Journal of Sensor and Actuator Networks	1
Materials Today : Proceedings	1
Procedia CIRP	1
Procedia computer science	1
Procedia Manufacturing	1
Progress in Industrial Mathematics at ECMI 2018	1
Technological Forecasting & Social Change	1
<b>Conférences :</b>	<b>6</b>
2016 International Conference on Computational Science and Computational Intelligence	1
2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications	1
2019 International conference on computational intelligence and knowledge economy	1
Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing	1
Annual International Conference on Industrial Engineering and Operations Management	1
Knowledge Management in Organizations : 13th International Conference	1
<b>Total</b>	<b>31</b>

**TABLE 4.3** – Table récapitulative des journaux de publication des articles



**FIGURE 4.1** – Histogramme des années de publication des articles retenus

### 4.1.3 Données chiffrées sélectionnées

Plusieurs données ont été trouvées sur Statista. Le premier graphique montre les principales raisons qu'ont les entreprises à implémenter des projets d'industrie 4.0 en 2022 (Laricchia, 2022). Ces données ont été récoltées auprès de 55 répondants en Italie, à propos d'entreprises italiennes, pour l'année 2022, entre février et mars. Le deuxième graphique montre quelles technologies sont adoptées pour automatiser des entrepôts en 2019 (Placek, 2022). Le graphique propose aussi une prédiction de l'adoption de ces technologies pour l'année 2030. Ces données ont été récoltées en 2019 au niveau mondial, il n'y a malheureusement pas plus d'informations quant à la provenance des données de ce deuxième graphique. Deux graphiques montrent le taux d'adoption de l'intelligence artificielle dans les entreprises de développement de produits et dans les entreprises de fabrication en 2022 et en prévision de 2025 (Statista, 2023). Ces deux graphiques sont construits sur base de 600 réponses à un questionnaire distribué au niveau mondial. Le cinquième graphique est un visuel qui montre les pays leaders dans l'industrie 4.0 (Statista Research Department, 2016). Cette étude a été menée en 2016 auprès de 559 entreprises de plus de 100 personnes en leur posant la question "Quelle nation, selon vous, est actuellement à la pointe de l'industrie 4.0?". Le dernier graphique sélectionné montre le taux de préparation à l'intelligence artificielle dans plusieurs pays du monde en 2018 (Capgemini, 2018).

## 4.2 Résultats

Cette section est destinée à l'analyse, c'est-à-dire la synthèse et la critique des articles scientifiques récoltés après leur lecture et leur mise en perspective avec les autres articles.

### 4.2.1 Mise en pratique du machine learning et de la data science dans l'industrie 4.0

Le machine learning et la data science ont beaucoup d'applications concrètes dans l'environnement de l'industrie 4.0. Il est ici question de citer ces applications, de les expliquer et de donner l'utilité des techniques de machine learning et de data science utilisées pour ces applications.

Le machine learning et la data science sont utilisés dans beaucoup d'industries de différents types. L'article de Balamurugan et al. (2019) cite les différents cas d'utilisation de l'intelligence artificielle (IA). L'IA peut être utilisée dans un cadre marketing pour faire des recommandations aux clients ou améliorer leur expérience durant leurs achats (Balamurugan et al., 2019). L'IA permet aussi d'augmenter l'efficacité des industries en rendant des robots autonomes, en permettant une gestion automatique des équipements de production ou en aidant les dirigeants à prendre des décisions, l'IA aide aussi à optimiser les systèmes de transports et de gestion de l'énergie (Balamurugan et al., 2019). L'IA est aussi utile dans l'environnement financier, elle peut aider à détecter des fraudes ou à scanner des flux monétaires (Balamurugan et al., 2019). Finalement, en matière de cybersécurité, l'IA peut détecter les activités et les schémas inhabituels en matière de données (Balamurugan et al., 2019). Dans la suite de l'analyse, l'industrie de fabrication et de production sera focalisée.

Les applications du machine learning et de la data science dans l'industrie 4.0 sont multiples. L'article de Cimini et al. (2017) cite différentes améliorations qu'apporte l'industrie 4.0 par rapport à des méthodes de production classiques. Instaurer l'industrie 4.0 permet d'améliorer la qualité des produits et l'efficacité de la production. Cela permet aussi de faire de la maintenance prédictive ainsi que d'optimiser la chaîne d'approvisionnement (Cimini et al., 2017). L'article de Jagatheesaperumal et al. (2021) apporte des précisions à propos de l'utilisation du machine learning pour ces applications. Il confirme notamment l'utilisation du machine learning pour la maintenance prédictive, l'analyse de la qualité et la chaîne d'approvisionnement. Pour cela, des techniques de machine learning classiques sont employées (machines à vecteurs de support, arbres de décision, régression logistique, k-nearest neighbor) ainsi que des techniques de deep learning (Jagatheesaperumal et al., 2021).

L'analyse des données passées de l'entreprise peut leur permettre de pouvoir faire diverses prédictions, c'est ce qui est appelé l'analyse prédictive (Candanedo et al., 2018). Les technologies impliquées dans l'analyse prédictive vont des plus simples aux plus complexes. Il est possible d'utiliser des modèles de régression linéaire, les arbres de décisions, des statistiques bayésiennes, des réseaux de neurones artificiels ou encore des machines à vecteurs de support (Xu & Duan, 2019). Ces modèles transforment un ensemble de don-

nées historiques de l'entreprise où la valeur cible est précisée en une collection de règles pour faire une prédiction sur base de nouvelles données sans valeur cible (Romeo et al., 2020). Le but est de prédire ce qu'il va se passer pour faire des économies. De plus, les entreprises peuvent prendre des décisions sur base des prédictions faites par le modèle de machine learning (Paolanti et al., 2018).

Une branche plus spécifique de l'analyse prédictive est la maintenance prédictive. La maintenance prédictive est une partie importante de l'industrie 4.0. Chaque machine tout au long du processus de production a une durée de vie limitée et a besoin d'entretiens réguliers pour rester performante et ne pas causer de perte de qualité des produits (Dalzochio et al., 2020). Pour cela, il faut pouvoir récolter des données sur l'état de l'équipement et il faut aussi pouvoir interpréter ces données (Dalzochio et al., 2020). Pour les modèles de diagnostic prévisionnel, ce sont les SVM (machines à vecteurs de support) et les ANN (réseaux de neurones artificiels) qui sont le plus souvent utilisés pour analyser les données (Diez-Olivan et al., 2019). Ces modèles sont utiles puisqu'ils peuvent permettre de faire la distinction entre une situation normale et anormale et donc de détecter les problèmes (Diez-Olivan et al., 2019). Les séries temporelles sont aussi un moyen convaincant de mettre en place un système de maintenance prédictive (Langone, Cuzzocrea, & Skantzos, 2020). L'article de Langone et al. (2020) montre un cas réel d'application d'une série temporelle aux données récoltées de divers capteurs d'une machine. Ils obtiennent une prédiction avec très bonne précision jusqu'à une heure avant l'occurrence de l'incident d'une machine (Langone et al., 2020). Cependant, ces méthodes ne fonctionnent pas avec tous les types de machines. Il y en a certaines qui ont une probabilité de ne plus fonctionner correctement à n'importe quel moment sans signes préventifs. Les algorithmes ne s'appliquent pas dans ce cas-ci car les capteurs n'émettent pas de signal précoce (Langone et al., 2020).

Une autre application des technologies de machine learning est la détection de défauts dans la production (Rai et al., 2021). Détecter les produits défectueux permet de limiter les pertes liées au renvoi des produits mais aussi celles liées au fait qu'un produit défectueux à l'origine va continuer dans tout le cycle de production inutilement (Villalba-Diez et al., 2019). L'article de Villalba-Diez et al. (2019) dit que pour effectuer du contrôle de la qualité optique, une méthode intéressante est d'utiliser des deep neural network. Ce sont des algorithmes de deep learning (branche du machine learning) qui permettent de traiter presque n'importe quel type de données en apprenant par eux-mêmes (Ahmed et al., 2023). Ceux-ci permettent de détecter des défauts par l'analyse d'images prises au cours du processus de production (Villalba-Diez et al., 2019). Les algorithmes de machine learning (réseaux de neurones artificiels, deep learning, réseaux de neurones convolutifs) sont de plus en plus utilisés pour détecter des défauts dans la production ainsi que les problèmes relatifs à la fiabilité des machines et des outils (Angelopoulos et al., 2019).

