



INTRODUCTION À L'IA

RENDRE ACCESSIBLE L'IA EN PRATIQUE MÉDICALE

HCL
HOSPICES CIVILS
DE LYON

28/03/2023

ANTOINE RICHARD, CHARGÉ DES APPLICATIONS IA, HCL, DSN GIÈRES

www.chu-lyon.fr

INTRODUCTION

CONTEXTE ET DÉFINITIONS

QU'EST-CE QUE « L'INTELLIGENCE ARTIFICIELLE » ?

3

UN LABYRINTHE DE DÉFINITIONS ¹

« Ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine » ^{2 3}

- N'englobe pas tous les sous-domaines de l'IA (ex. intelligence collective, algorithmes bio-inspirés, ingénierie de la connaissance, etc.) ⁴
- L'intelligence humaine a des spécificités que l'on ne souhaite pas reproduire (ex. biais cognitifs) ^{5 6}

1. [Rehak, R. \(2021\) – « The Language Labyrinth: Constructive Critique on the Terminology Used in the AI Discourse »](#)
2. [Larousse – Intelligence Artificielle](#)
3. [Robert – Intelligence Artificielle](#)
4. [Gao and Ding \(2022\) – « The research landscape on the artificial intelligence: a bibliometric analysis of recent 20 years »](#)
5. [Haselton M., Nettle D. and Andrews P. W. \(2015\) – « The Evolution of Cognitive Bias »](#)
6. [O'Sullivan E. and Schofield S. \(2018\) – « Cognitive Bias in Clinical Medicine »](#)

QU'EST-CE « L'INTELLIGENCE » ?

UN PROBLÈME DE DÉFINITION

En français ¹:

- Faculté de connaître, de comprendre; qualité de l'esprit qui comprend et s'adapte facilement
- L'ensemble des fonctions mentales ayant pour objet la connaissance rationnelle
- L'intelligence de qqch.: acte ou capacité de comprendre (qqch.)

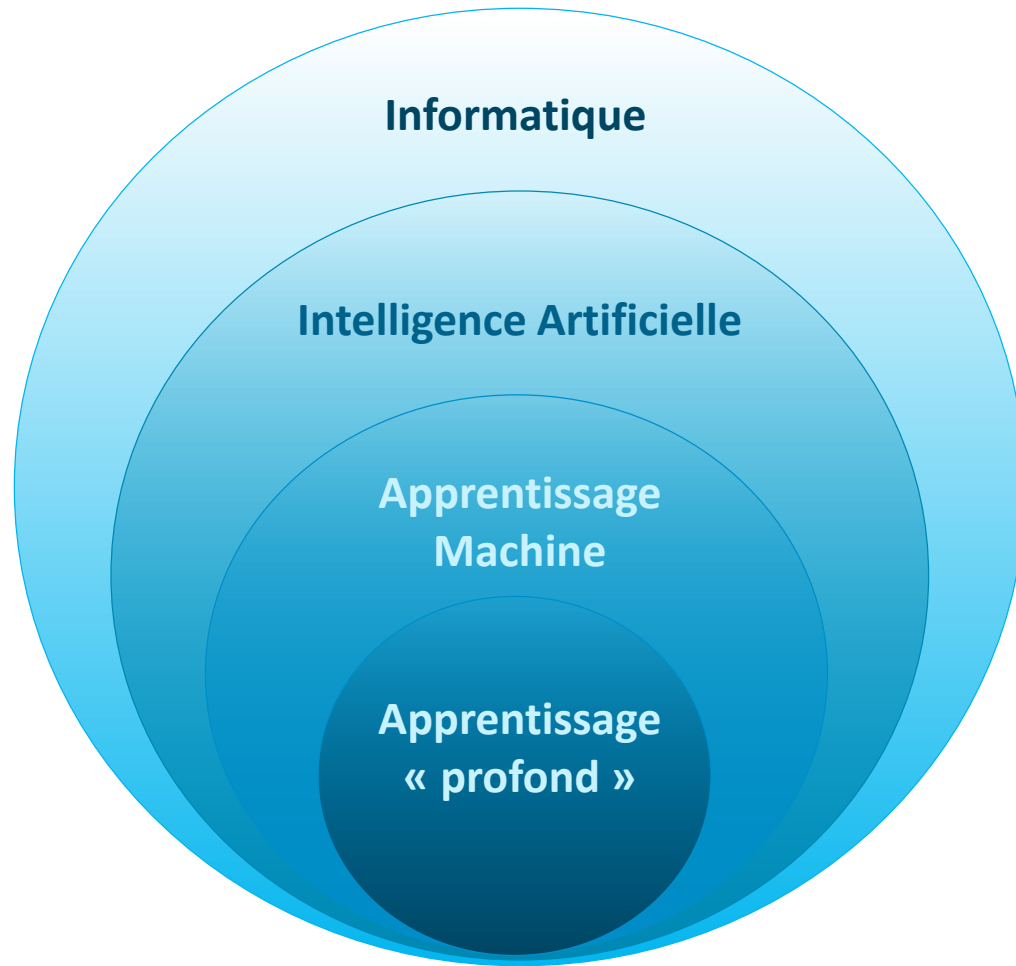
En anglais ²:

- The ability to understand and learn well, and to form judgments and opinions based on reason.
 <=> "Intelligence" en français
- A government department or other group that gathers information about other countries or enemies, or the information that is gathered.
 <=> "Renseignement" en français

1. <https://dictionnaire.lerobert.com/definition/intelligence>

2. <https://dictionary.cambridge.org/dictionary/english/intelligence>

INFORMATIQUE, IA, MACHINE LEARNING, AND DEEP LEARNING



Algorithmes:

Méthodes mathématiques pour formaliser un processus, utilisés en informatique pour créer des logiciels

Intelligence Artificielle (IA):

Sous-domaine de l'informatique dédié à l'élaboration de compétences « intelligentes » dans des logiciels¹

Machine Learning (ML):

Sous-domaine de l'IA consacré à la création d'algorithmes capable « d'apprendre » comment effectuer une tâche à partir de données²

Deep Learning (DL):

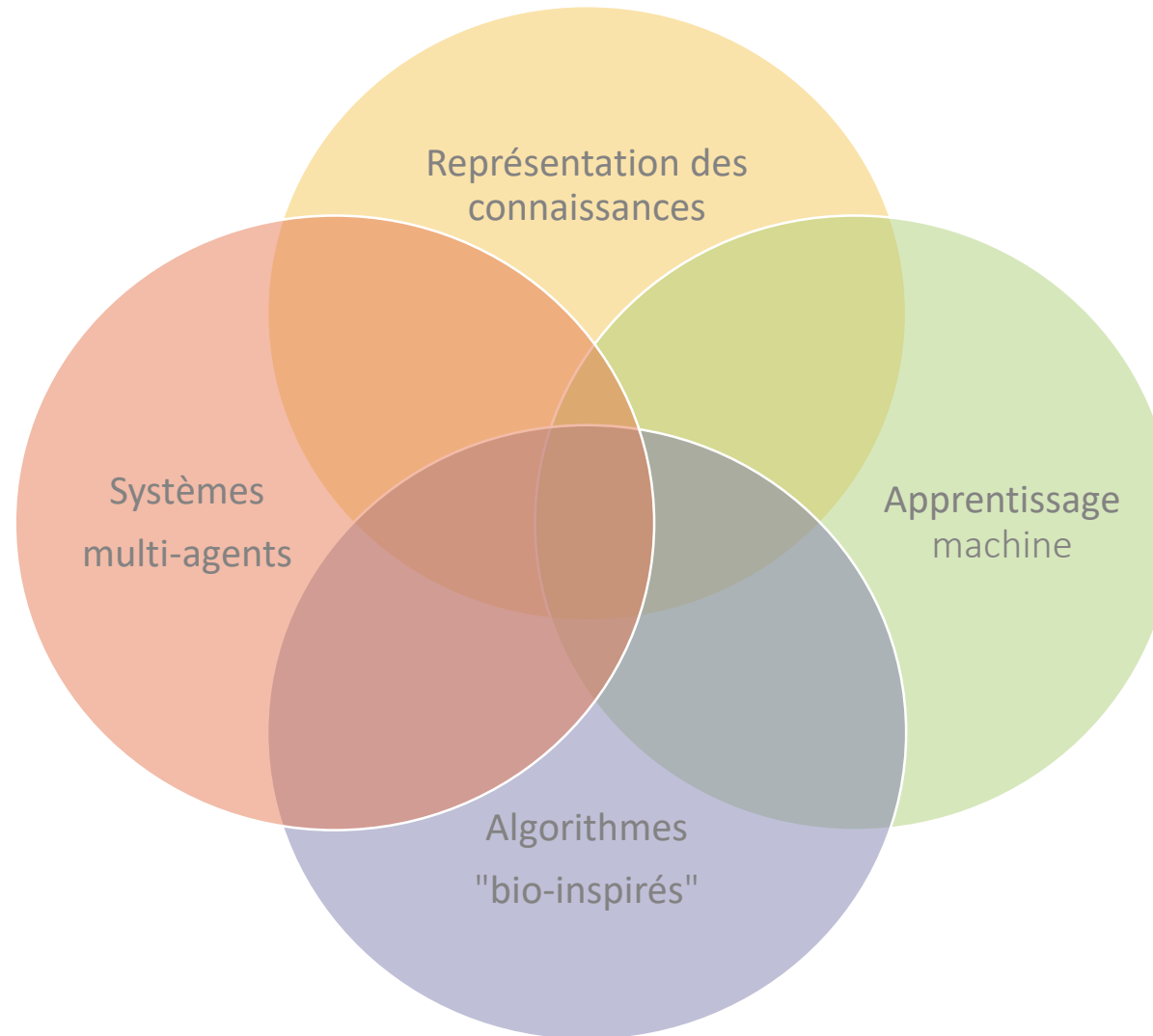
Sous-domaine du ML consacré à l'étude et aux développements de réseaux de neurones artificiels multi-couches³

1. [Gao and Ding \(2022\) - The research landscape on the artificial intelligence: a bibliometric analysis of recent 20 years](#)
2. [Mahesh \(2020\) – Machine Learning Algorithms: a review](#)
3. [Dong, Wang and Abbas \(2021\) – A Survey on Deep Learning and its Applications](#)

QU'EST-CE QUE « L'INTELLIGENCE ARTIFICIELLE » ?

6

UN VASTE DOMAINE DE RECHERCHE ¹

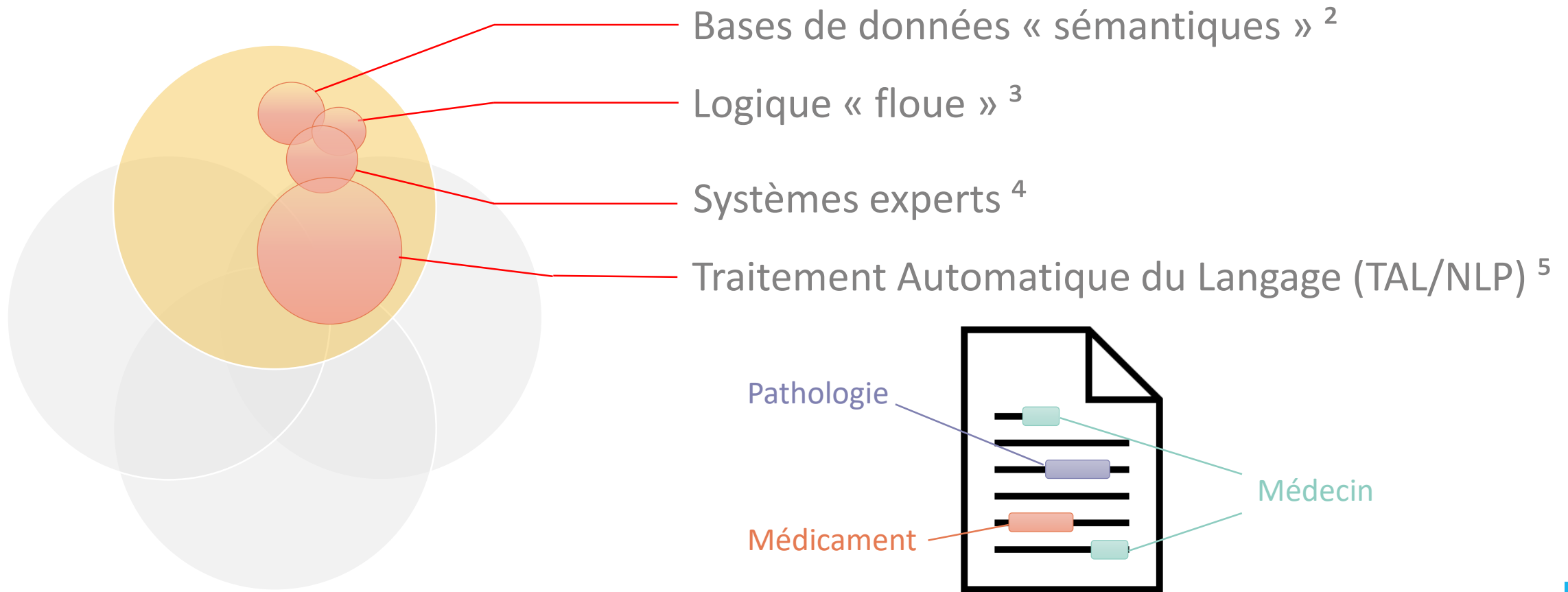


○ Sous-domaine de recherche

QU'EST-CE QUE « L'INTELLIGENCE ARTIFICIELLE » ?

7

LA REPRÉSENTATION DES CONNAISSANCES ¹

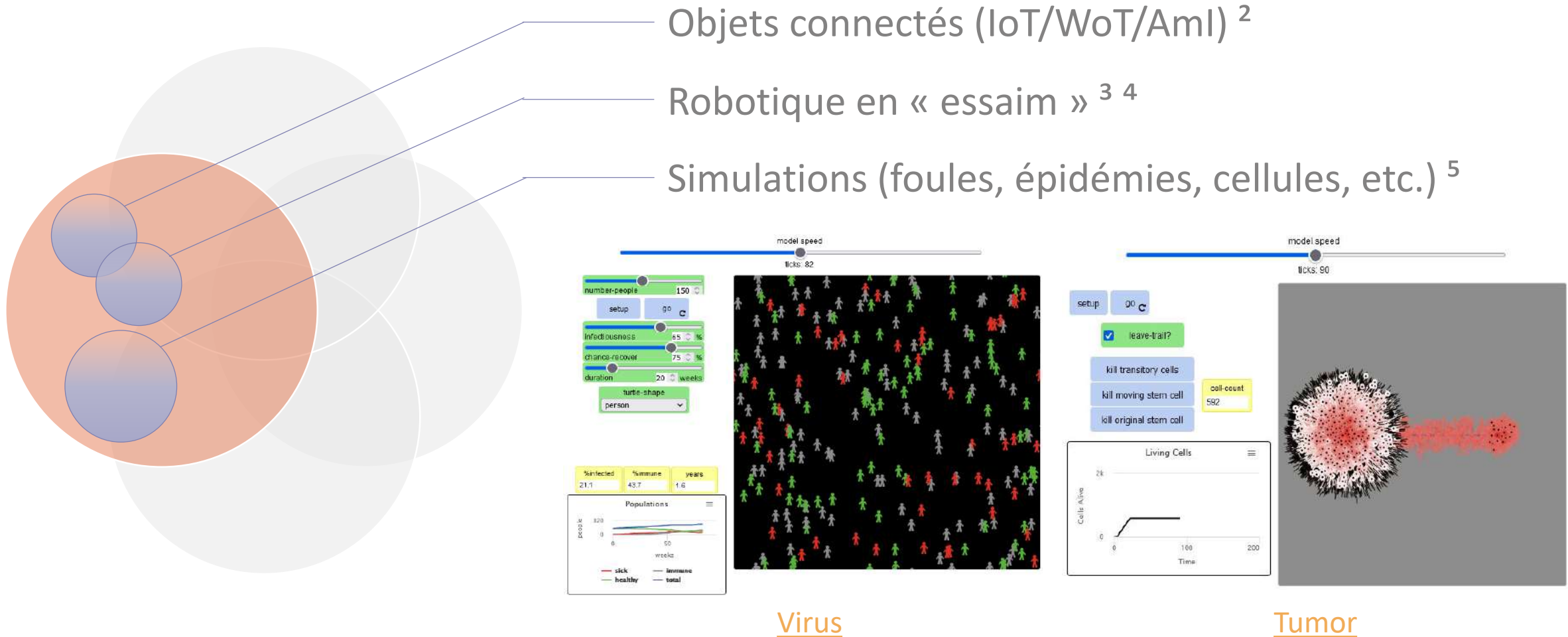


1. [Kendal S. and Creen M. \(2007\) – « An introduction to knowledge engineering »](#)
2. [Decker *et al.* \(2000\) - Framework for the semantic web: an RDF tutorial](#)
3. [L. A. Zadeh \(1977\) - Fuzzy Logic](#)
4. [Jackson \(1986\): Introduction to expert systems](#)
5. [Chowdhary \(2020\) - Natural Language Processing](#)

QU'EST-CE QUE « L'INTELLIGENCE ARTIFICIELLE » ?

8

LES SYSTÈMES MULTI-AGENTS (SMA/MAS) ¹

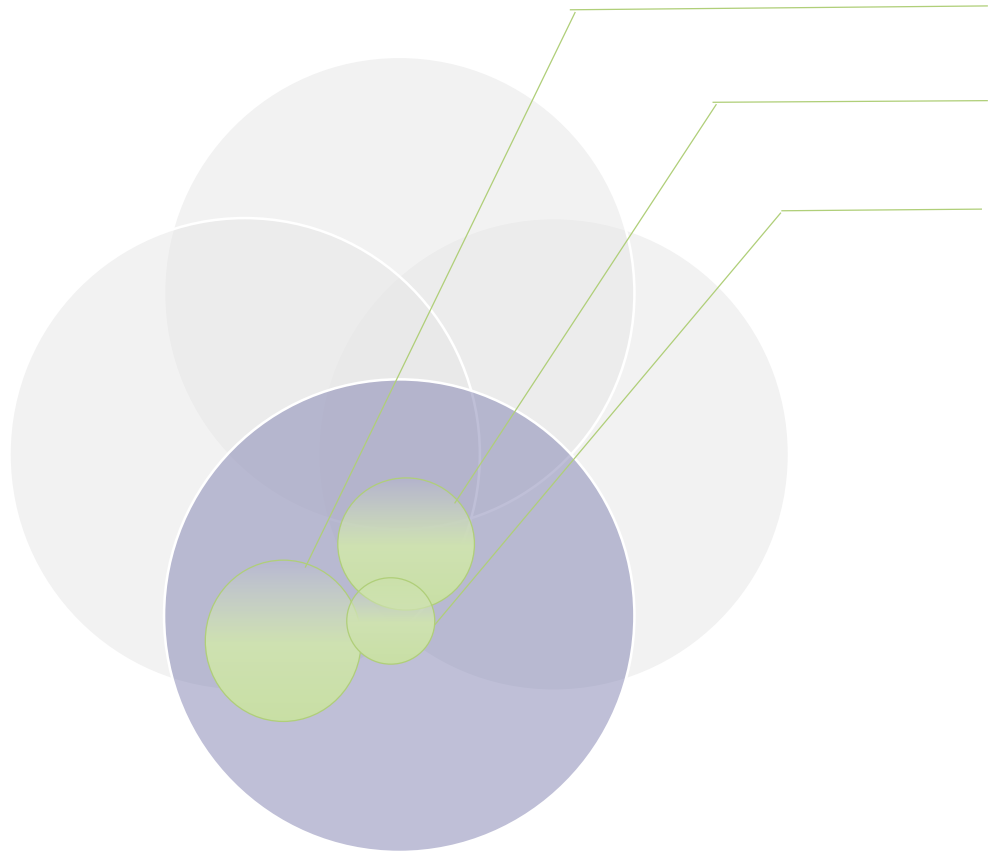


1. [Balaji P. G. and Srinivasan D. \(2010\) – « An Introduction to Multi-Agent Systems »](#)
2. [Darshan and Anandakumar \(2015\) - A comprehensive review on usage of Internet of Things \(IoT\) in healthcare system](#)
3. [Dorigo et al. \(2013\) - Swarmanoid: a novel concept for the study of heterogeneous robotic swarm](#)
4. [Fouloscopie – Le grand tournoi de robotique collective](#)
5. [Varenne and Silberstein \(2013\) - Modéliser & simuler \(tome 1\)](#)

QU'EST-CE QUE « L'INTELLIGENCE ARTIFICIELLE » ?

