

**UNIVERSITE TOULOUSE III PAUL SABATIER**

**FACULTE DE SANTE**

**DEPARTEMENT DES SCIENCES**

**PHARMACEUTIQUES**

**ANNEE : 2023**

**THESE 2023 / TOU3 / 2001**

# **THESE**

**POUR LE DIPLOME D'ETAT DE DOCTEUR EN PHARMACIE**

Présentée et soutenue publiquement  
par

**M'RAD SAMI**

**LES OUTILS DE L'INDUSTRIE 4.0 ET LEUR APPLICATION A L'INDUSTRIE  
PHARMACEUTIQUE**

Date de soutenance

27 Janvier 2023

Directeur de thèse : Arellano Cécile

## **JURY**

Président : Peggy Gandia  
1er assesseur : Arellano Cécile  
2ème assesseur : Serain Stéphane

**PERSONNEL ENSEIGNANT**  
**du Département des Sciences Pharmaceutiques de la Faculté de santé**  
**au 4 avril 2022**

**Professeurs Emérites**

Mme BARRE A.	Biologie Cellulaire
M. BENOIST H.	Immunologie
M. BERNADOU J.	Chimie Thérapeutique
M. CAMPISTRON G.	Physiologie
Mme NEPVEU F.	Chimie analytique
Mme ROQUES C.	Bactériologie - Virologie
M. ROUGE P.	Biologie Cellulaire
M. SALLES B.	Toxicologie

**Professeurs des Universités**

**Hospitolo-Universitaires**

Mme AYYOUB M.	Immunologie
M. CESTAC P.	Pharmacie Clinique
M. CHATELUT E.	Pharmacologie
Mme DE MAS MANSAT V.	Hématologie
M. FAVRE G.	Biochimie
Mme GANDIA P.	Pharmacologie
M. PARINI A.	Physiologie
M. PASQUIER C.	Bactériologie - Virologie
Mme ROUSSIN A.	Pharmacologie
Mme SALLERIN B.	Pharmacie Clinique
M. VALENTIN A.	Parasitologie

**Universitaires**

Mme BERNARDES-GENISSON V.	Chimie thérapeutique
Mme BOUTET E.	Toxicologie - Sémiologie
Mme COUDERC B.	Biochimie
M. CUSSAC D. (Vice-Doyen)	Physiologie
M. FABRE N.	Pharmacognosie
Mme GIROD-FULLANA S.	Pharmacie Galénique
M. GUIARD B.	Pharmacologie
M. LETISSE F.	Chimie pharmaceutique
Mme MULLER-STAUMONT C.	Toxicologie - Sémiologie
Mme REYBIER-VUATTOUX K.	Chimie analytique
M. SEGUI B.	Biologie Cellulaire
Mme SIXOU S.	Biochimie
M. SOUCHARD J-P.	Chimie analytique
Mme TABOULET F.	Droit Pharmaceutique

## Maîtres de Conférences des Universités

Hospitalo-Universitaires		Universitaires	
M. DELCOURT N.	Biochimie	Mme ARELLANO C. (*)	Chimie Thérapeutique
Mme JUILLARD-CONDAT B.	Droit Pharmaceutique	Mme AUTHIER H.	Parasitologie
Mme KELLER L.	Biochimie	M. BERGE M. (*)	Bactériologie - Virologie
M. PUISSET F.	Pharmacie Clinique	Mme BON C. (*)	Biophysique
Mme ROUCH L.	Pharmacie Clinique	M. BOUJILA J. (*)	Chimie analytique
Mme ROUZAUD-LABORDE C	Pharmacie Clinique	M. BROUILLET F.	Pharmacie Galénique
Mme SALABERT A.S.	Biophysique	Mme CABOU C.	Physiologie
Mme SERONIE-VIVIEN S (*)	Biochimie	Mme CAZALBOU S. (*)	Pharmacie Galénique
Mme THOMAS F. (*)	Pharmacologie	Mme CHAPUY-REGAUD S. (*)	Bactériologie - Virologie
		Mme COLACIOS C.	Immunologie
		Mme COSTE A. (*)	Parasitologie
		Mme DERAËVE C. (*)	Chimie Thérapeutique
		Mme ECHINARD-DOUIN V.	Physiologie
		Mme EL GARAH F.	Chimie Pharmaceutique
		Mme EL HAGE S.	Chimie Pharmaceutique
		Mme FALLONE F.	Toxicologie
		Mme FERNANDEZ-VIDAL A.	Toxicologie
		Mme GADEA A.	Pharmacognosie
		Mme HALOVA-LAJOIE B.	Chimie Pharmaceutique
		Mme JOUANJUS E.	Pharmacologie
		Mme LAJOIE-MAZENC I.	Biochimie
		Mme LEFEVRE L.	Physiologie
		Mme LE LAMER A-C. (*)	Pharmacognosie
		M. LE NAOUR A.	Toxicologie
		M. LEMARIE A.	Biochimie
		M. MARTI G.	Pharmacognosie
		Mme MONFERRAN S	Biochimie
		M. PILLOUX L.	Microbiologie
		M. SAINTE-MARIE Y.	Physiologie
		M. STIGLIANI J-L.	Chimie Pharmaceutique
		M. SUDOR J. (*)	Chimie Analytique
		Mme TERRISSE A-D.	Hématologie
		Mme TOURRETTE-DIALLO A. (*)	Pharmacie Galénique Pharmacognosie
		Mme VANSTEELANDT M.	Mathématiques
		Mme WHITE-KONING M. (*)	

(\*) Titulaire de l'habilitation à diriger des recherches (HDR)

## Enseignants non titulaires

Assistants Hospitalo-Universitaires		Attaché Temporaire d'Enseignement et de Recherche (ATER)	
M. AL SAATI A	Biochimie	Mme AMRANE Dyhia	Chimie Thérapeutique
Mme BAKLOUTI S.	Pharmacologie		
Mme CLARAZ P.	Pharmacie clinique		
Mme CHAGNEAU C.	Microbiologie		
Mme LARGEAUD L	Immunologie		
M. LE LOUEDEC F.	Pharmacologie		
Mme STRUMIA M.	Pharmacie clinique		
Mme TRIBAUDEAU L.	Droit Pharmaceutique		

## REMERCIEMENTS

Je remercie avant tout Mme Arellano Cécile, ma directrice de thèse de m'avoir suivi, et conseillé tout au long de la préparation de la thèse et d'avoir fortement contribué à la réalisation de ce document

Je tiens aussi à remercier Mme Gandia Peggy d'avoir eu la gentillesse de se libérer et de présider la soutenance de ma thèse.

Un grand merci à Mr Stéphane Serain avec qui j'ai pu beaucoup apprendre dès mon stage de 5ème année de pharmacie au sein de Sanofi Ambarès-et-Lagrange et qui a gentiment accepté de prendre part à ma thèse en tant que membre du jury.

Je remercie l'ensemble des enseignants de la faculté ainsi que les services administratifs pour l'ensemble de mes années d'études. Ils m'ont permis de traverser les études de pharmacie sans encombre et de rendre cet enseignement toujours attractif.

Je remercie les enseignants de mon Master 2 PIF2P et principalement Mr Bakri Abdelaziz et Khalef Nawel qui de par : leurs expériences et les cours proposés m'ont permis d'en apprendre encore plus sur l'industrie pharmaceutique. C'est en effet grâce à cela que j'ai choisi l'industrie 4.0 comme sujet d'étude.

Je remercie la pharmacie Borderouge pour les années de plaisir et de formation durant mes années étudiantes.

### **À ma famille :**

Je remercie infiniment ma famille, ma mère Odile, mon père Fadhel, et ma sœur Ines, pour votre soutien inconditionnel et votre amour à toute épreuve. Merci pour tout, votre soutien émotionnel et financier depuis le début. J'espère que ce travail vous rendra fiers. Merci à toi Maman, de m'avoir aidé sur mes projets, encouragé et poussé quand il le fallait. Merci, papa de m'avoir poussé a donné le meilleur de moi-même et soutenu dans mes décisions. Merci Ines, d'avoir été à l'écoute et d'avoir été là pour m'aider à relativiser quand je m'emballais. C'est en grande partie grâce à vous que je suis ce que je suis aujourd'hui. Je ne vous remercierai jamais assez.

À ma tante Isabelle, merci pour ton amour, pour ton accueil, ta gentillesse, ta rigueur et ton sens de l'organisation grâce auquel je dois une partie de ma réussite.

À ma tante Muriel, pour ton amour et ta gentillesse et pour tous les moments de soutien.

**À mes amis :**

À Khalil, pour ces moments de geekage intense et pour toutes les heures de joies.

À Jean-Baptiste, pour les ses journées passées incroyables et ces moments de réflexions qui nous font grandir ensemble.

À Paul, pour le moment de magie, les soirées de folie, ton humour sans limite et ta loyauté pure, merci d'être toujours là.

À Laura, ma sœur de cœur, capable de me comprendre, merci pour ton soutien et ta gentillesse.

À Côme, pour tes conseils, ton amitié sans faille et ta positivité. Merci de me tirer toujours vers le haut.

À Remi, mon binôme incroyable, merci pour toutes ces années de bonheur et cette amitié en béton.

À Marion, pour ta gentillesse et ta générosité et pour cette amitié qui ne fait qu'embellir.

À Clément, merci pour cette jeune amitié solide. Aussi, merci d'être mon médecin traitant. Tu vas m'accompagner longtemps.

À Malek, mon frère qui comme Elyes a toujours été présent. Merci pour ton soutien, et ton « positive mindset ».

À Amine, mon frère, pour tes conseils, ta folie et ton humour. Merci d'être présent depuis plus de 15 ans.

À Atef, merci pour cette incroyable fraternité qui nous apporte réciproquement et pour tous les moments de bonheur passé ensemble à partager et profiter. Grand Hippie

A Elyes, mon meilleur ami, merci de me soutenir depuis ma tendre enfance, toi qui fais partie entièrement partie de ma famille, je te remercie pour ton aide durant tous les obstacles rencontrés, je suis fier aujourd'hui de partager ce moment avec toi.

À ma chère et tendre Anaëlle pour le soutien, la patience, l'amour qui me donne quotidiennement le courage nécessaire à ma réussite. Merci pour tes conseils et ton oreille attentive.

## Table des matières

Introduction .....	17
I- L'évolution de l'industrie .....	18
A. Trois révolutions industrielles .....	18
1. L'industrie 1.0.....	18
2. L'industrie 2.0.....	18
3. L'industrie 3.0.....	19
B. L'industrie 4.0.....	19
II- Les outils de l'industrie 4.0 .....	21
A. « Le Cloud Computing » .....	21
B. Big Data .....	23
C. Internet industriel des objets (IIot) .....	24
D. Le Machine Learning .....	26
1. L'apprentissage Supervisé .....	27
a. Les Arbres décisionnels .....	29
b. La classification naïve Bayésienne.....	29
c. Les machines à vecteur de support (SVM) .....	30
2. L'apprentissage non supervisé .....	32
a. Analyse en composante principale (ACP).....	32
b. Le positionnement en K-moyenne .....	33
3. Apprentissage Semi-Supervisé .....	34
a. Self-Training.....	34
b. Machine a Vecteur Support Transductif (TSVM).....	34
c. Modèle Génératif .....	35
4. Apprentissage par renforcement .....	36
5. Réseau de neurone artificiel .....	37
6. L'apprentissage ensembliste .....	38
a. Bagging .....	40
b. Boosting.....	40
c. Stacking .....	41
d. Moyenne du modèle Bayésien.....	41
7. L'apprentissage Multitâches .....	41
8. Apprentissage basé sur les exemples.....	42
a. Méthode des K plus proches voisin.....	42

E.	Les logiciels de pilotage de Production (MES).....	43
F.	Les Jumeaux Numériques.....	47
G.	La réalité augmentée (RA).....	48
H.	L'impression 3D.....	49
I.	La robotique collaborative .....	51
J.	Cybersécurité.....	52
K.	Le quantum computing .....	54
L.	La Blockchain .....	55
III-	L'application des outils 4.0 à l'industrie pharmaceutique.....	56
A.	Etat des lieux de l'industrie pharmaceutique .....	56
1.	L'évolution du modèle économique .....	56
2.	La maturité Digitale .....	57
B.	Stratégie de digitalisation.....	57
1.	La production .....	59
2.	La qualité .....	60
3.	La logistique.....	65
4.	La Maintenance .....	66
5.	Les autres fonctions supports .....	67
C.	L'application des outils de l'industrie 4.0.....	68
1.	L'IIOT et les Big Data.....	68
a.	En production .....	68
b.	Dans le Magasinage.....	69
c.	Dans la chaine d'approvisionnement et le transport.....	69
2.	Le Cloud computing.....	70
a.	En R&D.....	71
b.	En entreprise .....	71
3.	Le Machine Learning et l'IA .....	72
a.	En R&D.....	72
b.	En production .....	73
c.	En qualité.....	74
d.	En logistique .....	74
e.	En marketing .....	75
f.	En Santé publique et pharmacovigilance .....	75
4.	Le MES.....	76
a.	Intérêt du MES en Production.....	76
b.	Intérêt du MES en Qualité .....	77
c.	Les Ressources humaines .....	77



5.	Les jumeaux numériques.....	78
6.	La réalité augmentée.....	79
7.	L'impression 3D .....	79
8.	Les robots collaboratifs .....	80
a.	En logistique .....	81
b.	En production .....	81
c.	En Qualité .....	81
9.	La Blockchain .....	82
IV-	Les Défis et les impacts de l'industrie 4.0.....	82
A.	Les défis.....	83
1.	Les défis réglementaires.....	83
a.	Gestion de la qualité et le Quality by Design .....	84
b.	La validation des procédés .....	86
2.	Les défis techniques .....	87
a.	La gestion des données .....	87
b.	L'intégrité des données et la cybersécurité .....	88
c.	Les défis technologiques .....	88
3.	Les défis logistiques.....	89
a.	La culture d'entreprise .....	89
B.	Impact sur les métiers.....	90
1.	Evolution des métiers et création des métiers .....	90
a.	Evolution des métiers.....	90
b.	Suppression des métiers .....	90
c.	Création de nouveaux métiers .....	91
2.	Discussion sur l'évolution du métier de pharmacien.....	92
V-	Conclusion.....	94
VI-	Bibliographie .....	96

## LISTE DES FIGURES

**Figure 1 :** *Les différents algorithmes composant le Machine Learning d'après la revue internationale des sciences et de la recherche.*

**Figure 2 :** *Expression mathématique du Théorème de Bayes*

**Figure 3 :** *Exemple de graphique SVM de séparation de données linéaires pour différencier une leucémie aigüe lymphoblastique d'une leucémie aigüe myéloïde en utilisant l'expression du gène MARCKSL1*

**Figure 4 :** *Donnée classé en Cluster Grace au K-means d'après la revue internationale des sciences et de la recherche.*

**Figure 5 :** *Schéma d'un réseau neuronal artificiel en propagation avant selon lebigdata.fr*

**Figure 6 :** *pyramide de la **productique** (PERA) en 5 niveaux*

**Figure 7 :** *Les domaines d'activités fonctionnels d'un système MES selon la norme ISA S95*

**Figure 8 :** *Schéma représentatif de la chaine d'approvisionnement de l'industrie pharmaceutique*

## LISTE DES TABLEAUX

**Tableau 1 :** Tableau des niveaux de maturité digitale des industries pharmaceutiques selon Biophorum

**Tableau 2 :** Les 7 principes fondamentaux du SMQ.

**Tableau 3 :** l'intégrité des données selon l'ALCOA+

**Tableau 4 :** Descriptions des fonctions supports en fonction du niveau de digitalisation de l'industrie d'après Biophorum

## GLOSSAIRE

**Algorithme** : C'est la description d'une suite d'étapes permettant d'obtenir un résultat à partir d'éléments fournis en entrée.

**Automate programmable industriel (API)** : C'est un dispositif informatique programmable qui permet de contrôler et d'automatiser des processus de commande de machine dans les usines.

**Attrition**: En marketing, l'attrition est le phénomène de perte de clientèle ou d'abonnés, phénomène contraire à la fidélisation

**Balilage ou Etiquetage** : consiste à assigner une étiquette à un type de données. Cette étiquette (ou balise) permet à l'ordinateur d'apprendre à distinguer les différentes catégories de données. Par exemple sur un ensemble de mails (sous forme de données), on balise chaque mail afin de définir s'il correspond à un spam ou un mail valide.

**Base de données relationnelle** : La base de données relationnelle stocke les données dans un tableau à deux dimensions (ligne/colonne). Le logiciel Excel est un parfait exemple de base de données relationnelle. On appellera système de gestion de base de données relationnelles, tous les logiciels permettant de créer et utiliser une base de données.

**Classifieur** : Les classifieurs sont des algorithmes permettant de faire du classement. Il classe les différentes données dans des groupes en fonction des caractéristiques similaires.

**Communication en champ proche** : technologie de communication sans fil à courte portée et à haute fréquence, permettant l'échange d'informations entre des périphériques jusqu'à une distance d'environ 10 cm

**Clustering** : Le clustering ou partitionnement est une méthode d'analyse de données qui cherche à regrouper les données en différents « clusters » homogènes qui vont regrouper des caractéristiques communes.

**Deep Learning** : C'est une des technologies de Machine Learning qui repose sur le modèle des réseaux de neurones artificiels et qui apporte de la complexité à l'établissement des règles.

**Données structurées** : Il s'agit des données organisées dans des bases de données relationnelles en fonction de leur format, de leur typologie ou d'autres paramètres prédéfinis. Les données récupérées et saisies automatiques ou manuelles sur une feuille de calcul sont

un exemple de données structurées regroupées dans un tableur. Ceci permet leur interprétation et leur traitement par des machines et algorithmes.

**Données non structurées** : Toutes les autres données qui ne sont pas organisées comme des mails, imagerie satellite, des vidéos, des fichiers vocaux etc. Les données des capteurs utilisés avec l'internet, des objets et les données de vidéosurveillance sont deux exemples de données non structurées.

**Données semi-structurées** : Les données semi-structurées consistent en des données qui ne se peuvent pas être structurées par des paramètres prédéfinis, mais qui comportent des caractéristiques communes qu'on peut hiérarchiser ou auxquelles on peut rajouter une balise. Concrètement, à une photo prise par un téléphone, correspond une géolocalisation, un identifiant, et l'heure de la prise de la photo. A cela, on peut rajouter des balises en fonction du thème de la photo (photo de maison, photo de lac, photo de chien etc.). On crée ainsi une structure aux données.

**Hardware** : Correspond aux matériels et composants physique d'un appareil informatique comme la carte mère et les processeurs d'un ordinateur.

**Interaction Homme-Machine (IHM)** : L'IHM fait référence au tableau de bord utilisé par l'Homme pour communiquer avec une machine ou un programme informatique. Les IHM affichent les données en temps réel ce qui permet aux utilisateurs de piloter les machines

**Interopérabilité** : Capacité que possède un système ou produit à fonctionner avec d'autres systèmes ou produits sans restriction d'accès ou de mise en œuvre.

**Jeu de données** : Un ensemble de données organisé dans des lignes et des colonnes sous forme de tableau. Chaque colonne décrit une variable et chaque ligne correspond à un élément donné de l'ensemble de données

**Lac de donnée (Data Lake)** : méthode de stockage massif des Big Data sans transformation des données. Les données sont stockées brutes et ne sont pas compréhensibles en tant que telles. Cette méthode est surtout utilisée pour un stockage rapide et volumineux de différents types de données.

**Middleware** : Un middleware est un logiciel qui agit comme une passerelle entre différentes applications informatiques, outils et bases de données pour offrir aux utilisateurs des services unifiés.

**Modèle** : C'est un fichier entraîné pour reconnaître certain type de données. Un algorithme peut fournir plusieurs modèles de Machine Learning en fonction par exemple des données d'entrée ou d'autres paramètres.

**Modèle de Classification** : Ce modèle permet principalement d'identifier des groupes similaires et d'organiser les données en fonction de ces classes auxquelles elles appartiennent.

**Modèle de Régression** : Un modèle de régression est un des types d'algorithmes supervisés (l'autre étant la classification). Ces algorithmes exploitent des valeurs numériques pour dégager une tendance ou une évolution prévisible dans le temps. Comme exemple de modèle de régression on retrouve les arbres décisionnels ou les réseaux de neurones artificiels.

**Photopolymères** : un photopolymère correspond à une substance synthétique dont les propriétés physiques changent au contact de la lumière. Ils durcissent au contact de la lumière ultraviolet.

**Planification chirurgicale** : La planification à pour but d'allouer du temps de bloc opératoire a des praticiens afin d'optimiser une opération et de minimiser les effets secondaires. Elle est souvent effectuée sur des images radio acquise avant l'opération. Durant cette phase cherchent à définir le mieux possible l'objet de l'opération.

**Productique** : La productique regroupe l'ensemble des applications de l'automatique et de l'informatique aux processus de production industrielle.

**Product Life Management** : ou Gestion du cycle de vie des produits rassemble des concepts, méthodes, et des logiciels capable de gérer les produits industriels tout au long de leur cycle de vie en passant par le maintien en condition opérationnel depuis la conception et le développement du produit jusqu'à son utilisation en passant par sa fabrication.

**Profil du Produit Cible** : Target Product Profil en anglais, décrit le « profil » souhaité ou les caractéristiques d'un produit cible. Il indique l'utilisation prévue, les populations cibles et d'autres attributs souhaités des produits, y compris les caractéristiques liées à l'innocuité et à l'efficacité du médicament.

**Scalabilité** : La scalabilité d'un Cloud fait référence à la capacité d'augmenter ou de réduire la quantité de ressources informatiques selon les besoins pour répondre à l'évolution de la demande.

**Système Cyber Physique** : sont des systèmes formés d'entités collaboratives, dotées de capacité de calcul, qui sont en connexion intensive avec le monde physique environnant et les phénomènes s'y déroulant, fournissant et utilisant à la fois les services de mise à disposition et de traitement de données disponibles sur le réseau.<sup>1</sup>

**Système de contrôle industriel** : Au sein d'une usine, il intègre et gère les équipements et logiciels utilisés pour la surveillance et le contrôle des machines et appareils de l'environnement industriel.

**Tâche** : Une tâche représente le type de prédiction qui est demandé à un algorithme comme une classification ou un clustering

**Technologies de l'information et de la communication** : techniques utilisées dans le traitement et la transmission des informations, principalement l'informatique, l'Internet et les télécommunications pour produire, manipuler, convertir, stocker, gérer, transmettre et retrouver l'information et pour communiquer.

**Workflow** : Le workflow ou flux de travaux décrit les suites de tâches ou d'opérations effectuées par une personne un groupe ou un organisme.

## ABRÉVIATIONS

**API** : Automate Programmable Industriel

**AR** : Augmented Reality

**CPS** : Cyber Physical System

**FDM** : Fused Deposition Modeling

**GMAO** : Gestion de Maintenance Assisté par Ordinateur

**HRI** : Human Robot interaction

**IA** : Intelligence artificiel

**IAAS** : Infrastructure As A Service

**ICS** : Industrial Control System

**IOT** : Internet Of Things

**IIOT** : Industrial Internet Of Things

**LIMS** : Laboratory Information Management System

**PAAS** : Platform As A Service

**PAT** : Process Analytical Technology

**PLM** : Product Life Management

**PME** : Petite ou Moyenne Entreprise

**MES** : Manufacturing Execution System

**SAAS** : Software As A Service

**SCS** : Systèmes de Contrôle Industriel

**SLA** : Stereolitography

**SLS** : Selective Laser Sintering

**SMQ** : Système de Management de la Qualité

**TI** : Technologie de l'information

**TO** : Technologie Operationnelle

**TIC** : Technologie de l'Information et de la Communication



## Introduction

Le secteur industriel a toujours subi des changements en lien avec les innovations et les améliorations technologiques dans le but de maintenir un haut niveau de qualité et de compétitivité. Aujourd'hui, le monde industriel est touché par une nouvelle révolution rattachée à la création d'outils digitaux qui permettent le passage des industries à l'ère du numérique. Cette révolution résulte de la demande en compétitivité de l'ensemble de celles-ci. La qualité étant devenue un standard pour les clients et consommateurs, les entreprises cherchant à se démarquer de la concurrence ont besoin d'une stratégie qui offre une valeur ajoutée tel qu'une grande flexibilité, des délais de livraison plus court, ou encore une large gamme de variantes.

L'industrie du médicament ne fait pas exception à cette révolution et on retrouve de plus en plus cette volonté d'automatiser et d'optimiser les différents secteurs de la production pharmaceutique, mais aussi les services dit « supports » qui permettent de fournir aux patients des médicaments de qualité en temps et en heures.

L'industrie 4.0 étant soutenue par une large gamme d'outils numériques, cette thèse a pour but de rassembler au mieux les connaissances sur les différents outils qui permettent le passage à une industrie 4.0 ou industrie autonome, de montrer leurs intérêts et leur possible application à l'industrie pharmaceutique.

Le secteur pharmaceutique connaît une transition du modèle économique vers une organisation numérique. Ce secteur étant en retard par rapport à d'autres comme le secteur automobile ou aéronautique, de nombreuses questions se posent autour de notre secteur. La transition vers l'industrie 4.0 est-elle possible ? Quels sont les défis liés aux secteurs pharmaceutiques ? Quels outils implémenter à la chaîne de valeurs pharmaceutiques et comment les implémenter ? Comment protéger les données numériques générées ?

Autant de questions pour lesquelles ce document va nourrir la réflexion et tenter d'apporter des éléments de réponses.

## I- L'évolution de l'industrie

### A. Trois révolutions industrielles

#### 1. L'industrie 1.0

L'industrie 1.0 est définie par le passage, à partir du XVII<sup>ème</sup> siècle, de la fabrication manuelle en petite quantité à un volume de fabrication augmenté d'un facteur 8. Cette transition est bien entendu permise par l'utilisation d'outils mécaniques, qui via l'utilisation de la vapeur, ou du charbon comme source d'énergie pour les machines<sup>2</sup>, ont révolutionné les pratiques. Au XIX<sup>ème</sup> siècle, la production à grande échelle de médicaments par l'utilisation de machines mécaniques s'est développée à partir de deux sources : la production de médicaments pour des pharmacies officinales et les industries du colorant et des produits chimiques en plein essor à cette époque.<sup>3</sup> Aujourd'hui, certaines machines dont le principe date de cette révolution industrielle, telles que les broyeurs pneumatiques et les presses à comprimés sont toujours utilisées.<sup>3</sup>

#### 2. L'industrie 2.0

C'est à partir du XIX<sup>ème</sup> siècle, avec la découverte de l'électricité que naît la deuxième révolution industrielle. Les machines électroniques ont permis la production de masse sur des lignes d'assemblage qui permettait d'avoir des commandes pré-réglées incorporant une automatisation de base et des commandes de processus. Au niveau pharmaceutique, on a vu apparaître des machines électroniques pour le broyage, le mélange, le pesage, permettant de produire plus et également de mieux contrôler ; même si les contrôles à ce stade sont statiques et limités à des paramètres prédéterminés. Comme pour l'industrie 1.0, certaines machines dérivées de cette ère sont toujours utilisées aujourd'hui, comme la presse à comprimé rotative moderne qui peut fabriquer jusqu'à 1 million de comprimés à l'heure.<sup>3</sup>

### **3. L'industrie 3.0**

L'industrie 3.0 naît de la création dans les années 1970 des ordinateurs, et plus tard dans les années 2000, des outils de communication comme Internet ; les communications sans fils (Bluetooth®) et l'énergie nucléaire. C'est grâce à ces technologies que l'automatisation des procédés et des équipements est possible en utilisant par exemple les contrôles à mémoire programmables. Cette révolution a permis la création de nouveaux concepts comme la production continue nécessitant de nouvelles approches de contrôle en cours grâce à divers d'outils désignés par le terme « process analytical technology » ou « PAT » (également très utile dans les approches de « quality by design »). Les contrôles sont maintenant plus sophistiqués et de meilleure qualité. L'automatisation a réduit le besoin d'opérateurs humains en fabrication et a facilité le suivi des paramètres tout le long de la production.

