

ACADÉMIE D'ORLÉANS-TOURS

UNIVERSITÉ DE TOURS

FACULTE DE PHARMACIE « Philippe-Maupas »

Année 2020

N° 74

THÈSE D'EXERCICE

Pour le

DIPLÔME D'ÉTAT DE DOCTEUR EN PHARMACIE

Par

HOERDT Aurélien, né le 07/11/1994

PRÉSENTÉE ET SOUTENUE PUBLIQUEMENT LE 16/12/2020

Intelligence Artificielle et Industrie Pharmaceutique :  
De nouveaux défis pour de nouvelles innovations

JURY

Président : **M. GIRAUDEAU Bruno**, Professeur - Faculté de Pharmacie, Tours

Membres :

- **M. MASSOT Victor**, Docteur en Pharmacie, Praticien Hospitalier - Tours
- **M. MUREAU Charles**, Docteur en Pharmacie – SH Consulting, Parçay-Meslay
- **M. PAPON Géraud**, Docteur en Pharmacie – SH Consulting, Parçay-Meslay
- **Mme VIERRON Emilie**, Maître de conférences - Faculté de Pharmacie, Tours

---

(1) Indiquer le millésime de l'année civile au cours de laquelle la thèse a été soutenue

(2) Le numéro de la thèse attribué par la faculté sera porté par le service de la bibliothèque après soutenance

**ANNEE : 2020 - 2021**

**Directrice : Pr Véronique MAUPOIL**

**Directeur Adjoint : M. Hervé MARCHAIS**

**Assesseurs : Pr Daniel ANTIER, M. Matthieu JUSTE, Pr Karine MAHEO, Mme Audrey OUDIN**

## **ENSEIGNANTS**

### **16 PROFESSEURS**

ALLOUCHI	Hassan	CHIMIE PHYSIQUE
ANTIER	Daniel	PHARMACIE CLINIQUE
BRAND	Denys	MICROBIOLOGIE-IMMUNOLOGIE-BIOEPIDEMIOLOGIE
CHEVALIER	Stéphane	BIOCHIMIE GENERALE & BIOTHERAPIE
CHOURPA	Igor	CHIMIE ANALYTIQUE & HYDROLOGIE
CLASTRE	Marc	BIOLOGIE CELLULAIRE & BIOCHIMIE VEGETALE
DIMIER-POISSON	Isabelle	IMMUNOLOGIE PARASITAIRE
EMOND	Patrick	BIOPHYSIQUE & MATHEMATIQUES
ENGUEHARD-GUEIFFIER	Cécile	PHARMACOGNOSIE
GIRAUDEAU	Bruno	BIOPHYSIQUE & MATHEMATIQUES
LANOTTE	Philippe	MICROBIOLOGIE-IMMUNOLOGIE-BIOEPIDEMIOLOGIE
MAHEO	Karine	PHYSIOLOGIE
MAUPOIL-DAVID	Veronique	PHARMACOLOGIE
POUPLARD	Claire	HEMATOLOGIE
THIBAUT	Gilles	MICROBIOLOGIE-IMMUNOLOGIE-BIOEPIDEMIOLOGIE
VIAUD-MASSUARD	Marie-Claude	CHIMIE ORGANIQUE

### **3 PROFESSEURS EMERITES**

AGAFONOV	Viatcheslav	CHIMIE PHYSIQUE
GUILLOTEAU	Denis	BIOPHYSIQUE & MATHEMATIQUES
BARIN	Francis	MICROBIOLOGIE-IMMUNOLOGIE-BIOEPIDEMIOLOGIE

### **38 MAITRES DE CONFERENCES**

ALLARD-VANNIER	Emilie	PHARMACIE GALENIQUE
ARLICOT	Nicolas	BIOPHYSIQUE & MATHEMATIQUES
AUBREY	Nicolas	BIOCHIMIE GENERALE & BIOTHERAPIE
BAKRI	Françoise	HYGIENE SANTE PUBLIQUE & TOXICOLOGIE
BESSON	Pierre	PHYSIOLOGIE
BIRER-WILLIAMS	Caroline	BIOLOGIE CELLULAIRE & BIOCHIMIE VEGETALE
BONNIER	Franck	CHIMIE ANALYTIQUE & HYDROLOGIE
BORDY	Romain	PHARMACOLOGIE
BOUDESOCQUE-DELAJE	Leslie	PHARMACOGNOSIE
BOUVIN-PLEY	Mélanie	MICROBIOLOGIE-IMMUNOLOGIE-BIOEPIDEMIOLOGIE
BRAIBANT	Martine	MICROBIOLOGIE-IMMUNOLOGIE-BIOEPIDEMIOLOGIE
BREDELOUX	Pierre	PHARMACOLOGIE
DAVID	Stéphanie	PHARMACIE GALENIQUE
DEBIERRE-GROCKIEGO	Françoise	IMMUNOLOGIE PARASITAIRE

DELAYE  
 DENEVAULT  
 DOUZIECH-EYROLLES  
 DUMAS  
 GERMON  
 GLEVAREC  
 HERVE-AUBERT  
 JUSTE  
 LAJOIE  
 LANOUE  
 MARC  
 MARCHAIS  
 MAVEL  
 MUNNIER  
 OMBETTA-GOKA  
 OUDIN  
 PASQUALIN  
 PRIE  
 RESPAUD  
 SOUCE  
 TAUBER  
 VELGE-ROUSSEL  
 VERCOUILLIE  
 VERGOTE  
 VIERRON  
 ZHANG

Pierre-Olivier  
 Caroline  
 Laurence  
 Jean-François  
 Stéphanie  
 Gaëlle  
 Katel  
 Matthieu  
 Laurie  
 Arnaud  
 Jillian  
 Hervé  
 Sylvie  
 Emilie  
 Jean-Edouard  
 Audrey  
 Côme  
 Gildas  
 Renaud  
 Martin  
 Clovis  
 Florence  
 Johnny  
 Jackie  
 Emilie  
 Bei-Li

CHIMIE THERAPEUTIQUE  
 CHIMIE THERAPEUTIQUE  
 AFFAIRE REGLEMENTAIRE ET MANAGEMENT DE LA QUALITE  
 BIOCHIMIE GENERALE ET BIOTHERAPIE  
 IMMUNOLOGIE PARASITAIRE  
 BIOLOGIE CELLULAIRE & BIOCHIMIE VEGETALE  
 CHIMIE ANALYTIQUE & HYDROLOGIE  
 IMMUNOLOGIE PARASITAIRE  
 MICROBIOLOGIE-IMMUNOLOGIE-BIOEPIDEMIOLOGIE  
 BIOLOGIE CELLULAIRE & BIOCHIMIE VEGETALE  
 BIOMOLECULES ET BIOTECHNOLOGIES VEGETALES  
 PHARMACIE GALENIQUE  
 CHIMIE THERAPEUTIQUE  
 PHARMACIE GALENIQUE  
 CHIMIE ORGANIQUE  
 BIOLOGIE CELLULAIRE & BIOCHIMIE VEGETALE  
 PHARMACOLOGIE  
 CHIMIE ORGANIQUE  
 CHIMIE ANALYTIQUE & HYDROLOGIE  
 CHIMIE ANALYTIQUE & HYDROLOGIE  
 BIOPHYSIQUE & MATHEMATIQUES  
 IMMUNOLOGIE PARASITAIRE  
 BIOPHYSIQUE & MATHEMATIQUES  
 AFFAIRE REGLEMENTAIRE ET MANAGEMENT DE LA QUALITE  
 BIOPHYSIQUE & MATHEMATIQUES  
 PHARMACOLOGIE

### 1 CONTRAT D'ENSEIGNEMENT

VANIER  
 Antoine  
 BIOPHYSIQUE & MATHEMATIQUES

### 1 DIRECTEUR DE RECHERCHE

CHALON  
 Sylvie  
 INSERM

### 2 CHARGES DE RECHERCHE

MEVELEC  
 MOIRE  
 Marie-Noëlle  
 Nathalie  
 INRA  
 INRA

### 1 PRAG

WALTERS-GALOPIN  
 Susan  
 ANGLAIS

### 1 PAST

JOYEUX  
 VINCENT  
 Filière Pharmacie

### 3 AHU

FOUCAULT  
 FOUCAULT-FRUCHARD  
 MARLET  
 Amélie  
 Laura  
 Julien  
 HEMATOLOGIE  
 PHARMACIE CLINIQUE  
 MICROBIOLOGIE-IMMUNOLOGIE-BIOEPIDEMIOLOGIE



## SERMENT DE GALIEN

*En présence des Maitres de la Faculté, je fais le serment :*

*D'honorer ceux qui m'ont instruit(e) dans les préceptes de mon art et de leur témoigner ma reconnaissance en restant fidèle aux principes qui m'ont été enseignés et d'actualiser mes connaissances ;*

*D'exercer, dans l'intérêt de la santé publique, ma profession avec conscience et de respecter non seulement la législation en vigueur, mais aussi les règles de Déontologie, de l'honneur, de la probité et du désintéressement ;*

*De ne jamais oublier ma responsabilité et mes devoirs envers la personne humaine et sa dignité ;*

*En aucun cas, je ne consentirai à utiliser mes connaissances et mon état pour corrompre les mœurs et favoriser des actes criminels ;*

*De ne dévoiler à personne les secrets qui m'auraient été confiés ou dont j'aurais eu connaissance dans l'exercice de ma profession ;*

*De faire preuve de loyauté et de solidarité envers mes collègues pharmaciens ;*

*De coopérer avec les autres professionnels de santé ;*

*Que les Hommes m'accordent leur estime si je suis fidèle à mes promesses. Que je sois couvert(e) d'opprobre et méprisé(e) de mes confrères si j'y manque.*

Date : 16/12/2020

L'étudiant

M Aurélien HOERDT

Le Doyen de la Faculté

Professeur Véronique Maupoil

# Remerciements

## **A Monsieur Bruno GIRAudeau, Président du Jury,**

Je vous remercie d'avoir accepté de présider ce jury, pour votre accompagnement durant tout ce travail de recherche, de réflexion et de rédaction.

## **A Monsieur Géraud PAPON, mon directeur de thèse industriel,**

Je vous remercie grandement pour la confiance que vous me témoignez depuis mon stage de 5<sup>ème</sup> année à vos côtés. Votre expérience dans le domaine pharmaceutique, votre implication et vos conseils m'ont beaucoup aidé dans la réussite de ce travail. Une très belle rencontre.

## **A Madame Emilie VIERRON, ma directrice de thèse universitaire,**

Je vous remercie d'avoir accepté de m'épauler durant cette année de travail. Votre disponibilité et votre curiosité ont été de précieux alliés au cours de cette année.

## **Aux membres du jury,**

Merci à Victor MASSOT, d'avoir accepté de participer à mon jury de thèse. Je n'oublierai pas notre expérience commune à l'UBCO !

Merci enfin à mon comparse du premier jour, Charles MUREAU, d'avoir complété mon jury. Une amitié à laquelle je tiens énormément.

## **A ma famille,**

Merci Papa et Maman pour tous les sacrifices que vous avez fait pour me permettre de faire des études aussi longues, et d'avoir cru en moi dans tous les projets que j'ai entrepris ! Ma réussite est la vôtre. Merci Estelle pour ton aide et d'être la meilleure sœur qu'un frère pourrait espérer. J'ai de la chance de vous avoir, je vous aime fort.

## **A Inès,**

Merci Inès pour tout, ton soutien, nos moments de complicité. Tu me supportes au quotidien et sans toi je ne serai pas le même. Une nouvelle vie commence pour nous, faite de nouvelles aventures.

## **A mes amis Tourangeaux,**

Merci à Alex, Manon, Dylan, Amandine, Coco, Laurine, Benjamin, Adrien, Sophie, Greg, les autres Kiwis, et enfin tous les Fanfarons de la meilleure fanfare de Pharmacie de France :



l'Atropine, pour ces 5 années de bons moments partagés, ces soirées, ces voyages à travers l'Europe et ces souvenirs qui persisteront toute ma vie !

**A mes amis d'Albi,**

Je remercie également tout particulièrement Emilie, Vince et Ben pour votre soutien dans la rédaction de cette thèse, Mendy, Paul, Fanny, Anaïs, Lisa et Edouard, pour ces deux belles années au bord du Tarn.

**A mes amis Orléanais de toujours,**

Enfin, j'ai une énorme pensée à mes amis de toujours, un peu éloignés par la vie, mais qui ont une place particulière dans mon cœur : JB, Louis, Pierre, Corentin, Aurel, Guigui, Arthur, Camille...

**A l'ASM et à tous ceux que j'oublie.**

# Table des Matières

Table des Matières .....	7
Table des Illustrations .....	11
Liste des Abréviations .....	12
Lexique .....	14
Préambule .....	15
Introduction .....	16
Partie 1 – L’Intelligence Artificielle .....	19
<b>1. Définitions</b> .....	19
1.1. Introduction.....	19
1.2. Définition du Dictionnaire .....	20
1.3. Définition du LEEM .....	20
1.4. Définition scientifique .....	20
<b>2. Histoire de l’IA</b> .....	21
2.1. Début du XXe siècle .....	21
2.2. 1930 – 1956 : Turing et premiers balbutiements de l’IA .....	21
2.2.1. Fonctionnement de la machine Enigma .....	21
2.2.2. Résolution par Turing .....	22
2.2.3. Premiers travaux sur l’IA.....	23
2.3. 1956 – 1974 : L’âge d’or de l’IA .....	24
2.4. Depuis 1987 .....	25
<b>3. Principes essentiels</b> .....	26
3.1. Les bases théoriques de l’IA .....	27
3.1.1. Les algorithmes.....	27
3.1.2. Les réseaux neuronaux artificiels.....	28
3.1.2.1. Présentation .....	28
3.1.2.2. La règle de Hebb .....	29
3.1.2.3. Le Perceptron .....	30
3.1.3. Les réseaux convolutifs .....	30
3.1.4. Les techniques statistiques.....	31
3.2. L’apprentissage de l’IA.....	32
3.2.1. Apprentissage Automatique – <i>Machine Learning</i> .....	32
3.2.2. Apprentissage Profond – <i>Deep Learning</i> .....	35
3.3. Le test de Turing .....	37
<b>4. L’ère du Big Data</b> .....	38
4.1. Définition du Big Data.....	38
4.2. Evolution du volume des Big Data .....	39
4.3. Variété des Big Data .....	41
4.4. Vitesse des Big Data .....	42
4.5. Conclusion .....	42
<b>5. Place de la France et exemples d’application dans la santé</b> .....	43

5.1.	Place de la France dans la recherche en IA .....	43
5.2.	Diabeloop.....	44
5.3.	Synapse Médecine .....	45
Partie 2 - Nouveaux défis et Limites de l'IA dans l'Industrie Pharmaceutique .....		47
<b>1.</b>	<b>Point de vue Gouvernemental et Législatif .....</b>	<b>47</b>
1.1.	Point de vue Législatif .....	47
1.1.1.	En France.....	47
1.1.2.	Dans le Monde.....	48
1.2.	Établir des responsabilités.....	49
1.3.	Propriété intellectuelle et Brevets .....	50
1.3.1.	Définitions .....	50
1.3.1.1.	La Propriété Intellectuelle .....	50
1.3.1.2.	Les Brevets.....	50
1.3.2.	Considérations Légales .....	51
1.3.2.1.	IA, Innovation et Collaboration.....	51
1.3.2.2.	La protection des Algorithmes et autres solutions d'IA .....	52
1.3.2.3.	L'adaptation de la Législation.....	53
1.4.	Limites .....	53
1.4.1.	Inégalités entre les différents pays.....	53
1.4.2.	Aggravation de la Fracture Numérique entre les pays .....	54
1.4.3.	Difficultés dans l'établissement de normes ou de règles .....	54
<b>2.</b>	<b>Point de vue Ethique.....</b>	<b>54</b>
2.1.	Définition.....	54
2.2.	Ethique & IA.....	54
2.3.	Augmenter la transparence et l'auditabilité des systèmes.....	55
2.4.	Démocratiser le débat sur les usages de l'IA .....	56
2.5.	Le Règlement Général sur la Protection des Données (RGPD) .....	56
2.5.1.	Définition et périmètre d'application.....	56
2.5.1.1.	Définitions.....	56
2.5.1.2.	Périmètre d'application .....	57
2.5.2.	Les modifications pour les industriels .....	57
2.5.3.	Les sanctions en cas de non-conformité .....	57
2.6.	Limites .....	58
<b>3.</b>	<b>Point de vue Organisationnel .....</b>	<b>58</b>
3.1.	Stratégie de l'entreprise .....	58
3.1.1.	Agilité de l'entreprise .....	59
3.1.2.	Collaboration .....	60
3.1.3.	Challenges liés à l'intégration de l'IA dans les entreprises et gestion des risques .....	61
3.2.	Connectivité universelle.....	62
3.3.	Gestion des coûts .....	62
3.4.	Qualification & Validation.....	62
3.5.	Limites .....	64
3.5.1.	Manque de connaissances sur le sujet.....	64
3.5.2.	Manque d'infrastructures .....	64



3.5.3.	Manque d'alignement des entreprises.....	64
<b>4.</b>	<b>Point de vue Humain .....</b>	<b>65</b>
4.1.	Etat des lieux de l'industrie pharmaceutique en France.....	65
4.2.	Le monde du Travail.....	65
4.2.1.	Mutation des Organisation et Gestion du Changement.....	65
4.2.2.	L'Automatisation.....	66
4.2.3.	Les Emplois .....	66
4.2.3.1.	Les suppressions d'emploi et licenciements.....	67
4.2.3.2.	Les créations et modifications d'emploi.....	67
4.2.3.2.1.	Problématiques liées aux modifications d'emplois .....	67
4.2.3.2.2.	La mutation des compétences .....	68
4.3.	Coopération entre la force de Travail et l'IA .....	69
4.4.	La compréhension des techniques.....	69
4.5.	L'accompagnement et la formation .....	69
4.5.1.	La Gestion prévisionnelle de l'Emploi et des Compétences .....	69
4.5.2.	La Formation .....	70
4.5.2.1.	Formation des Managers .....	70
4.5.2.2.	Formation des Employés.....	71
4.5.2.3.	Formation des Étudiants.....	72
4.6.	Limites .....	72
4.6.1.	Point de vue du grand public.....	72
4.6.2.	La rareté des experts .....	73
<b>5.</b>	<b>Point de vue Technique .....</b>	<b>74</b>
5.1.	La Gestion des données .....	74
5.1.1.	Les données utilisées dans l'Industrie Pharmaceutique .....	74
5.1.2.	La cybersécurité et protection des données.....	74
5.1.3.	La qualité et quantité des données .....	75
5.1.3.1.	Qualité des Données.....	75
5.1.3.2.	Quantité des Données.....	76
5.1.4.	L'accès aux données .....	76
5.2.	Explication des résultats .....	76
5.3.	Limites techniques .....	77
5.3.1.	L'accès aux données.....	77
5.3.2.	La qualité des données.....	77
5.3.3.	La compréhension des techniques .....	77
5.3.4.	Les biais et erreurs de modèle.....	78
5.3.5.	Peut-on tout automatiser ? .....	78
<b>6.</b>	<b>Point de vue Environnemental.....</b>	<b>78</b>
6.1.	Impact de l'IA sur l'environnement .....	78
6.2.	Réduire les impacts environnementaux des médicaments.....	79
Partie 3 - Exemples d'utilisations actuelles et futures de l'IA dans l'industrie pharmaceutique.....		80
<b>1.</b>	<b>Introduction - Cycle de vie du Médicament .....</b>	<b>80</b>
<b>2.</b>	<b>Intelligence Artificielle dans la Recherche &amp; Développement .....</b>	<b>81</b>
2.1.	Définition de la Recherche & Développement .....	81

2.2.	Projets en cours de déploiement .....	82
2.2.1.	En recherche fondamentale - Recherche de la cible moléculaire .....	82
2.2.1.1.	Etude bibliographique .....	82
2.2.1.2.	Identification des cibles moléculaires .....	84
2.2.1.3.	Prédiction de la structure tridimensionnelle d'une cible d'intérêt .....	86
2.2.2.	En recherche pharmaceutique - Criblage de molécules .....	86
2.2.2.1.	Criblage de molécules d'intérêt .....	86
2.2.2.2.	Prédiction de caractéristiques d'une molécule .....	89
2.2.2.3.	Synthèse de la molécule .....	90
2.2.2.4.	Recyclage de la molécule .....	91
2.2.3.	Sécuriser la recherche .....	92
<b>3.</b>	<b>Intelligence Artificielle dans la Recherche Clinique .....</b>	<b>92</b>
3.1.	Définition de la Recherche Clinique .....	92
3.2.	Projets en cours de déploiement .....	93
3.2.1.	Identification & recrutement des patients .....	93
3.2.2.	Amélioration de l'observance des patients pendant les essais cliniques .....	94
3.2.2.1.	AiCure .....	94
3.2.2.2.	Catalia Health – Mabu .....	96
3.2.2.3.	Sensely .....	96
3.2.3.	Sécuriser les données issues des essais cliniques .....	97
<b>4.</b>	<b>Intelligence Artificielle dans la Production .....</b>	<b>97</b>
4.1.	L'IA et la Production Pharmaceutique (fabrication / conditionnement) .....	98
4.1.1.	L'optimisation des processus .....	98
4.1.2.	Le contrôle des paramètres de production en temps réel .....	99
4.2.	L'IA et le Contrôle Qualité des Produits .....	100
4.2.1.	Définition & problématiques actuelles .....	100
4.2.2.	Contrôles en cours de production et des produits finis .....	100
4.3.	La Gestion des Risques .....	101
<b>5.</b>	<b>Intelligence Artificielle dans la Logistique Pharmaceutique .....</b>	<b>101</b>
5.1.	La logistique pharmaceutique .....	101
5.2.	Gestion des stocks et des commandes .....	102
5.3.	Optimisation des flux (livraisons, flux internes) .....	104
5.3.1.	Automatisation du transfert des produits dans l'entrepôt .....	104
5.3.2.	Sélection des itinéraires et optimisation des déplacements .....	104
<b>6.</b>	<b>Autres applications de l'Intelligence Artificielle .....</b>	<b>105</b>
6.1.	Pharmacovigilance .....	105
6.2.	Biotechnologies .....	106
6.3.	Marketing .....	107
	Conclusion .....	108
	Bibliographie .....	110

# Table des Illustrations

Tableau 1- Comparaison entre l'industrie traditionnelle et l'industrie 4.0 .....	17
Figure 1 - Reconstitution du chariot inventée par Léonard de Vinci, Château du Clos Lucé, Amboise, 2001 (13) .....	21
Figure 2 - Machine Enigma allemande (16) .....	21
Figure 3 - Schéma de fonctionnement d'une machine Enigma (17) .....	22
Figure 4 - Principe de la machine de Turing (18) .....	23
Figure 5 - Schéma de fonctionnement des Agents Intelligents (26).....	26
Figure 6 -Représentation d'un réseau neuronal artificiel (33) .....	28
Figure 7 - Schéma du mécanisme d'activation d'un neurone artificiel (34).....	29
Figure 8 - Schéma de fonctionnement d'un perceptron (38) .....	30
Figure 9 - Schéma d'un réseau convolutif multicouche (42) .....	31
Figure 10 - Lien entre l'IA et l'Apprentissage (44) .....	32
Figure 11 - Principe de fonctionnement de l'Apprentissage Automatique ou Machine Learning (44) .....	33
Figure 12 - Evolution de la teneur en PA d'un médicament en fonction de la force de mélangeage .....	33
Figure 13 - Extrapolation de données à une situation inconnue.....	34
Figure 14 - Principe de fonctionnement de l'apprentissage supervisé (45) .....	34
Figure 15 - Principe de fonctionnement de l'apprentissage non supervisé (45).....	35
Figure 16 - Principe de fonctionnement de l'apprentissage par renforcement (46).....	35
Figure 17 - Principe de fonctionnement de l'Apprentissage Profond ou Deep Learning (44) .....	36
Figure 18 - Exemple d'application d'apprentissage profond sur de la reconnaissance d'images (48) .....	36
Figure 19 - Schéma du test de Turing (52) .....	37
Figure 20 - Evolution du volume de la Datasphere entre 2010 et 2025 (57) .....	39
Figure 21 - Evolution du nombre d'objets connectés entre 2007 et 2025 (61) .....	40
Figure 22 - Comparaison des vitesses de téléchargement en fonction des technologies (63).....	40
Figure 23 - Evolution du prix d'un capteur depuis 2005 (64).....	41
Figure 24 - Prévisions de la croissance du marché de l'e-santé en France (milliards d'euros)(68).....	44
Figure 25 - Schéma de fonctionnement de Diabeloop (5) .....	45
Figure 26 - Module d'analyse d'ordonnances (78) .....	46
Figure 27- Activités liées aux technologies de l'IA dans l'industrie pharmaceutique (102) .....	58
Figure 28 - Schéma de l'agilité d'une entreprise (104).....	59
Figure 29 – Evolution du nombre de partenariats et leur valeur (106) .....	61
Figure 30 - Evolution des employés en entreprises du médicament entre 1997 et 2017 (109) .....	65
Figure 31- Cycle de vie du Médicament simplifié (132).....	80
Figure 32 - Evolution du retour sur investissement dans la R&D pharmaceutique depuis les années 1990 (134) .....	81
Figure 33 - Principe de DeepDTnet (146) .....	85
Figure 34 - Différentes structures des protéines (147) .....	86
Figure 35 - Exemple de la retranscription d'une molécule avec un graphe réduit moléculaire (150) .....	87
Figure 36 - Exemple d'une voie de rétrosynthèse et d'un arbre de recherche (155) .....	91
Figure 37 - Schéma des différentes phases de la recherche clinique (159).....	92
Figure 38 - Interface de Mendel.ai (162).....	93
Figure 39 - Concept d'Industrial Data Lake (173).....	98

# Liste des Abréviations

AIPD : Analyses d'Impact relatives à la Protection des Données

AMM : Autorisation de Mise sur le Marché

ANSM : Agence Nationale de sécurité du Médicament et des produits de santé

ASMR : Amélioration du Service Médical Rendu

BPF : Bonnes Pratiques de Fabrication

CA : Chiffre d'Affaires

CCNE : Comité Consultatif National d'Ethique pour les sciences de la vie et de la santé

CE : Conformité Européenne

CESE : Comité Economique et Social Européen

CNEDiMTS : Commission Nationale d'Evaluation des Dispositifs Médicaux et des Technologies de Santé

COMEST : Commission Mondiale d'Ethique des Connaissances Scientifiques et des Technologies

CRM : Customer Relationship Management

DARPA : Defense Advanced Research Projects Agency

DMP / DPP : Dossier Médical /Patient Partagé

EC : Essai Clinique

EI : Effet Indésirable

ERP : Enterprise Ressource Planning

FDA : Food and Drug Administration

GPEC : Gestion prévisionnelle de l'Emploi et des Compétences

HAS : Haute Autorité de la Santé

IA : Intelligence Artificielle

INPI : Institut National de la Propriété Industrielle

INSEE : Institut National de la Statistique et des Etudes Economiques

LEEM : Les Entreprises du Médicament

MIT : Massachusetts Institute of Technology

OCDE : Organisation de Coopération et de Développement Economiques

OMS : Organisation Mondiale de la Santé

PA : Principe Actif

PI : Propriété Intellectuelle

PIB : Produit Intérieur Brut

PV : Pharmacovigilance

R&D : Recherche et le Développement

RGPD : Règlement Général sur la Protection des Données

RSE : Responsabilité Sociale des Entreprises

SMR : Service Médical Rendu

TAP : Technologies Analytiques des Procédés

TIC : Technologies de l'Information et de la Communication

UNESCO : Organisation des Nations unies pour l'éducation, la science et la culture

# Lexique

**Apprentissage automatique** : apprentissage réalisé par un système à partir d'une base de données, menant à l'extraction de règles générales appliquées ensuite à d'autres jeux de données afin de les analyser plus efficacement.

**Connexionnisme** : Modélisation des phénomènes mentaux ou comportementaux comme des processus émergents de réseaux d'unités simples interconnectées (1).

**CRM** : Systèmes aidant les entreprises dans la gestion de leur relation avec la clientèle (prise d'information, résolution des problèmes, etc...).

**ERP** : Logiciel permettant de gérer différents processus au sein d'une entreprise (prise de commande, gestion des stocks, monitoring de la production, etc...).



# Préambule

Dans la multitude de sujets intéressants que le monde pharmaceutique propose, j'ai décidé de réfléchir sur la place actuelle et à venir de l'Intelligence Artificielle (IA) dans l'industrie pharmaceutique.

L'Intelligence Artificielle est apparue il y a bien longtemps, mais son implication dans le secteur de la santé ne cesse de grandir depuis plus de 20 ans. Elle est un vecteur d'espoir dans de nombreux domaines : du diagnostic à l'accompagnement du patient dans son parcours de soin en passant par les robots effectuant des opérations complexes, etc... J'ai toujours eu une affection toute particulière pour les nouvelles technologies et le milieu de la santé, dans lequel je baigne depuis mon enfance. C'est ce qui m'a poussé à y travailler et à m'orienter vers l'industrie pharmaceutique.

En tant que futur pharmacien-ingénieur, ce choix de sujet me semble très pertinent, car il me permet de faire le lien entre ma formation en **Sciences Pharmaceutiques**, afin d'appréhender l'Intelligence Artificielle dans son versant biologique et ses applications, avec celle **d'Ingénieur Généraliste**, afin de comprendre les mécanismes qui sous-tendent ces phénomènes, futur acteur majeur de notre quotidien tant professionnel que personnel.

Le but de cette thèse est de distinguer la fiction de la réalité, afin d'imaginer ce que pourra être l'industrie pharmaceutique de demain. Les thèmes, projets, et perspectives ne sont pas exhaustifs, certains projets sont volontairement mis de côté pour ne traiter que les plus pertinents.

Je vous souhaite une excellente lecture.

Aurélien

# Introduction

## *Pharma 4.0*

L'**Industrie 4.0** ou encore industrie du futur est apparue en Allemagne en 2012. Elle correspond à une nouvelle philosophie pour l'organisation des moyens de production autour des innovations liées aux objets connectés (*IoT - Internet of Things*) et aux technologies numériques (robots collaboratifs, réalité augmentée, ou l'IA notamment). L'essor de ces technologies a fait naître le concept de **système Cyber-physique** (CPS pour *Cyber-physical Systems*), défini comme 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*" (2).

Ce système extrait donc des informations de son environnement, qu'il traite, afin de contrôler et piloter des processus physiques et ainsi les modifier. Ces technologies sont amenées à être coordonnées au sein d'usines intelligentes ("*Smart Factories*") mais également entre elles, afin d'augmenter la flexibilité et l'efficacité des processus industriels en limitant l'intervention humaine due aux maintenances, réparations, ou réglages des machines (3).

L'intégration de ces nouvelles technologies au sein de l'industrie pharmaceutique a donné naissance à la **Pharma 4.0**. Elle se base sur plusieurs concepts :

- **Numérisation de l'entreprise**

C'est la convergence de la conception numérique, du monde virtuel avec les produits et objets du monde réel. Elle intervient tout au long de la chaîne de production : conception du produit, contrôle et pilotage (capteurs, etc...), fabrication et maintenance...

- **Modélisation et simulation des procédés**

La numération de l'entreprise et l'acquisition des données issues de la chaîne de production permettent de la modéliser virtuellement, à partir de quoi des simulations de tests ou de processus sont réalisables. Ces simulations sont de plus en plus fidèles et sont d'une grande aide dans les processus de prise de décision.

- **Flexibilisation des usines et personnalisation des produits**

L'utilisation d'objets intelligents précédemment évoqués va permettre d'accroître la flexibilité de la production : les produits fabriqués pourront être personnalisés à souhait par le client. On passe dès lors d'une production de masse à une personnalisation de masse, ce qui va profondément changer l'usine et toute la chaîne d'approvisionnement.

- **Big Data et flux logistiques**

De très grandes quantités de données vont être générées par les nombreux objets connectés à Internet (capteurs, équipements, opérateurs...), qu'il faudra savoir analyser et pouvoir échanger le plus vite possible. C'est une des priorités de l'industrie de demain.

- **Economies d'énergie et de ressources**

L'industrie 4.0 n'est pas épargnée par les problématiques de gestion des ressources, de l'énergie et des déchets. Les échanges immédiats et les données en temps réel permettent de coordonner de façon plus efficace les besoins réels actuels, les disponibilités des éléments et d'adapter les commandes en conséquence. Le surstock est limité, et les commandes des fournisseurs sont optimisées, ce qui limite les transports et les flux de marchandises et donc *in fine* le coût et la consommation énergétique (4).

- **Relation avec la clientèle**

La relation avec le client est un des enjeux majeurs de l'industrie digitale : transparence des processus et suivi de commande en temps réel, de centralisation des informations ou interopérabilité.

L'objectif final de cette industrie, est de prendre des décisions plus rapidement et plus efficacement, d'avoir des processus réactifs, capables d'identifier instantanément les alertes, les problèmes, ou les risques pouvant perturber les opérations. Une fois identifiée, une décision ou une action pourra être mise en place pour corriger cette situation de façon plus efficace et rapide.

Les principales différences entre l'industrie traditionnelle et l'industrie 4.0 sont présentées dans le tableau ci-dessous :

*Tableau 1- Comparaison entre l'industrie traditionnelle et l'industrie 4.0*

	<b>Procédés de Production "traditionnels"</b>	<b>Industrie 4.0</b>
Processus	Manuels et Rigides	Automatisée et agile
Produit	Standardisé	Personnalisé et customisé
Echelle des usines	Grandes usines dans des lieux centralisés	Petites usines dans des lieux décentralisés
Supply Chain	Planning basé sur les stocks	Dynamique et prédictive
Mesure de la réussite	Haute efficacité, bas prix	Rentabilité des capitaux investis
Relation avec la clientèle	Faible et indirecte	Forte et directe

On attribue donc à l'industrie connectée trois principaux avantages :

- Augmenter la compétitivité de l'entreprise (et prendre un avantage vis à vis de la concurrence) ;
- Augmenter la flexibilité de production ;
- Et permettre au client de suivre l'ensemble du processus de fabrication de son produit en temps réel.

On parle beaucoup de l'aspect financier de cette industrie de demain, mais l'impact social est très important également : quelle sera la place de l'humain dans cette industrie ?

Bien qu'à l'heure actuelle la Pharma 4.0 ne soit pas encore un standard, les perspectives qu'elle offre sont très nombreuses. Les nombreuses avancées technologiques - l'intelligence artificielle notamment - accompagnant le développement de l'industrie 4.0 peuvent être un tremplin pour cette industrie, dont le rôle dans notre société n'est que plus grand de jour en jour (5).

C'est justement ce que cette thèse va traiter :

*Quels sont les enjeux de l'IA dans l'industrie pharmaceutique et les applications possibles ?*

- La première partie sera axée sur l'Intelligence Artificielle son histoire et les personnages clés, et énoncera les principes essentiels la régissant. Quelques exemples d'utilisations actuelles de l'IA dans le domaine de la santé seront ensuite présentés, afin d'avoir une idée des applications possibles.
- La seconde partie abordera les défis qui accompagnent le développement de l'IA dans l'industrie pharmaceutique : adaptation des lois, considérations éthiques, modifications des conditions de travail, et les modifications tant organisationnelles que légales qui seront nécessaires.
- La dernière partie présentera des exemples de projets d'utilisation de l'IA dans différents secteurs de l'industrie et à différents stades du cycle de vie du médicament.

# Partie 1 – L'Intelligence Artificielle

## 1. Définitions

### 1.1. Introduction

Le grand public connaît de l'Intelligence Artificielle (IA) ce que les œuvres (notamment cinématographiques) du vingtième siècle lui ont présenté : les robots intelligents Skynet de *Terminator* cherchant à éliminer toute menace les empêchant d'asseoir leur supériorité sur les humains, les machines de *Matrix* qui, après s'être rebellées, utilisent l'espèce humaine comme source d'énergie, ou encore l'antagoniste de *2001 : L'Odyssée de L'espace*, HAL 9000 qui mène son équipage à sa perte...

L'IA semble correspondre à des entités dangereuses ayant pris conscience de leur nature et souhaitant anéantir l'espèce humaine, pour des raisons obscures. Cela n'est que de la fiction. L'IA nourrit les fantasmes des individus les plus pessimistes, et beaucoup lui attribuent des traits qui ne lui correspondent pas.

C'est pourquoi, avant de définir le concept d'IA, nourri par les lieux communs des œuvres de fiction, il est important de préciser ce que l'IA n'est pas :

L'IA n'est pas synonyme de "robot" : certains robots fonctionnent sans IA et les applications de l'IA ne se résument pas aux robots (6).

L'IA n'est pas "vivante" : Bien qu'elle soit composée de programme la dotant d'une certaine intelligence, elle reste irrémédiablement tributaire de son concepteur qui détermine ses capacités et fonctions.

L'Homme est un être doué de raison, de conscience vis à vis de son environnement et des personnes qui l'entourent, et il est très complexe. L'imiter dans son ensemble est impossible comme l'explique Luc Julia (ingénieur français spécialisé dans l'IA) « *C'est en comparant l'IA à l'intelligence humaine qu'on s'aperçoit à quel point l'IA se concentre sur un domaine d'activité et néglige le vécu, la sensibilité, l'assimilation d'expériences, en un mot, la multidisciplinarité.* » (7). Cette machine n'a aucune notion spontanée de conscience : si un tel "sentiment" existe chez elle, cela ne provient que d'un programme implanté, donc décidé en amont par l'humain. Le robot raisonne uniquement en fonction des données auxquelles il a eu accès, sans vraiment comprendre ce qu'il fait.

L'IA n'est pas non plus autonome à l'heure actuelle. Les programmes ont toujours besoin de l'Homme en amont pour fonctionner.

L'IA n'est pas la solution à tous les problèmes que connaît le monde d'aujourd'hui : un robot ne pourra pas remplacer un humain dans toutes ses tâches du quotidien ;

L'IA n'est pas non plus quelque chose dont il faut avoir peur. Sa finalité n'est pas de mettre tout le monde au chômage. Non, même si des métiers seront très sûrement

supprimés, de nouveaux vont naître simultanément, comme cela a toujours été le cas, au fil des révolutions technologiques.

Dans la société actuelle et malgré les progrès scientifiques, il a été démontré que les capacités de l'IA aujourd'hui ne sont pas celles vendues par les fictions littéraires ou cinématographiques connues du grand public. Son essence est beaucoup plus complexe et la place de l'Homme est indéniable. Après la présentation de ses écueils, nous pouvons nous pencher sur la définition propre de l'IA.

## 1.2. Définition du Dictionnaire

L'IA, prise dans son sens strict, repose sur un oxymore : l'*intelligence* est propre à l'être Humain, tandis qu'est *artificiel* tout ce qui ne l'est pas. Le dictionnaire Larousse définit l'IA comme « *l'Ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine* » (8). L'homme touche ici du doigt l'un de ses rêves les plus ambitieux : concevoir des outils équipés d'un « esprit » semblable au sien.

C'est cette ambition qui explique pourquoi il faudra à la fois s'intéresser aux progrès purement technologiques et techniques de la machine, mais également à la meilleure compréhension du fonctionnement biologique de l'intelligence humaine afin de la "traduire" au mieux.

## 1.3. Définition du LEEM

Définir précisément l'IA est très complexe. Selon le site du LEEM (Les Entreprises du Médicament), « *l'IA désigne moins un champ de recherche bien défini qu'un programme, fondé autour d'un objectif ambitieux : comprendre comment fonctionne la cognition humaine et la reproduire ; créer des processus cognitifs comparables à ceux de l'être humain* » (9). Son champ d'application est extrêmement vaste, puisqu'elle fait appel à des techniques et disciplines très diverses (mathématiques, sciences cognitives, etc...), et ses méthodes sont très variées.

## 1.4. Définition scientifique

C'est le scientifique américain John McCarthy qui a été le premier à utiliser le terme "d'Intelligence Artificielle", comme la science de "*rendre des machines intelligentes*" (10). L'un de ses créateurs, Marvin Lee Minsky, la définit comme « *la construction de programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains car elles demandent des processus mentaux de haut niveau tels que : l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique* »(11).

La prochaine partie traitera des grandes dates ayant abouti à l'IA que l'on connaît aujourd'hui.



## 2. Histoire de l'IA

Au départ, les premières machines étaient de simples automates, entièrement codées par l'Homme : celles-ci ne faisaient que les tâches élémentaires (*calculs notamment*) programmées par son créateur, ni plus ni moins. Ces modèles ont été perfectionnés au fil des siècles.

La première évocation de l'IA dans l'Histoire ne date pas d'hier, bien que celle-ci ne mentionne pas ce terme exactement. Raymond Lulle, prétendait par exemple au 14<sup>ème</sup> siècle avec son **Ars Magna** (Le Grand Art en latin) pouvoir démontrer automatiquement la véracité ou l'inexactitude d'un postulat(12). Léonard de Vinci a également apporté sa pierre à l'édifice, en imaginant des automates au 15<sup>e</sup> siècle, dont on ne sait pas s'ils ont en effet été utilisés : le chariot programmable (*cf. Figure 1*) dont les trajets semblaient pouvoir être automatisés ou l'automate chevalier, capable de s'asseoir ou de bouger ses membres notamment(13).



Figure 1 - Reconstitution du chariot inventée par Léonard de Vinci, Château du Clos Lucé, Amboise, 2001 (13)

### 2.1. Début du XX<sup>e</sup> siècle

Les grands progrès dans le domaine de l'IA ont surtout été rendus possibles grâce aux travaux réalisés au début du vingtième siècle par de nombreux scientifiques.