9

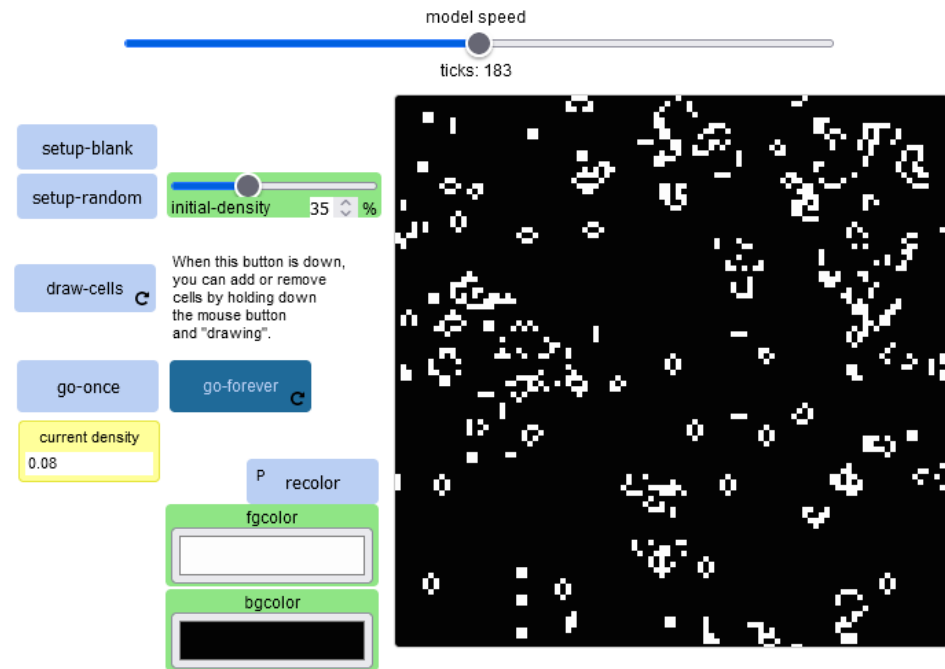
ALGORITHMES « BIO-INSPIRÉS » ¹



Colonies d'insectes « sociaux » ²

Algorithmes génétiques ³

« Vie artificielle » ^{4 5}



Game of Life

1. [Fan X., Sayers W., Zhang S. et al. – « Review and Classification of Bio-inspired Algorithms and Their Applications »](#)

2. [Dorigo and Stützle \(2018\) - Ant Colony Optimization: Overview and Recent Advances](#)

3. [Katoch, Chauhan and Kumar \(2021\) - A review on genetic algorithm: past, present, and future](#)

4. [Sarkar \(2000\) - A brief history of cellular automata](#)

5. [ScienceEtonnante – « LENIA: Une nouvelle forme de vie mathématique »](#)

QU'EST-CE QUE « L'INTELLIGENCE ARTIFICIELLE » ?

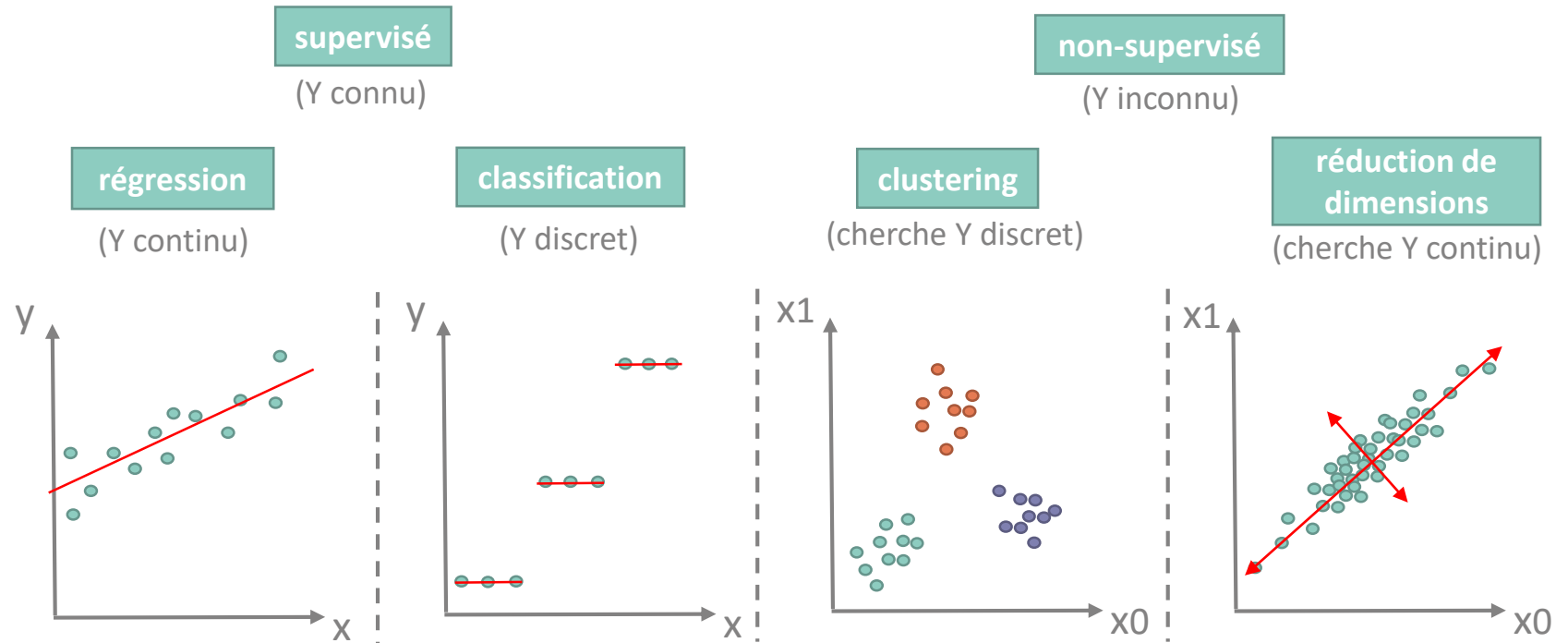
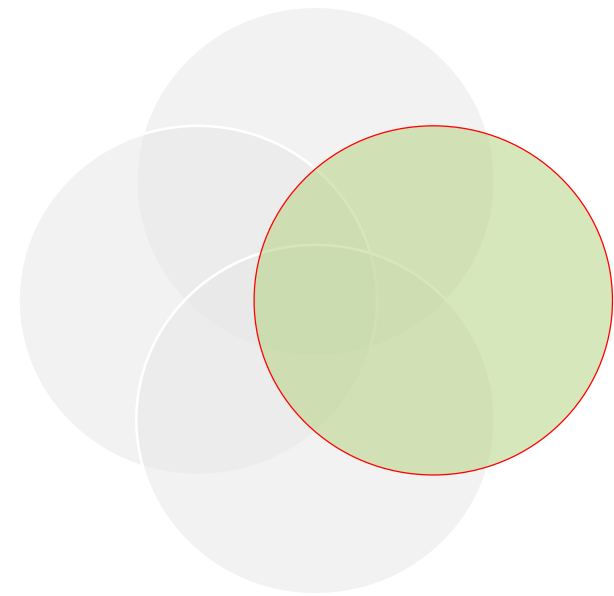
10

APPRENTISSAGE MACHINE (ML) ^{1 2 3}

On a des données
X et (parfois) Y

Apprentissage

On cherche à approximer
 $F: X \rightarrow Y$

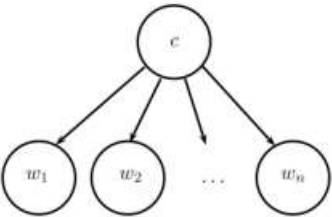
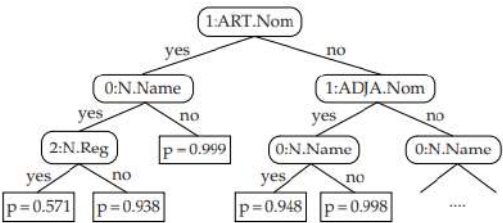
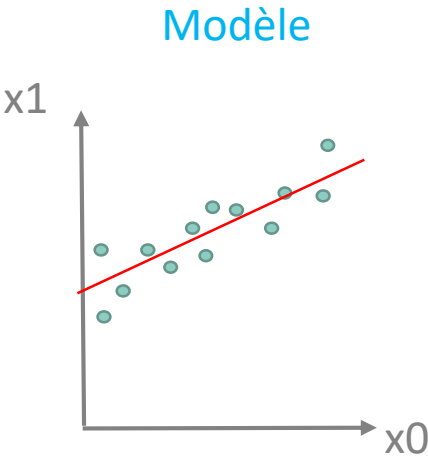


1. [Alloghani et al. \(2019\) - A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science](#)
2. [Shailaja, Seetharamulu and Jabbar \(2018\) - Machine Learning in Healthcare: A Review](#)
3. [Dhillon and Singh \(2019\) - Machine Learning in Healthcare Data Analysis: A Survey](#)

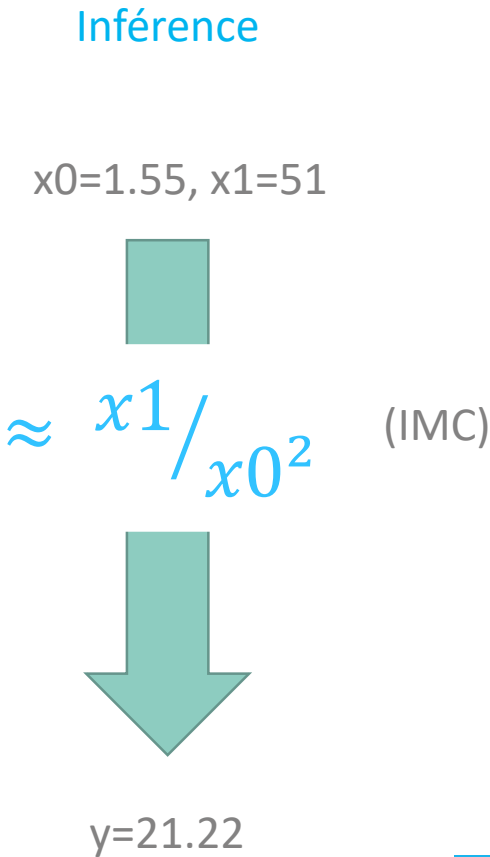
EXEMPLE

Objectif: $F: X \mapsto Y$

I		S
x0	x1	y
1.98	107	27.29
1.52	60	25.97
1.56	49	20.13
1.96	143	37.22
1.82	74	22.34
1.67	64	22.95
1.91	55	15.08
...
1.86	98	28.32



...

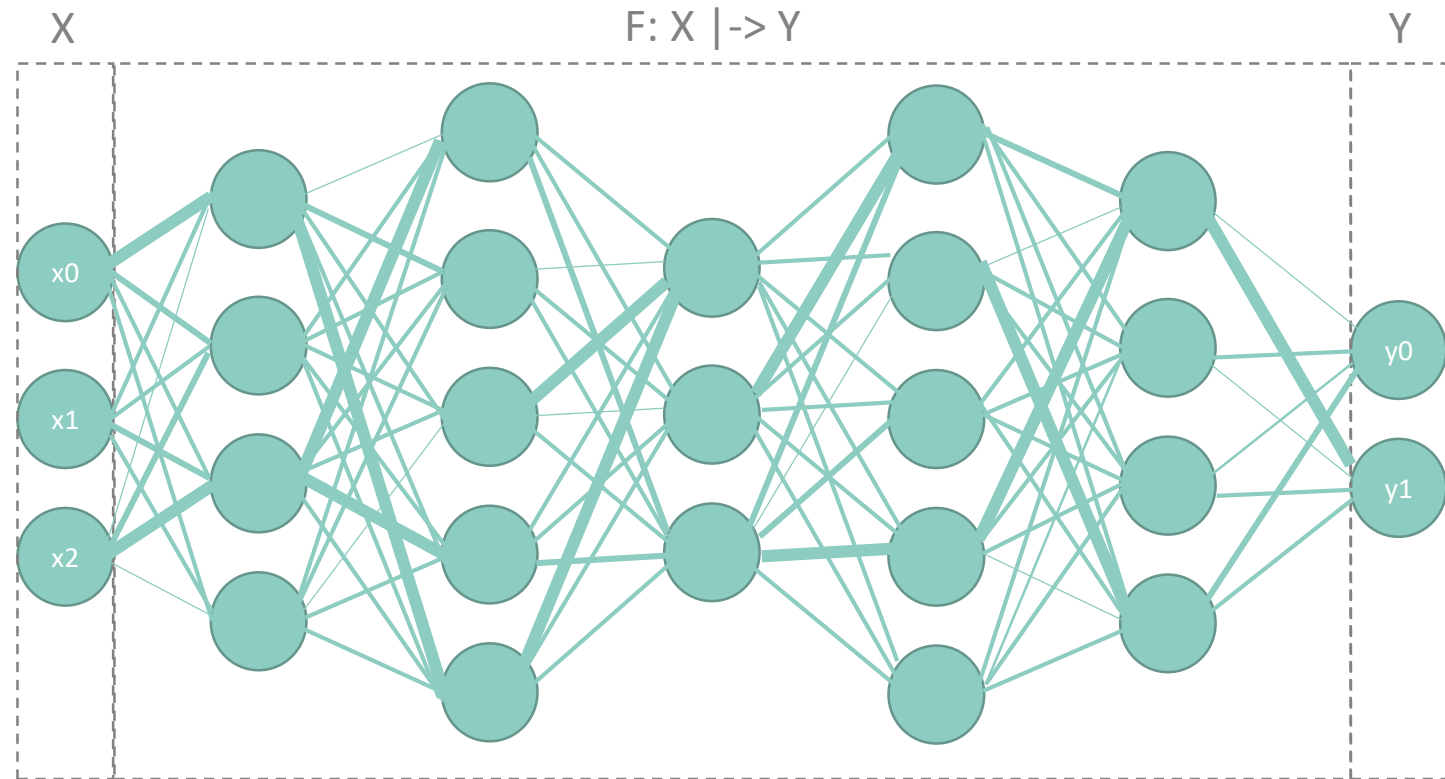


QU'EST-CE QUE « L'INTELLIGENCE ARTIFICIELLE » ?

12

APPRENTISSAGE MACHINE

Apprentissage « profond » (DL) ^{1 2 3 4}



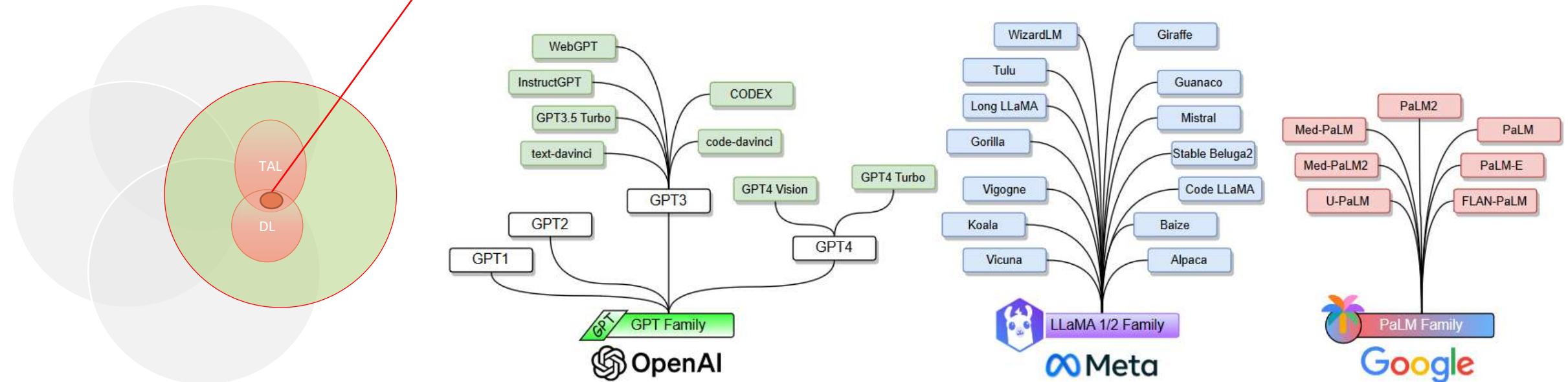
1. [LeCun, Bengio and Hinton \(2015\) - Deep Learning](#)
2. [Esteva et al. \(2019\) - A guide to deep learning in healthcare](#)
3. [Scarselli and Tsoi \(1998\) - Universal Approximation Using Feedforward Neural Networks: A Survey of Some Existing Methods, and Some New Results](#)
4. Pour aller plus loin: [Formation Fidle au Deep Learning](#)

QU'EST-CE QUE « L'INTELLIGENCE ARTIFICIELLE » ?

13

APPRENTISSAGE MACHINE

Large Language Models (LLM) ^{1 2}

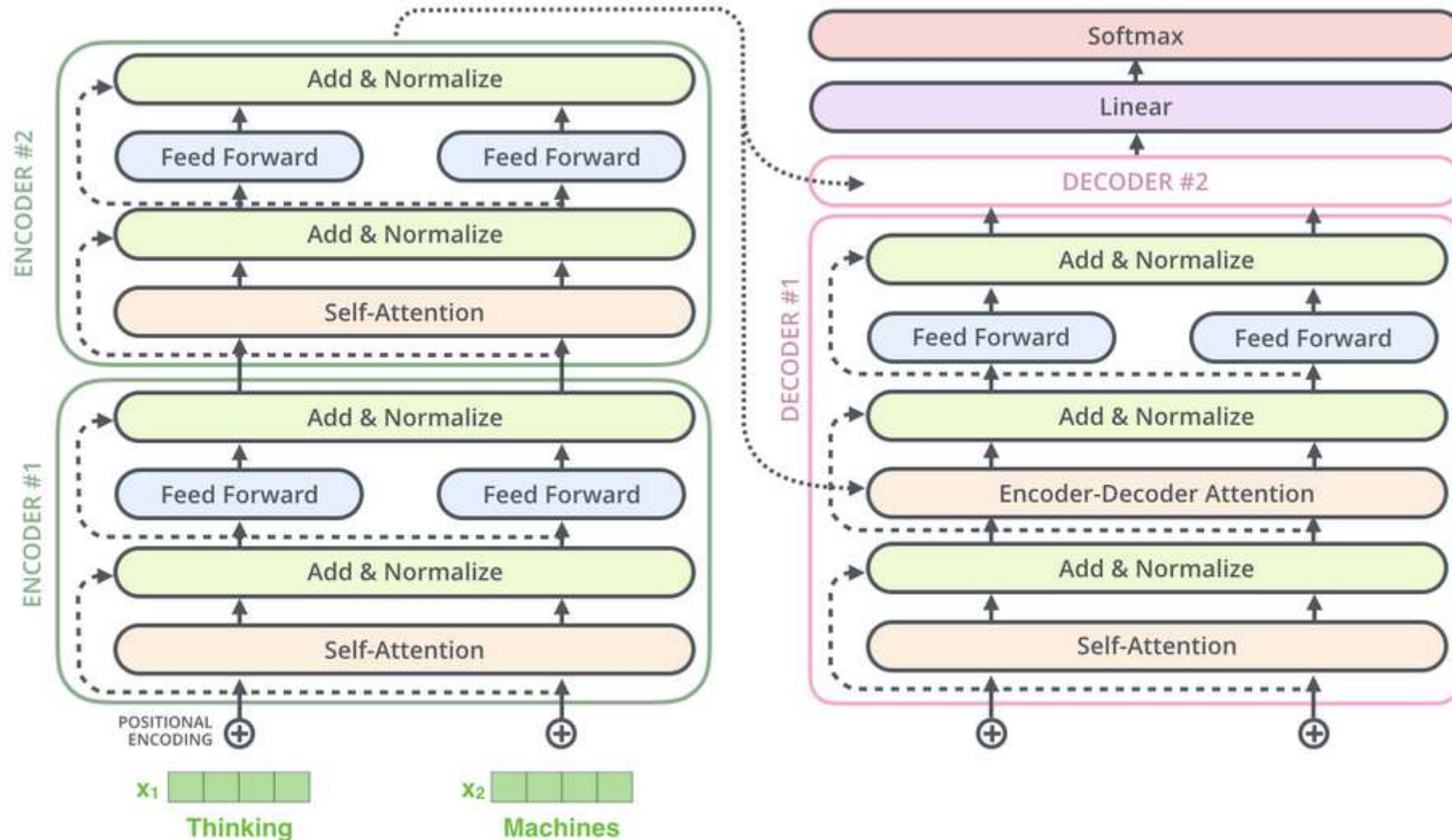


1. [Zhao, Zhou, Li et al. \(2023\) – « A Survey of Large Language Models »](#)
2. [Minaee, Mikolov, Nikzad et al. \(2024\) – « Large Language Models: A Survey »](#)