Une grande partie des industries actuelles est bien ancrée dans l'industrie 3.0, comme l'aéronautique ou l'automobile. Les industries pharmaceutiques quant à elles, sont toujours en transition.

Par exemple, la production continue, qui permet le passage des produits/matières d'une étape de fabrication à l'autre automatiquement et continuellement, est encore peu adoptée. Ainsi l'utilisation de la méthode six sigma basée sur des analyses statistiques et la capabilité des procédés permettant de fabriquer, du premier coup, des produits de qualité en réduisant les variations observées sur le produit, est encore peu intégrée en production pharmaceutique. En revanche, les industries de l'aéronautique et de l'automobile, deux secteurs où la qualité du produit est aussi indispensable pour la sécurité des consommateurs, le principe y est commun et indispensable.

### **B. L'industrie 4.0**

C'est en Allemagne, en 2011, que le concept d'industrie 4.0 est né dans le plan d'action stratégique Allemand.<sup>4</sup> L'idée principale était de promouvoir la numérisation, revaloriser l'industrie européenne et créer de l'emploi.<sup>5</sup> Cette révolution devrait permettre, à partir de nouveaux outils technologiques, de créer une « usine autonome ». Cela correspond entre autres à un système de fabrication intégré, autonome et autoorganisé qui fonctionne indépendamment de l'intervention humaine. Cette révolution promet donc l'automatisation

de systèmes entiers ou l'industrie 3.0 permettait l'automatisation des opérations. L'industrie 4.0 suggère donc la récolte et l'utilisation des données en temps réel, l'utilisation d'algorithmes et d'intelligences artificielles pour une prise de décision commerciale et opérationnelle décentralisée, une cyberassistance, une production autonome et **l'interopérabilité** grâce aux systèmes **cyber physiques**.

Comme l'analyse Monsieur E. Mercier-Laurent, l'expérience de la crise sanitaire due à la COVID-19 nous a prouvé que le besoin d'une usine autonome capable d'apporter une plus grande flexibilité, une optimisation de processus de fabrication et faciliter l'innovation à toutes les étapes est nécessaire.<sup>5</sup>

L'industrie 4.0 se base sur plusieurs outils numériques qui vont permettre entre autres la collecte de données et la prise de décision en temps réel en production, faciliter et améliorer le développement ou l'amélioration des procédés, aider à la recherche et au développement (R&D) ou encore à la transposition d'échelle (Scale Up). Un des grands challenges de l'industrie 4.0 est l'utilisation concomitante de données provenant de plusieurs sources distinctes et qui permettraient une optimisation des services de la production et de la qualité des produits fabriqués. Par exemple rassembler les données externes à la fabrication comme l'expérience du patient, le marché et la demande, les stocks des fournisseurs et les urgences de santé publique, et les fusionner avec des informations telles que la gestion de l'énergie et des ressources, les résultats de modélisation et de simulation, et les données de laboratoire.<sup>3</sup> Ainsi, L'intégration de sources de données internes et externes permettrait une réactivité, une surveillance, un contrôle et une prédiction en temps réel sans précédent.<sup>3</sup>

Comme dit plus haut, un des objectifs de cette révolution est d'augmenter la flexibilité et l'efficacité grâce à une gestion optimisée des données. D'un autre côté, il s'agit aussi perfectionner la fabrication, la logistique et la qualité grâce à la mise en place d'un système cyber physique, d'une intelligence artificielle (IA), de la robotique et d'autres nouvelles technologies qui seront présentées plus tard.<sup>5</sup>

Globalement, il est possible de dire que les challenges de l'industrie 4.0 sont nombreux et touchent globalement à tous les services d'une usine. En effet, de la production à la qualité en passant par la logistique ou la formation du personnel, les nouvelles technologies ont pour but de révolutionner la façon de faire dans l'ensemble des domaines d'une entreprise. Aussi,

ces technologies modernes sont utilisables par l'ensemble des secteurs industriels. Il s'agit donc d'une révolution de l'organisation industrielle.

## **II- Les outils de l'industrie 4.0**

C'est grâce aux innovations technologiques et numériques telles qu'Internet en 1989 ou encore l'intelligence artificielle dont le concept est né avec Alan Turing en 1950, que l'idée d'une révolution industrielle est née. Le but premier était d'améliorer la compétitivité des industries en Allemagne, en implémentant de nouveaux outils informatiques et numériques. A partir de là sont nées des recommandations et guidelines pour la mise en place de cette nouvelle industrie. Il existe de nombreux outils utilisés qui pourrait permettre la création d'une usine intelligente. Cette thèse a pour objectif de définir l'ensemble des outils ayant un rôle majeur dans le développement de l'industrie 4.0. Ces outils sont détaillés ci-dessous.

### **A. « Le Cloud Computing »**

Le « cloud computing » est la mise à disposition de ressources et services informatiques à la demande, par Internet malgré une indisponibilité géographique. Il permet de stocker et de livrer des données sur Internet. Le « cloud computing » permet un modèle économique de gestion informatique pour les entreprises qui est de plus en plus utilisé.

Il existe trois types de Cloud :

- l'Infrastructure as a service (IaaS),
- Le Platform as a Service (PaaS)
- le Software as a Service (SaaS).

L'IaaS permet un contrôle total de l'infrastructure numérique d'un site. L'avantage est que l'infrastructure physique comprenant les centrales et les matériaux numériques sont loués et ne se trouve pas sur le site. Dans le modèle IaaS, un fournisseur tiers héberge le **hardware** et les softwares (logiciels), les connexions réseaux et les bandes passantes, le stockage et toute autre composant formant l'infrastructure. Toutes les données sont stockées physiquement

dans des serveurs et des réseaux nombreux et sont répartis sur différents Data Center pour assurer la sécurité des données et éviter les crashes. Bien entendu, tout le cloud est entretenu par le fournisseur et ce dernier donne accès aux clients les composants virtualisés pour qu'ils puissent développer leur propre structure de plateforme informatique. Un réseau local ou un cloud privé peuvent utiliser une infrastructure en IaaS. Cela permet par exemple aux jeunes entreprises d'adapter leur infrastructure en fonction de leur besoin. De plus les cloud privés garantissent une sécurité de stockage des données.<sup>6</sup>

Le second service proposé est le « Platform as a service » ou PaaS. Ici, en plus de l'infrastructure, le fournisseur propose des instances de middleware et des contextes d'exécutions comme des serveurs d'application ou des bases de données. Le développement des applications, leur installation et leur gestion sont à la charge du client. « Les PaaS sont surtout utilisées pour le développement de logiciels et d'applications.

Le SaaS permet, sous forme d'un abonnement, de mettre les logiciels à disposition des entreprises. Comme pour les autres systèmes cloud, les logiciels sont installés dans des serveurs physiques, gérés et maintenus par le fournisseur. Le SaaS est le service le plus popularisé en ce qui concerne le Cloud et est utilisé par les entreprises comme par les particuliers (messagerie, traitement de texte, stockage en ligne). Grâce au SaaS, les entreprises n'ont plus besoin d'investir dans la mise en place d'un nouveau logiciel qui implique l'implémentation de la sécurisation des données, la maintenance, la licence etc. L'installation d'un nouveau logiciel demande beaucoup de temps, de personnel et donc d'argent.

Il existe aussi trois types d'hébergement de Cloud : le Cloud public, le Cloud privé et le Cloud hybride.

Pour le Cloud privé, les installations sont souvent gérées en interne. Le Cloud privé est entièrement dédié à l'entreprise. Ce modèle offre une plus grande flexibilité, plus de contrôle et de sécurité du réseau, et un approvisionnement rapide, mais il est plus cher.<sup>7</sup>

Le Cloud public est un service fourni par un tiers, par l'intermédiaire d'Internet. Il n'est pas nécessaire ici d'acheter les installations qui peuvent coûter très cher. Ces Clouds contiennent des applications et logiciels nécessaires aux entreprises (gestion des ressources humaines, des transactions outils d'analyse, etc.)<sup>7</sup>

Le Cloud Hybride est un mélange entre le Cloud privé et le Cloud public. Dans ce cas précis, les entreprises vont stocker par exemple les informations sensibles ou des tâches très importantes dans un Cloud privé et les tâches nécessitant une **scalabilité** des ressources dans un Cloud public. De plus en plus d'entreprises penchent vers le Cloud Hybride qui permet de bénéficier des logiciels de traitement les plus récents (Saas) et qui offrent davantage de choix en termes de modèle de déploiement.<sup>7</sup>

Ainsi, de nombreux secteurs ont déjà adopté le « Cloud computing » qui permet plus d'agilité, de simplicité et d'accessibilité aux ressources informatiques, aujourd'hui indispensables au bon fonctionnement de l'entreprise. Cependant, les industries pharmaceutiques ont encore peine à mettre en place un modèle informatique basé sur le Cloud pour des raisons de confidentialité et de sécurité des données qui restent beaucoup plus contraignantes pour les entreprises du médicament.

## **B. Big Data**

Le Big data décrit la collecte et le traitement d'un très grand nombre de données numériques générées par la population mondiale. Communément, ces données sont collectées par des capteurs à partir de nos outils numériques utilisés tous les jours (ordinateur, téléphone, montre connectée, dossiers médicaux numériques, etc.). Elles sont donc très variées (données de santé, géolocalisation, tranche d'âge, comportement etc.)<sup>8</sup>

Une entreprise privée peut collecter jusqu'à des centaines de pétaoctets (1 048 576 Go) de données issues des capteurs enregistrant sur les machines et les logiciels. Le Big Data est donc indispensable pour une industrie 4.0 afin de collecter et traiter leurs données.

Aucune définition précise n'a été établie pour le Big data. Mais la communauté scientifique est d'accord pour la définir en fonction des caractéristiques des données, à travers la définition des « 6Vs »<sup>8</sup> :

1. *Le volume* : Comme dit précédemment on ne récolte pas le même volume de données chez une start-up et chez une multinationale.
2. *La Vitesse* : Toutes les données ne sont pas reçues et traitées à la même vitesse. Les données haute vitesse sont transmises à la mémoire vive là où les autres sont écrites

sur le disque dur d'un ordinateur. Cela a son importance lorsqu'on veut avoir l'aperçu d'un processus en temps réel.<sup>9</sup>

3. *La variété* : Cette caractéristique fait référence aux types ou formats de données récoltées. S'il s'agit de **données structurées, non structurées** ou **encore semi structurées**. En effet, les données non structurées et semi structurées comme les fichiers texte, audio, vidéo nécessiteront un prétraitement pour leur prise en charge.
4. *La valeur* : La valeur prend en compte la valeur ajoutée par la récolte des données et du traitement de ces dernières. Il est important de comprendre l'objectif avant même de récolter et de traiter les données.<sup>9</sup>
5. *La véracité* : Ce point implique la fiabilité des données. Toutes les données et surtout les données non structurées comme les mails ou publications sur les réseaux sociaux sont sujettes à un manque de fiabilité et il reste nécessaire de le prendre en compte lors du traitement de celles-ci.<sup>10</sup>
6. *La variabilité* : La variabilité caractérise le changement des données récoltées. À quelle fréquence les données changent-elles ? À quelle fréquence le type de données change. Ce point prend en compte le flux de données récoltées par une même source.<sup>9</sup>

L'application de cette technologie est devenue primordiale d'un point de vue commercial et marketing pour la plupart des entreprises. Le Big data est aussi devenu un nouvel outil novateur dans la transformation des processus, de la chaîne de logistique et dans la prise de décision. Il sont à la base de nombreux outils technologiques comme l'internet des objets, le Machine Learning et bien d'autres.

### **C. Internet industriel des objets (IIoT)**

L'internet des objets décrit entre autres ce qu'on appelle communément les objets connectés. Elle permet d'interconnecter des objets, des environnements physiques avec Internet. Elle permet aussi de récolter des données sur des objets physiques et de les



retransmettre numériquement à travers un objet électronique (smartphone, ordinateurs). Plus globalement, l'internet des objets permet de connecter les objets entre eux avec Internet et avec plusieurs milliards de personnes. Il est défini par L'Union Internationale des télécommunications comme étant « un réseau de réseaux qui permet, via des systèmes d'identification électronique normalisés et unifiés, et des dispositifs mobiles sans fil, d'identifier directement et sans ambiguïté des entités numériques et des objets physiques et ainsi de pouvoir récupérer, stocker, transférer et traiter, sans discontinuité entre les mondes physiques et virtuels, les données s'y rattachant. »<sup>11</sup>

La montre connectée est le parfait exemple pour décrire l'internet des objets. Cette montre peut en effet, grâce à des capteurs, déterminer la fréquence cardiaque et d'autres paramètres physiques en temps réel et les transmettre numériquement sur la montre ou sur le téléphone en passant par un Cloud.

L'internet des objets est en partie responsable de l'augmentation exponentielle du volume de données générées sur le réseau qu'on appelle aujourd'hui Big Data. L'ilot existe grâce à l'utilisation d'un grand nombre de nouvelles technologies dont les plus importantes sont les systèmes Cyber-Physiques (CPS) et le Cloud computing.

Comme le propose Monsieur Hohmann Christian, un système cyber-physique interagit avec son environnement et collecte des données, les traite, et ainsi contrôle ou influence le processus auquel il est associé.<sup>12</sup> Par définition ce système peut interagir avec d'autres systèmes de manière filaire ou sans fil. C'est lorsque la connexion entre ce système est sans fil et utilise internet que le CPS devient la base de l'internet des objets.

L'internet industriel des objets fonctionne sur le même principe. Une machine est équipée de capteurs qui collectent les données en temps réel et qui vont communiquer à d'autres machines ou humains en passant par le Cloud ou un réseau interne à l'entreprise. L'IIOT, associé à un système cyber physique, constitue un des principaux fondements des industries du futur concernant la gestion et l'analyse des données. Le CPS peut, en effet, collecter et analyser les données issues des machines et de process pour faciliter la prise de décision qui devient plus réactive et optimisée. Il permet de maximiser la production et le rendement en améliorant entre autres les performances opérationnelles.

Par exemple, si une machine à un rendement diminué, les détecteurs placés directement sur la machine vont initier une requête de service humain, robotique ou directement lié à l'équipement pour résoudre la problématique.<sup>3</sup>

## D. Le Machine Learning

Avant de parler du Machine Learning, il semble indispensable de définir ici l'intelligence artificielle. John McCarthy, un des pères fondateurs de l'IA le définit comme « *La construction de programmes informatiques capables d'accomplir des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisantes par des êtres humains.* ». <sup>13</sup> Il s'agit donc de rendre des machines intelligentes. Depuis sa création deux approches se sont opposées pour « rendre les machines plus intelligentes ». La première se base sur l'idée que la pensée humaine peut être assimilée à un traitement de l'information obéissant à des règles. Dans ce modèle, créer une IA correspond à la modélisation de règles à travers du codage. Cette méthode était dominante jusque dans les années 1990. Depuis, l'approche du Machine Learning est née. <sup>13</sup>

Le « Machine Learning » appelé en français : l'apprentissage automatique est donc un sous-domaine de l'intelligence artificielle. D'après Arthur Samuel, pionnier de l'intelligence artificielle basée sur l'apprentissage, le Machine Learning est défini comme le domaine d'étude qui donne aux ordinateurs la capacité d'apprendre sans les programmer explicitement à partir de modèles statistiques et surtout d'algorithmes. <sup>14</sup> L'idée est d'alimenter au maximum les algorithmes avec des données informatiques récoltées en réel. Le traitement et la corrélation de ces données par les algorithmes vont permettre d'établir un schéma décisionnel prédictif. C'est avec cette technologie que sont développées par exemple les voitures autonomes.

Des capteurs présents sur des voitures réelles enregistrent un très grand nombre de données qui vont être envoyées à un algorithme en développement. Cette technologie est très utile pour les entreprises afin de corréler un grand nombre de données afin de prendre la meilleure décision. Cela montre aussi l'importance du « Big Data » évoqué plus haut.

Avant de traiter des différents types de Machine Learning, il est important d'établir que les explications qui suivent cherchent à vulgariser les concepts. En effet, le Machine Learning est un domaine vaste, qui pour être décrit dans les détails, nécessite des connaissances profondes

en mathématiques en informatique et en programmation. Ainsi, les applications mathématiques et de développement ne seront pas abordées dans cette thèse.

Il existe différents types d'algorithmes composant le Machine Learning utilisés différemment en fonction du besoin et des données à traiter<sup>14</sup>, comme le montre la figure 1 ci-dessous.

## 1. L'apprentissage Supervisé

L'apprentissage supervisé consiste à superviser l'apprentissage de la machine. A partir d'exemples **balisés** (on peut aller jusqu'à des millions de données), la machine va créer une fonction de prédiction permettant de prédire un résultat recherché. Par exemple, si l'on cherche à classer des photos de chat et de chien, on va fournir à l'algorithme un grand nombre de photos représentant ces deux animaux. Ces photos seront balisées par chien ou chat. C'est à partir de ces exemples que l'algorithme va créer sa fonction de prédiction permettant ici de classer de nouvelles photos.

Il existe des algorithmes de régression et des algorithmes de classement. Les algorithmes de classement vont prédire des variables qualitatives comme photos de chien ou de chat, alors que les algorithmes de régression vont prédire des variables quantitatives comme une probabilité qu'une machine tombe en panne.<sup>15</sup>

À partir des exemples balisés fournis, il va y avoir une phase d'apprentissage de l'algorithme. Ce dernier va s'entraîner sur une grande quantité de données, d'où l'importance des Big Data. Dans un second temps, on va fournir à l'algorithme des nouvelles données avec lesquelles il devra interagir de la même manière. Les algorithmes vont pouvoir prédire ou classer les données que l'on va lui donner. Ces algorithmes nécessitent une action humaine externe pour baliser les données d'apprentissage.

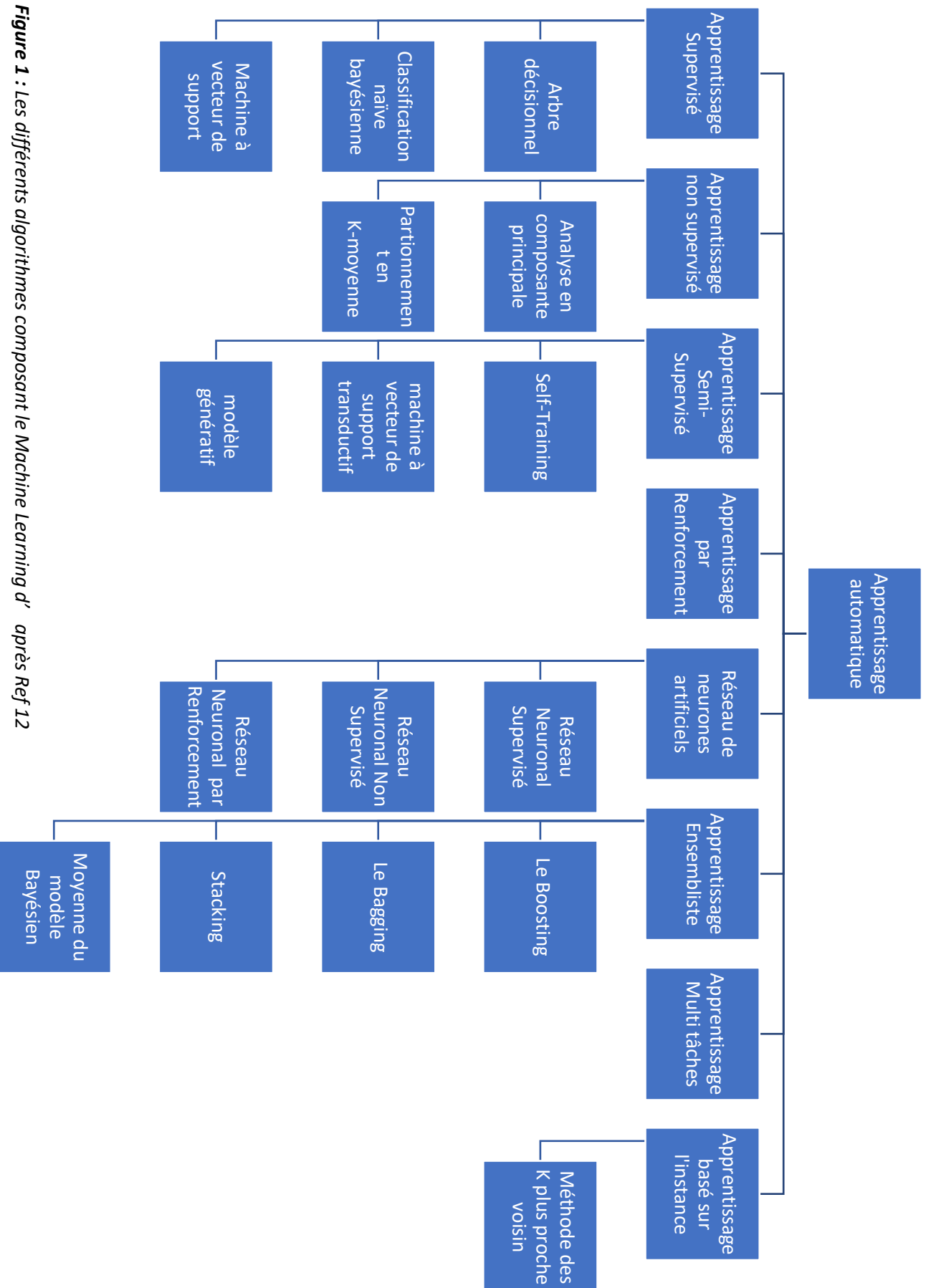


Figure 1 : Les différents algorithmes composant le Machine Learning d' après Ref 12

## a. Les Arbres décisionnels

L'arbre décisionnel est un graphique ou logigramme permettant de représenter des résultats en fonction de choix. C'est communément un outil simple pour faciliter la prise de décision et pouvant être appliqué dans tous les domaines. Le développement des mathématiques appliquées a permis la création d'algorithmes capables de créer des arbres décisionnels sans introduction des règles de façon manuelle.<sup>16</sup>

Les algorithmes ont pour but de déterminer la meilleure caractéristique (les choix possibles de l'arbre de décision) à partir des données fournies. Il divise les données en sous-ensembles contenant les valeurs possibles de la meilleure caractéristique. Il génère ensuite indéfiniment de nouveaux arbres décisionnels en utilisant les sous-ensembles créés et s'arrête lorsqu'il n'est plus possible de classer les données.<sup>16</sup>

## b. La classification naïve Bayésienne

La classification bayésienne est basée sur le théorème de Bayes. Celui-ci est un théorème de probabilité utilisé en statistique qui permet de déterminer la probabilité qu'un événement se produise sachant qu'un autre s'est déjà déroulé, notamment quand les deux événements sont interdépendants.<sup>17</sup>

$$P(Y/X) = \frac{P(X/Y) P(Y)}{P(X)}$$

Figure 2 : Expression mathématique du Théorème de Bayes

Où X et Y sont les événements :

$P(Y/X)$  est la probabilité conditionnelle de Y sachant que X se produit (Probabilité postérieure)

$P(X/Y)$  est la probabilité conditionnelle de X sachant que Y se produit (Probabilité postérieure)

$P(Y)$  et  $P(X)$  sont les probabilités antérieures d'apparition de Y et X indépendamment l'un de l'autre

La probabilité antérieure est la probabilité qu'un événement se passe avant la collecte de nouvelles données. Lorsque l'on reçoit de nouvelles données, la probabilité antérieure est révisée pour obtenir une probabilité plus précise.

La classification bayésienne regroupe les algorithmes **classifieurs** qui permettent de hiérarchiser les entrées en fonction des caractéristiques. Par exemple, si on cherche à classer les individus par sexe en fonction de la taille, du poids et de la pointure ; l'algorithme va permettre pour une entrée inconnue de déterminer les probabilités que le sexe soit féminin ou masculin.

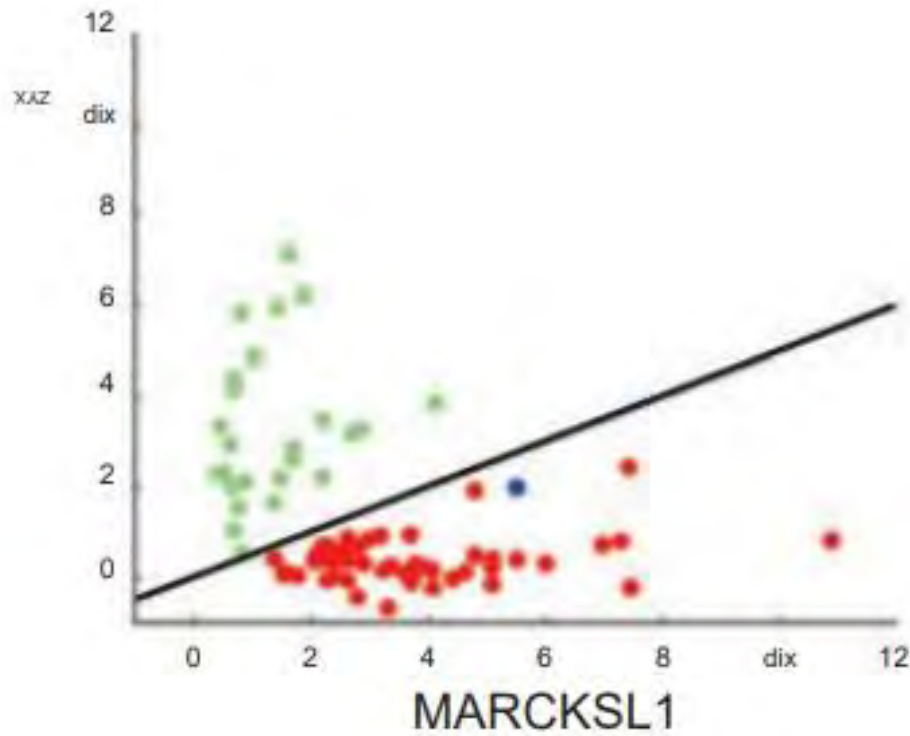
La classification bayésienne est dite naïve car l'hypothèse émise définit les variables du modèle comme indépendantes et simples. Par exemple, on va déterminer si un objet est une voiture en fonction de son poids de sa longueur et du nombre de roues. L'avantage de cet algorithme est qu'il ne nécessite pas un grand volume de données pour la phase d'apprentissage. Son exécution est donc rapide.<sup>18</sup>

### **c. Les machines à vecteur de support (SVM)**

Les SVM correspondent à des modèles de Machine Learning supervisé composé d'algorithmes qui analysent les données pour de la classification ou de l'analyse par régression.

Le but du SVM est de séparer les données en différentes classes avec une frontière, appelée hyperplan, en faisant en sorte que la distance entre les groupes de données et la frontière soit maximale.<sup>19</sup>

La figure ci-dessous montre les profils d'expression bidimensionnels de la leucémie lymphoblastique (LAL) et de la leucémie myéloïde aiguë, elle illustre l'hyperplan : ligne droite séparant les points rouges des points verts.