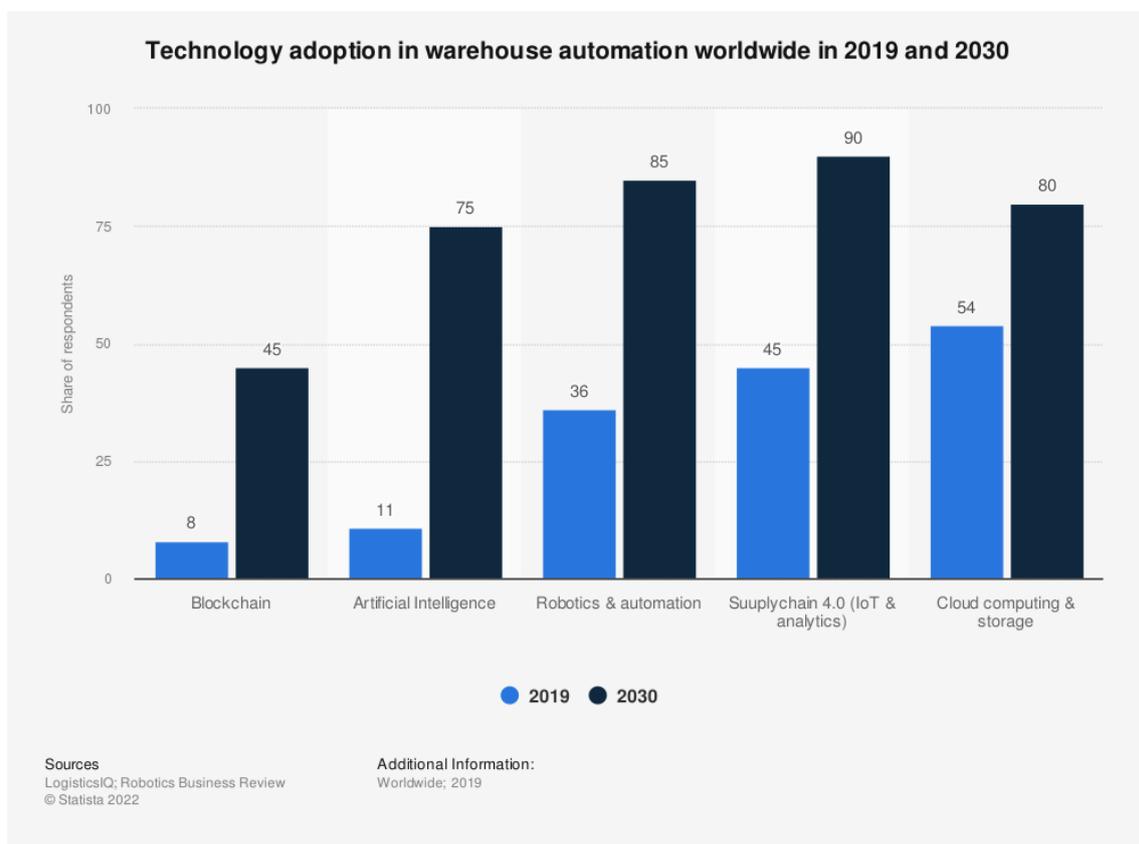
La gestion et le contrôle de la chaîne d’approvisionnement constituent un enjeu majeur pour toutes les entreprises, quelles qu’elles soient (Nagar et al., 2021). Il est intéressant pour elles de l’optimiser le plus possible afin de profiter de délais plus courts et/ou d’une meilleure gestion du temps et des matériaux (Heydarbakian & Spehri, 2022). Cavalcante et al. (2019) proposent une méthode de machine learning qui utilise différents algorithmes tels que le k-nearest neighbor (kNN) et la régression logistique pour prédire la résilience des fournisseurs. Ils montrent qu’un modèle de machine learning sans biais permet de choisir des fournisseurs sur base des résultats de celui-ci (Cavalcante et al., 2019). L’article de Heydarbakian et Spehri (2022) a un but très similaire, c’est-à-dire classer les fournisseurs en fonction de leur performance. Ils utilisent des algorithmes bayés naïfs et des arbres de décision. Les algorithmes parviennent à classer les fournisseurs correctement dans plus de 80% des cas, ce qui confère un avantage pour les entreprises en matière de sélection de fournisseurs (Heydarbakian & Spehri, 2022).

La planification de la production est le fait de prévoir les quantités nécessaires de matériaux pour pouvoir produire (Usuga Cadavid et al., 2020). C’est un enjeu essentiel pour les entreprises. Il leur est nécessaire de décider des besoins de leurs différentes installations (entrepôts, chaînes de production, logistique) en ce qui concerne les matériaux nécessaires, l’espace disponible, les coûts liés aux installations, etc. (Zäpfel & Missbauer, 1993). Usuga Cadavid et al. (2020) disent que de manière classique, la planification de la production se fait via un ensemble de données qui sont ensuite analysées de manière systématique et utilisées dans des simulations qui ne peuvent pas prendre en compte toutes les variables. Pour résoudre ce problème, il est envisageable d’utiliser des modèles de machine learning pour améliorer la planification de la production (Usuga Cadavid et al., 2020). Les réseaux de neurones artificiels à plusieurs couches peuvent être utilisés dans ce cas étant donné la complexité de la réalité (Almeida, 2002) que même des simulations ont du mal à prendre en compte (Gunawan et al., 2021). Une entreprise pourrait récolter et stocker des données historiques, dans ce cas, les variables nécessaires pour avoir une représentation de l’environnement et ainsi prédire la quantité d’actifs à prévoir dans la production (Usuga Cadavid et al., 2020).

Toutes les applications vues jusqu’à présent permettent de prendre des décisions opérationnelles de manière automatique ou presque automatique, faites dans le cadre d’opérations au jour le jour à des niveaux inférieurs de l’organisation et qui n’ont qu’un impact à court terme (Angelopoulos et al., 2019). Il existe aussi des méthodes de visualisation qui permettent aux opérateurs de prendre ce type de décisions par eux-mêmes (Lodgaard & Dransfeld, 2020). Les informations capturées par les capteurs sont directement affichées sur des tableaux de bords dynamiques qui comportent des indicateurs clés de performance en temps réel (Zheng et al., 2018). Grâce à cela, les opérateurs peuvent observer si la performance d’une machine diminue ou s’il faut prévoir un entretien prochainement

(Shafiq, Szczerbicki, & Sanin, 2019). Selon Shafiq et al. (2019), ceci permet de réagir plus rapidement aux changements et aux imprévus .

Actuellement, les applications du machine learning et de la data science sont assez peu répandues dans les entreprises. Les résultats d’une enquête (Placek, 2022), disponibles sur la FIGURE 4.2, montrent que 45% des répondants disent utiliser la chaîne d’approvisionnement 4.0 en 2019, ce qui représente tout de même une partie non négligeable des entreprises. En revanche, l’intelligence artificielle était utilisée par seulement 11% des répondants. Les prévisions de Placek (2022) pour l’année 2030 sont de 75% pour l’intelligence artificielle et de 90% pour la chaîne d’approvisionnement 4.0. Cela représenterait une évolution de 582% et de 100%, respectivement.



**FIGURE 4.2** – Adoption des technologies d’automatisation des entrepôts dans le monde en 2019 et 2030 (Placek, 2022)

#### 4.2.2 Intérêts en termes de performance de passer à l’industrie 4.0

L’industrie 4.0 a comme pour objectif d’améliorer les performances des entreprises. Il est question de détailler l’impact au niveau des performances de l’implémentation de l’industrie 4.0 pour les entreprises.

Mettre en place l’industrie 4.0 dans les entreprises leur est globalement bénéfique en termes de performances (Tortorella, Giglio, & Van Dun, 2019). L’étude de Tortorella et

al. (2019) stipule que certains secteurs bénéficient significativement de l'implémentation de l'industrie 4.0, ce sont plus particulièrement les industries où le lean management et la lean production sont très utilisés. Il existe une corrélation positive entre l'introduction de technologies de l'industrie 4.0 et la performance des processus de production (Tortorella et al., 2019). En revanche, l'industrie 4.0 n'a pas d'impact bénéfique significatif sur les tâches plus créatives comme le développement de produits ou l'innovation. Pour généraliser les résultats de l'étude de Tortorella et al. (2019), avec l'introduction de l'industrie 4.0 ce sont les tâches qui impliquent des matériaux, physiques ou virtuels, qui voient leur performance augmenter alors que les tâches dites d'information ne subissent pas d'augmentation de performance. Cependant, Tortorella et al. (2019) mentionnent qu'implémenter des techniques d'industrie 4.0 n'améliorera pas des processus qui ne fonctionnent pas à l'origine.

Toutefois, les résultats de Tortorella et al. (2019) sont à mettre en perspective avec ceux de Lin, Wu, et Song (2019) qui disent que la stratégie d'implémenter l'industrie 4.0 fait augmenter le nombre de demandes de brevets, et par conséquent, cela fait augmenter le comportement d'innovation des entreprises (Lin et al., 2019). Cette différence peut venir du fait que l'étude de Tortorella et al. (2019) est faite sur des entreprises au Brésil et que l'étude de Lin et al. (2019) porte sur des entreprises chinoises. De plus, l'étude de Tortorella et al. (2019) utilise des valeurs issues de questionnaires alors que celle de Lin et al. (2019) reprend des valeurs quantifiables objectivement, cela peut être une autre explication de l'incohérence.