MODÈLES DE LANGAGE

14

LES TRANSFORMERS ^{1 2}



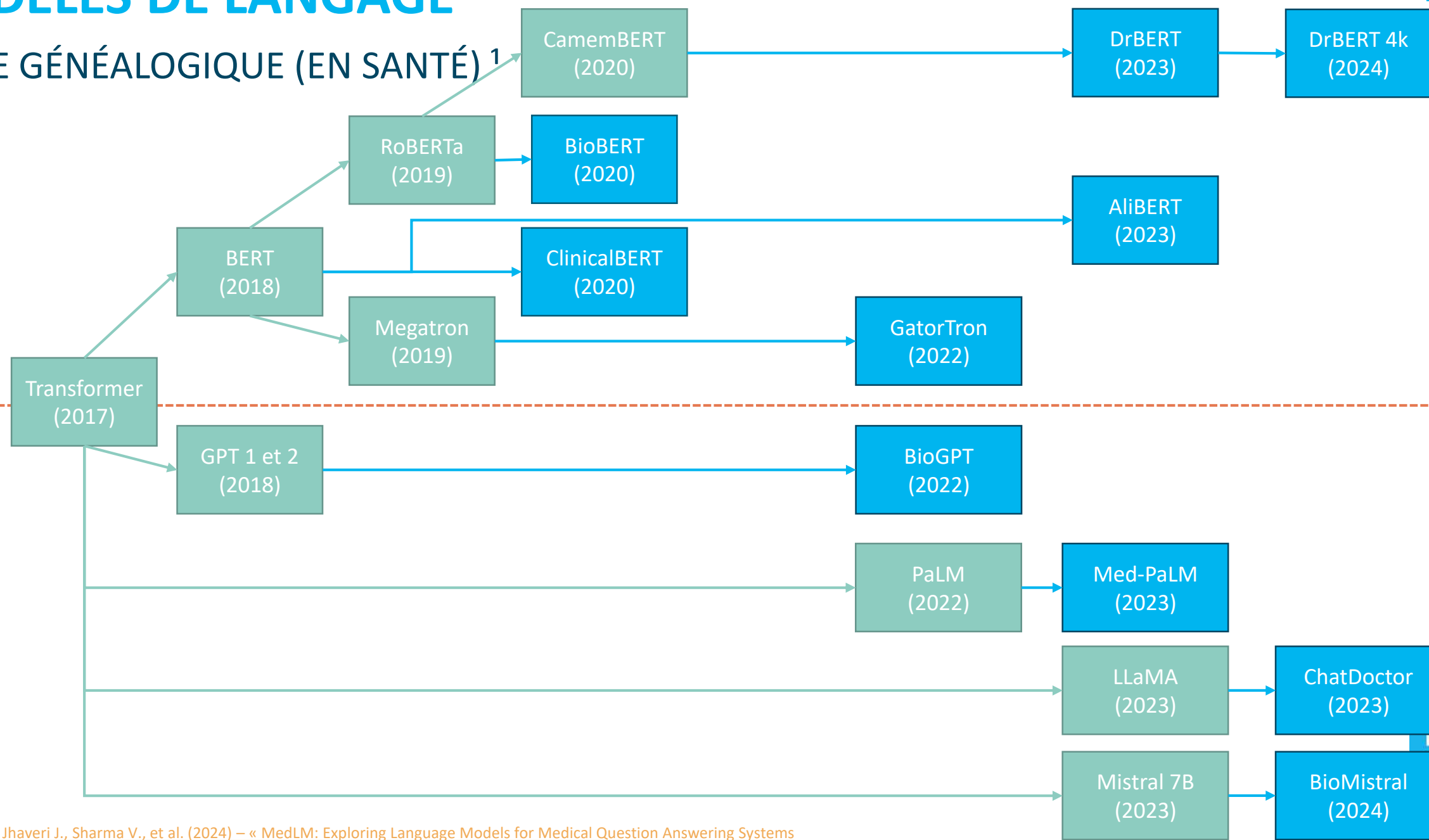
1. [Vaswani A., Shazeer N., Parmar N. et al. \(2017\) – « Attention Is All You Need »](#)
2. <https://jalammar.github.io/illustrated-transformer/>

MODÈLES DE LANGAGE

ARBRE GÉNÉALOGIQUE (EN SANTÉ) ¹

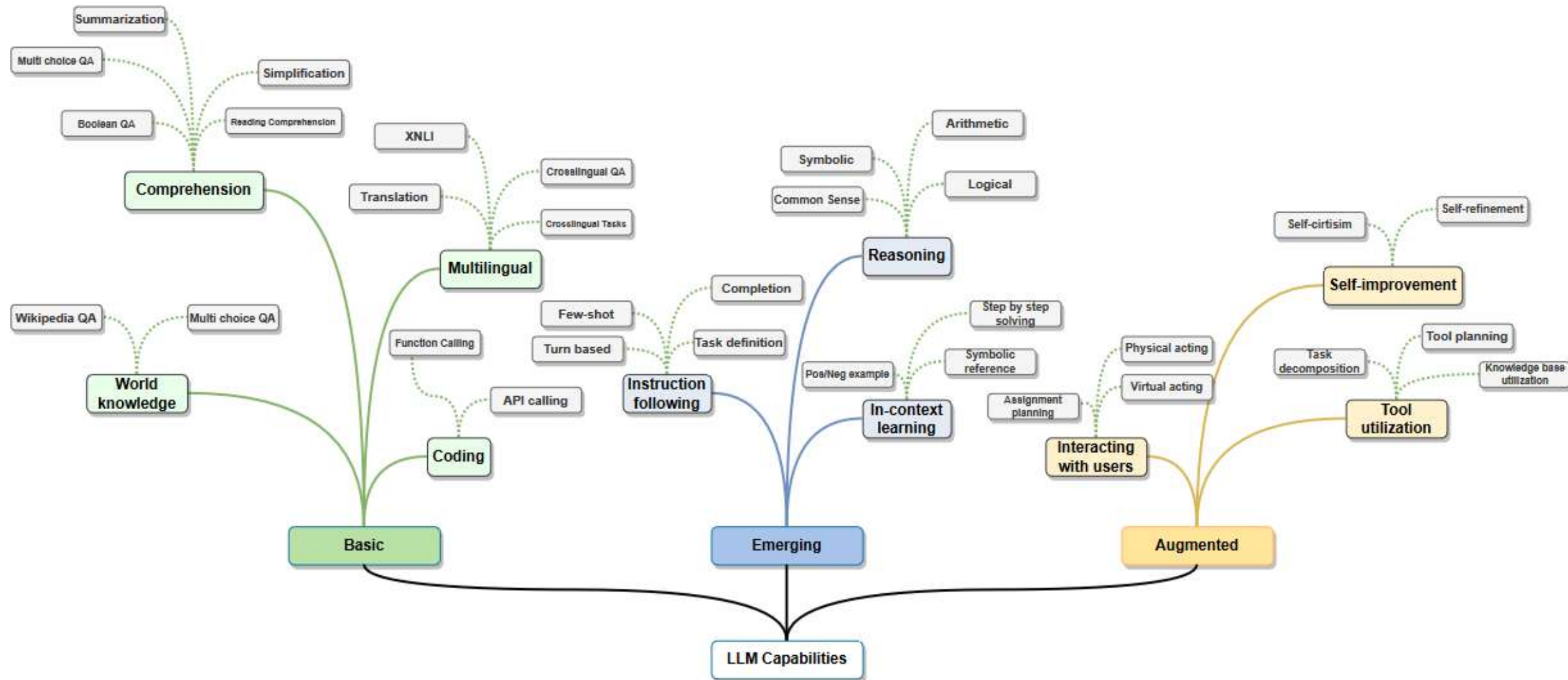
Encoder

Decoder



1. Yagnik N., Jhaveri J., Sharma V., et al. (2024) – « MedLM: Exploring Language Models for Medical Question Answering Systems »

UN ÉVENTAIL DE POSSIBILITÉS ¹



1. [Minaee, Mikolov, Nikzad et al. \(2024\) – « Large Language Models: A Survey »](#)

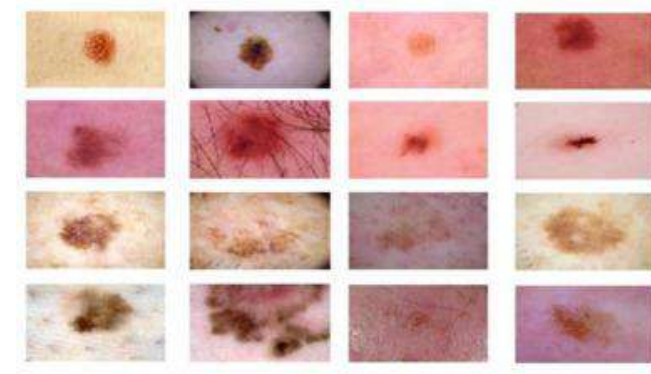
POURQUOI ? COMMENT ? ¹

Raisons ² :

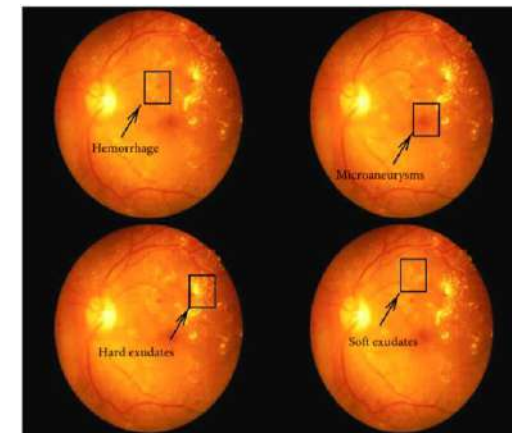
- Des hôpitaux largement informatisés
- De large bases de données disponibles
- Des modèles pouvant:
 - Compléter l'avis des médecins ³
 - Éviter des tests invasifs pour les patients

Néanmoins :

- Des données sensibles
- Risque d'atteinte à la vie privée
- Besoin de puissance de calculs
- Risque d'impact négatif sur les parcours de soins



Détection de tumeurs de peau⁴



Détection de rétinopathies diabétiques⁵

1. [Yu K., Beam A., and Kohane I. \(2018\) – « Artificial Intelligence in Healthcare »](#)
2. [Davenport T. and Kalakota R. \(2019\) – « The Potential for Artificial Intelligence in Healthcare »](#)
3. [Haenssle H.A., Winkler J.K., Fink C. et al \(2021\) – « Skin lesions of face and scalp – Classification by a market-approved convolutional neural network in comparison with 64 dermatologists »](#)
4. [Choudhary P., Singhai J., and Yadav J.S. \(2022\) – « Skin lesion detection based on deep neural network »](#)
5. [Senapati A., Tripathy H.K., Sharma V. et al. \(2024\) – « Artificial intelligence for diabetic retinopathy detection: a systematic review »](#)

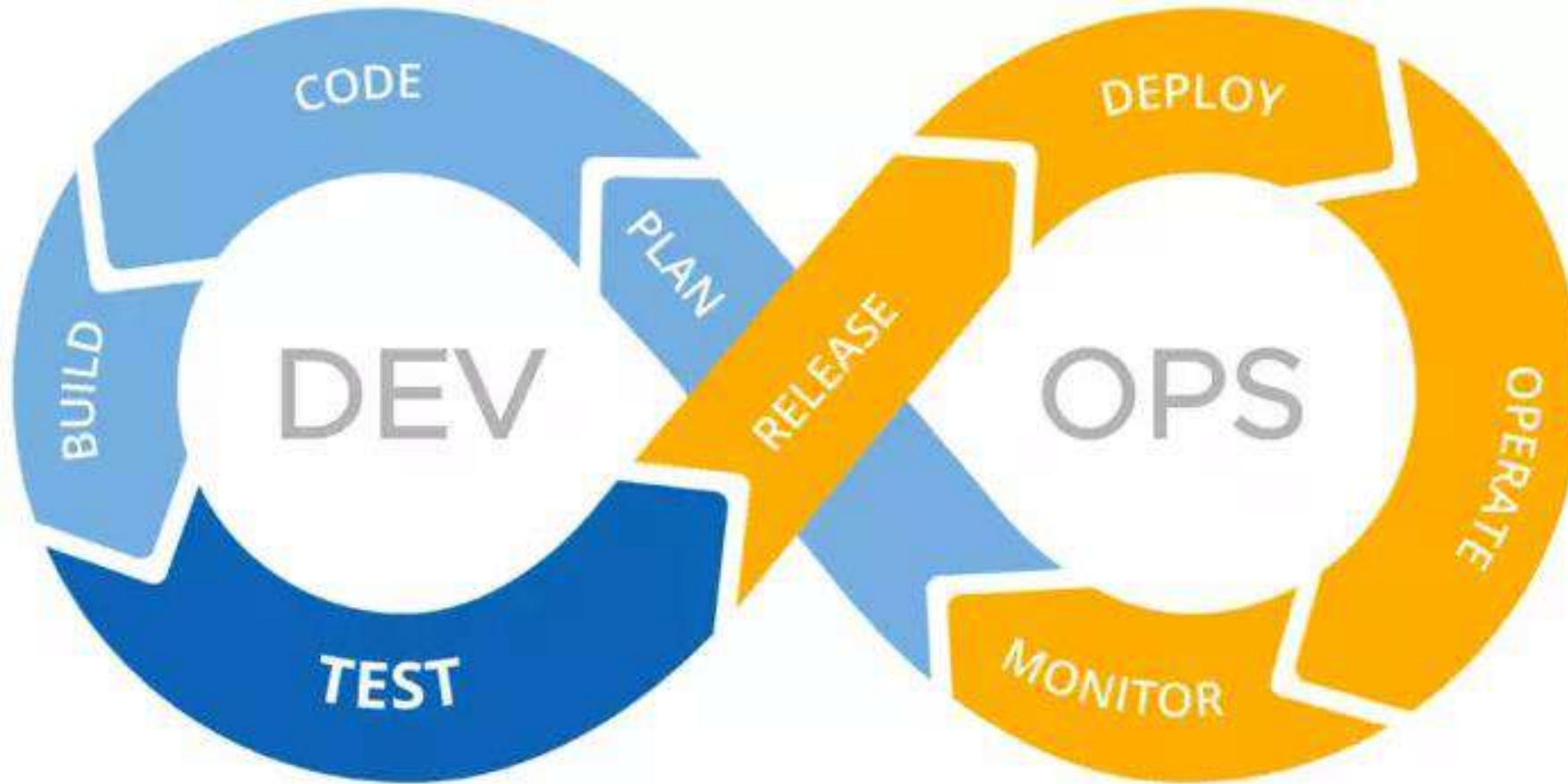
DÉVELOPPER ET INDUSTRIALISER DU ML

PRINCIPES DE BASE DU MLOPS

LE DEVOPS ET L'INTÉGRATION CONTINUE (CI/CD)

19

PRINCIPES DE BASE ^{1 2}

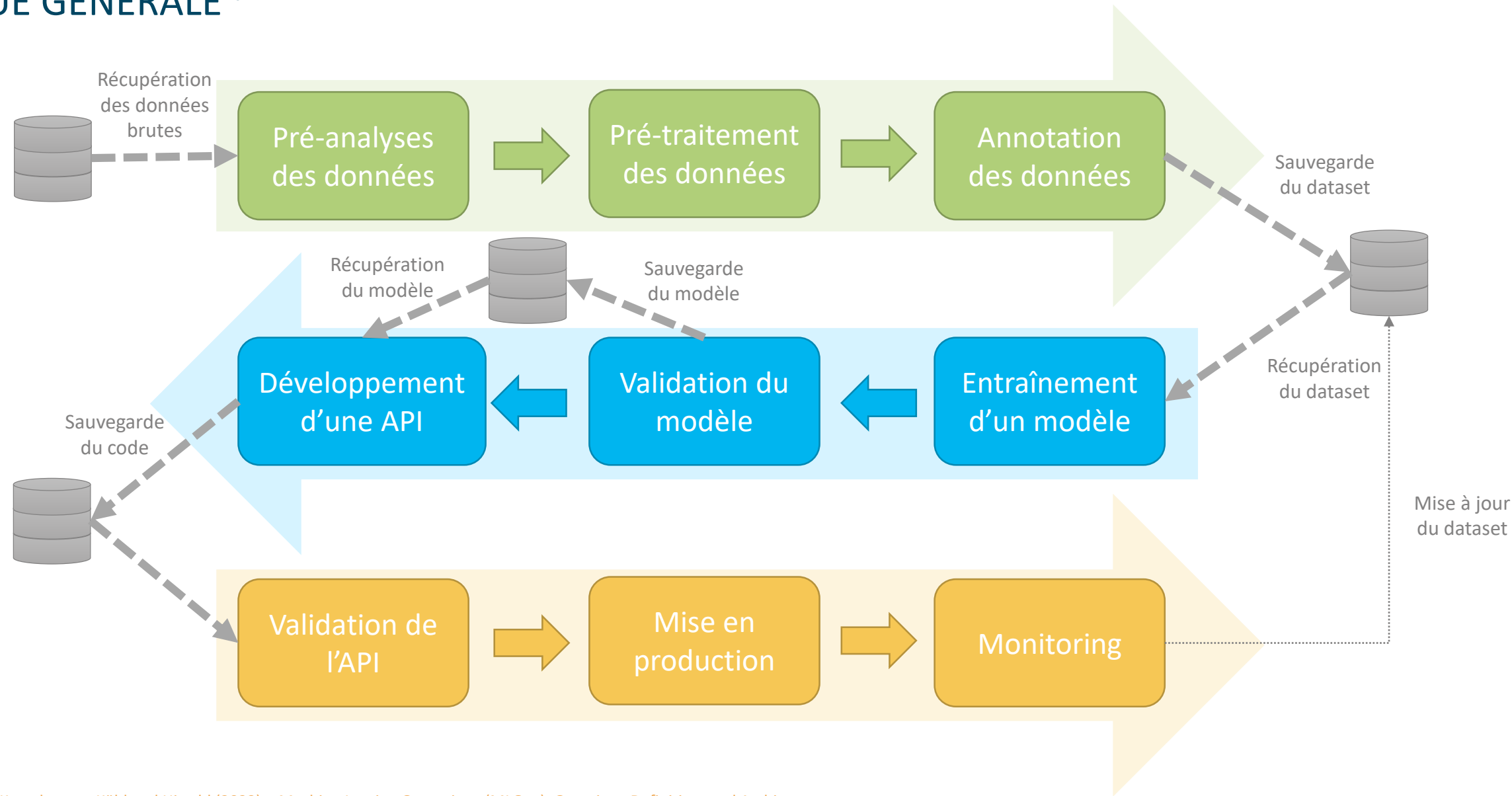


1. [Humble J., and Farley D. \(2010\) – « Continuous Delivery: Reliable Software Releases through Build, Test, and Deployment Automation »](#)

2. [Leite L., Rocha C., Kon F. et al. \(2019\) – « A Survey of DevOps Concepts and Challenges »](#)

LE MLOPS

VUE GÉNÉRALE ¹

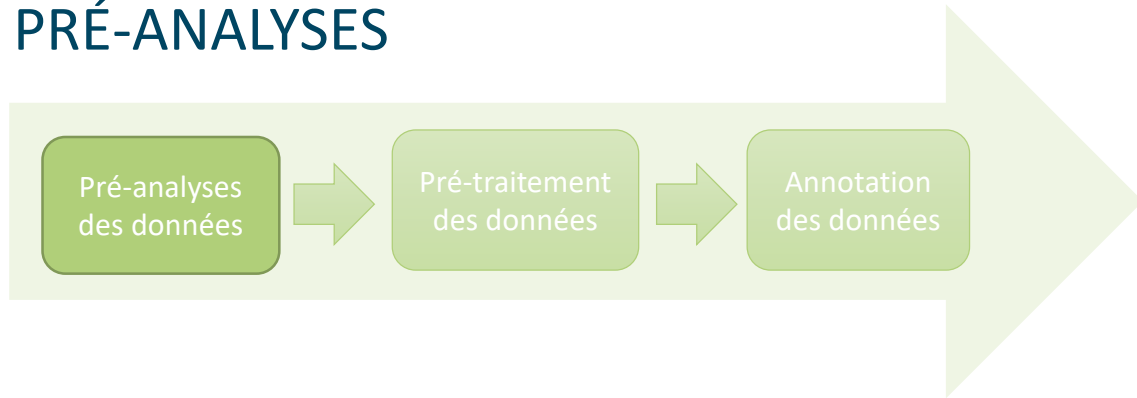


1. [Kreuzberger, Kühn and Hirschl \(2023\) – Machine Learning Operations \(MLOps\): Overview, Definition, and Architecture](#)

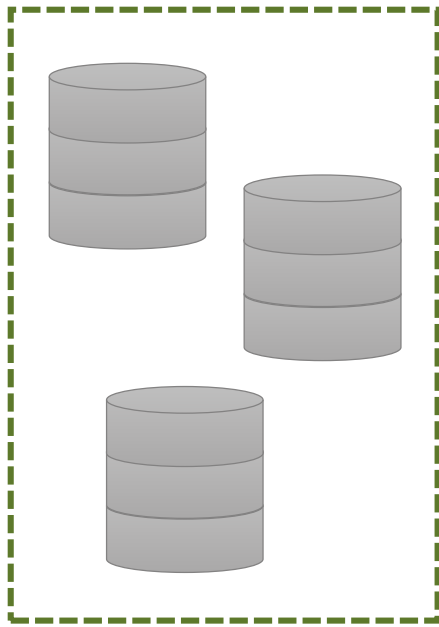
LE PROCESSUS DE PRÉPARATION DES DONNÉES ¹

21

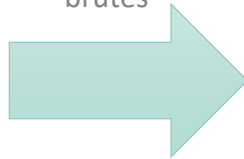
PRÉ-ANALYSES



Bases de Données

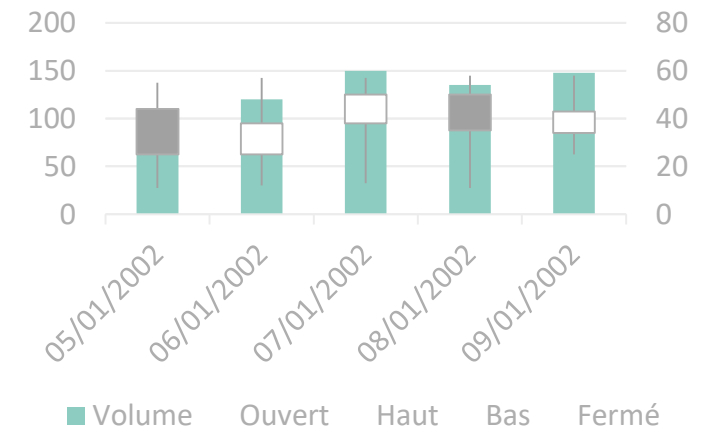
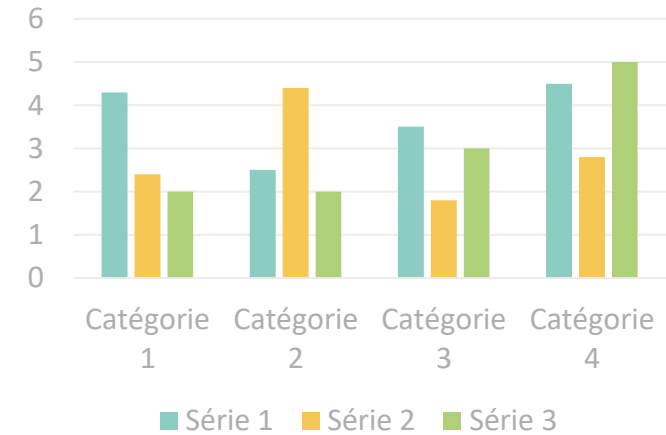