**Figure 3** : Exemple de graphique SVM de séparation de données linéaires pour différencier une leucémie aigüe lymphoblastique d'une leucémie aigüe myéloïde en utilisant l'expression du gène MARCKSL1 chez plusieurs échantillons. Ref<sup>19</sup>

Les points verts représentent les leucémies aigües lymphoblastiques, les points rouges représentent les leucémies aigües myéloïdes. Le point bleu représente un inconnu. Autrement dit un profil qui n'a pas pu être déterminé par l'algorithme. On observe bien un hyperplan entre les deux secteurs (rouge et vert) représentant les deux classes.

Quand les données sont linéairement séparables, on peut obtenir ce genre de graphique qui permet de séparer les données en deux classes par une droite (hyperplan). Mais ces cas sont rares. La plupart du temps, pour les données non linéaires, on utilise la fonction noyau. Ce noyau apporte une solution au problème en ajoutant des dimensions supérieures (par exemple mettre les valeurs d'origine au carré). On projette donc les données vers un espace vectoriel de plus grandes dimensions.<sup>19</sup>

Cet algorithme permet par exemple de reconnaître l'utilisation frauduleuse de carte bancaire de l'utilisation légale d'une carte bancaire<sup>17</sup>. En biologie, le SVM a aussi déjà été utilisé dans différents processus très variés; par exemple : pour classer des séquences de protéines ou d'ADN. Il a aussi été utilisé pour examiner le profil d'expression génique et en classant le

différent profil comme le montre l'exemple de la *figure 3*<sup>19</sup>, afin de déterminer un diagnostic ou un pronostic.

## **2. L'apprentissage non supervisé**

Ces algorithmes ne nécessitent pas d'interventions humaines. Les données fournies ne sont pas étiquetées. Par exemple, si l'algorithme cherche à différencier des emails valides des spam, en fournissant à l'algorithme des e-mails sans préciser s'il s'agit de spam ou d'e-mail valide, cette fois l'algorithme va de lui-même créer des classes présentant des caractéristiques communes (s'apercevoir de la différence et faire le tri entre spam et mail valide) là où pour le Machine Learning supervisé, une action humaine a permis d'élaborer au préalable les classes (étiquettes/balises). Il existe deux principales méthodes d'apprentissage non supervisé, les algorithmes de K-moyenne et l'analyse en composante principale.<sup>20</sup>

### **a. Analyse en composante principale (ACP)**

L'analyse en composante principale est une méthode statistique qui permet de transformer des données corrélées entre elles en données non corrélées appelées composantes principales. Le but est de diminuer le nombre de variables et donc de synthétiser l'information. L'idée est de créer de nouvelles variables à partir des variables existantes tout en restant pertinent dans l'analyse.<sup>14</sup>

Si on essaye par exemple de décrire les différentes personnes présentes dans une classe, on peut se baser sur un grand nombre de dimension comme la taille, l'âge, le poids, la couleur de cheveux etc. L'analyse en composante principale va par exemple associer le poids et la taille afin de créer une nouvelle dimension appelée la corpulence.<sup>21</sup>

Un trop grand nombre de dimensions est en effet problématique pour la communauté scientifique pour 3 raisons principales :

- La visualisation des données devient difficile
- Le calcul mathématique par les ordinateurs prend beaucoup plus de temps
- Les algorithmes font de la surinterprétation quand les dimensions sont trop grandes. C'est-à-dire que l'algorithme ne va pas pouvoir généraliser son classement car les

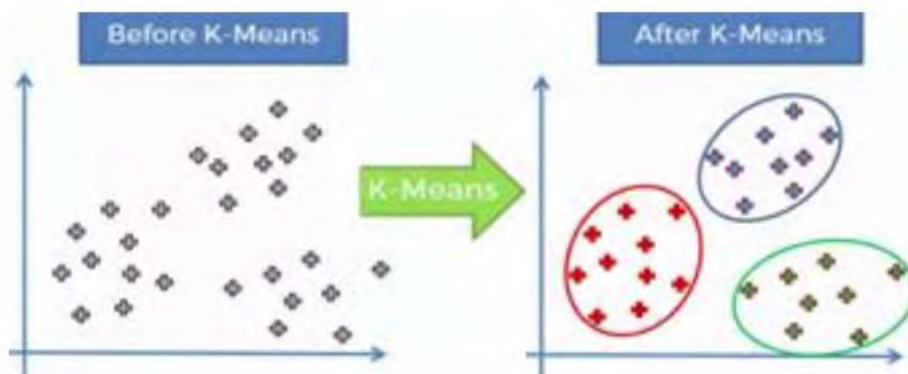


données reçues sont trop spécifiques (trop de dimensions). Il ne pourra pas reproduire le résultat avec un autre **jeu de données**.

Finalement l'ACP est utilisé pour plusieurs raisons : la principale raison est de faciliter la visualisation des données, Il permet aussi de décorrélérer les variables et de diminuer le **bruit**. Enfin l'ACP permet d'effectuer une réduction de dimension des données pour éviter le sur-apprentissage des algorithmes lors de la phase d'apprentissage du Machine Learning.

## b. Le positionnement en K-moyenne

Cette méthode est principalement utilisée pour le positionnement de données elle consiste à regrouper ensemble des données similaires dans des clusters (Clustering)<sup>20</sup>. Parmi différentes familles d'algorithmes de clustering ; le K-means est une méthode de classement non hiérarchique en K clusters où K est prédéfinis.<sup>22</sup>



**Figure 4** : Données classées en Cluster Grace au K means Ref 12

Le but est de classer les données en fonction des caractéristiques communes dans un cluster.

Pour ce faire, l'algorithme va effectuer 3 étapes à répétition :

- 1- Prendre K données au hasard du jeu de données qu'on va appeler centroïde.
- 2- Il va ensuite placer chacune des données en fonction de leurs distances au centroïde.
- 3- Calculer le centre de gravité (barycentre) des groupes qui devient le nouveau centre

Cela revient donc à étiqueter les données. Il va répéter ces trois étapes jusqu'à ce que les nouveaux centroïdes ne diffèrent pas des précédents. Cela correspond à la distance la plus faible entre les données et les centroïdes.<sup>23</sup>

### **3. Apprentissage Semi-Supervisé**

L'apprentissage semi-supervisé est un mélange entre le Machine Learning supervisé et non supervisé.<sup>24</sup> Il utilise un grand nombre de données non étiquetées et un petit nombre de données étiquetées. Il a été démontré que ce mélange de données permet d'améliorer significativement l'étape d'apprentissage des algorithmes. Le Machine Learning semi supervisé peut être transductif ou inductif. S'il est transductif, l'algorithme a pour but de déduire les étiquettes pour les données non étiquetées. L'algorithme se base sur des données d'entraînement spécifique pour résoudre une problématique composée d'un dataset spécifique. On ne cherche pas à créer un modèle général qui pourrait traiter un grand nombre de nouvelles données.<sup>25</sup> S'il est inductif, il étiquette les données non étiquetées et les classifie.<sup>26</sup> Inversement, on raisonne à partir de cas spécifiques vers des règles générales qui sont ensuite appliquées au cas de test.<sup>25</sup>

#### **a. Self-Training**

Ici, la phase d'apprentissage est supervisée et donc appliquée sur des données étiquetées. L'algorithme est ensuite appliqué sur des données non étiquetées. L'algorithme va alors prédire les « pseudo-étiquettes » des données. On parle de pseudo-étiquette tant quand les étiquettes ne sont pas définitives. On va ensuite sélectionner les étiquettes qui ont le plus de probabilité d'être correctes. On ajoute les données « pseudo étiquetées » aux données d'entraînement initiales étiquetées. On relance une phase d'apprentissage avec le nouveau jeu de données. On peut répéter ces étapes jusqu'à qu'à ne plus avoir de pseudo-étiquette probable. Enfin, on peut utiliser cet algorithme pour étiqueter des données similaires non étiquetées.

#### **b. Machine à Vecteur Support Transductif (TSVM)**

Comme décrit plus haut le SVM a pour but de déterminer la limite de décision pour l'appartenance des données étiquetées à un groupe ou à un autre, le TSVM utilisant des données non étiquetées cherche à baliser les données de façon à ce que la distance entre la limite (hyperplan) et les données soit maximale.<sup>27</sup> La présence des données non étiquetées va aider l'algorithme à classer les données.

### c. Modèle Génératif

Un modèle génératif est avant tout un modèle de classification. Il est dit génératif, car il permet de créer des données à partir des exemples de données lors de la phase d'apprentissage. Il se doit de comprendre la distribution des données afin de générer des données ressemblantes. C'est dans ce sens qu'on peut recréer des images à partir d'un ensemble d'images fournies au modèle<sup>28</sup> ; le domaine des modèles génératifs profonds, et son intérêt dans la reconnaissance d'images, en font, par exemple l'un des domaines prometteurs de l'imagerie médicale.

Il existe principalement deux types de modèle génératif. Le premier est le **réseau adversarial génératif (GAN)**. Le GAN implique l'utilisation de deux modèles d'algorithmes différents. Le premier est le générateur d'image qui va recréer des images à partir des données d'entraînement. Ce premier modèle est souvent un réseau neuronal artificiel. Le deuxième est le discriminateur. Il apprend à distinguer les vraies images des fausses images créées par le générateur. Ce dernier sera donc un **classifieur**. Au fur et à mesure de leur confrontation, le générateur va de plus en plus souvent, réussir à tromper le discriminateur en réussissant à créer des images de plus en plus proches de la réalité. De son côté, le discriminateur va de mieux en mieux discerner les différences entre les images réelles et les fausses.<sup>29</sup>

Le deuxième modèle est l'**autoencodeur**. C'est un modèle en réseaux de neurone artificiel composé de deux parties : l'encodeur et le décodeur. L'encodeur traite les données d'entrée. Il les compresse pour garder les informations essentielles et diminue la dimensionalité. Le décodeur décompresse ces données et reconstitue les données. En se basant sur les informations sélectionnées par l'encodeur, le décodeur va essayer de reconstituer l'ensemble des données le plus fidèlement possible à la réalité. Il permet donc de créer une nouvelle représentation de données à partir des données d'entrée non étiquetées. Ainsi, le nombre de neurones sur la couche d'entrée est le même que sur la couche de sortie. Comme pour l'ACP, il permet aussi de synthétiser l'information en gardant les données essentielles.

## 4. Apprentissage par renforcement

Contrairement aux apprentissages supervisés et non supervisés, on cherche à influencer le comportement d'un algorithme d'intelligence artificielle par des récompenses ou des pénalités selon que l'algorithme se comporte comme on le souhaite ou non.<sup>20</sup>

C'est une méthode qui demande à un algorithme d'atteindre un objectif dans un environnement incertain.<sup>30</sup>

L'ordinateur va chercher toutes les combinaisons possibles pour atteindre les objectifs et va apprendre de ses erreurs grâce au système de récompense. Ce type d'apprentissage est par exemple utilisé pour apprendre aux algorithmes à jouer à des jeux (en choisissant les bons coups à jouer).

Le système se base sur le processus de décision Markovien.

Le processus de décision Markovien est un formalisme mathématique. Pour un processus donné, à chaque étape, le processus est dans un état **E** et l'IA va prendre une décision **D**. Lors de la prise de décision, l'état va alors passer dans l'état **E'**. A chaque étape passée **E'**, l'agent va recevoir une grande récompense ou une récompense plus petite. Pour que le processus soit Markovien, la décision prise ne dépend pas de l'état précédent mais de l'état **E** présent. La récompense perçue par l'ordinateur est un nombre réel, qui peut être plus ou moins élevé en fonction des décisions.

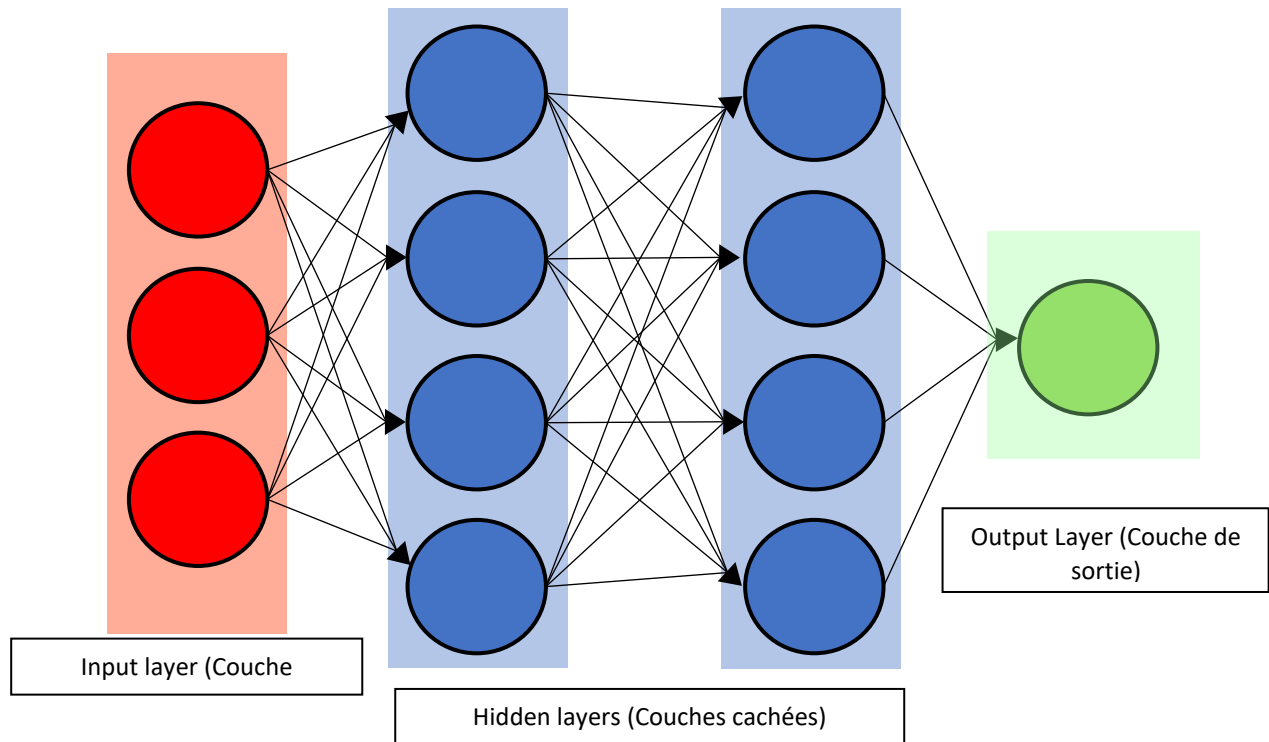
Dans un modèle d'apprentissage par renforcement, l'agent (ou l'IA) ne connaît pas le modèle. Il doit connaître les conséquences de ses actions en interagissant avec l'environnement.<sup>31</sup>

Prenons l'exemple d'une machine à sous avec 3 bras différents. L'algorithme ne connaissant pas son modèle va tirer au maximum sur tous les bras et comparer lequel va donner la plus grande récompense mathématique.

À force de s'entraîner l'IA va chercher à passer toutes les étapes en récoltant la récompense la plus élevée possible. On dit qu'il génère une « politique optimale ».<sup>31</sup>

## 5. Réseau de neurone artificiel

Un réseau neuronal artificiel est un sous-domaine du **Deep Learning**. Ce réseau de neurones fonctionne sur le même principe que le système neuronal du cerveau humain. Ici, c'est une série d'algorithmes qui s'efforcent de reconnaître les relations sous-jacentes à un ensemble de données.<sup>14</sup>



**Figure 5 :** schéma d'un réseau neuronal artificiel en propagation avant

Schématiquement, un réseau de neurones se divise en plusieurs couches. La première est « l'Input Layer » (couche d'entrée de données). La dernière couche, « l'Output Layer » est une couche de sortie de données. Le nombre de couches intermédiaires pouvant se situer entre deux et l'infini sont des « Hidden Layers » (couches cachées) et correspondent au réseau neuronal.<sup>32</sup>

Chaque tiers va recevoir les données des couches précédentes traitées et ainsi de suite jusqu'à sortir les résultats dans l'Output Layer. Toutes les couches sont reliées entre elle par des connexions appelées des poids. Ces poids sont des valeurs mathématiques qui permettent de transmettre l'information d'un neurone à l'autre. A chaque neurone est associée une fonction d'activation qui va plus ou moins activer le neurone. En fin de parcours au niveau de « l'Output Layer », on va retrouver pour chaque neurone final, correspondant chacun à un résultat : une

probabilité. Au début de la phase d'apprentissage, les résultats seront aléatoires. Il faut alors faire varier les valeurs de poids après chaque test afin d'obtenir le meilleur résultat. Pour ce faire, on compare mathématiquement les résultats donnés avec les résultats attendus. C'est ce qu'on appelle le coup. Le but est de répéter l'expérience un grand nombre de fois jusqu'à obtenir le coup le plus faible.

Comme pour tout algorithme, il faut une phase d'entraînement longue avec un grand nombre d'essais afin que l'algorithme puisse fonctionner correctement.

Ces algorithmes sont utilisés par exemple pour différencier des images. Prenons, par exemple, 10 images de différents chiffres numériques allant de 0 à 10. Les inputs seront donc les caractéristiques de l'image. A chaque neurone correspond un pixel de l'image. L'output Layer sera composé de 10 neurones correspondant aux probabilités que l'image représente un des dix chiffres. Les « Hidden Layers » représentent le traitement de l'information transmis de tous les pixels.<sup>32</sup>

A ce système, on peut appliquer les 3 méthodes définies précédemment qui sont les algorithmes supervisés, non supervisés et l'apprentissage par renforcement.

Pour un système supervisé, les couches d'entrée et de sortie sont connues et comme expliqué précédemment, on va changer les paramètres pour se rapprocher du résultat attendu.<sup>14</sup>

Pour un système non supervisé, l'algorithme doit catégoriser les données en fonction de certaines similitudes. Le réseau vérifie la corrélation entre les différentes entrées et les regroupe.<sup>14</sup>

Pour un système d'apprentissage par renforcement, il s'agit d'un algorithme orienté vers un objectif. Le but du réseau est de maximiser les points et sera « pénalisé » lorsqu'il prend de mauvaises décisions et récompensé quand il prendra les bonnes.<sup>14</sup>

## **6. L'apprentissage ensembliste**

L'idée principale de l'apprentissage ensembliste est d'utiliser plusieurs algorithmes ensemble afin d'obtenir de meilleures prédictions. Il est important de comprendre qu'un même algorithme peut créer plusieurs modèles de prédiction. Il est possible alors d'utiliser un seul algorithme, mais de combiner l'ensemble des modèles de prédiction afin d'affiner le

résultat. En effet, en regroupant les modèles, on peut garder la prédiction majoritaire de l'ensemble des modèles. On augmenterait ainsi la probabilité d'obtenir une bonne réponse. C'est ce qu'on appelle « **the wisdom of the crowd** ». <sup>33</sup> Dans ce cas, la force d'un algorithme compensera la faiblesse des autres et vice-versa. Aussi, il est important de souligner que selon la **loi des grands nombres**, il existe 3 exigences pour que le résultat apporté par les modèles se rapproche de 100% de véracité :

- Plus on augmente le nombre de modèles, plus on va se rapprocher d'un résultat de 100%
- Il est nécessaire que la prédiction de chaque modèle soit supérieure ou égale à 50% de véracité. S'ils sont inférieurs à 50%, on se rapproche d'une prédiction à 0%
- Il est important que l'ensemble des résultats obtenus par les modèles sur les différents échantillons (ou sous-échantillons) soit varié. Il faut donc que les modèles soient variés. Les faiblesses des uns seront compensées par les forces des autres.

Cette méthode permet d'entraîner des modèles différents sur un même ensemble de données grâce aux moyennes du modèle bayésien. Mais il permet aussi d'entraîner un même modèle d'algorithme sur différents ensembles de données grâce aux approches « Bagging » ou au « Boosting », décrites un peu plus loin. On combine ensuite ces deux approches afin d'avoir un résultat le plus prédictif possible. <sup>34</sup>

Afin de mieux expliquer l'utilisation des deux méthodes ensemblistes, il est nécessaire de définir deux concepts statistiques importants lors de la phase d'apprentissage. Le premier est le biais. Le biais est la différence entre la prédiction moyenne du modèle et la valeur correcte à prédire ou lorsque la prédiction est différente de la réalité. Très souvent, il est dû à une hypothèse de départ trop forte qui ne prend pas en compte assez de caractéristiques. Par exemple, si on tente de prédire à quelle classe appartiennent les molécules en fonction de leur structure. Dans ce cas, ne prendre en compte que le nombre de carbone fournira un modèle pauvre. Un biais trop important va entraîner un sous-apprentissage de l'algorithme qui ne se généralisera pas bien sur un nouvel échantillon de données. <sup>35</sup>

La variance est le deuxième concept important. La variance prend en compte la fluctuation de l'échantillon. <sup>36</sup> Si l'échantillon de données a une variance trop forte lors de la phase d'apprentissage, l'algorithme s'adaptera trop fortement à l'échantillon. Il prendra en compte le bruit. En effet, il ne pourra pas se généraliser à un autre échantillon de données, c'est le sur-apprentissage. <sup>35</sup>

L'idée pour tous les algorithmes est de trouver un compromis entre la variance et le biais. C'est ce qu'on appelle le *dilemme biais/variance*. On cherche à obtenir un résultat le plus juste et prédictif possible. Les connaissances actuelles prédisent que la variance du modèle augmente et que le biais diminue en fonction de la complexité du modèle.<sup>36</sup>

### a. Bagging

Le « Bagging » ou « *Bootstrap agrégation* » correspond à une des méthodes ensemblistes parallèle les plus connues. Une méthode parallèle a pour but d'entraîner les algorithmes ou modèles créés en simultanée.<sup>37</sup> Il consiste à sous échantillonner l'ensemble des données utilisées pour l'entraînement de l'algorithme. A partir de ces sous ensemble, un même algorithme va pouvoir générer différents modèles. Ensuite, il faut ;

- Soit moyenner les résultats sortants des différents modèles (pour des **modèles de classification**), pour obtenir un résultat qualitatif.

- Soit voter (pour des **modèles de régression**).

Imaginons que l'on veuille déterminer si une molécule donnée est un alcool fort ou faible en fonction de sa structure. Si la moyenne des réponses est « alcool fort » alors la réponse sortante sera « alcool Fort ». En moyennant les résultats, on va diminuer la variance de l'échantillon sans augmenter le biais puisque en sommant les classifieurs, on obtient l'ensemble des sous-ensembles des échantillons.<sup>34</sup>

Le Bagging est utilisé pour les algorithmes à forte variance et en surapprentissage puisqu'il va permettre de diminuer le surapprentissage et la variance.

### b. Boosting

Le « Boosting » est une méthode ensembliste séquentielle. C'est-à-dire que des modèles, qui sont plutôt faibles, sont entraînés à la suite, et vont tenter de réparer les erreurs commises par les modèles précédents. Le But de Boosting est de pondérer les classifieurs créés.<sup>37</sup> Cette méthode permet d'augmenter l'importance des classifieurs donnant de bons résultats. Le classifieur de sortie finale est une moyenne pondérée des classifieurs construits à chaque itération avec un poids plus important pour les classifieurs ayant une précision prédictive plus



élevée (taux d'erreur plus faible sur les données d'apprentissage).<sup>34</sup> Sachant que le boosting sert à combiner plusieurs modèles chacun ayant un biais élevé mais différent, le résultat final permet d'obtenir un ensemble de modèle ayant de meilleur résultat avec un biais global beaucoup plus faible.

### c. Stacking

Aux modèles d'Ensemble Learning décrits précédemment comme le Bagging ou le Boosting, on peut ajouter le Stacking. Il s'agit ici de former un nouveau modèle qui va prédire le résultat final en se basant sur les prédictions des premiers modèles. Ce modèle va déterminer quels modèles sont fiables et lesquels ne le sont pas.<sup>38</sup> Très souvent, lors de l'utilisation de cette technique, on divise la base de données initiale en deux avec une partie pour entraîner les premiers modèles qui forment les apprenants de base. Et le second appelé le meta-dataset qui est utilisé pour entraîner le modèle de Stacking.

### d. Moyenne du modèle Bayésien

Le modèle Bayésien est un système ensembliste qui permet d'entraîner des modèles différents sur un même ensemble de données. Le modèle Bayésien a pour source les statistiques Bayésiennes qui, comme vu précédemment, sont un modèle statistique qui permet de prédire à posteriori la distribution de données en fonction des connaissances *a priori*. L'approche de cette méthode est de mettre en place des probabilités *a priori* que certains modèles sont plus efficaces que d'autres et de pondérer ces modèles.<sup>39</sup>

## 7. L'apprentissage Multitâches

Comme pour le réseau de neurones artificiels, c'est un sous-domaine du **Deep Learning**. Le principe de ce domaine est de résoudre plusieurs **tâches** distinctes, en se basant sur leurs similitudes. L'apprentissage automatique multitâche aidera à améliorer l'apprentissage d'un modèle en se basant sur les connaissances contenues dans les différentes tâches présentes.<sup>14</sup> Ce type de Machine Learning est encore peu utilisé mais s'avère très efficace lorsque l'on est

face à une très grande base de données ou les tâches des modèles utilisés présentent des degrés élevés de similitude.<sup>40</sup>

## **8. Apprentissage basé sur les exemples**

Ces algorithmes ont pour objectif d'apprendre par cœur les exemples d'apprentissage. Une fois la phase de test lancée avec de nouvelles données, l'algorithme va utiliser les informations d'apprentissage stockées pour répondre à la demande. Ces algorithmes sont utilisés pour la classification et la régression.<sup>41</sup>

### **a. Méthode des K plus proches voisins**

La méthode des K plus proches voisins est un algorithme supervisé qui nécessite donc des données d'entrées et de sorties étiquetées. L'algorithme va classer ou prédire le résultat des observations en se basant sur les observations d'entraînement mémorisées.<sup>42</sup> Pour une nouvelle observation, le travail de l'algorithme est de déterminer dans quelle classe (ou quelle valeur pour la régression) elle appartient. Le nombre de classes est prédéterminé puisqu'il s'agit d'un algorithme supervisé. Pour ce faire, il va comparer la nouvelle observation aux observations d'entraînement. L'algorithme va ensuite déterminer la distance entre la nouvelle observation et les observations d'entraînement. Parmi les observations d'entraînement similaires, l'algorithme va sélectionner les k voisins les plus proches de la nouvelle observation. K correspond aux nombres de données les plus proches de la donnée de référence appartenant à la base de données d'apprentissage choisie. K est prédéterminé afin de savoir combien nous voulons de donnée par classe.<sup>43</sup>

Si on fait de la classification, l'algorithme va associer la donnée à la classe contenant le plus grand nombre de k voisins. Si on fait de la régression on calcule la moyenne des k plus proches voisines et on attribue cette valeur à la nouvelle donnée.