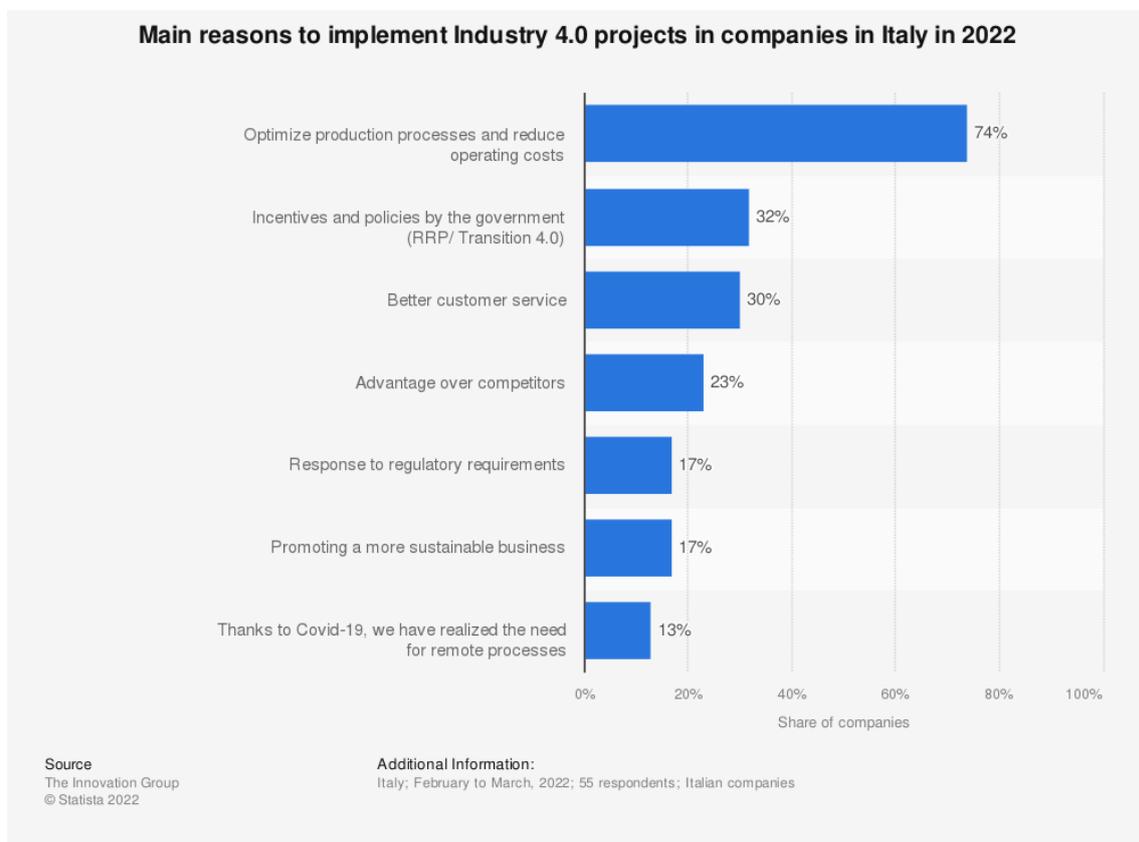
Un autre aspect qui peut être vu comme de la performance est l'amélioration de la qualité globale des produits. Une des conclusions d'une étude réalisée sur des entreprises turques est "l'un des principaux avantages de l'I4.0 est l'augmentation de la qualité des produits" (Duman & Akdemir, 2021). Cela se traduit par un avantage concurrentiel pour les entreprises. Par conséquent, l'investissement dans l'industrie 4.0 peut être considéré comme une stratégie de performance pour les entreprises cherchant à améliorer la qualité de leurs produits et à rester compétitives (Dos Santos et al., 2021).

En plus de la performance opérationnelle, l'industrie 4.0 permet d'améliorer les performances financières des entreprises. Lin et al. (2019) démontrent que "la mise en œuvre de l'industrie 4.0 a un impact positif sur le rendement des capitaux propres des entreprises" en conduisant une étude sur des entreprises chinoises (Lin et al., 2019). Lin et al. (2019) montrent aussi que le gain de performances financières est plus élevé dans les petites et moyennes entreprises que dans les grandes entreprises étant donné qu'instaurer des nouvelles technologies demande un changement de processus, ce qui est plus facile à réaliser dans les petites compagnies.

L'article de Szász et al. (2021) résume bien la situation de l'industrie 4.0 et ses effets

positifs sur les entreprises. Il est dit qu'instaurer les technologies de l'industrie 4.0 sert à "améliorer les quatre dimensions traditionnelles de la performance opérationnelle (coût, qualité, livraison et flexibilité)" (Szász et al., 2021).

L'impact positif que l'industrie 4.0 a sur les performances des entreprises est corroboré par les compagnies qui disent que le principal intérêt pour elles d'implémenter l'industrie 4.0 est l'optimisation des processus et la réduction des coûts (Laricchia, 2022). Il est possible de voir à la FIGURE 4.3 que cet intérêt est désirable pour près de trois quarts des entreprises en Italie, le deuxième étant les politiques gouvernementales qui ne représentent que 32% et le troisième étant un meilleur service client pour 30% des répondants.



**FIGURE 4.3** – Principales raisons de mettre en œuvre des projets d'industrie 4.0 dans les entreprises en Italie en 2022 (Laricchia, 2022)

### 4.2.3 Amélioration des performances de l'industrie 4.0 grâce au machine learning et à la data science

En combinant les résultats obtenus aux sections 4.2.1 et 4.2.2, il devient possible de formuler des interprétations relatives à l'effet bénéfique de l'intégration du machine learning et de la data science dans le contexte de l'industrie 4.0. Ces interprétations mettent en évidence l'amélioration des performances ainsi que la réduction des coûts associés à cette mise en œuvre. Cette section concerne les implications managériales de

l'implémentation des techniques de machine learning et de data science dans l'industrie 4.0.

Il est possible de remarquer que beaucoup d'améliorations de la performance grâce à l'industrie 4.0 se font au niveau des processus de production et de la chaîne de valeur (Tortorella et al., 2019). Les entreprises essaient toujours d'optimiser leur production. Une des considérations est le contrôle de la qualité (Coleman, 2019). Pour cela, elles peuvent implémenter du machine learning (dont du deep learning) pour effectuer des contrôles de qualité fréquents tout au long d'une chaîne de production simplement grâce à des caméras et/ou des capteurs (Villalba-Diez et al., 2019). Les images sont généralement traitées par des réseaux de neurones dont les méthodes de réseaux de neurones artificiels ou convolutifs (Ahmed et al., 2023). Après le traitement, les résultats sont interprétés par la machine de tri ou par un opérateur et les produits défectueux peuvent être éliminés automatiquement ou manuellement (Rai et al., 2021) (Benbarrad et al., 2021). Cela fait gagner du temps et le travail est très qualitatif, pour peu que les algorithmes soient bien entraînés (Villalba-Diez et al., 2019).

Pour que le processus de production soit le plus fiable possible, il est intéressant pour les entreprises de choisir les meilleurs fournisseurs (Cavalcante et al., 2019). Pour ce faire, il est possible d'analyser leurs performances à la main. Une méthode plus robuste consiste à utiliser des algorithmes de machine learning, comme le kNN, la régression logistique (Cavalcante et al., 2019) ou même des arbres de décision (Heydarbakian & Spehri, 2022) pour sélectionner les meilleurs fournisseurs (Cavalcante et al., 2019). Cela permet de diminuer les risques liés aux retards de livraisons, à la livraison produits défectueux, à la réception d'une quantité inexacte (Szász et al., 2021). Améliorer la résilience des fournisseurs leur confère donc un avantage en termes de régularité de la part des fournisseurs, ce qui induit de la régularité dans la production (Tortorella et al., 2019). L'utilisation de techniques de machine learning et de data science donne aux entreprises l'opportunité d'améliorer leur productivité en ce qui concerne les temps de production et la constance de la production (Dalenogare et al., 2018).

En plus du choix optimal des fournisseurs, les entreprises peuvent planifier leur production via des systèmes autonomes. Ces systèmes sont simplement un assemblage de capteurs, d'objets connectés et d'algorithmes qui prennent en compte une multitude de variables pour en sortir des décisions relatives à la gestion des flux de production (Usuga Cadavid et al., 2020). L'utilisation de réseaux de neurones artificiels (Gunawan et al., 2021) permet aux entreprises de créer un modèle de la réalité plus performant qu'une simple simulation pour prévoir les flux de matériaux, le stockage et la logistique (Usuga Cadavid et al., 2020). Cela amène les entreprises vers un système se rapprochant des méthodes de lean manufacturing sans avoir besoin de l'organiser manuellement

(Tortorella et al., 2019). Optimiser la planification de la production a pour objectif managérial de réduire les temps morts dû à des délais provoqués par des processus vieillissants et de moins en moins optimaux avec l'arrivée de technologies (Tortorella et al., 2019). En réduisant les temps morts, la production peut être plus rapide et constante, ce qui maximise leur rentabilité et par conséquent leur profit (Tortorella et al., 2019). Cependant, le machine learning n'est pas une méthode qui va transformer les flux de production, les algorithmes vont seulement optimiser les processus existants. Comme le mentionne l'étude de Tortorella et al. (2019), les processus ne sont pas transformés mais simplement optimisés.

Toujours dans le but de minimiser les temps morts causés par des problèmes soudains, les entreprises ont la possibilité de faire de la maintenance prédictive (Szász et al., 2021). Une fois de plus, ce sont les réseaux de neurones artificiels qui sont utilisés pour analyser les données de capteurs et ainsi générer des modèles sur base des données existantes et faire des prédictions (Diez-Olivan et al., 2019). Ces prédictions sont ensuite utilisées pour savoir à quel moment les opérateurs doivent faire la maintenance des machines (Dalzochio et al., 2020). L'application du machine learning et de la science des données pour la maintenance préventive permet d'éviter les risques de sous-entretien ou de sur-entretien des équipements, ce qui est souvent fait pour garder la même fiabilité que si les machines étaient entretenues périodiquement sans les laisser tomber en panne (Diez-Olivan et al., 2019). En prévoyant à quel moment une machine tombe en panne ou perd en performance, les opérateurs peuvent effectuer la maintenance au moment opportun (Dalzochio et al., 2020) et ainsi éviter de perdre de l'argent dans des coûts de maintenance trop réguliers. Les opérateurs peuvent également éviter que les machines soient inactives pendant un certain temps, ce qui pourrait réduire la productivité et pourrait freiner la production globale (Paolanti et al., 2018).