Récupération
des données
brutes



x0	x1	...	xN
198	107	...	27.29
152	60	...	25.97
156	49	...	20.13
196	143	...	37.22
182	74	...	22.34
167	64	...	22.95
191	55	...	15.08
...
186	98	...	28.32

Analyses
des
données

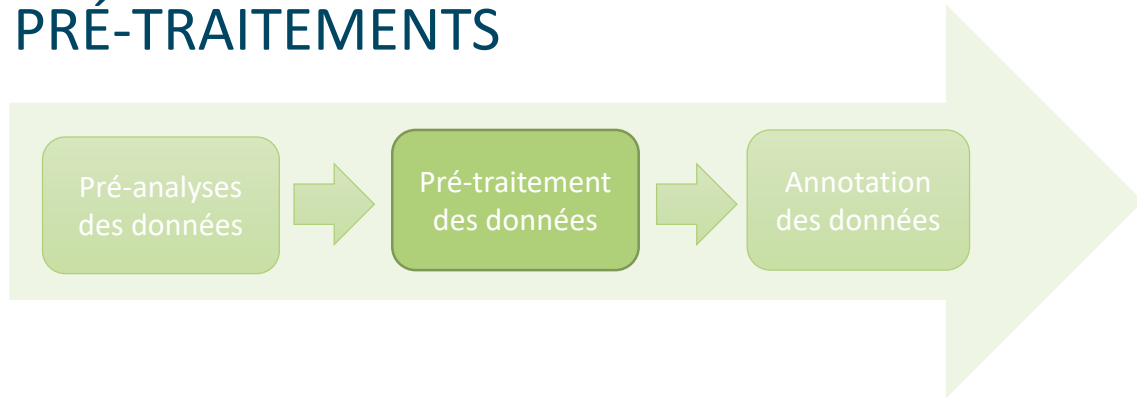


1. [Bergh C., Benghiat G., and Strod E. \(2019\) – « The DataOps Cookbook »](#)

LE PROCESSUS DE PRÉPARATION DES DONNÉES¹

22

PRÉ-TRAITEMENTS



x0	x1	...	xN
198	107	...	27.29
152	60	...	25.97
156	49	...	20.13
196	143	...	37.22
182	74	...	22.34
167	64	...	22.95
191	55	...	15.08
...
186	98	...	28.32

Identifications
des
variables



x0	x1	y
198	107	27.29
152	60	25.97
156	49	20.13
196	143	37.22
182	74	22.34
167	64	22.95
191	55	15.08
...
186	98	28.32

Transformations



Règles
de
transformations

x0	x1	y
1.98	107	27.29
1.52	60	25.97
1.56	49	20.13
1.96	143	37.22
1.82	74	22.34
1.67	64	22.95
1.91	55	15.08
...
1.86	98	28.32

1. [Bergh C., Benghiat G., and Strod E. \(2019\) – « The DataOps Cookbook »](#)

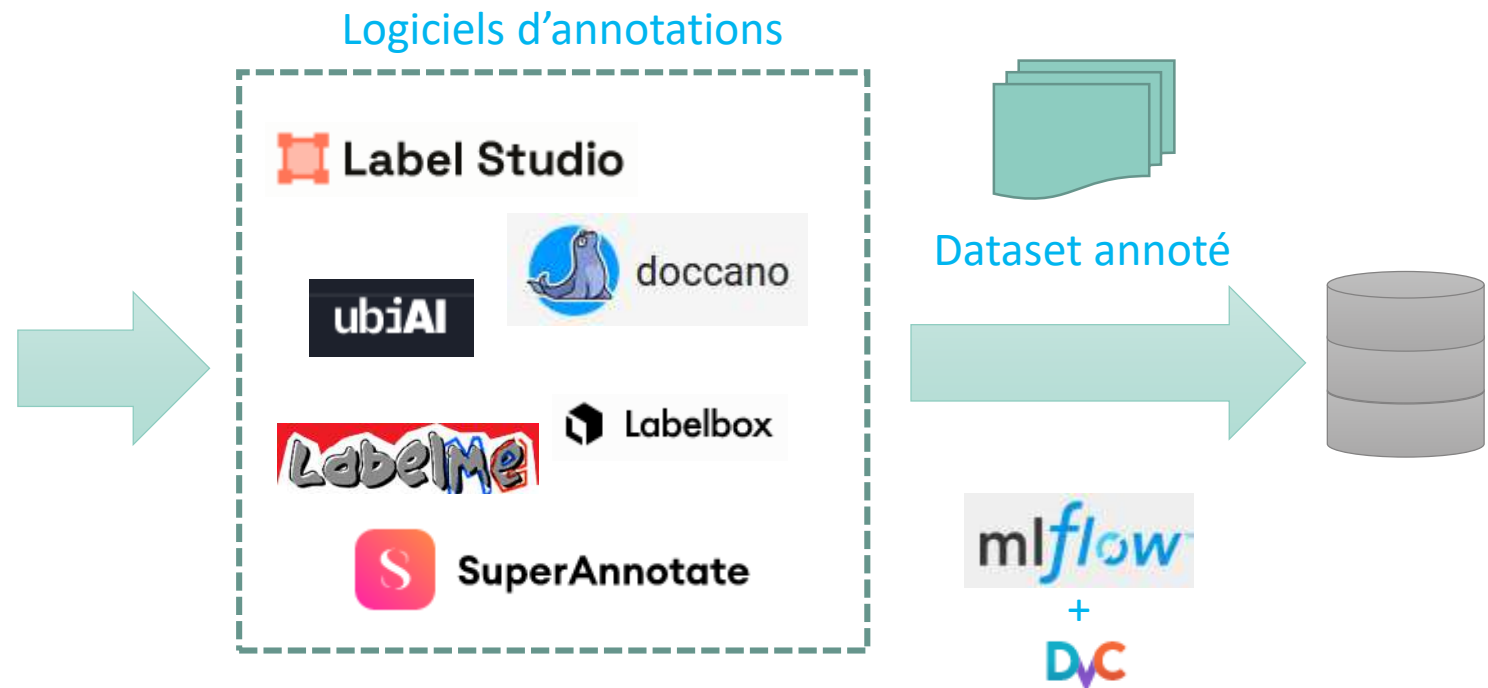
LE PROCESSUS DE PRÉPARATION DES DONNÉES¹

23

ANNOTATION DES DONNÉES²



img	txt	y
810.png	Lorem ipsum	A
17.png	Dolor sit	B
187.png	Amet consectetur	B
88.png	Adipiscing elit	A
22.png	Vestibulum enim	C
738.png	Diam hendrerit	A
361.png	Id est sed	B
...
42.png	Sollicitudin nulla	C



1. [Bergh C., Benghiat G., and Strod E. \(2019\) – « The DataOps Cookbook »](#)
2. [Grohmann R., and Fernandes Araújo W. \(2021\) – « Beyond Mechanical Turk: The Work of Brazilians on Global AI Platforms »](#)

ANNOTATION DES DONNÉES

24

EXEMPLE ¹

Text Classification

To have faith is to trust yourself to the water

Choose text sentiment

☒ Positive^[1] ☐ Negative^[2] ☐ Neutral^[3]

Entity

Nothing selected

Entities (0)
No Entities added yet

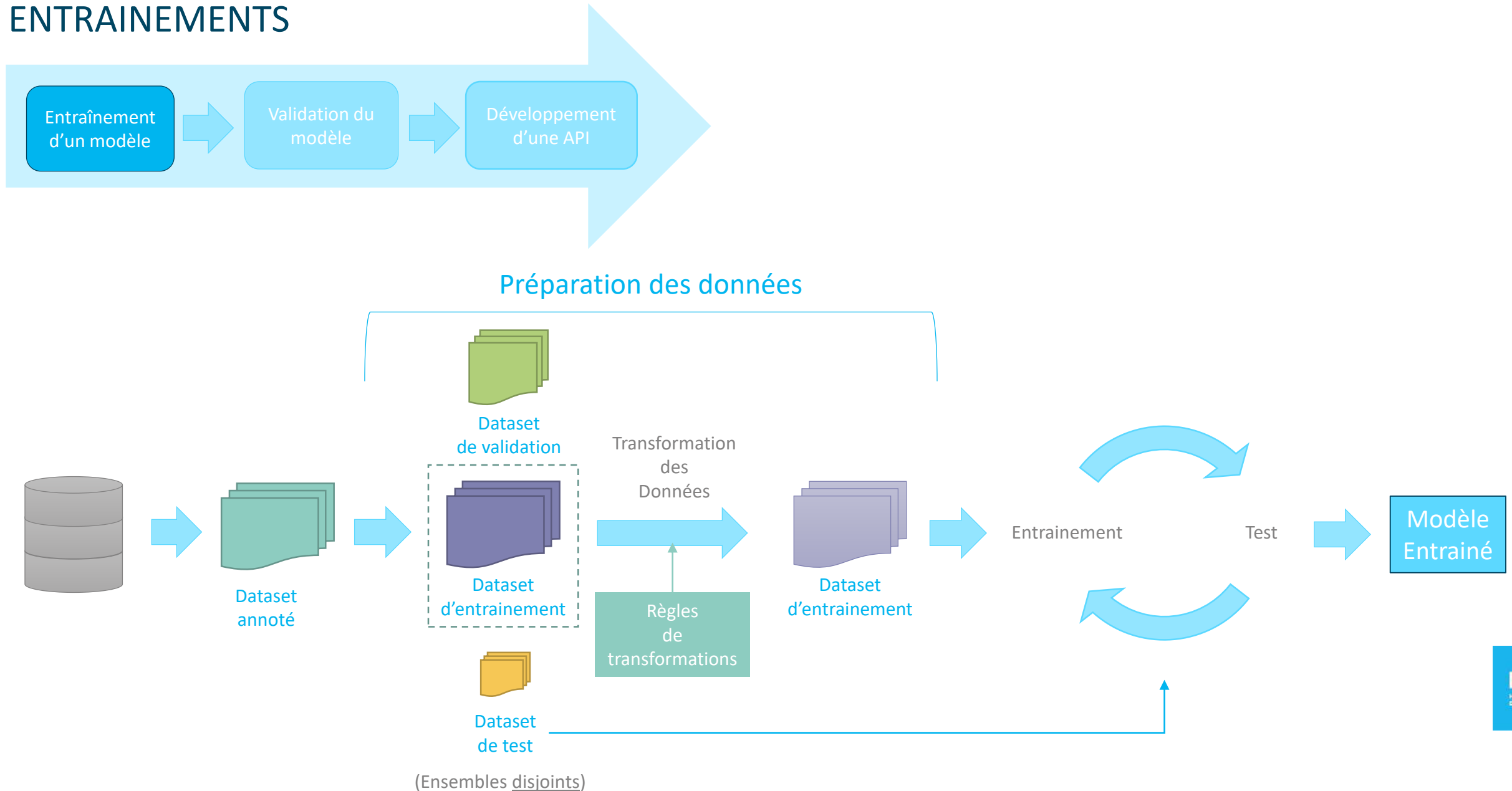
Relations (0)
No Relations added yet

1. <https://github.com/HumanSignal/label-studio>

LE PROCESSUS DE PRÉPARATION DU MODÈLE

25

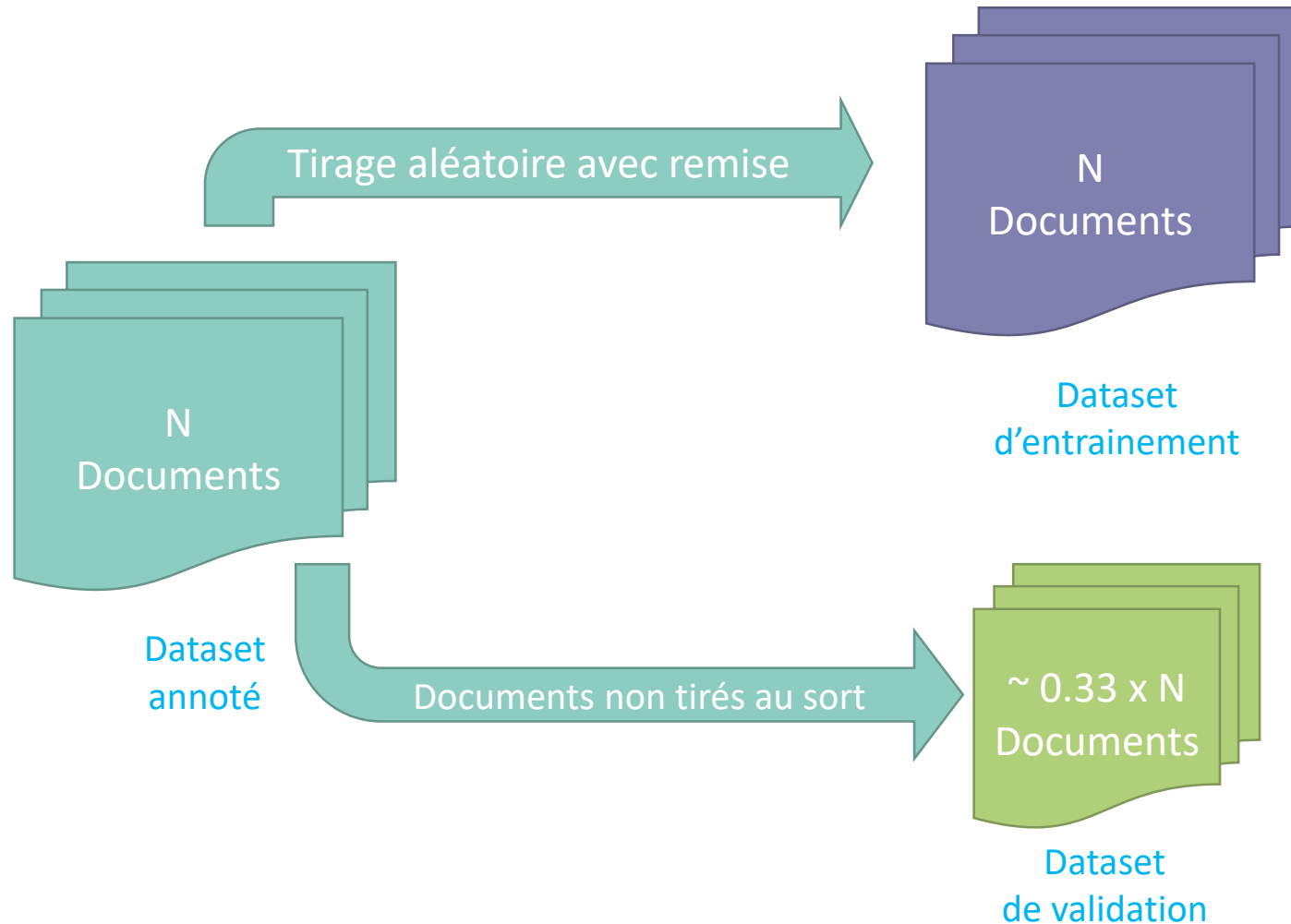
ENTRAÎNEMENTS



PRÉPARATION DES DONNÉES

BOOTSTRAPPING

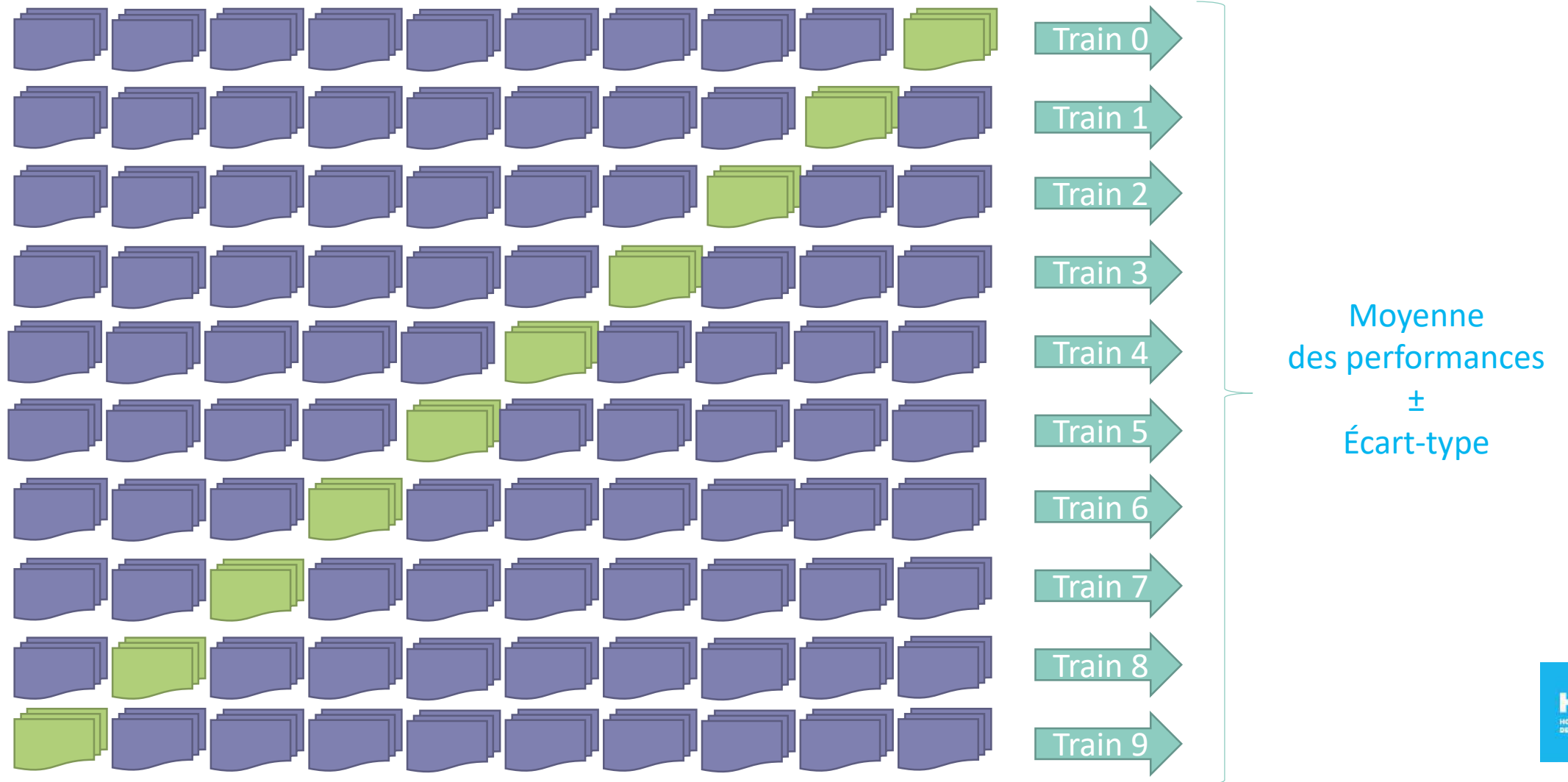
26



PRÉPARATION DES DONNÉES

27

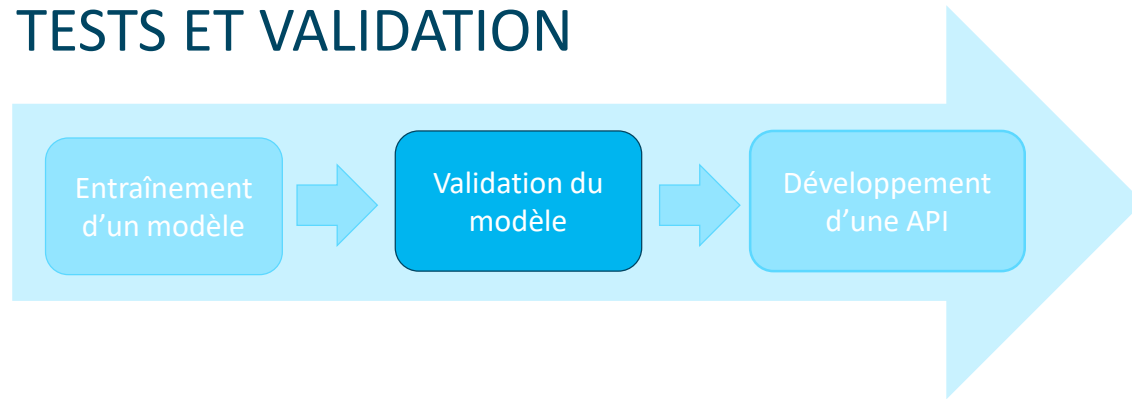
CROSS-VALIDATION



LE PROCESSUS DE PRÉPARATION DU MODÈLE

28

TESTS ET VALIDATION



	Prédiction Positive	Prédiction Négative
Valeur attendue Positive	Vrai Positifs (TP)	Faux Négatif (FN)
Valeur attendue Négative	Faux Positifs (FP)	Vrai Négatifs (TN)