## E. Les logiciels de pilotage de Production (MES)

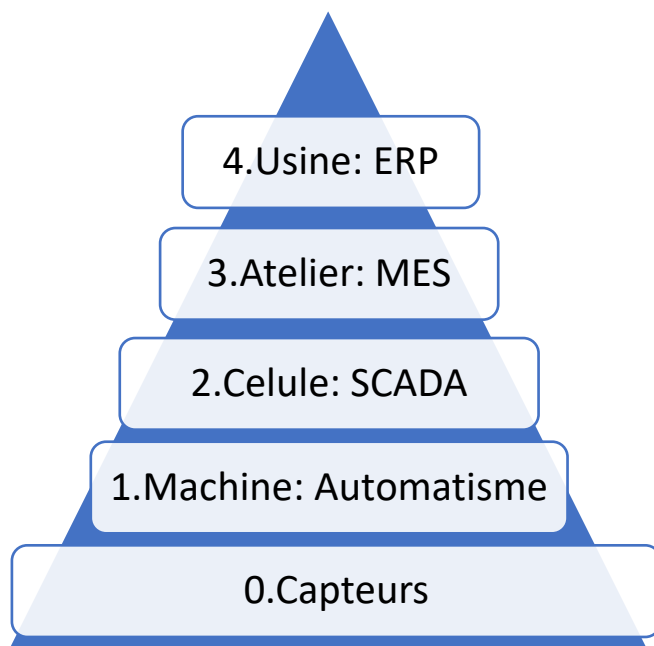
Le rassemblement et le traitement des données est devenu important pour les industries depuis que les nouveaux challenges industriels imposent de gagner en transparence, en réactivité et en rentabilité.<sup>44</sup> Pour ce faire, l'utilisation d'outils informatiques est devenue indispensable.

L'introduction de ces logiciels a débuté dans les années 1990.<sup>44</sup> Tout démarre dans les années 1980 avec la volonté des entreprises de collecter les données dans les différents services d'une usine, comprenant les services : RH, planification et logistique, fabrication, et qualité. Au début des années 1990, les logiciels de récolte des données évoluent ajoutant de plus en plus de fonctionnalités comme l'évaluation des données. Jusqu'alors, chaque logiciel de récolte de données était indépendant et ne pouvait effectuer qu'une seule tâche. Leur combinaison demandait un long travail humain. De ces logiciels, sont nées des systèmes de collecte combinée qui peuvent désormais exécuter plusieurs tâches. Ces logiciels sont ainsi devenus des modules pour l'« Entreprise Ressource Planning » (ERP).

L'ERP est l'outil que l'on retrouve le plus communément aujourd'hui, dans l'ensemble des secteurs industriels y compris le secteur pharmaceutique. C'est le **progiciel** qui permet de regrouper et évaluer les données des différents logiciels combinés ou modules (**computer-aided quality assurance** pour l'assurance qualité, **PDA** pour la production) dans une interface numérique. L'ERP est composé de différents modules intégrés (ou applications) qui communiquent. Il est multifonctionnel, et comme pour le MES, récolte les données des différents secteurs d'activité tels que la logistique, la production, l'assurance qualité etc.<sup>45</sup> Ce progiciel rassemble l'ensemble des données de l'entreprise afin d'améliorer l'organisation opérationnelle, accroître la réactivité par rapport aux clients, et fournir un accès des données à travers un système commun.<sup>45</sup> Cependant, depuis plus d'une vingtaine d'années, les exigences des clients sont de plus en plus strictes. Ceci oblige les fabricants industriels à être de plus en plus réactifs aux instabilités naturelles des processus de fabrication, de planification, de livraison, de stockage, de qualité etc.<sup>44</sup> Or, le système ERP ne collecte pas les informations en temps réel et les retours d'informations sont dans certains cas indisponibles. En effet, la collecte d'informations est de l'ordre de l'heure au minimum. Ceci s'explique par la nécessité dans certains cas comme en fabrication d'entrer les informations manuellement. C'est pourquoi la demande d'un système informatique permettant la récolte et le retour

d'information en temps réel est devenu de plus en plus importante. Le MES est une des réponses à cette demande.<sup>44</sup>

Globalement, un système de production peut être découpé en 5 niveaux de matériel informatique et automatique<sup>46</sup> :



**Figure 6** : pyramide de la **productique** (PERA) en 5 niveaux.

La pyramide de la productique (figure 6) est un modèle hiérarchisant les différents systèmes informatiques et le passage des informations, retrouvé entre une usine et le service de fabrication. Jusqu'à aujourd'hui, c'est le modèle le plus fréquemment rencontré dans les différents secteurs industriels même si la volonté future est de passer à un modèle en réseau grâce au Cloud.

L'ERP joue un rôle primordial dans la planification, le contrôle des stocks et la planification des ordres de fabrication, la gestion des clients et des commandes à long terme. Il permet d'avoir une vue globale d'une usine.

Le MES devient alors un atout pour l'exécution et le contrôle des tâches au quotidien. Il permet entre autres, l'exécution des ordres de fabrication, la gestion des lots, des mouvements physiques de ces lots ou de la localisation en temps réel des produits.

La cellule correspond au contrôle des machines. Les Systèmes SCADA nés dans les années 1980 sont des logiciels utilisés en fabrication pour la supervision et le contrôle des commandes des appareils automatisés comme l'augmentation du chargement, la visualisation des dernières mesures, la détection et l'affichage des différentes alarmes.

L'automatisme représente les actions automatiques programmées par un ordinateur pour le bon fonctionnement d'un type de machine. Ce sont des ordres précis dictés à une machine pour qu'elle puisse opérer. C'est un **automate programmable industriel** ou **API** qui est programmé pour commander une machine. <sup>46</sup>

Ainsi, le MES est un système de pilotage conçu pour être utilisé en routine dans le secteur de fabrication industrielle. Il permet de regrouper et de contrôler les données de fabrication. On l'utilise particulièrement pour contrôler l'ensemble des variables pouvant interagir dans la fabrication du produit, en temps réel, à l'aide de capteurs directement branchés sur les machines ou reliés à des ordinateurs pour la collecte de données.<sup>47</sup> C'est principalement le rôle du MES d'interconnecter et évaluer l'ensemble des données récoltées indépendamment, et d'avoir un rôle dans la fabrication du produit pour répondre aux problématiques en temps réel. De plus, le système MES est un acteur intermédiaire entre la planification à long terme et la fabrication à court terme. En ce sens, il permet la transmission des informations depuis l'ERP jusqu'aux ateliers, mais aussi le retour des informations de fabrication depuis l'atelier jusqu'à l'ERP.

En 2004, la norme ISA S95, publié comme norme international ISO/ IEC62264, décrit les fonctions de la gestion de la production, ainsi que les domaines d'activité d'un système MES, qui sont :

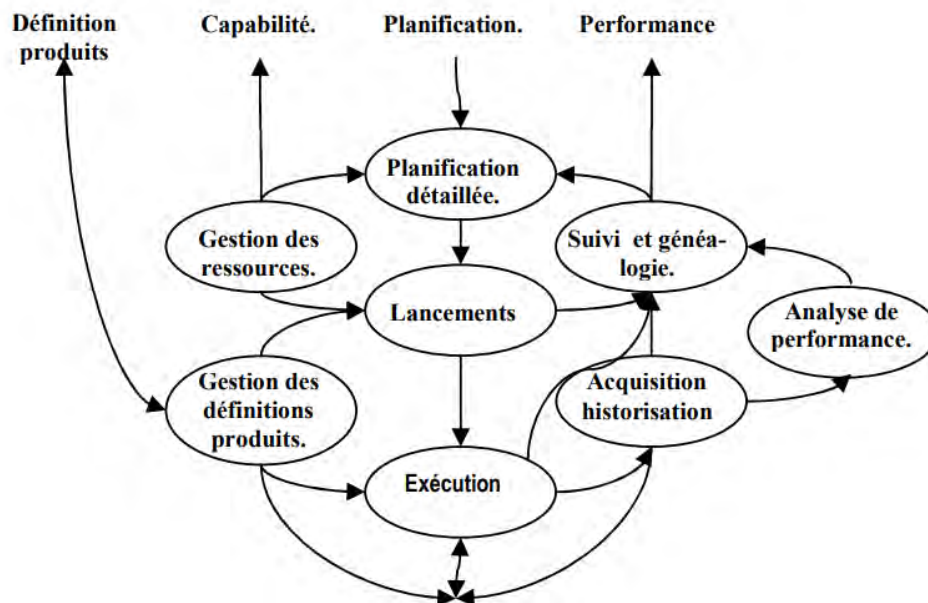
- **La planification détaillée** pour l'optimisation du temps et des séquences de fabrication en fonction des ordres de fabrication émis, mais aussi de la performance des machines et des ressources nécessaires
- **La gestion des ressources** permettant de déterminer et gérer les charges et ressources nécessaires aux ordres de fabrication
- **L'enregistrement et l'affichage de l'état actuel des ressources** renseignent en temps réel les opérateurs et dirigeants du statut de chacune des ressources nécessaires
- **La gestion des documents** est primordiale pour l'application des process, pour diminuer les risques d'erreurs, augmenter la traçabilité ou encore améliorer les

formations des nouveaux opérateurs. Un système MES permet alors d'améliorer la distribution des documents comme les instructions, les ordres de fabrication, les dossiers de lots etc.

- **La gestion des matières premières** et des produits intermédiaires utilisés en fabrication.
- **L'analyse de performance** permettant la comparaison entre les valeurs réelles de production et les valeurs théoriques planifiées, les objectifs clients et opérationnels.
- **La gestion des commandes** qui a pour but de définir les priorités de fabrication, et la répartition du travail entre les différents postes de travail.
- **La gestion de la maintenance** planifiant les maintenances préventives et alarmes sur une demande corrective des machines de fabrication afin d'optimiser leur utilisation et se rapprocher de la cible de performance recherchée.
- **La gestion des Processus** correspondant à la gestion du **Workflow**, conformément aux charges et spécification prévues.
- **La gestion de la qualité** qui analyse, suit et contrôle le produit et le processus tout le long du workflow en fonction des spécifications qualités demandées
- **La collecte et acquisition de données** permettant le rassemblement, l'analyse et la visualisation des données produites par l'ensemble des étapes de fabrication et des ressources matérielles ou humaines utilisées
- **Le suivi des produits et généalogie**, correspondant à l'enregistrement de l'historique de l'état des produits de fabrication depuis les matières premières jusqu'au produit finis et de tous les événements liés à ce produit.

Toujours en se basant sur la norme ISA S95, on peut regrouper l'ensemble de ces domaines dans un schéma modélisant le passage du niveau 4 au niveau 3 de la pyramide productique

(Le MES étant le messenger intermédiaire) :



**Figure 7 :** Les domaines d'activité fonctionnels d'un système MES selon la norme ISA S95

Ce modèle représente le flux des activités d'une usine depuis la demande du client jusqu'à l'exécution de la commande. Chaque ovale représente une activité et les arcs représentent le flux d'informations.<sup>48</sup>

Tout commence avec la commande du client qui est traitée par l'usine sous forme d'un plan de fabrication. Depuis la commande, on répartit et planifie le travail. On définit les ordres de fabrication, on alloue les ressources etc. On exécute ensuite les opérations planifiées, on collecte les données, puis les retourne comme des réponses à la planification. A cette boucle doit s'ajouter un système de gestion des ressources et des matières premières, un suivi et une bonne distribution de la documentation (dossier de lots etc.) et une analyse du travail effectué pour la traçabilité et une possible amélioration future.<sup>48</sup>

## F. Les Jumeaux Numériques

Un jumeau numérique est défini comme étant une réplique d'un objet ou d'un système réel physique en un système identique numérique. Cet exploit est possible grâce aux nouvelles technologies digitales comme les Big Data, l'ILOT ou encore l'IA (Machine Learning ou Deep Learning) décrit plus haut. En utilisant un modèle qui a pour but d'être prédictif, descriptif, ou

perspectif et les nouvelles technologies, on peut alors créer un répliquant numérique qui s'adapte en temps réel aux informations reçues par le système réel.<sup>49</sup>

On peut alors différencier trois stades d'utilisation du jumeau numérique :

- Le plus basique est le scénario où le monde physique envoie en continu les données au monde virtuel, permettant une visualisation et une analyse en temps réel des processus de fabrication, mais surtout pour l'expérimentation au stade Recherche et Développement (comme le changement de variables). L'environnement, le processus et les équipements sont fixes.
- Le second degré d'utilisation plus complexe, demande au système virtuel par la communication entre les deux systèmes physique-virtuel de rester opérationnel en cas de changement d'environnement comme, par exemple, le changement d'un équipement.
- La troisième possibilité est la simulation dans le monde virtuel de phénomènes qui permettront de prédire l'évolution d'un système physique, avec une modification directe des paramètres sur le monde physique en fonction de ces simulations depuis la modification des variables processus jusqu'au changement de l'approvisionnement en matières premières.<sup>49</sup>

## **G. La réalité augmentée (RA)**

La réalité augmentée est une forme d'**interaction Homme-Machine** basée sur l'association d'objet virtuel et physique. Le principe de la réalité augmentée est d'enrichir le monde réel par des informations virtuelles. L'expression « réalité augmentée » a été introduite en 1992 par Caudell et Mizell<sup>50</sup> qui présentaient un concept de lunette ajoutant des informations virtuelles sur les verres afin d'aider les opérateurs industriels à être plus performants, en ayant à leurs dispositions l'ensemble des informations nécessaires. Cette information virtuelle forme alors ce qu'on appelle « Head Up Display » (HUD).<sup>50</sup> Les informations sont produites par un ordinateur et superposées sur les lunettes. Ces informations peuvent être d'origine totalement numérique (tableaux ou graphiques) ou être digitalisées depuis la réalité (comme un objet en 3 dimensions). Aussi, la nature des informations peut être différente, elle peut



être visuelle ou auditive, provenir de graphiques digitalisés ou encore de graphiques représentant en temps réel l'environnement d'une machine grâce à un capteur.<sup>51</sup>

Aujourd'hui, il existe différents outils de réalité augmentée. Les plus communs sont les lunettes ou les vidéocasques qui sont des dispositifs portables. On peut aussi retrouver des dispositifs de projection comme des écrans.<sup>52</sup>

Pour utiliser la réalité augmentée, trois équipements sont nécessaires : des capteurs (un appareil photo ou vidéo pour l'environnement spatial par exemple), un processeur (ordinateur qui fait les calculs) et un écran de retransmission (écran).

Les capteurs ont pour rôle de transmettre les informations du monde réel au processeur. Les principales informations à transmettre sont la localisation et l'orientation du sujet sous RA, mais aussi les autres informations importantes pour l'application de la RA ou sur l'environnement. Les capteurs optiques sont les plus communs. Une caméra enregistre ce qu'elle « voit » et le retransmet à l'application de RA.

La notion d'enregistrement spatial et temporel est très importante en RA. En effet, quoi que l'on digitalise dans un environnement donné, la localisation de l'objet digitalisé restera initialement toujours la même et peut être perçue en 3 dimensions, quel que soit la localisation du sujet. La notion de temporalité se concentre sur le fait qu'à chaque mouvement du ou des sujets en AR, l'ordinateur recalcule à chaque mouvement pour que la représentation reste adéquate au modèle défini. Par exemple, si l'objet digitalisé est un livre en 3D, lorsque le sujet tourne autour du livre, l'ordinateur calcule en continu pour permettre au sujet de visualiser le livre sous tous ses angles.<sup>51</sup>

## **H. L'impression 3D**

L'impression 3D est de plus en plus utilisée en industrie. Elle permet de créer physiquement en 3 dimensions des objets créés numériquement, en ajoutant de la matière, couche après couche. Les matières utilisables sont diverses. On peut utiliser les plastiques, les métaux, la céramique, les matières organiques etc. C'est une technique de production additive. La production additive est la création à partir d'un fichier numérique, d'un objet réel. Elle est dite additive car la fabrication repose sur la formation de couches de matière superposées une à une.<sup>53</sup> Le standard ISO/ASTM 52900, la définit comme étant « un processus d'assemblage de

matériaux pour fabriquer des objets à partir de données de modèles 3D, généralement couche après couche, par opposition aux méthodes de fabrication soustractive. »

C'est grâce à un ordinateur que la machine crée les objets. Il va en effet, découper numériquement en une multitude de couches (de 0.05 à 0.1 mm pour la stéréolithographie par exemple) un objet numériquement créé et transmettre à l'appareil de fabrication, les ordres d'addition de matière par couche.

On reconnaît six grands types d'impression 3D dans le secteur de l'industrie pharmaceutique :

- *La Photopolymérisation* : une source lumineuse solidifie un matériau liquide formant un objet couche après couche. Par exemple, la stéréolithographie (SLA), la première imprimante 3D créée en 1986, permet de solidifier des **photopolymères** liquides grâce à un rayon laser ultraviolet.<sup>53</sup>
- *Le Dépôt de fil fondu ou Extrusion (FDM)* : C'est la principale technique utilisée aujourd'hui en industrie pharmaceutique. Elle consiste à fondre un polymère thermoplastique sous forme de filaments et de le faire passer mécaniquement dans une filière dont la section est choisie préalablement (c'est l'extrusion). Cet extrudât est alors déposé couche par couche afin de fabriquer l'objet.<sup>54</sup>
- *L'extrusion semi-solide* : c'est le même principe que le FDM, mais l'extrusion se fait à l'aide d'une seringue contenant un semi-solide composé de solvant, de principe actif et de polymère. Ce semi-solide va être extrudé couche après couche pour construire l'objet 3D.
- *La fusion de poudre (SLS)* : le procédé le plus commun, le frittage laser sélectif, consiste à consolider des particules poudreuses à l'aide d'un laser.<sup>55</sup> Pour l'industrie pharmaceutique, cela consiste à consolider une poudre contenant excipient et principe actif grâce à l'énergie fournie par le laser couche par couche jusqu'à obtenir le produit fini ( par exemple un comprimé).<sup>54</sup>
- *Le jet d'Encre* : C'est l'écoulement d'un liquide sous forme de gouttelettes, contenant des photopolymères, sur un substrat spécifique (le support) en fonction du modèle numérique. Dans un second temps, un traitement ultraviolet est appliqué sur la couche pour durcir la résine.<sup>54</sup>

- *Le jet de liant* : C'est une technologie dérivée de l'impression 3D par jet d'encre. Ici, c'est un liant sous forme de gouttelettes qui va être déposé sur un lit de poudre et qui va consolider les particules en fonction des jets. On répète cette opération couche après couche. La machine va réinstaller un lit de poudre pour chacune des couches jusqu'à la formation de l'objet. Les liaisons entre les particules de poudre se font soit par liaisons interparticulaires, soit par dissolution et recristallisation des particules humidifiées par les gouttelettes.<sup>54</sup>

-

## **I. La robotique collaborative**

La robotique collaborative est l'utilisation de robot conçu pour fonctionner et aider l'humain dans sa tâche. Concrètement, là où un robot industriel traditionnel va arrêter sa tâche lors de la présence de l'Homme pour des raisons de sécurité, un « cobot », ou dispositif d'assistance intelligente, définissant les robots collaboratifs, va travailler à proximité ou en relation directe avec l'opérateur.

Il existe 3 types de cas de travail pour les cobots :

- L'Homme et le robot partagent l'espace de travail, mais réalisent des tâches distinctes.
- Le robot et l'Homme travaillent à tour de rôle à la réalisation d'une même tâche. C'est une collaboration indirecte.
- L'Homme et le robot travaillent simultanément. C'est une collaboration directe.<sup>56</sup>

Les robots collaboratifs sont surtout utilisés pour réaliser ou aider à la réalisation des tâches pénibles au sein des usines de fabrication. Ils sont utiles pour aider les Hommes à réaliser des tâches qui demandent des connaissances, de l'expertise ou de la dextérité humaine, comme l'assemblage de pièces fines ou la découpe manuelle spécifique de métal. Ces tâches ne sont donc pas réalisables par des robots industriels. Il s'agit donc de mélanger le savoir humain avec la force l'endurance et la vitesse d'un robot.<sup>57</sup>

Ces cobots peuvent être utilisés pour des tâches pénibles comme porter des charges lourdes, ou répétitive visser ou dévisser. Ils peuvent aussi être utilisés en collaboration comme vu avec l'exemple au-dessus lors de la découpe d'une pièce de métal, le cobot articulé va porter la charge lourde et le matériel de découpe et l'homme va simplement le guider.

Dans un contexte industriel utilisant de plus en plus l'IA, la robotique et les nouvelles technologies, la communauté scientifique se tourne vers l'importance de centrer les processus autour de l'humain. Pour pouvoir agir en collaboration parfaite avec l'humain, les cobots les plus évolués doivent être en mesure de détecter l'humain, de comprendre ses émotions et d'interagir avec lui afin principalement d'assurer la sécurité de l'opérateur.<sup>58</sup> L'aspect et l'efficacité du robot ont aussi un rôle à jouer pour obtenir une collaboration fluide. C'est l'interaction Humain-Robot (HRI).<sup>59</sup>

Pour améliorer les caractéristiques du robot, et reconnaître les émotions, on utilise différents capteurs allant de la reconnaissance faciale ou vocale aux électrocardiogrammes et électroencéphalogrammes. L'idée est donc d'ajuster les opérations du robot en fonction du profil des travailleurs.

## **J. Cybersécurité**

Comprenant que l'industrie 4.0 se développe autour du numérique et des systèmes cyber physique, il devient alors important de protéger les données et les systèmes. Quand on parle de cybersécurité, on parle de sécurité de l'information et donc de la technologie de l'information (TI) mais aussi de la sécurité des systèmes de contrôle industriel (SCI) et des opérations et donc de la technologie opérationnelle (TO). L'IT est lié au stockage, la confidentialité, la récupération, la manipulation, et la transmission des données numériques au sein d'une entreprise. Il peut s'agir des mails internes comme les ressources réseaux, la capacité de stockage, la gestion de données, etc. L'OT, elle, se concentre sur le contrôle et la disponibilité des équipements et des processus. Elle gère les systèmes tels que l'ERP, le MES, les SCADA etc. Les deux pôles travaillent donc tous les deux sur des systèmes informatiques de manière parallèle sans pour autant collaborer. Ils utilisent tous deux des méthodes différentes.<sup>60</sup> L'utilisation de l'informatique devenant plus globale dans les usines industrielles, les données transitent entre le TI et le TO. Il faudrait donc faire converger les deux pôles pour centraliser la sécurité informatique. On parle de convergence TI/TO.

La globalisation de l'informatique et de la mise en réseau des données (IIOT et Cloud) rend les cyberattaques plus fréquentes et surtout plus dangereuses. Les raisons sont variées et peuvent

être du sabotage industriel, du vol de propriété intellectuelle ou de données, de l'espionnage informatique etc. Les attaques elles aussi peuvent être différentes:<sup>61</sup>

- Logiciels malveillants : Très commun dans le secteur industriel, c'est un logiciel qui permet d'infiltrer un système mal protégé et qui ensuite se rend aux autres systèmes connectés entre eux.
- Hacking : un individu malveillant va obtenir une autorisation d'accès à des données sensibles.
- Denial of Service: Il s'agit ici de rendre un système inaccessible de diverses manières comme rendre un capteur non fonctionnel, corrompre un système, saturant un serveur etc.
- Perte ou corruption de données rendant un système non fonctionnel.

Une des premières dispositions à prendre face aux cyberattaques est la prévention. Il s'agit par exemple<sup>62</sup> :

- De former les collaborateurs aux dangers des cyberattaques, des risques et des actions qui peuvent être menés individuellement pour diminuer le risque.
- La protection des logiciels en formant et modifiant régulièrement des comptes, pseudos, et mots de passes.
- La gestion des comptes est aussi importante il s'agit par exemple de supprimer un compte qui n'est plus utilisé.
- La cartographie du système industriel afin de savoir où les systèmes sont faibles et où ils sont forts, et de repérer rapidement une brèche lors d'une attaque.
- Augmenter la fréquence de sauvegarde
- Analyse de risques avant la mise en place d'un nouveau système industriel.
- Test de cybersécurité afin d'évaluer la protection des systèmes informatiques et des données.
- Contrôle d'accès physique aux installations.
- Le cloisonnement des différents systèmes industriels qui permet de limiter les conséquences des attaques et la diffusion de logiciels malveillants.
- Maitriser la sécurité des terminaux nomades comme les tablettes ou ordiphones.

- La prise en compte de la cybersécurité dans les projets internes de l'entreprise.
- Maîtrise de la cybersécurité chez les fournisseurs et prestataires.

*L'International conference on competitive manufacturing* a déterminé en 2019 que pour la mise en place d'une industrie 4.0 dans des conditions sécurisées deux points importants sont à prendre en compte. La première est l'implémentation de la Security by Design (sécurité par conception). Il s'agit de mettre la cybersécurité au cœur de la conception d'un logiciel ou d'un produit avant même sa mise en marche dans une usine. C'est une démarche préventive, qui cherche à diminuer le risque d'attaque.<sup>61</sup> Le deuxième point est que les stratégies de sécurité, et l'architecture des Technologie de l'information (TI) doivent être développés et implémentés à des degrés élevés de confidentialité entre les systèmes et les processus interconnectés.

De plus avec l'avènement de l'intelligence artificielle, on voit apparaître de nouveaux logiciels spécialisés dans la cybersécurité. *L'automatic exploit generation* est par exemple un logiciel intelligent permettant de détecter les failles de sécurité d'un système et de générer, en fonction des possibilités, des contre solutions. Aussi, avec l'arrivée du Cloud, on voit apparaître un nouveau service qu'est le « Security as a service ». Le système le plus basique est l'accès à des antivirus depuis internet. Le logiciel est souvent mis à jour par les fournisseurs. Le client est donc toujours équipé avec les protections les plus optimales.

## **K. Le quantum computing**

Le quantum computing est un ordinateur quantique. Avec le développement des connaissances physiques et informatiques, Le concept est d'utiliser la physique quantique comme base d'un ordinateur ce qui permettrait d'avoir des vitesses de calcul supérieures aux ordinateurs classiques. Deux lois basiques de la physique quantique sont à la base de la conception des ordinateurs quantiques:

- *La superposition d'état* : Un objet quantique existe dans plusieurs états différents tant que l'observation n'a pas été faite.
- *Le principe d'indétermination de Heisenberg* : on ne peut pas mesurer différentes grandeurs caractéristiques d'un objet simultanément. On ne peut pas connaître avec une bonne précision la position d'un objet quantique sans avoir une précision médiocre de sa vitesse.

Un ordinateur classique fonctionne avec des bits. Un bit est une unité élémentaire informatique pouvant prendre la valeur de 0 ou 1. C'est un code binaire. Le code binaire est le langage qui permet de transmettre les instructions de processeurs et qui font donc fonctionner l'ordinateur. Un ordinateur quantique fonctionne avec des qubits. Les qubits peuvent représenter un 0 ou un 1 simultanément. Chaque qubit a des probabilités différentes d'être dans un état ou un autre. Cette fonction permet globalement d'augmenter la capacité de calcul de l'ordinateur.<sup>63</sup>

## L. La Blockchain

La blockchain est avant tout une technologie de stockage de certification et de transmission de l'information. Elle permet aux utilisateurs de partager des informations sans intermédiaire. La mission d'information commune de l'assemblée nationale sur les usages de chaînes de blocs et autres technologies de certification de registre la définit comme étant un registre, une grande base de données qui a la particularité d'être partagée simultanément avec tous ses utilisateurs, tous également détenteurs de ce registre, et qui ont également tous la capacité d'y inscrire des données, selon des règles spécifiques fixées par un protocole informatique très bien sécurisé grâce à la **cryptographie**.<sup>64</sup>

Une blockchain comme son nom l'indique est constituée de blocs. Chaque bloc contient des centaines ou milliers de transactions numériques communes. Par exemple, un même bloc va contenir l'ensemble des contrôles qualité validés de l'ensemble des lots d'un même produit de fabrication. Les blocs sont liés les uns aux autres formant une chaîne, d'où l'appellation blockchain. Chaque transaction numérique ajoutée est vérifiée et contrôlée par des algorithmes appelés « mineurs » qui vont déchiffrer les données, identifier la traçabilité, vérifier les transactions connexes, et les comparer avec les opérations précédentes. La décentralisation de la sécurisation des données, empêche la falsification et augmente donc la transparence des données.