L'implémentation des technologies de machine learning et de data science dans l'industrie 4.0 permet aussi d'améliorer les performances financières des entreprises (Lin et al., 2019). L'analyse de données historiques d'une entreprise permet de prédire l'occurrence de certains événements dans le futur (Jagatheesaperumal et al., 2021). Les modèles de régression linéaire, les arbres de décisions et les réseaux de neurones artificiels qui sont le plus souvent employés dans ce cas (Xu & Duan, 2019). Il est assez difficile de donner des exemples génériques d'analyse prédictive en entreprise, chaque entreprise dispose de ses données propres et peut faire des prédictions sur les ventes, la consommation électrique, les retours de produits, les essais demandés, etc. (Candanedo et al., 2018). Les entreprises doivent adapter les modèles de machine learning à leurs besoins. Les prédictions permettent alors aux entreprises de s'adapter aux situations qu'elles vont probablement rencontrer comme acheter plus de stock, prévoir du personnel, répartir la demande (Candanedo et al., 2018). Cela leur permet de faire le moins d'erreurs de

calcul possible et de minimiser les pertes d'argent (Tortorella et al., 2019). D'un autre côté, l'analyse prescriptive utilise des techniques de data science pour faire des inférences entre les événements passés et des résultats pour déduire des règles entre des actions et leurs conséquences (Diez-Olivan et al., 2019). Ces règles sont des prescriptions de ce que doivent faire les entreprises, les actions, pour atteindre certains objectifs, les conséquences (Jagatheesaperumal et al., 2021).

Pour aider les opérateurs qui sont disposés à des postes stratégiques au long de la production, il est nécessaire de leur donner des informations relatives à l'état des machines et des outils de production (Shafiq et al., 2019). Les méthodes de visualisation permettent d'améliorer les performances financières des entreprises en donnant des informations clés le plus tôt possible aux opérateurs (Lodgaard & Dransfeld, 2020). Au plus vite ils peuvent réagir à des problèmes ou à des alertes, au moins il y aura de conséquences de délais de production et donc des avantages de performance et financiers non négligeables (Shafiq et al., 2019).

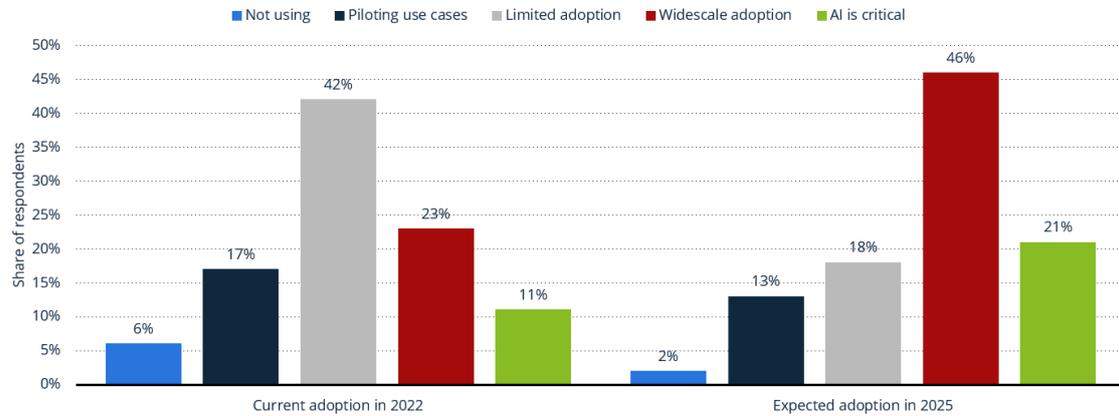
#### **4.2.4 Maturité du machine learning et de la data science dans l'industrie 4.0**

Cette section concerne la mesure de la maturité de l'implémentation des technologies de machine learning et de data science dans l'industrie 4.0. Dans cette section, il est question de prendre en compte certains articles scientifiques mais majoritairement des données chiffrées pour estimer le développement de ces technologies.

Selon les données de Statista (2023), il est possible de remarquer qu'en 2022, au niveau mondial, le taux d'adoption de l'intelligence artificielle ("adoption à grande échelle" et "l'IA est critique") était de 34% pour les entreprises de développement de produits (FIGURE 4.4) et de 45% pour les entreprises de fabrication (FIGURE 4.5) (Statista, 2023). En revanche, en 2025, toujours au niveau mondial, le taux d'adoption attendu de l'intelligence artificielle serait de 67% pour les entreprises de développement de produits et de 68% pour les entreprises de fabrication (Statista, 2023). Cela représente une évolution d'environ 97% pour les entreprises de développement de produits et de 51% pour les entreprises de fabrication. Il est aussi important de noter que le taux d'utilisation de l'intelligence artificielle quand l'IA est critique passe de 11% en 2022 à 38% attendus en 2025 pour les entreprises de fabrication (Statista, 2023).

## Artificial intelligence (AI) adoption rate in product development businesses worldwide in 2022 and 2025

AI adoption rate in global product development 2022-2025



Description: The adoption rate of artificial intelligence (AI) is expected to gain considerable importance in product development companies worldwide between 2022 and 2025. Currently, companies operating in that sector were mostly reporting limited adoption of AI in their production cycles. Technology executives expected this to change considerably by 2025. [Read more](#)

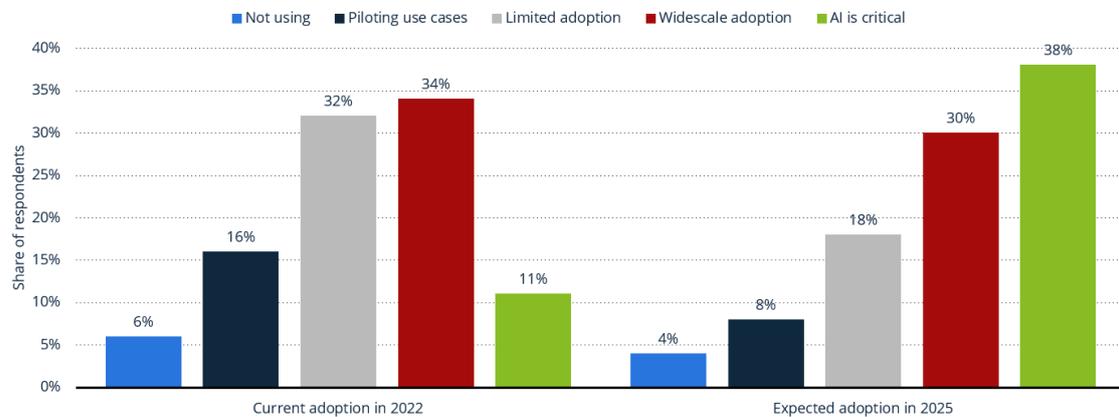
Notes: Worldwide, 2022: 600 respondents; Individuals holding senior technology roles

statista

**FIGURE 4.4** – Taux d’adoption de l’intelligence artificielle dans les entreprises de développement de produits dans le monde en 2022 et 2025 (Statista, 2023)

## Artificial intelligence (AI) adoption rate in supply chain and manufacturing businesses worldwide in 2022 and 2025

Adoption rate of AI in global supply chain business 2022-2025



Description: The adoption rate of artificial intelligence (AI) is expected to grow in companies operating in supply chains and manufacturing industries from 2022 to 2025. In 2022 over a third of executives expected their companies to have a widescale adoption of AI in their companies. [Read more](#)

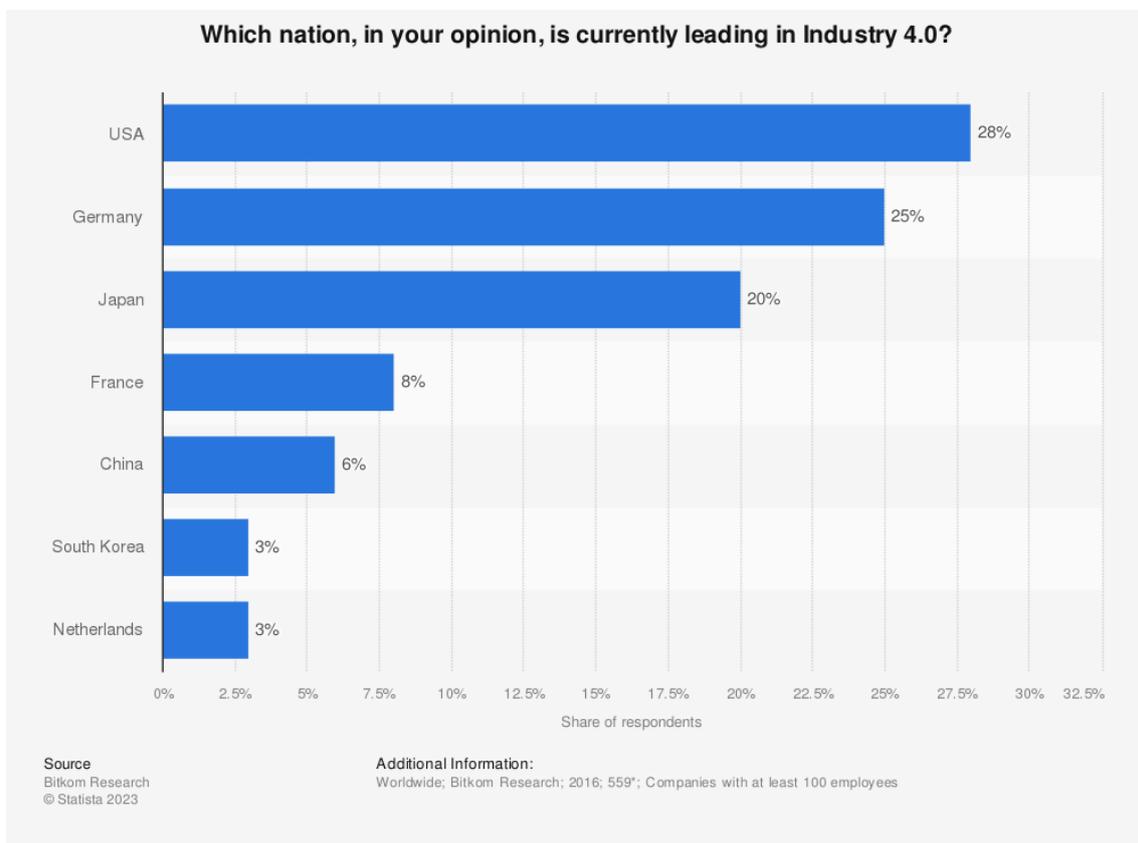
Notes: Worldwide, 2022: 600 respondents; Individuals holding senior technology roles

statista

**FIGURE 4.5** – Taux d’adoption de l’intelligence artificielle dans les entreprises de la chaîne d’approvisionnement et de la fabrication dans le monde en 2022 et 2025 (Statista, 2023)