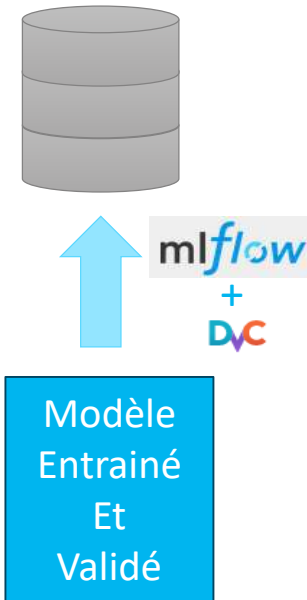
Prédiction	Valeur Attendue
24.89	25.12
22.15	22.10
35.23	31.5
24.12	27.82
...	...
32.81	30.22

Calculs de Performances

$$F = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

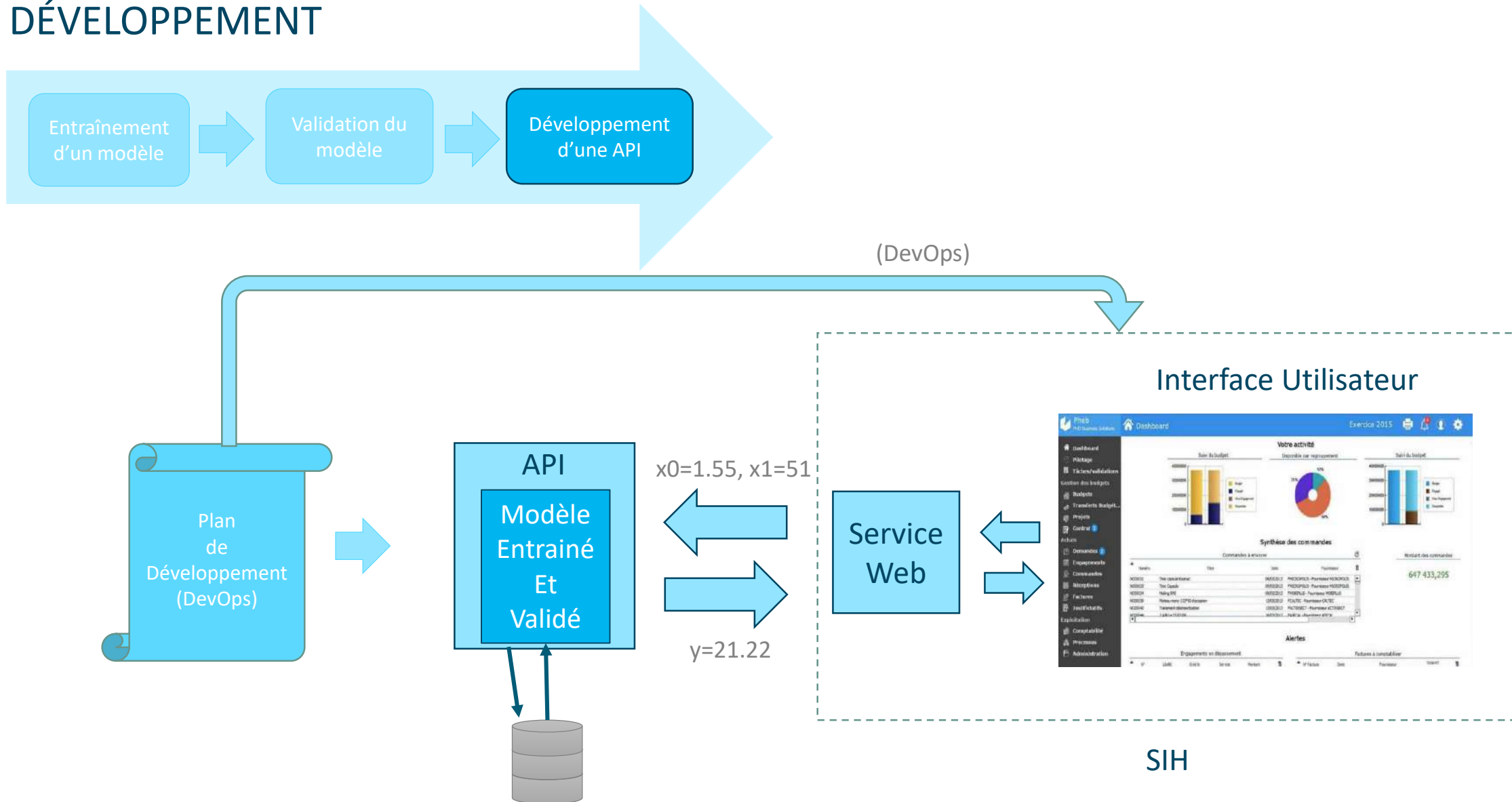
$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$



LE PROCESSUS DE PRÉPARATION DU MODÈLE

DÉVELOPPEMENT

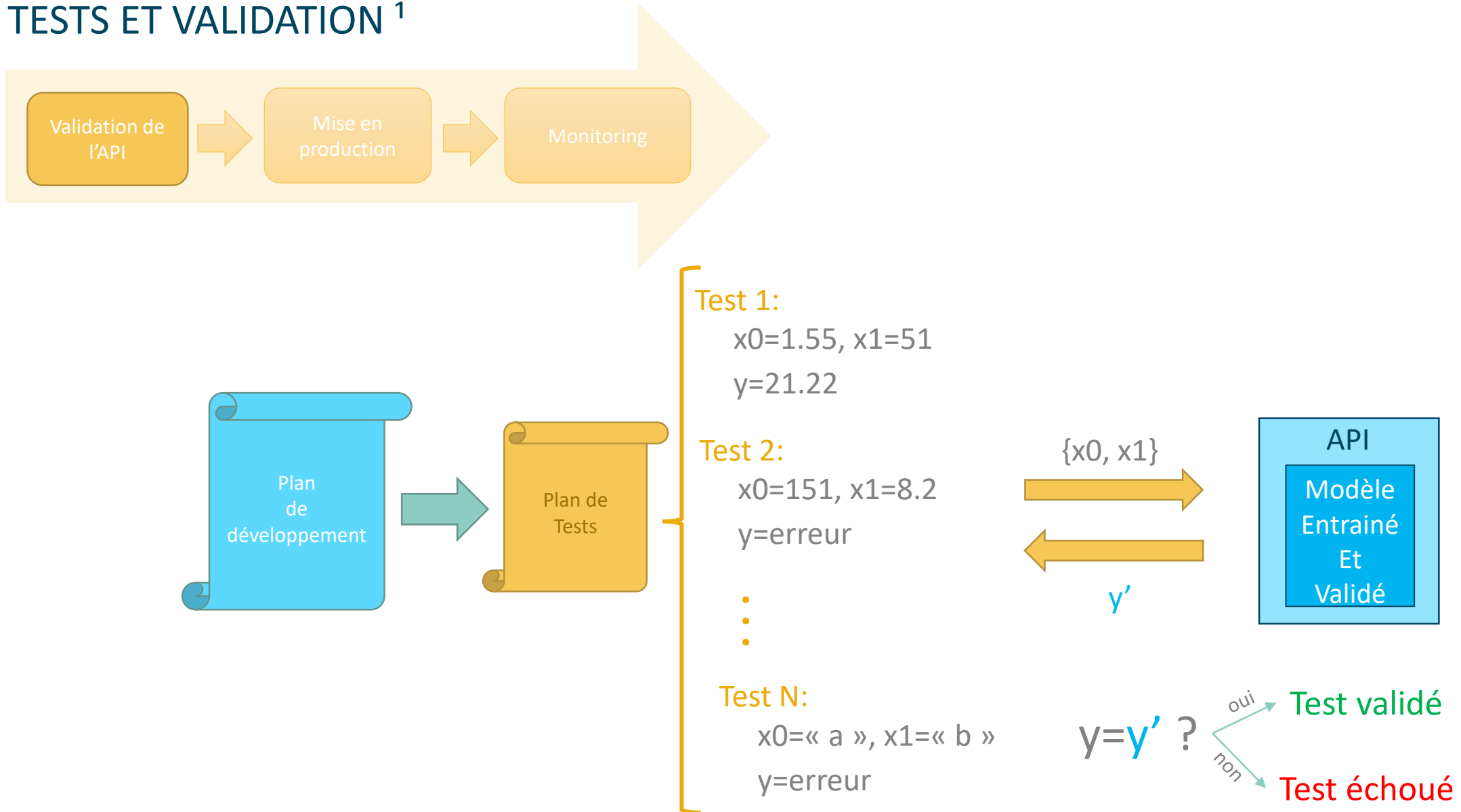
29



LE PROCESSUS DE DÉPLOIEMENT DU MODÈLE

30

TESTS ET VALIDATION ¹



1. [Beck K. \(2022\) – « Test-Driven Development: By Example »](#)

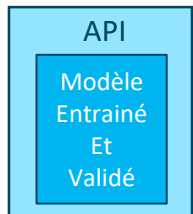
LE PROCESSUS DE DÉPLOIEMENT DU MODÈLE

31

INTÉGRATION ET MISE EN PROD



Serveurs de développement



Serveurs de Tests



Serveurs d'intégration



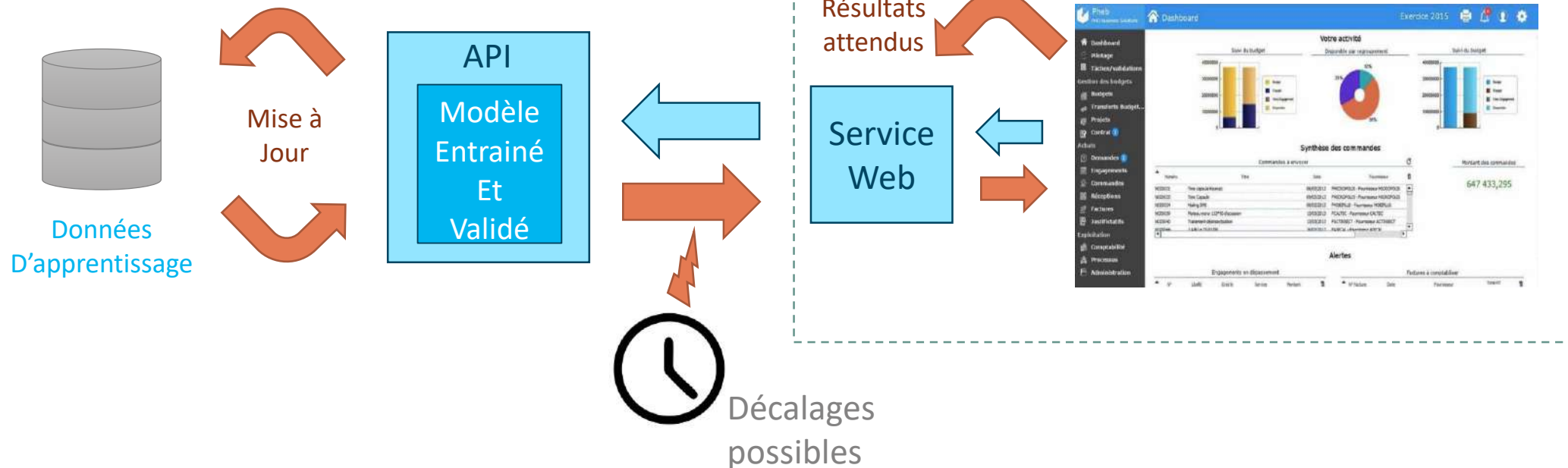
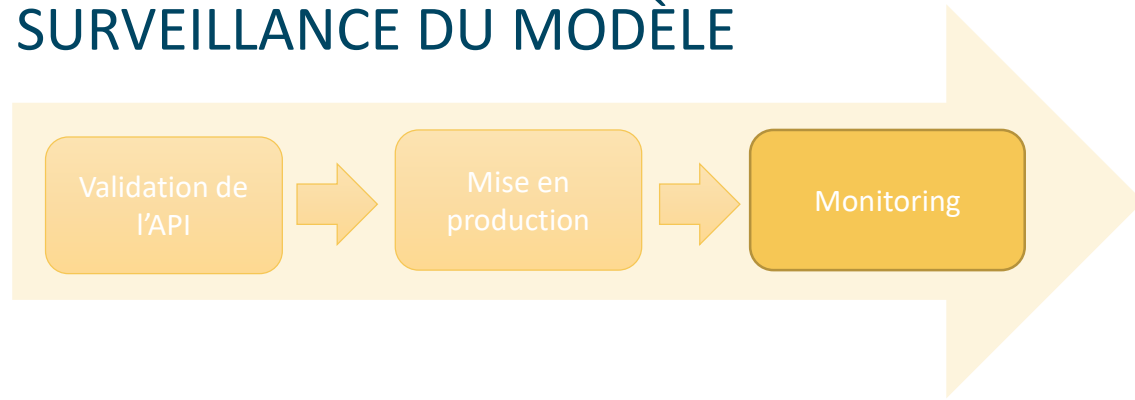
Serveurs de Production



LE PROCESSUS DE DÉPLOIEMENT DU MODÈLE

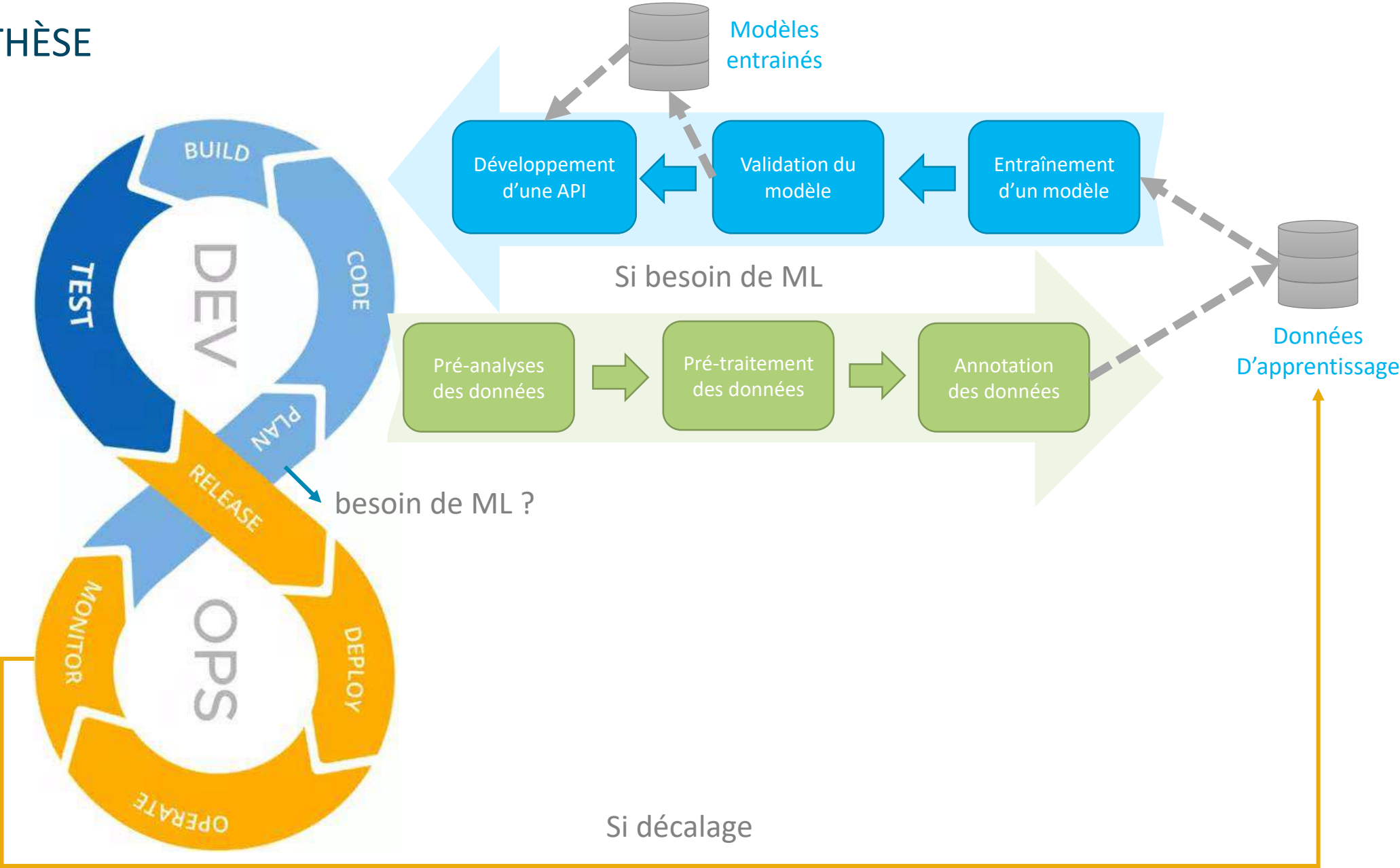
32

SURVEILLANCE DU MODÈLE



LE MLOPS

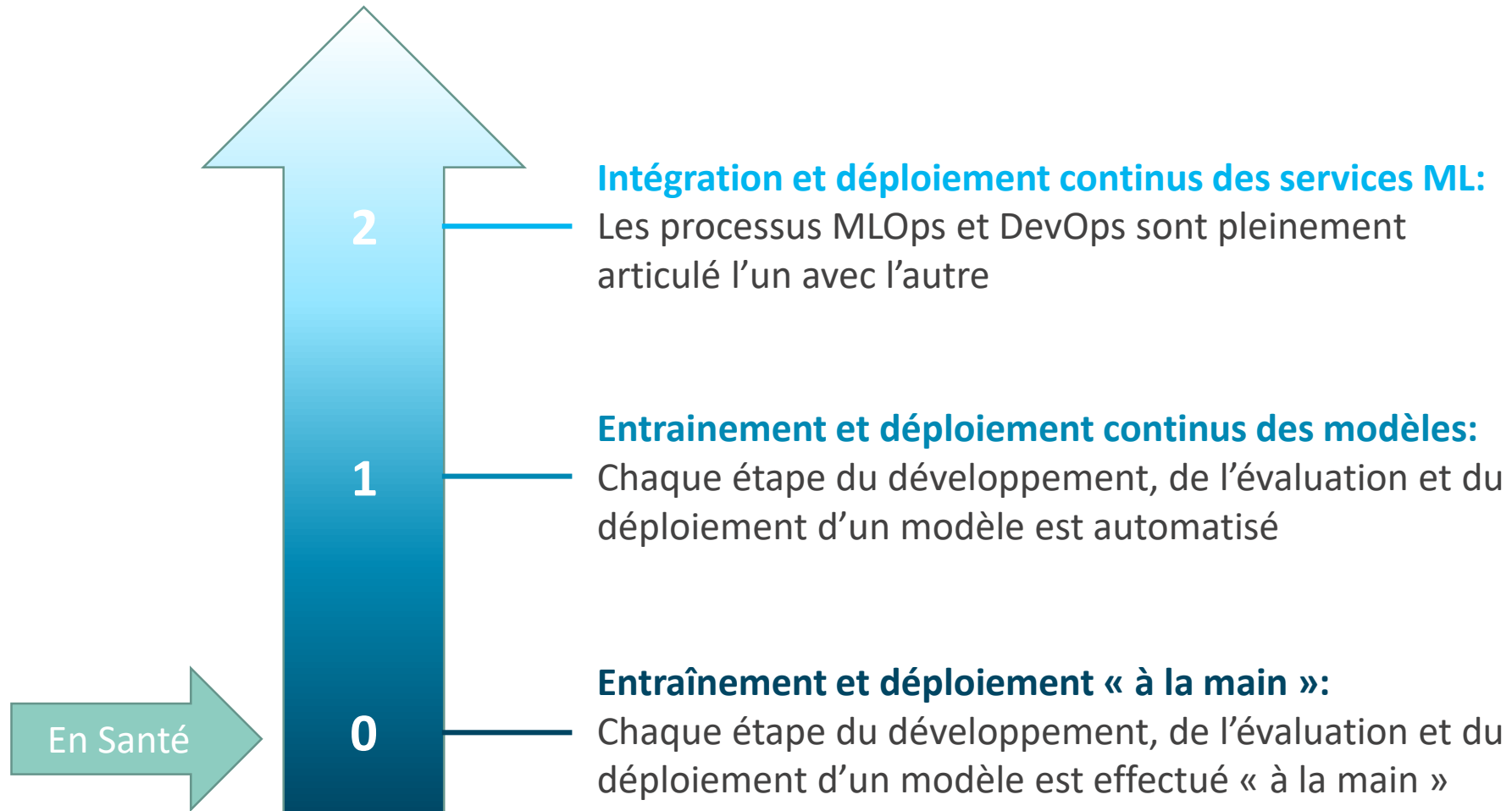
SYNTHÈSE



METTRE EN PLACE LE MLOPS

34

LES NIVEAUX DE MATURITÉ ¹



1. [Kreuzberger, Kühl and Hirschl \(2023\) – Machine Learning Operations \(MLOps\): Overview, Definition, and Architecture](#)