Ce système est avantageux pour les industries car la transmission des données est rapide et sécurisée ce qui augmente l'efficacité et la productivité globale.

### **III- L'application des outils 4.0 à l'industrie pharmaceutique**

#### **A. Etat des lieux de l'industrie pharmaceutique**

##### **1. L'évolution du modèle économique**

L'ère moderne de l'industrie pharmaceutique est caractérisée par des grandes entreprises, et ce, depuis la moitié du 20<sup>ème</sup> siècle. Jusqu'alors, le modèle économique était basé sur les « blockbusters », médicaments qui coûtent très cher à développer, mais qui possèdent un brevet d'une vingtaine d'année permettant d'obtenir des retours sur investissement. Cependant, ce modèle est aujourd'hui révolu, car les imprévus sont trop grands, la recherche du médicament coûte trop cher et les recherches basées sur la chimie tendent aujourd'hui à s'essouffler.<sup>65</sup> À partir des années 1980, les entreprises pharmaceutiques ont commencé à investir dans la biotechnologie et la pharmacogénomique. Les dirigeants croient aujourd'hui en la possibilité d'une médecine personnalisée qui serait lucrative et qui donnerait rapidement des résultats tangibles. Comme l'explique Xavier Deroy, désormais, un médicament véritablement efficace est relatif à un individu et à son patrimoine génétique.<sup>65</sup> Toutefois, les médicaments classiques sont toujours prédominants grâce à l'emploi de technologies informatiques qui permettent d'améliorer la R&D, comme le criblage à haut débit ou encore le criblage virtuel. Aujourd'hui, les « Big pharma » et les plus petites entreprises de biotechnologies cohabitent dans un environnement économique en évolution continue. En effet, les grandes entreprises pharmaceutiques financent ou concluent des contrats de licence avec les start-ups de biotechnologies externalisant ainsi la R&D.<sup>65</sup>

Dans un contexte de croissance continue de la demande, d'ouverture de marché vers des pays en développement, un vieillissement global des populations, et une demande de plus en plus exigeante des clients, des consommateurs et des acteurs réglementaires, la compétitivité et la flexibilité des industries sont devenues des enjeux primordiaux.<sup>44</sup>

Il en résulte une évolution de leur modèle économique, mais aussi de leur structure interne avec l'arrivée des nouvelles technologies. Ce contexte explique l'arrivée du concept d'industrie 4.0. C'est un changement de paradigme.<sup>65</sup>



## 2. La maturité Digitale

Comme dit précédemment, de nombreuses industries pharmaceutiques sont toujours en transition vers l'industrie 3.0 qui a pour but de digitaliser les données, les processus et l'automatisation de masse. L'industrie 4.0 se basant sur le traitement des données de masse (Big Data, IIoT, Cloud, ML), il est difficile pour les industries d'évoluer rapidement sans passer par un changement massif de la structuration des usines. Ces changements impliquent la digitalisation des données et l'abandon du support papier, la robotisation et l'automatisation des processus.

On parle de maturité digitale pour décrire le statut de transformation digitale d'une entreprise.<sup>66</sup> La plus grande partie des industries, tous secteurs confondus, s'intéressent à la digitalisation. En 2016 déjà, la commission Européenne « *Digitising European Industry* » établissait un plan d'action afin d'obtenir une entraide entre les entreprises pour optimiser la transition vers une digitalisation des données en industrie.<sup>67</sup> Différents modèles peuvent être utilisés pour évaluer la maturité digitale, mais ils ne seront pas tous détaillés dans cette thèse. Ces outils peuvent être établis par des cabinets de conseil dans le cadre d'accompagnement des entreprises dans une transformation digitale ; comme par exemple : « *le quotient digital* » proposé par la firme *Mckinsey* ou le *Deloitte Digital Maturity Model* proposé par le cabinet *Deloitte*. D'autres modèles sont établis par des associations. On peut citer le modèle *ACATECH* créé par l'Académie Allemande des sciences et de la technologie (association privée sans but lucratif).

En 2018, la maturité digitale des industries manufacturières a été évaluée par plusieurs groupes de consultants. Le Boston Consultant Group a alors déterminé que 70 % des industries pharmaceutiques se situaient dans le groupe des « retardataires ».<sup>66</sup>

Avant tout changement structurel, il est donc important de digitaliser les usines.

### B. Stratégie de digitalisation.

L'industrie 4.0 est finalement présentée comme un ensemble d'outils applicable afin d'améliorer la productivité, la rentabilité et la personnalisation de la production de masse<sup>66</sup>. Pour appliquer ces outils, les industries se doivent d'appliquer une stratégie permettant d'assurer une bonne transition vers une usine numérique puis intelligente. Différentes

stratégies peuvent être choisies qui selon les industries et le secteur orienteront le choix des outils optimaux qui permettront d'améliorer la qualité et le flux des processus. Il est par exemple possible d'associer certains outils de l'industrie 4.0 comme le RFID ou le MES, à une approche Lean afin d'atteindre un niveau de performance plus élevé.<sup>68</sup>

Cette partie se base sur une stratégie de digitalisation proposée par *Biophorum* en 2019. Ce modèle est pris comme exemple pour expliquer les différentes étapes à suivre pour arriver à un système digitalisé.<sup>69</sup> *Biophorum* est une collaboration mondiale de leaders de l'industrie et d'experts en la matière travaillant ensemble pour mettre en commun les connaissances, les pratiques et les idées. Cette collaboration a réuni 20 experts industriels de 11 industries pharmaceutiques différentes pour proposer un modèle qui décrit la maturité digitale en cinq niveaux, depuis l'utilisation du papier majoritairement jusqu'à l'obtention d'un système informatique autonome et adaptable. Ces 5 niveaux sont présentés dans le tableau suivant :

Niveau de maturité Digitale	Nom du Niveau	Description
1	Usine pré-digitale	Processus et manuel décrits et contrôlés sur papier
2	Silos numériques	Automatisation des systèmes, mais chaque machine est un îlot indépendant, il y a peu de transmission de l'information.
3	Usine connectée	Les systèmes sont interconnectés et les machines sont gérées par un système de gestion de ligne.
4	Usine prédictive	Optimisation des processus grâce à l'acquisition et l'archivage des données de production, utilisation de nouvelles technologies pour la prédiction de problématiques qualités et des déviations.
5	Usine adaptative	Numérisation depuis le fournisseur jusqu'au patient, les processus s'autorégulent. Usine centrée sur le patient.

**Tableau 1 :** Tableau des niveaux de maturité digitale des industries pharmaceutiques selon *Biophorum*<sup>69</sup>

*Biophorum* propose un outil décrivant les étapes à suivre pour digitaliser l'entreprise pour les secteurs de la production, de la qualité, de la logistique, et de la maintenance. Le passage de la phase 1 à la phase 4 correspondrait au passage à l'industrie 4.0. Le niveau 5, lui, est une ambition hypothétique de l'usine du futur qui ne peut encore être installée en vue des avancées technologiques actuelles, elle ne sera donc pas détaillée ici.

## 1. La production

La production est le maillon le plus difficile à digitaliser. En effet, chaque changement de process, de nettoyage, d'équipement ou de maintenance doit être documenté et tracé selon les réglementations en vigueur. Digitaliser la production équivaut à numériser tous les dossiers papiers, c'est le « Paper on glass »<sup>70</sup> :

- *Les documents opérationnels* : instructions de fabrication, de nettoyage et de maintenance, procédure, etc.
- *Les documents d'enregistrements* : dossiers de lot, formulaires, rapports, cahiers de route.
- *Les documents organisationnels* : manuel qualité et procédure d'organisation générale.

Il s'agit aussi de digitaliser les process correspondant à la réception des commandes, la fabrication, la livraison, la gestion des stocks, et la centralisation des données. Pour les entreprises en retard, il s'agit donc de mettre à jour les réseaux industriels et les automates obsolètes (usine en silos numériques).<sup>66</sup>

Pour passer à une usine connectée (niveau 3), un tableau de bord automatisé de révision et d'élimination des lots prend en charge la publication des lots, en tirant parti de l'intégration entre l'ERP, le MES et l'automatisation. L'enregistrement automatique des échantillons et la récupération des résultats font partie de l'exécution de la recette. Un logiciel appelé « *Data Historian* » collecte et analyse les données sur tous les sites. L'ERP et le MES permettent de numériser les processus de fabrication correspondant à :

- La configuration automatique de la ligne basée sur des paramètres de recettes prédéfinis.

- Les instructions de travail prédéfinies pour la gestion des exceptions.
- La gestion des alarmes critiques.

Enfin, l'intégration des systèmes entre MES et ERP et le **Product Life Management** (PLM) est indispensable au bon fonctionnement de l'usine connectée.

L'usine prédictive (4.0) est basée sur une automatisation de l'ensemble de la chaîne de production. Les actions des opérateurs sont limitées à la prise de note d'évènements exceptionnels, car l'ensemble des évènements dus à l'environnement sont pris en charge par les machines. Les informations provenant des instruments analytiques sont intégrées, traitées et utilisées par le MES pour mettre à jour les paramètres de processus. C'est à ce stade que sont utilisées la robotique collaborative et les nouvelles technologies, ce qui permet au matériel d'être adapté, suivant les normes mentionnées précédemment, pour être indépendant du fournisseur.

## 2. La qualité

La qualité en entreprise est aujourd'hui régie par le Système de Management de la Qualité (SMQ). La norme 9000/9001 définit le SMQ comme un ensemble de politique et de processus mise en place pour aider une entreprise à satisfaire aux exigences des parties prenantes.<sup>71</sup> il est basé sur 7 principes fondamentaux qui sont :

<b>1</b>	<b>L'orientation client</b>	<ul style="list-style-type: none"> <li>- Les exigences du client ainsi que les exigences légales et réglementaires applicables sont déterminées, comprises et satisfaites en permanence.</li> <li>- La priorité d'accroissement de la satisfaction du client est préservée.</li> </ul>
<b>2</b>	<b>Le leadership</b>	Les dirigeants établissent la finalité et les orientations et créent des conditions dans lesquelles le personnel est impliqué pour atteindre les objectifs qualité de l'organisme.
<b>3</b>	<b>Implication du personnel</b>	Il est important de respecter et d'impliquer l'ensemble du personnel à tous les niveaux. La reconnaissance, l'habilitation et l'amélioration des compétences facilitent l'implication du personnel dans l'atteinte des objectifs qualité de l'organisme.
<b>4</b>	<b>L'approche processus</b>	Les activités sont comprises et gérées comme des processus corrélés fonctionnant comme un système cohérent.
<b>5</b>	<b>L'amélioration continue</b>	Le succès d'un organisme repose sur une volonté constante d'amélioration pour conserver ses performances et sa compétitivité tout en stabilisant la qualité.
<b>6</b>	<b>La prise de décision fondée sur des preuves</b>	Les décisions fondées sur l'analyse et l'évaluation de données et d'informations sont davantage susceptibles de produire les résultats escomptés.
<b>7</b>	<b>Le management des relations avec les parties prenantes</b>	Les parties intéressées pertinentes ont une influence sur les performances d'un organisme. Pour obtenir des performances durables, les organismes gèrent leurs relations avec les parties intéressées pertinentes, telles que les prestataires.

**Tableau 2 : Les 7 principes fondamentaux du SMQ (ISO 9000)**

Le SMQ est maintenant bien établi dans l'industrie pharmaceutique, car requis pour respecter les Bonnes Pratiques de Fabrication (BPF). Le SMQ est indispensable au passage à l'industrie 4.0.<sup>72</sup>

Comme pour la fabrication, il est indispensable ici aussi de rassembler l'ensemble des données et de numériser tout ce qui est encore sous format papier pour passer à une usine en silos

numérique (niveau 2). Une fois l'ensemble des documents numérisés, il reste indispensable de gérer et de tracer les dossiers à l'aide de logiciels de gestion électronique des documents. Ensuite, un logiciel d'analyse des données permet de standardiser et d'optimiser l'analyse des données.<sup>66</sup>

L'étape suivante vers la digitalisation fait partie du SMQ. Il s'agit de conserver l'intégrité des données et des systèmes pour l'ensemble du site. (Data integrity et system integration).

L'intégrité des données vise à garantir la fiabilité et la précision des données émanant des différentes opérations ayant conduit à l'obtention du produit fini.<sup>73</sup> Rappelons que les Bonnes Pratiques de Fabrication (BPF), exigent que les données soient gérées et qu'un système soit mis en place afin de limiter les risques liés à la perte de l'intégrité des données.

On décrit souvent l'intégrité d'une donnée selon l'acronyme ALCOA+<sup>74</sup>:

<b>Attribuable</b>	Chaque donnée doit être attribuable à une personne. Il est primordial de s'assurer de l'authenticité de la source des données.	
<b>Lisible</b>	Comme expliqué dans les paragraphes précédents, une donnée peut être encodée et incompréhensible sans traduction informatique. Ce point insiste sur la lisibilité des données par l'humain.	
<b>Contemporaine</b>	Il s'agit de programme permettant d'horodater et de classer dans l'ordre chronologique l'enregistrement des données de manière non falsifiable. (Synchronisation à un serveur de temps, contrôle humain).	
<b>Originale</b>	Certains outils permettent de vérifier qu'une donnée est bien originale et n'a pas été modifier ou copier.	
<b>Accurate (Fiable)</b>	Vérifier la qualité d'une donnée. Une donnée ne doit pas être modifiée par un tiers. Elle doit décrire l'exactitude des évènements. C'est ce point qui cherche à traduire la transparence des données et de leur modification.	
<b>+</b>	<b>Complète</b>	Une donnée modifiée peut comporter un ensemble de versions. Le caractère complet exige la sauvegarde de l'ensemble des fichiers ou version d'un même fichier modifié.
	<b>Reproductible</b>	Le caractère reproductible implique qu'une donnée générée par un processus de création (process de fabrication) devrait générer la même à tout moment. Ce point est influencé par la qualification des équipements et la validation de systèmes informatisés.
	<b>Permanente</b>	Ce point se focalise sur l'archivage des données. Une donnée doit être conservée pour une durée minimale de six ans selon les BPF. Cela est facilité avec le Cloud qui permet d'avoir une grande base de stockage.
	<b>Accessible</b>	Une donnée doit être accessible à tout moment.

**Tableau 3 : l'intégrité des données selon l'ALCOA+ <sup>74</sup>**

C'est donc à l'aide de ces outils comme l'ALCOA+ que l'on va vérifier l'intégrité des données et assurer la traçabilité et la protection des informations. Il est important de rappeler que la traçabilité et l'intégrité des données sont des obligations réglementaires pour l'industrie pharmaceutique.

L'intégration des systèmes est définie par *Seddon et al* comme étant l'unification des processus, système et/ou des données depuis de multiples systèmes informatisés, et ce pas nécessairement dans une seule organisation.<sup>75</sup> Il s'agit donc de centraliser l'information et la standardiser afin d'éviter les problèmes de compatibilité. La SI permet alors de gagner du temps entre le passage des données d'un système à l'autre. Par exemple, si on implémente un système qualité comme le *Laboratory Information Management System (LIMS)*, l'intégration des systèmes permet une communication entre ce système et l'ERP ou encore avec le MES. C'est l'usine connectée (niveau 3).

Le LIMS est un système de gestion des données de laboratoire. Il est utilisé par le service de contrôle qualité. Il est conçu pour augmenter la productivité des laboratoires. Il assure un suivi des échantillons et de l'ensemble des informations cruciales qui leur sont associées comme le flux de travail, les résultats d'assurance et contrôle qualité, etc.<sup>76</sup>

Enfin, la qualité du futur (usine prédictive) correspondrait à une qualité intégrée au processus de fabrication, et ce sur tout le cycle de vie du produit, depuis les matières premières jusqu'à la libération des lots pour l'industrie pharmaceutique. Le concept existe déjà avec les PAT (*Process Analytical Technology*). En 2004, la FDA lance l'initiative PAT. Ce guide le définit comme un système permettant de concevoir, d'analyser, et de contrôler la production grâce à des mesures opportunes, des spécifications critiques (**CTQ**) et des caractéristiques de performance sur les matières premières ou sur le produit en cours de fabrication dans le but d'assurer la qualité du produit final. Le PAT inclut des analyses physiques, chimiques, microbiologiques, mathématiques, et l'analyse de risque de manière intégrée.<sup>75</sup> Il s'agit donc de contrôler, en continu, l'environnement de la fabrication et les paramètres probablement définis dans le design d'espace (approche QbD) et de gérer les erreurs. L'homme restera dans ce concept indispensable à la gestion des erreurs exceptionnelles et le PAT gèrera les variations dues à l'environnement.<sup>66</sup>



### 3. La logistique

La première étape du passage au numérique, comme pour tous les autres services, est un état des lieux et un recueil de l'information des systèmes informatiques ainsi que des systèmes de gestion de stock et de planification. Afin de passer d'un site de gestion par le papier au numérique en logistique, la mise en place de modules d'ERP pour les systèmes basiques est nécessaire. Ces modules permettent la gestion des finances, la gestion des ressources, de l'inventaire, des opérations. <sup>69</sup>

Pour passer à une usine en silos numériques (de niveau 2), le rôle de l'ERP sera étendu pour inclure certains éléments de gestion du matériel et des capacités de commande à fabriquer. De plus, au niveau opérationnel, l'intégration des systèmes sera présente même si elle reste limitée au système de gestion de la chaîne d'approvisionnement. L'échange d'information automatique est quasi inexistant à ce niveau, car les systèmes ne sont pas intégrés et connectés, et se fera essentiellement de manière directe (mail, réunions).<sup>69</sup> La sérialisation des boîtes de médicament est implémentée pour améliorer la traçabilité digitale.<sup>66</sup>

La logistique d'une usine connectée (niveau 3) est caractérisée par un ERP complété par les modules les plus utilisés en industries pharmaceutiques : planning de production, finance et contrôle de gestion, inventaire et mouvements de matières, maintenance et distribution.<sup>66</sup> De même, c'est à cette phase que l'on intègre les systèmes pour faciliter le flux de l'information entre les différents systèmes d'opération comme le MES, le LIMS, l'ERP, etc. La planification de la chaîne d'approvisionnement fonctionnera à un niveau plus complet, avec des informations actualisées sur les progrès des systèmes opérationnels.

Une usine prédictive aura une chaîne d'approvisionnement numérique centrée sur le réseau. Il existe une connectivité profonde à la fois au sein de l'installation, avec les tiers extérieurs comme les fournisseurs. Les systèmes intégrés ERP, de fabrication et de chaîne d'approvisionnement permettent une planification et une réponse plus flexible. Les parties externes sont intégrées au niveau du flux de matériaux, offrant une visibilité accrue des performances opérationnelles, la capacité de visualiser l'impact sur l'approvisionnement et permettant la gestion des stocks par les fournisseurs.

L'utilisation de nouvelles technologies comme les jumeaux numériques permettra de simuler et d'optimiser la chaîne d'approvisionnement en temps réel. Ce système permettra de planifier la production, l'ordonnancement et la modélisation des capacités ainsi que la gestion

des risques en temps réel. On a, à ce niveau, une visibilité en temps réel sur les stocks des fournisseurs et l'intégration des systèmes de qualité des fournisseurs.

#### **4. La Maintenance**

Dans une usine en silos numériques indépendants, les dossiers de maintenance sont stockés dans un système autonome, la planification de la maintenance et les pièces de rechange sont gérées à travers une feuille de calcul (Excel). À ce niveau, il n'y a pas d'intégration du système à l'ERP. Dans une usine connectée, la maintenance est gérée par un système de gestion de la maintenance assisté par ordinateur (GMAO). Un fournisseur externe gère le stockage des pièces détachées directement à travers ce logiciel. La planification de la maintenance et de la fabrication est intégrée à un système unique pour faciliter l'organisation de la production.

Dans une usine prédictive, l'utilisation de nouvelles technologies comme l'internet des objets permet la mise en place de capteurs sur les équipements, et la réception des données en temps réel. Ceci optimise la maintenance prédictive et l'étalonnage des machines, ce qui maximise leur disponibilité et la productivité. La maintenance prédictive permet aussi de minimiser le stockage de pièce de rechange et de diminuer les coûts de maintenance. Grâce à l'IIoT les opérateurs peuvent surveiller à distance les équipements.

## 5. Les autres fonctions supports

Fonction support	Silos numériques	Usine connectée	Usine Prédictive
<b>Conception d'installation</b>	La documentation est stockée électroniquement, mais reste indépendante à la gestion des changements. Le changement est géré par des processus décrits dans des manuels avec une évaluation des risques pour évaluer l'impact du changement.	Les jumeaux numériques permettent une représentation en 3D de l'usine pour faciliter et optimiser l'installation d'équipement et de changement dans l'usine.	La 3D est combinée à la réalité virtuelle permettant une immersion totale pour vérifier les installations.
<b>Formation</b>	Les fichiers de formations sont stockés dans un ordinateur.	Les formations sont adaptées au besoin et aux utilisateurs grâce aux technologies mobiles, à l'utilisation de la réalité augmentée/virtuelle.	La 3D est combinée à la réalité virtuelle pour fournir une formation immersive
<b>Surveillance environnementale</b>	Suivi des échantillons et surveillance de l'environnement enregistré sur une feuille de calcul et sont stockées dans un Logiciel de gestion de l'information de laboratoire (LIMS)	Le suivi des échantillons collectés est possible grâce aux appareils mobile et code barre ou encore la <b>communication en champ proche</b>	L'analyse prédictive (ML) permet à partir de données obtenues automatique grâce à des capteurs de gérer en temps réel l'utilisation des ressources (eau, électricité) et va adapter l'approvisionnement en fonction des plages de sécurité pour le produit. L'intervention humaine reste indispensable pour l'échantillonnage microbiologique et pour les événements exceptionnels.

**Tableau 4** : Descriptions des fonctions supports en fonction du niveau de digitalisation de l'industrie d'après Biophorum<sup>74</sup>

## C. L'application des outils de l'industrie 4.0.

### 1. L'IOT et les Big Data

Ces deux outils présentés plus haut sont les principales technologies permettant de mettre en place un système cyberphysique dans une entreprise. À l'aide des grandes quantités de données récoltées et à la connexion entre les différentes machines et systèmes grâce à l'IOT, on peut créer une usine connectée à travers un réseau commun. C'est dans ce contexte que la digitalisation des données de l'industrie est primordiale pour pouvoir adopter l'IOT. Aujourd'hui l'utilisation de l'IOT en est à ses débuts dans l'industrie pharmaceutique. Ce secteur étant en digitalisation, on voit tout de même apparaître de plus en plus en plus d'équipements connectés afin de digitaliser les données lors des processus de fabrication.<sup>78</sup> L'utilisation de l'IOT génère une très grande quantité de données, les Big Data. Il est donc important de stocker et de gérer ces données grâce au *Cloud computing*.

L'IOT a des applications possibles dans de nombreux secteurs de l'industrie pharmaceutique.

#### a. En production

Elle permet aux équipements connectés de transmettre facilement les données opérationnelles vers d'autres appareils ou humains en temps réel, permettant de répondre à une demande de flexibilité et de gérer l'industrie afin d'améliorer la productivité. Elle permet aussi d'avoir une vision sur l'ensemble du flux de fabrication depuis les premières étapes de fabrication du principe actif jusqu'au conditionnement. Les capteurs sont au cœur de cette technologie fournissant des informations réelles sur le processus et transmettant des décisions intelligentes au réseau central pour garantir la qualité du produit. Ces capteurs peuvent être variés et permettent le suivi de l'environnement de production (températures, quantité d'eau, lumière, etc) mais aussi du suivi de la qualité en temps réel. Donc, dans un premier temps l'IOT permet un contrôle optimisé de la production et une flexibilité, permettant de réagir rapidement au changement puisque l'accès aux données est plus rapide, mais aussi de standardiser les process. C'est à ces fins qu'est utilisé l'IOT avec le MES.

Dans un deuxième temps, les Big Data générées par l'IOT sur les produits en cours de fabrication pourraient aussi être analysées afin d'améliorer le processus de fabrication et la qualité du médicament pour les bienfaits du patient.

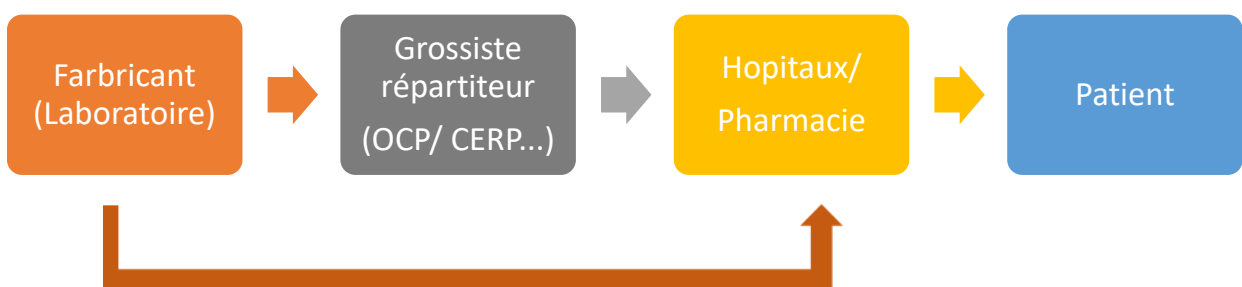
## b. Dans le Magasinage

La gestion du magasinage a pour but d'optimiser le flux de distribution des matières en interne et en externe. D'après une étude de l'entreprise de conseil *McKinsey*, le magasinage représente 95 % des frais de la logistique en industrie pharmaceutique.<sup>78</sup> L'IOT permettrait alors de faciliter le suivi des marchandises, leur identification, leur stockage et l'inventaire. En effet, la technologie « **Radio identification** », base de l'IOT est une étiquette qui permet de recevoir et de mémoriser des données à distance. Grâce à ces étiquettes, le taux d'identification et de suivi des produits est optimal en entrée comme en sortie de magasin. Ces étiquettes contiennent les informations clefs concernant le produit en question et permettent une augmentation de la réactivité et de la prise de décision, grâce par exemple, à une alerte sur la date d'expiration, limitant ainsi les coûts et le gaspillage.

L'analyse des données reçues en temps réel permet de déterminer les problématiques principales du magasin et les zones de blocage afin, toujours dans la même optique, d'être plus flexible.<sup>78</sup>

## c. Dans la chaîne d'approvisionnement et le transport

Rappelons que la chaîne d'approvisionnement du marché du médicament suit le schéma commun suivant :



**Figure 8** : Schéma représentatif de la chaîne d'approvisionnement de l'industrie pharmaceutique

Comme on peut le voir sur ce schéma, la chaîne d'approvisionnement se divise en 4 étapes. Le laboratoire envoie les produits finis aux grossistes/répartiteurs qui vont traiter avec les hôpitaux et les pharmacies. Enfin, ce sont ces derniers qui vont délivrer directement les

traitements aux patients. D'un autre côté, les fabricants peuvent directement transférer les produits aux hôpitaux et aux pharmacies.