Une autre étude quantitative menée par Statista Research Department (2016) recense à la FIGURE 4.6 un classement des pays où l’industrie 4.0 (par conséquent, les technologies de machine learning et de data science) est la plus développée. En première position

se trouvent les États-Unis qui représentent 28% des réponses, en deuxième position se trouve l'Allemagne avec 25% des réponses et en troisième position se trouve le Japon avec 20% des réponses (Statista Research Department, 2016). Dans ces sept pays où l'industrie 4.0 est a priori la plus développée, il y a trois pays européens, l'Allemagne, la France et les Pays-Bas. Ces trois pays totalisent 36% des réponses (Statista Research Department, 2016). Selon cette étude, l'Europe ne peut pas vraiment être considérée comme un continent leader de l'industrie 4.0 mais c'est, malgré tout, un continent majeur dans le développement des technologies de machine learning et de data science dans l'industrie 4.0. Il est quand même nécessaire de notifier qu'il ne s'agit pas d'une étude qui se base sur des données objectives mais sur l'avis de 559 entreprises de plus de 100 employés auxquelles la question suivante a été posée : "Quelle nation, selon vous, est actuellement à la pointe de l'industrie 4.0?". Il peut donc y avoir certains biais (d'échantillonnage, entre autres) qui peuvent mener à des résultats approximatifs.



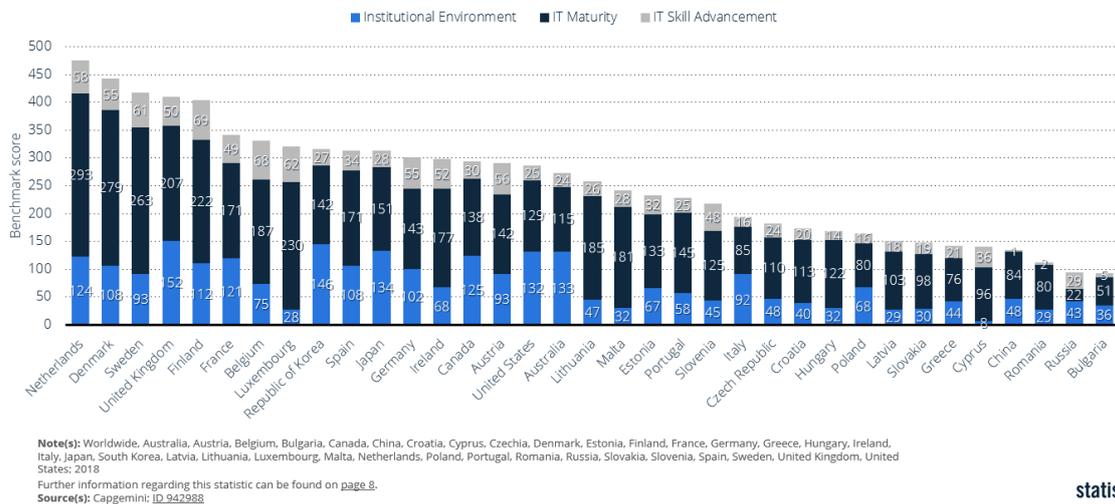
**FIGURE 4.6** – Pays leaders de l'industrie 4.0 dans le monde en 2016 (Statista Research Department, 2016)

L'étude de Capgemini (2018) montre un classement des pays en fonction de leur degré de préparation à l'intégration de l'intelligence artificielle à la FIGURE 4.7. Sur les 10 premiers pays, 9 sont européens (Capgemini, 2018), ce qui montre une bonne préparation aux technologies d'intelligence artificielle qui pourront être utilisées dans le contexte de

l'industrie 4.0. La Belgique est située en septième position (Capgemini, 2018), ce qui est une très bonne performance au niveau mondial et une performance dans la moyenne des pays européens.

## Artificial intelligence (AI) readiness benchmark by country as of 2018

AI country readiness benchmark 2018



**FIGURE 4.7** – Benchmark de l'état de préparation à l'intelligence artificielle (IA) par pays en 2018 (Capgemini, 2018)

L'article de Niewiadomski, Stachowiak, et Pawlak (2019) donne un indice de connaissance et de maturité des employés à propos de 25 outils, applications et systèmes d'IA appliqués aux entreprises sur une échelle de 1 à 5 allant de "Très faible connaissance de la maturité des outils d'IA" à "Très bonne connaissance de la maturité des outils d'IA". Il en ressort que la connaissance de la maturité des outils d'IA est de 3,96 sur 5 en moyenne sur les 25 objets (Niewiadomski et al., 2019). Cette étude a été menée auprès de 71 entreprises du secteur de la production agricole. Ce n'est pas exactement le domaine d'activité qui fait l'objet de l'analyse de ce mémoire mais cela donne une idée de la manière dont la maturité de l'IA est perçue.

## 5 Discussion

Il est maintenant nécessaire de répondre aux questions de recherches posées à la section d'introduction. Pour ce faire, il faut rappeler les questions de recherche et les mettre en perspective avec un résumé des différents apprentissages réalisés au cours de l'analyse des articles scientifiques trouvés. Ensuite, il est nécessaire de comparer les résultats obtenus et les réponses aux questions de recherche avec la littérature existante. Il sera aussi nécessaire de parler des diverses limitations de ce travail pour finir par aborder des recommandations si un travail similaire devait être réalisé.

## 5.1 Réponses aux questions de recherche

Pour répondre à la première question de recherche (Quelles sont les applications pratiques du machine learning et de la data science dans le monde de l'industrie 4.0?), le machine learning et la data science sont vastement utilisés dans le monde de l'industrie 4.0 (Gokalp et al., 2016). Les algorithmes de machine learning tels que les réseaux de neurones artificiels, convolutifs, les machines à vecteurs de support et le deep learning permettent aux entreprises de faire des prédictions (Diez-Olivan et al., 2019). Les techniques de data science comme le k-nearest neighbor ou les régressions linéaires sont utiles pour détecter des outliers, c'est-à-dire, des occurrences qui ne sont pas dans la norme (Cavalcante et al., 2019).

Ces disciplines sont mises en pratique dans l'industrie 4.0 pour permettre de tirer de l'information et de la connaissance de données brutes récoltées par les capteurs ou toute entité de l'internet des objets (Cimini et al., 2017) (Nasution, 2021). Ces nouvelles connaissances générées permettent aux entreprises gagner du temps et de la productivité au long des chaînes de production par divers moyens (Tortorella et al., 2019). Les algorithmes peuvent réaliser des prédictions sur l'état futur de la production, des machines, la maintenance des outils de production, la détection de défauts, la gestion de la chaîne d'approvisionnement, la planification de la production (Candanedo et al., 2018) (Dalzochio et al., 2020). Les techniques de data science sont aussi utilisées pour supporter les algorithmes de machine learning en détectant des cas particuliers (Langone et al., 2020).

Pour répondre à la deuxième question de recherche (Quels sont les intérêts en termes de performance de mettre en place ces technologies pour des entreprises voulant passer à l'industrie 4.0?), les algorithmes de machine learning et les techniques de data science ont globalement un impact positif sur les performances des entreprises qui ont instauré l'industrie 4.0 (Tortorella et al., 2019) (Lin et al., 2019).

Le premier avantage d'inclure du machine learning et de la data science est de mieux contrôler la maintenance des machines. Il est possible de prédire à quel moment une machine va tomber en panne grâce à des algorithmes de machine learning (Szász et al., 2021). Cela fait gagner du temps et de la disponibilité et, donc, évite certains coûts (Dalzochio et al., 2020).

Un autre aspect positif est le contrôle de la qualité des produits. Les algorithmes de machine learning permettent de détecter si un produit est défectueux en analysant des images prises au cours de la production de produits (Villalba-Diez et al., 2019). Les produits défectueux ne sont alors pas vendus et cela évite des coûts liés à des retours éventuels (Villalba-Diez et al., 2019).

Un troisième bienfait de l'utilisation du machine learning et de la data science est le contrôle de la chaîne d'approvisionnement (Heydarbakian & Spehri, 2022). Via des algorithmes, il est faisable de trier les fournisseurs sur base des données historiques et des résultats que les fournisseurs ont proposés (Cavalcante et al., 2019). Cela offre des gains de performance si les fournisseurs sont plus rapides, ce qui produit aussi des gains financiers, plus la production est rapide, plus les entreprises peuvent vendre (Tortorella et al., 2019).

Comme quatrième avantage d'implémenter le machine learning, il y a la planification de la production (Usuga Cadavid et al., 2020). Le machine learning permet de voir ce qui était bon et mauvais dans les flux de matériaux précédents et de créer un modèle de gestion logistique optimal (Usuga Cadavid et al., 2020). Avoir des flux de matériaux qui sont optimisés au maximum donne aux entreprises la possibilité d'avoir des performances opérationnelles et financières maximales (Tortorella et al., 2019).