ADAPTER LE MLOPS À LA SANTÉ

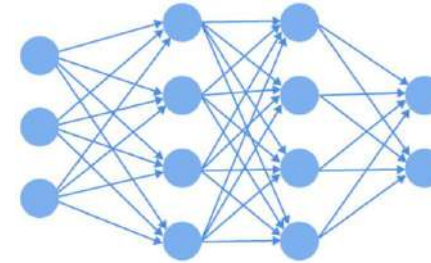
CONTRAINTES ET SOLUTIONS

LE MLOPS EN SANTÉ

COMMENT L'ADAPTER? ¹



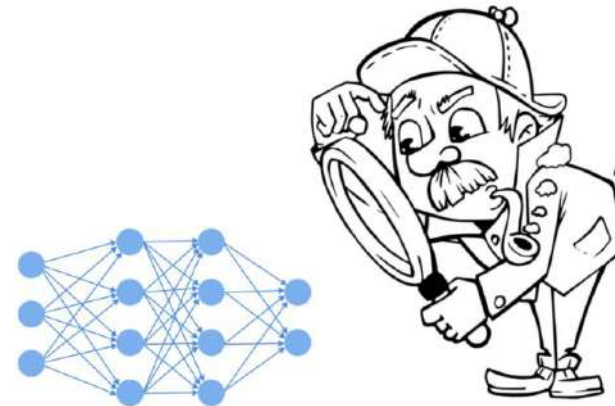
Données de santé



Modèles



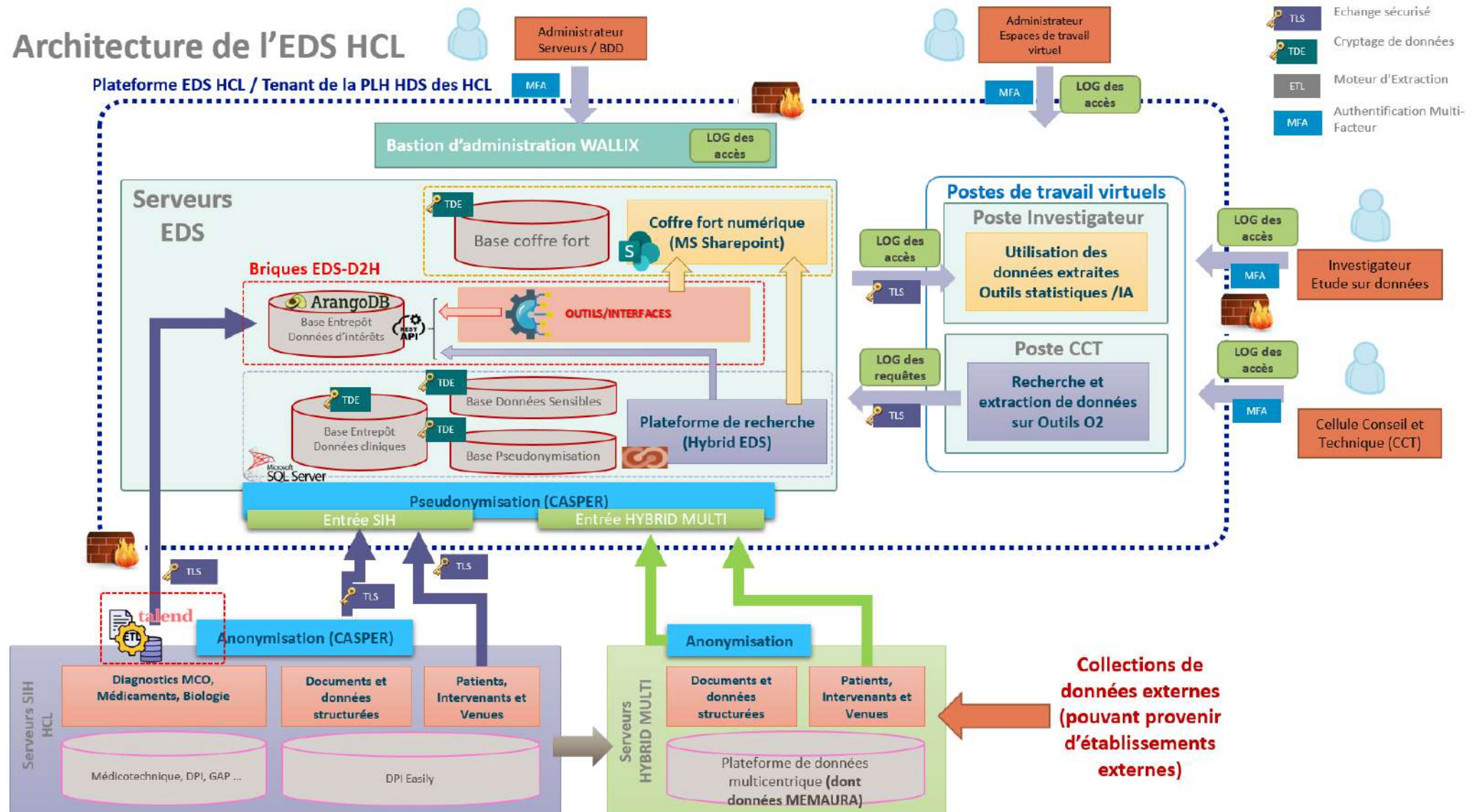
Infrastructures



Monitoring

1. [Khattak et al. \(2023\) – MLHOps: Machine Learning for Healthcare Operations](#)

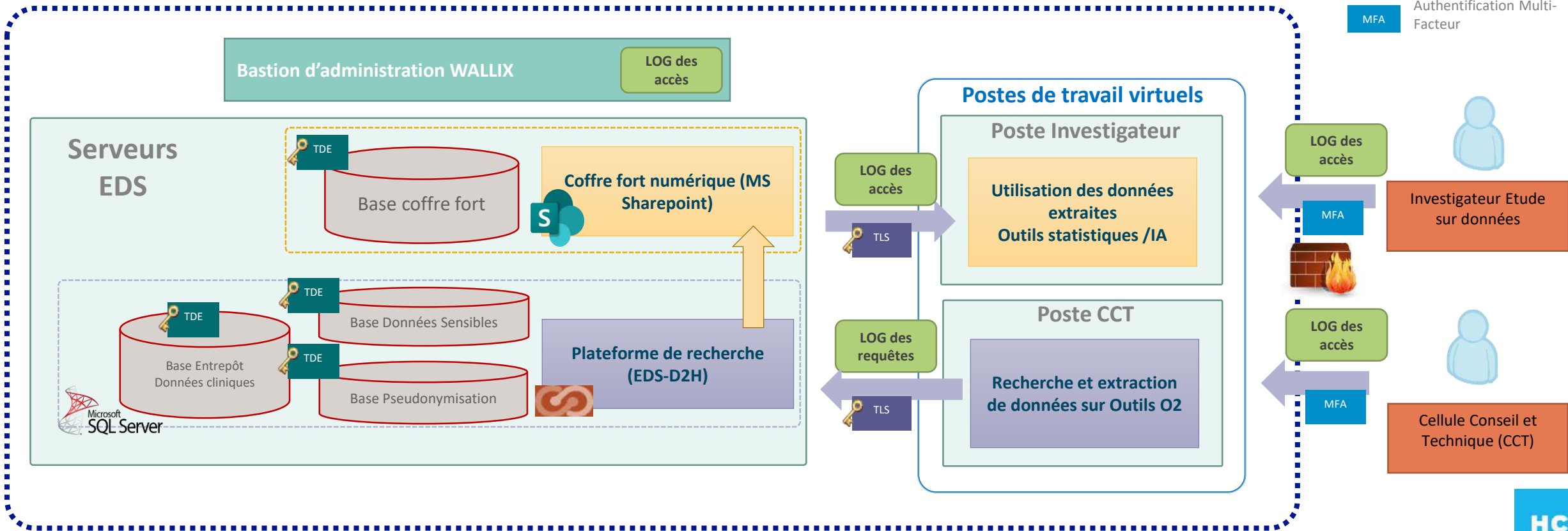
SÉCURISER LES ACCÈS – ENTREPÔT DE DONNÉES DE SANTÉ (EDS) ¹



1. TIC Santé (2024) – « Auvergne-Rhône-Alpes: lancement d'un entrepôt de données de santé commun aux quatre CHU pour près de 10 M€ »

SÉCURISER LES ACCÈS – MACHINE VIRTUELLES (VM)

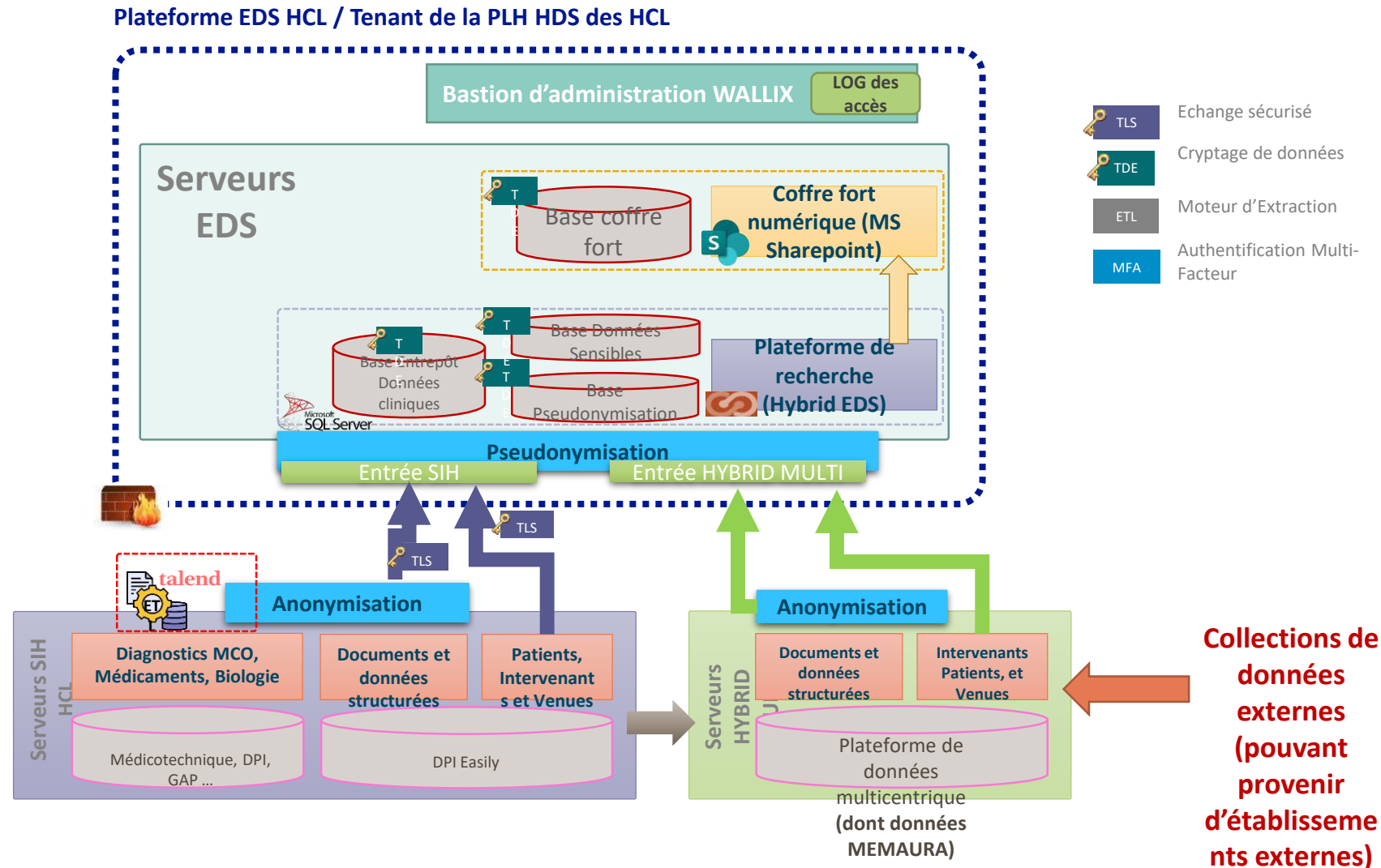
Plateforme EDS HCL / Tenant de la PLH HDS des HCL



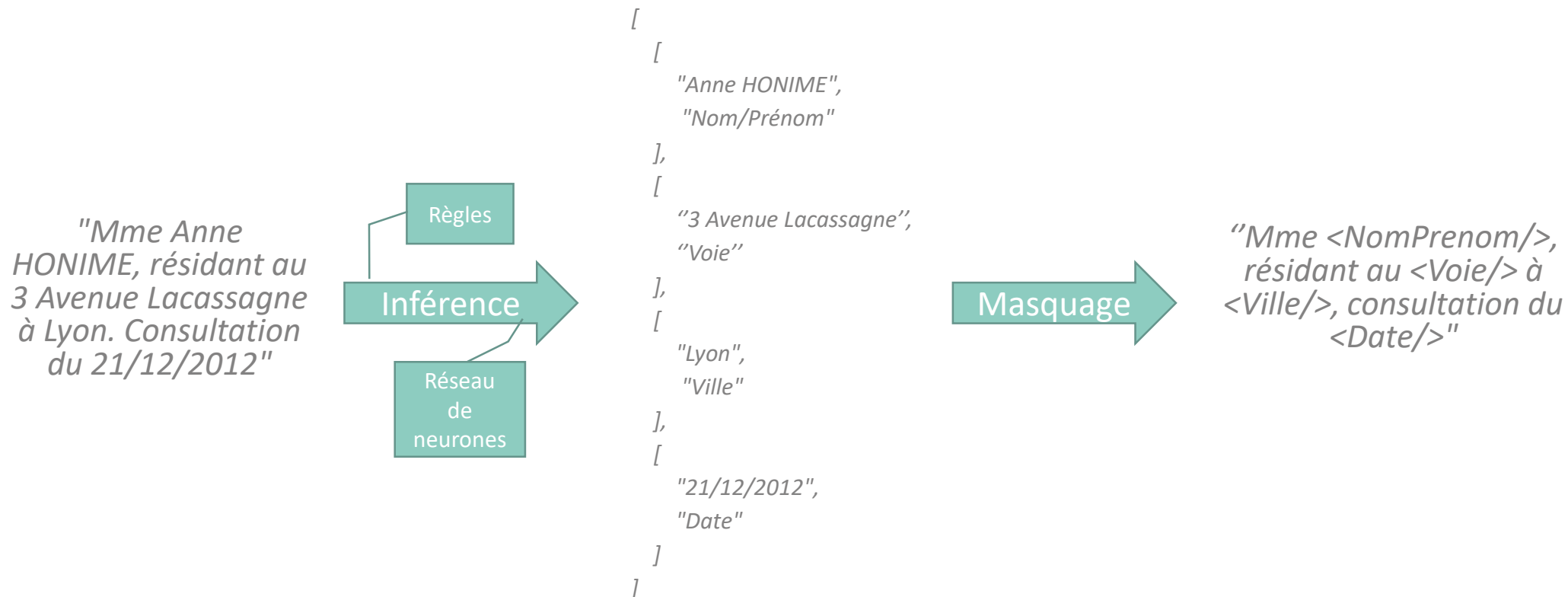
DONNÉES DE SANTÉ

39

MASQUAGE DES DONNÉES AVANT TRAITEMENT



MASQUAGE DES ÉLÉMENTS IDENTIFIANTS ^{1 2}



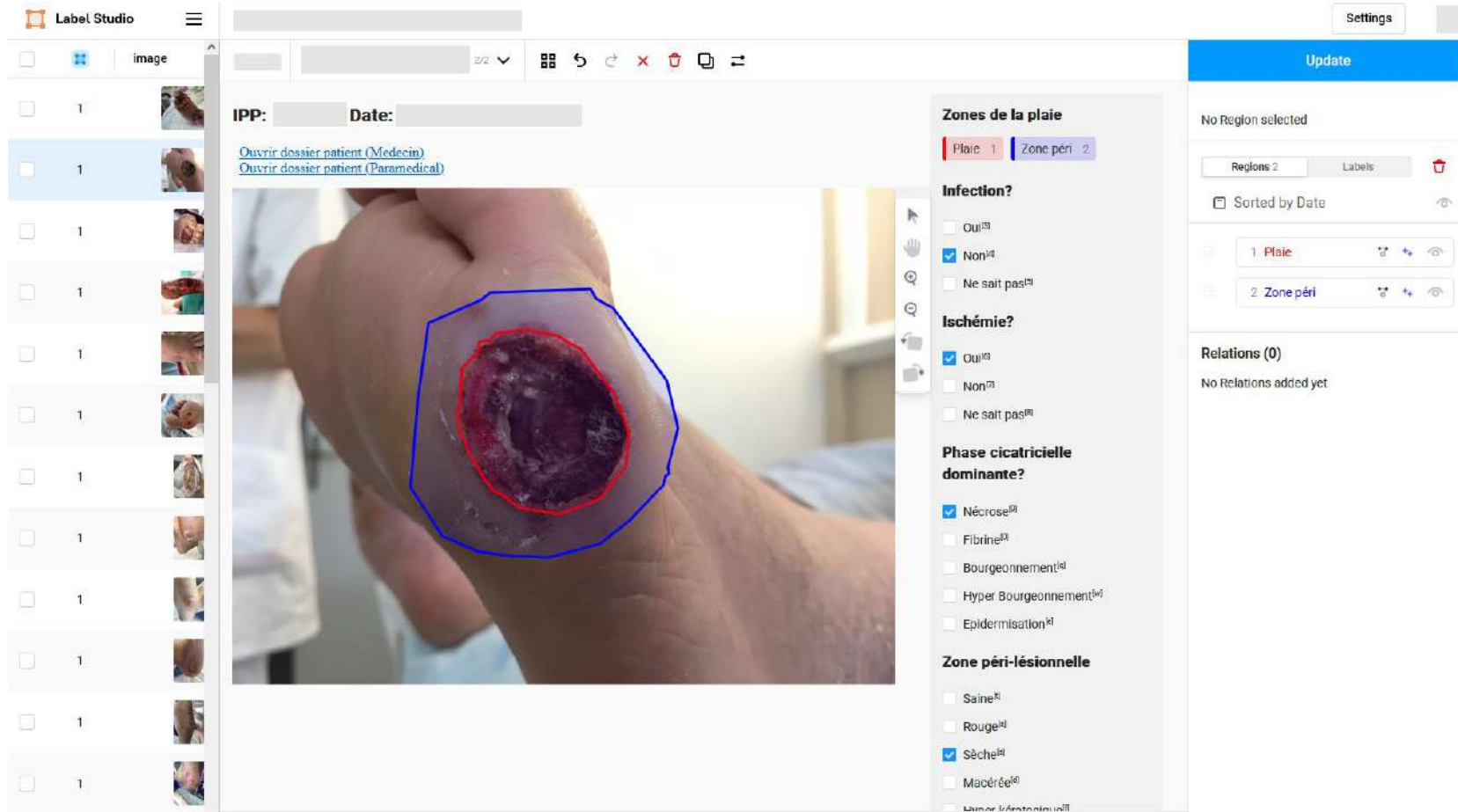
1. [Richard A., Talbot F. and Gimbert D. \(2023\) – « Anonymisation de documents médicaux en texte libre et en français via réseaux de neurones »](#)

2. [Tannier X., Wajsbürt P., Calliger A., et al. \(2023\) – « Development and validation of a natural language processing algorithm to pseudonymize documents in the context of a clinical data warehouse »](#)