Russell et Whitehead ont par exemple suggéré, dans leur chef d'œuvre *Principia Mathematica* en 1913, que toute forme de raisonnement mathématique pouvait être mécanisée(14). Alonzo Church lui, stipule dans sa thèse, qu'un appareil mécanique manipulant des symboles tels que des 0 et des 1 pouvait aisément imiter tout processus concevable de déduction mathématique, précurseur du langage informatique actuel.

### 2.2. 1930 – 1956 : Turing et premiers balbutiements de l'IA

La première utilisation de l'IA à grande échelle a eu lieu durant le plus grand conflit de l'Histoire, la seconde Guerre Mondiale. Le britannique Alan Turing a permis, grâce à sa création décodant la machine Enigma des Allemands, de changer le cours de la Guerre.

#### 2.2.1. Fonctionnement de la machine Enigma

Les machines Enigma (*cf. Figure 2*) représentent une famille de machines utilisées par l'armée allemande au cours de la Seconde Guerre Mondiale (1939 - 1945) afin de crypter leurs communications. Ces machines étaient réputées inviolables : en effet, même si les messages étaient interceptés par l'ennemi, ils étaient illisibles à leurs yeux(15).



Figure 2 - Machine Enigma allemande (16)

Le fonctionnement de cette machine est très simple et astucieux (cf. Figure 3). Elle est composée d'un clavier physique, et d'un autre uniquement composé de lampes. Comme tout algorithme de cryptage, une lettre est substituée par une autre (si on appuie sur le D, la lettre K s'affiche sur le clavier supérieur par exemple)(16).

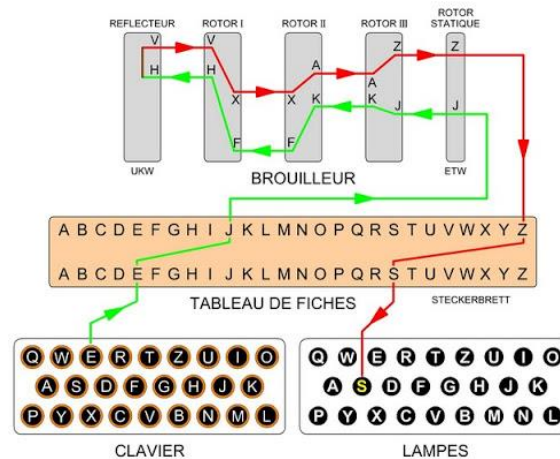


Figure 3 - Schéma de fonctionnement d'une machine Enigma (17)

La différence ici, et principal atout de cette machine, est que la substitution varie d'une lettre à une autre. Le codage serait bien évidemment trop aisé si la même lettre s'illuminait à chaque fois que l'on utilise une lettre. C'est pour cela qu'à chaque fois qu'une lettre est entrée au clavier, au moins l'un des rotors (*pièces circulaires rotatives, au nombre de 4*) tourne, si bien que le circuit électrique change et aboutit à une lettre différente.

Il a été calculé qu'il existe plus de  $10^{16}$  possibilités de permutation pour une seule lettre de cette façon, soit autant de configurations différentes de cette machine.

*Ainsi, pour réussir à décrypter un message codé par la force brute (tester toutes les possibilités), il faut faire tourner un ordinateur actuel pendant un an !*

### 2.2.2. Résolution par Turing

Le déchiffrement du code repose sur la connaissance de la **configuration initiale** de la machine par le destinataire : en effet, cette configuration permet à la fois le chiffage et le déchiffrement des messages. Chaque jour, à minuit, la configuration initiale de l'ensemble des machines Enigma allemandes était réinitialisée (une seule et même configuration était utilisée par l'ensemble des machines), et les permutations étaient ainsi modifiées. C'est ce qui a permis de percer le mystère de cette machine, et de changer le cours de la Guerre.

Les premiers à avoir tenté de déchiffrer cette machine furent les polonais, qui ont remarqué une redondance de la clé de codage au début de chaque transmission, ce qui facilitait le déchiffrement (qui se faisait au départ à la main). Après l'invasion du territoire polonais, ce sont d'autres pays, dont les Royaume-Unis, qui ont poursuivi les recherches, avec Alan Turing à leur tête.

Turing a conçu une **bombe cryptologique**. Pour schématiser, cela correspondait à un ensemble de machines Enigma fonctionnant simultanément afin d'analyser de nombreuses configurations possibles de la machine à une très grande vitesse. Ces essais consistaient à effectuer des tests de boucles. En effet, les permutations sont réversibles : un remplacement de la lettre P par T équivaut à celle de T par P. Chaque configuration testée permet de visualiser si l'on obtient bien une boucle ou non, si cela n'est pas le cas, la bombe étudie la configuration suivante, jusqu'à trouver la bonne(15).

Grâce à cette méthode, la clef de chiffrage était désormais identifiée en vingt minutes (*au lieu de 5 milliards d'années si le déchiffrement devait se faire à la main*). Une fois cette clé identifiée, il suffit de paramétrer la machine et de retranscrire le message tel qu'il a été intercepté. De nombreux historiens estiment que le déchiffrement de la machine Enigma et donc des communications allemandes a permis d'écourter la Guerre de près de deux ans. Le déchiffrement de la machine Enigma est donc l'une des premières utilisations de l'IA (utilisation de tests de solutions réalisables par tâtonnement, de boucles et d'algorithmes...).

Avant que l'IA ne soit conceptuellement définie en 1956, des premiers travaux concrets ont posé les bases de cet outil. Alan Turing a créé le concept de **Machine de Turing**. Cet outil mathématique est l'ancêtre des ordinateurs que nous connaissons aujourd'hui et a défini le concept d'algorithme.

Un ruban était divisé en une multitude de cases successives, dans lesquelles étaient inscrits des symboles définis. Sur ce ruban se déplaçait une tête de lecture vers la gauche ou la droite. L'état courant de la machine est mémorisé dans un registre d'état, et une table d'actions permet d'indiquer à la machine les actions qu'elle doit réaliser selon les symboles qu'elle lisait dans les cases du ruban (sens de déplacement, symboles à écrire, etc...)(18).



Alan Turing a également inventé en 1950 le **Test de Turing**, dont l'objectif est de faire la distinction entre l'Homme et la Machine (*ce test sera détaillé ultérieurement*).

Toujours au cours des années 50, les scientifiques Walter Pitts et Warren McCulloch présentent le premier travail sur les **réseaux neuronaux** dans l'article « *What the frog's eye tells the frog's brain* ». Ils stipulent que les “yeux communiquent avec le cerveau dans un langage déjà très organisé, bien que non interprété, au lieu de ne transmettre qu'une copie plus ou moins fidèle de la distribution de la lumière sur les récepteurs”(19).

C'est à la suite de ce travail qu'ils créèrent la notion de **neurone formel**, qui sera la base d'une notion détaillée plus tard : les réseaux de neurones (ou connexionnisme).

### 2.3. 1956 – 1974 : L'âge d'or de l'IA

La naissance de l'IA telle qu'on la connaît de nos jours remonte à la **conférence de Dartmouth** (*Dartmouth Summer Research Project on Artificial Intelligence*), qui a eu lieu à l'été 1956, aux Etats-Unis, sous l'impulsion de Marvin Minsky et de John McCarthy. C'est au cours de cette conférence que l'IA sera définie, et que ses objectifs seront précisés : ils estimaient que “chaque aspect de l'apprentissage ou toute autre caractéristique de l'intelligence peut être si précisément décrit qu'une machine peut être conçue pour le simuler”(20). Leur objectif principal était donc simple : la simulation, sur des machines, de chacune des différentes facultés de l'intelligence humaine.

C'est par la suite en outre Atlantique que cette discipline s'y développera, en particulier au “Massachusetts Institute of Technology” (MIT), qui reçoit de nombreuses subventions de la DARPA (“Defense Advanced Research Projects Agency”) pour financer ses recherches. De nombreuses découvertes primordiales ont été réalisées au cours de cette période, dont :

- **Le raisonnement par tâtonnements**

La réflexion se fait pas à pas jusqu'à la solution (en effectuant un mouvement ou une déduction à la fois), en revenant en arrière dès que l'on se heurte à une impasse.

- **Le langage naturel**

Le premier agent conversationnel, ELIZA, a été créé entre 1964 et 1966. Celui-ci était capable de mener des conversations si réalistes que certains utilisateurs se sont laissé abuser en croyant communiquer avec un être humain et non un programme. C'est ce qui a été à la base de la fondation du Traitement Automatique du Langage Naturel (TALN), discipline étudiant la capacité qu'ont certains programmes de comprendre et d'analyser le langage humain.

- **La Loi de Moore**

Gordon Moore, directeur de la Recherche et du Développement à Fairchild Semiconductor observe, en 1965, que la puissance de calcul des processeurs double tous les deux ans à prix

constants, témoignant une avancée technologique essentielle. Cette loi est toujours d'actualité à l'heure actuelle pour expliquer les révolutions technologiques(21).

L'apparition des **capteurs** vers la fin des années 60 a permis l'apparition de la cybernétique, c'est-à-dire la *“Science qui étudie les mécanismes de communication et de régulation dans les machines et chez les êtres vivants”*(22). Les robots peuvent désormais interagir avec leur environnement.

L'année 1967 marque la première victoire d'un programme d'IA sur un Homme au jeu d'échec : le Max Hack 6 de Richard Greenblatt.

Cette période fait naître un très grand optimisme chez les scientifiques : *“des machines seront capables, d'ici vingt ans, de faire tout travail que l'homme peut faire”* disait Herbert Simon, en 1965(23). Marvin Minsky avouera plus tard que cet objectif n'a pas été atteint : *“les choses les plus simples à faire par un humain se sont avérées extrêmement complexes à reproduire sur une machine, tandis que des choses difficiles pour un humain (comme résoudre un problème d'échec) se sont avérées très simples”*(24).

Cependant, après cette longue période de découvertes et de progrès, survint une période plus compliquée pour l'IA, où très peu de projets aboutissent : *l'AI Winter*, ou la traversée de l'hiver de l'IA (entre 1974 et 1980)(12).

De nombreux projets sont abandonnés par manque de technologie, conduisant à de grosses pertes d'argent. De nombreuses critiques relatives à ce manque de progrès émanent également de la communauté scientifique et le grand public commence à avoir une vision négative de cette technologie. L'optimisme des chercheurs est mis à mal par l'arrêt des subventions de recherche fondamentale en IA par les gouvernements américains et britanniques par manque de progrès.

Un regain d'intérêt a eu lieu dans les années 80, et des projets en hibernation depuis de nombreuses années ont pu renaître de leur cendre : le connexionnisme notamment, sous l'impulsion des travaux du physicien John Hopfield en 1982. Les financements reprennent finalement dans le Monde.

Ce retour de l'IA sur le devant de la scène lui a permis d'être ce qu'elle est aujourd'hui.

#### 2.4. Depuis 1987

Un des derniers paradigmes ayant vu le jour dans le monde de l'IA a été celui des **Agents Intelligents** (cf. Figure 5). Leurs créateurs, Stuart Russel et Peter Norvig, les définissent dans leur ouvrage *IA* comme des entités autonomes capables de percevoir leur environnement et d'entreprendre des actions maximisant leurs chances de réussite. Ce paradigme a ouvert la voie pour de nouvelles découvertes très prometteuses(25).

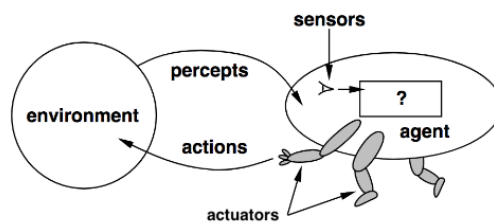


Figure 5 - Schéma de fonctionnement des Agents Intelligents (26)

Depuis la fin des années 80, le secteur de l'IA s'est fractionné en de nombreuses branches et une véritable **révolution technologique** a eu lieu : smartphone, Internet, ou ordinateurs personnels deviennent plus faciles d'accès, la connectivité explose : le terme d'ère de l'*Homo Numericus* est apparu.

Les possibilités qui s'offrent à ces technologies semblent infinies : création d'œuvres d'art, véhicules autonomes, traduction instantanée, etc...c'est pourquoi de grandes entreprises informatiques, telles que Microsoft ou IBM investissent énormément dans ce domaine pour poursuivre les progrès.

Une entreprise britannique, DeepMind Technologies, a créé un logiciel d'IA en 2010, Alpha Go, qui a réussi à vaincre au jeu de Go non seulement d'autres programmes (une seule défaite face à d'autres programmes en 500 matchs), mais également les meilleurs joueurs mondiaux (27). Cette entreprise a depuis été rachetée par Google.

De nombreuses notions sont apparues dans les années 2000 : les concepts d'IA faible et forte (vers 2011), l'interconnectivité, ou les analyses de comportement, et de nombreuses autres restent à découvrir. L'IA fait désormais partie intégrante de notre quotidien, elle n'appartient plus qu'aux seuls scientifiques : tout le monde se fait son avis sur ces technologies.

### 3. Principes essentiels

L'IA est un sujet complexe : ce n'est pas une discipline établie comme les mathématiques ou les sciences biologiques. Bien au contraire, l'IA, bien que très mécanique à sa création, fait désormais appel à de nombreuses notions provenant d'horizons différents : logique statistique et mathématique, informatique et neurobiologie computationnelle (*discipline étudiant les algorithmes qui permettent d'expliquer l'implémentation de nos fonctions cognitives dans notre système nerveux central* (28)), tout en se perfectionnant jour après jour. A ce titre il est possible de dire que l'IA est un processus très dynamique.

L'IA use de plusieurs phases, plusieurs algorithmes pour aboutir à un résultat. Aujourd'hui, cette discipline utilise de grandes bases de données, à partir desquelles elle crée des règles générales par **induction**. Une fois ces règles établies, il est possible d'en faire des **déductions** sur d'autres données.



Il existe différents types d'IA :

- L'IA faible

L'IA faible correspond à des systèmes parvenant à reproduire un seul comportement spécifique, sans en comprendre le fonctionnement (*si l'on demande au même programme d'effectuer une autre tâche il en sera incapable*). Ils analysent des mégadonnées et en extraient des modèles à partir desquels ils effectuent des prédictions.

Cela concerne la majorité des outils utilisés à l'heure actuelle : l'assistant virtuel d'Apple "Siri" en est un bon exemple, puisque ses capacités se limitent à une gamme prédéfinie de tâches. Les suggestions de ventes personnalisées que l'on peut retrouver sur Internet également : en connaissant vos préférences d'achat les logiciels peuvent identifier les cibles les plus adaptées (les fameux *cookies* d'Internet).

- L'IA forte

Par opposition à la faible, l'IA forte *"peut non seulement reproduire des aptitudes à la réflexion et à l'interaction intelligente (analyser, raisonner, effectuer des actions rationnelles), mais également avoir une conscience, des émotions et comprendre ses propres raisonnements"*(29).

Ces systèmes capables de réaliser de nombreuses activités différentes sont toujours en développement à l'heure actuelle.

L'exemple le plus connu de ce type d'IA est sans nul doute la voiture autonome. La conduite automatique prend un grand nombre de facteurs en compte, qu'il faut gérer de façon simultanée (environnement, capteurs internes, etc...) et savoir quel comportement adopter en cas d'urgence, différentes tâches sont à effectuer en même temps.

- L'IA super-intelligente

Ce concept est défini par les prévisionnistes en IA comme des systèmes plus performants que les humains dans tous les domaines, sans en nécessiter l'assistance. On parle de **"singularité technologique"** pour définir le moment où une IA plus intelligente que l'Homme émerge. Mais de tels systèmes n'existent pas encore, et risquent de ne pas exister avant un certain temps, notamment dû aux nombreuses problématiques éthiques ou techniques qu'ils soulèvent(30).

*L'exemple envisagé serait un système mêlant robotique et IA à un point tel que l'Humain ne serait plus nécessaire une fois la conception achevée : acquisition de conscience, d'un système de récompense, de la notion de survie, etc...*

### 3.1. Les bases théoriques de l'IA

#### 3.1.1. Les algorithmes

Un algorithme correspond à un *"ensemble de règles opératoires dont l'application permet de résoudre un problème énoncé au moyen d'un nombre fini d'opérations"*(31). Cela correspond à la mécanisation en plusieurs opérations de la résolution d'un problème.

Il est composé d'une succession finie d'instructions élémentaires (calculs), réalisées dans un ordre précis, puis traitées par un ordinateur. Ce dernier a la puissance adaptée à la réalisation des calculs et au stockage des données. La finalité de l'algorithme est de fournir une solution de façon systématique

*Quel est le lien entre les algorithmes et l'IA ?*

Les algorithmes sont massivement utilisés afin d'automatiser des calculs et obtenir des résultats très rapidement. Les algorithmes sont à la base de nombreuses autres notions relatives à l'IA, notamment les réseaux neuronaux artificiels.

### 3.1.2. Les réseaux neuronaux artificiels

#### 3.1.2.1. Présentation

En IA, le souhait est de développer des machines ayant conscience de leur environnement, et qui vont pouvoir avoir des actions dépendant de celui-ci. L'objet va interpréter des signes, du langage et avec raisonnement produire une solution, en imitant le fonctionnement du cerveau humain. Ce raisonnement se base sur une notion fondamentale en IA : le neurone formel et les réseaux neuronaux artificiels.

Le **neurone formel** (défini par Pitts et McCulloch) correspond à un modèle de calcul simplifié, sous la forme d'un neurone mathématique et informatique mimant le fonctionnement d'un neurone biologique humain.

*Un neurone biologique fonctionne en émettant des signaux électriques : les influx nerveux. Chacun possède un axone qui sert à transmettre cet influx vers le neurone voisin, ce qui permet l'activité fonctionnelle du cerveau. La zone de contact entre deux neurones est appelée une synapse. L'activation des neurones voisins fonctionne selon une loi du tout ou rien, selon la nature des signaux reçus.*

Conçus pour reproduire certaines caractéristiques de la mémoire, les neurones formels ont la capacité d'apprendre, de mémoriser l'information dans les connexions entre les neurones ou même de traiter des informations incomplètes (32).

Ces neurones sont associés en **réseaux** (cf. Figure 6) : chaque neurone est un nœud du réseau, recevant de l'information entrante, la traitant avec une fonction de transfert et émettant de l'information en sortie.

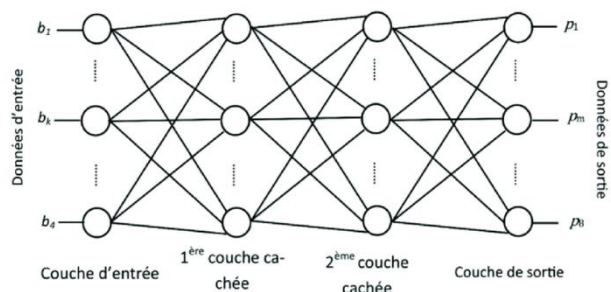


Figure 6 -Représentation d'un réseau neuronal artificiel (33)

Les neurones de la couche d'entrée (*assimilables à une couche sensorielle*) sont chargés de capter les données extérieures (données brutes), telles que la couleur des pixels d'une image, ou les lettres d'un document. Par la suite, le signal va être transmis par des synapses, à des

neurones de la seconde couche qui réagissent directement au signal perçu par les neurones de la première couche.

Chacun de ces neurones va diviser le signal qu'il reçoit en deux catégories selon la nature des synapses transmettant les signaux : activatrice ou inhibitrice. La contribution du signal à l'activation du neurone sera la multiplication des **coefficients d'activation de la synapse** (ou poids synaptiques) par l'**intensité du signal** (cf. Figure 7).

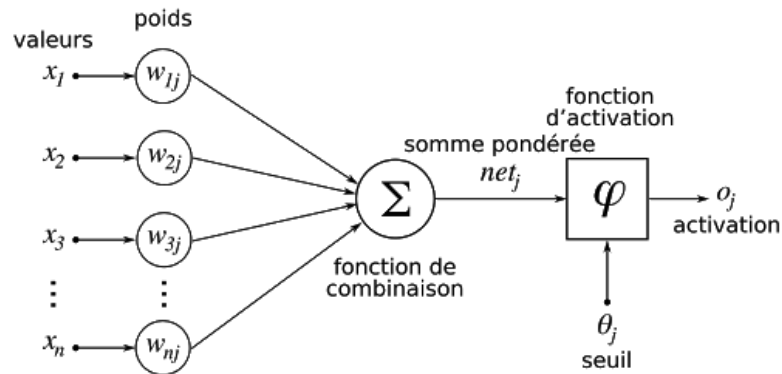


Figure 7 - Schéma du mécanisme d'activation d'un neurone artificiel (34)

La topologie des connexions est variable, tout comme l'efficacité de transmission du signal d'un neurone formel à un autre. Cette dernière est régie par le « **poids synaptique** » et peut être contrôlée par des règles d'apprentissage mathématique.

Une **phase d'apprentissage** du réseau a lieu en début d'exploitation : le but est de calculer et de faire tendre ces poids synaptiques vers la valeur la plus adaptée à chaque situation. Le principal défaut de cette méthode est qu'à sa création, aucune méthode d'adaptation des coefficients n'existe.

Finalement, le neurone s'active, si la somme pondérée des signaux (inhibiteurs et activateurs) est supérieure au seuil d'activation du neurone, en appliquant une **fonction d'activation** aux valeurs d'entrée (le signal d'entrée est ainsi modifié en sortie).

Différents types de fonctions d'activation existent en IA :

- Les neurones ReLU pour *Rectifier Linear Unit*. Ils ne s'activent que si les contributions activatrices l'emportent, et son niveau d'activation sera proportionnel à l'écart entre les contributions activatrices et inhibitrices.
- Les neurones Sigmoides : ces neurones s'activent légèrement si les contributions inhibitrices l'emportent, à 50% s'il y a égalité entre les deux contributions, et à près de 100% si les contributions activatrices l'emportent(35).

### 3.1.2.2. La règle de Hebb

C'est Donald Hebb, qui va par la suite exposer une règle aidant dans la modification de ces valeurs de coefficients synaptiques, selon l'activité des neurones connectés : la **Règle de**

**Hebb.** Celui-ci la résume de la façon suivante : “des neurones qui s'excitent ensemble se lient entre eux” (36), évoquant la création d'une connexion (ou le renforcement d'un lien existant) lors de l'excitation conjointe de deux neurones. Si deux neurones connectés sont actifs simultanément, le poids de leur connexion est augmenté ou diminué (37). Cette règle sera la base de l'apprentissage associatif, et est un des standards, toujours présent dans les modèles actuels.

### 3.1.2.3. Le Perceptron

En 1957, le scientifique Frank Rosenblatt s'inspire des travaux de McCulloch et Pitts pour créer le premier réseau de neurone, nommé le **Perceptron** (cf. Figure 8), au sein du *Cornell Aeronautical Laboratory*. Les neurones sont disposés sur une couche unique, sans que la topologie des connexions ne puisse évoluer. Cependant, les poids synaptiques relatifs à la transmission du signal peuvent être modifiés par rétroaction afin d'obtenir le résultat final désiré. De façon similaire à la plasticité synaptique de nos réseaux neuronaux, l'efficacité de la transmission des signaux d'un neurone à un autre peut ainsi varier dans ce système.

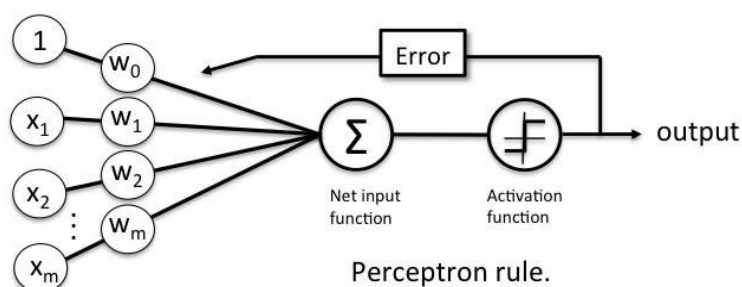


Figure 8 - Schéma de fonctionnement d'un perceptron (38)

L'apprentissage s'inspire donc d'un modèle visuel, et fonctionne par expérience, même dans le cas où l'instructeur a commis une erreur au cours du processus (corrigée par rétroaction), ce qui n'était pas le cas dans les réseaux neuronaux formels (39). L'information circule de la couche d'entrée vers la couche de sortie. Ici, les sorties sont multiples, à la différence du réseau neuronal formel(40).

### 3.1.3. Les réseaux convolutifs

En 1969, Marvin Minsky et Seymour Papert dénoncent les limites du concept dans *Perceptrons*(41). Des études menées par la suite sur le cortex visuel d'animaux ont démontré qu'il existait des arrangements complexes de cellules se chevauchant lors du quadrillage du champ visuel. Cela a donné naissance au principe de **réseau neuronal convolutif** (cf. Figure 9). Ce réseau consiste en la superposition de multicouches de perceptrons qui effectuent un prétraitement des informations qui y transitent, afin d'en extraire des caractéristiques essentielles.

L'image va être divisée en différentes parties lors du quadrillage, et chaque cadran sera traité individuellement par un neurone artificiel, qui en analyse les pixels, grâce à des paramètres de

réglages. Les différents cadrans se chevauchent légèrement les uns sur les autres, afin d'obtenir une représentation plus fidèle de l'image.

Tous les neurones partageant les mêmes paramètres (effectuant le même traitement) forment le **noyau de convolution**. C'est un des avantages du réseau convolutif vis à vis du perceptron multicouche que de pouvoir affecter un poids unique aux signaux d'entrée entrant dans un noyau de convolution. En effet, le réseau multicouche considère chaque neurone indépendamment.

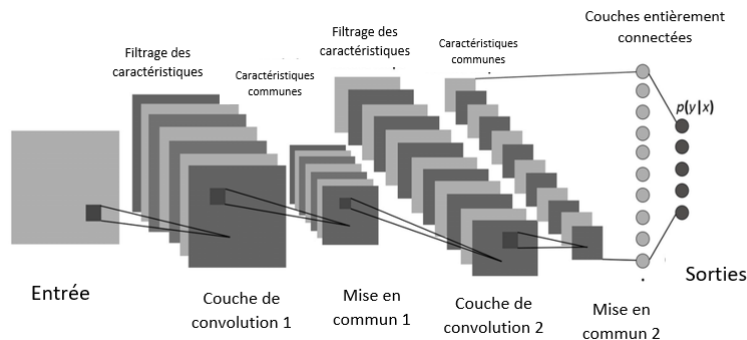


Figure 9 - Schéma d'un réseau convolutif multicouche (42)

La première couche de neurones est propre au réseau convolutif, et permet d'**extraire des caractéristiques** de l'image en entrée. Des opérations dites de filtrage par convolution permettent de faire correspondre des caractéristiques communes de différents noyaux de convolution et les normaliser (étape de mise en commun ou "*pooling*"). Au départ, les premiers filtres appliqués sont aléatoires, mais au fur et à mesure des mises en commun, ils se perfectionnent (les filtres donnant les meilleures correspondances sont conservés et appris).

Un seul de ces noyaux va pouvoir analyser **une caractéristique de l'entrée**. Ainsi, pour en analyser plusieurs, des strates de noyaux seront empilés, et c'est cet ensemble de couches qui forme la "couche de traitement convolutif". Ce procédé peut être réitéré à plusieurs reprises, jusqu'à la sortie du premier bloc de neurones, sous la forme d'une image intermédiaire.

Cette découverte a eu un impact très important dans le développement des technologies de reconnaissance d'images, ou de reconnaissance vocale, mais surtout dans le développement des techniques d'apprentissage automatique (*présentées ci-après*).

#### 3.1.4. Les techniques statistiques

Les statistiques jouent un rôle prépondérant dans l'IA. Voici quelques exemples non exhaustifs d'outils utilisés :

- Arbres de classification et de régression

Ces arbres sont très utilisés dans les processus d'aide à la prise de décision. Ils se présentent sous la forme d'arborescences de situation et leurs conséquences possibles.

Le résultat d'un arbre de classification, ou de décision, est une valeur discrète (telle que la classe à laquelle appartiennent les données) tandis que les arbres de régressions prédisent une valeur numérique réelle (prix d'un logement, etc...).

- Apprentissage par instance

L'apprentissage par instance, ou par l'exemple, qui, plutôt que de faire des généralisations explicites, compare les nouvelles observations avec des cas observés au cours d'une formation qui ont été sauvegardés en mémoire.

### 3.2. L'apprentissage de l'IA

Yann Le Cun est un chercheur en IA français, considéré comme l'un des fondateurs de l'apprentissage profond. Il a déclaré en 2016 au Collège de France “*Il n'y a pas d'intelligence sans apprentissage*”. L'Homme apprend de ses expériences passées, tandis que les machines en suivent les instructions. L'Homme ne peut-il donc pas entraîner les machines afin qu'elles apprennent directement des données passées qu'elles ont eu à leur disposition ? (43)

Au commencement de l'IA, l'apprentissage n'était pas considéré comme essentiel : l'outil intelligent était programmé « à la main » pour effectuer des fonctions simples (jouer aux échecs, reconnaître des caractères, etc...) sans prendre en compte l'environnement dans lequel il évolue. Cette approche était donc limitée, et peu adaptée à certaines tâches pourtant simples (reconnaissance d'objet, etc...). En effet, les données extérieures sont très complexes, variables, et ternies de bruit (pixels d'une image, échantillons de son...).

Une machine interprète une image d'une façon totalement différente de l'Homme : pour elle, une image est un tableau de chiffres précisant la couleur ou la luminosité de chaque pixel. Ainsi, comment peut-elle identifier un objet précis ou un animal, quand leur apparence peut varier à l'infini ?

L'IA englobe en son sein un grand nombre de notions, dont l'apprentissage automatique (cf. Figure 10). Cette thématique englobe elle-même les réseaux de neurones et l'apprentissage profond. Par conséquent, l'apprentissage est fortement imbriqué dans le concept d'IA.

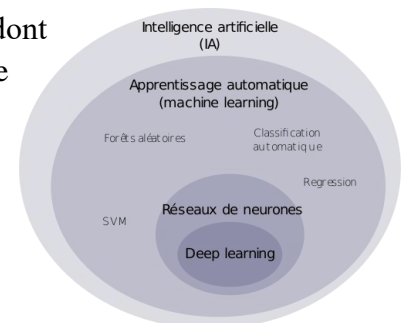


Figure 10 - Lien entre l'IA et l'Apprentissage (44)

#### 3.2.1. Apprentissage Automatique – Machine Learning

Depuis quelques années, le fonctionnement des réseaux neuronaux est optimisé par des méthodes d'apprentissage automatique plus communément appelés « Machine Learning ».



C'est Arthur Samuel, un informaticien américain qui fut le premier à faire usage de l'expression *machine learning* en 1959. Il a créé un programme pour IBM dont le but était de jouer au jeu de dames. Sa particularité était de s'améliorer au fur et à mesure des parties jouées.

Le machine learning (cf. Figure 11) vise à permettre aux ordinateurs d'apprendre à partir de données, et donc d'améliorer leurs performances à résoudre des tâches sans pour autant être programmés pour chacune.

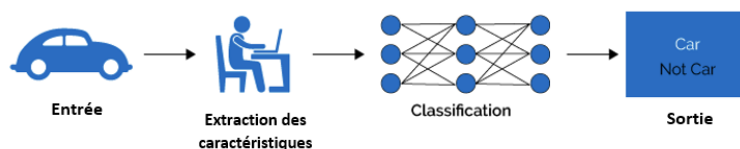


Figure 11 - Principe de fonctionnement de l'Apprentissage Automatique ou Machine Learning (44)

Ce principe est très utilisé dans la reconnaissance d'images. Un **algorithme intermédiaire**, qui n'appartient pas à un réseau de neurones, est créé et va permettre d'extraire les caractéristiques intéressantes de l'image étudiée (c'est l'abstraction d'image). Ces caractéristiques seront transmises à un réseau de neurone pour effectuer la reconnaissance. Cela permet de réduire drastiquement la quantité de données en entrée du réseau (*on passe d'une image de quelques millions de pixels, à une dizaine de caractéristiques à analyser*). Le concepteur de l'algorithme joue donc un rôle primordial.

L'apprentissage automatique comporte généralement deux phases :

La première est la **phase d'apprentissage**. Elle consiste à estimer le lien dans un modèle entre une entrée et une sortie à partir de données issues d'observations afin de résoudre un problème pratique : reconnaître la présence d'un animal sur une photographie, traduire un énoncé, etc...

*Un exemple consiste à mesurer la teneur en Principe Actif (PA) d'un médicament obtenu pour différentes forces de mélangeage utilisées (la plupart des utilisations de ce modèle se fait avec plus de deux variables). Les données obtenues peuvent être représentées sur un graphique (cf. Figure 12) :*

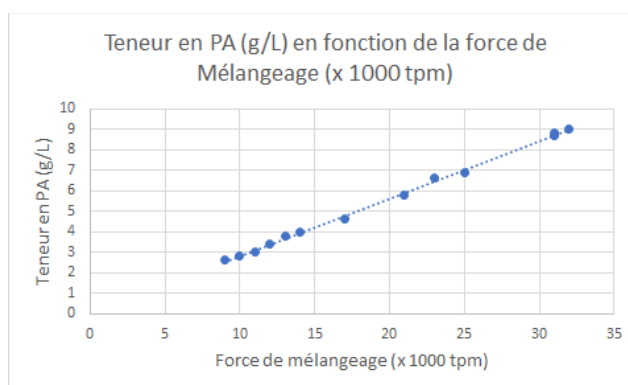


Figure 12 - Evolution de la teneur en PA d'un médicament en fonction de la force de mélangeage

Le but du machine learning est de trouver la relation mathématique modélisant la relation entre l'entrée (forces de mélangeage) et la sortie (teneur en Principe Actif) le plus fidèlement. Cela correspond donc à déterminer l'équation de la droite  $ax + b$  qui passe au plus près des points. Ici la finalité de notre modèle est d'estimer l'un des paramètres en en connaissant un autre (déduire la teneur en PA en en connaissant la force appliquée au mélangeur par exemple).

La seconde est la **phase de prédiction**. Une fois le modèle déterminé, de nouvelles données peuvent être utilisées afin d'estimer le résultat de la tâche souhaitée.

Une fois le modèle établi, il est possible pour l'outil d'extrapoler ces résultats à des situations qu'il ne connaît pas (cf. Figure 13). En demandant à l'algorithme de nous donner la teneur en PA pour une force de mélangeage de 20.000 tours par minute, il nous répondra dans notre exemple environ 5.6 grammes par litre, grâce au modèle précédemment établi.

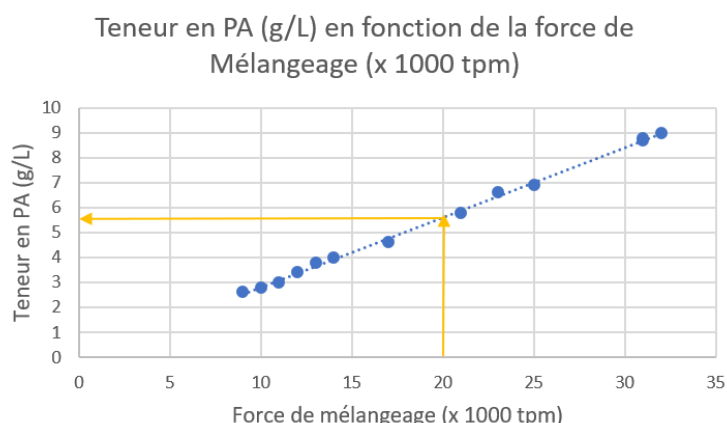


Figure 13 - Extrapolation de données à une situation inconnue

Ce procédé d'apprentissage automatique regroupe trois grandes méthodes d'apprentissage :

- L'Apprentissage Supervisé (cf. Figure 14)

Avec cette méthode, on connaît la réponse à la tâche pour les données utilisées. Cet apprentissage se fait par l'exemple.

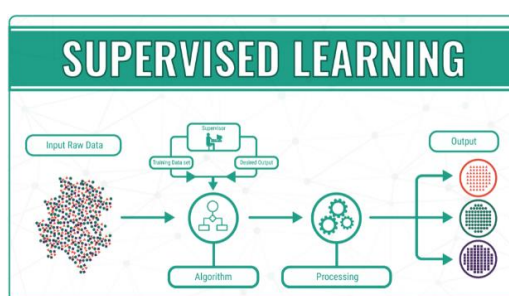


Figure 14 - Principe de fonctionnement de l'apprentissage supervisé (45)

L'Homme derrière l'algorithme connaît donc les différentes classes de résultats possibles des exemples qu'il fournit à l'algorithme.

Google a par exemple utilisé cette méthode pour développer une IA capable d'identifier des chats sur des images. Pour ce faire, l'outil a été exposé à un très grand nombre d'images, en lui indiquant lesquelles étaient celles de félins, jusqu'à ce qu'elle soit capable de les reconnaître par analogie.

- L'Apprentissage Non Supervisé (cf. Figure 15)

Cette fois, l'Homme ne fournit que les données en entrée, sans en préciser la réponse attendue. C'est l'outil qui va classer les données présentant des caractéristiques communes, à l'aide de critères statistiques, dans le but de discriminer les futures nouvelles entrées.

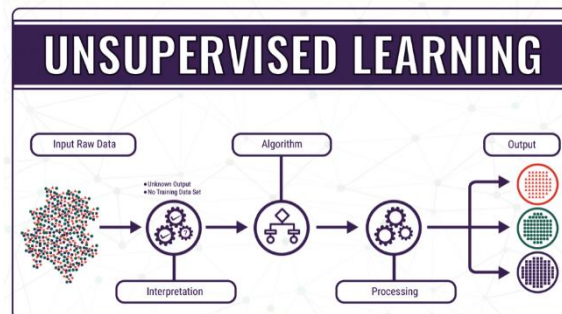


Figure 15 - Principe de fonctionnement de l'apprentissage non supervisé (45)

L'Homme ne connaît pas le nombre et les classes rattachées aux données, c'est au modèle de les découvrir. Cette méthode est moins développée que la précédente à l'heure actuelle.

- L'Apprentissage Par Renforcement (cf. Figure 16)

Cette dernière méthode est considérée comme la plus faible des techniques d'apprentissage automatique. Le système possède une entrée et va tester les différentes sorties par essai/erreur, selon la probabilité qu'elles auront d'être de bonnes solutions, définies comme étant celles donnant une récompense au long terme (*la victoire d'une partie par exemple*). La découverte de la solution optimale est donc incrémentale.

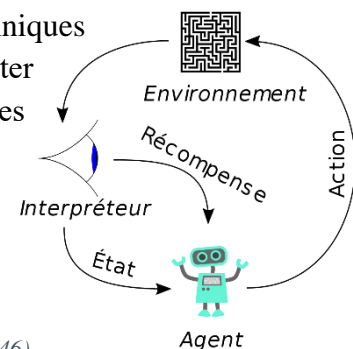


Figure 16 - Principe de fonctionnement de l'apprentissage par renforcement (46)

C'est une méthode très lente, qui nécessite de faire tourner les options des millions de fois, en récompensant les bonnes actions et en punissant les mauvaises, jusqu'à découvrir la meilleure solution.

### 3.2.2. Apprentissage Profond – Deep Learning

C'est un système d'apprentissage automatique qui s'appuie également sur les réseaux de neurones artificiels et les réseaux convolutionnels précédemment évoqués.

Plutôt que d'utiliser un algorithme pour extraire les caractéristiques essentielles de l'entrée, on va chercher à entraîner le réseau, afin que les couches supérieures extraient ces caractéristiques. Ce sera directement l'algorithme qui aura découvert les caractéristiques de l'entrée, sans qu'aucune intervention humaine ne soit nécessaire.

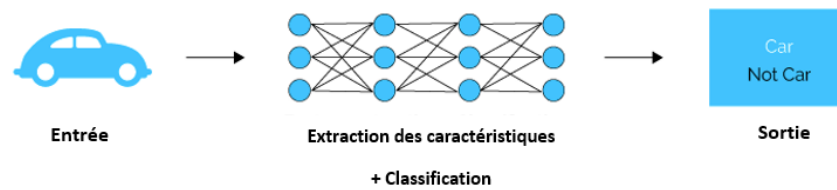


Figure 17 - Principe de fonctionnement de l'Apprentissage Profond ou Deep Learning (44)

L'apprentissage profond (cf. Figure 17) se base sur les réseaux de neurones artificiels imitant le cerveau humain et sur une multitude d'algorithmes simples. Ce système va pouvoir adapter ses comportements et ses analyses selon les données contenues dans des bases de données massives (ou des capteurs).

Ce réseau est composé de plusieurs couches de neurones, chacune recevant, transformant et interprétant les informations de la couche précédente. Pour chaque étape, les réponses jugées "mauvaises" sont éliminées et renvoyées vers les couches en amont, afin de modifier et d'adapter le modèle mathématique. Les informations sont donc réorganisées en blocs plus complexes au cours des étapes successives (47).

Le "deep learning" s'inspire du fonctionnement cérébral de cette façon :

- Il débute par effectuer des **corrélations** entre des actions et des conséquences, afin d'identifier les situations les plus probables ;
- Il émet ensuite des **hypotheses** et les teste ; si elles sont validées, il les intègre dans son modèle ;
- Il regroupe des éléments simples et récurrents dans des **concepts**. C'est la phase d'imitation ;
- Enfin, il recommence, et **affine** son modèle.

Une fois le modèle validé, il peut être utilisé dans d'autres situations.