L'avant-dernier bénéfice de l'ajout de techniques de data science est la possibilité d'avoir de la visualisation avancée (Shafiq et al., 2019). Cela sert à avoir des données sur la production, les équipements, etc. en temps réel pour détecter les éventuels problèmes (Lodgaard & Dransfeld, 2020). Identifier des anomalies le plus tôt possible évite d'interrompre la production et donc cela accroît la productivité (Shafiq et al., 2019).

Les derniers avantages de l'implémentation des techniques de machine learning et de data science sont l'analyse prédictive et l'analyse prescriptive. Les entreprises peuvent prévoir ce qu'il va se passer en utilisant des algorithmes de machine learning et en faisant de l'analyse prédictive (Candanedo et al., 2018). Elles peuvent aussi détecter quelles actions entreprendre pour obtenir certains résultats via des méthodes de data science d'analyse prescriptive (Romeo et al., 2020). Cela leur offre l'opportunité de ne pas devoir faire d'essais qui n'ont pas de bons résultats. Par ce biais, elles améliorent leurs performances financières (Paolanti et al., 2018).

Pour répondre à la troisième question de recherche (Quelle est la maturité de l'intégration du machine learning et de la data science dans l'industrie 4.0 ?), le développement des techniques de machine learning et de data science dans l'industrie 4.0 est relativement nouveau (Cimini et al., 2017). L'intégration du machine learning et de la data science dans l'industrie 4.0 est relativement récente (Jagatheesaperumal et al., 2021). Les technologies sont utilisées pour des applications assez basiques où le machine learning et la data science aident à faire des prédictions et donnent des résultats numériques (Candanedo et al., 2018). L'adoption de ces technologies dans le futur est assez prometteuse étant donné le degré d'utilisation attendu à l'avenir (Statista, 2023). Les pays qui ont le plus de capacité à développer ces technologies sont majoritairement européens, nord-américains et les pays les plus développés d'Asie (Capgemini, 2018) (Statista Research Department, 2016).

## 5.2 Comparaison avec la littérature existante

Les articles scientifiques qui traitent du même sujet ont, en général, des conclusions similaires. L'article de Oikonomidi (2020) dit que "l'utilisation du big data dans l'industrie 4.0 a des effets bénéfiques" comme première conclusion. L'auteur a effectué une revue de littérature systématique, dans tous les articles scientifiques pris en compte, une conclusion majeure est qu'au sein de la communauté académique, une attention particulière est accordée aux avantages considérables relatifs au big data dans le but de stimuler les organisations à adopter ces technologies (Oikonomidi, 2020).

L'article de Szász et al. (2021) a des conclusions similaires à celles de ce travail en ce qui concerne l'augmentation de la performance grâce à l'industrie 4.0. Cette étude qualitative démontre que les technologies émergentes de l'industrie 4.0 offrent la possibilité d'améliorer la performance opérationnelle. Cette amélioration se traduit par une augmentation de la qualité des produits et, dans un second temps, une baisse des coûts. Cela est dû au fait que réduire les coûts n'est pas le but premier de l'implémentation de l'industrie 4.0 (Szász et al., 2021). L'article n'aborde cependant pas les sujets concernant le machine learning et la data science.

De manière générale, les articles trouvés via la méthodologie de recherche d'articles scientifique ou via de simples recherches d'articles ne parlent pas des effets directs du machine learning et de la data science sur les performances des entreprises ayant intégré l'industrie 4.0. Il n'y a aussi que très peu de littérature scientifique abordant le sujet de la maturité des outils de machine learning et de data science dans l'industrie 4.0. Ce présent travail est donc pertinent dans ce contexte.

## 5.3 Limitations

Analyser l'impact direct du machine learning et de la data science sur la performance des entreprises est très complexe étant donné la quantité très limitée de ressources (littérature, données) qui étudient cet effet direct. Pour tirer des conclusions, il faut étudier l'implication de ces technologies dans l'industrie 4.0 et ensuite étudier l'impact de l'implémentation des techniques de l'industrie 4.0 sur les performances des entreprises. Cela ajoute une étape supplémentaire susceptible de ne pas totalement mettre en évidence l'importance du lien à démontrer.

Sans mener une analyse avec des données chiffrées, il est difficile de déduire des conclusions objectives et mathématiquement fondées. Il est seulement possible de voir quelles technologies ont un impact sur les performances des entreprises utilisant les outils de l'industrie 4.0. Malheureusement, il n'a pas réellement été possible de mesurer cet impact.

## 5.4 Recommandations

Les travaux similaires sont en partie des études quantitatives portant sur les effets de l'implémentation de l'industrie 4.0 sur la performance. Il aurait alors été intéressant de mener une étude quantitative auprès d'entreprises ayant déjà instauré des techniques de machine learning et de data science durant leur passage à l'industrie 4.0. Cette étude aurait pu avoir pour but de démontrer s'il existe un effet positif ou non de chaque technologie et de chaque application sur les performances opérationnelles et financières. Elle aurait aussi pu aider à démontrer la force de cet effet en prenant en compte un grand ensemble de variables.

Ce mémoire étant essentiellement une revue de littérature et, par conséquent, un travail majoritairement théorique, une autre possibilité aurait été de mener une étude qualitative auprès de consultants ou de responsables en transformation digitale. Cette étude aurait pu permettre d'avoir un point de vue pratique sur l'implémentation des technologies dans l'industrie 4.0.

S'il fallait à nouveau faire un travail portant sur ce sujet, les recommandations seraient alors les suivantes : mener une étude qualitative pour pouvoir démontrer mathématiquement l'impact des technologies étudiées sur la performance opérationnelle et financière et mener une étude quantitative pour avoir des avis de personnes mettant en pratique ces technologies et dans quel but elles sont implémentées.

## 6 Conclusion

Pour répondre à la première question de recherche (Quelles sont les applications pratiques du machine learning et de la data science dans le monde de l'industrie 4.0 ?), une revue de littérature intégrative a été menée. Bien que ne soit pas une revue systématique, une méthodologie rigoureuse a été développée. Les articles scientifiques trouvés ont directement permis de répondre à la première question de recherche.

En ce qui concerne la deuxième question de recherche (Quels sont les intérêts en termes de performance de mettre en place ces technologies pour des entreprises voulant passer à l'industrie 4.0 ?), c'est aussi la revue de littérature qui a partiellement permis d'y répondre. Pour cette deuxième question, il a fallu réaliser une combinaison de connaissances au sujet des technologies implémentées et de l'impact de l'industrie 4.0 sur les performances. Les différents articles scientifiques pris en compte dans cette analyse montrent par différents moyens (questionnaires qualitatifs, quantitatifs, données, statistiques, etc.) que la mise en place du machine learning et de la data science dans l'industrie 4.0 a bel et bien un impact positif. Les entreprises améliorent la qualité de leurs produits, les temps de production, l'utilisation de leurs machines tout en diminuant les coûts. Cependant, cet impact n'a

pas pu être mesuré à cause du type de recherche donc ce travail n'a pas complètement pu permettre de répondre à cette question de recherche. En effet, il n'y a pas assez de littérature sur ce sujet précis que pour avoir une réponse claire. Il a fallu utiliser une méthode pour générer de la connaissance en interpolant des articles sur des sujets liés.

Ce travail a aussi pu apporter des éléments de réponse à la troisième question de recherche (Quelle est la maturité de l'intégration du machine learning et de la data science dans l'industrie 4.0?). Une réponse a pu être apportée par une recherche de données chiffrées qui a permis de mesurer la maturité des technologies de machine learning et de data science dans l'industrie 4.0. Cette mesure a été appuyée par une revue de la littérature pour répondre à la question de recherche.

Finalement, il pourrait être intéressant de mener des études qui mesurent la performance des entreprises avant et après le passage à l'industrie 4.0 pour déterminer s'il y a un changement significatif au niveau des performances. Ces études pourraient aussi inclure la dimension des technologies utilisées dans le cadre du passage à l'industrie 4.0 pour déterminer lesquelles ont le plus d'impact positif au niveau des performances de l'entreprise.

## Références

- Ahleroff, S., Xu, X., Lu, Y., Aristizabal, M., Velásquez, J. P., Joa, B., & Valencia, Y. (2020). Iot-enabled smart appliances under industry 4.0 : A case study. *Advanced engineering informatics*, *43*, 101043.
- Ahmed, S. F., Alam, M. S. B., Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., ... Gandomi, A. H. (2023). Deep learning modelling techniques : current progress, applications, advantages, and challenges. *Artificial Intelligence Review*, 1–97.
- Allen, L., Atkinson, J., Jayasundara, D., Cordiner, J., & Moghadam, P. Z. (2021). Data visualization for industry 4.0 : A stepping-stone toward a digital future, bridging the gap between academia and industry. *Patterns*, *2*(5), 100266.
- Almeida, J. S. (2002). Predictive non-linear modeling of complex data by artificial neural networks. *Current opinion in biotechnology*, *13*(1), 72–76.
- Angelopoulos, A., Michailidis, E. T., Nomikos, N., Trakadas, P., Hatziefremidis, A., Voliotis, S., & Zahariadis, T. (2019). Tackling faults in the industry 4.0 era—a survey of machine-learning solutions and key aspects. *Sensors*, *20*(1), 109.
- Balamurugan, E., Flaih, L. R., Yuvaraj, D., Sangeetha, K., Jayanthiladevi, A., & Kumar, T. S. (2019). Use case of artificial intelligence in machine learning manufacturing 4.0. In *2019 international conference on computational intelligence and knowledge economy (iccike)* (pp. 656–659).
- Benbarrad, T., Salhaoui, M., Kenitar, S. B., & Arioua, M. (2021). Intelligent machine vision model for defective product inspection based on machine learning. *Journal of Sensor and Actuator Networks*, *10*(1), 7.
- Candanedo, I. S., Nieves, E. H., González, S. R., Martín, M. T. S., & Briones, A. G. (2018). Machine learning predictive model for industry 4.0. In *Knowledge management in organizations : 13th international conference, kmo 2018, žilina, slovakia, august 6–10, 2018, proceedings 13* (pp. 501–510).
- Capgemini. (2018). *Artificial intelligence (ai) readiness benchmark by country as of 2018*. <https://www.statista.com/statistics/942988/ai-readiness-benchmark-by-country/>. (Consulté le 25-05-2023)
- Cavalcante, I. M., Frazzon, E. M., Forcellini, F. A., & Ivanov, D. (2019). A supervised machine learning approach to data-driven simulation of resilient supplier selection in digital manufacturing. *International Journal of Information Management*, *49*, 86–97.
- Cimini, C., Pinto, R., Pezzotta, G., & Gaiardelli, P. (2017). The transition towards industry 4.0 : business opportunities and expected impacts for suppliers and manufacturers. In *Advances in production management systems. the path to intelligent, collaborative and sustainable manufacturing : Ifip wg 5.7 international conference, apms 2017, hamburg, germany, september 3-7, 2017, proceedings, part i* (pp. 119–

126).