DONNÉES DE SANTÉ

ANNOTATION

41

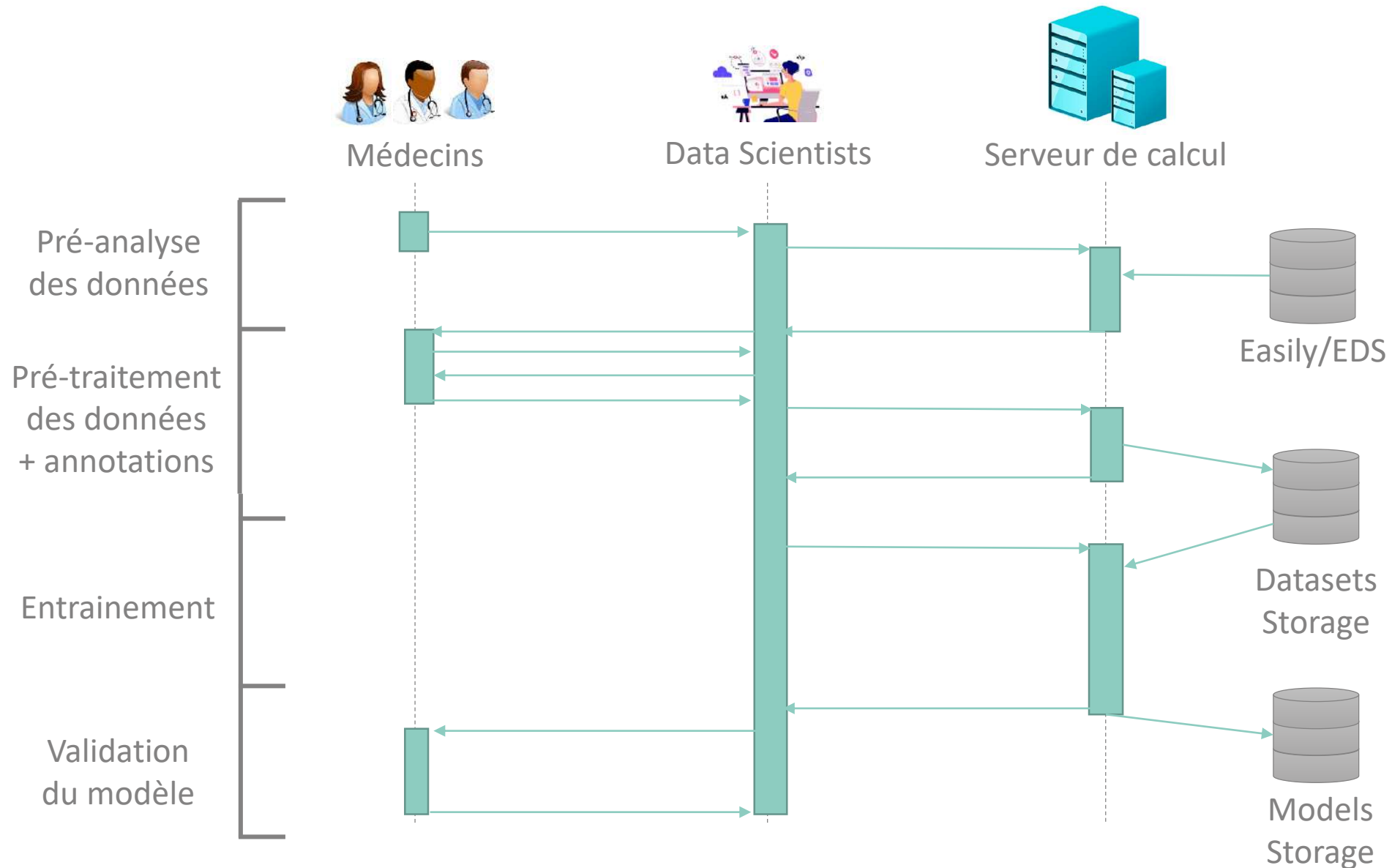


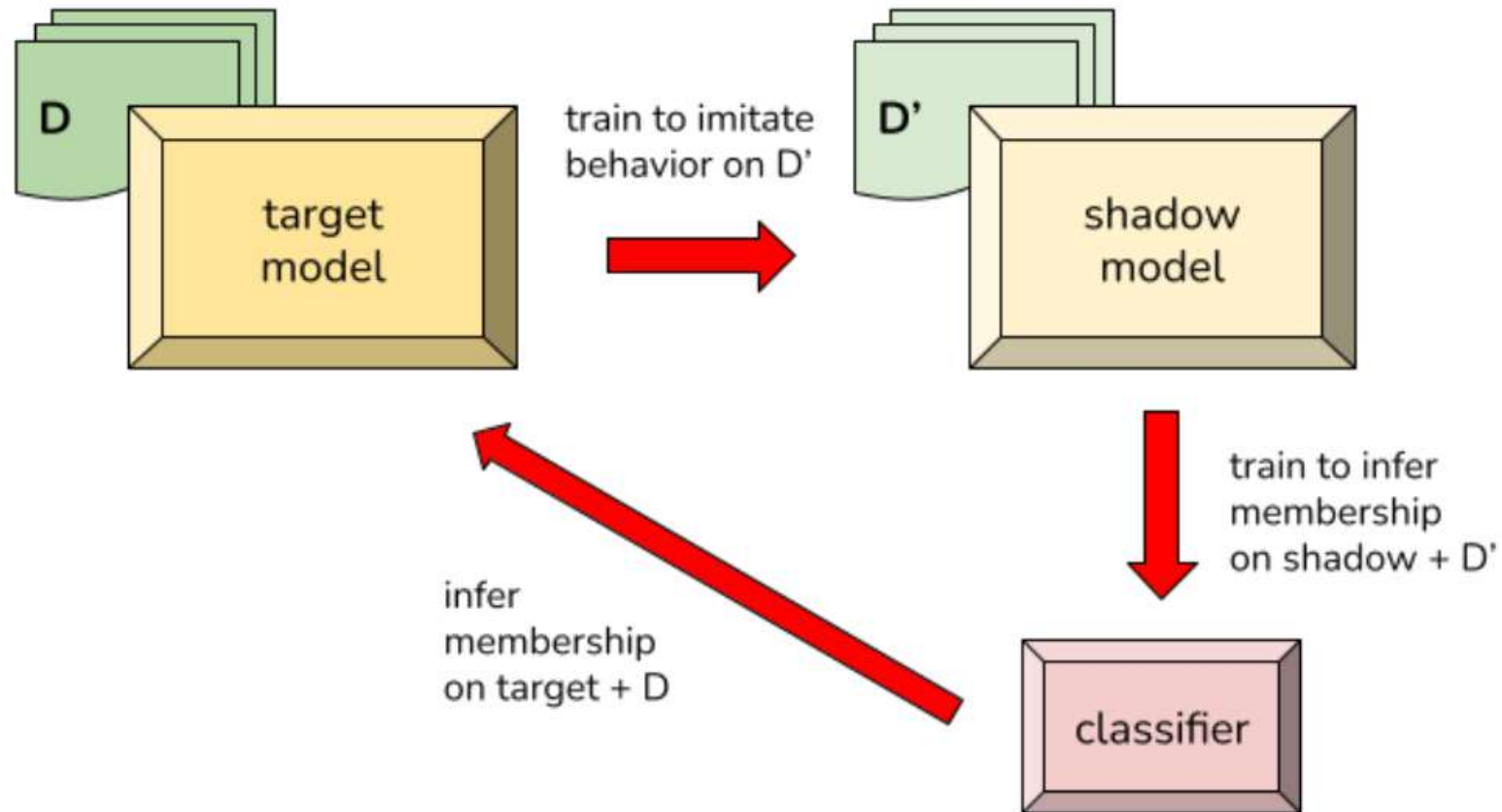
Besoins:

- De personnel soignant (au moins 2 ou 3)
- D'un outil d'annotation « user-friendly »
- D'un protocole d'annotation bien établi

MODÈLES

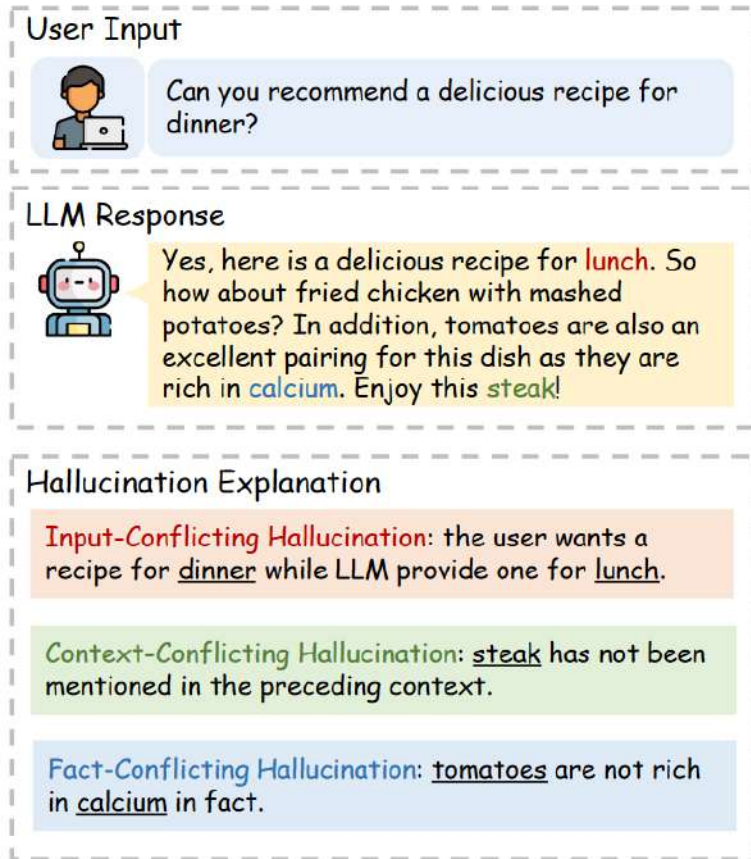
PROCESSUS DE DÉVELOPPEMENT



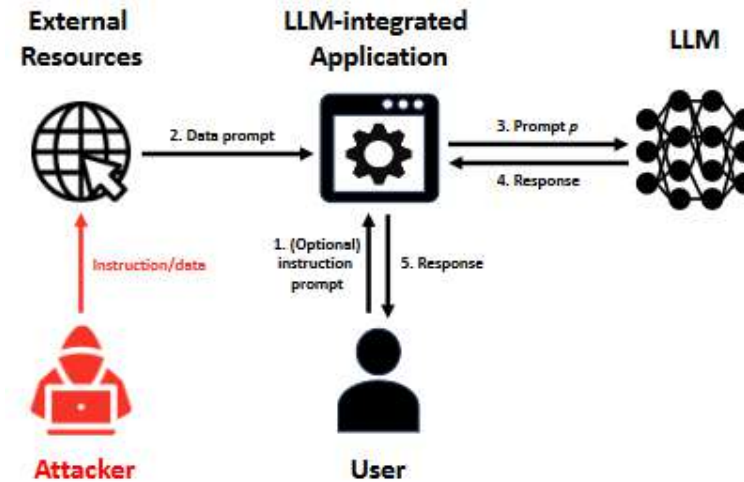


1. [Carlini N., Tramer F., Wallace E., et al. \(2021\) – « Extracting training data from Large Language Models »](#)
2. [Bertheliet G., Boutet A., and Richard A. \(2023\) – « Toward training NLP models to take into account privacy leakages »](#)

FAILLES TECHNIQUES



Hallucinations ^{1 2}



Attaques par injection de prompt ³

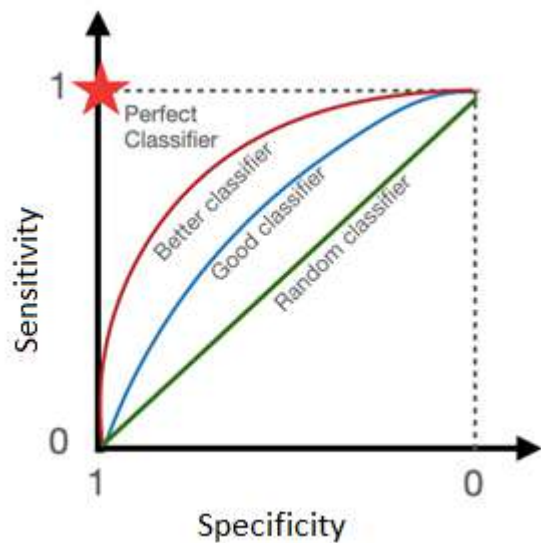
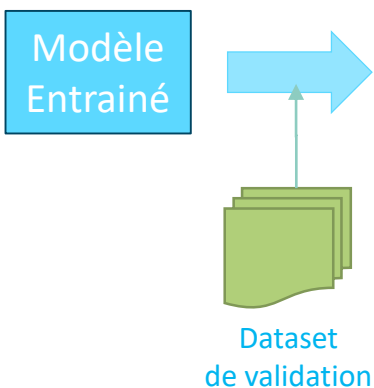


Jailbreaking ⁴

- Huang L., Yu W., Ma W., et al. (2023) – « A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges and Open Questions »
- Zhang Y., Li Y., Cui L., et al. (2023) – « Siren's Song in the AI Ocean: A Survey on Hallucination in Large Language Models »
- Lui Y., Jia Y., Geng R., et al. (2023) – « Prompt Injection Attacks and Defenses in LLM-Integrated Applications »
- Wei A., Haghtalab N., and Steinhardt J. (2023) – « Jailbroken: How Does LLM Safety Training Fail? »

SEUILS DE VALIDATION

Sensibilité	Spécificité	
0.55	0.55	✗
0.98	0.55	✗
0.72	0.98	✗
0.97	0.98	✓?
1.0	1.0	Modèle « parfait »



Comment déterminer si un modèle est valide?

- Définir un seuil minimum pour chaque métrique
- Dépend de la pré-valence, comme tout test médical
- À définir au cas par cas lors de la mise en place du projet

ENTRAINEMENTS ET INFÉRENCES

BESOIN D'INFRASTRUCTURES DÉDIÉES

Pour les entraînements:



Serveurs
de calcul

- Fermes de serveurs avec des GPU
- Mutualisation des ressources
- Ordonnancements des calculs

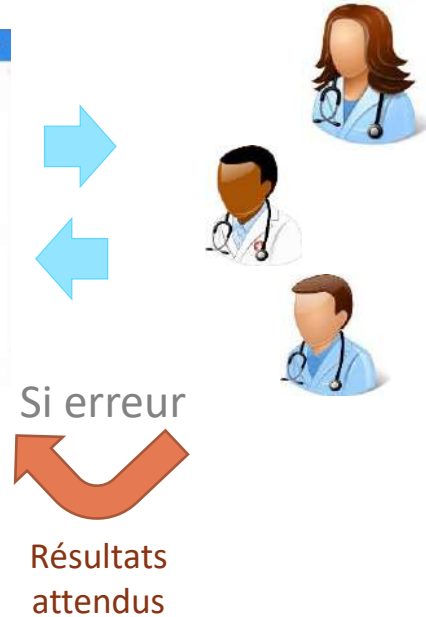
Pour l'inférence:



Serveurs
de
Développement
d'Intégration
et Production

- Avec des GPU
- Beaucoup de mémoire pour charger plusieurs modèles en parallèle
- Capacité de traiter plusieurs appels en parallèle

INCLURE LES RETOURS UTILISATEURS ¹

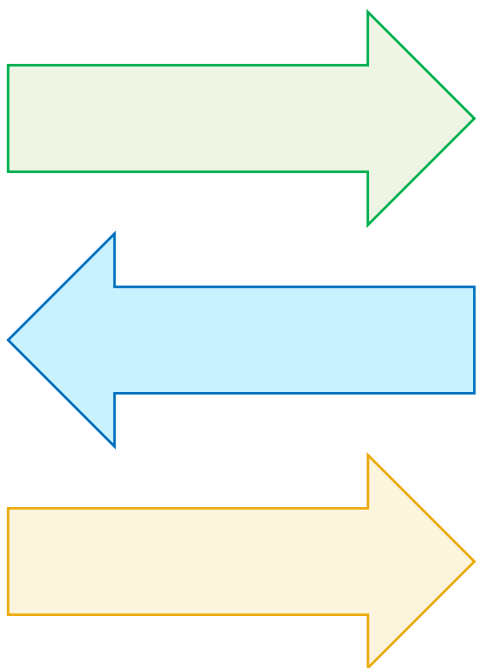


Besoins:

- Inclure des fonctionnalités de « feedback » dans les interfaces
- Former les soignants à détecter les erreurs des modèles ²
- Anticiper la charge de travail nécessaire

1. [Henry K.E., Kornfield R., Sridharan A., et al. \(2022\) - « Human-machine teaming is key to {AI} adoption: clinicians' experiences with a deployed machine learning system »](#)
2. [Tsai T., Fridsma D., and Gatti G. \(2003\) – « Computer Decision Support as a Source of Interpretation Error: The Case of Electrocardiograms »](#)

SYNTHÈSE



Le MLOps:

- Adapte le DevOps au développement de fonctionnalités ML
- Permet d'industrialiser efficacement ces fonctionnalités
- Doit s'articuler avec un processus de DevOps classique

Le MLOps en Santé:

- Nécessite de mettre en place de protocoles et des environnements sécurisant la vie privée des patients
- Nécessite d'inclure les soignants dans le développement et la surveillance des outils basés sur du ML ¹
- Nécessite de former *a minima* les soignants sur le ML et sur l'utilisation d'outils basé sur du ML
- Nécessite d'inclure le développement d'outils basé sur du ML dans des processus organisationnels plus globaux ²



Quid de l'impact sur les patients, les soignants et les parcours de soin ?

1. [Henry K.E., Kornfield R., Sridharan A., et al. \(2022\) - « Human-machine teaming is key to {AI} adoption: clinicians' experiences with a deployed machine learning system »](#)
2. [Kim J.Y., Boag W., Gulamali F., et al. \(2023\) – « Organizational Governance of Emerging Technologies: AI Adoption in Healthcare »](#)

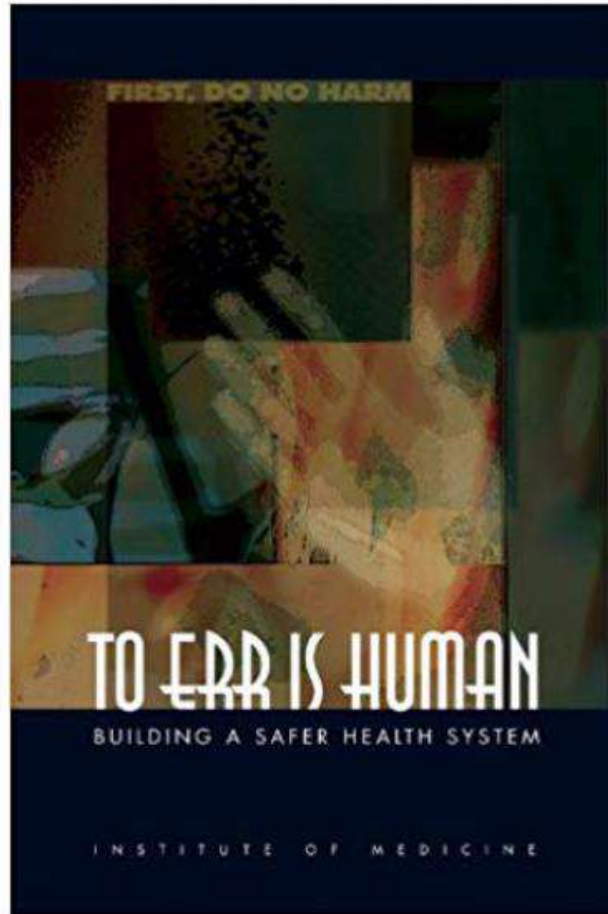
INTÉGRER LE ML EN SANTÉ

RISQUES POSSIBLES

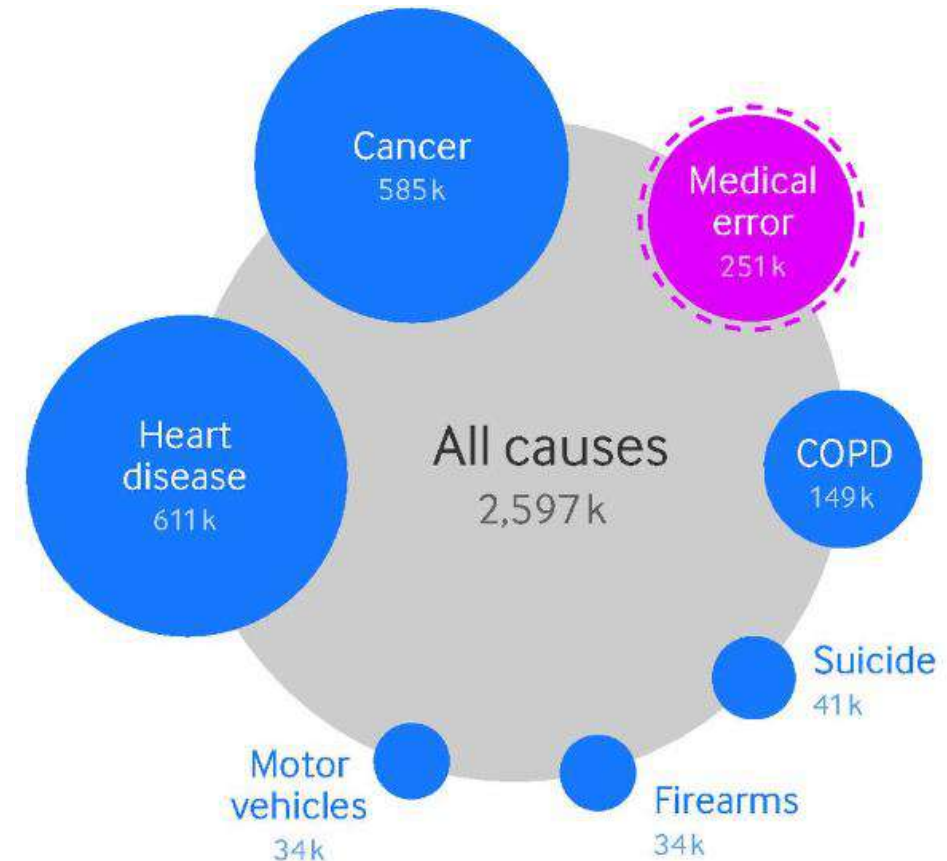
POURQUOI DÉVELOPPER DES SIH ?