Dans ce secteur, L'IOT permet d'assurer la sécurité de la chaîne d'approvisionnement. En effet, il est important de rappeler que la chaîne d'approvisionnement et le transport de médicaments se doit d'être stable et sécurisée durant toutes les étapes de cette chaîne, pour la qualité du médicament et la sécurité du patient. Comme pour le magasinage, l'utilisation de puce RFID, ou de code barre 2D, permet d'implémenter des informations numériquement et uniques pour chacune des boîtes de médicament, c'est la sérialisation. Cette technologie déjà utilisée par de nombreuses entreprises, peut avoir quatre grandes utilités dans ce secteur :

- Le premier est pour le patient. En effet, ces puces contiennent par exemple la date de fabrication, de péremption, la composition et les consignes de sécurité, facilitant l'utilisation de médicament et diminuant le mésusage.<sup>78</sup>

- Dans un deuxième temps, ces technologies permettent l'identification et le suivi en continu des produits limitant les vols et la contrefaçon, deux préoccupations majeures pour les industries pharmaceutiques.<sup>78</sup>

- Dans un troisième temps, des capteurs connectés permettent de suivre en continu l'environnement de stockage comme la température, la pression ou l'humidité assurant la qualité du produit sur tout son parcours.

- Enfin, ces puces permettent de suivre l'inventaire et le stockage des matières et produits finis, ainsi que la date d'envoi et la date d'arrivée des produits afin d'aider à la planification et à coordonner le transfert des matériaux et produits finis.

## **2. Le Cloud computing**

Comme expliqué précédemment, le Cloud permet d'avoir un espace de stockage et une force de calcul importante et adaptable sans investir des ressources monétaires et humaines dans la mise en place d'un réseau et serveur physique. Le Cloud est déjà largement utilisé dans l'industrie ainsi que dans le secteur pharmaceutique, mais est surtout utilisé pour le stockage de données et beaucoup pour leur capacité de calcul.<sup>79</sup> En effet, le Cloud permet d'utiliser des logiciels et des outils à la pointe de la technologie sans installation physique.

## **a. En R&D**

D'après *Preffitt*, en 2009 déjà, des laboratoires comme *Pfizer* ou *Johnson & Johnson* ont utilisé le Cloud computing en R&D permettant d'utiliser les outils novateurs comme la bio-informatique, les statistiques et les essais cliniques au design adaptatif avec des coûts et des délais plus courts et prévisibles.<sup>80</sup>

Le Cloud permet aussi aux Petites ou Moyennes Entreprises (PME) d'utiliser des outils, qui coûtent normalement cher à implémenter, grâce à un accès adaptable de ces outils depuis internet.

Aussi la scalabilité du Cloud permet aux PME comme aux plus grosses firmes d'éviter les pertes de coût dû à l'installation des serveurs physiques qui pourraient ne pas être adaptés à la demande (trop ou pas assez de force de calcul et de stockage).<sup>80</sup>

Le Cloud se définissant par un accès aux mêmes sources indépendamment du lieu d'utilisation, *McCarthy* en 2001, affirme que cette technologie augmenterait la collaboration internationale des chercheurs, permettant à tous les chercheurs concernés de contribuer aux recherches quelle que soit leurs moyens technologiques.<sup>80</sup>

Enfin, durant les étapes de développement, *Shurell* a expliqué que l'utilisation du Cloud facilite la gestion des données et la mise en place des essais cliniques qui deviennent de plus en plus décentralisés et internationaux. Le Cloud permet d'améliorer le stockage et l'accessibilité de l'ensemble des documents réglementaires et le suivi de l'essai, de faciliter le transfert des données entre différents sites, de faciliter la gestion des parties prenantes, et d'accélérer le recrutement des patients.<sup>80</sup>

## **b. En entreprise**

Le Cloud devenu largement utilisé à révolutionné le secteur de la TI, permettant une gestion globale plus efficace des volumes de données de plus en plus important (Big Data). Comme pour la R&D le Cloud permet une utilisation de ressources importante sans investir dans du matériel et des ressources humaines, diminuant ainsi les coûts de mise en place et d'entretien. Le Cloud permettrait ainsi d'améliorer le stockage et l'analyse de données, d'améliorer la gestion du changement, la collaboration entre les services et une flexibilité d'accès aux ressources informatiques.<sup>80</sup>

### 3. Le Machine Learning et l'IA

Le domaine de l'IA étant très vaste, son application sera variée. En effet, cette technologie peut être exploitée depuis la R&D jusqu'à la production pharmaceutique en passant par la logistique, le marketing, et d'autres secteurs.

#### a. En R&D

L'intelligence artificielle a son plus large potentiel dans la recherche fondamentale et la recherche pharmaceutique. La recherche fondamentale étudie les mécanismes pathologiques et les cibles moléculaires permettant une action pharmacologique. La recherche pharmaceutique cherche à déterminer quelle molécule pourrait agir sur les cibles préalablement déterminées.<sup>81</sup>

Dans les deux cas, l'IA a un grand rôle à jouer dans des recherches bibliographiques, sur la maladie, les cibles et les connaissances actuelles, prémices indispensables de la recherche. En effet, des modèles algorithmiques permettent de rechercher, de trier et d'analyser des très grandes quantités de données bibliographiques afin de ressortir des informations primordiales pour le chercheur : les cibles moléculaires les plus pertinentes, les molécules ayant échoué, les expériences réalisées, etc. Ces recherches permettent aux chercheurs de déterminer ce qui a déjà été fait et d'éviter de réitérer les mêmes erreurs.<sup>81</sup>

Par exemple, en recherche fondamentale, *Open Target* est une plateforme qui permet de déterminer pour chaque pathologie les cibles potentielles pour un médicament. Pour ce faire, ils ont utilisé un algorithme de *Machine Learning* qui permet de classer chaque cible aux pathologies et d'établir une relation entre la cible et la maladie. Toujours en recherche fondamentale, Le *Machine Learning* permet la prédiction de structure tridimensionnelle des protéines.<sup>81</sup> *Alpha fold* a créé un modèle (basé sur un réseau de neurones artificiels) utilisant de grands ensembles de données génomiques pour prédire la structure des protéines.<sup>82</sup>

Le criblage à haut débit, la synthèse, la prédiction de caractéristiques d'une molécule sont autant de domaines que l'IA peut aider. En criblage, la capacité de prédiction du *Machine Learning* permet de sélectionner dans une large banque de données les molécules les plus aptes à être des cibles pertinentes en tant que médicament. IKTOS est un exemple de start-up utilisant l'IA pour aider au criblage moléculaire. Ils utilisent des modèles génératifs



reposant sur des réseaux de neurones artificiels. Ce modèle permet de créer virtuellement des molécules optimales conçues par l'IA qui vont correspondre au **profil du produit cible**.<sup>83</sup>

En synthèse, L'IA peut être utilisée pour déterminer la voie de synthèse la plus optimale, ou pour estimer la difficulté de synthèse de la molécule. Cela va donc faciliter le processus de synthèse de la molécule.<sup>81</sup> L'IA permet aussi le recyclage de molécules déjà créées afin de les réutiliser dans des pathologies où elles seront plus actives. Par exemple, en 2020 le *Massachusetts Institute of Technology* a développé grâce à l'IA, un antibiotique : l'halicine, médicament initialement utilisé contre le diabète.<sup>84</sup>

Ainsi l'IA est un grand avantage en R&D puisqu'elle permet d'optimiser la recherche, de gagner du temps, et de diminuer les coûts. C'est un bénéfice majeur lorsque l'on sait que la découverte et la commercialisation d'un nouveau médicament peut prendre jusqu'à 15 ans.

## **b. En production**

En production, l'IA peut être révolutionnaire dans l'optimisation des processus. Comme décrit plus haut, l'optimisation de la production vient de la collecte des données en temps réel qui devient possible grâce à la mise en place d'un système cyberphysique caractérisé en industrie par l'IIOT, le MES, le Cloud et L'IA. En effet, l'IA a un rôle à jouer dans l'analyse des données réceptionnées. L'IA pourrait alors analyser les données permettant de contrôler les procédés et apporterait des prédictions sur les événements en fonction des paramètres. En effet, en stockant l'ensemble de ses données en temps réel sur un **lac de donnée (Data Lake)**, l'IA combinée à cette méthode de stockage rapide et sans transformation des données pourrait les décrypter, les analyser afin d'être utilisés directement par les managers et data analysts.<sup>81</sup>

Aujourd'hui, l'IA est un enjeu stratégique pour l'ensemble des pays, mais n'est toujours pas une technologie mature.<sup>85</sup> Dans le futur, une IA plus « forte », permettrait d'automatiser certaines tâches de production répétitive et pénible pour l'humain. Une IA plus forte aura des capacités de prise de décision. La chaîne de production serait alors entièrement automatisée. L'IA planifierait la production à partir du bon de commande, générerait les ressources nécessaires, surveillerait et contrôlerait la production, le conditionnement et le stockage. Une IA forte, définie comme ayant au moins une intelligence supérieure à l'homme<sup>85</sup>, aura une meilleure utilisation du matériel et des matières, une production plus rapide et efficace, et un

respect constant des attributs qualité critiques comme l'explique Hoerdt Aurélien dans sa thèse.<sup>81</sup>

### **c. En qualité**

Le PAT défini précédemment permet le contrôle de la qualité des produits en cours des étapes de fabrication. Ceci est permis par des IA, qui aidées par l'intermédiaire de données reçues par l'IOT permettent de contrôler différemment les produits. Par exemple, la mise en place de détecteurs visuels avec des caméras pourrait contrôler visuellement les produits et détecter les non-conformités et les défauts, parfois microscopiques et invisibles à l'œil nu. Le PAT permet aussi de contrôler les paramètres physiques, chimiques, ou encore microbiologiques. Aussi, une IA basée sur du *Machine Learning* pourrait classer l'ensemble de ces non-conformités (en fonction du type, de la gravité etc). Les algorithmes pourront aussi analyser les anomalies afin de déterminer leurs causes et améliorer les procédés en fonction.<sup>81</sup>

L'IA peut aussi révolutionner l'analyse de risque, pratique indispensable pour la sécurité du patient. En effet, l'utilisation d'algorithmes permettrait de détecter, classer et analyser, ainsi qu'évaluer les risques de manière plus sûre, se rapprochant le plus possible du risque « zéro ». Pour ce faire, le modèle utiliserait les données du système de production pour cartographier la fabrication et détecter l'ensemble des risques sur l'ensemble de la chaîne pour ensuite proposer des contres mesures.<sup>81</sup>

### **d. En logistique**

La logistique peut aussi être dirigée par l'IA grâce à des modèles de *Machine Learning* qui vont analyser les données provenant des ERP, des tendances du marché, des données de gestion de relation client, et déterminer les problématiques ainsi que recommander des solutions.<sup>81</sup> Par exemple, *Tracelink*, utilisé par l'entreprise *Merck* est un logiciel IA développé par *Aera Technology* qui permet d'augmenter les commandes fournisseurs, réduire la fabrication, réacheminer les produits en fonction des données analysées.<sup>81</sup> Cet outil permet d'automatiser les tâches répétitives permettant aux ressources humaines de se concentrer sur le plus important.

De même, le transport des produits d'une machine à l'autre peut aussi être dirigé par une intelligence artificielle avec l'aide de robots, limitant ainsi l'utilisation de l'Homme pour ses tâches pénibles pouvant avoir des répercussions sur la santé. Pour ce faire, une IA optimise la coordination du transport des produits et des matières premières pour la logistique interne (entreposage, et trajet de transport dans les locaux industriels) et externe (proposer des itinéraires de livraison et les modifier en temps réel).<sup>81</sup>

### **e. En marketing**

L'IA est déjà largement utilisée dans le marketing. La publicité ciblée est un bon exemple. La publicité ciblée se base sur la réception de Big Data à travers les cookies des utilisateurs internet. Ces données vont être analysées par un modèle algorithmique pour déterminer les produits les plus enclins à être achetés par le consommateur afin de lui attribuer des publicités ciblées.<sup>81</sup> Certaines entreprises ont conçu des modèles qui permettent d'aider les entreprises dans leur stratégie marketing. *BAYESIA* conçoit par exemple des modèles à base de réseaux bayésien pour l'aide à la décision.<sup>86</sup> En marketing, l'application de ces modèles permet l'élaboration des profils de la clientèle, l'analyse de satisfaction, la prévision des parts de marché ou encore la détection anticipée de **l'attrition**.<sup>86</sup>

### **f. En Santé publique et pharmacovigilance**

La pharmacovigilance, stade IV du développement d'un médicament, permet de recueillir et d'évaluer les effets indésirables dans le but de la prévention et de la réduction des risques.<sup>87</sup> La quantité de données recueillies est de plus en plus importante provenant de différentes sources comme les déclarations de patients, les dossiers médicaux, les études de cohortes, etc. L'IA aurait un rôle majeur en pharmacovigilance puisqu'elle pourrait recueillir et classer les données de l'ensemble des bases de données relatives aux effets indésirables. Par exemple, avec la crise de la COVID-19, l'Agence National de Santé du Médicament et des produits de santé (ANSM) a lancé une application dotée d'un module d'IA permettant la « catégorisation et un précodage des effets indésirables, le tout facilitant et fluidifiant le travail d'expertise mené par les centres régionaux de pharmacovigilance (CRPV) ».<sup>88</sup>

Aussi, en santé publique, l'IA pourrait prévenir la diffusion d'épidémie. Toujours en recueillant les big data publiques, une IA pourrait suivre l'incidence des maladies. Cet algorithme pourrait alors alerter les autorités de santé en cas d'augmentation majeure d'une incidence afin de faciliter la réponse des organismes et de gagner en flexibilité. Par exemple, *HealthMap*, est une application qui cartographie un grand nombre de maladies dans le monde à partir des données recueillies auprès des organismes de santé locaux, d'internet, ou des rapports officiels.<sup>81</sup>

## **4. Le MES**

Le MES est un outil informatique indispensable dans la production 4.0. Un système MES a pour but d'améliorer la production en délivrant en temps réel des informations relatives aux performances, les temps d'arrêt des machines, mais aussi sur la qualité des produits. Il fournit des informations sur les processus, la traçabilité, l'historique des événements. Une étude réalisée par IDG en 2020 sur 108 entreprises a cherché à déterminer quelle technologie de l'industrie 4.0 à mettre en place est prioritaire. Selon cette étude, 25 % des répondants ont prévu d'implémenter un système MES dans un avenir proche. De même, Les répondants ont établi que la technologie MES est classée première technologie à mettre en œuvre pour l'industrie 4.0. Un des raisons à ce résultat est la promesse de générer des économies grâce à la surveillance des temps de travail des équipements et du personnel et en diminuant les temps d'arrêts et de panne.<sup>89</sup>

En mettant en place un MES au sein d'une usine, on permet de connecter les machines avec le système de gestion qu'est l'ERP. Comme expliqué précédemment, le MES est une interface entre les deux niveaux hiérarchiques présentés dans la pyramide de la productique.

Dans l'idéal, un système MES facilite la gestion en temps réel de la production, de la qualité et des ressources humaines.<sup>44</sup>

### **a. Intérêt du MES en Production**

L'acquisition de données en temps réel se fait par l'intermédiaire des capteurs qui sont envoyés au MES. À travers différents modules de suivi de production, comme le « Machine Data collection » (MDC), le « Production Data Acquisition » (PDA), le « Tools and Ressources

Management » (TRM), et autres modules, le MES gère les ressources utilisées, le personnel assigné, la surveillance des machines et du process. Il est capable de repérer un dysfonctionnement d'une machine ou du process et d'établir les causes lors des temps d'arrêts, de gérer l'organisation et la logistique des matières premières (circulation des matières, stockage provisoire). L'analyse des données facilite la maintenance préventive qui se fera donc en amont de la défaillance.<sup>44</sup>

### **b. Intérêt du MES en Qualité**

La maîtrise statistique des procédés peut directement être intégrée au MES. Ce faisant, les outils statistiques (carte de contrôles, droite d'Henry, khi deux) seront utilisés sur les données obtenues en temps réel et permettront de suivre les tendances de production plus rapidement, de comparer les résultats en temps réel aux spécifications et de créer une alerte en cas de dérégulation.<sup>44</sup>

Un module de gestion des non-conformités permet de tracer les produits concernés afin de déterminer les conditions de fabrication, le matériel utilisé et les aspects techniques de fabrication.<sup>44</sup>

Un autre module permet de gérer les entrées et les sorties des matières avec l'enregistrement spécifique des marchandises livrées ou des marchandises expédiées, la vérification de la numérotation des lots et une fonction d'alarme en ligne si certaines valeurs dépassent les spécifications.<sup>44</sup>

Un dernier module, le process data processing (PDP) permet de gérer les données relatives à l'environnement de fabrication. Il s'agit ici d'assurer la qualité du produit à travers les conditions de fabrication. Par exemple, ce module permet de récolter les températures et les pressions de fabrication d'un produit et de les vérifier par rapport à des limites de tolérances.<sup>44</sup>

### **c. Les Ressources humaines**

Le MES peut cartographier la gestion du personnel lié à la production afin d'optimiser l'affectation des opérateurs, de gérer les pointages d'arrivée et de sortie, les absences, les performances du personnel, et les salaires.<sup>44</sup>

## 5. Les jumeaux numériques

Recréer un objet réel et physique virtuellement aboutit à la création d'un jumeau numérique. Son utilisation est vaste. Durant la crise du Coronavirus 19, l'utilisation de jumeau numérique de patient virtuel a permis de simuler les effets de vaccins avant l'utilisation sur l'homme. De même, *Exact cure*, une start-up niçoise a créé un jumeau numérique de l'hydroxychloroquine et a testé son efficacité contre le coronavirus sur des humains virtuels.<sup>90</sup>

En R&D tout comme en production, le jumeau numérique permet de créer virtuellement un processus et tester la fabrication d'un produit en modifiant les paramètres. Il permet par exemple<sup>91</sup> :

- En R&D, de définir les paramètres optimaux, pour une performance de fabrication maximale.
- En production, de surveiller en continu des variables pour piloter en temps réel la fabrication, en implémentant en temps réel les données depuis le monde physique vers le monde virtuel. On peut alors simuler la modification des variables afin de se rapprocher au mieux des performances maximales.
- De surveiller l'état d'avancement de la production et simuler des décisions concernant la planification ou l'affectation du personnel à courts termes.
- Il permet aussi de modéliser à moyen et long terme, le comportement d'un système de fabrication et de modéliser différents scénarios comme le changement du taux de production ou de matière première.
- D'évaluer en amont l'implémentation de nouveaux équipements afin de déterminer si ces derniers seront bien intégrés au système et d'évaluer leur interférence avec d'autres machines ou opérateurs.
- Toujours dans le cas de la mise en place d'un nouvel équipement, un jumeau numérique permet d'anticiper les modes de dégradation de ladite machine, les possibilités de panne et ainsi anticiper la maintenance prédictive et curative.
- Il permet de faciliter la formation des opérateurs à distance ou lorsque la ligne de production n'est pas encore physiquement présente.

Globalement, l'utilisation de jumeaux numériques consent à la transition vers des traitements médicaux préventifs et personnalisés en modélisant la réalité avec des techniques d'analyse avancées afin que les problèmes puissent être prédits.<sup>91</sup> Cette technologie fournit un environnement pour tester l'impact des changements sur les performances et la qualité de fabrication d'un produit et d'un système. Enfin, le jumeau numérique accélère la recherche et le développement du médicament et réduit les coûts.

*Sanofi* a démontré l'efficacité de l'utilisation de jumeau numérique avec son entreprise à Framingham aux Etats Unis. Dans ce site, les jumeaux numériques sont utilisés au quotidien par le responsable du site qui peut à distance effectuer des ajustements en temps réel pour simuler les changements de procédé. De même, comme le déclare Philippe Luscan, vice-président et directeur général des affaires industrielles, ils utilisent aussi les jumeaux numériques pour former leurs nouveaux collaborateurs.<sup>92</sup>

## 6. La réalité augmentée

La réalité augmentée a de nombreux avantages parmi lesquels on peut compter l'amélioration de la formation des travailleurs, le contrôle de l'usine de production, le diagnostic des erreurs, l'amélioration de la sécurité en production, la conception et la reconfiguration du produit, et la facilitation de la fabrication, en fournissant des informations essentielles au niveau opérationnel. Ainsi, elle permet aux entreprises de réduire les coûts de conception et de production, maintenir la qualité du produit et réduire le temps nécessaire pour passer du produit concept à la fabrication.<sup>93</sup>

*Takeda Pharmaceuticals*, industrie pharmaceutique japonaise, utilise aujourd'hui la réalité augmentée afin de guider les opérateurs dans leurs tâches, limitant ainsi l'utilisation des procédures papier et les erreurs.<sup>94</sup>

## 7. L'impression 3D

La fabrication additive est une nouvelle solution comme procédé de fabrication des médicaments. L'impression 3D offre la possibilité de fabriquer des formes galéniques orales solides, y compris complexes comme les produits à libération modifiée.<sup>93</sup> Dans le domaine de la santé, l'imprimante 3D a d'abord été utilisée dans la **planification chirurgicale** et la

production d'implants.<sup>55</sup> En 2015, on voit apparaître la première commercialisation d'un médicament produit en 3D, le *Spiritam* (Lévétiracetam), développé par *Aprecia Pharmaceuticals*. La fabrication additive s'inscrit dans la volonté de plus en plus grande de satisfaire le patient, et donc dans la volonté d'entrée dans la médecine personnalisée. La médecine personnalisée se définit comme étant l'adaptation du traitement aux caractéristiques génétique du patient et de sa maladie.<sup>95</sup> En effet, l'impression 3D peut s'adapter au patient grâce à la caractéristique du procédé de fabrication. En fabrication additive chaque unité est fabriquée individuellement. Ainsi, cette technologie permettrait de s'adapter aux clients (pour les sous-traitants) et aux patients en modifiant le dosage, la forme de libération, ou la géométrie de la forme galénique d'une unité à l'autre.<sup>55</sup>

En R&D, lors des essais préclinique et clinique, l'imprimante 3D peut être utilisée pour créer des tissus pour les tests in vitro ou encore pour la fabrication de petit lot moins coûteux et directement sur site pour les formes placebos.<sup>55</sup>

L'imprimante 3D offre aussi aux industries pharmaceutiques la possibilité d'une production décentralisée. La production des médicaments pourra alors se faire à l'hôpital, en pharmacie ou directement chez le patient en fonction du patient.<sup>55</sup> La décentralisation permettrait alors une diminution des coûts de transport et de stockage et de fabrication pour les industriels.<sup>93</sup> Le modèle économique se baserait alors sur la création du médicament numériquement qui serait envoyé aux différents centres de fabrication en fonction du besoin.

Enfin, l'impression 3D étant précise lors du dépôt de la fabrication, elle permettrait une diminution du coût lié aux quantités de principes actifs et d'excipients achetés et diminuerait le gaspillage.<sup>55</sup>

## 8. Les robots collaboratifs

Le marché des **cobots** est en expansion grâce à une diminution des prix sur le marché de l'industrie robotique.<sup>96</sup> L'application des robots collaboratifs est vaste. On peut les utiliser pour automatiser les processus, faciliter la logistique, optimiser le contrôle de la qualité des produits, ou encore pour le nettoyage.



### **a. En logistique**

Les cobots sont très utilisés dans les entrepôts pour le prélèvement et le placement d'objet. Ces robots peuvent entreposer et prélever les matières premières et produits finis dans le magasin de manière autonome. Les robots logistiques collaboratifs comme les robots mobiles autonomes permettent d'augmenter la productivité, réduit les mouvements inutiles, diminuer les besoins en ressources humaines, et améliorer la sécurité d'une installation.<sup>96</sup>

### **b. En production**

Les cobots peuvent s'occuper intégralement les tâches répétitives. Le chargement de matière dans une machine et sa sortie à la fin de la transformation peut être dirigé par des cobots qui élimineront alors les tâches répétitives, limiteront les risques de blessures pour l'Homme, et minimiseront les erreurs humaines dans l'exécution des tâches de routine. Des tâches comme la manipulation stérile, l'assemblage de petites pièces pour la fabrication de dispositifs médicaux, le tri, le mélange, la palettisation pourront être complétés par des cobots. Cela offre de nombreux avantages parmi lesquels on peut compter un déploiement rapide, une programmation faible, un besoin en espace réduit, et des coûts d'acquisition de plus en plus faible. Globalement, l'utilisation de cobots est avantageuse par le fait que les robots collaboratifs ont la capacité d'effectuer des tâches répétitives avec précision à un rythme plus rapide que les humains.<sup>96</sup>

### **c. En Qualité**

Le contrôle de la qualité peut être automatisé dans certains cas. Par exemple, les cobots peuvent être utilisés pour l'inspection visuelle sur les produits et les assemblages à différentes étapes du processus de fabrication. Ces robots sont équipés de vision 2D ou 3D capable d'inspecter une même unité avec justesse et précision. Ils peuvent aussi être utilisés en laboratoire. À l'hôpital universitaire de Copenhague, des cobots créés par *Universal Robot* sont utilisés en laboratoire pour trier (à l'aide d'un système de caméras) les échantillons sanguins en fonction de 4 couleurs possibles. Un deuxième robot place ensuite les échantillons dans une machine pour qu'ils soient centrifugés et analysés.<sup>97</sup> De même, les cobots facilitent

la traçabilité, ils génèrent des enregistrements fiables et continus de toutes les opérations afin de certifier la validité et la qualité des produits pharmaceutiques.<sup>98</sup>

## **9. La Blockchain**

La blockchain est une solution à la protection et à la traçabilité des données. Comme expliqué précédemment, elle permet de stocker et transférer des données et informations de manière transparente et sécurisée. Chaque information intégrée dans la blockchain est validée numériquement.

C'est surtout dans le domaine de la supplychain (Chaîne d'approvisionnement) que l'utilité de cette technologie s'est fait sentir. Comme expliqué par Manuela M. Schöner, d'après l'OMS, la vente de médicament de contrefaçon dans le monde représentait déjà 75 milliards de dollars en 2010, et a vu une augmentation de 90% en 5 ans.<sup>99</sup>

Cette technologie, applicable à l'ensemble de la chaîne d'approvisionnement, permettrait de sécuriser les informations, et la transparence permettant aux consommateurs de pister leurs médicaments et vérifier leur validité. En effet, les informations de fabrication et de transport seraient intégrées en temps réel à la Blockchain. Les clients intermédiaires et les patients, par un simple scan d'un code barre, pourrait alors tracer l'historique du médicament.<sup>99</sup> C'est dans ce contexte que Manuela M. Schöner et son équipe au sein de la *Frankfurt School Blockchain Center*, ont créé un prototype appelé *Life Crypter*. Ce prototype est une application permettant de tracer les médicaments.

## **IV- Les Défis et les impacts de l'industrie 4.0**

Comme nous avons vu, les outils de l'industrie 4.0 sont totalement applicables à l'industrie pharmaceutique. Cependant, ce secteur a du retard sur sa transition numérique. Il y a donc plusieurs défis à réussir pour compléter la transition.

## A. Les défis

### 1. Les défis réglementaires

L'industrie pharmaceutique est restreinte par une réglementation beaucoup plus stricte que pour les autres secteurs industriels. En effet, l'ensemble des usines pharmaceutiques se doivent de répondre aux autorités réglementaires nationales, européennes, ou encore américaines pour certains produits afin de fabriquer un médicament de qualité, inoffensif et efficace aux patients. Aussi, La qualité des médicaments étant très stricte, la totalité de ce qui englobe la fabrication des produits pharmaceutiques est guidé par des recommandations telles que les normes ISO ou ISA qui devront, souvent, être obligatoirement appliquées sous l'autorité des organismes réglementaires.