- Coleman, S. Y. (2019). Data science in industry 4.0. In *progress in industrial mathematics at ecmi 2018* (pp. 559–566). Springer.
- Dalenogare, L. S., Benitez, G. B., Ayala, N. F., & Frank, A. G. (2018). The expected contribution of industry 4.0 technologies for industrial performance. *International Journal of production economics*, *204*, 383–394.
- Dalzochio, J., Kunst, R., Pignaton, E., Binotto, A., Sanyal, S., Favilla, J., & Barbosa, J. (2020). Machine learning and reasoning for predictive maintenance in industry 4.0 : Current status and challenges. *Computers in Industry*, *123*, 103298.
- Diez-Olivan, A., Del Ser, J., Galar, D., & Sierra, B. (2019). Data fusion and machine learning for industrial prognosis : Trends and perspectives towards industry 4.0. *Information Fusion*, *50*, 92–111.
- Dos Santos, L. M. A. L., da Costa, M. B., Kothe, J. V., Benitez, G. B., Schaefer, J. L., Baierle, I. C., & Nara, E. O. B. (2021). Industry 4.0 collaborative networks for industrial performance. *Journal of Manufacturing Technology Management*, *32*(2), 245–265.
- Duman, M. C., & Akdemir, B. (2021). A study to determine the effects of industry 4.0 technology components on organizational performance. *Technological Forecasting and Social Change*, *167*, 120615.
- Gokalp, M. O., Kayabay, K., Akyol, M. A., Eren, P. E., & Kocyiğit, A. (2016). Big data for industry 4.0 : A conceptual framework. In *2016 international conference on computational science and computational intelligence (csci)* (pp. 431–434).
- Gröger, C. (2018). Building an industry 4.0 analytics platform. *Datenbank-Spektrum*, *18*(1), 5–14.
- Gunawan, P. A., Gozali, L., Widodo, L., Daywin, F. J., & Doaly, C. (2021). Production planning and capacity control with demand forecasting using artificial neural network (case study pt. dynaplast) for industry 4.0. In *Proceedings of the 11th annual international conference on industrial engineering and operations management singapore*.
- Hermann, M., Pentek, T., & Otto, B. (2016). Design principles for industrie 4.0 scenarios. In *2016 49th hawaii international conference on system sciences (hicss)* (pp. 3928–3937).
- Heydarbakian, S., & Spehri, M. (2022). Interpretable machine learning to improve supply chain resilience, an industry 4.0 recipe. *IFAC-PapersOnLine*, *55*(10), 2834–2839.
- Jagatheesaperumal, S. K., Rahouti, M., Ahmad, K., Al-Fuqaha, A., & Guizani, M. (2021). The duo of artificial intelligence and big data for industry 4.0 : Applications, techniques, challenges, and future research directions. *IEEE Internet of Things Journal*, *9*(15), 12861–12885.
- Kagermann, H., Wahlster, W., Helbig, J., et al. (2013). Recommendations for implemen-

- ting the strategic initiative industrie 4.0 : Final report of the industrie 4.0 working group. *Forschungsunion : Berlin, Germany*.
- Lampropoulos, G. (2023). Artificial intelligence, big data, and machine learning in industry 4.0. In *Encyclopedia of data science and machine learning* (pp. 2101–2109). IGI Global.
- Langone, R., Cuzzocrea, A., & Skantzos, N. (2020). Interpretable anomaly prediction : Predicting anomalous behavior in industry 4.0 settings via regularized logistic regression tools. *Data & Knowledge Engineering, 130*, 101850.
- Laricchia, F. (2022). *Main reasons to implement industry 4.0 projects in companies in italy in 2022*. <https://www.statista.com/statistics/1313615/reasons-to-implement-industry-40-projects-companies-italy/>. (Consulté le 13-04-2023)
- Lin, B., Wu, W., & Song, M. (2019). Industry 4.0 : Driving factors and impacts on firm's performance : An empirical study on china's manufacturing industry. *Annals of Operations Research, 1–21*.
- Lodgaard, E., & Dransfeld, S. (2020). Organizational aspects for successful integration of human-machine interaction in the industry 4.0 era. *Procedia cirp, 88*, 218–222.
- Madakam, S., Ramaswamy, R., & Tripathi, S. (2015). Internet of things (iot) : A literature. *Journal of Computer and Communications, 3*, 164–173.
- Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).[Internet], 9*, 381–386.
- Nagar, D., Raghav, S., Bhardwaj, A., Kumar, R., Singh, P. L., & Sindhvani, R. (2021). Machine learning : Best way to sustain the supply chain in the era of industry 4.0. *Materials Today : Proceedings, 47*, 3676–3682.
- Nasution, M. K. (2021). Industry 4.0 : Data science perspective. In *Iop conference series : Materials science and engineering* (Vol. 1122, p. 012037).
- Niewiadomski, P., Stachowiak, A., & Pawlak, N. (2019). Knowledge on it tools based on ai maturity–industry 4.0 perspective. *Procedia Manufacturing, 39*, 574–582.
- Oikonomidi, S. (2020). *Impact of big data analytics in industry 4.0*.
- Paolanti, M., Romeo, L., Felicetti, A., Mancini, A., Frontoni, E., & Loncarski, J. (2018). Machine learning approach for predictive maintenance in industry 4.0. In *2018 14th ieee/asme international conference on mechatronic and embedded systems and applications (mesa)* (pp. 1–6).
- Placek, M. (2022). *Technology adoption in warehouse automation worldwide in 2019 and 2030*. <https://www.statista.com/statistics/1169394/global-warehouse-automation-technology-adoption/>. (Consulté le 13-04-2023)
- Rai, R., Tiwari, M. K., Ivanov, D., & Dolgui, A. (2021). *Machine learning in manufacturing and industry 4.0 applications* (Vol. 59) (N° 16). Taylor & Francis.
- Riveret, R., Gao, Y., Governatori, G., Rotolo, A., Pitt, J., & Sartor, G. (2019). A pro-

- babilistic argumentation framework for reinforcement learning agents. *Autonomous Agents and Multi-Agent Systems*, 33(1), 216–274.
- Romeo, L., Loncarski, J., Paolanti, M., Bocchini, G., Mancini, A., & Frontoni, E. (2020). Machine learning-based design support system for the prediction of heterogeneous machine parameters in industry 4.0. *Expert Systems with Applications*, 140, 112869.
- Sagiroglu, S., & Sinanc, D. (2013). Big data : A review. In *2013 international conference on collaboration technologies and systems (cts)* (pp. 42–47).
- Sajid, S., Haleem, A., Bahl, S., Javaid, M., Goyal, T., & Mittal, M. (2021). Data science applications for predictive maintenance and materials science in context to industry 4.0. *Materials today : proceedings*, 45, 4898–4905.
- Shafiq, S. I., Szczerbicki, E., & Sanin, C. (2019). Proposition of the methodology for data acquisition, analysis and visualization in support of industry 4.0. *Procedia computer science*, 159, 1976–1985.
- Song, I.-Y., & Zhu, Y. (2016). Big data and data science : what should we teach ? *Expert Systems*, 33(4), 364–373.
- Statista. (2023). *Artificial intelligence (ai) in business*. <https://www.statista.com/study/133541/artificial-intelligence-ai-in-business/>. (Consulté le 25-05-2023)
- Statista Research Department. (2016). *Leading countries in industry 4.0 worldwide as of 2016*. <https://www.statista.com/statistics/667634/leading-countries-industry-40-worldwide/>. (Consulté le 25-05-2023)
- Szász, L., Demeter, K., Racz, B.-G., & Losonci, D. (2021). Industry 4.0 : a review and analysis of contingency and performance effects. *Journal of Manufacturing Technology Management*, 32(3), 667–694.
- Tortorella, G. L., Giglio, R., & Van Dun, D. H. (2019). Industry 4.0 adoption as a moderator of the impact of lean production practices on operational performance improvement. *International journal of operations & production management*.
- Usuga Cadavid, J. P., Lamouri, S., Grabot, B., Pellerin, R., & Fortin, A. (2020). Machine learning applied in production planning and control : a state-of-the-art in the era of industry 4.0. *Journal of Intelligent Manufacturing*, 31, 1531–1558.
- Villalba-Diez, J., Schmidt, D., Gevers, R., Ordieres-Meré, J., Buchwitz, M., & Wellbrock, W. (2019). Deep learning for industrial computer vision quality control in the printing industry 4.0. *Sensors*, 19(18), 3987.
- Weber, R. H. (2010). Internet of things—new security and privacy challenges. *Computer law & security review*, 26(1), 23–30.
- Xu, L. D., & Duan, L. (2019). Big data for cyber physical systems in industry 4.0 : a survey. *Enterprise Information Systems*, 13(2), 148–169.
- Zäpfel, G., & Missbauer, H. (1993). New concepts for production planning and control.