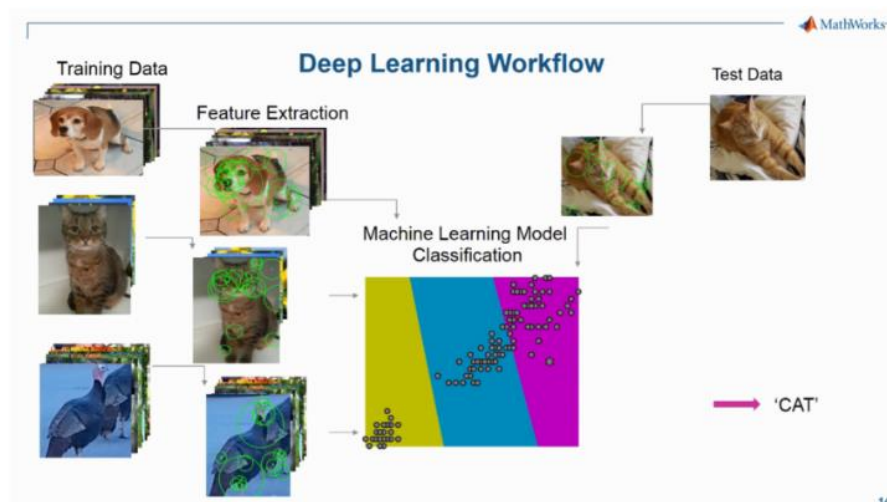


Figure 18 - Exemple d'application d'apprentissage profond sur de la reconnaissance d'images (48)

Dans le cas de l'exemple ci-dessus (cf. Figure 18), un grand nombre de données sont utilisées pour entraîner le système (photo de chien, de chats...). Cela lui permet d'en extraire des caractéristiques qui leur sont propres, classifiées dans le modèle d'apprentissage. L'algorithme peut dès lors être utilisé sur une image inconnue, afin de reconnaître ce qui se trouve dessus.

Le *deep learning* est en pleine expansion à l'heure actuelle. Il nécessitait des puissances de calcul très importantes auparavant irréalisables, qu'il est désormais possible d'atteindre.

La majeure avancée dans ce domaine a été l'accès facilité aux données. Des bases de données d'images ont en effet été mises à disposition de professionnels gratuitement, notamment par un laboratoire de Stanford, *Imagenet* (49). Ce site contient des millions d'images classifiées (avec la description de ce qu'elles sont) permettant de nourrir les algorithmes(50).

L'intérêt de ce modèle est d'être **adaptatif**. En effet, les programmes peuvent s'adapter lorsqu'ils sont confrontés à de nouvelles données. C'est de là que vient l'importance des Big Data mentionnées prochainement dans le domaine de l'IA : elles permettent de "nourrir" les programmes, et de les améliorer, en les rendant plus fidèles à la réalité. Plus un modèle a été confronté à un nombre de données importante et plus il est réaliste.

Aujourd'hui, des réseaux constitués de centaines de couches, avec plus d'un million de neurones arrivent à être créés, ce qui reste encore bien loin de notre architecture cérébrale, mais les performances de reconnaissance d'image sont spectaculaires. Les débouchés sont pléthores : reconnaissance d'images pour les non-voyants, détection de situations dangereuses, conduite de véhicule autonome, adaptation de machines à diverses situations, etc...

### 3.3. Le test de Turing

Dès la découverte de l'IA, il a fallu trouver un moyen de définir un standard permettant de qualifier une machine de « consciente ». C'est Alan Turing qui, en 1950 avec sa publication *Computing Machinery and Intelligence*(51) a décrit ce test. Son objectif était de répondre à la question : "Est-ce que les machines pensent ?"

Le principe de ce test est extrêmement simple : il consiste à mettre un individu Humain (*noté C sur la figure 19*) en contact verbal à l'aveugle avec un autre Humain (*noté B sur la figure 19*) et un Ordinateur (*noté A sur la figure 19*). Les échanges demeurent simples, et se limitent à des échanges de messages textuels entre les individus A et C, et B et C (cf. Figure 19).

Un logiciel passe avec succès le test de Turing si l'interlocuteur C n'arrive **pas à distinguer l'Homme de l'ordinateur**. Cette expérience correspond ainsi à un "jeu d'imitation" pour la machine (*Imitation Game*, comme nommé par Alan Turing), qui essaye d'avoir la sémantique la plus "humaine" possible.

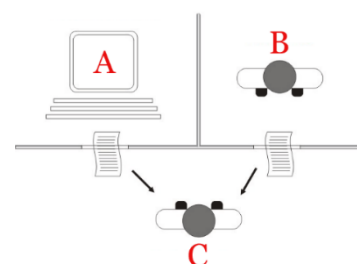


Figure 19 - Schéma du test de Turing (52)

Des tests de Turing sont organisés à des échéances régulières, un exemple de machine ayant réussi à passer ce test a été un robot russe dénommé Eugène Goostman. Pour cela, il a remporté le prix Loebner en 2012, qui récompense les machines arrivant le mieux à simuler l'esprit humain.

La stratégie des robots a évolué au fil du temps. Les premières machines se 'contentaient' de comprendre les questions des jurés, tandis qu'à l'heure actuelle, celles-ci posent des questions, changent de sujet, font des fautes de frappe... La réussite d'Eugène Goostman repose beaucoup sur la personnalité qui lui a été attribuée : celle d'un jeune homme de 13 ans. Un des créateurs du robot, Vladimir Veselov, explique : *"Notre idée conductrice était que son âge rend parfaitement plausible qu'il ne sache pas tout. Cela rend également ses attributions comme les fautes d'orthographe plus vraisemblables que venant d'un adulte"*(53).

Pour que le test soit valide, il faut que l'ordinateur soit confondu avec un humain plus de 30% du temps pendant une série de 5 minutes d'échanges manuscrits. Le robot a réussi à convaincre 33% des juges de la Royal Society de Londres en 2014 qu'il était humain(53).

Cependant, ce test est remis en cause par bon nombres de spécialistes, et il a depuis été démontré inadéquat, une machine pouvant facilement réussir le test sans pour autant être consciente. La principale limite de ce test est qu'il ne teste pas directement si l'ordinateur se comporte de façon intelligente, uniquement s'il parvient à imiter un être humain. Or, tous les comportements intelligents ne sont pas Humains, de la même façon que certains comportements correspondant à une intelligence humaine ne sont pas parfaits ou rationnels (mensonge, tromperie, fautes de frappe...).

De plus, certaines conversations buggées peuvent conduire à un succès de la machine, en effet, dès lors qu'un doute subsiste, la machine est déclarée vainqueur. Malgré toutes les critiques qui lui sont faites, aucun autre test n'a réussi à détrôner celui de Turing à l'heure actuelle.

## **4. L'ère du Big Data**

### **4.1. Définition du Big Data**

Le Big Data est également nommé mégadonnées ou données massives en français. La Commission Générale de Terminologie et de Néologie, a défini ces données comme étant les *"données structurées ou non dont le très grand volume requiert des outils d'analyse adaptés"*(54).

De cette définition découle la notion d'abondance de données, qui surpassent à la fois les capacités humaines d'analyse et d'intuition, mais aussi les outils informatiques classiques de gestion de base de données. Le Big Data est communément défini par les "3V"(55) :

- Volume ;
- Variété ;
- Vitesse.

Ces "3V" sont complétés par deux autres :



- Véracité

Elle correspond à la fiabilité des données, qui est menacée par les comportements falsifiés, la multiplication des formats de données ou par l'activité de faux profils sévissant sur Internet.

- Valeur

Dans un environnement où tant de données transitent, il faut savoir distinguer les données ayant une réelle valeur de celles qui n'en ont pas afin de mener des analyses pertinentes.

L'IA et les mégadonnées sont intimement liées : sans Big Data, l'IA ne pourrait tout simplement pas apprendre. En effet, plus un outil utilisant l'IA rencontre de données, plus il peut en apprendre et donc devenir "intelligent"(56). En ce sens, les modèles d'IA fonctionnent de la même façon que le cerveau humain : ils se développent en accumulant des données à la suite de diverses expériences.

#### 4.2. Evolution du volume des Big Data

Les données sont partout dans nos vies quotidiennes. Des mails échangés entre collègues, aux données GPS de nos téléphones, en passant par les vidéos mises en ligne, tout est créateur de *data*. Selon l'International Data Corporation (cf. Figure 20) et son "Data Age 2025", le volume de données mondiales (*Datasphere*) va passer de 33 zettabytes ( $10^{21}$ ) en 2018 à 175 d'ici 2025(57).

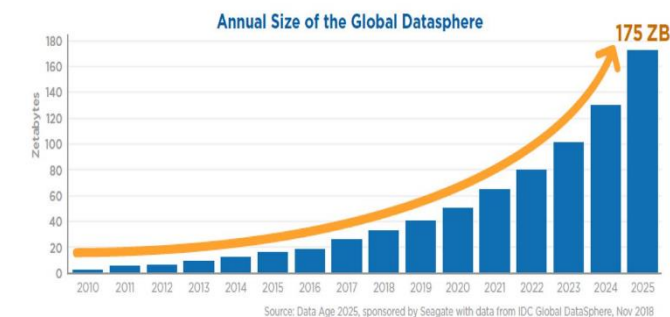


Figure 20 - Evolution du volume de la Datasphere entre 2010 et 2025 (57)

Selon l'œuvre de Yann Houry, *Un centaure à l'école : Comment le numérique change (et doit encore changer)*, ce sont près de 2,5 trillions ( $10^{18}$ ) d'octets de données qui sont générées tous les jours(58). En une minute, près de 350.000 appels ont lieu sur Skype, 700 vidéos sont ajoutées sur YouTube, et plus de 160 millions d'emails sont échangés dans le Monde. Tout cela génère de la donnée.

En ce qui concerne les données relatives à la santé, selon une projection établie en 2014 dans le cadre de l'étude EMC Digital Universe d'IDC, 2.300 exaoctets ( $10^{18}$ ) de données de santé devraient être produites et enregistrées sur des serveurs informatiques à travers le monde en 2020 contre "seulement" quinze fois moins en 2013(59).



Cette explosion du nombre de données est directement liée à celle des objets connectés qui génèrent un grand nombre de data.

Les **objet connecté** (*IoT - Internet of Things*) sont définis comme étant “des objets ayant la particularité d’être, comme leur nom l’indique, connectés entre eux via Wi-Fi, Bluetooth ou tout autre protocole de communication leur permettant d’échanger des informations”(60). Ils existent sous toutes les formes (montre, téléphone, balance de pesée, machine à laver, etc...) : à vrai dire tout objet du quotidien semble destiné à le devenir un jour.

Dans le milieu pharmaceutique, de plus en plus de capteurs sont utilisés au quotidien dans de nombreux domaines divers (en production pour suivre les déviations, en logistique pour gérer les stocks, etc...).

Selon une étude menée en Mai 2019 par *Strategy Analytics*, “près de 22 milliards d’objets connectés à Internet ont été déployés dans le monde à la fin de l’année 2018, tandis que 17 milliards supplémentaires s’y ajouteront d’ici 2025”(61). Il y a donc environ trois fois plus d’objets connectés que d’êtres humains sur Terre.

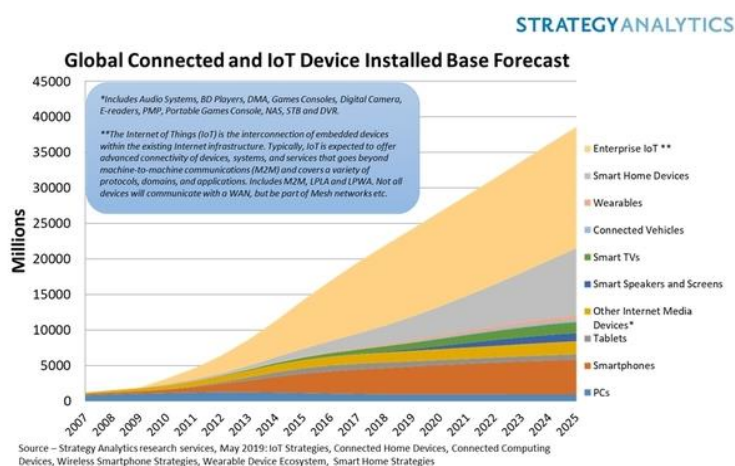


Figure 21 - Evolution du nombre d’objets connectés entre 2007 et 2025 (61)

Le principal secteur (cf. *Figure 21*) concerné par les Big Data est le milieu professionnel, avec près de 50% des données générées, mais un autre est en plein essor : la maison connectée(62).

Plusieurs facteurs expliquent cette explosion du nombre d’objets connectés.

- Hausse de la connectivité

Les réseaux 4G sont très développés sur les territoires, tandis que le réseau 5G est en cours de déploiement. Ce dernier est très prometteur en termes de latences, de trafic de données ou de vitesse de connexion pour la progression des objets connectés (cf. *Figure 22*).

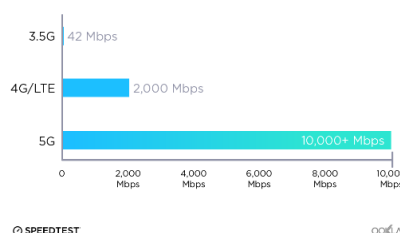


Figure 22 - Comparaison des vitesses de téléchargement en fonction des technologies (63)

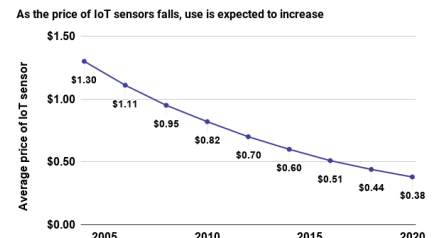
- Miniaturisation des processeurs

Les processeurs, véritables moteurs des objets connectés sont de plus en plus petits, ce qui facilite leur intégration et leur exploitation à grande échelle.

- Baisse du prix des composants (capteurs, etc...)

Selon un rapport de Microsoft, *2019 Manufacturing Trends*, le prix moyen d'un capteur est passé de \$1.3 en 2005 à \$0.38 en 2020(64) (cf. Figure 23). Cela entraîne une diminution du prix des objets connectés, et de fait une diffusion plus large.

Figure 23 - Evolution du prix d'un capteur depuis 2005 (64)



- Explosion du cloud

Le stockage sur le cloud est une option économique, offrant des capacités de stockage presque infinies. Tout cela conduit à une intégration, à un croisement et à une distribution facilitée de ces données.

#### 4.3. Variété des Big Data

Les données générées sont de deux grands types :

- Les données Structurées

Ce sont les données contenues dans les champs. Structurée, car leur nature et leur fonction sont identifiées par les **balises des métadonnées**. Les informations structurées sont disposées à être traitées automatiquement et efficacement par un logiciel.

- Les données Non structurées :

Il s'agit des données représentées ou stockées **sans format prédéfini**, (documents textes ou multimédias). Cette absence de format entraîne des irrégularités et des ambiguïtés qui peuvent rendre difficile l'analyse ou la compréhension des données.

*Dans un e-mail, les données structurées correspondent à toutes les informations inscrites dans les champs prédéfinis (adresse e-mail, objet du message, date et heure, etc...) destinées à être traitées par une machine, tandis que le corps du message, destiné à être lu par un humain, sera composé de données non structurées, car sans format prédéfini.*

Les données sont d'une grande variabilité :

- Horodatées

Les données horodatées sont des données ordonnancées permettant de définir la séquence selon laquelle chaque point de données a été enregistré. Ce type de données est généralement utilisé lors de la collecte de données comportementales afin d'avoir une représentation des actions au

fil du temps. Cela sert notamment pour étudier les actions d'un utilisateur sur un site Web, afin d'en définir le parcours.

*Ce type de données nous donne par exemple l'heure de conditionnement d'un produit.*

- Data spatio-temporelles

Les données spatio-temporelles décrivent, en plus de l'heure, le lieu d'un événement et peuvent également nous montrer comment les phénomènes dans un lieu physique évoluent avec le temps. C'est ce type de données qui sont générées dans le suivi de véhicules en mouvement.

- Machine Data

Les données machine sont les informations digitales générées par les systèmes qui alimentent les entreprises modernes (*données de capteurs, historiques d'appels, etc...*). Certaines entreprises analysent par exemple les données machine sur les performances en temps réel de leurs équipements, en les croisant avec les historiques de performances afin de mieux appréhender les problèmes et d'effectuer de la maintenance préventive(65).

- Open data

Ces données sont librement accessibles à quiconque, pour utilisation et analyse. Ces données peuvent être publiées sans restriction de droits d'auteur ou de brevets. C'est la grande variété des données générées qui en explique la complexité de gestion. C'est un des enjeux majeurs de l'industrie de demain.

#### 4.4. Vitesse des Big Data

La vitesse correspond à la fréquence à laquelle les données sont générées, captées, partagées et mises à jour. Du fait de leur grande vitesse de génération, les données doivent à tout prix être traitées en temps réel ou à *minima* le plus vite possible.

#### 4.5. Conclusion

La quantité grandissante des données à analyser conduit à plusieurs enjeux : optimiser le traitement des données, assurer la mise en relation de différentes professions, et le plus important, sécuriser les données. La Directive Européenne 2016/680 stipule qu' *"Afin de préserver la sécurité entourant le traitement et de prévenir tout traitement effectué en violation de la présente directive, il convient que les données à caractère personnel soient traitées de manière à garantir un niveau de **sécurité** et de **confidentialité** approprié, notamment en empêchant l'accès non autorisé à ces données et à l'équipement servant à leur traitement ainsi que l'utilisation non autorisée de ces données et de cet équipement, et à tenir compte de l'état des connaissances et de la technologie disponible, des coûts de mise en œuvre au regard des risques et de la nature des données à caractère personnel à **protéger**"*(66).

Il est donc obligatoire de mettre en place des mesures de sécurité nécessaires pour éviter le traitement illégal ou la perte accidentelle des données personnelles. C'est un des enjeux majeurs entourant l'IA et le Big Data.

## 5. Place de la France dans la recherche en IA

### 5.1. Place de la France dans la recherche en IA

La France joue un grand rôle dans la recherche en IA. Avec près de 270 équipes de recherches comptabilisant plus de 5000 chercheurs et plus de 400 million d'euros de financement public, elle figure dans les 4 pays du Monde produisant le plus d'articles scientifiques à son sujet (après les Etats-Unis, la Chine et le Royaume-Uni)(9).

Selon le LEEM(9), deux approches d'analyse et d'usage de l'IA en santé sont complémentaires aujourd'hui :

- La première approche s'attache à appliquer des algorithmes afin de faire apparaître des **conclusions** sur des données (efficacité des résultats obtenus, valeur médicale de biomarqueurs de prévention, etc...) ;
- Tandis que la seconde vise à **modéliser** à l'échelle cellulaire, tissulaire, voire organique, dans le but de guider l'apprentissage algorithmique pour gagner en efficacité, et ainsi réduire la masse de données requises pour l'apprentissage.

L'IA a donc un potentiel très important dans l'ensemble du secteur de la santé, pour améliorer la qualité des soins pour le patient, en réduisant leur coût et en augmentant sa sécurité et sa traçabilité. De nombreuses start-up françaises ont anticipé l'importance de l'IA dans le système de santé français du futur. En 2018, Bpifrance publiait la première cartographie des startups françaises en santé qui utilisent l'IA(67).

L'institut a ainsi identifié 104 startups françaises créées depuis 2010 (dont 60% créées depuis 2016) utilisant des technologies d'IA (systèmes experts, machine learning, deep learning) à travers 8 segments marchés :

- L'aide au diagnostic et à l'imagerie ;
- Les objets connectés et le monitoring ;
- L'analyse du comportement et la prévention ;
- Les chatbots médicaux ;
- La gestion du parcours patient et hospitalier ;
- La recherche biologique et pharmaceutique ;
- Le handicap ;
- Et autres projets divers

Il y a eu une explosion du nombre de startup dans le monde liées à ce domaine ces dernières années (cf. Figure 24) et des prévisions estiment que les revenus de ce marché passeront de 633 millions en 2014 à plus de 6,6 milliards en 2021 !

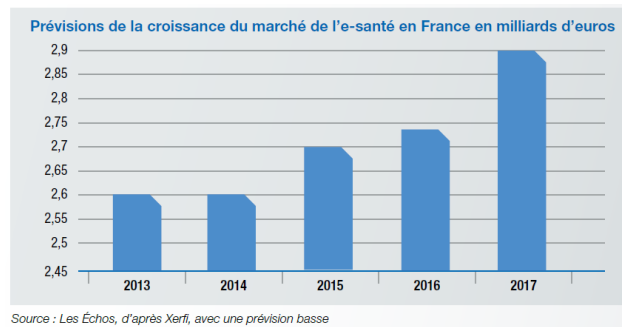


Figure 24 - Prévisions de la croissance du marché de l'e-santé en France (milliards d'euros)(68)

Le marché de l'e-santé en France, comme dans le Monde, est en pleine expansion. Le nombre d'emplois concernés est en croissance constante depuis près de 10 ans, et cette filière est soutenue par le programme *Territoires de soins numériques*, dont l'objectif est de “moderniser le système de soins en expérimentant, dans certaines zones pilotes, les services et les technologies les plus innovants en matière d'e-santé”(69). Son budget annuel est de 80 millions d'euros pour soutenir des projets sélectionnés.

Voici quelques exemples d'utilisation françaises de l'IA dans le secteur de la santé, auprès des professionnels de santé (pharmaciens, médecins, etc...) et des utilisateurs finaux : patients, accompagnants.

## 5.2. Diabeloop

Cette start-up française basée à Grenoble (Auvergne-Rhône-Alpes) a été créée en 2015, par le Dr Guillaume Charpentier et Erik Huneker. Elle est associée au CEA-LETI au sein d'un laboratoire de recherche commun.

Le dispositif médical de Diabeloop permet **d'automatiser la gestion du diabète de type 1** en reproduisant les fonctions du pancréas annihilées par la pathologie.

Dans cette pathologie, le système immunitaire des patients atteints produit des anticorps qui détruisent les cellules bêta du pancréas productrices de l'insuline. Ce déficit en insuline doit être comblé par des injections quotidiennes d'insuline (en stylo, ou sous forme de pompe) : on dit que ce diabète est **insulino-dépendant**, à l'inverse du diabète de type 2.

Le système *Diabeloop* intègre deux éléments (70) (cf. Figure 25) :

- Un capteur en temps réel du taux de glucose ;
- Et une pompe qui délivre l'insuline (*traitement de référence actuel*).

Ces deux éléments sont portés sur la peau et communiquent en Bluetooth avec un terminal dédié et sécurisé qui héberge l'IA. Un algorithme développé par Diabeloop détermine et commande automatiquement la pompe pour injecter les doses d'insuline optimales selon la glycémie.

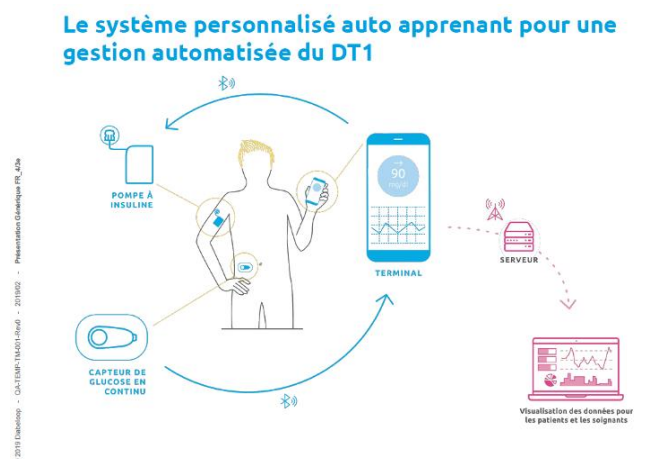


Figure 25 - Schéma de fonctionnement de Diabeloop (5)

Les essais cliniques réalisés en lien avec plus de dix centres hospitaliers, auprès de 150 personnes, ont délivré des résultats très prometteurs. Le dispositif médical a obtenu le marquage CE (*conformité européenne*), lui autorisant l'accès au marché européen en Novembre 2018. La Commission Nationale d'Evaluation des Dispositifs Médicaux et des Technologies de Santé (CNEDiMTS) de la Haute Autorité de Santé (HAS) a estimé dans un avis publié le 7 février 2020 que projet de Diabeloop apportait un service attendu suffisant avec un niveau d'amélioration du service attendu modéré (de niveau III), ce qui ouvre la porte d'un remboursement du dispositif par la Sécurité Sociale(71). La population cible estimée est de 7300 patients.

En Avril 2020, Diabeloop a débuté une étude mondiale, en vue d'obtenir l'approbation de la FDA (*Food and Drug Administration*), après avoir réussi à lever près de 30 millions d'euros en Décembre 2019 pour aider à son lancement dans les marchés internationaux (Etats-Unis principalement)(72). Cet essai clinique inclus 180 individus atteints du diabète de type 1, âgés de 14 ans au minimum et doit se clôturer en Octobre 2020(73).

### 5.3. Synapse Médecine

Synapse Médecine a été créé en 2017 dans la région bordelaise par deux docteurs et un ingénieur, en lien avec l'INSERM et le CHU de Bordeaux. Son objectif est de promouvoir le bon usage du médicament, à la fois auprès des professionnels de santé, que du grand public.

Le **mésusage du médicament** est une problématique majeure du système de santé au cours des activités médicales. En effet, ce sont près de 10 à 30 000 décès par an qui sont imputables au mésusage du médicament et 130 000 hospitalisations, dont entre 45 et 70% pourraient être évitées chaque année en France(74). Les personnes âgées de plus de 75 ans, prenant plus de 4 médicaments par jour sont tout particulièrement à risque.

Un essai clinique a été mené aux urgences du CHU de Bordeaux en 2019 sur la pertinence du module d'interactions médicamenteuses de Synapse et une étude sera lancée en 2020 dans le

service gériatrique du CHU de Bordeaux, avant le lancement de l'application au grand public(75).

La plateforme est hébergée en ligne, et est accessible en ligne depuis un ordinateur ou via une application installable sur tablette et smartphone. Elle est dotée d'une première fonctionnalité d'assistant virtuel, qui fonctionne, selon son créateur Clément Goehrs "*un peu comme un Google Assistant du médicament qui peut répondre très précisément à des questions simples comme 'quelle est la posologie de tel médicament ?'*"(76). Le fonctionnement de cet outil repose une fois encore sur l'apprentissage automatique par machine learning afin de comprendre la question posée, et ainsi l'information recherchée. Toujours selon le cocréateur de Synapse Médecine, "*Les algorithmes se mettent à jour toutes les nuits, en consultant la littérature scientifique et en assimilant les dernières informations parues sur le médicament, comme celles publiées sur le site de l'Agence nationale de sécurité du médicament et des produits de santé (ANSM)*"(77).

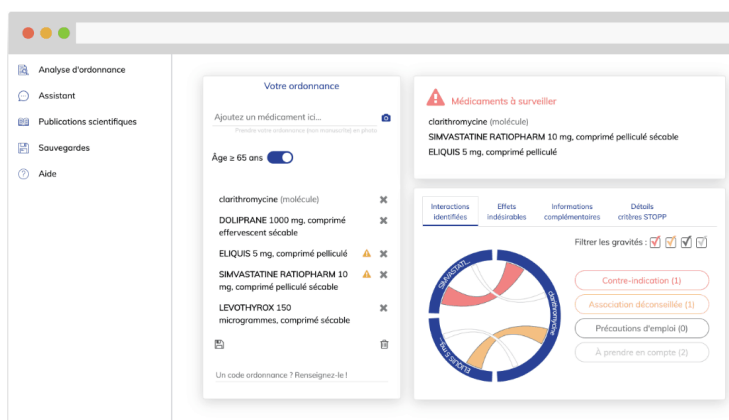


Figure 26 - Module d'analyse d'ordonnances (78)

La deuxième fonctionnalité de l'outil est très prometteuse pour les médecins prescripteurs et les pharmaciens : il permet d'analyser une ordonnance à partir d'une photo prise par un smartphone (cf. Figure 26). Le logiciel détecte instantanément s'il y a des interactions médicamenteuses ou un risque d'effets indésirables trop important et en informe le professionnel, afin d'éviter des erreurs de prescriptions ou de dispensation(76).

## Conclusion

Les principes essentiels de l'IA ont pu être exposés dans cette première partie, afin de mieux comprendre les utilisations possibles à l'avenir. Dans la suite de mon travail, je m'intéresserai plus à la portée industrielle de l'IA : comment cette dernière peut faciliter, et changer le travail des industriels ? Quels sont les enjeux de son utilisation, ses limites ?

La portée de ce travail est donc multiple : ingénieurs, pharmaciens, techniciens, opérateurs... tout le monde est concerné par cette révolution de l'industrie pharmaceutique.



# Partie 2 - Nouveaux défis et Limites de l'IA dans l'Industrie Pharmaceutique

La multiplication de l'offre en médicaments a mené à la création de la législation pharmaceutique que l'on connaît aujourd'hui : l'informatisation des données patients (antécédents, etc...) dans le dossier patient/médical partagé (DPP/DMP) a sensibilisé à la sécurisation des données, la multiplication des sources d'approvisionnement en produits a conduit à la recrudescence de fabrication de contrefaçons et donc pour les industriels à imaginer de nouveaux moyens de maîtrise de ce risque (contrôles, apposition de scellés d'inviolabilité, etc...). Tant de modifications qui montrent la formidable habileté de l'industrie pharmaceutique à s'adapter et intégrer les révolutions qui peuvent la toucher.

## **Quels sont les défis à relever pour intégrer au mieux l'Intelligence Artificielle dans le travail quotidien de pharmacien industriel ?**

L'importance que l'IA va avoir dans l'industrie pharmaceutique du futur ne fait aucun doute et les principaux défis qui en accompagnent la démocratisation sont les mêmes qui touchent les autres industries (automobile, finance, environnement, etc..) :

- Aspects Législatifs
- Aspects Organisationnels
- Aspects Humains
- Aspects Techniques
- Aspects Environnementaux
- Aspects Éthiques

*Un focus est réalisé sur l'industrie pharmaceutique, bien que bon nombre d'enjeux mentionnés dans cette seconde partie soient également applicables à l'ensemble des industries.*

## **1. Point de vue Gouvernemental et Législatif**

### **1.1. Point de vue Législatif**

#### **1.1.1. En France**

Cédric Villani, ancien député La République En Marche et lauréat 2010 de la médaille Fields (plus haute distinction mondiale en mathématiques), a mené entre septembre 2017 et mars 2018 une étude sur l'IA en France à la suite de la demande du premier ministre, Edouard Philippe. Il estime que *"la France a une très grande capacité en recherche en matière d'intelligence artificielle"*(79) ce qui incite des grandes entreprises américaines ou asiatiques à implanter des laboratoires de recherche sur notre territoire.

Il identifie quatre secteurs prioritaires dans lesquels les efforts de développement de l'IA doivent se concentrer : les transports, l'environnement, la défense et la **santé**. Toujours selon M. Villani, les pouvoirs publics français doivent s'adapter rapidement à la démocratisation des usages liés à l'IA et à sa vitesse d'évolution, sous peine d'assister *"impuissant à une*

*reformulation complète des enjeux de santé publique et de pratiques médicales” (79). Enfin, il conclut sur l’importance de l’IA dans le monde de demain : “L’intelligence artificielle est donc une des clés du pouvoir de demain dans un monde numérique”.*

Le chef de l’Etat envisage une **enveloppe de 1,5 milliard d’euros** d’ici 2022 destinés à la recherche et l’innovation en faveur de l’IA et faciliter ainsi la création de start-ups dans ce domaine. Pour lui, les recherches en IA ne peuvent se restreindre à une vision française, mais doivent bien disposer d’une **dimension européenne**, à commencer par une forte collaboration avec l’Allemagne : *“Ce serait vain de le faire dans un seul pays”*(80). Le développement de cette technologie dans l’industrie française repose sur quatre piliers :

- Conforter l’écosystème de l’IA et les talents ;
- Engager une politique d’ouverture des données ;
- Avoir une stratégie à la fois de financement et de projets ;
- Et penser les termes d’un débat politique et éthique.

Un des points phare du discours du Président est la volonté de **simplification administrative** du processus d’innovation, en utilisant des « bacs à sable réglementaires spécifiques à l’IA » permettant de faciliter et d’encourager la mise sur le marché d’idées innovantes. Cela rendra possible de tester et d’obtenir des garanties dans la technologie plus rapidement, en en testant la sécurité et l’efficacité d’utilisation dans un environnement réel. *« Le temps aujourd’hui nécessaire pour tester une innovation en France, qu’il s’agisse d’un algorithme ou d’un médicament, n’est pas toujours et tout à fait celui d’une économie de l’innovation »* ajoute M. Macron (80), qui affirme le besoin de la France d’accélérer la vitesse des délais d’autorisation d’expérimentation.

#### 1.1.2. Dans le Monde

Pour beaucoup de chercheurs, hommes politiques ou économistes, l’IA pourrait représenter la nouvelle course à l’espace. Un chercheur en IA, Yoshua Bengio, a déclaré à l’UNESCO (Organisation des Nations unies pour l’éducation, la science et la culture) en 2018 : *“les économistes prévoient que d’ici à une dizaine d’années, [les produits de l’IA] atteindront jusqu’à 15 % de la production totale des biens”* (81).

Vladimir Poutine, Président de la Fédération de Russie, a estimé en 2017 dans Russia Today que *“l’intelligence artificielle représente l’avenir, non seulement pour la Russie, mais aussi pour l’ensemble de l’humanité [...] La personne qui prendra les rênes dans ce domaine **dirigera le monde**”*(82).

Les premiers à avoir établi une stratégie nationale en matière d’IA ont été le Japon et le Canada, dès 2017. Bien d’autres pays ont depuis fait de même, mais selon Tim Dutton, un chercheur en sciences politiques, *“aucune stratégie ne se ressemble, chacune se focalise sur différents aspects de la politique de l’IA : recherche scientifique, développement des talents...”*(81). Lors de l’établissement des stratégies nationales, il est primordial de réaliser des analyses des effets

socio-économiques de l'IA afin de mettre en place des mesures permettant d'assurer la responsabilisation, la transparence ou la protection des données.

Des réglementations à la fois nationales mais aussi (et surtout) internationales devront exister afin que l'IA ne devienne pas “*un agent de l'exploitation et du non-respect de l'être humain et de l'environnement*”(81). L'Université de Stanford, aux Etats-Unis, a créé un “incubateur de politiques numériques” afin de favoriser la collaboration internationale en vue d'établir des normes communes assurant la sécurité, le développement économique et le droit humain.

La part de l'IA dans l'industrie ne va cesser d'augmenter au cours des prochaines années, mais leur encadrement juridique actuel n'est pas pour autant adapté. Que cela soit en termes de responsabilité ou d'éthique, aucune disposition légale n'existe encore à ce jour (83).

## 1.2. Établir des responsabilités

Une des principales problématiques soulevées par les gouvernements est celle des **responsabilités**. En cas de défaillance d'un système d'IA, qui est responsable ? L'entreprise ayant commandé le programme, celle l'ayant réalisée ou une tierce personne ? Cette question très complexe ne dispose pas de réponse à l'heure actuelle : aucune règle ou loi ne légifère encore cela.

*Exemple : Si un programme d'IA utilisé dans l'industrie pharmaceutique dysfonctionne, et conduit à la libération d'un lot de produits pharmaceutiques non conformes sur le marché, à l'origine d'un scandale sanitaire : comment établir les responsabilités ?*

De plus, les technologies d'IA incluent les notions d'apprentissage autonomes, qui peuvent les conduire à disposer de nouvelles fonctions, parfois totalement opposées à celles pour lesquelles elles avaient été conçues à la base, sans que le concepteur n'en ait eu la volonté. Ainsi se pose la question : quelles sont les innovations souhaitables et celles qui ne le sont pas ? Et de fait est ce que le créateur d'un programme peut être tenu responsable d'utilisations qu'il n'aurait pas imaginé ?

*Par exemple un programme se servant d'une base de données massive relative aux liens entre malformations humaines et dysfonctionnements génétiques à partir de photos de patients pourrait à la fois être utilisé par un médecin afin d'établir un diagnostic, mais également par des compagnies d'assurance pour effectuer de la discrimination envers certaines catégories de clients “à risque”.*

Chaque entreprise devra désigner des responsables légaux, en contact avec les autorités, et s'assurant de la maîtrise des utilisations de l'IA. Parallèlement à cela, un Délégué à la Protection des données sera nommé pour surveiller l'utilisation des données le cas échéant. Les responsabilités sont donc partagées au sein des entreprises (manufacturières comme sous-traitantes).

C'est pourquoi des politiques sont nécessaires pour énoncer clairement les responsabilités de chacun et tracer le périmètre d'application de l'IA dans tous les domaines, afin d'en éviter les usages abusifs.

### 1.3. Propriété intellectuelle et Brevets

Un point également très important dans le monde industriel pharmaceutique : la propriété intellectuelle et les brevets.

#### 1.3.1. Définitions

##### 1.3.1.1. La Propriété Intellectuelle

Selon l'Institut national de la statistique et des études économiques (INSEE), la **propriété intellectuelle** (PI) permet de protéger le propriétaire d'une invention de toute utilisation non autorisée par un tiers (contrefaçon, plagiat...) et de tirer les bénéfices de cette invention (financiers, technologiques...) (84). Elle inclut en son sein les notions de propriété industrielle (brevets) et de droit d'auteur.

##### 1.3.1.2. Les Brevets

La propriété industrielle est obtenue à la suite du dépôt d'un Brevet, qui offre un monopole dans l'exploitation d'une découverte ou invention : nouvelle molécule, nouvelle technologie de formulation, etc... L'Institut National de la Propriété Industrielle (INPI) définit comme brevetable toute innovation qui est :

- Nouvelle (elle ne doit pas correspondre à une innovation déjà publique) ;
- Susceptible d'application industrielle ;
- Inventive (elle ne doit pas être évidente) ;
- Et licite(85).

*Une problématique apparaît pour les **logiciels** : tous ne sont pas brevetables. En effet, un logiciel permettant "seulement" d'automatiser informatiquement des activités déjà faites en routine manuellement ne peut pas être considéré comme une innovation inventive, car évidente pour un Homme du métier concerné. Des systèmes de surveillance de procédés industriels, ou de traitement de données physiques (température, taille, ...) sont considérés comme relevant d'un caractère technique, et sont donc brevetables (86).*

Ainsi les enjeux pour les laboratoires pharmaceutiques sont critiques : breveter une molécule ou une nouvelle technologie offre un monopole pendant 20 ans pour le titulaire, qui peut profiter des retombées financières, avant que cette découverte ne tombe dans le domaine public et que d'autres laboratoires soient autorisés notamment à produire des médicaments génériques utilisant cette molécule.

L'industrie pharmaceutique a toujours considéré les Brevets comme la méthode de protection des innovations la plus forte. Leur nombre lié à des technologies d'IA est en constante augmentation depuis quelques années.

*La société AiCure (mentionnée en partie 3) a par exemple déposé un Brevet en Février 2020 pour protéger une méthode et un appareil de vérification de l'observance médicale (87).*

Mais certaines de ces découvertes sont considérées comme étant des idées abstraites donc non brevetables. Comment les protéger ?

### 1.3.2. Considérations Légales

#### 1.3.2.1. IA, Innovation et Collaboration

Au vu de l'importance capitale que relèvent ces notions, elles sont **très réglementées**. De la même manière que l'IA change le visage de l'industrie pharmaceutique, de nombreux industriels se posent la question suivante : à qui appartient la découverte, quand elle provient d'un programme d'IA ?

En effet, par sa capacité à analyser de très vastes bases de données en des temps records, l'IA prend une part de plus en plus importante dans la Recherche et le Développement (R&D) de nouveaux médicaments : dans l'identification de gènes, de cibles moléculaires, ou de nouvelles molécules...

A ce jour, il est très rare que les processus de R&D soient effectués par une seule personne. Une **collaboration** existe la plupart du temps entre un laboratoire et un partenaire propriétaire d'une solution d'IA. Ce dernier "loue" sa solution à l'industriel, afin qu'il l'utilise sur ses données propres. Deux situations peuvent dès lors exister :

- La première correspond à une collaboration traditionnelle entre un client et un prestataire. Celui-ci offre une plateforme technologique et un conseil adapté à l'usage de son client. Ni plus, ni moins.
- La seconde, change du modèle traditionnel. Le prestataire en IA peut également souhaiter effectuer sa propre R&D à partir des données de ses clients, et conserver les droits d'utilisation des découvertes issues de cette collaboration.

Ainsi la question de la propriété intellectuelle se pose réellement ici : qui est propriétaire d'une découverte faite à la suite d'un tel partenariat ? L'industriel ayant commandé la solution, ou le prestataire l'ayant fourni ?

De nouvelles tensions apparaissent dans l'analyse de telles collaborations. Le modèle entraîné obtenu est une amélioration certaine de celui initialement fourni par le partenaire en IA. Cependant, il repose indubitablement sur les données du client pharmaceutique. Le partenaire en IA peut-il réclamer la PI de cette découverte et ses données dérivées, ou, être autorisé à utiliser le modèle entraîné et les données en dehors de cette collaboration, sans n'avoir aucune obligation envers l'industriel pharmaceutique ?

**Aucune règle précise n'existe à l'heure actuelle**, et l'éventuel partage de PI (ou l'exclusivité de sa jouissance) repose uniquement sur une **négociation commerciale** entre les deux parties, et l'établissement d'un contrat (88). Certains accords relatifs aux données d'entrée, ou aux découvertes restreignent l'usage de la technologie d'IA à la collaboration initiale. Cela empêche dès lors le partenaire pharmaceutique de bénéficier des nouvelles données ou des améliorations du modèle artificiel obtenu grâce à des données provenant d'autres collaborations du prestataire avec des industriels.