Les nouvelles technologies sont encore très peu déployées, coûteuses à développer, et les instances réglementaires n'ont pas, dans certains cas, émis de guides, de recommandations, ou de réglementations. Ces causes incitent les industries à privilégier une approche « premier à être second »<sup>3</sup>. Cette approche cherche à suivre l'innovation sans en être le pionnier pour apprendre des erreurs des autres et éviter les flous réglementaires. En effet, comme l'explique N. Sarah Arden et al, l'absence de précédent réglementaire peut conduire l'industrie à maintenir les processus conventionnels même lorsque de nouveaux processus peuvent réduire le fardeau réglementaire global et améliorer la qualité à long terme<sup>3</sup>. De plus, les juridictions réglementaires étant indépendantes (ANSM, FDA, EMA), il faut dans certain cas créer un dossier de demande de changement pour chaque autorité. A noter que pour une procédure d'autorisation de mise sur le marché centralisé permettant de commercialiser un médicament dans tous les états membres de l'Union Européen, le changement est plus facile<sup>3</sup>.

Les réglementations strictes pour les industries pharmaceutiques touchent surtout à la gestion de la qualité. Or, certaines autorités comme la FDA ou des guides et normes internationales comme l'ICH Q8 <sup>66</sup>, se concentrent de plus en plus sur une réglementation basée sur les performances avec par exemple l'adoption du Quality By Design (QbD).

## a. Gestion de la qualité et le Quality by Design

L'approche de la qualité par la conception, est défini dans l'ICH Q8 comme étant « Une approche systématique du développement qui commence par des objectifs prédéfinis et met l'accent sur la compréhension et le contrôle des produits et des processus, sur la base d'une science solide et d'une gestion des risques qualité. »<sup>100</sup>

Ainsi, ces mêmes recommandations définissent plusieurs termes sur lesquels se base cette approche.

Tout d'abord, un produit est caractérisé par son profil qualité cible (en quelque sorte la projection du produit idéal que l'on souhaite développer), en anglais le Quality Target Product Profile (QTPP). Le QTPP, détaillera donc toutes les caractéristiques telles que son utilisation, sa forme posologique, son mode d'administration, le dosage, les conditionnements, le mode de libération du principe actif et les attributs affectant les caractéristiques pharmacocinétiques, les critères de qualité du produit (pureté, stérilité, stabilité, libération).<sup>101</sup>

Un produit est aussi caractérisé par ses attributs de qualité Critique ou Critical Quality Attributes (CQA) en anglais. Ces CQAs sont les plages, ou les limites des paramètres chimiques, physiques, biologiques, et microbiologiques à respecter pour garantir la qualité du produit.

Le Qbd a pour but d'améliorer la conception de la qualité pharmaceutique grâce à <sup>100</sup> :

- Une évaluation, une compréhension et un raffinement systématiques du processus de formulation et de fabrication :
  - o Identifier, par exemple, par le biais de connaissances préalables, d'expérimentations comme les plans d'expériences et d'évaluations des risques, les attributs matériels et les paramètres de processus qui peuvent avoir un effet sur les Attributs de qualité critiques des produits.
  - o Déterminer les relations fonctionnelles qui lient les attributs matériels et traiter les paramètres des CQA des produits.
- Utiliser la compréhension approfondie des produits et des processus en combinaison avec la gestion des risques qualité pour établir une stratégie de

contrôle appropriée qui peut, par exemple, inclure une proposition d'espace(s) de conception (Design Space) et/ou des tests de version en temps réel.

Le Design Space correspond, pour un produit donné, à la combinaison et l'interaction multidimensionnelle des variables d'entrée (par exemple, les caractéristiques des matières premières) et les paramètres de processus qui ont été définis/validés pour fournir une maîtrise de la qualité. C'est l'espace à l'intervalle duquel chaque paramètre peut être modifié tout en préservant la qualité, la sécurité et l'efficacité du produit.<sup>101</sup>

Ainsi, comme l'explique Janet Woodcock, commissaire de la FDA : « En améliorant la compréhension scientifique des produits et des procédés, l'approche QbD rend possible la conformité fondée sur le risque. Son but n'est pas d'éliminer la variabilité dans les processus, mais de développer un processus qui peut supporter la gamme de variabilité acceptable pour le maintien de la qualité du produit. À partir d'un profil cible du produit (PPT) sur la base des attributs qualité critiques (AQC), il est alors possible d'utiliser des méthodes analytiques appropriées et des outils pour comprendre la conception spatiale (Design Space).<sup>101</sup>

L'approche QbD, moins contraignante, est complémentaire à l'industrie 4.0 qui révolutionne les processus, le traitement de l'information et se base sur une amélioration des performances. Une réglementation basée sur les performances se concentre sur les résultats attendus ou recherchés plutôt que sur les procédés. Ainsi, les actions réglementaires se concentrent sur l'identification de mesures de performance qui garantissent une marge de sécurité et incitent à améliorer la sécurité sans intervention réglementaire formelle. De ce fait, les processus de fabrication sont moins dépendants du contrôle par les autorités de santé tant que la qualité du produit est assurée. Le système de production peut donc évoluer de manière continue. Dans ce contexte, le défi des régulateurs sera alors de surveiller les risques associés au changement continu des paramètres et procédés. Pour ce faire, il peut être nécessaire d'associer des technologies de fabrication avancées et des systèmes de réglementation et de contrôle avancés pour permettre une assurance en temps réel des performances du système de production.<sup>3</sup>

## **b. La validation des procédés**

Les réglementations de la santé sont strictes pour assurer la qualité et la sécurité du produit. Dans ce sens, Les autorités réglementaires demandent aux industriels de valider les procédés de fabrications et l'installation de nouveaux équipements au sein d'une usine pharmaceutique ou pour tout changement. Ainsi, il n'est pas rare pour les dirigeants industriels d'hésiter à transiter vers l'industrie 4.0 tant la validation des technologies insérées dans les processus de production peut coûter cher.<sup>102</sup> Selon l'étude de Felipe Silva et son équipe, certains industriels estiment que la validation de l'implémentation d'une nouvelle technologie peut coûter jusqu'à deux à trois fois plus cher que la technologie elle-même.

Aussi, l'application des nouvelles technologies et principalement l'IA pourrait accélérer la validation et la qualification des processus de fabrication. La validation des processus se fait en trois étapes<sup>103</sup> :

- La conception des processus de fabrication.
- Mise en œuvre des processus à l'échelle commerciale
- Suivi tout au long du cycle de vie du produit

En effet, depuis la conception des processus jusqu'à l'amélioration continue de ce dernier, l'IA, les jumeaux numériques, et les nouvelles technologies peuvent accélérer ou encore permettre de réaliser simultanément les étapes de validation des processus de fabrication. Ainsi, on passerait d'un modèle sporadique et discontinu de la validation d'un processus à un modèle continu grâce à l'IA et à une meilleure capture et transmission des données augmentant les performances et la robustesse des processus. Cependant, comme pour la qualité, les autorités réglementaires doivent s'adapter à un processus continu de validation avec peu de surveillance active compte tenu de la très grande quantité de données générées par les nouvelles technologies.<sup>3</sup>

Aussi, les différentes industries pharmaceutiques ne sont pas au même stade de transition vers l'industrie 4.0. En fonction de la taille des entreprises, du modèle de l'entreprise, des moyens ou de la structure, on retrouve des entreprises dans les différents stades industriels : 2.0, 3.0 et 4.0. Aussi, au sein même d'une entreprise qui transite vers un nouveau paradigme (3.0 ou 4.0) certaines usines seront dotées de nouvelles technologies là où d'autres se tiendront à leur ancien modèle. Là aussi le challenge des autorités réglementaires sera

d'arriver à réguler l'ensemble des fabricants de médicament en prenant en compte les différents stades d'évolution de ces derniers.

## **2. Les défis techniques**

### **a. La gestion des données**

Le déploiement d'algorithmes repose fortement sur l'accès aux données et la présence d'environnements de test. Or, dans l'industrie pharmaceutique, l'accès aux données est très restreint, pour des raisons de protection de la propriété intellectuelle, pour des raisons techniques ou pour des raisons de sécurité. Les environnements virtuels de sites de fabrications (jumeau numérique) n'offrent pas de données suffisantes pour le développement d'algorithmes. Les algorithmes sont alors fréquemment entraînés sur des données réelles exportées.<sup>47</sup> Or, en production, les environnements opérationnels et les logiciels de fabrication ne sont généralement pas développés en tenant compte de la science des données (Data Science) et des algorithmes. De plus, appliquer tout changement dans un environnement déjà validé est difficile d'un point de vue réglementaire. Enfin, développer un algorithme qui fonctionne parfaitement sur un système de production demande beaucoup de temps et de données.<sup>47</sup>

Aussi, les données générées sont variées et de format différent tout au long du cycle de vie d'un médicament (fichiers Excel, Data Lake, Data Historian, SCADA, MES, LIMS, données de logistiques, etc).<sup>47</sup> Ainsi, la gestion de toutes ces données est un défi technique pour les industries de la santé qui se doivent de déterminer comment standardiser le traitement des données pour obtenir des informations utilisables. En effet, ces données se doivent d'être exploitables par l'humain, mais aussi par les différents systèmes. L'autre défi consiste à gérer l'accès et la transmission des données afin que les informations naviguent au bon endroit au bon moment pour la bonne personne ou la bonne machine. Il s'agit par exemple de savoir quelles données pourraient être utilisées à des fins d'audit interne, de décisions de lancement de produits, de marketing ou de partage avec les organismes de réglementation.<sup>3</sup>

## b. L'intégrité des données et la cybersécurité

En vue de l'adoption des technologies qui nécessitent les Big Data, la protection des données devient primordiale pour chacune des usines pharmaceutiques. La protection des données a plusieurs fonctions. Premièrement, les industriels protègent leurs données pour des raisons de propriétés intellectuelles et pour limiter l'espionnage industriel. Deuxièmement, l'autorité réglementaire veille à ce que les données soient intégrées, tracées et protégées pour la qualité de la fabrication du médicament et pour la sécurité des patients. Le défi pour les futures usines intelligentes est d'assurer que la masse de données générées (BIG DATA) ainsi que les liaisons entre les différentes machines seront traçables et sécurisées pour éviter des défaillances de sécurité. Une défaillance de sécurité risquerait d'augmenter les falsifications, les vols et la contrefaçon. Pour ce faire, certains outils ont déjà été mis en œuvre pour assurer la sécurité et la qualité des médicaments commercialisés, comme les étiquettes RFID ou encore les « Virtual Private Networks » (VPN).

## c. Les défis technologiques

Comme nous l'avons vu, l'intégration des nouveaux systèmes, applications ou appareil au système informatique d'une entreprise est indispensable pour une transition vers le numérique. Il existe 2 types d'intégration à assurer : l'intégration horizontale et l'intégration verticale.<sup>104</sup>

L'intégration verticale se définit comme l'intégration des systèmes de **technologies de l'information et de la communication** (TIC) à différents niveaux hiérarchiques d'une organisation, notamment entre les niveaux de production et le niveau de gestion (entre fabrication et ERP).<sup>104</sup>

L'intégration horizontale est la connexion entre les différentes machines et systèmes, des unités de production et le tout entre les entreprises, au niveau d'une chaîne d'approvisionnement et de fabrication, avec un échange de ressources et d'informations en temps réel.<sup>104</sup>

Or, pour un grand nombre d'entreprises, les infrastructures actuelles des TIC ne sont pas suffisamment développées pour supporter la transformation numérique que visent l'intégration verticale et horizontale<sup>105</sup>



En ce qui concerne le « Cloud computing », comme nous l'avons vu précédemment, un grand nombre d'objets physiques vont se connecter générant une quantité astronomique de données à gérer et à stocker. Entre les données à stocker et les réseaux de fabrication qui sont utilisés pour la transmission d'une grande et volumineuse quantité de données, il se peut que la scalabilité du Cloud ne suive pas.<sup>105</sup>

### **3. Les défis logistiques**

#### **a. La culture d'entreprise**

L'adaptation des outils de l'industrie 4.0 demande des changements de la culture d'entreprise pour les fabricants et les régulateurs. La culture d'entreprise est l'ensemble des éléments qui régissent le fonctionnement d'une entreprise. Selon le travail d'Eric Delavallée et son équipe<sup>106</sup>, la culture d'entreprise est « un ensemble d'évidences, partagées dans l'entreprise, construites tout au long de son histoire, en réponse aux problèmes rencontrés par l'entreprise. ». Ainsi, il régit un système de fonctionnement. Or, l'avènement des nouvelles technologies oblige le système à être muté pour s'adapter aux nouveaux outils numériques, constituant un défi de gestion de l'entreprise pour créer une usine du futur. En effet, il est bien connu que les changements au sein des entreprises font face à des réticences chez les collaborateurs. L'accroissement des inégalités, le changement des activités des employés et de l'organisation globale, le déplacement géographique, ou la dépréciation des compétences demandant des nouvelles formations sont des facteurs pouvant freiner la transition vers l'usine numérique.<sup>66</sup> La transition doit donc être largement étudiée afin de connaître l'impact des changements sur les employés, de les accompagner et de les former au mieux.<sup>66</sup>

Aussi, certains outils comme l'IA demandent des larges banques de données durant la phase d'apprentissage pour devenir efficace. Or, les modèles de partage des données existant actuellement entre les entreprises sont très stricts et limités pour des raisons préalablement citées comme la sécurité et la propriété intellectuelle.<sup>3</sup> Ceci représente donc un défi à régler pour les industriels afin de profiter complètement des capacités des différents outils numériques.<sup>3</sup>

## **B. Impact sur les métiers.**

### **1. Evolution des métiers et création des métiers**

#### **a. Evolution des métiers**

Le changement des infrastructures et de l'organisation des entreprises vont obligatoirement avoir un impact sur les métiers. C'est dans ce sens que les métiers se voient évoluer indispensablement pour la mise en place d'une usine numérique.

Par exemple, la maintenance prédictive permet de planifier les temps d'arrêt des machines. L'opérateur de maintenance devra donc s'adapter pour passer d'une logique d'inspection physique, à une logique de diagnostic et de résolution de problèmes.<sup>107</sup>

L'installation des systèmes cyberphysiques de production permet la personnalisation de masse, la réadaptation du plan de production à une variation de la demande ou à un besoin de réactivité.<sup>107</sup> Ce système va modifier les métiers de la planification et la logistique qui devront s'adapter afin de passer d'une logique de production en push (produire pour stocker) à une logique de production en pull (produire sur commande).

Aussi, l'impact des nouvelles technologies sur la fabrication est tel que le personnel nécessaire au fonctionnement de l'usine sera réduit de moitié environ.<sup>107</sup> Cependant, comme l'explique Max Blancher <sup>107</sup>, l'humain est centré dans ce nouveau modèle. En effet, les risques de santé liés à la pénibilité et aux tâches répétitives seront largement réduits (par les cobots par exemple), et sa relation avec le client interne, sa compréhension du besoin, sa résolution du problème seront les axes principaux des métiers.<sup>107</sup> Il s'agit donc de redéfinir les postes de travail qui vont tendre vers plus d'autonomie et de responsabilités.<sup>108</sup>

#### **b. Suppression des métiers**

Inéluctablement, l'arrivée de nouvelles technologies supprime des emplois, mais les remplace par de nouveaux métiers. Ainsi, d'après l'analyse de PWC datant de 2018, l'intelligence artificielle pourrait supprimer jusqu'à 7 millions d'emplois, mais en créera 7,2 millions.<sup>81</sup> L'idée est donc de supprimer les postes opérationnels qui sont pénibles et répétitifs et les remplacer par des robots automatisés qui pourront être plus précis. Bien entendu, même si les emplois opérationnels sont dans certains cas supprimés, une personne en charge de surveillance et de

contrôle reste indispensable. Aussi, il est important de notifier que la plupart du temps les employés ne seront pas licenciés mais plutôt formés puis mutés à un autre poste.<sup>81</sup>

### **c. Création de nouveaux métiers**

L'adoption et l'installation des nouvelles technologies demandent de nouvelles compétences jusque-là inexistantes. La mise en place de nouveaux systèmes informatiques ainsi que leur intégration nécessiteront des connaissances poussées en informatique, en mathématique et en ingénierie. La création de données en masse nécessite des compétences de gestion et d'analyse des données et donc la création des métiers de Data Manager, Data Analyst ou encore Data scientist. Le premier, le data manager, ou gestionnaire de données, recueille et organise les informations de l'entreprise.<sup>109</sup> Le deuxième travaille sur des données déjà extraites et organisées. Il analyse ces données pour répondre aux questions soumises par son entreprise afin de fournir des solutions.<sup>110</sup>

Le troisième a une vision plus globale et nécessite des connaissances profondes en informatique et mathématique. Il collecte, trie et organise les données, les croise et les analyse dans le but de résoudre des problèmes, Il se charge de traduire les problèmes business en problèmes mathématiques et statistiques, afin de fournir des rapports permettant d'orienter les prises de décision du management et d'améliorer les performances et les stratégies marketing.<sup>111</sup> Idéalement, un data scientist est capable de créer des modèles d'IA pour aider à la prise de décision en autonomie.

Il peut être pertinent de mentionner que l'IA, à elle seule, a fait apparaître de nouveaux métiers autour de sa construction. Le « psydesigner » qui joue un rôle important dans l'architecture de l'IA, ou encore l'ergoteller qui façonne la personnalité de cette dernière.

Toujours dans le numérique, la sécurisation des données et des applications demande des compétences numériques. C'est l'expert en sécurité informatique qui sécurise les systèmes collectant les données, les données informatiques en elles-mêmes et les programmes d'intelligence exploitant ces données.<sup>112</sup> Il contribue globalement à l'ensemble de la sécurité numérique d'une entreprise.

La réalité virtuelle et augmentée demande aussi des animateurs et formateurs qui pourront former les collaborateurs à l'utilisation de la RA ou des jumeaux numériques.<sup>109</sup>

Dans le futur, lorsque l'application des ordinateurs quantiques se fera de manière industrielle, un informaticien quantique qui pourra utiliser les fonctionnements et les algorithmes de ces ordinateurs, fondamentalement différents, pour l'entreprise sera indispensable.<sup>109</sup>

Aussi, au niveau managérial, de nouveaux métiers sont créés pour diriger les opérations de digitalisation et les stratégies d'application, comme les « Chief digital Officer » ou le responsable de l'usine du futur.

## **2. Discussion sur l'évolution du métier de pharmacien**

Comme nous l'avons expliqué plus haut, tous les métiers de l'industrie vont se voir évoluer dans ce nouvel environnement numérique et de plus en plus autonome. De ce fait, il est normal de se poser la question sur les possibles changements liés au métier de pharmacien industriel.

Il semble logique de penser que comme pour tous les métiers, le pharmacien devra s'adapter afin de comprendre le système de fonctionnement des outils et de l'organisation générale de l'entreprise. Tout en étant le garant de la sécurité et de la qualité des médicaments fabriqués, il est possible que la portée du métier nécessite d'être redéfinie. Par exemple, l'arrivée du PAT, permettant de contrôler la qualité des médicaments en gérant de manière optimisée et en temps réel les processus de fabrication, impactera obligatoirement la façon de faire actuelle du contrôle qualité des médicaments fabriqués en lot. Le pharmacien aura par exemple pour rôle le contrôle et la surveillance du logiciel ainsi que sa mise à jour en fonction des changements de la réglementation ou des spécifications.

Dans tous les cas, les pharmaciens industriels devront mettre à jour leurs connaissances et leurs compétences en informatique et en mathématiques pour faciliter les interactions avec ces outils novateurs afin d'assurer continuellement la qualité de fabrication des médicaments.

Enfin, il est possible, en fonction de l'ampleur de l'intégration de ces outils dans le monde pharmaceutique, que la fonction de pharmacien soit transformée au point que la formation aux études supérieures de pharmacien industriel soit fortement impactée. Aujourd'hui, il est commun lors de la 6<sup>ème</sup> année de pharmacie de s'orienter vers un double diplôme de pharmacien/ingénieur. Cette évolution pédagogique est une conséquence de la demande en connaissances et compétences d'ingénierie dans le secteur industriel. Il n'est donc pas

improbable que demain, pour intégrer l'industrie 4.0, des compétences en mathématiques, statistiques, informatiques encore plus poussées ainsi que des connaissances sur le rôle et le fonctionnement des différents outils de l'industrie pharmaceutiques, soient intégrées dans le cursus de la formation de pharmacien industriel.

## **V- Conclusion.**

Le secteur industriel a connu de nombreuses révolutions poussées par des innovations technologiques ou par le changement de modèle économique et social. Souvent, les innovations sont telles, qu'elles révolutionnent les infrastructures et la gestion des entreprises. L'industrie 4.0 est, dans ce sens, un concentré de révolution technologique basé sur un modèle économique numérique, de mondialisation de l'information, et de la personnalisation de l'offre. En effet, la digitalisation et les nouvelles technologies représentent un avantage concurrentiel fort pour l'ensemble des entreprises du secteur.

Depuis plus de 50 ans, ces outils détaillés plus haut, sont développés afin d'améliorer la productivité, l'efficacité, la sécurité des différents produits de fabrication. Dans l'industrie pharmaceutique 4.0, ces technologies sont utilisées pour centrer la chaîne de valeurs autour du patient et des données de santé. En effet, ils transforment fondamentalement l'ensemble des étapes de fabrication du médicament, depuis sa conception jusqu'à sa délivrance en passant par toutes les fonctions supports comme la qualité et la maintenance. Leur utilité est indéniable et applicable à l'industrie pharmaceutique comme à beaucoup d'autres secteurs. C'est dans ce sens que l'industrie 4.0 et ses technologies sont appelées : révolution industrielle.

Cependant, comme nous l'avons vu, de nombreux défis sont toujours présents. Comme pour tous les secteurs, ce changement de paradigme implique des métamorphoses au sein même de l'organisation d'une usine. L'implémentation de nouvelles technologies implique des moyens, des compétences, des structures physiques ainsi qu'une mise à jour des compétences des collaborateurs ou dans certains cas la suppression de certains postes de travail.

Les avancées technologiques ne sont dans certains cas pas encore au point. En effet, que ce soit le quantum computing, les blockchains ou encore certaines IA, toutes les nouvelles technologies ne sont pas assez développées pour s'intégrer de manière optimale dans l'ensemble de la chaîne de valeurs de l'industrie pharmaceutique. Aussi, les nouvelles technologies étant génératrices d'une grande quantité de données, la question de la gestion de ces données est de vigueur.

Un des grands défis du secteur pharmaceutique réside dans la rigidité de la réglementation, beaucoup plus stricte que pour les autres secteurs. De ce fait, elle ralentit les entreprises pharmaceutiques dans la décision de passer vers un modèle digital et innovant encore peu

réglementé. Du côté des institutions réglementaires, les grands défis de l'industrie 4.0 sont de réussir à adapter les textes, les recommandations, et les audits aux nouvelles technologies qui se basent plus sur une production en continu que sur une production par lots de produits.

Aussi, il est important de noter que cette révolution risque de bouleverser les métiers. Ces nouvelles technologies font naître de nouveaux métiers corrélés au besoin de nouvelles compétences, principalement autour de la gestion et l'analyse de données, l'informatique et les statistiques. Aussi, la robotique remplacera certaines tâches répétitives et pénibles pour l'Homme supprimant ainsi certains métiers. Tous les métiers devront s'adapter aux nouveaux outils implémentés impliquant des formations et l'acquisition de nouvelles compétences.

Ainsi il est donc possible d'implémenter les nouvelles technologies et la transition vers la pharma 4.0 est possible si les défis sont résolus.

Enfin, beaucoup reste à faire. Toutes les questions ne sont résolues et de nouvelles questions se posent avec l'avancée vers La transition 4.0 :

La réglementation pourra-t-elle s'adapter ? Quel avenir pour les métiers supprimés ? Comment être sûr de la sécurité des données ? La personnalisation des traitements en masse est-elle réellement possible ?.

Beaucoup de ces problématiques restent pour l'instant partiellement résolues ou sans réponses et dépendront des décisions des industriels et des institutions réglementaires à l'avenir.