*European Journal of Operational Research*, 67(3), 297–320.

Zheng, P., Wang, H., Sang, Z., Zhong, R. Y., Liu, Y., Liu, C., ... Xu, X. (2018). Smart manufacturing systems for industry 4.0 : Conceptual framework, scenarios, and future perspectives. *Frontiers of Mechanical Engineering*, 13, 137–150.

## A Annexes

---

### Citation de la source

---

#### Journaux scientifiques :

Angelopoulos, A., Michailidis, E. T., Nomikos, N., Trakadas, P., Hatziefremidis, A., Voliotis, S., & Zahariadis, T. (2019). Tackling faults in the industry 4.0 era—a survey of machine-learning solutions and key aspects. *Sensors*, 20(1), 109.

Balamurugan, E., Flaih, L. R., Yuvaraj, D., Sangeetha, K., Jayanthiladevi, A., & Kumar, T. S. (2019). Use case of artificial intelligence in machine learning manufacturing 4.0. 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), 656–659. IEEE.

Benbarrad, T., Salhaoui, M., Kenitar, S. B., & Arioua, M. (2021). Intelligent machine vision model for defective product inspection based on machine learning. *Journal of Sensor and Actuator Networks*, 10(1), 7.

Candanedo, I. S., Nieves, E. H., González, S. R., Martín, M. T. S., & Briones, A. G. (2018). Machine learning predictive model for industry 4.0. *Knowledge Management in Organizations : 13th International Conference, KMO 2018, Žilina, Slovakia, August 6–10, 2018, Proceedings 13*, 501–510. Springer.

Cavalcante, I. M., Frazzon, E. M., Forcellini, F. A., & Ivanov, D. (2019). A supervised machine learning approach to data-driven simulation of resilient supplier selection in digital manufacturing. *International Journal of Information Management*, 49, 86–97.

Cimini, C., Pinto, R., Pezzotta, G., & Gaiardelli, P. (2017). The transition towards industry 4.0 : business opportunities and expected impacts for suppliers and manufacturers. *Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing : IFIP WG 5.7 International Conference, APMS 2017, Hamburg, Germany, September 3-7, 2017, Proceedings, Part I*, 119–126. Springer.

Coleman, S. Y. (2019). Data science in industry 4.0. In *progress in industrial mathematics at ecmi 2018* (pp. 559–566). Springer.

Dalenogare, L. S., Benitez, G. B., Ayala, N. F., & Frank, A. G. (2018). The expected contribution of Industry 4.0 technologies for industrial performance. *International Journal of Production Economics*, 204, 383–394.

Dalzochio, J., Kunst, R., Pignaton, E., Binotto, A., Sanyal, S., Favilla, J., & Barbosa, J. (2020). Machine learning and reasoning for predictive maintenance in Industry 4.0 : Current status and challenges. *Computers in Industry*, 123, 103298.

- Diez-Olivan, A., Del Ser, J., Galar, D., & Sierra, B. (2019). Data fusion and machine learning for industrial prognosis : Trends and perspectives towards Industry 4.0. *Information Fusion*, 50, 92–111.
- Dos Santos, L. M. A. L., da Costa, M. B., Kothe, J. V., Benitez, G. B., Schaefer, J. L., Baierle, I. C., & Nara, E. O. B. (2021). Industry 4.0 collaborative networks for industrial performance. *Journal of Manufacturing Technology Management*, 32(2), 245–265.
- Duman, M. C., & Akdemir, B. (2021). A study to determine the effects of industry 4.0 technology components on organizational performance. *Technological Forecasting and Social Change*, 167, 120615.
- Gokalp, M. O., Kayabay, K., Akyol, M. A., Eren, P. E., & Koçyiğit, A. (2016). Big data for industry 4.0 : A conceptual framework. 2016 International Conference on Computational Science and Computational Intelligence (CSCI), 431–434. IEEE.
- Gunawan, P. A., Gozali, L., Widodo, L., Daywin, F. J., & Doaly, C. O. (2021). Production planning and capacity control with demand forecasting using artificial neural network (Case Study PT. Dynaplast) for Industry 4.0. *Proceedings of the 11th Annual International Conference on Industrial Engineering and Operations Management Singapore*.
- Heydarbakian, S., & Spehri, M. (2022). Interpretable Machine Learning to Improve Supply Chain Resilience, An Industry 4.0 Recipe. *IFAC-PapersOnLine*, 55(10), 2834–2839.
- Jagatheesaperumal, S. K., Rahouti, M., Ahmad, K., Al-Fuqaha, A., & Guizani, M. (2021). The duo of artificial intelligence and big data for industry 4.0 : Applications, techniques, challenges, and future research directions. *IEEE Internet of Things Journal*, 9(15), 12861–12885.
- Langone, R., Cuzzocrea, A., & Skantzos, N. (2020). Interpretable anomaly prediction : Predicting anomalous behavior in industry 4.0 settings via regularized logistic regression tools. *Data & Knowledge Engineering*, 130, 101850.
- Lin, B., Wu, W., & Song, M. (2019). Industry 4.0 : Driving factors and impacts on firm's performance : An empirical study on China's manufacturing industry. *Annals of Operations Research*, 1–21.
- Lodgaard, E., & Dransfeld, S. (2020). Organizational aspects for successful integration of human-machine interaction in the industry 4.0 era. *Procedia Cirp*, 88, 218–222.
- Nagar, D., Raghav, S., Bhardwaj, A., Kumar, R., Singh, P. L., & Sindhwani, R. (2021). Machine learning : Best way to sustain the supply chain in the era of industry 4.0. *Materials Today : Proceedings*, 47, 3676–3682.

Niewiadomski, P., Stachowiak, A., & Pawlak, N. (2019). Knowledge on IT tools based on AI maturity–Industry 4.0 perspective. *Procedia Manufacturing*, 39, 574–582.

Paolanti, M., Romeo, L., Felicetti, A., Mancini, A., Frontoni, E., & Loncarski, J. (2018). Machine learning approach for predictive maintenance in industry 4.0. 2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA), 1–6. IEEE.

Rai, R., Tiwari, M. K., Ivanov, D., & Dolgui, A. (2021). Machine learning in manufacturing and industry 4.0 applications. *International Journal of Production Research*, Vol. 59, pp. 4773–4778. Taylor & Francis.

Romeo, L., Loncarski, J., Paolanti, M., Bocchini, G., Mancini, A., & Frontoni, E. (2020). Machine learning-based design support system for the prediction of heterogeneous machine parameters in industry 4.0. *Expert Systems with Applications*, 140, 112869.

Shafiq, S. I., Szczerbicki, E., & Sanin, C. (2019). Proposition of the methodology for Data Acquisition, Analysis and Visualization in support of Industry 4.0. *Procedia Computer Science*, 159, 1976–1985.

Szász, L., Demeter, K., Racz, B.-G., & Losonci, D. (2021). Industry 4.0 : a review and analysis of contingency and performance effects. *Journal of Manufacturing Technology Management*, 32(3), 667–694.

Tortorella, G. L., Giglio, R., & Van Dun, D. H. (2019). Industry 4.0 adoption as a moderator of the impact of lean production practices on operational performance improvement. *International Journal of Operations & Production Management*.

Usuga Cadavid, J. P., Lamouri, S., Grabot, B., Pellerin, R., & Fortin, A. (2020). Machine learning applied in production planning and control : a state-of-the-art in the era of industry 4.0. *Journal of Intelligent Manufacturing*, 31, 1531–1558.

Villalba-Diez, J., Schmidt, D., Gevers, R., Ordieres-Meré, J., Buchwitz, M., & Wellbrock, W. (2019). Deep learning for industrial computer vision quality control in the printing industry 4.0. *Sensors*, 19(18), 3987.

Xu, L. D., & Duan, L. (2019). Big data for cyber physical systems in industry 4.0 : a survey. *Enterprise Information Systems*, 13(2), 148–169.

Zheng, P., Wang, H., Sang, Z., Zhong, R. Y., Liu, Y., Liu, C., ... Xu, X. (2018). Smart manufacturing systems for Industry 4.0 : Conceptual framework, scenarios, and future perspectives. *Frontiers of Mechanical Engineering*, 13, 137–150.

---

**Statistiques :**

Capgemini. (2018). Artificial intelligence (AI) readiness benchmark by country as of 2018. Retrieved from <https://www.statista.com/statistics/942988/ai-readiness-benchmark-by-country/>

Laricchia, F. (2022). Main reasons to implement industry 4.0 projects in companies in italy in 2022. <https://www.statista.com/statistics/1313615/reasons-to-implement-industry-40-projects-companies-italy/>. (Consulté le 13-04-2023)

Placek, M. (2022). Technology adoption in warehouse automation worldwide in 2019 and 2030. <https://www.statista.com/statistics/1169394/global-warehouse-automation-technology-adoption/>. (Consulté le 13-04-2023)

Statista Research Department. (2016). Leading countries in Industry 4.0 worldwide as of 2016. Retrieved from <https://www.statista.com/statistics/667634/leading-countries-industry-40-worldwide/>

Statista. (2023). Artificial Intelligence (AI) in business. Retrieved from <https://www.statista.com/study/133541/artificial-intelligence-ai-in-business/>

---

**ANNEXE 1** – Table récapitulative des sources retenues pour l'analyse