Dans d'autres situations, des logiciels d'IA sont directement utilisés pour créer des molécules actives. Ainsi, la découverte serait ici entièrement imputable à une machine, et non à un être humain. Cependant, pour être protégée par les droits d'auteur, une œuvre doit être **originale**, donc refléter la personnalité de son créateur, et disposer de la **personnalité juridique** (c'est-à-dire disposer de droits et de devoirs). A ce titre, la loi française ne permet pas (à l'heure actuelle) à un logiciel d'IA d'être propriétaire d'une découverte, et donc protégée par les droits d'auteur (89).

L'existence de ce **vide juridique** force la législation française à s'adapter, en précisant le statut juridique des IA en France. Le Parlement Européen propose de créer une **personnalité juridique "électronique"** propre aux IA, les rendant responsables pour tout dommage causé. Mais beaucoup sont contre, et trouvent que cette solution soulèverait des problèmes éthiques.

#### 1.3.2.2. La protection des Algorithmes et autres solutions d'IA

Comme énoncé précédemment, les modèles d'IA sont capables d'apprendre et de s'améliorer en temps réel en étudiant de larges bases de données, dont la principale caractéristique est d'être en **perpétuelle évolution**. Cette inconstance pose des problèmes de PI. En effet, un algorithme évoluant constamment peut ne pas être protégé s'il diffère de sa description dans le dossier de dépôt de Brevet. De plus, de tels algorithmes peuvent être considérés comme des processus mentaux, donc **non brevetables**.

Puisque les algorithmes ne sont pas brevetables pour l'instant, l'industrie pharmaceutique doit avoir une approche diversifiée sur la protection de la PI, incluant non seulement les brevets, mais aussi les droits d'auteurs et les secrets d'affaires.

La **protection des droits d'auteur** peut être appliquée aux codes sources et aux bases de données, mais cela nécessite l'existence d'une notion de parenté humaine pour ces œuvres. Des codes ou bases de données créées par l'Homme pour être utilisées par des applications d'IA pourront être protégées par le droit d'auteur, tandis que de telles œuvres entièrement créées par l'IA ne le seraient pas.

La **protection des secrets d'affaires** peut, elle, être adaptée aux technologies évoluant rapidement, et difficiles à décrire. Au contraire du Brevet, avec les secrets d'affaires, la concurrence n'est pas au courant de la découverte. Il est légiféré depuis décembre 2018, avec le décret d'application n° 2018-1126 mais exige tout de même certaines garanties



et pose certaines questions : Comment faire lorsque le secret d'affaires est créé par une IA ? L'entreprise qui possède l'IA possède-t-elle le secret d'affaires ?

*Dans le cas d'une entreprise pharmaceutique qui déposerait un Brevet sur une formulation pharmaceutique, la formule complète serait rendue publique, bien que son exploitation par des tiers serait interdite. Les autres laboratoires concurrents auraient connaissance de cette formule, et devront attendre que le Brevet ne tombe dans le domaine public pour l'exploiter. En gardant la formule et la méthode de fabrication secrète, sans déposer de Brevet (le Coca-Cola par exemple), seules les personnes choisies seraient au fait de la formule et de son caractère confidentiel.*

### 1.3.2.3. L'adaptation de la Législation

Le Bureau Américain des Brevets et des Marques (*United States Patent and Trademark Office*) a récemment proposé d'établir des nouvelles formes de protection de la PI pour les innovations liées à l'IA et d'adapter la législation américaine à ces nouvelles problématiques. Des juristes français ont également commencé à réfléchir à des ajustements de lois existantes et à la création de nouvelles pour encadrer ces nouvelles pratiques.

Il est donc certain que le paysage législatif français, européen, et mondial va devoir s'adapter et trouver des solutions aux nombreuses questions que soulève l'importance accrue des solutions d'IA dans les découvertes.

## 1.4. Limites

### 1.4.1. Inégalités entre les différents pays

Des observateurs suggèrent que la capacité à innover (en particulier dans l'IA) et d'acquérir les compétences humaines nécessaires seront parmi les catalyseurs les plus importants influençant la **croissance future du Produit Intérieur Brut (PIB)**. Selon le cabinet britannique PwC, l'IA pourrait permettre au Produit Intérieur Brut mondial d'augmenter de 14% d'ici 2030 (90). Tous les pays ne sont pas au même niveau dans leur utilisation de l'IA ou dans la préparation aux changements qu'elle va entraîner :

- Les leaders dans ce domaine sont la Chine et les Etats-Unis, responsables pour la grande majorité des recherches dans ce domaine.
- Le second groupe rassemble des pays tels que l'Allemagne, le Japon, le Canada, le Royaume-Uni ou la France, en retrait.
- Un troisième groupe concerne les pays en retard vis-à-vis des autres, bien qu'ils disposent de potentiels dans certains domaines précis : l'Italie, le Brésil ou l'Inde.

D'autres pays n'ont pas encore débuté leur transition, principalement les pays en développement ou en Afrique, par manque de moyen ou d'intérêt. Ils sont plus souvent des terrains d'expérimentation, des fournisseurs de données ou des consommateurs d'IA, que des véritables parties prenantes dans la recherche en IA.



La France a donc du retard sur d'autres pays, du fait des faibles budgétaires allouées à la recherche, la dissémination des équipes, l'absence de politique cohérente ou la faiblesse des salaires en comparaison d'autres pays. Retard qu'il est primordial de combler, au vu de l'importance que va avoir l'IA dans l'ensemble des métiers, activités économiques et industrielles.

#### 1.4.2. Aggravation de la Fracture Numérique entre les pays

Le principal risque découlant de cette inégalité d'investissement dans l'IA est **d'aggraver la fracture numérique** entre les pays. Elle correspond à des inégalités d'accès et d'utilisation des technologies de l'information et de la communication (TIC), faute d'accès aux équipements, ou de formation pour les maîtriser. Mais les pays en développement ne sont pas les seuls concernés : ceux industrialisés peuvent également prendre du retard.

En effet, 93% des brevets en IA étaient détenus en 2014 par des multinationales basées dans 7 économies, avec par ordre décroissant : le Japon, la Corée du Sud, les Etats-Unis, Taiwan, la Chine, l'Allemagne, et la France (qui détenait 2% de l'ensemble des brevets)(81).

#### 1.4.3. Difficultés dans l'établissement de normes ou de règles

Puisque les pays ne placent pas la même importance dans l'IA, il est très complexe d'établir des règles communes permettant de normaliser les législations. Il est pourtant certain *“qu'une gouvernance claire doit être établie afin d'encadrer la conception, l'exploitation et la maintenance des modèles”* selon Jean-David Benassouli (91), responsable de l'activité Data Intelligence, chez PwC France, un cabinet d'audit. A l'heure actuelle, aucune norme mondiale n'existe encore, bien que ce projet soit en gestation depuis quelque temps.

## 2. Point de vue Ethique

### 2.1. Définition

Selon le dictionnaire Larousse, l'Éthique est la *“partie de la philosophie qui envisage les fondements de la morale. Ensemble des principes moraux qui sont à la base de la conduite de quelqu'un”* (92).

Rob Kitchin, professeur en sciences sociales, définit dans son œuvre *“The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequence”*, le terme d'**Éthique des Données** comme la *“défense et la recommandation de concepts de bonne ou de mauvaise conduite en relation avec les données, et surtout les données personnelles”* (93). C'est cette éthique qui concerne principalement l'IA.

### 2.2. Ethique & IA

La notion d'Éthique est **indissociable de l'IA** et est au cœur de toutes les problématiques actuelles. L'IA va être amenée à “manipuler” un grand nombre de données (personnelles ou non), et les analyser avec des capacités dépassant de loin celles des Humains. Il n'est pas possible de faire tout et n'importe quoi avec celles-ci. Elle va également utiliser des logiciels

amenés à prendre des décisions à la place d'Hommes (ou tout du moins à l'aider dans sa prise de décision)

Cédric Villani estime dans son rapport qu'il *“faut une instance pour émettre des avis, donner des jugements en toute indépendance, qui puisse être saisie par le gouvernement comme par les citoyens, et qui nous dise ce qui est acceptable ou non acceptable”* (79).

Emmanuel Macron annonce dans son discours d'octobre 2019 qu'il *“faudra mettre à la fois de la transparence et de la loyauté dans le système. La clé est de mettre partout de la transparence publique sur les algorithmes”* (80). Le gouvernement a également pour projet de définir des normes pour le contrôle et la certification des algorithmes au niveau local, européen et mondial.

La Directrice Générale de l'UNESCO, Audrey Azoulay, a annoncé que *« les problématiques posées par l'intelligence artificielle ne sont pas d'ordre technologique. Elles touchent à notre propre humanité et soulèvent des questions scientifiques, politiques, philosophiques et éthiques [...]. Il est plus que temps de définir les principes éthiques qui devront servir de fondement et de cadre à ce bouleversement, pour garantir que l'IA soit au service des choix collectifs, en fonction de valeurs humanistes »*(81). La Commission mondiale d'éthique des connaissances scientifiques et des technologies (COMEST) de l'UNESCO s'est donnée pour objectif d'énoncer des règles visant à fixer les bases éthiques de ces technologies. Ainsi des **bases éthiques communes** seront les fondations des utilisations de l'IA de demain.

### 2.3. Augmenter la transparence et l'auditabilité des systèmes

Les systèmes utilisés sont très “opaques” : le résultat est certes visible, mais le cheminement ne l'est pas. Or l'Humain souhaite comprendre comment il fonctionne, pourquoi il donne un tel résultat. L'IA est en ce sens souvent comparée à une boîte noire : on voit les entrées et les sorties mais pas ce qu'il y a entre.

L'objectif est **d'accroître la transparence des systèmes**, afin de pouvoir en expliquer les décisions prises, et de démontrer que l'Humain conserve la maîtrise dessus. Il faut assurer l'équité, et l'absence de discriminations et de biais dans la prise de décisions.

Des prestations **d'audit des systèmes d'IA** et de leur intégration au sein des entreprises pharmaceutiques devront être développées, afin d'assurer leur conformité avec le Règlement Général sur la Protection des Données (RGPD) et les autres normes qui régiront les systèmes informatiques (*annexe 11 des Bonnes Pratiques de Fabrication ou 21 CFR art 11*). Les algorithmes en eux-mêmes seront audités, mais aussi la façon dont ils sont intégrés et utilisés dans les organisations.

Il sera nécessaire d'analyser minutieusement les complications relatives à la confidentialité des projets d'IA et s'assurer que les responsabilités réglementaires et de transparence sont bien établies. Des lignes directrices relatives à l'éthique dans le monde pharmaceutique devraient être publiées dans les mois à venir pour guider les industriels dans cette transition.

## 2.4. Démocratiser le débat sur les usages de l'IA

Afin de conforter la population et d'en conserver la confiance, intégrer toutes les parties prenantes dans l'établissement des règles est nécessaire. Les décisionnaires devront être transparents et justifier les règles de confidentialité, d'usages qui auront été décidées.

Une gouvernance spécifique de l'éthique en IA va être établie, grâce aux travaux du comité consultatif national d'éthique pour les sciences de la vie et de la santé (CCNE). Son objectif principal est de *“faire participer les citoyens à la réflexion éthique et leur permettre de comprendre les enjeux éthiques que soulèvent certaines avancées scientifiques dans le domaine des sciences de la vie et de la santé”* (94).

Deux principes de base ont été édictés par la Commission nationale de l'informatique et des libertés (CNIL) pour le développement des algorithmes et de l'IA : la **vigilance** et la **loyauté**. Tous les acteurs de la chaîne “algorithmique”, concepteurs, utilisateurs professionnels ou grand public, doivent être formés à l'éthique. Ce thème devra également être renforcé au sein des entreprises, par la création de comités d'éthique, l'établissement de bonnes pratiques éthiques ou l'adaptation de chartes déontologiques...

L'éthique aura donc une place centrale dans l'industrie de demain, de sa conception à son utilisation finale.

## 2.5. Le Règlement Général sur la Protection des Données (RGPD)

L'Europe a été pionnière dans la législation de la gestion des données, avec le Règlement Général sur la Protection des Données, qui a introduit des consignes plus strictes sur la collecte de données, donné le droit aux utilisateurs de décider ce que les entreprises peuvent faire avec les données collectées et introduit de très lourdes amendes en cas de non-respect.

### 2.5.1. Définition et périmètre d'application

#### 2.5.1.1. Définitions

Le Règlement numéro 2016/679, ou **Règlement Général sur la Protection des Données** (ou *General Data Protection Regulation, GDPR* en anglais) a été adopté en 2016, effectif depuis 2018. Applicable sur l'ensemble du territoire européen, son objectif est d'encadrer le traitement des données personnelles par des professionnels. Il permet de renforcer *“le contrôle par les citoyens de l'utilisation qui peut être faite des données les concernant”* et d'harmoniser *“les règles en Europe en offrant un cadre juridique unique aux professionnels”* (95) afin d'obtenir la confiance des utilisateurs.

La CNIL définit deux notions primordiales sur son site internet.

Une **donnée personnelle** qui est *“toute information se rapportant à une personne physique identifiée ou identifiable”* (96) directement (nom, prénom) ou indirectement (numéro de téléphone, identifiant...).

Un **traitement de données personnelles**, soit une “*opération, ou ensemble d'opérations, portant sur des données personnelles, quel que soit le procédé utilisé*” (96). Sa finalité est multiple : collecte, conservation ou communication d'informations.

Les objectifs d'une telle pratique doivent être légaux, légitimes et clairement définis *a priori* : il n'est pas permis de conserver des données dans l'attente d'en avoir le besoin dans le futur.

#### 2.5.1.2. Périmètre d'application

Ce règlement concerne **toutes les organisations**, qu'elles soient de nature publique ou privée, de petite ou grande taille, qui traitent des données personnelles, à partir du moment où elles sont implantées sur le territoire européen ou que les données étudiées proviennent de citoyens européens. Toutes les activités sont concernées, y compris les organismes faisant de la sous-traitance pour d'autres.

#### 2.5.2. Les modifications pour les industriels

Toutes les organisations concernées par ce règlement doivent y être conformes depuis 2018 (97). Elles doivent désormais se référer à **l'autorité de protection des données de l'État membre**, unique interlocuteur à ce sujet. Désormais, plusieurs obligations touchent les responsables de traitement :

- Ils doivent limiter au maximum le nombre de données traitées ;
- Assurer la protection (sécurité et confidentialité) des données dès la conception et en continue ;
- Notifier les individus en cas de fuite de données ;
- Et pouvoir démontrer leur conformité au règlement à tout moment.

Ils devront également mettre en place des analyses d'impact relatives à la protection des données (AIPD), permettant de faire ressortir les caractéristiques, ou les risques du traitement de données. Si des risques sont identifiés, l'entreprise devra consulter son autorité de protection des données, et ce, avant toute mise en œuvre du traitement. Cette dernière pourra refuser l'exploitation des données.

Le CNIL énonce 6 points primordiaux à destination des industriels : ne collecter que les données **pertinentes**, être **transparent** dans l'utilisation qui en est faite, respecter le **droit des personnes**, garder la **maîtrise** des données, identifier les **risques** et **sécuriser** les données (98).

#### 2.5.3. Les sanctions en cas de non-conformité

Tout organisme qui n'est pas conforme au RGPD s'expose à de **lourdes sanctions**, qui vont d'un simple avertissement à une amende, pouvant atteindre jusqu'à 20 millions d'euros ou 4% du chiffre d'affaires (CA) annuel mondial pour une entreprise (97).

L'Europe a été une vitrine en matière de protection des données avec le RGPD, les Etats-Unis ont par la suite signé en 2018 un accord avec le Canada et le Mexique se calquant sur les

principes européens. A l'heure actuelle, ce sont plus de 100 pays qui disposent d'une législation sur le traitement des données (99).

## 2.6. Limites

Le risque de détournement et d'usage malveillant des données est réel (100). Certaines données capitales d'entreprises (procédés ou recettes de fabrication, données de production, etc..) pourront être sujettes à des attaques à grande échelle mettant en péril leur fonctionnement : espionnage industriel, campagnes de désinformation ciblées, etc...

C'est pourquoi il est primordial d'adapter la protection des droits et des libertés au regard des abus potentiels et de surveiller l'aspect éthique de l'utilisation des données. Il faut avoir conscience des risques dès l'établissement des normes, pour garder la main sur ces technologies.

## 3. Point de vue Organisationnel

### 3.1. Stratégie de l'entreprise

L'IA promet des bénéfices réels et accessibles à toutes les entreprises pharmaceutiques qui seront prêtes à l'adopter. La réussite de son intégration au sein des entreprises pharmaceutiques repose avant tout sur la **stratégie et le positionnement** que chacune d'entre elles va adopter.

De plus en plus d'entreprises pharmaceutiques (*Merck, Pfizer et Sanofi notamment*) investissent (101) dans ces technologies de demain et modifient leur stratégie, afin d'inclure la transition numérique dans leurs opérations, en désignant des directeurs numériques. Leur rôle est de lier le développement des technologies numériques et le Big Data, au développement des médicaments et aux activités du quotidien. Assez logiquement, les plus grands laboratoires pharmaceutiques sont ceux investissant le plus dans le domaine de l'IA. Mais plutôt que d'investir directement dans ces technologies, ils priorisent des achats ou partenariats avec des start-up proposant des solutions innovantes. Une étude a été réalisée par Forbes en Juillet 2020 (*cf. Figure 27*) afin d'étudier les activités liées à l'IA par les plus grandes entreprises pharmaceutiques mondiales : partenariat avec des start-ups, investissements en R&D, publications scientifiques. Novartis ressort comme le leader des investissements internes, tandis que AstraZeneca est le leader en termes de publications.

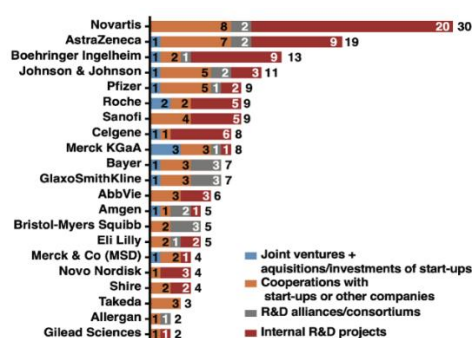


Figure 27- Activités liées aux technologies de l'IA dans l'industrie pharmaceutique (102)

## Pourquoi investir dans l'IA ?

Les technologies d'IA sont idéales pour effectuer des tâches répétitives, touchées par des manques d'efficacité, et ainsi améliorer les performances des entreprises. Ses applications seront possibles sur l'ensemble du cycle de vie du produit (maintenance prédictive améliorée, optimisation de la chaîne logistique ou de la relation clientèle, etc...). *Ces mêmes applications seront détaillées dans la troisième partie de ce document.*

Leur intégration dans l'industrie pharmaceutique découle de trois facteurs : l'accès facilité aux données, la croissance du nombre de produits et de solutions d'IA et la puissance de calcul disponible. En moyenne, les entreprises disposent d'un **délai de trois ans** pour intégrer l'IA sans prendre le risque de perdre de la part de marché au profit de concurrents nouveaux ou existants dans leur domaine. Pour simplifier, plus les entreprises adoptent l'IA tôt et plus leurs profits vis-à-vis de leur concurrence augmente.

Une étude réalisée en 2017, *Artificial Intelligence for the Real World*, a interrogé 250 managers industriels sur les bénéfices qu'ils espèrent tirer de l'IA dans leurs entreprises (103). La moitié d'entre eux attendent une amélioration notable des fonctionnalités et des performances de leurs produits ou services, un tiers une optimisation des opérations internes, une meilleure prise de décision ou la hausse de la créativité des employés en les déchargeant de tâches à faible valeur ajoutée. Ainsi, en simplifiant au maximum, la finalité de l'IA dans l'industrie pharmaceutique sera (comme dans toutes les autres industries) d'assurer un gain de temps, d'argent et donc de productivité.

### 3.1.1. Agilité de l'entreprise

Pour implémenter des nouveaux outils ou méthodes technologiques, l'approche recommandée est **l'agilité**. Elle correspond à la mise en place d'actions testées et approuvées successivement afin de faciliter l'adoption des changements et d'améliorer l'efficacité et la rapidité des cycles de renouvellement des produits ou des méthodes. Plutôt que d'implémenter des changements de façon brutale et totale, une mise en place incrémentale permet d'éduquer de façon plus efficace et plus "douce" la force de travail.

Des boucles (*scrums*) successives (cf. Figure 28) d'une durée prédéterminée seront utilisées pour tester des méthodes, des réunions quotidiennes seront organisées afin de modifier ce qui doit l'être, et une réunion finale sera planifiée pour conclure sur les bons et mauvais points de la méthode, afin de la perfectionner.

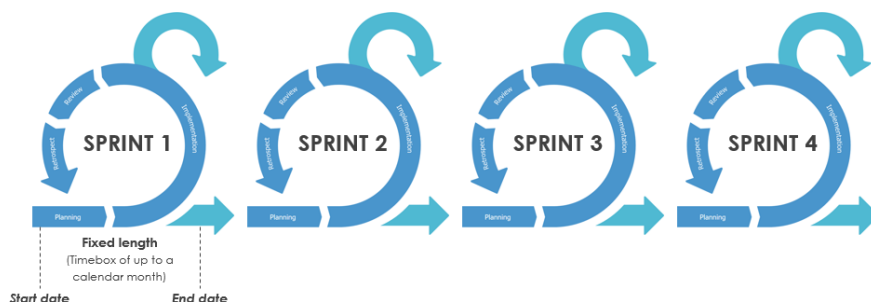


Figure 28 - Schéma de l'agilité d'une entreprise (104)



Une entreprise agile intègre l'ensemble de ses employés dans le processus de changement, ce qui, dans le cas de l'IA, est une formidable opportunité pour comprendre ces nouveaux outils, qui peuvent être effrayants au premier abord.

L'Humain a une place primordiale dans cette “révolution technologique” pour les entreprises. En effet, l'évolution est constante, et les méthodes seront amenées à s'y adapter, les salariés doivent donc être préparés à ces changements, pour y adhérer. Les entreprises sont obligées d'intégrer la transition digitale dès à présent dans leur stratégie pour prendre un avantage concurrentiel certain. Plus un changement est anticipé, plus la méthodologie agile fera sens : les modifications seront étudiées, testées, et perfectionnées.

*A contrario, une entreprise n'ayant pas investi dans l'IA et se rendant compte de son erreur devra mettre en place un grand nombre d'actions en peu de temps pour rattraper son retard vis-à-vis de la concurrence, ce qui sera brutal et traumatisant pour ses employés.*

Le déploiement de ces nouvelles technologies n'est donc pas réalisable du jour au lendemain, il nécessite une **vision à long terme et globale** de l'entreprise. Cette dernière doit également s'assurer que sa philosophie est appliquée à tous les niveaux hiérarchiques.

### 3.1.2. Collaboration

Bien heureusement, les entreprises ne sont pas seules dans cette démarche de changement. Il est primordial d'inclure dans la stratégie de transition numérique l'établissement de **collaborations** avec des organisations spécialisées dans l'IA, qui pourront leur fournir leur expertise et leur expérience dans l'utilisation de ces outils. Des partenariats avec des équipes de recherche universitaires devront également être envisagés, ainsi que la promotion de l'expertise au sein des entreprises par la mise à disposition de ressources adaptées aux équipes, afin de faire croître l'utilisation de ces méthodes dans l'industrie pharmaceutique

Le guide des Bonnes Pratiques de Fabrication (BPF) inclut déjà dans son annexe 11 relative aux systèmes informatisés la notion de partenariat entre entreprises. L'ANSM stipule qu'un “*contrat formel doit être établi dès lors que le fabricant fait appel à un tiers [...] qui interviendrait dans [...] l'installation, la configuration, l'intégration, la validation, la maintenance, la modification ou la conservation d'un système informatisé*” (105). Ce contrat doit impérativement définir les responsabilités de chacune des parties. Comme pour toute collaboration (comme avec les fournisseurs ou sous-traitants), ces tierces parties doivent être **évaluées** à échéance régulière pour s'assurer de leur conformité et toute la documentation conservée doit être disponible à tout moment.

Une étude réalisée par *CB Insights* en Novembre 2019 (106), a montré que le nombre de partenariats entre des entreprises de santé et des start-ups d'IA a atteint un niveau record au troisième trimestre 2019, avec une centaine d'accords signés pour un montant record d'1.5 milliard de dollars (cf. *Figure 29*).



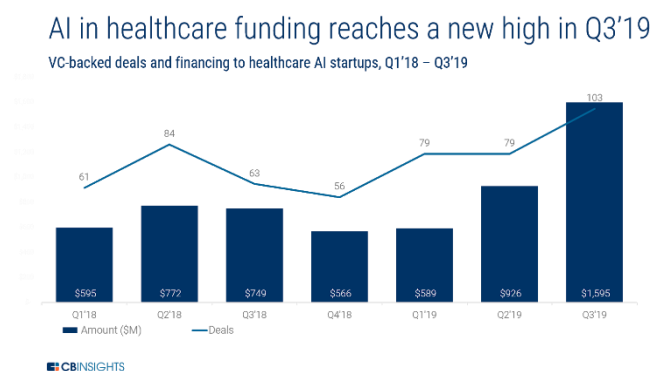


Figure 29 – Evolution du nombre de partenariats et leur valeur (106)

Il est certain que cette collaboration bénéficiera à l'ensemble des parties prenantes, mais elle présente quelques désavantages. Si les utilisateurs et l'entreprise ne disposent pas des connaissances et compétences pour élaborer la solution eux-mêmes, une **dépendance** envers l'outil qu'ils s'offriront naîtra. Il est le plus souvent moins onéreux pour une entreprise de créer ses propres solutions plutôt que de les acheter à un prestataire. Cela permet également d'affûter les compétences de ses salariés en créant de la valeur.

### 3.1.3. Challenges liés à l'intégration de l'IA dans les entreprises et gestion des risques

De nombreux responsables d'entreprises annoncent avoir des difficultés dans l'inclusion de technologies d'IA au sein de processus ou de systèmes préexistants. Toutes les entreprises pharmaceutiques n'auront pas les mêmes difficultés à surmonter lors de l'intégration de l'IA dans leurs procédés. Les producteurs de médicaments seront très fortement touchés, du fait du grand nombre de tâches automatisables dont ils disposent, tandis que les grossistes-répartiteurs par exemple devraient être moins affectés par cette transition numérique.

Les challenges liés à l'intégration de l'IA dépendent de nombreux facteurs, dont la **préparation technique** (disponibilité des systèmes et données, etc) et la **préparation organisationnelle** (la capacité et culture du changement de l'entreprise et sa capacité à utiliser ces nouveaux outils). Cette dernière apparaît comme la plus importante. La transition numérique ne se fera pas du jour au lendemain et doit s'effectuer en douceur : de nombreux procédés seront entièrement modifiés, afin d'intégrer de nouvelles contraintes, tester ces nouveaux procédés et solutions, et s'assurer que le résultat final est inchangé. Les outils d'IA doivent être mis en place, paramétrés, utilisés et maintenus correctement pour être bénéfiques à l'entreprise.

Réussir sa transition passera par l'intégration de l'ensemble du personnel dans le processus et de leur compréhension des modèles, à la fois à leur installation mais surtout dans leur utilisation quotidienne. Une intégration stratégique bien pensée et réfléchie est moins sujette au risque d'échec.

Comme pour toute autre technologie, la **gestion des risques** doit être appliquée à l'ensemble du cycle de vie du système informatique, afin d'assurer la qualité du produit final, l'intégrité des données et la sécurité de l'utilisateur final le cas échéant. Selon l'annexe 11 des BPF, "les décisions relatives à l'étendue de la validation et aux contrôles d'intégrité des

*données doivent être basées sur une évaluation justifiée et documentée des risques liés au système informatisé*” (105). L'accès des systèmes informatiques doit être restreint au seul personnel qualifié et autorisé, par l'intermédiaire de moyens physiques ou logiques.

Tous les incidents doivent être déclarés, évalués et investigués. Cela concerne à la fois ceux liés à des défaillances des systèmes d'IA et ceux relatifs aux erreurs de données. Comme pour tout type d'incident, l'enquête doit permettre d'identifier la cause racine et mener à des actions préventives et correctrices.

### 3.2. Connectivité universelle

Un grand nombre de données sera amené à être échangé au sein des organisations. Il faudra donc disposer d'infrastructures numériques fournissant une connectivité universelle. La 5G disposera de vitesses de téléchargement encore inégalées, rendant possible le partage quasi-instantané de grandes quantités de données.

Les technologies cloud (informatique en nuage) permettent de stocker en ligne des données, des applications ou des logiciels, qui peuvent être mis à la disposition de tous ceux qui le veulent, instantanément et peu importe leur localisation. Le cloud et la 5G seront ainsi à la base d'une utilisation de l'IA efficace et économiquement viable dans l'industrie pharmaceutique.

### 3.3. Gestion des coûts

Le coût de mise en place de la stratégie “Pharma 4.0” au sein des usines pharmaceutiques a été cité comme un frein, mais celui-ci sera très vite compensé par la hausse de la productivité, la diminution des temps morts et des rejets. Alasdair Gilchrist parle du “pouvoir des 1%” dans son livre *Industry 4.0: The Industrial Internet of Things* : une entreprise a seulement besoin de **réduire ses dépenses opérationnelles ou sa non-productivité d'1%** en utilisant l'industrie 4.0 pour faire des économies significatives (107). Ainsi, si le taux de rebut d'une entreprise arrive à être drastiquement réduit, les bénéfices en découlant absorberont totalement les frais engagés dans les solutions utilisées.

Les solutions d'IA à destination de l'industrie pharmaceutique vont offrir des bénéfices économiques à long terme à ses utilisateurs liés à des gains de productivité (par la baisse des coûts opérationnels automatisés, l'optimisation des flux de production, ou l'automatisation de la coordination entre différents services et acteurs, ...). L'investissement initial sera certes conséquent, mais s'il est planifié, il pourra s'étaler sur une période plus longue et donc être plus aisément absorbé.

### 3.4. Qualification & Validation

La politique très stricte de validation des procédés dictée par les autorités de régulation est indispensable pour assurer la sécurité des destinataires finaux des produits : les patients. Un des principaux challenges pour le monde pharmaceutique est d'adapter les usages des technologies de l'IA à la très stricte régulation qui l'entoure. Ce sont ces considérations réglementaires qui peuvent freiner certains à mettre en place la politique de changement.

Mais l'ANSM ne souhaite pas, avec son guide des BPF, freiner l'innovation dans l'industrie pharmaceutique : « *Ce guide ne devrait, en aucune façon, freiner le développement de nouveaux concepts ou de nouvelles technologies, dès lors qu'ils sont validés et procurent un niveau de gestion de la qualité au moins équivalent à celui défini dans ce guide* » (105).

Toute application doit être validée et l'infrastructure informatique qualifiée avant son utilisation. La constance dans la qualité des produits, l'assurance de la qualité ou la maîtrise des procédés, ainsi que celle du risque général doivent être assurées dès lors qu'une tâche manuelle est automatisée. Des procédures claires doivent encadrer l'usage, la configuration et la maintenance des systèmes informatiques. Un inventaire de l'ensemble des outils d'IA et de leurs fonctionnalités est à conserver, et les spécifications utilisateurs (« *Users Requirements Specifications* » - URS) « *doivent décrire les fonctions requises du système informatisé et être basées sur une évaluation documentée du risque et de l'impact BPF. Les exigences de l'utilisateur doivent être traçables tout au long du cycle de vie* » (105).

Les systèmes utilisés doivent avoir démontré leur efficacité et leur adéquation aux situations d'usage, après qu'une analyse de risques traitant les *worst-cases* a été réalisée. Ainsi l'industrie pharmaceutique va devoir obligatoirement **adapter son organisation et ses procédures** aux nouvelles technologies d'IA qui vont l'assister dans ses activités quotidiennes.

Les méthodes de validation actuelles se basent sur des **procédés figés** : ils sont créés pour assurer des produits de qualité constante, avec des méthodes robustes et reproductibles, laissant très peu de place pour changer. Dès qu'un changement est effectué, une nouvelle validation du procédé est nécessaire.

*Les systèmes informatisés doivent périodiquement faire l'objet d'une évaluation afin de s'assurer qu'ils restent dans un état validé et conforme aux BPF* (105).

Cependant, l'inclusion des technologies d'IA dans ces processus changerait la donne. En effet, les modèles d'apprentissage profond sont utilisés pour leur capacités de perfectionnement perpétuel en apprenant au quotidien à partir d'expériences passées qui lui permettent d'adapter son modèle et son algorithme pour affiner son jugement. Ainsi l'Humain ne maîtrise plus réellement la modification des systèmes informatiques, donc comment assurer la validation d'un **procédé en constante évolution** ?

*Une machine d'inspection de produits en cours de fabrication se base sur des spécifications constantes pour distinguer les conformes des non-conformes (dimensions, forme, esthétique...). Les modèles de deep-learning se basent eux sur des bases de données en constante évolution, nourries par les précédentes analyses, les non-conformités passées, etc... afin d'adapter leur décision aux expériences passées.*

Afin de répondre à cette problématique, une **validation continue des procédés** est à envisager. Celle-ci se composerait de deux étapes principales : une première validation à la conception, et une deuxième spécifique au client et à l'utilisation qu'il en fera.

### 3.5. Limites

#### 3.5.1. Manque de connaissances sur le sujet

Une des principales limites à l'implémentation de l'IA dans l'industrie pharmaceutique est le manque de connaissance qu'ont de nombreux responsables. Ces technologies sont souvent assimilées à des boîtes noires, dont le fonctionnement n'est pas compréhensible. Comment dès lors intégrer ces technologies dans les activités quotidiennes sans les maîtriser ?

#### 3.5.2. Manque d'infrastructures

Les entreprises ne sont pas toutes dotées d'infrastructures leur permettant d'adopter l'IA dans un futur proche : lors du déploiement des systèmes informatisés, très peu de technologies d'IA existaient, ainsi ces derniers ne sont pas adaptés à leur utilisation. C'est un investissement très important (en temps et en argent) qui attend toutes les entreprises qui souhaitent intégrer l'IA.

#### 3.5.3. Manque d'alignement des entreprises

Cause ou conséquence du manque de connaissances ou d'infrastructures, de nombreuses entreprises n'ont pas encore pris le pas de l'IA. O'Reilly a rédigé un e-book en 2019, *AI Adoption in the Enterprise*, qui étudie les politiques entrepreneuriales relatives à l'IA, et s'attarde sur les raisons qui freinent certains à intégrer l'IA (108). La principale raison, avancée par près de 25% des interrogés, réside dans la **non-reconnaissance d'un besoin en IA dans la culture de l'entreprise**, témoignant d'un manque de connaissance sur le sujet. En plus du manque de ressources techniques, une vraie difficulté est celle de trouver des personnes qualifiées, en nombre très largement inférieur à la demande.

A contrario, certaines entreprises se lancent avec **trop d'optimisme**, trop vite, et sans réellement disposer d'une stratégie claire. Le déploiement des technologies d'IA requiert une stratégie partagée par tous, des objectifs réalistes, et des indicateurs pertinents. Sans toutes ces actions, il est impossible d'évaluer les résultats (bénéfices ou inconvénients) de l'utilisation de l'IA.

## 4. Point de vue Humain

### 4.1. Etat des lieux de l'industrie pharmaceutique en France

*Les données permettant d'établir un rapide état des lieux de l'industrie pharmaceutique en France proviennent du syndicat LEEM, qui rassemble 260 entreprises du médicament. En septembre 2019, ils publient une étude récapitulative très détaillée relative à leurs membres.*

Selon cette étude du LEEM, la France dispose d'un nombre d'employés en entreprises du médicament constant depuis près de 5 ans, autour des **100.000 salariés** (109), après avoir connu une baisse significative de ses effectifs entre 2007 et 2014 (cf. Figure 30). 57% de ces employés étaient des femmes, l'industrie pharmaceutique étant un des secteurs les plus féminisés.

Le secteur industriel pharmaceutique représente près de **0.33% de la population active française** (29.700.000 français actifs en 2019 selon l'INSEE (110)), employés qu'il faut rassurer sur l'essor de l'IA, et accompagner dans la transition vers l'industrie de demain.

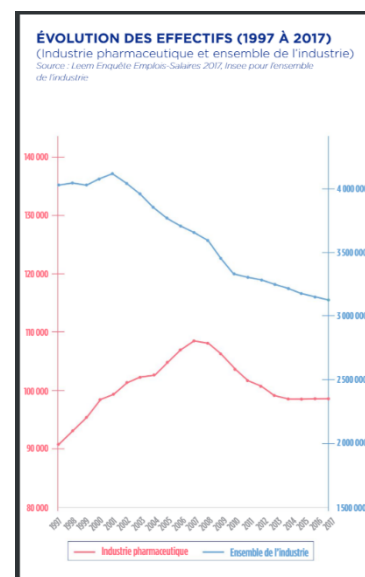


Figure 30 - Evolution des employés en entreprises du médicament entre 1997 et 2017 (109)

### 4.2. Le monde du Travail

L'IA va entraîner de **grandes transformations** du monde du travail, que le Code du Travail doit anticiper, afin de permettre une transition douce vers le monde du travail de demain.

#### 4.2.1. Mutation des Organisation et Gestion du Changement

Le monde du travail n'a pas attendu le développement de l'IA pour changer : il est en perpétuelle mutation au fil du temps. Dernièrement, c'est le développement des technologies numériques (et la crise mondiale du Coronavirus) qui a modifié en profondeur la façon de travailler (télétravail, etc...).

L'IA au sein des organisations pharmaceutiques va questionner sur les relations entre la machine et le travailleur : d'un simple instrument de travail, la machine sera désormais amenée à relayer voire remplacer ce dernier. Il faut y voir une opportunité selon Cédric Villani : *"l'automatisation des tâches et des métiers peut constituer une chance historique de désautomatisation du travail humain : elle permet de développer des capacités proprement humaines (créativité, dextérité manuelle, pensée abstraite, résolution de problèmes)"*(79).

Ainsi, ces nouvelles technologies pourront décharger l'Humain de tâches répétitives, à faible valeur ajoutée pour focaliser sa créativité sur des tâches à plus forte valeur ajoutée : c'est ce que le député nomme la *"complémentarité capacitante"*. Cela permettrait à la fois d'améliorer leurs conditions de travail et de valoriser leurs activités. Ils devront apprendre à travailler avec la machine, en synergie, pour atteindre une efficacité optimale.

Toutes ces nouvelles situations sur le lieu de travail obligent l'ouverture de nombreux **chantiers législatifs spécifiques**. En effet, le droit du travail actuel repose toujours sur le modèle traditionnel d'unité de lieu et de temps de travail, qui sera amené à changer dans les prochaines années (111).

Tous les changements que va engendrer l'IA obligent à se poser dès maintenant la question de la conduite du changement et de la façon de la mettre en place. Plus elle est anticipée, plus la transition s'effectuera en douceur.

#### 4.2.2. L'Automatisation

L'**indice d'automatisation** permet d'évaluer pour un emploi sa vulnérabilité vis à vis de l'automatisation, plus celui-ci tend vers 1 et plus le poste est susceptible d'être automatisé.

Il est donc vital de cibler dès maintenant les emplois présentant le plus grand risque d'automatisation, afin d'anticiper les changements inévitables qui vont en découler sans mettre de côté l'aspect humain.

Certaines catégories de métiers sont techniquement plus facilement automatisables que d'autres : les activités physiques réalisées dans des environnements structurés (manutention, conditionnement manuel...), la collecte de données ou leur exploitation..., ces postes représentent selon une étude du McKinsey Global Institute **près de la moitié des emplois** (tous secteurs confondus), dans la plupart des pays (112).

Les postes les moins concernés *a priori* sont ceux de management, d'expertise ou de relation avec la clientèle.

Malgré cela, il apparaît peu probable que des tâches soient intégralement automatisées. Beaucoup d'entre elles le seront partiellement.

#### 4.2.3. Les Emplois

L'IA et l'automatisation des tâches devraient permettre d'accroître la productivité des organisations. Mais de nombreuses craintes accompagnent ce processus : qu'en est-il des emplois ? Beaucoup sont inquiets d'une perte de travail au profit de solutions d'IA moins onéreuses, d'autres du rôle qu'ils auront à jouer dans leur nouveau travail. **Quelle est la place de l'Humain dans l'entreprise pharmaceutique de demain ?**

Les pays ne seront pas tous atteints de la même façon en ce qui concerne les emplois et le Monde du travail : les normes sociales, les facteurs législatifs, les dynamiques des marchés du travail seront des facteurs clés propres à chaque pays. Les pays fortement industrialisés comme la France, le Japon ou les Etats-Unis semblent plus enclins à souffrir de l'automatisation de certaines tâches que l'Inde par exemple.

#### 4.2.3.1. Les suppressions d'emploi et licenciements

Il est très difficile d'estimer précisément l'impact des technologies de l'IA sur les suppressions d'emplois. Le cabinet britannique PwC estime dans son étude *Will robots really steal our jobs ?* que l'automatisation ne va pas conduire à plus de chômage technique qu'il n'y en a déjà eu depuis le début de la révolution digitale. En effet, l'IA pourrait supprimer 7 millions d'emplois, mais en créer 7,2 millions dans le même temps (113).