## VI- Bibliographie

1. Monostori L. Cyber-physical Production Systems: Roots, Expectations and R&D Challenges. *Procedia CIRP*. 2014;17:9-13. doi:10.1016/j.procir.2014.03.115
2. Larousse É. machine à vapeur - LAROUSSE. Disponible sur : [https://www.larousse.fr/encyclopedie/divers/machine\\_%C3%A0\\_vapeur/100756](https://www.larousse.fr/encyclopedie/divers/machine_%C3%A0_vapeur/100756)
3. Duan Pinglia, Bala Anand Muthu, Seifedine Nimer Kadry. Business School, Northwest University of Political Science and Law, Xi'an, China bV.R.S. College of Engineering & Technology, Villupuram, India cBeirut Arab University, Beirut, Lebanon: Industry 4.0 for pharmaceutical manufacturing: Preparing for the smart factories of the future. 22 January 2020. <https://content.iospress.com/articles/work/wor203429>
4. Amina Hasani, L'industrie 4.0 et les facteurs clés de succès de projet, 98p, Mémoire de maîtrise en gestion de projet, Aout 2020, Quebec. <https://depot-e.uqtr.ca/id/eprint/9494/1/eprint9494.pdf>
5. Mercier-Laurent E. Intelligence artificielle 4.0 pour l'Industrie 4.0. *1024 Bull Société Inform Fr*. 2020;15:127-137. doi:10.48556/SIF.1024.15.127
6. Naren.J, Sowmya SK, Deepika P. Layers of Cloud – IaaS, PaaS and SaaS: A Survey. *Int J Comput Sci Inf Technol*. 2014;Vol. 5 (3). p4477-4480.
7. Goyal S. Public vs Private vs Hybrid vs Community - Cloud Computing: A Critical Review. *Int J Comput Netw Inf Secur*. 2014;6(3). p20-29. doi:10.5815/ijcnis.2014.03.03
8. Favaretto M, De Clercq E, Schneble CO, Elger BS. What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade. *PLoS ONE*. 2020;15(2):e0228987. doi:10.1371/journal.pone.0228987
9. Ristevski B, Chen M. Big Data Analytics in Medicine and Healthcare. *J Integr Bioinforma*. 2018;15(3):20170030. doi:10.1515/jib-2017-0030
10. Reimer AP, Madigan EA. Veracity in big data: How good is good enough. *Health Informatics J*. 2019;25(4):1290-1298. doi:10.1177/1460458217744369
11. Benghozi PJ, Bureau S, Massit-Folléa F. *L'Internet des objets / The Internet of Things: Quels enjeux pour l'Europe ? / What Challenges for Europe?* Les Editions de la MSH; 2015.
12. Hohmann Christian. Qu'est-ce qu'un système cyber-physique [En Ligne]. Référence du 29 Octobre 2017. Disponible sur : <http://christian.hohmann.free.fr/index.php/usine-du-futur/744-quest-ce-quun-systeme-cyber-physique>
13. Georges B. Intelligence artificielle : de quoi parle-t-on ? *Constructif*. 2019;54(3), p :5-10. doi:10.3917/const.054.0005
14. Mahesh B. *Machine Learning Algorithms -A Review.*; International Journal of Science and Research (IJSR).2020, 9(1), p 381-386. doi:10.21275/ART20203995
15. Massih-Reza Amini. Principes de base en apprentissage supervisé. Eyrolles. Machine Learning, , 2020. fahal-03049016f



16. Rakotomalala.R. Revue MODULAD. n°33 2005. Arbres de Décision. Lyon.p163-187. Disponible sur : <http://www.modulad.fr/archives/numero-33/tutorial-rakotomalala-33/rakotomalala-33-tutorial.pdf>
17. Cornfield J. Bayes Theorem. *Rev Inst Int Stat Rev Int Stat Inst.* 1967;35(1), p 34. doi:10.2307/1401634
18. Crochet-Damais.A, Classification naïve bayésienne : définition et principaux avantages. [En ligne] Référence de février 2022. Disponible sur : <https://www.journaldunet.fr/web-tech/guide-de-l-intelligence-artificielle/1501321-classification-naive-bayesienne-definition/>
19. Noble WS. What is a support vector machine? *Nat Biotechnol.* 2006;24(12), p: 1565-1567. doi:10.1038/nbt1206-1565
20. Sublime J. L'apprentissage non-supervisé et ses contradictions. *1024 Bull Société Inform Fr.* 2022;(19), p145-156. doi:10.48556/SIF.1024.19.145
21. Masson M. Analyse en composantes principales face à la malédiction de la dimensionnalité. Pensée Artificielle. [en ligne] référence du 14 juin 2021. Disponible sur : <https://penseeartificielle.fr/analyse-composantes-principales-face-malediction-dimensionnalite/>
22. JVC J. K-Means : fonctionnement et utilisation dans un projet de clustering. Data Transition Numérique. [en ligne] Référence du 30 novembre 2021. Disponible sur : <https://www.data-transitionnumerique.com/k-means/>
23. Thibault.V. K-means : Focus sur cet algorithme de Clustering & Machine Learning [En Ligne].Référence du 15 mai 2020. DataScientest.com. Disponible sur : <https://datascientest.com/algorithme-des-k-means>
24. Magnan C. Apprentissage Semi-Supervisé Asymétrique et Estimations d'affinités Locales Dans Les Protéines. Actes de CAP 05, Conférence francophone sur l'apprentissage automatique. 2005 p : 297-312.
25. Vapnik V. Estimation of Dependences Empirical Based pages. *eBook*, Springer New York, ed., 2006. doi:10.1007/0-387-34239-7
26. Sarkar A, Haffari G. Inductive Semi-supervised Learning. School of Computing Science. Simon Fraser University. Vancouver, BC, Canada. 69p.
27. Yu X, Yang J, Zhang J pei. A Transductive Support Vector Machine Algorithm Based on Spectral Clustering. *AASRI Procedia.* 2012;1. p384-388. doi:10.1016/j.aasri.2012.06.059
28. Safarpoor A, Kalra S, Tizhoosh HR. Generative models in pathology: synthesis of diagnostic quality pathology images†. *J Pathol.* 2021;253(2).p 131-132. doi:10.1002/path.5577
29. Wang K, Gou C, Duan Y, Lin Y, Zheng X, Wang FY. Generative adversarial networks: introduction and outlook. *IEEECAA J Autom Sin.* 2017;4(4). p 588-598. doi:10.1109/JAS.2017.7510583
30. L.Bastien. Reinforcement Learning : qu'est-ce que l'apprentissage par renforcement ? LeBigData.fr. [en ligne] référence du 30 Juin 2022. Disponible sur : <https://www.lebigdata.fr/reinforcement-learning-definition>
31. Filippi S. Stratégies optimistes en apprentissage par renforcement. 180p. Th.D. Mathématique. Ecole nationale supérieure des télécommunications (ENST).Paris. 2010

32. L B. Réseau de neurones artificiels : qu'est-ce que c'est et à quoi ça sert ? [en ligne] LeBigData.fr. référence du 5 Avril 2019. Disponible sur : <https://www.lebigdata.fr/reseau-de-neurones-artificiels-definition>
33. Wang L, Michoel T. Accurate wisdom of the crowd from unsupervised dimension reduction. *R Soc Open Sci.* 6(7):181806. doi:10.1098/rsos.181806
34. Bi Q, Goodman KE, Kaminsky J, Lessler J. What is Machine Learning? A Primer for the Epidemiologist. *Am J Epidemiol.* 2019;188(12), p 2222-2239. doi:10.1093/aje/kwz189
35. Briscoe E, Feldman J. Conceptual complexity and the bias/variance tradeoff. *Cognition.* 2011;118(1), p2-16. doi:10.1016/j.cognition.2010.10.004
36. Yang Z, Yu Y, You C, Steinhardt J, Ma Y. Rethinking Bias-Variance Trade-off for Generalization of Neural Networks. In: *Proceedings of the 37th International Conference on Machine Learning.* PMLR; 2020:10767-10777. Disponible sur : <https://proceedings.mlr.press/v119/yang20j.html>
37. Ensemble learning : définition, fonctionnement et algorithmes.[En ligne] Référence du 3 mai 2022. Disponible sur : <https://www.journaldunet.fr/web-tech/guide-de-l-intelligence-artificielle/1501883-methode-ensembliste-ensemble-learning/>
38. Sagi O, Rokach L. Ensemble learning: A survey. *WIREs Data Min Knowl Discov.* 2018;8(4). doi:10.1002/widm.1249
39. Domingos, Pedro, Bayesian Averaging of Classifiers and the Overfitting Problem Proceedings of the Seventeenth International Conference on Machine Learning, 2002. 8p
40. Hossain SMM, Khatun L, Ray S, Mukhopadhyay A. Pan-cancer classification by regularized multi-task learning. *Sci Rep.* 2021;11:24252. doi:10.1038/s41598-021-03554-8
41. Keogh, E. Instance-Based Learning. In: Sammut, C., Webb, G.I. (eds) Encyclopedia of Machine Learning. 2011, p 549 -550. Disponible sur : [https://doi.org/10.1007/978-0-387-30164-8\\_409](https://doi.org/10.1007/978-0-387-30164-8_409)
42. Zhang Z. Introduction to Machine Learning: k-nearest neighbors. *Ann Transl Med.* 2016;4 (11):218. doi:10.21037/atm.2016.03.37
43. Sun S, Huang R. An adaptive k-nearest neighbor algorithm. *2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery.* IEEE; 2010 : p91-94. doi:10.1109/FSKD.2010.5569740
44. Kletti J, *Manufacturing Execution Systems —MES.* Springer Berlin Heidelberg; 2007. doi:10.1007/978-3-540-49744-8
45. van Merode GG, Groothuis S, Hasman A. Enterprise resource planning for hospitals. *Int J Med Inf.* 2004;73(6). p 493-501. doi:10.1016/j.ijmedinf.2004.02.007
46. Sekkat S, Abadi A, Zemmouri E, Benazza H. Utilisation des ontologies pour satisfaire les exigences d'interopérabilité des systèmes informatiques de pilotage de production. Xème Conférence Internationale : Conception et Production Intégrées, Dec 2015, Tanger, Maroc. (hal-01260785)
47. Steinwandter V, Borchert D, Herwig C. Data science tools and applications on the way to Pharma 4.0. *Drug Discov Today.* 2019;24(9). p 1795-1805. doi:10.1016/j.drudis.2019.06.005
48. Sekkat S, Paris JL, Kouiss K. . Developpement du modele conceptuel d'un systeme d'indicateurs de performance pour le pilotage d'une cellule flexible d'usinage. Dans: "Evaluation et

optimisation des systèmes innovants de production de biens et de services". 8 e Conférence Internationale de Modélisation et Simulation - MOSIM'10 - 10 au 12 mai 2010 - Hammamet – Tunisie. 2010, P10.

49. Canzoneri M, De Luca A, Harttung J. Digital Twins: A General Overview of the Biopharma Industry. In: Herwig C, Pörtner R, Möller J, eds. *Digital Twins*. Vol 177. Advances in Biochemical Engineering/Biotechnology; 2020. p 167-184. doi:10.1007/10\_2020\_157
50. Caudell TP, Mizell DW. Augmented reality: an application of heads-up display technology to manual manufacturing processes. In: *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences*. IEEE; 1992 ; 2. p 659-669. doi:10.1109/HICSS.1992.183317
51. Craig AB. *Understanding Augmented Reality: Concepts and Applications*. Kindle Ed. 2013.
52. Anastassova MS, Burkhardt JM, Mégard C, Ehanno P. L'ergonomie de la réalité augmentée pour l'apprentissage : une revue. *Trav Hum*. 2007;70(2). p 97-125. doi:10.3917/th.702.0097
53. Berchon M, Luyt B. *L'impression 3D*. Editions Eyrolles. 2014.
54. Cazaubon V. Procédé de fabrication additive par dépôt de fil fondu: modèle, méthodes et stratégies pour la correction de défaut de pièce mécanique. p189. Th.D. Sciences physiques et de l'ingénieur spécialité mécanique. Institut de Mécanique et d'Ingénierie (I2M) Bordeaux. 2021
55. Gueche Y. Etude de l'impression 3D par frittage laser sélectif pour la conception et le développement de formes pharmaceutiques orales sèches. p277. Th.D. Génie des procédés. Université de Montpellier. 2021
56. Robots collaboratifs. De quoi parle-t-on ? INRS.2021.[En ligne] Référence du 8 Décembre 2020. Disponible sur : <https://www.inrs.fr/risques/robots-collaboratifs/de-quoi-parle-t-on.html>
57. Castro A, Silva F, Santos V. Trends of Human-Robot Collaboration in Industry Contexts: Handover, Learning, and Metrics. *Sensors*. 2021;21(12):4113. doi:10.3390/s21124113
58. Michalos G, Makris S, Tsarouchi P, Guasch T, Kontovrakis D, Chryssolouris G. Design Considerations for Safe Human-robot Collaborative Workplaces . *Procedia CIRP* 37. Understanding the life cycle implications of manufacturing. 2015. p248 – 253. Disponible sur :<https://www.sciencedirect.com/science/article/pii/S2212827115008550>
59. Toichoa Eyam A, Mohammed WM, Martinez Lastra JL. Emotion-Driven Analysis and Control of Human-Robot Interactions in Collaborative Applications. *Sensors*. 2021;21(14) p 4626. doi:10.3390/s21144626
60. Hahn, A. Operational Technology and Information Technology in Industrial Control Systems. Cyber-security of SCADA and Other Industrial Control Systems. 2016. , 66, p 58-61. [https://doi.org/10.1007/978-3-319-32125-7\\_4](https://doi.org/10.1007/978-3-319-32125-7_4)
61. Redelinghuys, AJH, Basson, AH & Kruger, K. Cybersecurity Considerations for Industrie 4.0. 7th International Conference on Competitive Manufacturing. pp. 266-271. 30 Jan - 1 Feb 2019, Stellenbosch, South Africa
62. La cybersécurité des systèmes Industriels .Mesure détaillé, Agence nationale de la sécurité des systèmes d'informations.2014. LIP V1 2010.04.02. Disponible sur : [https://www.ssi.gouv.fr/uploads/2014/01/securete\\_industrielle\\_GT\\_details\\_principales\\_mesures.pdf](https://www.ssi.gouv.fr/uploads/2014/01/securete_industrielle_GT_details_principales_mesures.pdf)

63. CEA. Le calcul et l'ordinateur quantiques. CEA/Découvrir & Comprendre. 2021.[En ligne].  
Référence du 18 mai 2021. Disponible sur : <https://www.cea.fr/comprendre/Pages/nouvelles-technologies/essentiel-sur-ordinateur-quantique.aspx>
64. Bercy Infos. Qu'est-ce que la blockchain ? [En ligne]. Référence du 12 avril 2022 Ministère de l'économie, des finances et de la souveraineté industrielle et numérique. 2022. Disponible sur : <https://www.economie.gouv.fr/entreprises/blockchain-definition-avantage-utilisation-application>
65. Deroy X. Le secteur pharmaceutique et l'histoire du contrôle de l'innovation. *Rev Fr Gest.* 2008;188-189(8-9), p 175-183. doi:10.3166/rfg.188-189.175-183
66. Dousson T. Digitalisation en production pharmaceutique : vers l'industrie 4.0. 113p, Th.D. Science Pharmaceutique. Grenoble. 2020
67. European Court of Auditors. Digitising European Industry.2019. Disponible sur : [https://www.eca.europa.eu/lists/ecadocuments/ap19\\_13/ap\\_\\_digitising\\_industry\\_en.pdf](https://www.eca.europa.eu/lists/ecadocuments/ap19_13/ap__digitising_industry_en.pdf)
68. Rosin F, Forget Pascal, Lamouri S, Pellerin R. Industrie 4.0 et Lean : Hiérarchie, Amélioration ou Complémentarité ;Conférence: CIGI QUALITA 2019, "Le numérique et l'intelligence industrielle: défis et opportunités, Montréal (Qc) Canada.
69. Digital plant maturity model (DPMM): assessment tool. BioPhorum. 2018. Disponible sur : <https://www.biophorum.com/download/digital-plant>
70. Hammouche D. Le lean management appliqué à la documentation opérationnelle dans une unité de production pharmaceutique. 132p. Th. D en Sciences pharmaceutiques. Marseille. 2017
71. International Conference on Harmonisation. ICH missions. Disponible sur: <https://www.ich.org/about/mission.html>
72. Buisine L. La qualité et son management en industrie pharmaceutique : s'imposer un cadre restrictif ou plutôt s'ouvrir à de nouveaux horizons ?. 108p, Th.D en Sciences pharmaceutiques. Université de Lorraine. 2016
73. Rumor O. Méthode d'analyse de la conformité des données et des équipements en accord avec les recommandations des autorités compétentes en termes de data integrity. 84p. Th D en Sciences pharmaceutiques. Université de Lorraine. 2019.
74. Manzano et Langer . Pharmaceutical Engineering . Getting ready for pharma 4.0. Technical Information Systems. p72-79. 2018. Disponible sur : [https://www.ispe.gr.jp/ISPE/02\\_katsudou/pdf/201812\\_en.pdf](https://www.ispe.gr.jp/ISPE/02_katsudou/pdf/201812_en.pdf)
75. De Corbière F, Rowe F, Wolff FC. De l'intégration interne du système d'information à l'intégration du système d'information de la chaîne logistique. Document de travail LEMNA, 2012.
76. McDowall RD. The role of laboratory information management systems (LIMS) in analytical method validation. *Anal Chim Acta.* 1999;391(2). p 149-158. doi:10.1016/S0003-2670(99)00107-5
77. FDA. Guidance for Industry PAT - A Framework for Innovative Pharmaceutical Development, manufacturing, and Quality Assurance. 2004, 19p.

78. Sharma A, Kaur J, Singh I. Internet of Things (IoT) in Pharmaceutical Manufacturing, Warehousing, and Supply Chain Management. *SN Comput Sci.* 2020;1(4), p232. doi:10.1007/s42979-020-00248-2
79. Collot M, Coutant M, Orsini M, Fallou O. EDEC. Industrie Santé. impacts compétences et métiers de 7 technologies numériques dans les industries de santé[En Ligne]. Référence du 15 Janvier 2021. 243p.2021 Disponible sur : [https://www.leem.org/sites/default/files/2021-06/EDEC%20Industrie%20Sant%C3%A9%20-%207%20technos%20num%C3%A9riques%20-%20Rapport%20270121\\_compressed.pdf](https://www.leem.org/sites/default/files/2021-06/EDEC%20Industrie%20Sant%C3%A9%20-%207%20technos%20num%C3%A9riques%20-%20Rapport%20270121_compressed.pdf)
80. Sommer T. Cloud Computing in Emerging Biotech and Pharmaceutical Companies. The International Information Management Association. Vol 13. Problématique n°3. 19p. USA. 2013. Disponible sur : <https://scholarworks.lib.csusb.edu/cgi/viewcontent.cgi?article=1218&context=ciima>
81. Hoerd A. Intelligence Artificielle et Industrie Pharmaceutique. 124p. Th.D. Sciences Pharmaceutiques . Tours. 2020. n°74
82. AlphaFold: Using AI for scientific discovery.[En ligne]. Référence du 15 janvier 2020. Disponible sur : <https://www.deepmind.com/blog/alphafold-using-ai-for-scientific-discovery-2020>
83. Perron Q et al, Institut Servier, IKTOS. Deep learning for ligand-based de novo design in lead optimization: a real life case study. Disponible sur : [https://iktos.ai/wp-content/uploads/2021/11/Poster-Servier-Deep-learning-for-ligand-based-de-novo-design-in-lead-optimization\\_-a-real-life-case-study.pdf](https://iktos.ai/wp-content/uploads/2021/11/Poster-Servier-Deep-learning-for-ligand-based-de-novo-design-in-lead-optimization_-a-real-life-case-study.pdf)
84. Talha Burki. The Lancet Digital Health, Mai 2020, Volume 2, Issue 5, Pages p 209-267. Mai 2020. Disponible sur [https://www.thelancet.com/journals/landig/issue/vol2no5/PIIS2589-7500\(20\)X0005-7](https://www.thelancet.com/journals/landig/issue/vol2no5/PIIS2589-7500(20)X0005-7)
85. Collet N, Jaremko S, Christine G, Matine A. AEC partners. L'intelligence artificielle dans les industries de santé. Rapport de l'AEC Partners. Septembre 2020, <https://www.leem.org/sites/default/files/2020-12/AEC%20Partners%20-%20Rapport%20Intelligence%20Artificielle%20en%20Sant%C3%A9%20-%2020201215.pdf>
86. Gérald PETITJEAN, l'Intelligence Artificielle et l'Aide à la Décision dans les entreprises, Bulletin de l'AFIA,2006. Disponible sur : <https://afia.asso.fr/wp-content/uploads/2016/10/AFIA62.pdf>
87. ANSM : Agence nationale de sécurité du médicament et des produits de santé. Qualité et sécurité des médicaments : les obligations.[En ligne]. Référence de 2017 Disponible sur : [https://archiveansm.integra.fr/Dossiers/Medicaments-generiques/Qualite-et-securite-des-medicaments-les-obligations/\(offset\)/2](https://archiveansm.integra.fr/Dossiers/Medicaments-generiques/Qualite-et-securite-des-medicaments-les-obligations/(offset)/2)
88. ANSM : Agence nationale de sécurité du médicament et des produits de santé. Actualité - Lancement de la nouvelle application nationale de pharmacovigilance [En Ligne]. Référence du 7 mai 2021. Disponible sur : <https://ansm.sante.fr/actualites/lancement-de-la-nouvelle-application-nationale-de-pharmacovigilance>
89. Zwolińska B, Tubis AA, Chamier-Gliszczyński N, Kostrzewski M. Personalization of the MES System to the Needs of Highly Variable Production. *Sensors.* 2020;20(22). p 6484. doi:10.3390/s20226484
90. ExactCure enables accurate dosage for drug trials. European Institute of Innovation & Technology (EIT)[En ligne]. Référence du 15 avril 2020. Disponible sur : <https://eit.europa.eu/our-activities/covid-19-response/solutions/drug-trials>

91. Zobel-Roos, S., Schmidt, A., Uhlenbrock, L., Ditz, R., Köster, D., Strube, J. (2020). Digital Twins in Biomanufacturing. Dans : Herwig, C., Pörtner, R., Möller, J. (eds) Digital Twins. Advances in Biochemical Engineering/Biotechnology, vol 176. Springer, Cham. Disponible sur : [https://doi.org/10.1007/10\\_2020\\_146](https://doi.org/10.1007/10_2020_146)
92. Sanofi. Digitalisation des Affaires Industrielles [En Ligne]. Référence du 09 avril 2020 Disponible sur : <https://www.sanofi.com/fr/nous-connaître/nos-recits/digitalisation-des-affaires-industrielles>
93. Reinhardt IC, Oliveira JC, Ring D. Industry 4.0 and the Future of the Pharmaceutical Industry. *Pharmaceutical Engineering*. 2021;41. Disponible sur : [https://www.researchgate.net/publication/350663515\\_Industry\\_40\\_and\\_the\\_Future\\_of\\_the\\_Pharmaceutical\\_Industry](https://www.researchgate.net/publication/350663515_Industry_40_and_the_Future_of_the_Pharmaceutical_Industry)
94. PTC. Le retour d'expérience de Takeda sur l'implémentation de la Réalité Augmentée dans l'industrie pharmaceutique [En Ligne]. Référence du 21 décembre 2022 Disponible sur : <https://www.ptc.com/fr/resources/augmented-reality/video/implementation-realite-augmentee-industrie-pharmaceutique>
95. Billaud M, Guchet X. L'invention de la médecine personnalisée: Entre mutations technologiques et utopie. *médecine/sciences*. 2015;31(8-9). p 797-803. doi:10.1051/medsci/20153108020
96. Lakshmanan S, Chandravadhana S. collaborative robots for industries. *International Research Journal in Global Engineering and Sciences*. Vol 4. N°3. p144-150. 2019. ISSN: 2456-172X <http://www.irjges.com/Volume4Issue3/paper18.pdf>
97. Universal robots. Nos Cobots à Copenhague - la Robotique Collaborative optimise les analyses au Centre Hospitalier Universitaire [En Ligne]. Référence du 21 décembre 2022. Disponible sur : <https://www.universal-robots.com/fr/études-de-cas/gentofte-hospital/>
98. Universal Robots. La cobotique au service du secteur pharmacie et produits médicaux.[En ligne]. Référence du 13 octobre 2021. Disponible sur : [https://www.faure-technologies.com/data/medias/1983/style/default/UR\\_Applications\\_Pharmaceutique.pdf#95](https://www.faure-technologies.com/data/medias/1983/style/default/UR_Applications_Pharmaceutique.pdf#95).
99. Schöner MM, Kourouklis D, Sandner P, Gonzalez E, Förster J. Blockchain Technology in the Pharmaceutical Industry. 9p.
100. Tietje C, Brouder A, eds. International Conference On Harmonisation Of Technical Requirements For Registration Of Pharmaceuticals For Human Use. Dans : *Handbook of Transnational Economic Governance Regimes*. Brill | Nijhoff; 2010:1041-1053. doi:10.1163/ej.9789004163300.i-1081.897
101. Roché Y. Les nouveaux concepts de gestion de la qualité pharmaceutique ICH Q8, Q9 et Q10, 2011 [en ligne] .. disponible sur : [https://www.acadpharm.org/dos\\_public/ROCHE\\_Yves\\_2011.03.02.pdf](https://www.acadpharm.org/dos_public/ROCHE_Yves_2011.03.02.pdf).
102. Silva F, Resende D, Amorim M, Borges M. A Field Study on the Impacts of Implementing Concepts and Elements of Industry 4.0 in the Biopharmaceutical Sector. *J Open Innov Technol Mark Complex*. 2020; 6. p 175. doi:10.3390/joitmc6040175
103. Amrih P, Damayanti RW . Pharma 4.0 Quality Management Challenge A Literature Review. Dans: International Conference on Industrial Engineering and Operations Management. Istanbul.7-10 Mars 2022. Disponible sur : <http://ieomsociety.org/proceedings/2022istanbul/106.pdf>

104. Zant CE. Industrie 4.0 processus – Quel modèle pour une intégration réussie. 249p.Th.D. École Nationale Supérieure d'Arts et Métiers. Paris. 2021
105. Li Da Xu, Eric L. Xu & Ling Li. Industry 4.0: state of the art and future trends, International Journal of Production Research, 2018, 56(8), p 2941-2962, DOI: 10.1080/00207543.2018.1444806
106. Delavallée - Culture d'entreprise la contribution de Herbert . IAE de Paris (Université Paris 1 Panthéon - Sorbonne ). GREGOR .1995.01. Disponible sur : <http://www.sietmanagement.fr/wp-content/uploads/2016/04/CultureDeLaValle%CC%81e.pdf>
107. Blanchet M. Industrie 4.0 : nouvelle donne industrielle, nouveau modèle économique. *Géoéconomie*. 2016 ;82(5). P 37-53. doi:10.3917/geoec.082.0037
108. Gaudron, P. & Mouline, A.. Compte rendu de [Les enjeux de l'industrie 4.0]. Management international. 2017 ;21(2), p165–169. <https://doi.org/10.7202/1052695ar>
109. Ficher Metier de Data manager[En ligne]. Référence du 21 décembre 2022 Onisep. [www.onisep.fr](http://www.onisep.fr). Disponible sur : [https://www.onisep.fr/view/content/770535/version\\_longue](https://www.onisep.fr/view/content/770535/version_longue)
110. Futura. Data analyst ou data scientist : quel métier choisir ?[En Ligne]. Référence du 19 Mai 2021. Disponible sur : <https://www.futura-sciences.com/sciences/questions-reponses/skillz-data-analyst-data-scientist-metier-choisir-15458/>
111. \_ Fiche metier de Data scientist. Onisep [En ligne]. Référence du 21 Décembre 2022. [www.onisep.fr](http://www.onisep.fr). Disponible sur : <https://www.onisep.fr/Ressources/Univers-Metier/Metiers/data-scientist>
112. Bourgognon C, Buisine S, Guillemot G, Kestelyn X, Müller Segard L, Saveuse M. Enseigner l'industrie du futur. L'institut de la réindustrialisation. 2018. 88p. Disponible sur : [https://www.researchgate.net/publication/331865875\\_ENSEIGNER\\_L'INDUSTRIE\\_DU\\_FUTUR\\_-\\_LIVRE\\_BLANC\\_DU\\_COLLOQUE](https://www.researchgate.net/publication/331865875_ENSEIGNER_L'INDUSTRIE_DU_FUTUR_-_LIVRE_BLANC_DU_COLLOQUE)

---

**TITLE :** Industry 4.0 tools and their application to the pharmaceutical industry

---

**ABSTRACT :**

Depending on the evolution of economic and social models, as well as technological innovations, industry has experienced successive revolutions to adapt to competitiveness and consumer demand. Even today, we are witnessing the 4th industrial revolution which is associated with digital technological tools, seeks to improve productivity, and the capacity of production to meet increasingly stringent customer demand. The objective of this thesis is to trace and define all the new technological tools that shape this 4th industrial revolution and to determine the benefits they could add for the pharmaceutical industrial sector. Secondly, this document seeks to define the challenges to take up in order to successfully implement these tools and obtain autonomous, flexible and efficient factories in the future.

---



**AUTEURS** : MRAD SAMI

**TITRE** : Les outils de l'industrie 4.0 et leurs applications à l'industrie pharmaceutique

**LIEU ET DATE DE SOUTENANCE** : Faculté des Sciences Pharmaceutique, le 27 Janvier 2023

---

**RESUME en français :**

En fonction de l'évolution des modèles économiques et sociaux, ainsi que des innovations technologiques, l'industrie a connu des révolutions successives pour s'adapter à la compétitivité et à la demande de ses consommateurs. Aujourd'hui encore, nous assistons à la 4<sup>ème</sup> révolution industrielle, qui, associé à des outils technologiques digitaux, cherche à améliorer la productivité, et la flexibilité de la production pour répondre à une demande des clients qui est de plus en plus stricte. L'objectif de cette thèse est de retracer et définir l'ensemble des nouveaux outils technologiques qui façonnent cette 4<sup>ème</sup> révolution industrielle et de déterminer les bénéfices qu'ils pourraient ajouter pour le secteur industriel pharmaceutique. Dans un deuxième temps, ce document cherche à définir les défis à relever pour réussir à implémenter ces outils parfaitement.

---

**Titre et résumé en Anglais** : voir au recto de la dernière page de la thèse

---

**DISCIPLINE administrative** : Science Pharmaceutique

---

**MOTS-CLES** : Industrie 4.0, Industrie pharmaceutique du futur, Outils technologiques et numériques, Intelligence Artificielle, Big Data, Digitalisation

---

**INTITULE ET ADRESSE DE L'UFR OU DU LABORATOIRE :**

Faculté des Sciences Pharmaceutiques-Université Paul Sabatier, Toulouse III  
35 chemin des maraîchers  
31400 Toulouse

**Directeur de thèse** : Arellano Cécile