Une étude réalisée en 2016 par la société de conseil McKinsey & Company prévoit que 15% de la force de travail globale (près de 400 millions d'ouvriers) pourrait être affectée par l'automatisation d'ici 2030. L'industrie pharmaceutique, employant beaucoup de personnes en France, sera également touchée par ces mesures, et si 15% de ces employés devraient être mutés, cela correspondrait à 15.000 personnes (112).

- Les licenciements

Les mutations technologiques à l'origine de suppression d'emplois sont-elles des motifs acceptables de licenciement au regard de la Loi ? L'article L.1233-3 du Code du Travail stipule que **oui** : *“Constitue un licenciement pour motif économique le licenciement effectué par un employeur pour un ou plusieurs motifs non inhérents à la personne du salarié résultant d'une suppression ou transformation d'emploi ou d'une modification, refusée par le salarié, d'un élément essentiel du contrat de travail, consécutives notamment : (...) 2° à des mutations technologiques”* (114). L'investissement dans les technologies d'IA pourrait donc *a priori* justifier des réductions d'effectifs.

Mais malgré cela, des protections existent afin de les limiter. L'employeur ne doit recourir au licenciement économique en réponse à des mutations technologiques importantes et rapides qu'après qu'un plan d'adaptation des salariés n'ait été réalisé, ou après avoir cherché à replacer les salariés sur des emplois disponibles dans l'entreprise sur le territoire.

La législation accordera une attention toute particulière à l'ensemble des licenciements causés par la transition technologique due à l'IA pour en étudier le bien-fondé.

#### 4.2.3.2. Les créations et modifications d'emploi

##### 4.2.3.2.1. Problématiques liées aux modifications d'emplois

Pour lutter contre les effets perturbateurs de l'automatisation (pertes d'emplois, etc...), il sera primordial d'assurer une croissance économique et de productivité robuste, qui sont des prérequis essentiels à la création d'emplois.

Une étude réalisée par McKinsey en 2018 a démontré que le nombre approximatif de postes créés d'ici 2030 serait de 555 à 890 millions dans le Monde, ou 21 à 33% de la force de travail actuelle (investissement dans de meilleures infrastructures, de nouvelles technologies ou dans la transition écologique) (115). La conclusion de cette étude était que la demande en main d'œuvre permettrait à elle seule de contrer les pertes de travail dus à l'automatisation. Le rapport



conjoint de Dell & l'Institut pour le Futur en juillet 2018 est lui sans appel : *"85% des métiers de 2030 n'existent pas aujourd'hui !" (116).*

Deux principales problématiques accompagnent la création et la modification d'emplois :

- Assurer un **cadre éthique**, pour protéger les catégories les plus vulnérables face au monde de demain.

Que ce soit la discrimination à l'embauche, lors des évaluations, ou dans le rapport quotidien avec ces technologies, les personnes les plus vulnérables doivent être protégées face à ces nouvelles technologies, pour éviter la survenue de risques psycho-sociaux (détaillé ci-après). Cédric Villani précise dans son rapport, que *"l'IA ne peut pas être une nouvelle machine à exclure"*(79). Pour cela, il propose la création d'un comité d'éthique des technologies numériques et de l'IA pour organiser le débat public autour de ces sujets. Le Comité économique et social européen (CESE) recommande pour sa part que les futures lignes directrices établissent *"une frontière claire dans l'interaction travailleurs-machines intelligentes, et incluent des principes de transparence dans l'utilisation des systèmes d'IA pour le recrutement, l'évaluation et le contrôle managérial des travailleurs"* (117).

- Et réfléchir à ce que sera **l'Homme employable à l'avenir**.

Depuis quelques années, et la complexité croissante des technologies utilisées, le renforcement des législations autour des produits de santé en termes de qualité ou de traçabilité et l'intensification de la concurrence intra-sectorielle, le **niveau de qualification du personnel de l'industrie pharmaceutique est en hausse**, du fait du besoin en nouvelles compétences. Selon l'étude du LEEM, plus de la moitié des employés dispose d'un niveau de qualification élevé (*supérieur ou égal au niveau 6 de la classification de la Convention collective*) : cadres, directeurs, etc... représentant des postes à responsabilités élevées (109).

Pour l'industrie pharmaceutique, il semble que la demande de force de travail en qualité, réglementaire, maintenance, gestion de données médicales ou en pharmacovigilance augmente dans les années futures. Certains métiers se verront profondément modifiés à la suite de l'intégration des nouvelles technologies. Cela exigera l'acquisition de nouvelles compétences par les scientifiques, les ingénieurs, les managers, les ouvriers : l'ensemble de la hiérarchie sera concerné.

#### 4.2.3.2.2. La mutation des compétences

Des millions d'employés de l'industrie pharmaceutique (comme d'autres domaines) risquent de changer de travail, du fait de la baisse du besoin dans des activités telles que la manutention ou d'analyse de données, de plus en plus automatisées. Cette transition risque de ne pas se faire en douceur, et pourrait mener à des hausses transitoires du taux de chômage dans les secteurs concernés.

Les employés de demain auront besoin de compétences différentes. Celles sociales et émotionnelles, telles que l'empathie ou la communication seront très recherchées, au moins

autant que les connaissances techniques avancées. Les **compétences cognitives** auront plus de place : créativité, traitement des informations complexes. La gestion de l'équilibre entre l'offre et la demande en compétences sera à surveiller, afin de ne pas avoir de pénuries de compétences dans certains secteurs.

#### 4.3. Coopération entre la force de Travail et l'IA

Plutôt que l'Homme ne soit au service de l'IA, une coopération synergique entre ces deux entités semble plus pertinente : chacun pourrait apprendre et se perfectionner grâce à l'autre. Le point essentiel est de trouver le juste équilibre entre l'IA et le travail humain. Cette méthode favorise également la culture de **l'innovation continue**, car les compétences humaines sont actualisées et adaptées aux besoins. Les employés et l'IA devront être actualisés régulièrement, et ainsi faire face à l'obsolescence toujours plus rapide des technologies et des connaissances.

#### 4.4. La compréhension des techniques

L'enjeu principal pour la bonne adhésion et utilisation des employés aux nouvelles technologies auxquelles ils seront confrontés en sera leur compréhension. C'est également pour cela que les développeurs doivent attacher une importance particulière à l'interface, l'ergonomie et le design de leurs solutions pour les rendre compréhensibles et faciles à utiliser.

L'environnement de l'industrie pharmaceutique de demain mêlera un grand nombre de nouvelles notions, de machines ou de procédés. Un individu seul ne sera pas suffisant pour tout gérer, au vu des nombreuses compétences nécessaires : le travail d'équipe sera plus que jamais mis en avant. Il faut tout d'abord maîtriser les technologies et les notions qui les nourrissent (statistiques, etc...), puis savoir interpréter les données et statistiques issues de ces technologies, afin d'en créer une visualisation et de pouvoir prendre une décision.

En plus de cet apprentissage technique, il sera important que les employés connaissent et maîtrisent le fonctionnement de leur organisation et de leurs procédés de production pour inclure l'ensemble des facteurs influant dans leur prise de décision. Il n'existe qu'une seule solution pour assurer la bonne maîtrise des outils d'IA par les salariés : **l'accompagnement et la formation aux nouvelles technologies**.

#### 4.5. L'accompagnement et la formation

##### 4.5.1. La Gestion prévisionnelle de l'Emploi et des Compétences

Les organisations pharmaceutiques doivent bénéficier de la révolution numérique et de l'IA en adaptant leur gestion des ressources humaines aux nouveaux besoins qui apparaissent. Pour ce faire, il faudra mettre les salariés dans les meilleures conditions pour maîtriser les nouvelles techniques et acquérir de nouvelles compétences permettant d'avoir une valeur ajoutée vis-à-vis des machines sur le poste de travail.

Un outil juridique oblige les entreprises à encadrer les salariés afin de les adapter à un nouveau poste : c'est la **Gestion prévisionnelle de l'Emploi et des Compétences** (GPEC). Elle

peut être résumée en deux termes : formation et mobilité, par l'implémentation de mesures de suivi de l'évolution des salariés (perspectives, mobilité géographique ou professionnelle), d'entretiens d'évaluation, de bilans de compétences, etc... Support essentiel dans la gestion des ressources humaines des organisations, sa finalité est d'anticiper les manquements futurs en force de travail et d'impliquer les salariés dans un projet d'évolution professionnelle.

Chaque responsable est obligé de proposer à ses employés une adaptation à l'évolution de leur emploi. L'article L.6321-1 du Code du travail annonce que *“l'employeur assure l'adaptation des salariés à leur poste de travail. Il veille au maintien de leur capacité à occuper un emploi, au regard notamment de l'évolution des emplois, des technologies et des organisations”* (118).

L'introduction régulière de nouvelles technologies dans le milieu pharmaceutique devra être prise en compte dans des revues et négociations régulières (annuelles) de la GPEC, afin d'adapter les salariés d'aujourd'hui à leurs emplois de demain (111).

#### 4.5.2. La Formation

La formation continue joue un rôle très important dans la vie professionnelle des employés de l'industrie pharmaceutique, qui y consacre en moyenne 3% de sa masse salariale selon le LEEM, afin de leur proposer des opportunités de développement des compétences ou d'évolution de carrière (109). La formation sera primordiale dans la réussite de l'intégration des nouvelles technologies dans le paysage pharmaceutique, mais les employés ne seront pas les seuls concernés.

##### 4.5.2.1. Formation des Managers

La formation du personnel commence par celle de l'exécutif. Les managers doivent comprendre et maîtriser les nouvelles technologies qui seront utilisées au sein de leurs organisations, en connaître les possibilités, limites, leurs forces et leurs faiblesses. En effet, cela leur permettra de transmettre à leurs employés les bénéfices qu'apportent ces nouvelles technologies, pour les motiver mais aussi comprendre et répondre à leurs inquiétudes : perte d'autonomie, d'importance, etc... Cela leur permettra également d'établir des stratégies et des objectifs raisonnables de transition numérique. Pour faire face aux nouveaux challenges de l'IA, les managers devront être formés afin de disposer de nouvelles compétences, en plus de celles de base (119):

- Compétences digitales ;
- Compétences d'Agilité, pour mettre en place des méthodes de travail collaboratives ;
- Compétences de *Design Thinking* (conception créative), pour implanter une philosophie innovante et différente dans l'entreprise.

Observer la concurrence sera aussi une méthode efficace pour apprendre des autres : les erreurs, comme les bonnes idées. Le soutien et le conseil d'experts du domaine sera également nécessaire. Il convient donc d'anticiper le plus tôt possible les plans de formation pour les managers de demain au vu de leurs nouveaux rôles et responsabilités.

#### 4.5.2.2. Formation des Employés

Les salariés sont les plus en contact avec les nouvelles technologies : leur formation sera le point d'orgue de l'intégration de l'IA dans l'industrie pharmaceutique. Un de ses freins est le **manque de connaissance du sujet**, qu'il faut combattre par l'éducation.

L'*Upskilling* est défini comme l'apprentissage de nouvelles compétences à des employés. C'est la clé de voûte de l'intégration de l'IA, permettant de contrer les risques du vieillissement de la population et de l'émergence des nouvelles technologies. En effet, l'IA se développe actuellement plus vite que la qualification des travailleurs, risquant de faire apparaître un écart important entre besoin et solution.

Selon une enquête réalisée par Malakoff Médéric et le Boston Consulting Group en décembre 2017, *“Intelligence artificielle et capital humain : quels défis pour les entreprises ?”*, seulement 44% des salariés se disent prêts à travailler avec une IA, et 20% d'entre eux se sentent à l'heure actuelle suffisamment accompagnés sur le sujet (120).

Mettre en place une formation initiale présentant les bases de l'IA et de ses applications industrielles sera nécessaire, en insistant sur ses bénéfices certains pour l'entreprise et surtout pour chaque employé.

*Une étude publiée par Forbes en 2017, “Empowering the firstline workforce” a démontré qu'une grande majorité des employés voyaient en l'IA plus une amélioration qu'un remplacement de leurs activités quotidiennes (121). Le travail des techniciens de l'industrie pharmaceutique par exemple peut se diviser en deux catégories :*

- *Des tâches quotidiennes simples, répétitives et chronophages, prioritaires (manutention, contrôles, etc...) ;*
- *Et des tâches plus complexes, mais “moins importantes” au vu du produit à fabriquer (aspect managérial, amélioration continue...).*

*Il reste ainsi souvent peu de temps aux employés pour se consacrer à ces tâches complexes, mais grâce à l'automatisation de ces dernières, de précieuses heures pourront être économisées. Les bénéfices pour les employés sont nombreux (122) : meilleure collaboration entre les équipes, amélioration de leur qualité de travail et de leur satisfaction, revalorisation de leur travail (augmentation de la part de tâche à valeur ajoutée), meilleure communication des idées et des informations à l'échelle de l'entreprise, amélioration des processus et des flux de travail...*

*La création automatique de supports de présentation (PowerPoint récupérant directement les données de production pour en extraire les indicateurs qualité par exemple, etc...), ou l'analyse simplifiée et automatisée des données pourront également faire gagner beaucoup de temps aux employés (et donc de l'argent aux entreprises).*

Il faudra faire preuve d'honnêteté dans les désavantages de l'IA, afin que les ouvriers ne les découvrent pas en cours d'utilisation, et perdent confiance à la fois en ces technologies mais également en leurs managers.

La formation continue ne devra pas être négligée : étant très dynamique, l'IA et ses technologies évoluent à un rythme très élevé, il ne faut pas prendre de retard, et toujours maintenir à jour les connaissances de ses employés, pour conserver la dynamique en place.

En effet, les compétences techniques ont vu leur durée de vie chuter depuis les années 1970. A cette époque, elle était de près de 20 ans, contre seulement 2,5 aujourd'hui. L'Organisation de coopération et de développement économiques (OCDE) prévoit même voir cette durée atteindre 1 an en 2025 (119). Ces compétences techniques ou *hard skills* seront donc à renouveler à fréquence régulière, grâce à la formation continue, en se basant sur les capacités de reconversion des salariés.

#### 4.5.2.3. Formation des Étudiants

Les employés de demain devront également être prêts à répondre aux nouvelles missions qui leur seront confiées. La demande en experts en IA est très forte actuellement, des emplois seront donc facilement atteignables pour des jeunes diplômés à partir du moment où une formation donnant les compétences requises est disponible. Cela n'est pas le cas à l'heure actuelle, les formations ne sont pas adaptées, et le nombre de talents n'est pas suffisant. Cédric Villani propose de multiplier par trois le nombre de personnes formées à l'IA pour pallier ce manque. Monsieur Macron concède que la France doit *“faire un effort particulier pour la formation”* mais il tempère : *“Nous doublerons le nombre d'étudiants formés à l'intelligence artificielle depuis la licence, jusqu'au doctorat, en passant par les formations professionnelles courtes et prévoirons les financements qui correspondent à ce doublement”* (80). Aucun délai n'est annoncé pour la mise en place de ces financements.

Pour aller plus loin, il est même envisageable que l'IA soit enseignée plus tôt dans le cursus scolaire : dès le collège ou le lycée, pour faire naître l'intérêt et des vocations pour les élèves. Le mode d'enseignement devrait être revu, et intégrer des compétences transversales, favoriser la créativité et de nouvelles formes de pédagogies, dès le plus jeune âge.

### 4.6. Limites

#### 4.6.1. Point de vue du grand public

La Résolution du Parlement européen du 12 février 2019 sur une politique industrielle européenne globale sur l'intelligence artificielle et la robotique annonce que *“l'avenir de cette technologie dépendra de son acceptation par la société et qu'il convient de mettre davantage l'accent sur la communication appropriée de ses avantages pour permettre une meilleure compréhension de la technologie et de ses applications”* et que *“si la société n'est pas suffisamment informée sur les technologies de l'IA, il y aura moins de motivation pour l'innovation dans ce secteur”* (83). La période de transition nécessaire devra comprendre des

formations générales pour comprendre l'intérêt de ces technologies dans l'intégration dans nos vies quotidiennes.

L'institut de sondages CSA a réalisé une étude, qui a conclu que **48% des Français considèrent l'IA comme une menace potentielle pour l'emploi**, tandis que 68% des jeunes estiment que leur travail pourra être remplacé par une machine (123). Ces chiffres témoignent d'une inquiétude d'une grande partie de la population envers ces technologies.

Beaucoup de personnes sont inquiètes de l'importance grandissante des technologies du numérique, à la fois sur l'impact qu'elles auront dans leur vie professionnelle mais aussi personnelle : place de l'Homme dans la société de demain, la gestion qui sera faite des données collectées, les usages malveillants de ces technologies, etc... Certains relayent des informations erronées, prêtant à l'IA des desseins fallacieux, qui peuvent atteindre l'opinion publique, et la "popularité" des nouvelles technologies.

Bien que les mécanismes qui forment l'IA peuvent être complexes, il est primordial d'arriver à **communiquer** simplement sur les bénéfices certains de ces technologies, afin de disposer du soutien de l'opinion publique. L'IA n'est pas destinée à remplacer l'Homme, mais bien à collaborer avec lui.

#### 4.6.2. La rareté des experts

L'IA repose sur les machines, certes, mais également sur les Humains qui l'utilisent. La construction et l'optimisation des modèles d'IA (réseaux de neurones, etc...) nécessite une réelle expertise. Les compétences en IA sont très recherchées et les formations en France ne sont pas encore adaptées.

A l'heure actuelle, **l'offre en experts dans le domaine de l'IA est très largement inférieure à sa demande**. Une étude réalisée par McKinsey en Octobre 2018 estimait à 10.000 le nombre d'experts en IA (124), capable de régler les problèmes les plus complexes. C'est très peu. Forcément, une concurrence existe dans l'acquisition de ces talents. Concurrence entre entreprises, mais également entre pays.

Les pays leaders dans la recherche en IA disposent d'atouts indéniables pour attirer ces experts, dont leur puissance d'investissement et de réseaux. Pour le moment, ils se recrutent hors d'Europe, principalement aux Etats-Unis, ou au Canada. De la même façon, les entreprises n'appartenant pas au groupe "FAMGA" (Facebook, Apple, Microsoft, Google, Amazon) ou de petite taille ont des difficultés à développer leur partie IA, par manque de compétences et de moyens face à la concurrence (81).

Pour lutter contre cette fuite des cerveaux et faire de la France un lieu d'attractivité et de compétitivité internationale, M. Villani propose de doubler les salaires des chercheurs en début de carrière. Un autre axe d'amélioration : la féminisation de la recherche en IA. En 2013, seuls 30% des chercheurs étaient des femmes. Ainsi, il ne suffit pas d'augmenter le nombre de chercheurs mais surtout de mieux exploiter le vivier de talents présents dans chaque pays.



Comme vu précédemment, l'industrie pharmaceutique est parmi les plus féminisée, donnant de l'espoir dans la diversification des profils de chercheurs de ce domaine.

Il ne suffira donc pas seulement de disposer d'outils et de nouvelles technologies, mais aussi et surtout de main d'œuvre qualifiée pour les manipuler.

## 5. Point de vue Technique

### 5.1. La Gestion des données

#### 5.1.1. Les données utilisées dans l'Industrie Pharmaceutique

Comme dit précédemment, **sans données il n'y a pas d'IA**. Ces données sont donc au centre de l'industrie pharmaceutique du futur. Quels en sont les types utilisés dans l'industrie pharmaceutique (et qui le seront à l'avenir) ?

Chacun de ces domaines utilisent et génèrent des données différentes (*liste non exhaustive*) :

- Recherche & Développement : données de littérature, spécifications des molécules candidates (morphologies, caractéristiques physico-chimiques...), cibles moléculaires, pathologies, données de formulation, etc...
- Recherche clinique : données des patients inclus dans les essais cliniques (âge, sexe, adresse, données médicales : cliniques, biologiques, d'imagerie, etc...)
- Production : conditions de production, monitoring des machines (spécifications, statistiques, statut en fonctionnement ou en panne...), données de capteurs et d'objets connectés, données de contrôles effectués (non-conformités, déviations, statistiques, etc...).
- Logistique : données de transport (itinéraires, consommation d'essence...), de consommation (statistiques d'utilisation de produits, de commandes fournisseurs, de délais de livraison, etc...)
- Marketing : cibles, intentions d'achats, marchés potentiels, etc...

Les données sont donc d'une **très grande diversité**, et leur maîtrise sera un point d'orgue à l'intégrité des systèmes. Si de telles informations tombaient entre les mauvaises mains cela pourrait être terrible : pertes financières catastrophiques, perte de monopole, fuite d'informations personnelles sensibles, etc...

#### 5.1.2. La cybersécurité et protection des données

La Pharma 4.0 rassemble en son sein des milliers d'appareils connectés, qui peuvent être des potentielles "faiblesses", d'un point de vue sécurité. Afin de profiter des avantages du concept d'industrie 4.0, et de limiter les risques liés à la sécurité, les industriels du médicament vont devoir **investir dans du personnel de sécurité qualifié**.

Les technologies d'IA (et leurs données principalement) sont sauvegardées dans des serveurs, qu'ils soient la propriété de l'entreprise ou du fournisseur du programme, ou sur le cloud. Des



**analyses de risques détaillées** doivent donc être réalisées dès l'installation des technologies, pour évaluer les flux de données et leur sécurité.

L'industrie pharmaceutique ne pourra réussir sa transition numérique si le public perd confiance dans les outils d'IA utilisés : en cas de violation d'intimité, d'usage malveillant ou de mauvaise qualité de résultats. C'est pourquoi des **bonnes pratiques d'usages de l'IA** propres à l'industrie pharmaceutique devront être rédigées et des **moyens de surveillance** (audits, certifications, etc...) mis en place afin d'assurer une utilisation adéquate et une détection efficace des mésusages de l'IA.

Voici quelques pistes pour assurer la sécurité d'usage des technologies d'IA :

- Renforcer la protection, la sécurité et la confidentialité des données et des utilisateurs

La spécificité de l'application de l'IA dans le monde pharmaceutique nécessite la mise en place d'un **régime de confidentialité des données particulier** (surtout pour les essais cliniques). Il faudra également réfléchir sur le statut du partenaire en IA (en tant que contrôleur, gestionnaire ou analyste des données), pour s'assurer que des données anonymisées ou pseudo-anonymisées seront utilisées, sans possibilité d'identifier des sujets anonymisés.

- Établir un cadre général et des principes pour une utilisation bénéfique et sûre de l'IA.
- Partager les meilleures pratiques et promouvoir l'innovation continue pour résoudre des problématiques telles que la sécurité, les biais et l'explicabilité.
- Trouver le bon équilibre entre les entreprises et la concurrence nationale afin de garantir que les avantages de l'IA sont largement disponibles et partagés (125).

Au vu du nombre colossal de données et de leur importance, cette étape est critique dans la validation des systèmes d'IA.

### 5.1.3. La qualité et quantité des données

#### 5.1.3.1. Qualité des Données

Les performances du système construit dépendent en grande partie des données qui lui sont fournies. Avoir des données fiables et de qualité irréprochable issues de sources crédibles est une condition *sine qua non* pour disposer d'outils d'IA adéquats et performants, donnant des résultats fiables. Des outils disposés à garantir la gestion de la qualité des données utilisées devront être mis en place. Une entreprise utilisant des données erronées ou non fiables (non harmonisées, etc...) passera plus de temps à les rendre utilisables qu'à les étudier.

Les données étaient auparavant pour la plupart sous forme structurée ou textuelle, mais une grande partie de celles issues de l'internet des objets l'est désormais sous forme d'images ou de vidéos. Les outils exploitant ces données nécessitent qu'elles soient étiquetées (comme le propose *Imagenet*), pour en permettre la meilleure compréhension. Le problème étant qu'à l'heure actuelle nous produisons plus de données qu'il n'existe de main d'œuvre pour les étiqueter. Cette tâche peut être réalisée en interne, ou sous-traitée à une autre entreprise, mais reste indispensable pour pouvoir disposer de données analysables.

#### 5.1.3.2. Quantité des Données

La quantité de données produite par l'industrie pharmaceutique est en hausse exponentielle depuis 10 ans, et cette tendance va se poursuivre de façon assez logique dans la prochaine décennie. Les problématiques de l'industrie pharmaceutique seront de mettre en place des méthodes pour exploiter ces bases de données toujours plus conséquentes, mais aussi de trouver des façons efficaces de récolter et de partager ces données massives.

Il ne s'agit pas que de disposer d'un grand nombre de données : encore faut-il qu'elles soient pertinentes. Donner à un modèle des données qui ne le sont pas risque de faire perdre du temps voir de diminuer la pertinence des conclusions du modèle. La question que l'Homme doit se poser est la suivante : *de quelles données disposons-nous actuellement, et quelles sont celles que le modèle requiert ?* La connaissance du modèle est donc ici aussi une absolue nécessité, et la finalité permet de savoir quelles données manquent au modèle.

#### 5.1.4. L'accès aux données

Un accès équitable devra être permis aux différentes organisations qui le souhaiteront, afin de s'assurer que toutes puissent profiter d'une façon juste aux bénéfices de l'IA. Cela vaut au sein des pays mais aussi entre les différents pays.

Le partage des données devra être facilité entre les différents pays, pour alimenter les bases de données plus efficacement. Aujourd'hui le taux de partage est nettement inférieur à ce qu'il pourrait être, et beaucoup de données sont mal ou sous-utilisées. Des normes communes devront être éditées pour encadrer ce partage de données. Lors du 8ème Conseil stratégique des industries de santé, des décisions ont été prises pour faciliter l'utilisation et l'exploitation des données de santé (recherche, pilotage, suivi) avec le déploiement du Health Data Hub.

Le partage et l'accès aux données devront donc être facilités, afin de garantir des processus plus efficaces et efficients. Mais des vides juridiques existent à l'heure actuelle sur la notion de propriété des données, et ont conduit à une prudence excessive de l'industrie à son égard. Ces vides devront être comblés, afin d'encourager le partage de données, tout en en assurant la sécurité. Angela Merkel, la chancelière allemande a milité pour la création d'un marché numérique commun au sein de l'UE, et à l'établissement de normes éthiques relatives à l'IA.

### 5.2. Explication des résultats

La démarche de lier une entrée à une sortie est nommée **l'explicabilité des résultats**. Avec le développement de l'IA, les modèles deviennent de plus en plus complexes, et il est de fait plus compliqué de comprendre précisément comment son résultat a été obtenu. C'est le concept de *boîte noire de l'IA* mentionné précédemment. Il faut disposer d'outils capables de fournir un résultat et de disposer de preuves robustes pour le justifier. Des méthodologies, telles que les *Explications locales interprétables, agnostiques au modèle* (LIME) tendent à augmenter la transparence des modèles (126). Le principe est simple en théorie : transposer les prédictions en un simple modèle mathématique interprétable (régression linéaire principalement).

L'Homme doit garder le contrôle sur ces technologies, et doit comprendre la démarche qui a amené le modèle à donner le résultat. La CNIL recommande dans son rapport *“Comment permettre à l'Homme de Garder la main”* de Décembre 2017 de *“Rendre les systèmes algorithmiques compréhensibles [...] [et de] travailler le design des systèmes algorithmiques au service de la liberté humaine, pour contrer l'effet « boîtes noires »”* (127).

### 5.3. Limites techniques

#### 5.3.1. L'accès aux données

Une des principales limites est l'accès aux données. Certaines sont difficiles à visualiser, ou collecter, les rendant impossibles à analyser.

La question de la **propriété des données** se pose également : les données de patients ne peuvent être utilisées qu'après avoir obtenu leur consentement. Si ces données sont utilisées, il faut obtenir le consentement une nouvelle fois. Le Big Data nécessite de disposer de grandes bases de données pour être fonctionnel, comment pouvoir accéder à celles-ci rapidement s'il faut obtenir l'accord des patients auparavant ? Un processus d'anonymisation sécurisé serait-il suffisant pour s'affranchir de cet accord ?

#### 5.3.2. La qualité des données

Comme dit précédemment, les résultats du modèle dépendent grandement de la qualité des données qui lui sont fournies. Certaines entreprises peuvent être amenées à travailler avec des données anciennes, qu'il est nécessaire de nettoyer avant de les utiliser. Selon le conseil national du Numérique, *“la valeur créée par l'intelligence artificielle provient des données nécessaires à l'apprentissage bien plus que de l'algorithme, dont les développements se font de manière ouverte (en open source). En conséquence, les organisations doivent prendre conscience de l'importance stratégique des données qu'elles détiennent”* (128).

Utiliser des données de mauvaise qualité (obsolètes, incomplètes, incorrectes...) va conduire à de mauvaises prédictions, des résultats erronés, un mauvais apprentissage du modèle et donc à l'apparition de biais.

#### 5.3.3. La compréhension des techniques

Un algorithme d'IA évolue constamment, aussi, il ne donnera pas toujours la même réponse à la même question : de mêmes données, avec le même modèle, donneront à deux instants différents potentiellement deux résultats différents (similaires toutefois). Comment assurer la **reproductibilité** des méthodes. Comprendre les techniques utilisées apparaît comme primordial pour parvenir à justifier les résultats obtenus et garder la main sur ces nouvelles technologies. A l'heure actuelle, très peu de salariés de l'industrie pharmaceutique maîtrisent ces notions d'IA : la formation en IA doit avancer à la même vitesse que la recherche fondamentale dans ce domaine.

Il est d'autant plus important pour le secteur pharmaceutique, au vu des nombreuses décisions cruciales que des programmes d'IA seraient amenés à prendre d'ici quelques années.

#### 5.3.4. Les biais et erreurs de modèle

Les modèles d'IA ne sont pas infallibles, une des inquiétudes envers ces technologies est l'existence de **biais**. Ils sont à l'origine d'erreurs systématiques, et sont le plus souvent évitables. Comme les modèles sont le reflet de l'esprit de l'Homme qui les ont créés, ces biais peuvent être induits dès leur naissance, mais ils peuvent également être induits au cours de leur apprentissage autonome ou apparaître avec les données utilisées. C'est pourquoi il faudra être précautionneux à la fois des outils et données utilisés, et d'utiliser pleinement ces systèmes pour ne pas restreindre leur développement.

*Si un modèle d'IA utilisé dans les contrôles visuels en cours de production de médicament est biaisé, il peut conduire à la destruction de produits finalement conformes (confusion entre des bulles d'air et des particules de verre...) et donc à une grande perte financière et temporelle.*

Tout biais ou erreur de modèle conduira à des résultats inexacts. Heureusement, des solutions sont développées pour identifier les biais et les sources potentielles d'erreurs. *Whitebox Testing for Deep Learning Systems* en est une : elle teste un réseau neuronal avec un grand nombre de données d'entrée et identifie les erreurs.

#### 5.3.5. Peut-on tout automatiser ?

On ne peut pas tout remplacer par un ordinateur. Notre **pensée** et notre **éthique**, nous en distinguent. Un programme peut éditer un planning, mais ne pourra jamais innover ou créer un esprit d'entreprise.

## 6. Point de vue Environnemental

### 6.1. Impact de l'IA sur l'environnement

L'industrie pharmaceutique n'est pas épargnée par les nouvelles problématiques liées à l'environnement, principalement le réchauffement climatique. Les industriels doivent ainsi les prendre en compte dans leurs activités quotidiennes pour satisfaire aux exigences de la législation. Les solutions d'IA entraînent une utilisation accrue de ressources (électricité, stockage de données sur des serveurs, etc...) et selon le Journal de l'environnement, la consommation énergétique du numérique augmente actuellement de 8,5% par an, risquant d'augmenter la demande en électricité de façon exponentielle dans les 20 prochaines années. Les équipements numériques nécessitent également des matériaux rares, tels que l'Iridium, dont on estime les réserves restantes à seulement 15 ans.

La question de **Responsabilité Sociale des Entreprises** (RSE) se pose pour de plus en plus d'industriels. Le gouvernement la définit comme *“l'intégration volontaire par les entreprises de préoccupations sociales et environnementales à leurs activités commerciales et leurs relations avec les parties prenantes”* (129).

Selon Cédric Villani néanmoins, *“l'IA peut être une véritable opportunité environnementale”* (130), principalement en optimisant la gestion des ressources énergétiques (serveurs, électricité,

etc...). Mais bien que le rôle de l'IA ne fasse pas de doute, de nombreuses actions pour se préparer sont nécessaires.

Une étude réalisée par PwC France a estimé que l'IA permettrait de réduire de 4% les émissions de gaz à effet de serre mondialement d'ici 2030 (131).

## 6.2. Réduire les impacts environnementaux des médicaments

Les produits de santé peuvent avoir un impact sur l'environnement tout au long de leur cycle de vie : de leur production (pollution de l'air, des sols et des eaux, gestion des résidus de médicaments) à la fin de leur vie (recyclage des déchets, etc...)

L'IA aura un grand rôle à jouer dans la **rationalisation du développement durable**, et elle se devra d'être responsable.

- Automatiser les analyses de consommation énergétique permettra d'optimiser l'utilisation des ressources par les entreprises. Cela permettra de réduire la pollution atmosphérique.
- Monitorer en temps réel la pollution à l'aide de capteurs permettra de mettre en place des actions pour la réduire au plus vite.
- L'IA pourrait également repenser notre façon de gérer les rebuts et déchets de production, afin d'en optimiser le recyclage.

La réussite de l'intégration de l'IA dans les problématiques de RSE repose sur deux principes : l'intégration au cœur de l'entreprise et l'établissement de partenariats stratégiques avec d'autres parties prenantes pour favoriser la collaboration sur ces sujets (91).

Ainsi, l'intégration des outils d'IA dans les entreprises pharmaceutiques englobe des problématiques très diverses, dont aucune ne doit être mise de côté, et doivent être anticipées afin de réussir cette transition numérique.

# Partie 3 - Exemples d'utilisations actuelles et futures de l'IA dans l'industrie pharmaceutique

## 1. Introduction - Cycle de vie du Médicament

L'industrie pharmaceutique concerne l'ensemble du cycle de vie du médicament (cf. Figure 31) : des premières études de cibles, au retrait du produit du marché, l'industrie pharmaceutique gravite toujours autour des produits de santé.

La phase initiale est celle de **Recherche et Développement**, au cours de laquelle une pathologie et une cible sont identifiées, et différentes molécules candidates sont retenues, un Brevet est ensuite déposé afin de protéger l'avancée des recherches. Ces molécules candidates vont ensuite être testées sur des animaux au cours des **études Pré-Cliniques**, dans le but d'évaluer leur efficacité et toxicité sur l'animal, avant de l'administrer à l'Homme.

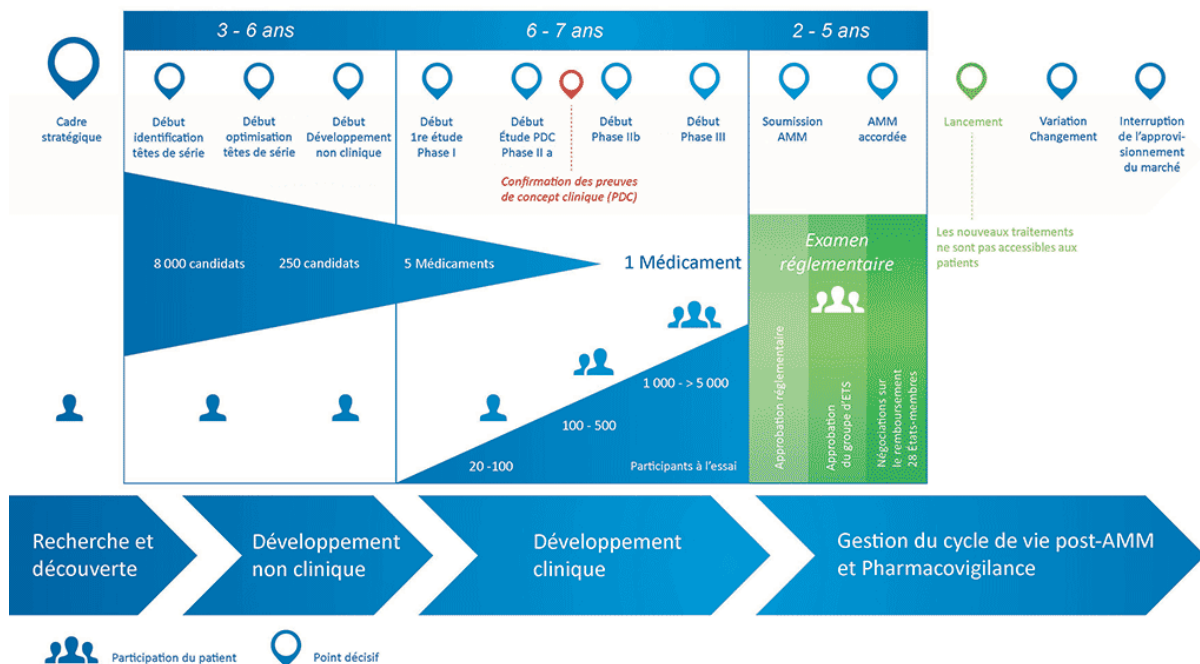


Figure 31- Cycle de vie du Médicament simplifié (132)

L'étape clé dans le développement d'un produit de santé est la phase d'**études cliniques**. Étant la première administration du produit chez l'Homme, elle est suivie de très près par les autorités et très réglementées (*Bonnes Pratiques de Recherche Clinique*). Constituées de 3 phases successives, elles servent à démontrer la sécurité et l'efficacité du Médicament pour l'Homme et leurs résultats sont utilisées pour obtenir l'Autorisation de Mise sur le Marché (AMM). Le dossier de **demande d'AMM** sera ensuite étudié par les autorités, qui statueront sur le rapport



bénéfice-risque du produit, et le comparerons avec des produits déjà existants, afin d'évaluer le Service Médical Rendu (SMR) et l'Amélioration du Service Médical Rendu (ASMR). Une fois l'AMM obtenue et les conditions d'usage du produit définies, la production industrielle débute et le produit est lancé sur le marché.

Mais la surveillance ne s'arrête pas là. Les éventuels nouveaux effets secondaires ou l'efficacité du produit est évaluée au cours des études post-AMM (ou de phase IV) : c'est la **Pharmacovigilance**. En cas de déséquilibre du rapport bénéfice-risque en faveur des risques, le produit peut voir ses modalités d'utilisations ajustées (modifications d'AMM...) ou même être retiré du marché. La fin du cycle de vie d'un produit de santé est atteinte lors de son **retrait du marché** (effets indésirables trop importants, efficacité remise en cause ou meilleure solution disponible).

*Cette dernière partie détaillera différentes phases du cycle de vie du produit afin d'en voir les projets innovants, de façon non-exhaustive.*

## 2. Intelligence Artificielle dans la Recherche & Développement

### 2.1. Définition de la Recherche & Développement

Le développement d'un médicament est un processus **très long** (15 ans sont nécessaires en moyenne pour passer d'une molécule d'intérêt à un médicament) et très coûteux (environ 1,2 milliards d'euros (133) de coûts moyens de recherche et développement, chiffre en forte croissance ces dernières années).

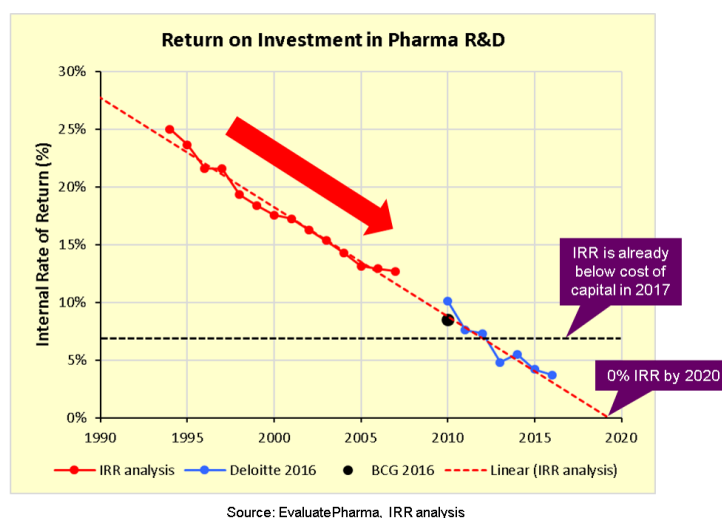


Figure 32 - Evolution du retour sur investissement dans la R&D pharmaceutique depuis les années 1990 (134)

Les entreprises dépensent plus de 170 milliards d'euros annuellement dans le développement, mais près de 90% de ces molécules échouent durant les essais cliniques. De plus, 75% des nouvelles molécules mises sur le marché ne parviennent pas à couvrir leur coût de développement, et les dernières prédictions parlent d'un retour sur investissement proche de 0 d'ici la fin de l'année 2020 (cf. Figure 32).



L'innovation pharmaceutique est en "crise" depuis quelques années. Plusieurs raisons s'ajoutent aux problématiques de coûts et de temps. Les produits vedettes (ou *blockbusters*) perdent en parts de marché, à cause du développement des médicaments génériques ("*copies des médicaments princeps tombés dans le domaine public, contenant la même quantité de principe actif et présentés sous la même forme pharmaceutique.*" selon l'Organisation Mondiale de la Santé (OMS)(135)). En effet, une fois sur le marché, le produit n'est plus protégé longtemps par son Brevet, et dès son expiration, son prix baisse de 30 à 50%, réduisant les revenus de l'entreprise, du fait que la molécule devient **générizable** (copiable) par un autre laboratoire en économisant plus de 40% des coûts (n'ayant que très peu de coûts de recherche à dépenser). Ces médicaments seront à la fois moins chers pour les patients, mais aussi pour le système de santé français.

Les **évaluations** des rapports bénéfices/risques des produits de santé par les autorités de régulation nécessaires à la mise sur le marché sont de plus en plus exigeantes et longues. De plus, les **remèdes** aux maladies faciles à traiter ont déjà été découverts, ne laissant que les pathologies plus complexes (cancers, maladies génétiques...) nécessitant des médicaments aux mécanismes d'action plus innovants.

Un des défis de l'industrie du futur pour améliorer ses marges est donc de réduire ces coûts et d'augmenter l'efficacité de développement (et de production) des médicaments et autres produits de santé, ou de se tourner vers des marchés de niche grâce aux biomédicaments.

L'IA et la puissance de calcul de ses ordinateurs pourraient résoudre ces contraintes de temps, et d'argent, et donc jouer un grand rôle dans la découverte de nouvelles molécules, de biomarqueurs, l'identification de cibles moléculaires, et devrait permettre de réduire la durée d'obtention de l'AMM, résultant en de grandes économies de coûts pour les développeurs (près de 30%), ainsi qu'une plus grande diversité de traitements disponibles (136). A l'heure actuelle, le marché de l'IA dans la R&D est évalué à 700 millions de dollars en 2018, avec une croissance estimée à 3,5 milliards d'ici 2027 (137).

## 2.2. Projets en cours de déploiement

La recherche se divise en deux phases :

- La **recherche fondamentale**, qui étudie les mécanismes de la maladie afin d'en identifier la cible moléculaire ;
- Et la **recherche pharmaceutique**, étape durant laquelle un grand nombre de molécules seront criblées pour sélectionner les molécules les plus efficaces au niveau des cibles identifiées précédemment.

### 2.2.1. En recherche fondamentale - Recherche de la cible moléculaire

#### 2.2.1.1. Etude bibliographique

La première étape de recherche scientifique repose sur une étude approfondie de la bibliographie relative à la maladie traitée, par l'exploration de rapports ou d'articles

scientifiques, de brevets, ou d'études. Cependant, leur nombre a explosé ces dernières années, et réaliser une étude bibliographique exhaustive est un processus très chronophage et fastidieux.

Les nouvelles techniques dérivées de l'IA, le **Text Mining** (*extraction automatique des sujets, concepts scientifiques et relations, ou descriptions à partir d'un ensemble de textes*) et le **Traitement du Langage Naturel** (*compréhension du langage humain par une IA*) sont des outils précieux dans l'extraction de l'information pertinente à partir de dizaines de milliers de sources diverses.

Le leader actuel dans ce domaine est *BenevolentAI* (138). Basé sur des systèmes de machine learning et de traitement du langage naturel, ce logiciel permet d'explorer et d'organiser des quantités de données très conséquentes provenant d'articles scientifiques, ou de données d'essais cliniques notamment, de façon totalement autonome (sans besoin d'intervention humaine en amont). Il analyse de façon autonome les bases de données médicales validées (Lancet, etc...), sans avoir à ajouter "manuellement" les articles à étudier.

BenevolentAI est capable de relier les informations sur les structures moléculaires aux données biomédicales des récepteurs et les pathologies associées afin de trouver les **cibles pertinentes potentielles** pour un traitement. BenevolentAI peut également identifier des molécules ayant échoué au cours d'essais cliniques précédents, et évaluer leur efficacité sur d'autres pathologies. Cela permet d'éviter de réitérer l'ensemble des tests de toxicité déjà réalisés et de passer directement à une phase plus avancée. Cela entraîne de grandes économies de temps et d'argent.

Cette entreprise a également créé une IA biomédicale, BenevolentBIO, qui a déjà identifié deux petites molécules apparemment efficaces contre la Sclérose Latérale Amyotrophique. Jackie Hunter, Présidente Directrice Générale de la start-up, a annoncé : *"Pour une industrie basée sur les preuves, nous n'en utilisons finalement que très peu. Une nouvelle publication scientifique paraît toutes les 30 secondes : il y a une grande quantité d'information qui n'est tout simplement pas utilisée pour la R&D de nouveaux médicaments"* (139).

L'efficacité de la recherche bibliographique dépend tout d'abord de la qualité, quantité et de la diversité de ses sources. Les techniques de Text Mining permettent d'extraire des informations et les concepts d'un texte et de les retranscrire, et, en lien avec le Machine Learning, des connexions peuvent être réalisées entre des documents de différentes sources : on nomme cela le **Linked-Data**. Des quantités astronomiques de documents peuvent être analysées et classifiées à une grande vitesse (environ 25 ms par document), grâce à l'identification de concepts clés et à la mise en relation de divers documents. Cette tâche chronophage a longtemps été réalisée par l'Homme, qui ne pouvait classer "que" 10.000 documents par mois (environ 10-15 minutes par documents) : le gain en efficacité est donc considérable (140).

Une publication d'Avril 2020 *"Can artificial intelligence replace manual search for systematic literature?"* (141) a comparé la méthode de screening documentaire traditionnelle à celle

réalisée grâce à un logiciel d'IA, afin d'extraire dans la littérature, les articles traitant d'un sujet particulier : les affections dermatologiques. Sur les 202 articles finalement retenus :

- 155 ont été identifiés par les deux méthodes ;
- 33 par la méthode manuelle uniquement ;
- Et 14 par la méthode d'IA uniquement.

Les résultats sont ainsi très proches entre les deux méthodes, mais l'IA a permis d'extraire sensiblement les mêmes informations en un temps beaucoup plus réduit. La solution optimale semble être la synergie entre ces deux méthodes.

La finalité de ce screening de documentation et de validation manuelle est d'optimiser la veille documentaire et d'obtenir de façon actualisée et automatique une base de données de l'ensemble du savoir concernant les interactions chimiques, les gènes ou la biologie. Cette base étant ensuite utilisée pour identifier des cibles existantes ou nouvelles, des candidats médicaments *in silico*, les synthétiser et actualiser les connaissances.

#### 2.2.1.2. Identification des cibles moléculaires

L'identification de la cible nécessite un grand nombre d'expériences, de modèles prédictifs, appliqués à diverses phases d'optimisation, demandant une durée très importante. Sa finalité est d'établir une relation de causalité entre la modulation d'une cible et la survenue d'une pathologie.

Grâce aux apports de **l'information génomique** (*lien entre une pathologie et l'ensemble du Génome*), les **attributs biochimiques** (*caractéristiques physico-chimiques de molécules*) et la **traçabilité des cibles**, l'IA est en train de changer notre façon d'aborder l'identification de cibles moléculaires dans le traitement de pathologies.

Une plateforme nommée *Open Targets* (142) permet d'associer pour chaque pathologie les gènes affectés. Elle utilise le machine learning et des données de génétique et génomique humaine afin d'identifier et de prioriser des cibles potentielles de médicaments par le listage de l'ensemble des cibles connues et l'étude de l'importance de la relation entre la cible et la pathologie, les interactions possibles avec un médicament (*druggability* en anglais), ou les risques liés à la sécurité de la cible (effets indésirables, etc...). Toutes les données proviennent de sources validées et référencées (publications scientifiques, etc...). Cette plateforme est mise à la disposition de scientifiques permettant son exploitation dans la recherche de nouveaux traitements.

L'identification et la validation des cibles peuvent être réalisées grâce à des outils précédemment mentionnés : le traitement du langage naturel permet de "scanner" de grandes bases de données de littérature scientifique, ou médicale afin d'identifier des relations entre des gènes et des pathologies, ou identifier ces cibles (143). Enfin, certaines études peuvent ne pas avoir été menées à leur terme, il peut donc être intéressant de vérifier si une idée similaire n'a pas déjà été testée auparavant, pour la faire progresser, ou ne pas perdre de temps dans un processus voué à l'échec (144).

L'ensemble des données extraites de l'identification de la cible seront ensuite intégrées à l'étape de *Drug Design* afin d'optimiser le(s) composé(s) avec les propriétés adéquates avant d'être synthétisées.

Les données de littérature jouent ici encore un rôle très important, et les processus automatiques de *data mining* rendent accessibles des données textuelles non-structurées qui ne l'étaient auparavant pas. Ces données reposent grandement sur la base de données des effets indésirables des médicaments de l'Union européenne (EU-ADR) annotée manuellement et sur une base de données des associations génétiques annotée semi-automatiquement.

Le machine learning pourra également aider à la compréhension de phénomènes biologiques sous-jacents à l'apparition d'une pathologie, afin d'identifier des cibles potentielles. Un exemple notable est la compréhension des variations génétiques dans les signaux d'épissage alternatif (*processus conduisant à la synthèse de plusieurs protéines différentes à partir du même gène codant*) ayant permis d'identifier des variants d'épissages spécifiques à la maladie d'Alzheimer.

Une fois identifiée, il faut étudier la possibilité de synthèse d'une molécule se liant à cette cible (145). Différents attributs de la cible seront passés en revue dont la taille et la forme des sites de liaison.

*DeepDTnet* (cf. Figure 33) est issu d'une étude parue dans le journal *Chemical Sciences* en Juillet 2020. Son objectif était de créer une méthodologie de deep learning basée sur des réseaux de neurones prenant en compte une grande diversité d'informations (chimiques, cellulaires, génomiques...) afin de générer des caractéristiques biologiques et pharmacologiques et d'identifier *in silico* des cibles moléculaires pour des médicaments connus (cf. Figure 33). Ce programme a permis d'identifier des mécanismes d'effets indésirables de certains traitements, découvrir de nouvelles cibles potentielles dans certaines pathologies, ou le recyclage de certains médicaments dans de nouvelles indications.

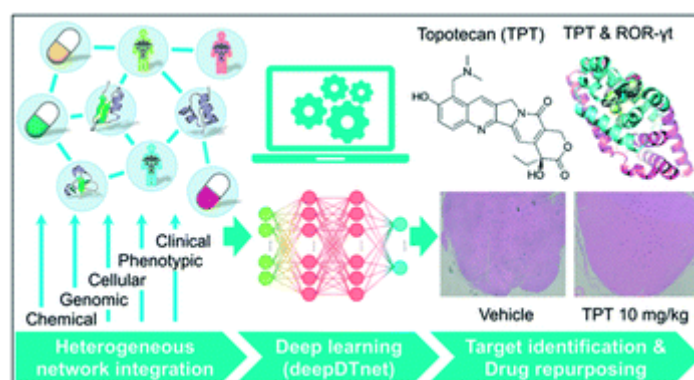


Figure 33 - Principe de DeepDTnet (146)

### 2.2.1.3. Prédiction de la structure tridimensionnelle d'une cible d'intérêt

La structure tridimensionnelle d'une protéine cible est essentielle, puisque ces dernières peuvent exister sous un grand nombre de conformations, qui permettront de comprendre leurs fonctions. Les molécules-médicaments seront ensuite modélisées en fonction de l'environnement chimique tridimensionnel du site de liaison au ligand d'une protéine cible.

Les protéines (cf. Figure 34) sont composées d'un assemblage d'acides aminés (Leucine, Glutamate, etc...), qui peuvent ensuite adopter différentes structures secondaires selon le repliement de leur chaîne principale (hélice alpha, feuillet bêta, etc...). La conformation tridimensionnelle d'une protéine correspond au repliement de l'ensemble de la chaîne dans l'espace. Différentes chaînes protéiques peuvent ensuite s'assembler pour former une structure quaternaire.

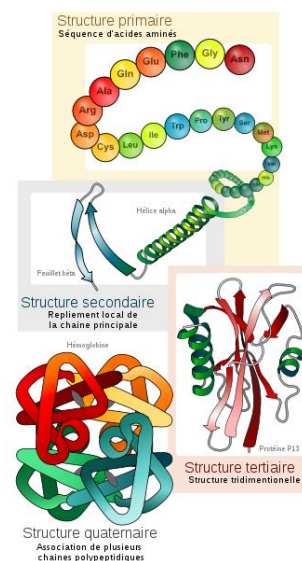


Figure 34 - Différentes structures des protéines (147)

La prédiction de la structure tridimensionnelle des protéines d'intérêts est devenue de plus en plus réaliste grâce au développement d'outils d'IA.

L'outil *AlphaFold*, créé en Janvier 2020, a été utilisé afin de déterminer la structure tridimensionnelle d'une protéine ciblée par un médicament. Le résultat a été bluffant : *AlphaFold* a prédit correctement 25 des 43 structures issues des séquences protéiques primaires qui lui avaient été données (148). Son fonctionnement se base sur des réseaux neuronaux entraînés à déduire les propriétés d'une protéine à partir de sa séquence primaire, en prédisant les distances entre deux paires d'acides aminés ainsi que les angles phi et psi ( $\phi-\psi$ ) entre les liaisons peptidiques voisines. Ces deux estimations sont ensuite combinées et utilisées pour estimer la vraisemblance d'un modèle de structure protéique tridimensionnel proposé. A partir de ces données, *AlphaFold explore* ensuite le paysage structural protéique afin de trouver des structures correspondant aux prédictions.

### 2.2.2. En recherche pharmaceutique - Criblage de molécules

#### 2.2.2.1. Criblage de molécules d'intérêt

Découvrir de nouvelles têtes de séries efficaces est un processus complexe. Cela est dû à la vaste étendue de l'espace chimique et ses  $10^{60}$  molécules parmi lesquelles il faut parvenir à sélectionner celle(s) ayant le meilleur rapport bénéfice/risque. Celles-ci seront ensuite modifiées pour en améliorer la sélectivité et affinité vis-à-vis de la cible, la pharmacodynamie ou la pharmacocinétique.

Le coût habituel d'un screening réalisé dans une bibliothèque contenant près d'un million de molécules est d'environ 50 millions de dollars, en plus du temps nécessaire à sa réalisation (149). Il est tout naturellement impossible de *screener* manuellement parmi

l'ensemble de cet espace, c'est pourquoi le machine learning et les modèles prédictifs seront utilisés afin d'identifier des molécules virtuelles se liant spécifiquement à des cibles prédéfinies, en optimisant la sécurité et efficacité de leurs caractéristiques. L'objectif final étant d'améliorer le taux de réussite de la R&D, afin de limiter les pertes d'argent dues à des échecs.

Takeda Pharmaceutical a par exemple réalisé un partenariat avec Recursion Pharmaceuticals, une start-up travaillant dans l'identification de nouveaux composés grâce au machine learning. Après un an et demi de collaboration, ils ont annoncé avoir déjà identifié des molécules efficaces dans plus de 60 indications, et déjà dans les stades d'évaluations précliniques (149).

Différents outils existent dans le criblage de molécules d'intérêt :

- IKTOS
- IN SILICO MEDECINE
- ATOMWISE

*Cette liste est non exhaustive.*

- IKTOS - Utilisation des graphes moléculaires et des outils de traduction linguistique

Une façon traditionnelle d'effectuer du *drug design* est de réaliser une description précise de la molécule que l'on souhaite synthétiser, sous la forme d'un **graphe réduit moléculaire** (cf. Figure 35). Cette représentation précise la structure chimique que doit avoir la molécule pour interagir avec son ligand. Elle joue un rôle important, car permet de réduire l'espace de recherche aux seules molécules répondant aux critères spécifiés. Les technologies d'IA peuvent assister dans la découverte de molécules correspondant au graphe réduit grâce aux technologies de **traduction linguistique**.

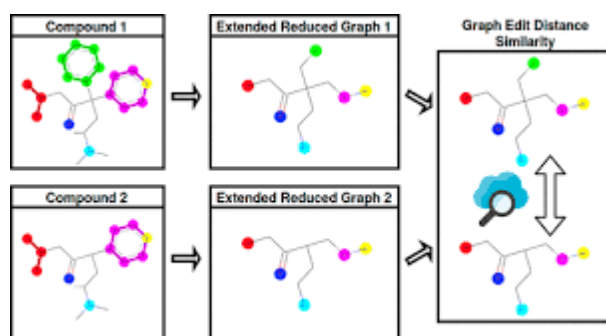


Figure 35 - Exemple de la retranscription d'une molécule avec un graphe réduit moléculaire (150)

Ce domaine mêle deux notions relatives aux réseaux de neurones :

- L'apprentissage séquence par séquence, qui permet la traduction d'un texte d'une langue à une autre ;
- Et les mécanismes d'attention, avec lesquels le modèle va se focaliser sur certains mots clés de la phrase, pour en capturer le sens global.



Ces deux notions permettent d'obtenir une traduction ayant sélectionné les bons mots et la bonne signification générale. Une molécule (ou sous graphique réduit) peut ainsi être représentée sous la forme d'un texte utilisant le langage SMILES (*Simplified Molecular Input Line Entry Specification*, permettant de décrire la structure de molécules avec des caractères simples), qui sera reconnue par l'IA, qui en captera les caractéristiques principales. En créant une base de données rassemblant des centaines de milliers de graphes réduits, il est possible d'identifier des structures répondant aux caractéristiques essentielles de la molécule recherchée (151).

Créée en 2016 par trois français, IKTOS utilise ce système pour générer des molécules synthétisables répondant à un cahier des charges précis. Il a été testé auprès d'un laboratoire pharmaceutique faisant des recherches en vain dans le domaine cardiovasculaire depuis près de 20 ans. Douze critères composaient le "cahier des charges" de la molécule idéale, aucune solution testée auparavant n'était parvenue à optimiser l'ensemble de ceux-ci. IKTOS a réussi à créer virtuellement une centaine de molécules ayant une chance plus élevée de réussite. Onze de celles-ci ont été sélectionnées pour leur meilleure "synthétisabilité" et testées : 3 ont maximisé 11 critères et une les 12, chose qu'aucune autre étude n'avait réussi à réaliser auparavant (152).

- IN SILICO MEDECINE - Generative Tensorial Reinforcement Learning (GENTRL) techniques

En Septembre 2019, une équipe conjointe de *InSilico Medecine* et de WuXi AppTec a publié une étude, *Deep learning enables rapid identification of potent DDR1 kinase inhibitors*, mettant en valeur l'efficacité des techniques d'Apprentissage par renforcement tensoriel génératif (GENTRL) dans la découverte de molécules efficaces dans la fibrose rénale. Ces techniques impliquent les réseaux antagonistes génératifs (GAN) et l'apprentissage par renforcement, et ont permis d'analyser le comportement d'un composé vis-à-vis de cellules et de sélectionner les plus efficaces en termes de bénéfices/effets indésirables (153).

Ce procédé leur aurait permis d'identifier et de synthétiser de nouveaux inhibiteurs de kinases, les inhibiteurs de la réponse aux dommages à l'ADN (DDR1) en seulement deux mois, réduisant les délais actuels de recherches et de développement précliniques de plus de 90% (*6 molécules synthétisées, dont 5 au moins aussi efficaces que les traitements existants*)(153). Ceux-ci étant réduits, de grandes économies seront réalisées sur les coûts globaux de R&D, les produits pourront obtenir l'AMM plus vite, générant des revenus pour les entreprises plus tôt. Il ne faut également pas oublier que toutes les pathologies actuelles n'ont pas encore de traitements, et que de nombreux patients attendent une solution à leurs problèmes. Eux-aussi pourront bénéficier de la mise sur le marché précoce de ces nouveaux produits.

- ATOMWISE - Utilisation des réseaux neuronaux profonds

La technologie d'*Atomwise* d'AtomNet repose sur les réseaux neuronaux convolutionnels, de grands jeux de données, et le domaine des statistiques afin d'extraire les informations de millions de mesures d'affinités expérimentales et de milliers de structures



protéiques afin de prédire la liaison d'une petite molécule à une protéine. Cela permet ainsi de découvrir de nouvelles molécules efficaces ou de prédire la toxicité de certains médicaments lorsqu'ils se lient à plusieurs cibles distinctes.

La priorité dans la découverte d'un nouveau Principe Actif est qu'il soit efficace et sûr. Les équipes testent des centaines de molécules afin de trouver un compromis entre bénéfices thérapeutiques et effets indésirables. Le résultat final découle souvent soit d'une déduction logique, soit de la chance, quand le développement arrive à son terme (ce qui n'est pas toujours le cas). Les ordinateurs ainsi que les technologies d'IA sont très efficaces dans la résolution de ce type de problèmes multifactoriels.

*Atomwise* peut analyser un très large espace chimique (et des milliards de composés), pour identifier des molécules hautement sélectives et spécifiques pour une cible donnée (154). Il représente le site de liaison entre une protéine et son ligand comme un espace tridimensionnel comprenant les différents atomes, afin d'apprendre de façon autonome les règles régissant la liaison moléculaire. Cette solution permet de s'affranchir des étapes traditionnelles manuelles d'ajustement faites dans les méthodes de calcul habituelles et ainsi prédire le comportement et la liaison d'un ligand avec une protéine. Cette technologie permet d'estimer la probabilité qu'un médicament soit efficace dans une certaine indication. Cela réduit le champ des possibilités et concentre la recherche et les tests sur les molécules ayant le plus de chance d'obtenir des résultats concluants.

De nombreuses entreprises ont donc établi des partenariats pour booster leur R&D pharmaceutique :

- GSK avait déjà établi un partenariat avec InSilico et ExScientia en 2017 afin de renouveler son portefeuille de molécules dans des secteurs thérapeutiques concurrentiels.
- Merck et Atomwise
- AstraZeneca et Novartis avec Benevolent AI ...

En outre, la première molécule issue du travail d'une IA, nommée DSP-1181, est entrée en Février 2020 en phase I des études cliniques, pour en étudier la sécurité d'usage dans le traitement des troubles obsessionnels compulsifs. Née de la collaboration entre ExScientia et Sumitomo Dainippon Pharma, elle n'a nécessité que 12 mois pour être découverte et optimisée.

#### 2.2.2.2. Prédiction de caractéristiques d'une molécule

Les outils d'IA peuvent également être utilisés afin de simuler *in silico* (numériquement) les caractéristiques des molécules (mode d'action, absorption, distribution, métabolisation, élimination, toxicité, et éventuellement efficacité) liées ou non à leur liaison à la cible, avant leur synthèse et leur expérimentation *in vivo* au cours des essais cliniques.

Une des principales caractéristiques à maîtriser au cours du développement d'un nouveau médicament est sa toxicité. Son optimisation est par ailleurs une des étapes les plus consommatrices en temps et en argent des étapes précliniques.

Des plateformes existent d'ores et déjà dans ce domaine :

- PrOCTOR, qui étudie la probabilité de toxicité de molécules utilisées au cours d'essais cliniques.
- DeepTox, qui utilise le machine learning afin de prédire la toxicité de nouvelles molécules ;

L'algorithme de Machine Learning DeepTox normalise tout d'abord la représentation chimique des composés, dont des caractéristiques chimiques (nombre d'atomes, surface, etc...) seront utilisées en entrée des méthodes de machine learning. Le logiciel étudie ensuite la structure et tente d'identifier d'éventuels **toxicophores**, terme anglais désignant les parties des molécules responsables des effets toxiques de la molécules, ou d'autres caractéristiques physicochimiques, en comparant cette structure avec leur base de données (contenant plus de 2500 toxicophores). *Les toxicophores s'opposent aux pharmacophores, responsables de l'action pharmacologique d'une molécule.*

Il a obtenu d'excellents résultats au Tox21 Data Challenge, événement au cours duquel les participants devaient prédire les effets toxiques de 12.000 produits (environnementaux, chimiques, ou médicaments)(149).

Cela permet de réaliser de grandes économies dans les délais et les coûts de développement des molécules et d'en réduire le taux d'échec. En effet, un nombre réduit de molécules seront finalement synthétisées, avec de plus grandes chances de réussite aux essais sur l'Homme. Ces nouvelles technologies ouvrent la possibilité de nouvelles molécules avec moins d'effets secondaires, rendant leur observance plus facile pour les patients.

#### 2.2.2.3.Synthèse de la molécule

L'IA peut également être un atout pour prioriser une molécule selon sa **facilité de synthèse**. Des réseaux de neurones peuvent en effet être utilisés pour prédire des **voies de rétrosynthèse** (cf. Figure 36), à la base de la capacité de synthèse d'un candidat médicament. L'approche rétrosynthétique se compose de deux phases principales :

- La première est l'analyse de la molécule d'intérêt et sa décomposition en de plus petits fragments ("blocs de constructions") plus faciles à fabriquer ou à se procurer (représenté sous la forme d'un *arbre de recherche*).
- La seconde correspond en l'analyse des réactions chimiques qui vont transformer ces fragments en la molécule recherchée.

Cette étape est la plus complexe au vu de la quantité des réactions chimiques existantes dans la nature, et parmi lesquelles il faut choisir la plus adaptée. L'IA pourrait aider ici à prédire les réactions les plus adaptées en comblant les vides (*plus souvent consécutifs d'effets stériques et électroniques imprédictibles ou d'un manque de maîtrise de la réaction*) causant un taux d'échec élevé lors de la synthèse de la molécule.

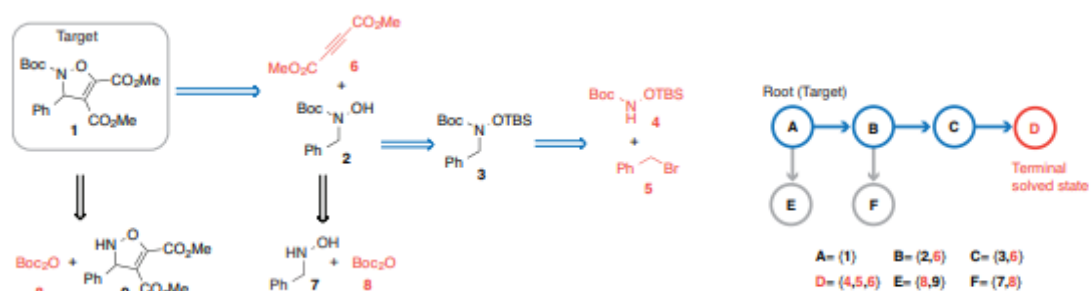


Figure 36 - Exemple d'une voie de rétrosynthèse et d'un arbre de recherche (155)

La finalité est d'estimer la difficulté de fabrication du médicament, en perfectionnant la planification et en éliminant les solutions non viables économiquement ou en pratique.

Des systèmes de synthèse de composés organiques assistés par ordinateur existent sur le marché pour assister les chimistes dans leur sélection des voies de synthèse. Parmi ceux-ci, une nouvelle plateforme, **3N-MCTS** (*3N-Monte Carlo Tree Search*), combinant trois réseaux de neurones différents, qui filtrent les "blocs de construction" les plus prometteurs et ne conservent que les réactions chimiques bien connues dans la synthèse des molécules d'intérêt (156). Le premier réseau neuronal transmet les connaissances relatives à la synthèse des molécules afin d'orienter la nouvelle. Le second réseau valide la faisabilité de la réaction proposée, tandis que le dernier réseau simule le résultat de ces réactions, et affecte à chacun une valeur. Cela permet d'orienter la recherche dans la direction la plus facilement et efficacement réalisable.

Le modèle a été entraîné à partir de la base de données Reaxys, contenant l'ensemble des réactions publiées dans l'histoire de la chimie organique (plus de 11 millions de réactions).

L'outil construit un arbre de recherche, composé de différents nœuds, auxquels sont affectés des valeurs, représentant leur "qualité" (*une position intéressante sera privilégiée au détriment d'une qui ne l'est pas*). Puis quatre étapes itératives se succèdent, en partant de la molécule d'intérêt et en testant progressivement un grand nombre de positions possibles pour des précurseurs que l'on sait préparer, jusqu'à obtenir la voie de synthèse ayant le plus de chance d'aboutir.

Cette méthode serait 30 fois plus rapide que les modèles traditionnels de recherche par ordinateur pour la majorité des molécules testées (155).

#### 2.2.2.4. Recyclage de la molécule

La réutilisation de médicaments existants pour une nouvelle pathologie est très avantageuse. En effet, ce dernier est directement autorisé à réaliser les **essais de phase II** des essais cliniques pour une nouvelle indication, sans forcément avoir à repasser les essais de phase I et de toxicologie (157). L'exploitation automatisée des données de littérature (dossiers médicaux, publications scientifiques, etc...) permettra d'identifier des ligands potentiels de cibles identifiées déjà utilisés dans d'autres indications actuellement. Cela permettra aux

entreprises propriétaires de ces molécules de disposer de revenus supplémentaires, sans grands efforts à fournir.

Cela donne également de l'espoir pour des pathologies orphelines, pour lesquelles développer des nouvelles molécules n'est pas économiquement profitable.

Le *Massachusetts Institute of Technology* a par exemple annoncé en Février 2020 avoir réussi à développer, grâce à une IA, un antibiotique (l'halicine) efficace contre une bactérie multirésistante, *Clostridium difficile* (158). Cette molécule dispose d'une structure chimique jamais envisagée auparavant dans ce domaine, et pour cause : elle avait initialement été envisagée dans le traitement du diabète, avant d'être abandonnée à cause de mauvais résultats lors des études.

### 2.2.3. Sécuriser la recherche

Les opérations de recherche obligent la manipulation de produits chimiques pouvant avoir des atteintes respiratoires, cutanées, ou digestives. Afin de limiter ces risques, des équipements de protection individuelle (lunettes, gants, masques...) doivent être portés par tous les opérateurs. Des outils de reconnaissance visuelle pourront être utilisés afin de s'assurer que tous portent ces EPI de façon correcte, que les mesures d'hygiène soient respectées et de prévenir les risques d'incidents potentiels (incendie, accident...).

## 3. Intelligence Artificielle dans la Recherche Clinique

### 3.1. Définition de la Recherche Clinique

La phase d'essais cliniques (EC) est une des phases les plus importantes dans le développement d'un nouveau produit de santé, en effet, elle seule permet l'obtention de l'Autorisation de Mise sur le Marché. En résumant au maximum : la phase I étudie la tolérance et les effets indésirables de la molécule, la phase II la dose à administrer, tandis que la phase III évalue son efficacité en la comparant à un placebo. La phase IV correspond à la surveillance post-AMM : la pharmacovigilance (*cf. Figure 37*).

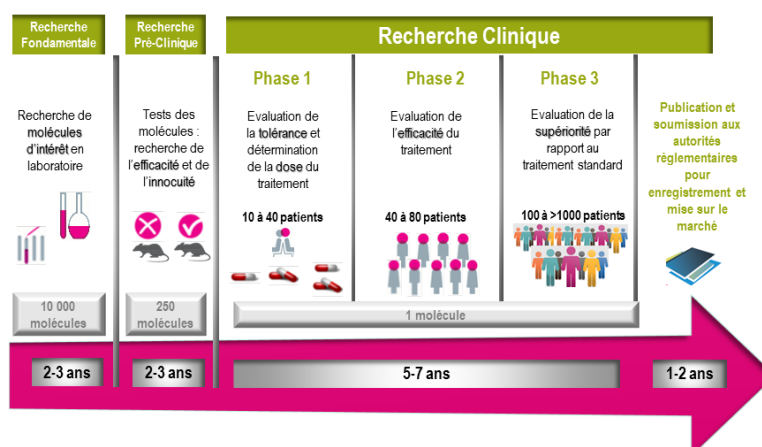


Figure 37 - Schéma des différentes phases de la recherche clinique (159)

Selon le Centre Tufts basé à Boston, **plus d'un tiers des essais cliniques ne vont pas à leur terme**, alors qu'ils représentent à eux seuls près de 40% des coûts de développement du médicament. De plus, le taux de réussite des essais cliniques pour les nouveaux médicaments n'oscille qu'autour de 10%, mettant en relief le grand nombre d'échecs (160).

Une des priorités pour l'industrie pharmaceutique est donc de limiter le nombre d'échecs au cours de ces essais pour réaliser de grandes économies de temps et d'argent.

Le machine learning est d'ores et déjà utilisé pour optimiser les critères de sélection et la taille des panels des essais, tandis que le traitement du langage naturel permet d'analyser automatiquement un grand nombre de sources diverses pour analyser les informations génétiques et cliniques d'une pathologie, et ainsi cibler de façon plus efficace les volontaires d'une étude, et améliorer l'efficacité du traitement (*en limitant le nombre de non-répondeurs*). La modélisation *in silico* pourra réduire le temps et les coûts des décisions prises dans la phase préclinique tout en augmentant la probabilité de réussite des composés testés.

### 3.2. Projets en cours de déploiement

#### 3.2.1. Identification & recrutement des patients

La première étape dans la réalisation d'un essai clinique est la plus importante : sélectionner des patients non-adaptés peut mettre en péril la fiabilité de l'étude avant même son début. Un patient était sélectionné après qu'un médecin et d'autres intermédiaires ont scrupuleusement analysé son dossier médical à la main, processus long et fastidieux. Une étude réalisée a conclu que 80% des essais cliniques prennent du retard au cours de cette phase, et qu'un tiers des interruptions des phases III sont dues à des difficultés de recrutement (161). L'intelligence artificielle aurait ici encore un rôle à jouer dans l'analyse des dossiers médicaux et l'aiguillage des patients vers les essais cliniques les plus adaptés à leur profil.

*MendelAI* (cf. Figure 38) est une entreprise créée par Karim Galil et Wael Salloum, qui utilise le machine learning pour analyser et comprendre les données contenues dans les dossiers médicaux de candidats à des essais cliniques et celles de dossiers d'essais cliniques extraits du site internet clinicaltrials.gov. L'application passe en revue les dossiers des patients, et met en valeur les profils les plus adaptés à un essai particulier selon les critères définis par le promoteur de l'EC (avec un pourcentage de compatibilité pour chaque test).

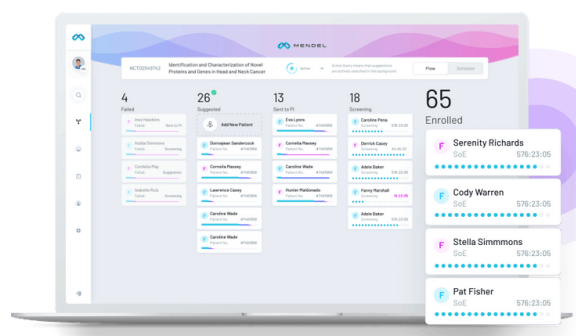


Figure 38 - Interface de Mendel.ai (162)

Les recherches qu'un patient doit réaliser pour trouver un essai clinique qui lui correspond sont très difficiles, et peuvent les décourager. Selon le Docteur Galil, directeur de *MendelAI*, « *un patient atteint d'un cancer du poumon, par exemple, pourrait trouver 500 essais potentiels sur clinicaltrials.gov, chacun d'entre eux ayant une liste unique et exhaustive de critères d'éligibilité qui doivent être lus et évalués. Comme ce pool d'essais change chaque semaine, il est humainement impossible de garder une trace de toutes les concordances* » (163).

La société *Clariness* a établi un partenariat avec Sanofi pour améliorer son recrutement de patients dans les essais cliniques. Celle-ci utilise le marketing cible afin de rediriger les patients vers la plateforme *ClinLife* selon le type de recherches qu'ils effectuent. Mais cette démarche est coûteuse (achat des encarts publicitaires) et exige une certaine participation du patient, ce que *MendelAI* ne demande pas.

Enfin, identifier la relation entre les biomarqueurs humains et les phénotypes in vitro conduit à une évaluation plus réaliste et quantifiable de l'incertitude dans la réponse thérapeutique d'un patient en particulier (en recrutant des populations spécifiques en phases II et III), ouvrant la voie à une possible prédiction *a priori* du succès dans la réponse thérapeutique.

Comme mentionné dans la partie précédente, la prédiction de la réussite aux essais cliniques sera intégrée dans les étapes précoces de R&D, afin de minimiser les risques d'échecs. Certains prédicteurs de succès ou d'échecs ont été découverts par des analyses scientifiques : l'expression génique témoignant de l'efficacité d'une cible par exemple, caractérisée par une faible expression d'ARN et une variance élevée entre les tissus (145).

### 3.2.2. Amélioration de l'observance des patients pendant les essais cliniques

La non-observance est une problématique de santé publique en expansion. Elle correspond en l'inadéquation entre une prescription médicale et le comportement d'un patient. L'OMS estime à 50% (164) la part des traitements prescrits non ou mal suivis par les patients, ce qui conduit à de conséquentes répercussions économiques et médicales : 9 milliards d'euros pourraient être économisés annuellement en France selon IMS Health (165). Les essais cliniques ne sont pas épargnés par ce fléau. Le suivi du protocole est primordial à son bon déroulement, et les organisateurs doivent se fier aux participants, ils n'ont que de peu de moyens de vérification (retours de blisters, dosages biologiques). Certaines start-up se sont penchées sur ce sujet et proposent dorénavant des solutions pour améliorer l'observance au cours des essais cliniques et donc limiter l'augmentation des biais statistiques.

#### 3.2.2.1. AiCure

La plateforme d'analyse de données *AiCure* permet de monitorer à distance le comportement des patients prenant part à une étude clinique. Elle réduit les délais et les coûts des études par le suivi à distance de l'engagement du patient ainsi que la mesure de biomarqueurs et la surveillance en temps réel de leur dosage.

Le logiciel utilise la caméra du smartphone ou de la tablette sur laquelle il est installé pour identifier le visage du patient, le produit à ingérer, confirme la prise du bon médicament et



transmet instantanément ces données au prescripteur ou à la compagnie réalisant l'essai clinique. Ed Ikeguchi, le responsable d'*AiCure*, a déclaré : *“de nos jours, les promoteurs sont obligés de recruter un nombre croissant de patients pour contrer les effets de la non-observance au cours des essais cliniques. Notre solution permet de rendre les essais plus rationnels et plus courts ce qui conduirait à de grandes économies financières et à une arrivée plus rapide sur le marché de médicaments vitaux”* (166).

En plus de surveiller la prise de traitement, *AiCure* peut envoyer des alertes aux patients (volontaires) si ces derniers sont sur le point d'oublier leur médicament (non-observance non-intentionnelle), les mettre en contact avec leur prescripteur en cas de problème, ou prévenir les investigateurs en cas de non-observance intentionnelle. Il est très facile pour l'outil de distinguer la non-observance intentionnelle de celle qui ne l'est pas. Un oubli involontaire sera caractérisé par une absence d'ouverture de l'application et d'enregistrement de la prise, tandis qu'une non-observance délibérée passera par une utilisation de l'application et une fausse ingestion du médicament (retrait du médicament de la bouche, etc...).

*AiCure* a réalisé un grand nombre d'études, sur près de 3000 participants, pour évaluer le phénomène de non-observance volontaire : 4% de l'ensemble des doses prescrites sont non administrées volontairement, et près de la moitié des sujets ont manqué une prise intentionnellement. Les patients en phase avancée d'essais cliniques (phases II et III) ont plus tendance à manquer des prises, par lassitude, ou scepticisme vis-à-vis des effets indésirables.

Les études ont également démontré que la première prise était la plus importante : en effet, si un sujet la manquait, son taux de non-observance intentionnelle était multiplié par 5. C'est pourquoi comprendre les raisons qui sous-tendent la non-observance est primordial pour les responsables d'essais cliniques, afin de les aider à développer des ressources éducatives ciblées et à réaliser des interventions personnalisées pour réduire son impact sur les résultats de l'étude.

Il sera possible d'évaluer les participants potentiels à une étude avant leur enrôlement via une utilisation de l'application avec un placebo pendant une certaine durée, afin de surveiller des comportements suspects (à risque de ne pas suivre le traitement). Cette stratégie conduira à un enrichissement des panels de test, puisque seuls seront randomisés les patients dont on se sera assuré la bonne adhérence au traitement *a priori*.

Les résultats obtenus n'en seront que plus représentatifs (*moins de variabilité*), et les coûts pour les organisateurs, réduits. En effet, à l'heure actuelle, les organisateurs ont tendance à augmenter la taille de leurs panels de 15% pour réduire l'impact de la non-observance sur la fiabilité des résultats statistiques, rendant les études plus longues et plus coûteuses (*surcoût évalué entre 5 et 7 million de dollar pour une étude clinique de phase III*)(167).

Cette application a déjà été utilisée au cours de certains essais cliniques, et a démontré des résultats encourageants dans le suivi du traitement. Elle a par exemple aidé la compagnie AbbVie à améliorer l'adhérence à son traitement contre la schizophrénie au cours d'une étude de phase II de 6 mois. Selon le directeur d'*AiCure*, la participation de patients atteints de schizophrénie tourne en moyenne autour de 50%, tandis que les sujets ayant utilisé *AiCure* ont



eu une adhérence moyenne de 90% (contre 70% pour la méthode alternative, d'observation directe de la thérapie, sans AiCure)(168).

En plus d'être utilisée au cours des essais cliniques, cette plateforme pourra tout à fait l'être pour suivre l'observance en pratique courante de pathologies chroniques.

#### 3.2.2.2.Catalia Health – Mabu

La start-up Catalia Health s'est également penchée sur ce problème d'observance thérapeutique avec son assistant virtuel Mabu (appelé également *chatbot*). Connaissant les patients, leur(s) pathologie(s) et traitement(s), l'interface peut effectuer le suivi du traitement, converser avec les patients, les conseiller et les avertir en cas d'oubli de prise. L'IA pourra, avec le consentement de l'utilisateur, transmettre des informations au prescripteur (effet indésirable, diminution de l'observance...), afin de les tenir au courant de l'avancée du traitement, sans qu'une consultation ne soit nécessaire.

L'interface (*robot physique ou application sur smartphone*) se compose ainsi d'un aspect médical mais également d'un aspect psychologique, afin d'accompagner les patients au quotidien dans leur pathologie chronique, ou lors d'un essai clinique. Des partenariats ont également été établis directement avec des compagnies pharmaceutiques, telles que Pfizer, afin de suivre les effets d'un traitement en temps réel, afin de mieux appréhender l'expérience du patient.

En plus d'optimiser l'observance du traitement, Mabu, au contraire d'*AiCure*, permet aux investigateurs des essais cliniques de comprendre les raisons pour lesquelles un patient n'a pas pris son traitement. Ils peuvent également mettre en valeur des biais ou des facteurs influant sur l'efficacité du traitement en collectant et analysant les données d'alimentation et d'activité des patients.

Avoir une interaction privilégiée avec un *chatbot* semble disposer de nombreux avantages : le délai entre l'apparition d'un effet indésirable et sa signalisation pourrait être réduit (le patient n'aurait plus à attendre un rendez-vous pour le mentionner), de plus, certains patients peuvent être plus enclins à discuter de certaines thématiques avec un interlocuteur considéré *non-humain* comme une étude parue en 2014 "*It's only a computer: Virtual humans increase willingness to disclose*" l'affirme (169).

#### 3.2.2.3.Sensely

Participer à une étude clinique ne signifie pas uniquement prendre un traitement. Cela implique également un grand nombre de déplacements chez des médecins pour réaliser des analyses sanguines régulières afin de suivre l'évolution du traitement. Arriver à récupérer et analyser ces données à distance permettrait de limiter ces déplacements au maximum.

*Sensely* est une petite entreprise basée à San Francisco, fondée en 2013 qui a réfléchi sur cette problématique. C'est une interface disponible sur smartphone, sous la forme d'une infirmière virtuelle qui peut réaliser un suivi personnalisé de patients en collectant en temps réel des

informations cliniques provenant à la fois de dispositifs médicaux (biocapteurs) ou du smartphone du patient afin d'en suivre les symptômes et d'alerter le médecin référent en cas de besoin. *Sensely* prétend permettre d'éviter plus de 24% des appels et rendez-vous inutiles, et améliore le suivi des essais cliniques par les patients (*test réalisé avec 91 patients, pour un taux de succès de 87%*)(170).

C'est cette efficacité qui a convaincu Novartis de les choisir pour leurs essais cliniques en cours de réalisation (171).

### 3.2.3. Sécuriser les données issues des essais cliniques

Selon *Protenus*, une entreprise d'analyse de la conformité des soins, plus de 41 millions de données de patients ont fuité en 2019 rien qu'aux Etats-Unis (*une brèche ayant concerné à elle seule 21 millions*), contre trois fois moins en 2018. Le nombre de brèches de sécurité dues à du piratage est en hausse depuis 2016 de 20% (*450 en 2016 contre 572 en 2019*), avec des *hackeurs* de plus en plus créatifs.

En plus de la problématique de l'observance au cours des essais cliniques, celle de la sécurité des données est donc plus que jamais d'actualité. Les systèmes actuels de protection fonctionnent avec des pare-feu et des antivirus qui exigent de nombreux audits chronophages et très coûteux. L'IA pourra permettre de réaliser une surveillance en temps réel de l'intégrité des données.

*Protenus* a créé une plateforme d'IA assurant la surveillance en temps réel des données contenues dans les dossiers médicaux électroniques (compte-rendu de rendez-vous, historique clinique...). Celle-ci permet **d'automatiser** des tâches auparavant manuelles (recueil et analyse de données, priorisation et résolution des menaces les plus importantes très rapidement), **l'audit** de chaque accès (analyser qui a eu accès à quelles informations, et si des connexions à risque ont été observées), de protéger l'intégrité des données les plus sensibles et **d'alerter** en cas de brèche de sécurité pour une résolution plus rapide.

Cette plateforme est déjà utilisée par l'Institut Johns Hopkins, pour surveiller les situations à risque.

## 4. Intelligence Artificielle dans la Production

Selon le site du LEEM, "*la production pharmaceutique regroupe l'ensemble des opérations de transformation des matières premières en produits finis (médicaments)*" (172). Cela inclut donc à la fois la phase de fabrication, celle de conditionnement mais également tout le domaine de la qualité qui gravite autour. La France est le quatrième pays producteur de médicaments en Europe (derrière la Suisse, l'Allemagne et l'Italie) après avoir été premier pendant plus de vingt ans avant 2008.

L'IA peut avoir un rôle à jouer dans l'optimisation des procédés, le contrôle qualité ou la maintenance préventive.

#### 4.1. L'IA et la Production Pharmaceutique (fabrication / conditionnement)

##### 4.1.1. L'optimisation des processus

La production digitalisée peut améliorer la **productivité**, la **robustesse** des procédés et des équipements, la prévision et planification des activités, grâce à des capteurs, ou des objets connectés collectant des données issues d'opérations en cours et de lots précédents intégrées dans des systèmes centralisés de collecte et d'analyse des informations en temps réel. Cela permet une analyse profonde, des prédictions réalistes et un contrôle des procédés plus performant.

De plus en plus d'outils innovants apparaissent : processeurs ultrarapides, solutions de stockage de données peu coûteuses et pratiques, algorithmes d'analyses et de contrôles, etc... afin de recueillir et analyser des données de diverses sortes. A partir de cela, les organisations peuvent savoir exactement ce qu'il se passe en temps réel et s'adapter à toute dérive de leurs procédés.

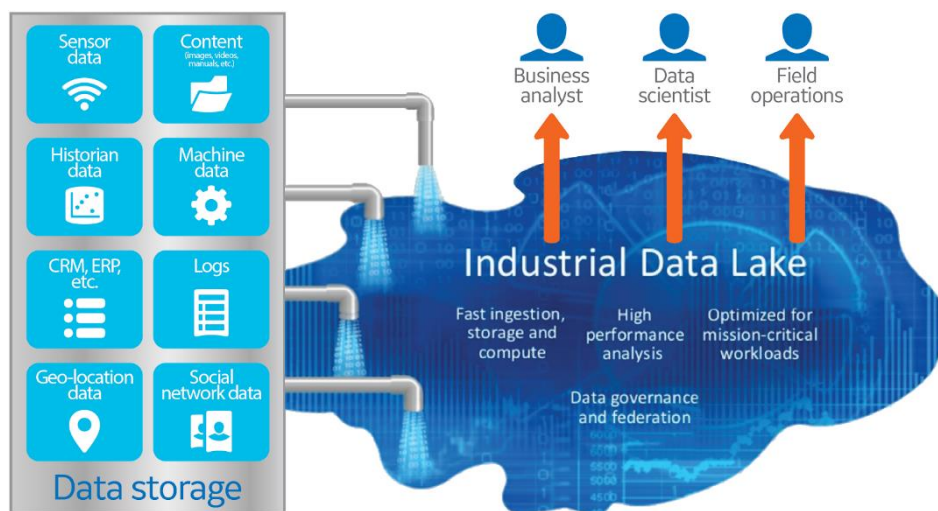


Figure 39 - Concept d'Industrial Data Lake (173)

Les données proviendront d'une grande diversité de sources (capteurs, données historiques, données des machines, progiciels, etc...) et seront rassemblées dans un "lac de données industrielles" (cf. Figure 39), nécessitant d'importantes capacités de stockage, d'analyse et de calcul. C'est à partir de ce lac que différentes informations pourront être extraites :

- Les managers pourront suivre la cadence de production, les goulots d'étranglement, les temps morts, etc....
- Les *Business Analysts* pourront suivre les tendances de commande et les comportements des clients ;

Un exemple de ce type de solution est la plateforme *Predix* de Cytiva Digital (174). Elle recueille l'information en continue grâce à des capteurs, la stocke dans le cloud, et l'analyse pour fournir une plateforme de contrôle et d'amélioration des procédés aux producteurs de médicaments. Elle prétend réduire de façon significative les stocks et les coûts de maintenance, et décharger les systèmes informatisés des tâches annexes pour se concentrer sur les activités de base, le tout en respectant les législations relatives au stockage de données. Reckitt Benckiser

et Lek Pharmaceuticals ont déjà adopté cette solution pour suivre et optimiser leurs procédés de fabrication

L'IA permettra aussi l'**automatisation** rigoureuse de certaines tâches de production répétitives, assurant une qualité continue (*en limitant les biais liés à l'intervention humaine*) et redirigeant la force de travail humaine vers des activités plus complexes et à plus forte valeur ajoutée.

La production pourrait être (presque) entièrement automatisée : une machine pourra analyser un bon de commande digital, en planifier la production et commander le matériel nécessaire, puis surveiller et gérer la fabrication, le conditionnement et le stockage, et enfin l'envoi du produit au client. Les bénéfices seront multiples : meilleure utilisation du matériel et des matières, production plus rapide et efficace, et respect constant des **attributs qualité critiques** (*propriété devant obligatoirement être comprise dans un intervalle de valeur pour assurer la qualité du produit fini*).

L'analyse des étapes de production rendue possible grâce aux capteurs permettra d'optimiser des procédés divers (compression, granulation ou synthèse chimique), en éliminant des opérations spécifiques inutiles ou en les combinant grâce à des algorithmes de deep learning dans des équipements capables de réaliser différentes tâches simultanément (*multi-tasking*). De plus, l'analyse des attributs de fabrication et des caractéristiques des produits de chaque lot rendra possible la sélection de la recette de fabrication donnant le dosage final répondant au mieux au cahier des charges.

Selon une étude réalisée par Siemens à l'occasion du forum mondial de l'industrie, les bénéfices de la digitalisation des chaînes de production pharmaceutiques dans le monde sont estimés à plus de 60 milliards d'euros de gain, et une économie d'un à deux milliards d'euros rien qu'en France ! (175)

L'entreprise française Sanofi a par exemple utilisé l'IA (*et dépensé par la même occasion 290 millions d'euros*) dans sa nouvelle unité basée à Framingham, aux États-Unis, afin de créer un **jumeau numérique**, c'est-à-dire une usine entièrement digitalisée (176). Elle se compose de 10.000 mètres carrés entièrement connectés et modulables : aucune ligne de production fixe, elles seront adaptables à chaque type de produits manufacturés, afin d'offrir une plus grande flexibilité de fonctionnement. L'objectif est de produire en continu, sans temps mort, en intégrant l'ensemble des données recueillies au cours des diverses étapes du processus (des premiers aux lots industriels).

#### 4.1.2. Le contrôle des paramètres de production en temps réel

Le **contrôle des paramètres critiques de production** est directement lié à l'optimisation des procédés de production ; par exemple l'utilisation de capteurs permettant de suivre en temps réel la température et le pH de certaines productions pour détecter tout écart avec la norme, et pouvoir rectifier le tir avant de dépasser le seuil à partir duquel le lot ne peut être libéré. Novartis a mis en place une telle démarche sur trois sites industriels pour tester ces technologies (177). La surveillance de ces paramètres s'étend aussi à la qualité de l'eau, matière

première massivement utilisée au cours de la production mais aussi et surtout lors du nettoyage des équipements et des outils.

Les coûts liés aux pannes ou dysfonctionnements des équipements sont très importants au sein de l'industrie pharmaceutique. Des outils de reconnaissance visuelle peuvent permettre de perfectionner les **systèmes de maintenance prédictive et de contrôle qualité** en surveillant en temps réel les machines (systèmes SCADA : *Supervisory Control And Data Acquisition*). Des caméras et des capteurs pourront monitorer en continue les chaînes de production et alerter les utilisateurs de toute pièce endommagée, usée, ou de défauts ponctuels ou récurrents sur les produits manufacturés. La détection serait instantanée, les temps d'arrêts réduits, optimisant les rendements des productions pharmaceutiques, tout en réduisant considérablement les frais de maintenance et ceux de production.

Une analyse réalisée par Siemens estimait à 30% la réduction des temps d'arrêts grâce à la numérisation des lignes de production (175).

#### 4.2. L'IA et le Contrôle Qualité des Produits

##### 4.2.1. Définition & problématiques actuelles

Les exigences essentielles d'un médicament sont simples : il doit être efficace, exempt de danger pour l'utilisateur et de qualité. C'est pourquoi des contrôles ont lieu tout au long de sa fabrication (arrivée et validation des matières premières, contrôles en cours et en fin de production...) pour en garantir la conformité avec les normes en vigueur. A l'heure actuelle, ces contrôles sont encore parfois faits par l'Homme, ce qui est très chronophage et ne garantit pas toujours une fiabilité complète.

##### 4.2.2. Contrôles en cours de production et des produits finis

L'intégration d'outils d'IA aux Technologies Analytiques des Procédés (TAP), permettra de limiter les tests destructifs en réalisant des tests directement sur les lignes de production, et de les adapter en cas de défauts récurrents. Les mêmes paramètres sont testés mais de manière différente (*uniformité de teneur en PA au travers de la qualité de mélangeage, etc...*).

Les outils d'IA pourront également être utilisés pour perfectionner les contrôles réalisés en cours de production. Des caméras, couplées à des modèles d'apprentissages automatiques, sont à la disposition des entreprises pharmaceutiques pour assurer :

- La détection automatique des défauts ou des non-conformités sur les lignes de production de façon automatique ;
- La détection des anomalies microscopiques, non détectables par l'œil humain ;
- Et la classification et l'analyse de ces anomalies, par l'élaboration de modèles prédictifs, permettant l'ajustement des procédés avant l'apparition de ces défauts.

Tout écart ou non-conformité détectée pourra ensuite être étudiée grâce à des algorithmes intelligents, capables d'identifier la(les) cause(s) racine(s), de mettre en place des actions

correctives nécessaires et de garder une trace écrite de ces anomalies. Les procédés liés à ces anomalies sont ensuite étudiés, afin d'étudier son impact sur la conformité réglementaire de l'entreprise au regard des normes et des procédures internes.

*A l'heure actuelle, la principale préoccupation est d'éliminer les produits non-conformes, tandis que les technologies d'IA s'attacheront plutôt à en supprimer les causes.*

#### 4.3. La Gestion des Risques

La gestion des risques est une étape primordiale, dans des systèmes qualité, pour la protection des patients. L'ICH Q9 (178), relatif à la gestion des risques, intègre différentes étapes dans ce processus : l'identification des risques, leur analyse, évaluation, leur contrôle (acceptation ou réduction du risque) et finalement leur réévaluation périodique.

Fondée en 2010, la société éditrice de logiciels *Novasecur* a créé un outil d'IA de gestion des risques. Selon son PDG, Cédric de Serpos, cette solution est capable d'identifier des risques autrement "invisibles", de lier différentes situations entre elles et d'en extraire des situations à risque et de proposer des solutions opérationnelles grâce aux prédictions de son système.

L'**identification des risques** est permise grâce aux importantes bases de données, créées et entretenues à partir des paramètres des procédés, des configurations des machines et des données historiques des situations précédentes. *Novasecur* est ensuite en mesure de créer une cartographie de l'ensemble des risques propre à l'entreprise, de les analyser et de proposer des actions correctives et préventives à mettre en place pour les éviter. L'IA met en place des arbres de décisions propres à chaque situation à risque, et recommande la solution ayant la probabilité d'obtenir le risque "zéro" la plus forte, sous la forme de tableaux de bord. Le modèle est régulièrement mis à jour avec les référentiels en vigueur, les procédures internes et autres documents complets et validés.

L'analyse de risque est le plus souvent réalisée "à la main", et ce sont beaucoup d'informations qu'il faut prendre en compte. Il peut être difficile de parvenir à établir des liens entre différentes situations qui ne paraîtraient pas liées au premier regard. L'IA permet de supprimer ce biais humain et de repousser les limites humaines, mieux identifier, et anticiper les risques.

Les analyses de risques sont souvent des processus chronophages, impliquant un grand nombre d'acteurs, souvent incomplets. Avoir l'aide d'un logiciel d'IA dans ce domaine permettra d'assurer la conformité réglementaire, de gagner du temps, de la robustesse dans les procédés, et d'assurer la qualité des produits finis.

## 5. Intelligence Artificielle dans la Logistique Pharmaceutique

### 5.1. La logistique pharmaceutique

La chaîne d'approvisionnement, ou *Supply Chain*, est définie par la norme AFNOR NF X50-600 comme une fonction "*dont la finalité est la satisfaction des besoins exprimés ou latents, aux meilleures conditions économiques pour l'entreprise et pour un niveau de service déterminé. Les besoins sont de nature interne (approvisionnement de biens et de services pour*



*assurer le fonctionnement de l'entreprise) ou externe (satisfaction des clients)'' (179). La finalité de la logistique pharmaceutique est de stocker et distribuer les médicaments du fournisseur au point de vente final en respectant les spécifications de conservation, impliquant donc une grande diversité de métiers.*

L'industrie pharmaceutique se heurte à de nombreuses problématiques diverses : nouvelles réglementations, pénuries de matières premières, variations non anticipées de la demande, etc.. auxquelles l'ensemble de la chaîne logistique doit s'adapter. Elle doit assurer une traçabilité complète et une gestion des stocks efficace, tout en respectant des délais de livraisons de plus en plus serrés.

La logistique pharmaceutique fait face à deux enjeux majeurs :

- La **lutte contre les médicaments contrefaits**, dont l'importance croît d'années en années. C'est pour cela que la Directive 2011/62/UE a été publiée, et impose l'utilisation de dispositifs de sécurité (sceaux d'inviolabilité) et traçabilité sur les boîtes de médicaments.
- La **sérialisation des produits pharmaceutiques**, permettant d'identifier une boîte d'une autre (et non plus un lot d'un autre) en leur attribuant un code Datamatrix 2D unique. Cette obligation s'étend à l'ensemble de l'espace européen et permet de limiter la circulation de contrefaçons.

Une fois encore l'IA fait naître bon nombre d'espoirs dans le domaine de la *Supply Chain* : réduction des erreurs de livraisons, automatisation des processus simples et répétitifs, augmentation de l'efficacité des employés, etc... Sa finalité sera d'anticiper le comportement des utilisateurs finaux afin de répondre exactement à leur besoin, sans rupture ni excès.

## 5.2. Gestion des stocks et des commandes

L'IA est un allié fidèle dans la gestion des stocks. Elle peut analyser et prévenir les ruptures de stocks en produits ou les goulots d'étranglement dans les chaînes d'approvisionnements. Des actions correctrices peuvent dès lors être mises en place : commande de plus de produits, augmentation de la production, déviation de certaines commandes, etc... Le tout résumé sur une interface simple d'utilisation. Quelques exemples sont présentés par la suite :

- Easy WMS

L'IA permet d'assouvir le besoin qu'a le client de tout savoir et tout maîtriser dans leur chaîne d'approvisionnement. Les processus logistiques se sont perfectionnés, et des logiciels de gestions d'entrepôt (*Warehouse management*), tels qu'*Easy WMS* (édité par la société française Mecalux), peuvent maintenant optimiser les opérations du quotidien :

- Suivre l'inventaire des stocks en temps réel (quantité, péremption...);
- Contrôler la préparation des commandes et prévenir les erreurs ;
- Optimiser l'utilisation des ressources humaines et matérielles ;



- Emettre et suivre les bons de commande ;
- Diminuer les délais entre la demande et la réception d'une commande.

Cette solution peut s'adapter aux modifications de l'environnement (augmentation des ventes, variation de la demande...), en fluidifiant les flux de produits et d'informations grâce à une intégration possible au sein des ERP (*Enterprise Resource Planning*) des entreprises. Easy WMS annonce réduire de 30% les activités de manutention, de 99% les erreurs de préparation de commande, et d'augmenter de plus de 40% les capacités de stockage (180).

- Aera Technology & Tracelink au service des flux logistiques d'une Big Pharma

Le géant pharmaceutique Merck a notamment utilisé deux solutions pour répondre à la complexité grandissante des flux de données et corriger un grand nombre de problèmes liés à sa Supply Chain, de plus en plus complexe : Aera Technology et Tracelink. Ces modèles de machine learning analysent un grand nombre d'informations afin de prévenir les incidents et les erreurs par la mise en place d'actions correctives au plus vite.

Aera Technology est un logiciel américain conçu pour l'industrie pharmaceutique capable de passer en revue la quarantaine d'ERP et CRM (*customer relationship management*) de Merck, plusieurs fois par jour, afin d'en recueillir des informations de divers types : gestion des stocks, commandes en cours, approvisionnement chez les fournisseurs, etc...

La technologie d'Aera se supplante aux systèmes existants chez Merck, et devient un “*cerveau numérique intelligent*” au sein de l'entreprise, pour en comprendre les opérations commerciales ou les flux (*goulots d'étranglement ou pénuries*), faire des recommandations en temps réel (augmenter les commandes fournisseurs, réduire la fabrication, réacheminer ses produits, etc...), prédire les résultats commerciaux, et même réaliser des actions de manière autonome. Cela permet de décharger les employés de certaines tâches répétitives, afin de se focaliser sur des tâches plus créatives et stratégiques. Le responsable de cette entreprise, M. Laluyaux évoque la possibilité pour les entreprises de fonctionner de façon “autonome” grâce à cette technologie.

Le système apprend la manière avec laquelle les humains prennent des décisions au sein de l'entreprise, à partir des grands jeux de données qui sont mis à sa disposition, afin d'aider dans la prise de décision future (*recommandations, prédictions, etc...*). Comme très souvent avec de telles technologies, les informations de sortie sont réintégrées dans le logiciel, permettant à Aera d'apprendre et de s'améliorer continuellement, en étant de plus en plus réaliste, efficace et décisif dans la stratégie de l'entreprise.

Ce logiciel permet d'agréger un grand nombre de données de diverses provenances (ERP, CRM, tendances du marché, réseaux sociaux...), afin de prévoir et suivre les variations de la demande en produit.

*Selon le directeur des Systèmes Informatiques de Merck, Alessandro De Luca, plus d'un milliard de lignes de données sont analysées tous les jours par près de 4000 employés grâce à cette technologie (181).*

Il prédit les volumes de vente et la demande pour un produit, les ruptures ou le surstockage de médicaments, les profils de clientèle en temps réel, et propose des recommandations pour la réponse à cette demande (planification de la production, ajustements de l'inventaire, stocks de sécurité nécessaires) aux décisionnaires de l'entreprise, qui peuvent les accepter ou les rejeter. Aera exécutera les actions en fonction de la réponse qui lui sera donnée.

Ces mêmes technologies permettront demain de comprendre les causes des sous-performances des entreprises afin de les corriger et d'éviter qu'elles ne se reproduisent. Ainsi, cette technologie promet de répondre aux besoins des clients et des entreprises en temps et en heure, d'améliorer les performances opérationnelles et financières et de prédire les opportunités et les risques commerciaux.

Merck n'est pas le seul Big Pharma à l'avoir adopté : Johnson & Johnson, Reckitt Benckiser ou Unilever ont également fait ce choix.

Tracelink (182) est un logiciel de gestion des stocks "intelligent" mettant l'accent sur la sécurisation des chaînes d'approvisionnement et donc la sérialisation. Il permet notamment :

- De visualiser en temps réel l'état des stocks (stocks de matières, de produits en cours et de produits finis) ;
- D'automatiser les processus logistiques : réception, retours, préparation, envoi des commandes ;
- De répondre aux exigences de la Directive sur les médicaments falsifiés sur la traçabilité des produits reçus et envoyés ;
- D'analyser et d'identifier des chaînes d'approvisionnement à risque (médicaments contrefaits, etc...) ;
- D'échanger l'information avec un grand nombre d'interlocuteurs (officines, hôpitaux, industriels...) sur toute la Supply Chain très rapidement ;
- De vérifier les numéros de série des produits suspects.

### 5.3. Optimisation des flux (livraisons, flux internes)

#### 5.3.1. Automatisation du transfert des produits dans l'entrepôt

La collaboration entre les logiciels de gestion et les robots permet de réaliser le transport de produits de façon entièrement autonome. Cela a de nombreux avantages : la suppression de tâches de manutention très souvent pénibles pour les employés, une meilleure maîtrise des stocks et des localisations des ressources, ou une adaptation des flux.

#### 5.3.2. Sélection des itinéraires et optimisation des déplacements

L'IA permet également d'optimiser la coordination du transport de produits ou de matières :

- D'une part les **transferts intra logistiques** (au sein des entreprises) : des logiciels peuvent effectuer une cartographie des installations d'un entrepôt ou d'une usine complète et des flux de personnes, de matières ou d'équipements qui s'y produisent. Ils peuvent ensuite optimiser les transports en utilisant les trajets les plus courts, en

assignant les ressources nécessaires à ces transferts, pour qu'ils se fassent le plus rapidement et efficacement possible (183).

- D'autre part les transferts **extra logistiques** (livraisons de commandes) : l'IA peut analyser les informations de trafic pour proposer les itinéraires de livraisons les plus rapides et sûrs, en les modifiant en temps réel en cas d'incidents (accident de la route, conditions climatiques extrêmes, etc...).

Tracelink permet également d'assurer le transport de marchandises pharmaceutiques en toute sécurité, par le suivi des transferts de produits et l'identification d'anomalies. Enfin, ce sont également les **rappels de produits** défectueux ou suspects qui peuvent être facilités grâce à l'IA de Tracelink (184). Ces mêmes rappels sont la cause de nombreux problèmes, à la fois pour la santé des patients (décès dus à des effets indésirables), mais aussi pour les industriels, avec des coûts logistiques, judiciaires très importants (plusieurs milliards de dollars annuels). Tracelink prétend pouvoir accélérer le processus de rappels des produits, en envoyant des notifications digitales instantanées aux pharmacies et hôpitaux concernés, leur permettant d'identifier les lots incriminés.

Selon IDC, un groupe d'études sur les marchés des technologies de l'information, ce serait la moitié des chaînes d'approvisionnement qui ont intégré des technologies d'IA en leur sein, avec des gains de productivité considérables estimés à 15% (181). Les bénéfices s'étendent à l'accroissement des capacités humaines, la sécurisation de la Supply Chain, la facilitation du travail pour les employés, etc... C'est donc un axe important de développement pour ces nouvelles technologies.

## 6. Autres applications de l'Intelligence Artificielle

### 6.1. Pharmacovigilance

La pharmacovigilance (PV) permet de « *procéder au recueil et à l'évaluation scientifique des effets indésirables dans un but de prévention et de réduction des risques* » (185). C'est un organisme de surveillance et d'investigation de l'apparition d'effets indésirables (EI) inconnus jusqu'alors, le tout correspondant aux études de phase IV du produit de santé. La tâche actuelle est majoritairement "manuelle" : des signalements d'EI peuvent être transmis par des professionnels de santé ou des patients directement, puis c'est l'organisme de PV qui va se charger d'enquêter (origines de l'EI, moyens de prévention, etc...).

L'usage d'outils d'IA pourra réduire la quantité de travail manuel lié à la transcription de données et laisser plus de place à l'évaluation scientifique et médicale des EI, qui bénéficiera au patient, dont la santé sera améliorée.

- Améliorer la détection et le suivi des EI

Le volume de données relatives aux effets indésirables est en constante augmentation depuis le siècle dernier, bien que bon nombre de ceux-ci soient encore sous-signalés. L'IA pourra permettre d'explorer de larges sources de données (*dont les réseaux sociaux liés à la santé, les forums ou parmi les 500 bases de données médico-économiques, cohortes, registres*

*et études en cours disponibles)* afin d'en extraire de nouvelles informations. Une source plus prometteuse serait l'étude des dossiers médicaux partagés, qui représentent l'état réel du patient, l'historique de ses traitements et conclusions d'entrevues avec des professionnels de santé. La détection d'effets indésirables pourrait être améliorée, en extrayant du texte des informations additionnelles : symptômes, pathologies, facteurs d'aggravation, etc... Selon des spécialistes de l'INSERM, *"l'analyse des données issues de cohortes ou des bases médico-économiques sur le long terme peut permettre d'observer beaucoup de phénomènes, et notamment de faire des rapprochements entre des traitements et la survenue d'événements en santé. Cette pratique permet de repérer des événements indésirables graves et d'alerter sur certains risques"* (186).

- Prévenir et limiter la diffusion d'épidémies

Dans la continuité, suivre l'état de santé d'une population d'une région ou d'un pays donné permet d'observer et de suivre une modification de l'incidence d'une pathologie en son sein, afin de prévenir les autorités de santé, pour qu'elles prennent des mesures correctrices (186). La finalité est de restreindre la propagation des épidémies, et de mieux en suivre la transmission. Des épidémiologistes américains ont créé une plateforme, *HealthMap*, qui agrège des données régulièrement actualisées de diverses provenances (organismes sanitaires et publics, rapports officiels ou d'Internet, etc...) afin de prédire l'apparition d'épidémies et leur probable dissémination.

Quoiqu'il en soit, ces outils risquent de bouleverser les méthodes de déclaration, d'investigation et de suivi des effets indésirables.

## 6.2. Biotechnologies

La *Big Pharma* tricolore Sanofi expérimente une fonctionnalité de l'IA appliquée à la culture cellulaire de son site de Framingham. En utilisant des bioréacteurs jusqu'à huit fois plus petits qu'à l'accoutumée, en extrayant les protéines synthétisées en continu au cours du processus et en alimentant le milieu en nutriments dans le même temps, le rendement pourrait être fortement accru. La productivité de l'usine pourrait ainsi être 80 fois plus importante que tout autre unité de production actuelle. Les bénéfices ne s'arrêteraient pas là : 80% de besoin en énergie et de rejet de dioxyde de carbone en moins, 91% d'eau et 94% de produits chimiques économisés pour le nettoyage des cuves (par l'utilisation de dispositifs à usage unique), etc...(176)

Une autre application possible dans le secteur biotechnologique est le développement de **modèles prédictifs** permettant de simuler le comportement de procédés de fermentation dans des milieux de culture. Ces modèles seront basés sur les données des lots précédents et constamment actualisés pour être plus réalistes et efficaces. Une entreprise a testé cette solution, la plateforme *Canvass*, sur leur production de Lysine (187). Elle a été en mesure de leur prédire la quantité d'acide aminé produite dans les délais choisis. Cette plateforme permet donc d'estimer le moment exact où la fermentation est terminée. Les opérateurs peuvent donc arrêter le processus dès ce moment-là, plutôt que de le laisser fonctionner 4 à 6 heures de plus, pour être sûr de ne pas l'interrompre avant son terme.

En plus d'une augmentation de la productivité estimée à 5% (et donc de la marge et des profits), les utilisateurs attendent bénéficier d'informations plus représentatives de la réalité, pour pouvoir reproduire les comportements ou les lots corrects (187). La même quantité de produit pourra être synthétisée avec moins de lots, en réduisant donc les délais et coûts de production.

### 6.3. Marketing

Le secteur de la vente et donc du marketing bénéficie déjà en partie des outils de l'IA : en utilisant les cookies d'internet, il est possible de connaître le profil de chaque utilisateur, afin de faire du **marketing ciblé** : c'est à-dire adapter les publicités selon les préférences du public, et ainsi maximiser les chances de vendre le produit.

Toujours à partir d'un grand nombre de données, il sera possible d'étudier le parcours du consommateur, et d'analyser ce qui le fait acquérir ce produit plutôt qu'un autre : cette application sera d'un intérêt tout particulier pour le secteur de la cosmétique ou de la vente de produits disponibles sans ordonnance. Les campagnes publicitaires précédentes pourront être scrutées, afin de comprendre leur réussite ou leur échec, et ainsi prédire la réussite (ou non) d'une campagne future.

# Conclusion

L'émergence de l'intelligence artificielle est assez ancienne, mais son intégration au sein de l'industrie pharmaceutique n'est que très récente.

Même si certaines phases du cycle de vie du médicament profitent plus à l'heure actuelle des technologies de l'IA (recherche et développement et recherche clinique) que d'autres, c'est l'ensemble de ce secteur qui devrait en bénéficier dans les années à venir.

En effet, de plus en plus d'entreprises prennent le virage de l'IA et de la révolution technologique qui l'accompagne, et intègrent des solutions toujours plus innovantes pour optimiser leurs procédés. De lourds investissements et de grandes réorganisations sont opérés par les entreprises pharmaceutiques dans cette optique.

Nous pouvons finalement nous demander quels seront les **facteurs de réussite** de l'intégration de l'IA au sein de l'industrie pharmaceutique.

Le développement de l'industrie pharmaceutique passera obligatoirement par une démocratisation des usages de l'IA en son sein, et son intégration au plus vite dans les stratégies des entreprises. Les technologies de l'IA sont complexes et leur adaptation dans l'organisation des procédés passera par une phase de formation du personnel à ces nouveaux outils et d'adaptation des processus, puis de validation, afin de s'assurer de la bonne efficacité et de l'exactitude des résultats obtenus.

L'adaptation des normes, des référentiels et des procédures internes aux entreprises sera inévitable. L'établissement de règles claires quant à la gestion des données, pierre angulaire de toutes ces solutions d'IA, est primordial afin d'en encadrer le recueil et l'usage.

C'est ainsi un grand nombre de défis auxquels l'industrie pharmaceutique est confrontée du fait de l'essor de l'IA. Mais aucun ne paraît insurmontable et les bénéfices sur le plan économique, humains, environnementaux dépassent largement tous les sacrifices qui pourront être faits.

Mais la dernière question que l'on pourrait se poser serait celle-ci : jusqu'où pourra aller l'IA au sein de l'industrie pharmaceutique ?

Il n'y a pas de limites à ce que l'intelligence artificielle peut apporter à l'industrie pharmaceutique : raccourcir les coûts et les délais de R&D, faciliter, sécuriser et améliorer la recherche clinique, optimiser les procédés de fabrication et les chaînes d'approvisionnement, etc... Des solutions apparaissent régulièrement et proposent des réponses à des problèmes que l'Homme ne parvient pas toujours à résoudre efficacement. Mais il est primordial de **garder le contrôle** sur ces technologies, et ne pas trop leur donner de pouvoir. La prise de décision devra toujours rester entre les mains de personnes qualifiées (*libération de produits, etc...*). En effet, aucun ordinateur, aussi puissant soit-il, ne peut reproduire la pensée humaine et sa grande complexité.

*« Vous aurez toujours besoin d'une vérification humaine. [...] A un moment donné, la machine pourra tout préparer, réduire les incertitudes jusqu'à les supprimer, et c'est une amélioration impossible autrement, mais à un moment donné, la décision d'agir ou de ne pas agir doit être une décision humaine, car il faut que quelqu'un en prenne la responsabilité. »*

Emmanuel Macron (188)



# Bibliographie

1. Connexionnisme. In: Wikipédia [Internet]. 2020 [cité 20 nov 2020]. Disponible sur: <https://fr.wikipedia.org/w/index.php?title=Connexionnisme&oldid=176424729>
2. Monostori L. Cyber-physical Production Systems: Roots, Expectations and R&D Challenges. *Procedia CIRP*. 31 déc 2014;17:9–13.
3. industrie 4.0 usine intelligente [Internet]. AmeliorAction. [cité 21 oct 2020]. Disponible sur: <https://www.amelioraction.net/wp-content/uploads/Usine-du-Futur-Energiency-4001.jpg>
4. GeSI [Internet]. [cité 21 oct 2020]. Disponible sur: <https://www.gesi.org/research/smarter2030-ict-solutions-for-21st-century-challenges>
5. DICOM\_Année.G, DICOM\_Année.G. Notre ambition pour les industries de santé : 8ème Conseil stratégique des industries de santé (CSIS) [Internet]. Ministère des Solidarités et de la Santé. 2020 [cité 21 oct 2020]. Disponible sur: <https://solidarites-sante.gouv.fr/actualites/presse/dossiers-de-presse/article/notre-ambition-pour-les-industries-de-sante-8eme-conseil-strategique-des>
6. Grinaker S. AI – what artificial intelligence is and is not [Internet]. Digital Leaders. 2019 [cité 16 mai 2020]. Disponible sur: <https://digileaders.com/ai-what-artificial-intelligence-is-and-is-not/>
7. Demichelis R. Pourquoi l'intelligence artificielle n'est pas intelligente | Les Echos [Internet]. [cité 16 mai 2020]. Disponible sur: <https://www.lesechos.fr/tech-medias/intelligence-artificielle/pourquoi-lintelligence-artificielle-nest-pas-intelligente-961135>
8. Larousse É. Encyclopédie Larousse en ligne - intelligence artificielle [Internet]. [cité 16 mai 2020]. Disponible sur: [https://www.larousse.fr/encyclopedie/divers/intelligence\\_artificielle/187257](https://www.larousse.fr/encyclopedie/divers/intelligence_artificielle/187257)
9. L'intelligence artificielle [Internet]. [cité 16 mai 2020]. Disponible sur: <https://www.leem.org/lintelligence-artificielle>
10. Artificial intelligence [Internet]. ScienceDaily. [cité 19 oct 2020]. Disponible sur: [https://www.sciencedaily.com/terms/artificial\\_intelligence.htm](https://www.sciencedaily.com/terms/artificial_intelligence.htm)
11. Futura la rédaction de. Qui sont les pionniers de l'intelligence artificielle ? [Internet]. Futura. [cité 16 mai 2020]. Disponible sur: <https://www.futura-sciences.com/tech/questions-reponses/intelligence-artificielle-sont-pionniers-intelligence-artificielle-4907/>
12. Histoire de l'intelligence artificielle. In: Wikipédia [Internet]. 2020 [cité 16 mai 2020]. Disponible sur: [https://fr.wikipedia.org/w/index.php?title=Histoire\\_de\\_l%27intelligence\\_artificielle&oldid=170788809](https://fr.wikipedia.org/w/index.php?title=Histoire_de_l%27intelligence_artificielle&oldid=170788809)
13. 1475-1495 Automates de Léonard de Vinci [La robotique aux cycles 3 et 4] [Internet]. [cité 19 oct 2020]. Disponible sur: [https://atelier-canope-95.canoprof.fr/eleve/Automates%20et%20robots/res/robot.dossierHtml/co/RenaissanceVinci\\_1.html](https://atelier-canope-95.canoprof.fr/eleve/Automates%20et%20robots/res/robot.dossierHtml/co/RenaissanceVinci_1.html)
14. Linsky B. Principia Mathematica (Stanford Encyclopedia of Philosophy) [Internet]. [cité 16 mai 2020]. Disponible sur: <https://plato.stanford.edu/entries/principia-mathematica/>
15. Cryptanalyse d'Enigma. In: Wikipédia [Internet]. 2020 [cité 16 mai 2020]. Disponible sur: [https://fr.wikipedia.org/w/index.php?title=Cryptanalyse\\_d%27Enigma&oldid=170397209](https://fr.wikipedia.org/w/index.php?title=Cryptanalyse_d%27Enigma&oldid=170397209)
16. Principe de fonctionnement de la machine Enigma [Internet]. [cité 19 oct 2020]. Disponible sur: <http://mathweb.free.fr/crypto/debvingt/enigmafonc.php3>

17. Canteaut A. Turing à l'assaut d'Enigma - Interstices [Internet]. Interstices. 2012 [cité 16 mai 2020]. Disponible sur: <https://interstices.info/turing-a-lassaut-denigma/>
18. Raynaud M. Construire une machine de Turing. Le Monde.fr [Internet]. 22 juin 2012 [cité 16 mai 2020]; Disponible sur: [https://www.lemonde.fr/sciences/article/2012/06/22/l-ordinateur-en-lego-inspire-par-alan-turing\\_1723058\\_1650684.html](https://www.lemonde.fr/sciences/article/2012/06/22/l-ordinateur-en-lego-inspire-par-alan-turing_1723058_1650684.html)
19. Lettvin J, Maturana H, McCulloch W, Pitts W. What the Frog's Eye Tells the Frog's Brain. Proc IRE. nov 1959;47(11):1940-51.
20. McCarthy J, Minsky ML, Rochester N, Shannon CE. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. AI Mag. 1956;27(4):12-12.
21. Raji J. Loi de moore : modélisation des évolutions technologiques ? [Internet]. Digital Corner. [cité 16 mai 2020]. Disponible sur: <https://www.digitalcorner-wavestone.com/2015/06/la-loi-de-moore-une-modelisation-toujours-fidele-des-evolutions-technologiques/>
22. Cybernétique : Définition simple et facile du dictionnaire [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.linternaute.fr/dictionnaire/fr/definition/cybernetique/>
23. Simon HA. The Shape of Automation for Men and Management [Internet]. Harper&Row. New York, USA; 1966 [cité 16 mai 2020]. 111 p. Disponible sur: <https://journals.sagepub.com/doi/abs/10.1177/002218566800800216>
24. CC A: I-A-L. L'intelligence artificielle (IA) fête ses 50 ans [Internet]. Futura. [cité 19 oct 2020]. Disponible sur: <https://www.futura-sciences.com/tech/actualites/informatique-intelligence-artificielle-ia-fete-50-ans-9276/>
25. Basa G. Intelligent Agents [Internet]. Medium. 2018 [cité 16 mai 2020]. Disponible sur: <https://gungorbasa.com/intelligent-agents-dc5901daba7d>
26. Pizziniaco L. Quarantine, Intelligent Agents and Vacuum Cleaner [Internet]. Medium. 2020 [cité 20 nov 2020]. Disponible sur: <https://medium.com/@lucpiz/quarantine-intelligent-agents-and-vacuum-cleaner-28cf8ae1359e>
27. AlphaGo. In: Wikipédia [Internet]. 2020 [cité 19 oct 2020]. Disponible sur: <https://fr.wikipedia.org/w/index.php?title=AlphaGo&oldid=173294011>
28. Neurosciences computationnelles. In: Wikipédia [Internet]. 2019 [cité 16 mai 2020]. Disponible sur: [https://fr.wikipedia.org/w/index.php?title=Neurosciences\\_computationnelles&oldid=165020851](https://fr.wikipedia.org/w/index.php?title=Neurosciences_computationnelles&oldid=165020851)
29. Dumont d'Ayot G, Barillet A. L'intelligence artificielle forte [Internet]. Inprincipio. [cité 16 mai 2020]. Disponible sur: <https://www.inprincipio.xyz/ia-forte/>
30. IA super intelligente : un risque faible mais « cataclysmique » [Internet]. [cité 19 oct 2020]. Disponible sur: <https://usbeketrica.com/fr/article/ia-super-intelligente-un-risque-faible-mais-cataclysmique>
31. Larousse É. Définitions : algorithme - Dictionnaire de français Larousse [Internet]. [cité 16 mai 2020]. Disponible sur: <https://www.larousse.fr/dictionnaires/francais/algorithme/2238>
32. Réseaux de neurones artificiels - Théorie [Internet]. intelligenceartificielle.com. [cité 19 oct 2020]. Disponible sur: <http://intelligence-artificielle-robotique.weebly.com/theacuterie.html>
33. Figure 1. Le réseau de neurones artificiels simplifié b avec k 1, , 4 k... [Internet]. ResearchGate. [cité 20 nov 2020]. Disponible sur: [https://www.researchgate.net/figure/Le-reseau-de-neurones-artificiels-simplifie-b-avec-k-1-4-k-Donnees-dentree\\_fig1\\_329528185](https://www.researchgate.net/figure/Le-reseau-de-neurones-artificiels-simplifie-b-avec-k-1-4-k-Donnees-dentree_fig1_329528185)
34. Théorie - intelligenceartificielle.com [Internet]. [cité 16 mai 2020]. Disponible sur: <https://intelligence-artificielle-robotique.weebly.com/theacuterie.html>

35. Les réseaux de neurones | Intelligence artificielle 41 [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.youtube.com/watch?v=8qL2ISQd9L8&t=583s>
36. Shaw GL. Donald Hebb: The Organization of Behavior. In: Palm G, Aertsen A, éditeurs. Brain Theory [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 1986 [cité 19 oct 2020]. p. 231-3. Disponible sur: [http://link.springer.com/10.1007/978-3-642-70911-1\\_15](http://link.springer.com/10.1007/978-3-642-70911-1_15)
37. Clergue M. Réseaux de Neurones Artificiels [Internet]. SlideServe. 2011 [cité 16 mai 2020]. Disponible sur: <https://www.slideserve.com/liam/r-seaux-de-neurones-artificiels>
38. Davies A. What is Perceptron | Simplilearn [Internet]. Simplilearn.com. [cité 16 mai 2020]. Disponible sur: <https://www.simplilearn.com/what-is-perceptron-tutorial>
39. Réseau de neurones artificiels. In: Wikipédia [Internet]. 2020 [cité 16 mai 2020]. Disponible sur: [https://fr.wikipedia.org/w/index.php?title=R%C3%A9seau\\_de\\_neurones\\_artificiels&oldid=170825918](https://fr.wikipedia.org/w/index.php?title=R%C3%A9seau_de_neurones_artificiels&oldid=170825918)
40. L B. Perceptron – Tout savoir sur le plus vieil algorithme de Machine Learning [Internet]. LeBigData.fr. 2019 [cité 16 mai 2020]. Disponible sur: <https://www.lebigdata.fr/perceptron-machine-learning>
41. Perceptrons. An Introduction to Computational Geometry. Marvin Minsky and Seymour Papert. M.I.T. Press, Cambridge, Mass., 1969. vi + 258 pp., illus. Cloth, \$12; paper, \$4.95 | Science [Internet]. [cité 19 oct 2020]. Disponible sur: <https://science.sciencemag.org/content/165/3895/780>
42. Blanc-Durand P. Figure 1.7-Architecture classique d'un réseau de neurones convolutif... [Internet]. ResearchGate. [cité 16 mai 2020]. Disponible sur: [https://www.researchgate.net/figure/Architecture-classique-dun-reseau-de-neurones-convolutif-Une-image-est-fournie-en\\_fig5\\_330995099](https://www.researchgate.net/figure/Architecture-classique-dun-reseau-de-neurones-convolutif-Une-image-est-fournie-en_fig5_330995099)
43. administrateur. « Pas d'Intelligence sans Apprentissage » - L'Actualité Nouvelle-Aquitaine — science et culture, innovation [Internet]. L'Actualité Nouvelle-Aquitaine — science et culture, innovation. 2020 [cité 16 mai 2020]. Disponible sur: <https://actualite.nouvelle-aquitaine.science/pas-dintelligence-sans-apprentissage/>
44. A.I. technical - Machine Learning vs. Deep Learning [Internet]. lawtomed. 2019 [cité 16 mai 2020]. Disponible sur: <https://lawtomed.com/a-i-technical-machine-vs-deep-learning/>
45. Why Unsupervised Machine Learning is the Future of Cybersecurity [Internet]. TechNative. 2020 [cité 16 mai 2020]. Disponible sur: <https://www.technative.io/why-unsupervised-machine-learning-is-the-future-of-cybersecurity/>
46. Apprentissage par renforcement. In: Wikipédia [Internet]. 2020 [cité 16 mai 2020]. Disponible sur: [https://fr.wikipedia.org/w/index.php?title=Apprentissage\\_par\\_renforcement&oldid=170893241](https://fr.wikipedia.org/w/index.php?title=Apprentissage_par_renforcement&oldid=170893241)
47. Deluzarche C. Deep Learning [Internet]. Futura. [cité 16 mai 2020]. Disponible sur: <https://www.futura-sciences.com/tech/definitions/intelligence-artificielle-deep-learning-17262/>
48. Brand J. Deep Learning for Image Recognition [Internet]. Medium. 2019 [cité 20 nov 2020]. Disponible sur: <https://medium.com/kwadigoai/deep-learning-for-image-recognition-1d612be00bbb>
49. ImageNet [Internet]. [cité 19 oct 2020]. Disponible sur: <http://www.image-net.org/>
50. Louapre D. Le deep learning — Science étonnante #27 [Internet]. [cité 16 mai 2020]. Disponible sur: <https://www.youtube.com/watch?v=trWrEWfhTVg>
51. Turing AM. I.—COMPUTING MACHINERY AND INTELLIGENCE. Mind. 1 oct 1950;LIX(236):433-60.
52. Turing test. In: Wikipedia [Internet]. 2020 [cité 20 nov 2020]. Disponible sur: [https://en.wikipedia.org/w/index.php?title=Turing\\_test&oldid=986039138](https://en.wikipedia.org/w/index.php?title=Turing_test&oldid=986039138)

53. Schmitt A. Un ordinateur passe le test de Turing : pourquoi c'est important [Internet]. L'Obs. [cité 16 mai 2020]. Disponible sur: <https://www.nouvelobs.com/les-internets/20140610.OBS0018/un-ordinateur-passe-le-test-de-turing-pourquoi-c-est-important.html>
54. Journal officiel de la République Française. Version électronique authentifiée publiée au JO n° 0193 du 22/08/2014 | Legifrance [Internet]. CTNX1419323X août 22, 2014 p. 1. Disponible sur: [https://www.legifrance.gouv.fr/jo\\_pdf.do?id=JORFTEXT000029388087](https://www.legifrance.gouv.fr/jo_pdf.do?id=JORFTEXT000029388087)
55. les 3 V du Big Data - Volume, Vitesse et Variété [Internet]. Le DataScientist. 2019 [cité 19 oct 2020]. Disponible sur: <https://le-datascientist.fr/les-3-v-du-big-data>
56. Oracle cloud. Quel est le lien entre Big Data et l'IA ? | Oracle France Blog [Internet]. [cité 16 mai 2020]. Disponible sur: <https://blogs.oracle.com/oracle-france/quel-est-le-lien-entre-big-data-et-ia>
57. DataAge 2025 - La numérisation du monde | Seagate France [Internet]. Seagate.com. [cité 19 oct 2020]. Disponible sur: <https://www.seagate.com/fr/fr/our-story/data-age-2025/>
58. Un Centaure à l'école: Comment le numérique change (et doit encore changer) l'école, l'élève et le professeur [Internet]. 2016 [cité 19 oct 2020]. Disponible sur: <https://bookshop.org/books/un-centaure-a-l-ecole-comment-le-numerique-change-et-doit-encore-changer-l-ecole-l-eleve-et-le-professeur/9781534918276>
59. Le chiffre du jour. D'énormes quantités de données sur notre santé devraient être produites en 2020 [Internet]. Courrier international. 2020 [cité 16 mai 2020]. Disponible sur: <https://www.courrierinternational.com/article/le-chiffre-du-jour-denormes-quantites-de-donnees-sur-notre-sante-devraient-etre-produites-en>
60. Statista Research Department. Thème: Les objets connectés [Internet]. www.statista.com. [cité 16 mai 2020]. Disponible sur: <https://fr.statista.com/themes/2972/les-objets-connectes/>
61. Laurence. Il y aurait 22 milliards d'objets connectés dans le monde (Strategy Analytics) [Internet]. [cité 16 mai 2020]. Disponible sur: [https://www.mac4ever.com/actu/143604\\_il-y-aurait-22-milliards-d-objets-connectes-dans-le-monde-strategy-analytics](https://www.mac4ever.com/actu/143604_il-y-aurait-22-milliards-d-objets-connectes-dans-le-monde-strategy-analytics)
62. Mercer D. Consumer Electronics Reports [Internet]. [cité 16 mai 2020]. Disponible sur: <https://www.strategyanalytics.com/access-services/devices/connected-home/consumer-electronics/reports/report-detail/global-connected-and-iot-device-forecast-update>
63. Everything You Need to Know about 5G [Internet]. Speedtest Stories & Analysis: Data-driven articles on internet speeds. 2018 [cité 20 nov 2020]. Disponible sur: <https://www.speedtest.net/insights/blog/what-is-5g-guide/>
64. Leonard M. Declining price of IoT sensors means greater use in manufacturing [Internet]. Supply Chain Dive. [cité 16 mai 2020]. Disponible sur: <https://www.supplychaindive.com/news/declining-price-iot-sensors-manufacturing/564980/>
65. Rouse M. What is machine data? - Definition from WhatIs.com [Internet]. IoT Agenda. [cité 16 mai 2020]. Disponible sur: <https://internetofthingsagenda.techtarget.com/definition/machine-data>
66. DIRECTIVE (UE) 2016/ 680 DU PARLEMENT EUROPÉEN ET DU CONSEIL - du 27 avril 2016 - relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécution de sanctions pénales, et à la libre circulation de ces données, et abrogeant la décision-cadre 2008/ 977/ JAI du Conseil. p. 43.
67. Gremy N. Panorama des startups santé françaises utilisant l'IA - Bpifrance Le Hub [Internet]. BPI France. [cité 16 mai 2020]. Disponible sur: <https://lehub.bpifrance.fr/panorama-startups-sante-francaises-ia/>

68. Entreprendre dans la santé [Internet]. [cité 20 nov 2020]. Disponible sur: <https://bpifrance-creation.fr/entrepreneur/actualites/entreprendre-sante>
69. Le programme Territoires de soins numériques [Internet]. ARS Santé. [cité 16 mai 2020]. Disponible sur: <https://www.ars.sante.fr/index.php/le-programme-territoires-de-soins-numeriques>
70. DIABELOOP - Site officiel [Internet]. Diabeloop. [cité 16 mai 2020]. Disponible sur: <https://www.diabeloop.fr/entreprise>
71. Haute Autorité de Santé - DBLG1 SYSTEM [Internet]. [cité 16 mai 2020]. Disponible sur: [https://www.has-sante.fr/jcms/p\\_3150658/fr/dblg1-system](https://www.has-sante.fr/jcms/p_3150658/fr/dblg1-system)
72. Diabeloop. Diabeloop Launched a Large Clinical Study to Support its iController's FDA Clearance [Internet]. [cité 16 mai 2020]. Disponible sur: <https://www.prnewswire.com/news-releases/diabeloop-launched-a-large-clinical-study-to-support-its-icontrollers-fda-clearance-301006825.html>
73. ClinicalTrials.gov. Safety Assessment of DBLUS System in Adolescent and Adult Patients With Type 1 Diabetes and Assessment of Its Clinical Efficacy (DIABELOOP SP8) (DIABELOOP SP8) [Internet]. ClinicalTrials.gov. 2019 [cité 16 mai 2020]. Disponible sur: <https://clinicaltrials.gov/ct2/show/NCT04190277>
74. Eviter les accidents liés aux médicaments [Internet]. Santé Magazine. 2018 [cité 19 oct 2020]. Disponible sur: <https://www.santemagazine.fr/actualites/actualites-traitement/mesusage-du-medicament-des-consequences-loin-detre-anodines-306925>
75. Synapse Medicine | L'IA dédiée au bon usage du médicament [Internet]. [cité 16 mai 2020]. Disponible sur: <https://synapse-medicine.com/fr/>
76. Zirar W. Synapse Medicine lance une plateforme d'IA pour l'aide à la prescription [Internet]. [cité 16 mai 2020]. Disponible sur: <https://www.ticsante.com/story/4533/synapse-medicine-lance-une-plateforme-d-ia-pour-l-aide-a-la-prescription.html>
77. Graindorge T. Synapse Medicine, une IA d'aide à la prescription de médicaments – After the Web [Internet]. [cité 19 oct 2020]. Disponible sur: <http://aftertheweb.com/synapse-medicine-une-ia-daide-a-la-prescription-de-medicaments/>
78. Découvrez Synapse Medicine, la nouvelle plateforme d'IA au service du bon usage du médicament [Internet]. Revue Pharma. 2019 [cité 20 nov 2020]. Disponible sur: <https://www.revuepharma.fr/2019/04/decouvrez-synapse-medicine-la-nouvelle-plateforme-dia-au-service-du-bon-usage-du-medicament/>
79. Rapport de Cédric Villani : donner un sens à l'intelligence artificielle (IA) [Internet]. Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation. [cité 19 oct 2020]. Disponible sur: <http://www.enseignementsup-recherche.gouv.fr/cid128577/rapport-de-cedric-villani-donner-un-sens-a-l-intelligence-artificielle-ia.html>
80. Discours du Président de la République sur l'intelligence artificielle [Internet]. elysee.fr. [cité 19 oct 2020]. Disponible sur: <https://www.elysee.fr/emmanuel-macron/2018/03/29/discours-du-president-de-la-republique-sur-lintelligence-artificielle>
81. L'Intelligence artificielle au service du développement durable: défis et perspectives pour les programmes de science et d'ingénierie de l'UNESCO - UNESCO Bibliothèque Numérique [Internet]. [cité 19 oct 2020]. Disponible sur: [https://unesdoc.unesco.org/ark:/48223/pf0000368028\\_fre](https://unesdoc.unesco.org/ark:/48223/pf0000368028_fre)
82. Vladimir Poutine : « le leader en intelligence artificielle dominera le monde » | La Revue du Digital [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.larevuedudigital.com/vladimir-poutine-le-leader-en-intelligence-artificielle-dominera-le-monde/>

83. Textes adoptés - Une politique industrielle européenne globale sur l'intelligence artificielle et la robotique - Mardi 12 février 2019 [Internet]. [cité 19 oct 2020]. Disponible sur: [https://www.europarl.europa.eu/doceo/document/TA-8-2019-0081\\_FR.html](https://www.europarl.europa.eu/doceo/document/TA-8-2019-0081_FR.html)
84. Définition - Propriété intellectuelle | Insee [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.insee.fr/fr/metadonnees/definition/c1684>
85. Les critères de brevetabilité [Internet]. INPI.fr. 2015 [cité 20 nov 2020]. Disponible sur: <https://www.inpi.fr/fr/comprendre-la-propriete-intellectuelle/le-brevet/les-criteres-de-brevetabilite>
86. Comment peut-on protéger un logiciel ? [Internet]. [cité 20 nov 2020]. Disponible sur: [https://www.cncpi.fr/ba-ba/FAQ\\_questions\\_reponses/Comment\\_peut-on\\_proteger\\_un\\_logiciel/](https://www.cncpi.fr/ba-ba/FAQ_questions_reponses/Comment_peut-on_proteger_un_logiciel/)
87. United States Patent: 10566085 [Internet]. [cité 20 nov 2020]. Disponible sur: <http://patft1.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=%2Fnetacgi%2FPTO%2Fsrchnum.htm&r=1&f=G&l=50&s1=10566085.PN.&OS=PN/10566085&RS=PN/10566085>
88. Key legal considerations in accelerating pharma using artificial intelligence | Pharmafile [Internet]. [cité 19 oct 2020]. Disponible sur: <http://www.pharmafile.com/news/527999/key-legal-considerations-accelerating-pharma-using-artificial-intelligence>
89. affiches-parisiennes.com. Intelligence artificielle : les défis juridiques en matière de propriété intellectuelle [Internet]. Affiches Parisiennes. [cité 19 oct 2020]. Disponible sur: <https://www.affiches-parisiennes.com/intelligence-artificielle-les-defis-juridiques-en-matiere-de-propriete-intellectuelle-8769.html>
90. PricewaterhouseCoopers. Utiliser la puissance de l'IA [Internet]. PwC. [cité 20 nov 2020]. Disponible sur: <https://www.pwc.fr/fr/vos-enjeux/data-intelligence/intelligence-artificielle.html>
91. adeline. IA et RSE : un lien pas si artificiel [Internet]. Big Data Paris 2020. 2020 [cité 19 oct 2020]. Disponible sur: <https://www.bigdataparis.com/2020/ia-et-rse-un-lien-pas-si-artificiel/>
92. Larousse É. Définitions : éthique - Dictionnaire de français Larousse [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.larousse.fr/dictionnaires/francais/%C3%A9thique/31389>
93. Kitchin R. The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences. SAGE; 2014. 271 p.
94. Présentation du comité consultatif national d'éthique pour les sciences de la vie et de la santé | Comité Consultatif National d'Éthique [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.ccne-ethique.fr/fr/pages/presentation-du-comite-consultatif-national-dethique-pour-les-sciences-de-la-vie-et-de-la>
95. RGPD : de quoi parle-t-on ? | CNIL [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.cnil.fr/fr/rgpd-de-quoi-parle-t-on>
96. CHAPITRE I - Dispositions générales | CNIL [Internet]. [cité 20 nov 2020]. Disponible sur: <https://www.cnil.fr/fr/reglement-europeen-protection-donnees/chapitre1>
97. Règlement européen sur la protection des données : ce qui change pour les professionnels | CNIL [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.cnil.fr/fr/reglement-europeen-sur-la-protection-des-donnees-ce-qui-change-pour-les-professionnels>
98. Adopter les six bons réflexes | CNIL [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.cnil.fr/fr/adopter-les-six-bons-reflexes>
99. RGPD : vers une harmonisation mondiale ou pourquoi vous ne pourrez pas l'éviter - ZDNet [Internet]. [cité 20 nov 2020]. Disponible sur: <https://www.zdnet.fr/actualites-partenaires/rgpd-vers-une-harmonisation-mondiale-ou-pourquoi-vous-ne-pourrez-pas-l-eviter-39884553.htm>



100. Cybercriminalité : Les hackers pourraient avoir recours à l'intelligence artificielle - Linfodrome [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.linfodrome.com/high-tech/57203-cybercriminalite-les-hackers-pourraient-avoir-recours-a-l-intelligence-artificielle>
101. Comment la pharmacie s'approprie le Big Data et l'intelligence artificielle | Les Echos [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.lesechos.fr/industrie-services/pharmacie-sante/comment-la-pharmacie-sapproprie-le-big-data-et-lintelligence-artificielle-1003161>
102. PhD AZ. Deep Dive Into Big Pharma AI Productivity: One Study Shaking The Pharmaceutical Industry [Internet]. Forbes. [cité 20 nov 2020]. Disponible sur: <https://www.forbes.com/sites/alexzhavoronkov/2020/07/15/deep-dive-into-big-pharma-ai-productivity-one-study-shaking-the-pharmaceutical-industry/>
103. Davenport T, Ronanki R. Artificial Intelligence for the Real World. Harvard Business Review [Internet]. 1 janv 2018 [cité 20 nov 2020]; Disponible sur: <https://hbr.org/2018/01/artificial-intelligence-for-the-real-world>
104. Scrum [Internet]. ArchiMetric. 2019 [cité 20 nov 2020]. Disponible sur: <https://www.archimetric.com/scrum-%e4%b8%ba%e4%bb%80%e4%b9%88sprint%e9%95%bf%e5%ba%a6%e6%98%af%e5%9b%ba%e5%ae%9a%ef%bc%9f/>
105. Bonnes pratiques de fabrication de médicaments à usage humain - ANSM : Agence nationale de sécurité du médicament et des produits de santé [Internet]. [cité 19 oct 2020]. Disponible sur: [https://www.ansm.sante.fr/Activites/Elaboration-de-bonnes-pratiques/Bonnes-pratiques-de-fabrication-de-medicaments-a-usage-humain/\(offset\)/4](https://www.ansm.sante.fr/Activites/Elaboration-de-bonnes-pratiques/Bonnes-pratiques-de-fabrication-de-medicaments-a-usage-humain/(offset)/4)
106. AI In Healthcare Q3'19 Funding Breaks Records - CB Insights Research [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.cbinsights.com/research/ai-artificial-intelligence-healthcare-funding-q3-19/>
107. Gilchrist A. Industry 4.0 [Internet]. Berkeley, CA: Apress; 2016 [cité 19 oct 2020]. Disponible sur: <http://link.springer.com/10.1007/978-1-4842-2047-4>
108. AI Adoption in the Enterprise [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.oreilly.com/data/free/ai-adoption-in-the-enterprise.csp>
109. L'industrie pharmaceutique en France : emploi et localisation | Leem [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.leem.org/emploi-et-localisation>
110. Population active – Tableaux de l'économie française | Insee [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.insee.fr/fr/statistiques/3676623?sommaire=3696937>
111. Le droit du travail face aux enjeux de l'intelligence artificielle : Tome 2 d'une trilogie - Employment and HR - France [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.mondaq.com/france/employee-rights-labour-relations/747986/le-droit-du-travail-face-aux-enjeux-de-lintelligence-artificielle-tome-2-d39une-trilogie>
112. Jobs lost, jobs gained: What the future of work will mean for jobs, skills, and wages | McKinsey [Internet]. [cité 20 nov 2020]. Disponible sur: <https://www.mckinsey.com/featured-insights/future-of-work/jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages#>
113. Will robots really steal our jobs?: an international analysis of the potential long term impact of automation | VOCEDplus, the international tertiary education and research database [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.voced.edu.au/content/ngv%3A78814>
114. Sous-section 2 : Définition du motif économique. (Article L1233-3) - Légifrance [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.legifrance.gouv.fr/codes/id/LEGIARTI000036762081/2018-04-01/>

115. AI, automation, and the future of work: Ten things to solve for (Tech4Good) | McKinsey [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.mckinsey.com/featured-insights/future-of-work/ai-automation-and-the-future-of-work-ten-things-to-solve-for#part2>
116. IFTF: The Next Era of Human|Machine Partnerships [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.iftf.org/humanmachinepartnerships/>
117. Artificial intelligence: anticipating its impact on jobs to ensure a fair transition (own-initiative opinion) [Internet]. European Economic and Social Committee. 2018 [cité 19 oct 2020]. Disponible sur: <https://www.eesc.europa.eu/en/our-work/opinions-information-reports/opinions/artificial-intelligence-anticipating-its-impact-jobs-ensure-fair-transition-own-initiative-opinion>
118. Section 1 : Obligations de l'employeur et plan de formation. (Article L6321-1) - Légifrance [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.legifrance.gouv.fr/codes/id/LEGIARTI000037385741/2019-01-01/>
119. L'intelligence artificielle, une nouvelle donne pour le monde du travail [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.escadrille.org/fr/blog/intelligence-artificielle-monde-du-travail>
120. Intelligence artificielle et capital humain, quels défis pour les entreprises ? [Internet]. France. [cité 20 nov 2020]. Disponible sur: <https://harris-interactive.fr/press/intelligence-artificielle-et-capital-humain-quels-defis-pour-les-entreprises/>
121. Empowering the Firstline Workforce [Internet]. Forbes Insights. [cité 20 nov 2020]. Disponible sur: <http://www.forbes.com/forbes-insights/our-work/empowering-the-firstline-workforce/>
122. Avantages de l'intelligence artificielle pour les travailleurs [Internet]. Element AI. [cité 20 nov 2020]. Disponible sur: <https://www.elementai.com/fr/news/2019/avantages-de-l-intelligence-artificielle-pour-les-travailleurs>
123. CSA - Les Français et l'Intelligence Artificielle [Internet]. csa.eu. [cité 19 oct 2020]. Disponible sur: <https://www.csa.eu/fr/survey/les-français-et-lintelligence-artificielle>
124. Notes from the AI frontier: Applications and value of deep learning | McKinsey [Internet]. [cité 20 nov 2020]. Disponible sur: <https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-applications-and-value-of-deep-learning>
125. Intelligence artificielle : le Parlement européen adopte des recommandations à l'attention de la Commission [Internet]. De Gaulle Fleurance & Associés. 2020 [cité 20 nov 2020]. Disponible sur: <https://www.degaullefleurance.com/intelligence-artificielle-le-parlement-europeen-adopte-des-recommandations-a-lattention-de-la-commission/>
126. Lauga N. IA et éthique : Comment comprendre son modèle ? 🤖 [Internet]. Medium. 2019 [cité 20 nov 2020]. Disponible sur: <https://medium.com/@nathan.lauga/ia-et-%C3%A9thique-comment-comprendre-son-mod%C3%A8le-ec217c92616d>
127. Comment permettre à l'Homme de garder la main ? Rapport sur les enjeux éthiques des algorithmes et de l'intelligence artificielle | CNIL [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.cnil.fr/fr/comment-permettre-lhomme-de-garder-la-main-rapport-sur-les-enjeux-ethiques-des-algorithmes-et-de>
128. Intelligence artificielle | CNNum | Traducteur et éclairer des transformations numériques [Internet]. [cité 19 oct 2020]. Disponible sur: <https://cnnumerique.fr/nos-travaux/intelligence-artificielle>
129. Qu'est-ce que la responsabilité sociétale des entreprises (RSE) ? [Internet]. [cité 19 oct 2020]. Disponible sur: <https://www.economie.gouv.fr/entreprises/responsabilite-societale-entreprises-rse>
130. Laramée de Tannenberg V. L'intelligence artificielle au service de l'environnement? [Internet]. Journal de l'environnement. [cité 19 oct 2020]. Disponible sur:

<https://259710f66da04a8e8d9ce8db5c2253d5.pages.ubembed.com/b224290e-92cf-4817-8dd4-f9cf2cd8effe/a.html?closedAt=0>

131. PricewaterhouseCoopers. L'intelligence artificielle au service du développement durable [Internet]. PwC. [cité 19 oct 2020]. Disponible sur: <https://www.pwc.fr/fr/publications/data/batir-avenir-durable-grace-a-intelligence-artificielle.html>
132. Fabrication d'un médicament. Étape 10 : gestion du cycle de vie - EUPATI Toolbox [Internet]. [cité 21 nov 2020]. Disponible sur: <https://toolbox.eupati.eu/resources/fabrication-dun-medicament-etape-10-gestion-du-cycle-de-vie/?lang=fr&print=print>
133. Average cost of developing a new drug could be up to \$1.5 billion less than pharmaceutical industry claims [Internet]. LSHTM. [cité 21 nov 2020]. Disponible sur: <https://www.lshtm.ac.uk/newsevents/news/2020/average-cost-developing-new-drug-could-be-15-billion-less-pharmaceutical>
134. PhD AZ. AI Hype Or AI Hope: When Will AI Disrupt The Pharmaceutical Industry? [Internet]. Forbes. [cité 21 nov 2020]. Disponible sur: <https://www.forbes.com/sites/cognitiveworld/2020/04/30/ai-hype-or-ai-hope-when-will-ai-disrupt-the-pharmaceutical-industry/>
135. Nouguez É. La définition des médicaments génériques entre enjeux thérapeutiques et économiques. Rev Française Aff Soc. 2007;(3):99-121.
136. "Bigger than DNA" — how AI is transforming the pharma industry [Internet]. Sifted. [cité 21 nov 2020]. Disponible sur: <https://sifted.eu/articles/ai-transforming-pharma/>
137. Inc GVR. Artificial Intelligence In Drug Discovery Market Size Worth \$3.5 Billion By 2027: Grand View Research, Inc. [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.prnewswire.com/news-releases/artificial-intelligence-in-drug-discovery-market-size-worth-3-5-billion-by-2027-grand-view-research-inc-301129855.html>
138. BenevolentAI | What We Do [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.benevolent.com/what-we-do>
139. Duncan M. First Steps Toward Decentralized Biomedical AI [Internet]. Medium. 2019 [cité 21 nov 2020]. Disponible sur: <https://blog.singularitynet.io/first-steps-toward-decentralized-biomedical-ai-b763503fd874>
140. Antidot. Comment l'intelligence artificielle améliore la recherche documentaire [Internet]. Logiciels présenté à; [cité 21 nov 2020]. Disponible sur: <https://fr.slideshare.net/AntidotNet/comment-lintelligence-artificielle-amliore-la-recherche-documentaire>
141. Orgeolet L, Foulquier N, Misery L, Redou P, Pers J-O, Devauchelle-Pensec V, et al. Can artificial intelligence replace manual search for systematic literature? Review on cutaneous manifestations in primary Sjögren's syndrome. Rheumatol Oxf Engl. 01 2020;59(4):811-9.
142. Our science - Open Targets [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.opentargets.org/science>
143. A breakthrough in imaginative AI with experimental validation to accelerate drug discovery [Internet]. EurekAlert! [cité 21 nov 2020]. Disponible sur: [https://www.eurekalert.org/pub\\_releases/2019-09/dka-abi090319.php](https://www.eurekalert.org/pub_releases/2019-09/dka-abi090319.php)
144. Medicine by machine: Is A.I. the cure for the world's ailing drug industry? [Internet]. Fortune. [cité 21 nov 2020]. Disponible sur: <https://fortune.com/longform/ai-artificial-intelligence-medicine-healthcare-pharmaceutical-industry/>
145. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. juin 2019;18(6):463-77.

146. Zeng X, Zhu S, Lu W, Liu Z, Huang J, Zhou Y, et al. Target identification among known drugs by deep learning from heterogeneous networks. *Chem Sci*. 19 févr 2020;11(7):1775-97.
147. SIMON M. Structures des protéines [Internet]. Cours Pharmacie. 2011 [cité 21 nov 2020]. Disponible sur: <https://www.cours-pharmacie.com/biochimie/structures-des-proteines.html>
148. AlphaFold: Using AI for scientific discovery | DeepMind [Internet]. [cité 21 nov 2020]. Disponible sur: <https://deepmind.com/blog/article/AlphaFold-Using-AI-for-scientific-discovery>
149. Chan HCS, Shan H, Dahoun T, Vogel H, Yuan S. Advancing Drug Discovery via Artificial Intelligence. *Trends Pharmacol Sci*. 1 août 2019;40(8):592-604.
150. Cleves AE, Jain AN. Structure- and Ligand-Based Virtual Screening on DUD-E<sup>+</sup>: Performance Dependence on Approximations to the Binding Pocket. *J Chem Inf Model*. 28 sept 2020;60(9):4296-310.
151. How Artificial Intelligence is Transforming Drug Design [Internet]. Drug Discovery World (DDW). 2019 [cité 21 nov 2020]. Disponible sur: <https://www.ddw-online.com/how-artificial-intelligence-is-transforming-drug-design-1530-201910/>
152. L'IA au service de la découverte de nouveaux médicaments [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.centraliens-lyon.net/technica/article/l-ia-au-service-de-la-decouverte-de-nouveaux-medicaments/111>
153. Deep learning enables rapid identification of potent DDR1 kinase inhibitors | Nature Biotechnology [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.nature.com/articles/s41587-019-0224-x>
154. Our Technology – Atomwise [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.atomwise.com/our-technology/>
155. Segler MHS, Preuss M, Waller MP. Learning to Plan Chemical Syntheses. *Nature*. mars 2018;555(7698):604-10.
156. Planning chemical syntheses with deep neural networks and symbolic AI | Nature [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.nature.com/articles/nature25978?proof=t>
157. Cha Y, Erez T, Reynolds IJ, Kumar D, Ross J, Koytiger G, et al. Drug repurposing from the perspective of pharmaceutical companies. *Br J Pharmacol*. 2018;175(2):168-80.
158. Stokes JM, Yang K, Swanson K, Jin W, Cubillos-Ruiz A, Donghia NM, et al. A Deep Learning Approach to Antibiotic Discovery. *Cell*. 20 févr 2020;180(4):688-702.e13.
159. La recherche clinique | ICM [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.icm.unicancer.fr/fr/recherche/la-recherche-clinique>
160. Wong CH, Siah KW, Lo AW. Estimation of clinical trial success rates and related parameters. *Biostat Oxf Engl*. 01 2019;20(2):273-86.
161. AI in Healthcare: The Future of the Clinical Trial [Internet]. CB Insights Research. [cité 21 nov 2020]. Disponible sur: <https://www.cbinsights.com/research/briefing/ai-in-healthcare-future-clinical-trial/>
162. Recruit [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.mendel.ai/recruit>
163. Mendel.ai nabs \$2 million to match cancer patients with the latest clinical trials | TechCrunch [Internet]. [cité 21 nov 2020]. Disponible sur: <https://techcrunch.com/2017/07/01/mendel-ai-nabs-2-million-to-match-cancer-patients-with-the-latest-clinical-trials/?guccounter=1>
164. OMS | L'observance des traitements prescrits pour les maladies chroniques pose problème dans le monde entier [Internet]. WHO. World Health Organization; [cité 21 nov 2020]. Disponible sur: <https://www.who.int/mediacentre/news/releases/2003/pr54/fr/>

165. APMnews - Observance médicamenteuse: au moins 9 milliards d'euros d'économies potentielles par an dans six pathologies (IMS/Crip) [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.apmnews.com/freestory/10/254296/observance-medicamenteuse-au-moins-9-milliards-d-euros-d-economies-potentielles-par-an-dans-six-pathologies-%28ims-crip%29>
166. New Research Sheds Light on Intentional Non-adherence in Clinical Trials [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.prnewswire.com/news-releases/new-research-sheds-light-on-intentional-non-adherence-in-clinical-trials-301075265.html>
167. Assessing the Scope and Predictors of Intentional Dose Non-adherence in Clinical Trials | SpringerLink [Internet]. [cité 21 nov 2020]. Disponible sur: <https://link.springer.com/article/10.1007/s43441-020-00155-x>
168. AI advanced analytics and IoT sensor technologies [Internet]. Wonder. [cité 21 nov 2020]. Disponible sur: <https://askwonder.com/research/ai-advanced-analytics-iot-sensor-technologies-owuyno36n>
169. Lucas G, Gratch J, King A, Morency L-P. It's only a computer: Virtual humans increase willingness to disclose. *Comput Hum Behav.* 1 août 2014;37:94–100.
170. Insurer - Sensely [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.sensely.com/insurer/>
171. Pharma Industry in the Age of Artificial Intelligence: The Future is Bright - Healthcare Weekly [Internet]. [cité 21 nov 2020]. Disponible sur: <https://healthcareweekly.com/artificial-intelligence-in-pharmacology/>
172. La production pharmaceutique c'est quoi ? [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.leem.org/la-production-pharmaceutique-cest-quoi>
173. Digital manufacturing of biologics [Internet]. Cytiva. [cité 21 nov 2020]. Disponible sur: <https://www.cytivalifesciences.com/en/us/solutions/bioprocessing/knowledge-center/digital-manufacturing-of-biologics>
174. New Predix Offering from GE Digital Brings Manufacturing Data to the Cloud | GE News [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.ge.com/news/press-releases/new-predix-offering-ge-digital-brings-manufacturing-data-cloud>
175. Nouvelle L. L'Intelligence Artificielle : la solution à la productivité des entreprises pharmaceutiques - Augustin Marty. 10 déc 2018 [cité 21 nov 2020]; Disponible sur: <https://www.usinenouvelle.com/blogs/augustin-marty/l-intelligence-artificielle-la-solution-a-la-productivite-des-entreprises-pharmaceutiques.N780804>
176. Nouvelle L. Sanofi inaugure sa première usine digitale au monde, près de Boston - L'Usine Santé. 16 oct 2019 [cité 21 nov 2020]; Disponible sur: <https://www.usinenouvelle.com/editorial/au-coeur-de-la-premiere-usine-digitale-au-monde-de-sanofi-pres-de-boston.N894809>
177. Industrie pharmaceutique. Accélération sur le big data et l'IA [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.gref-bretagne.com/Actualites/Revue-de-presse/Industrie-pharmaceutique.-Acceleration-sur-le-big-data-et-l-IA>
178. Anonymous. ICH Q9 Quality risk management [Internet]. European Medicines Agency. 2018 [cité 21 nov 2020]. Disponible sur: <https://www.ema.europa.eu/en/ich-q9-quality-risk-management>
179. NF X50-600 - Mars 2006 [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.boutique.afnor.org/norme/nf-x50-600/management-logistique-demarche-logistique-et-gestion-de-la-chaine-logistique/article/702739/fa137573>
180. Mecalux. Easy Mecalux WMS | Logiciel de gestion d'entrepôt [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.mecalux.fr/logiciel/wms>
181. Merck injecte de l'IA dans sa supply chain [Internet]. Enjeux Logistiques. 2019 [cité 21 nov 2020]. Disponible sur: <https://www.enjeuxlogistiques.com/merck-injecte-de-lia-dans-sa-supply-chain/>

182. Smart Inventory Tracker | TraceLink [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.tracelink.com/solutions/smart-logistics-and-iot/smart-inventory-tracking>
183. Mecalux. L'application de l'IA en Supply Chain : automatisation et efficacité [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.mecalux.fr/blog/intelligence-artificielle-supply-chain>
184. Digital Recalls for Pharmaceutical and Medical Products | TraceLink [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.tracelink.com/solutions/networked-health/digital-recalls>
185. Qualité et sécurité des médicaments : les obligations - ANSM : Agence nationale de sécurité du médicament et des produits de santé [Internet]. [cité 21 nov 2020]. Disponible sur: [https://www.ansm.sante.fr/Dossiers/Medicaments-generiques/Qualite-et-securite-des-medicaments-les-obligations/\(offset\)/2](https://www.ansm.sante.fr/Dossiers/Medicaments-generiques/Qualite-et-securite-des-medicaments-les-obligations/(offset)/2)
186. Big data en santé [Internet]. Inserm - La science pour la santé. [cité 21 nov 2020]. Disponible sur: <https://www.inserm.fr/information-en-sante/dossiers-information/big-data-en-sante>
187. AI-enabled Predictive Analytics Optimizes Fermentation Production [Internet]. ARC Advisory Group. [cité 21 nov 2020]. Disponible sur: <https://www.arcweb.com/blog/ai-enabled-predictive-analytics-optimizes-fermentation-production>
188. Emmanuel Macron Q&A: France's President Discusses Artificial Intelligence Strategy | WIRED [Internet]. [cité 21 nov 2020]. Disponible sur: <https://www.wired.com/story/emmanuel-macron-talks-to-wired-about-frances-ai-strategy/>



**ENGAGEMENT DE NON PLAGIAT**

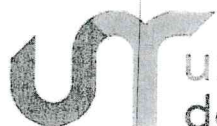
Je, soussign  (e) HOERDT Aur lien

D clare  tre pleinement conscient(e) que le plagiat de documents ou d'une partie d'un document publi s constitue une violation des droits d'auteur ainsi qu'une fraude caract ris e. (*D cret n 92-657 du 13 juillet 1992*)

En cons quence, je m'engage   citer toutes les sources que j'ai utilis es pour  crire ce m moire.

Signature :





université  
de TOURS

Faculté de pharmacie  
Philippe-Maupas

**SIGNATURES DU DIRECTEUR DE THESE ET DU DOYEN**

N° Étudiant : 21203416

N° Thèse : 74

Nom et Prénom : HOERDT Aurélien

Sujet : Intelligence Artificielle et Industrie Pharmaceutique :

De nouveaux défis pour de nouvelles innovations

Tours, le : 18/12/2020

Le(s) Directeur(s) de Thèse :

Emilie Viehron

Gérard PAPON

Vu et Transmis La Directrice de l'U.F.R.  
Le Doyen des Sciences Pharmaceutiques

Pr V. MAUPOIL-DAVID

## Intelligence Artificielle et Industrie Pharmaceutique - De nouveaux défis pour de nouvelles innovations

### RÉSUMÉ DE LA THÈSE

L'intelligence artificielle prend de plus en plus de place dans nos vies quotidiennes. Elle n'épargne pas le secteur pharmaceutique, qui commence à intégrer certaines de ces technologies afin d'optimiser ses processus.

Dans cette thèse, nous discutons des principes de base de l'intelligence artificielle, des défis que son utilisation dans le secteur de la santé implique, et enfin, d'exemples d'applications actuelles et en projet dans divers secteurs de l'industrie pharmaceutique (en recherche & développement, en recherche clinique, en production, assurance et contrôle qualité, ou encore en logistique).

### MOTS-CLÉS SIGNIFICATIFS DE SON CONTENU, ATTIBUÉS PAR LE CANDIDAT EN LIAISON AVEC LA BIBLIOTHÈQUE UNIVERSITAIRE ET LES MEMBRES DU JURY

Intelligence Artificielle, Big Data, Machine Learning, Deep Learning, Innovations

### JURY

PRÉSIDENT : M. Bruno GIRAUDEAU, enseignant de la Faculté de Pharmacie de Tours – HDR

#### MEMBRES :

- M. MASSOT Victor, Pharmacien Hospitalier - Tours
- M. MUREAU Charles, Pharmacien – Parçay-Meslay
- M. PAPON Géraud, Pharmacien – Parçay-Meslay
- Mme VIERRON Emilie, Maître de conférences - Tours

DATE ET LIEU DE SOUTENANCE

TOURS, le 16 / 12 / 2020