50

RÉDUIRE LE RISQUE D'ERREURS MÉDICALES



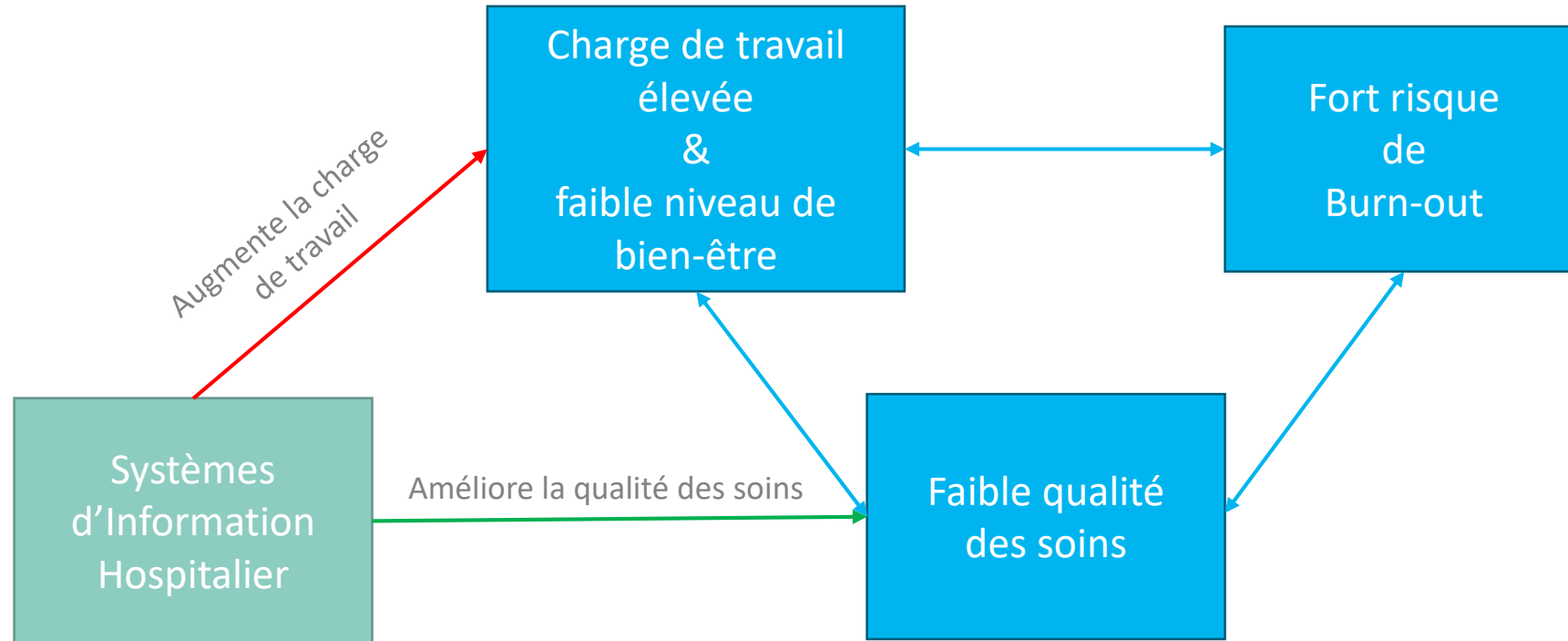
Entre 44k et 98k mort aux USA en 1997 ¹



La troisième principale cause de décès aux USA en 2013 ²

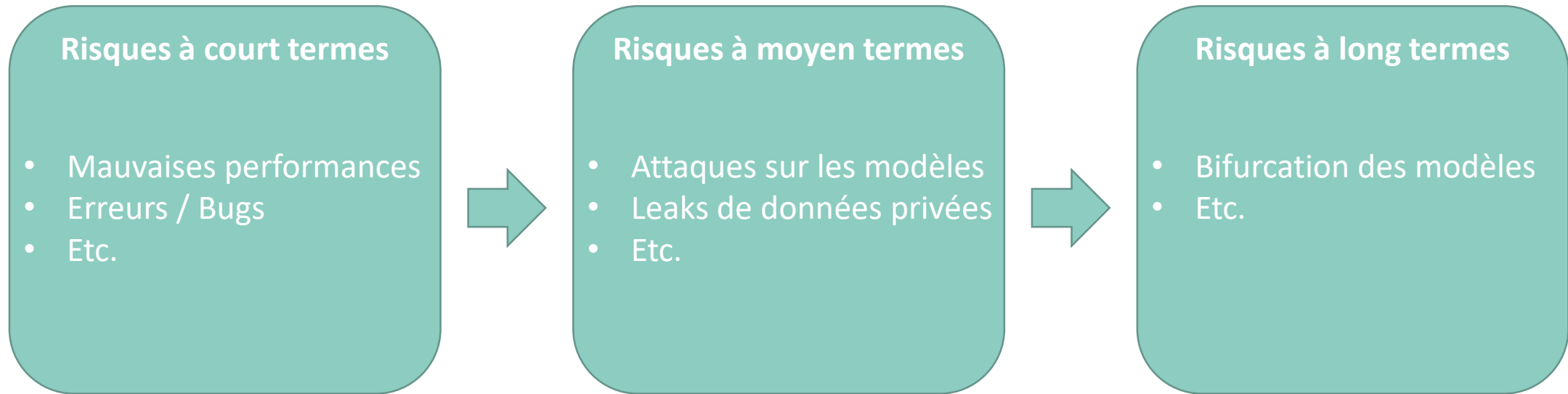
1. [Donaldson et al. \(2000\) – To err is human: building a safer health system](#)
2. [Makary and Daniel \(2016\) – Medical error : the third leading cause of death in the US](#)

CHARGE DE TRAVAIL ET QUALITÉ DES SOINS: UN CERCLE VICIEUX^{1 2 3 4}



1. [Hall et al. \(2016\) – Healthcare Staff Wellbeing, Burnout, and Patient Safety: A Systematic Review](#)
2. [Tawfik et al. \(2018\) – Physician Burnout, Well-being, and Work Unit Safety Grades in Relationship to Reported Medical Errors](#)
3. [West, Dybrye and Shanafelt \(2018\) – Physician burnout: contributors, consequences and solutions](#)
4. [Dutheil et al. \(2019\) – Suicide among physicians and health-care workers: A systematic review and meta-analysis](#)

DES RISQUES À PLUS OU MOINS LONG TERMES ^{1 2}



1. [Tan S., Taeihagh A., and Baxter K. \(2022\) – « The Risks of Machine Learning Systems »](#)

2. [Habehh H. and Gohel S. \(2021\) – « Machine Learning in Healthcare »](#)

REPRODUCTION DE COMPORTEMENTS DISCRIMINANTS^{1 2 3}

Probabilité de déclarer une expérience de discrimination dans le système de santé

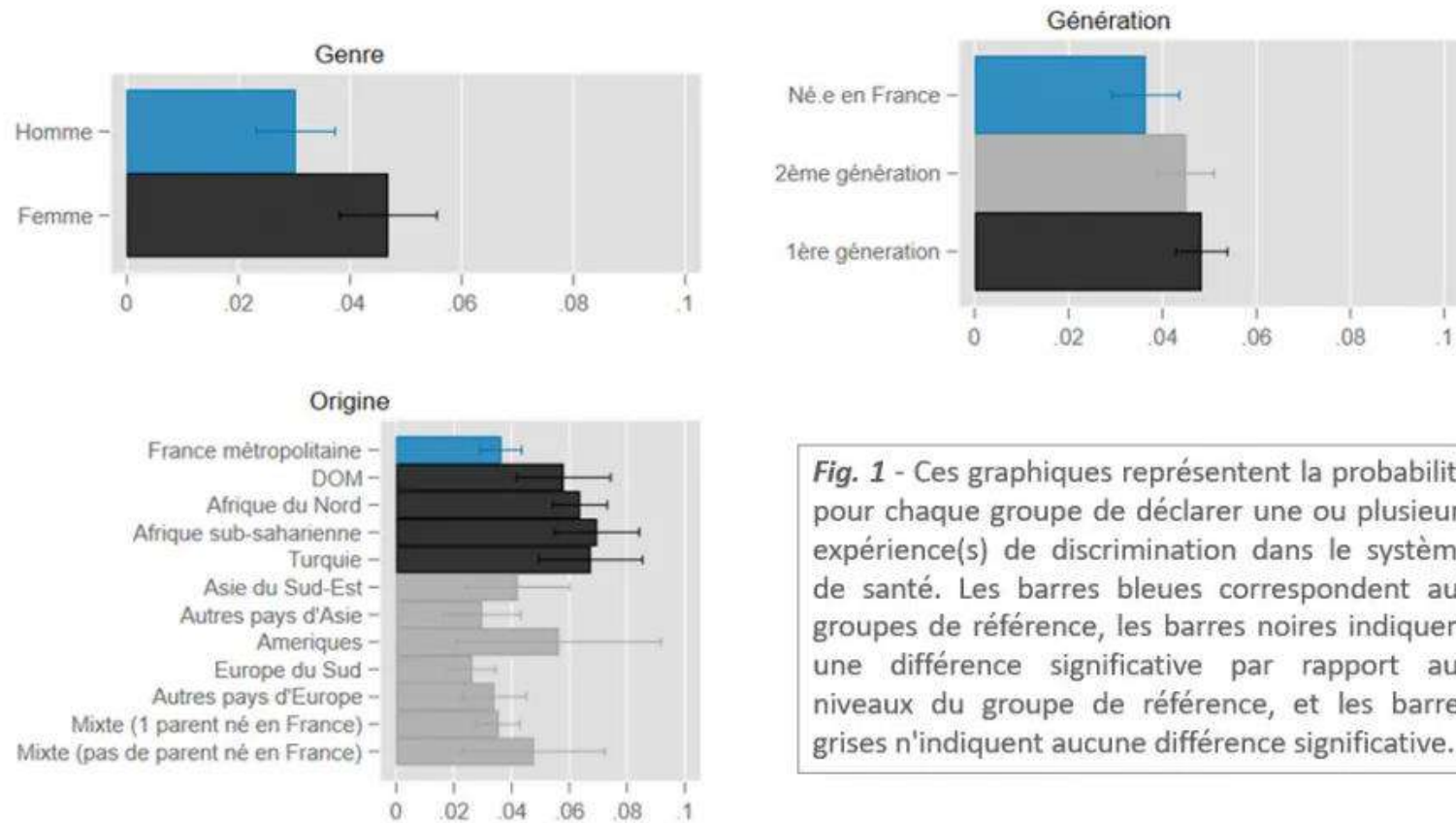


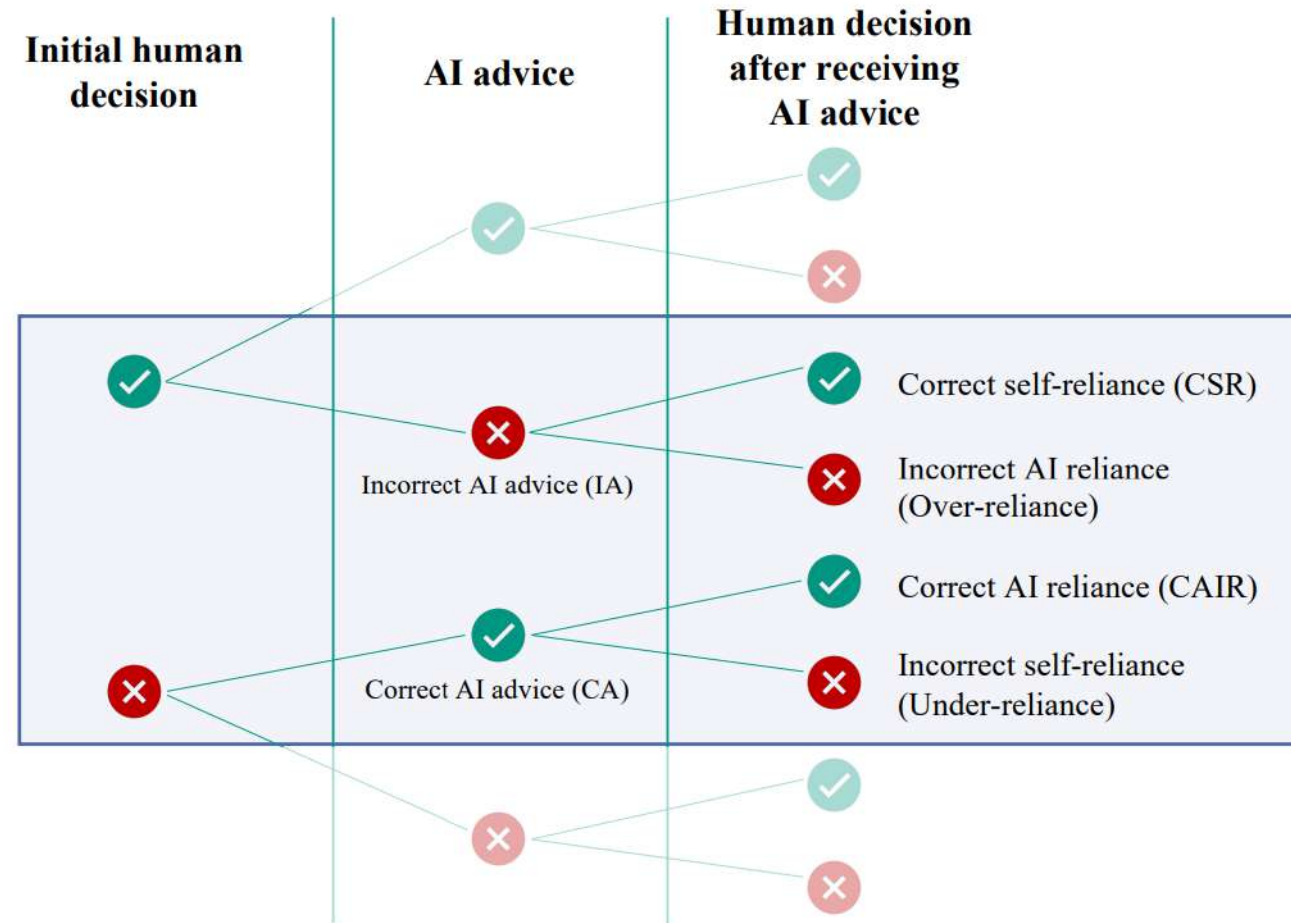
Fig. 1 - Ces graphiques représentent la probabilité pour chaque groupe de déclarer une ou plusieurs expérience(s) de discrimination dans le système de santé. Les barres bleues correspondent aux groupes de référence, les barres noires indiquent une différence significative par rapport aux niveaux du groupe de référence, et les barres grises n'indiquent aucune différence significative.

1. [Ined \(2020\) – « Les discriminations dans le système de santé français: un obstacle à l'accès aux soins »](#)
2. [Rivenbark J. G. and Ichou M. \(2020\) – « Discrimination in healthcare as a barrier to care: experiences of socially disadvantaged populations in France from a nationally representative survey »](#)
3. [Borgesius F. Z. \(2018\) – « Discrimination, artificial intelligence, and algorithmic decision-making »](#)
4. [Wang Q., Xu Z., Chen Z., et al. \(2021\) – « Visual Analysis of Discriminating in Machine Learning »](#)

RISQUES À MOYEN TERMES

54

AUTOMATISATION, PERTE DE SAVOIR-FAIRE ET DÉPENDANCE AUX OUTILS ^{1 2 3 4 5 6}

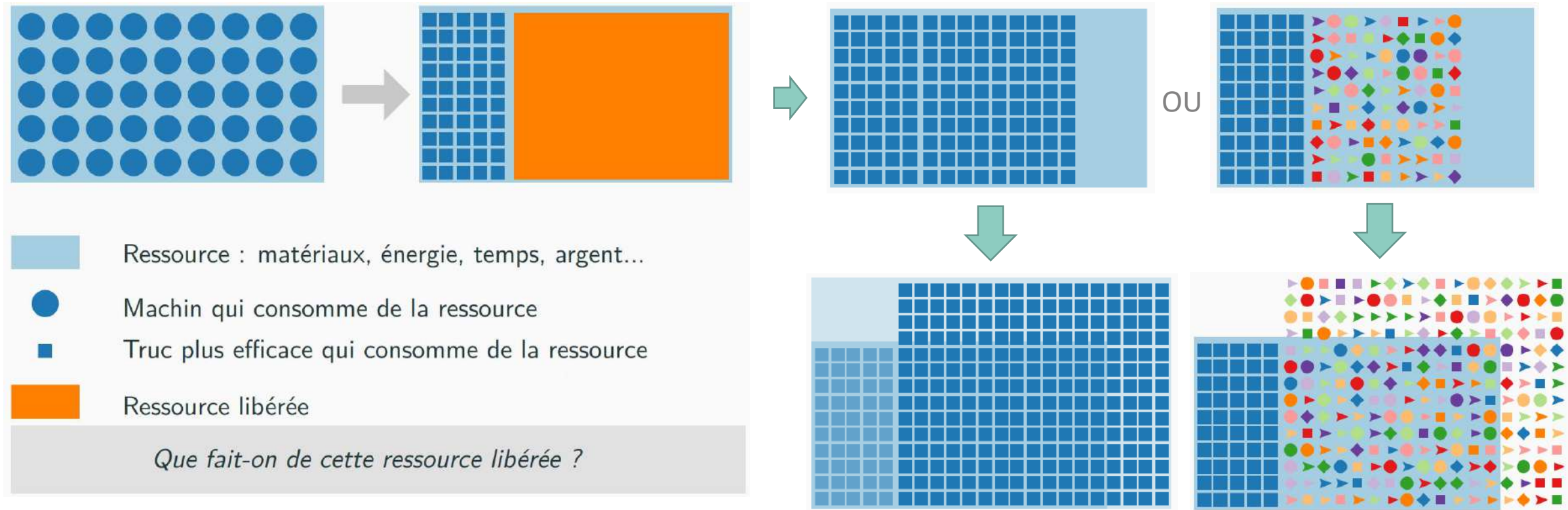


1. [Parasuraman R. and Manzey D. H. \(2010\) – « Complacency and Bias in Human Use of Automation: An Attentional Integration »](#)
2. [He G., Kuiper L., and Gadiraju U. \(2023\) – « Knowing About Knowing: An Illusion of Human Competence Can Hinder Appropriate Reliance on AI Systems »](#)
3. [Grissinger M. \(2019\) – « Understanding Human Over-Reliance On Technology »](#)
4. [Tsai, Fridsma and Gatti \(2003\) – « Computer decision support as a source of interpretation error: the case of electrocardiograms »](#)
5. [Povyakalo et al. \(2013\) – « How to discriminate between Computer-Aided and Computer-Hindered Decisions: A Case study in Mammography »](#)
6. [Schemmer M., Kuehl N., Benz C., et al. \(2023\) – « Appropriate Reliance on AI Advice: Conceptualization and the Effect of Explanations »](#)

RISQUES À LONG TERMES

55

« EFFETS REBOND » NÉGATIFS ^{1 2 3 4 5}

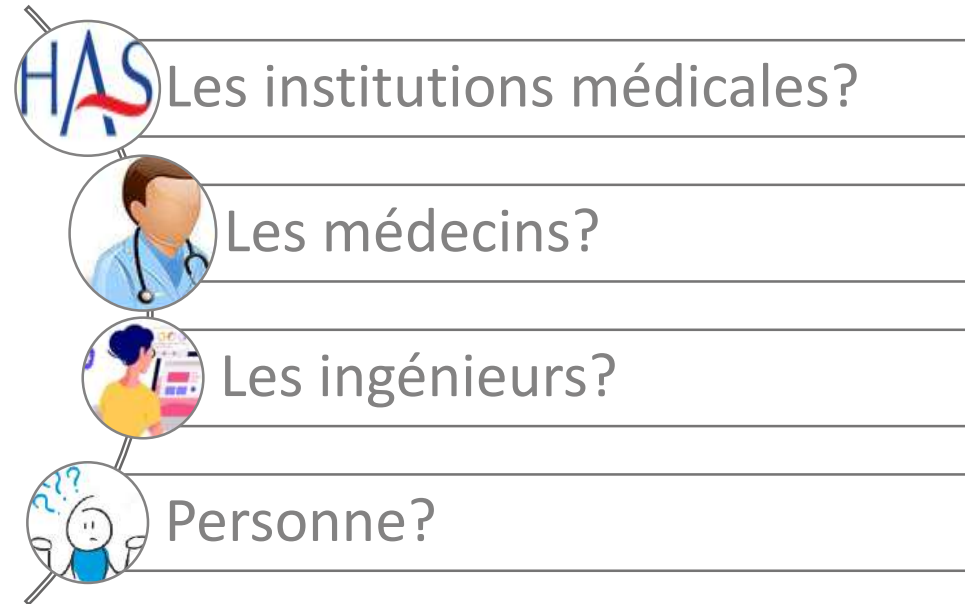


Retour de flamme !

1. [Berkhout P.H.G., Muskens J. C., and Velthuisen J. W. \(2000\) – « Defining the rebound effect »](#)
2. [Willenbacher M., Hornauer T., and Wohlgemuth V. \(2021\) – « Rebound Effects in Methods of Artificial Intelligence »](#)
3. [Ertel W. \(2019\) – « Artificial Intelligence, the spare time rebound effect and how the ECG would avoid it »](#)
4. [Bertillot \(2016\) – « Comment l'évaluation de la qualité transforme l'hôpital. Les deux visages de la rationalisation par les indicateurs »](#)
5. [Sylvain Bouveret \(2023\) – « Numérique : l'insoutenable matérialité du virtuel »](#)

PROBLÈMES DE RESPONSABILITÉ

Si un médecin utilise un SIH
basé sur de l'IA, et que
l'utilisation de ce SIH conduit à
une erreur médicale, qui est
responsable ?



Socialement, il y a une pression envers les médecins ¹



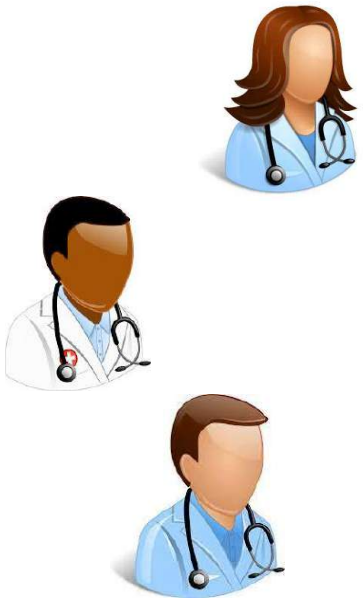
Légalement, les institutions sont tenues responsables et des
normes sont à prendre en comptes par les ingénieurs ^{2 3}

1. [Itani, Lecron and Fortemps \(2019\) – Specifics of medical data mining for diagnosis aid: A survey](#)
2. [Norme ISO 13485:2016 – Dispositifs médicaux – Systèmes de management de la qualité – Exigences à des fins réglementaires](#)
3. [Norme ISO 62304:2006 – Logiciels de dispositifs médicaux – Processus du cycle de vie du logiciel](#)

COMMENT FAIRE CONFIANCE AU ML ?

57

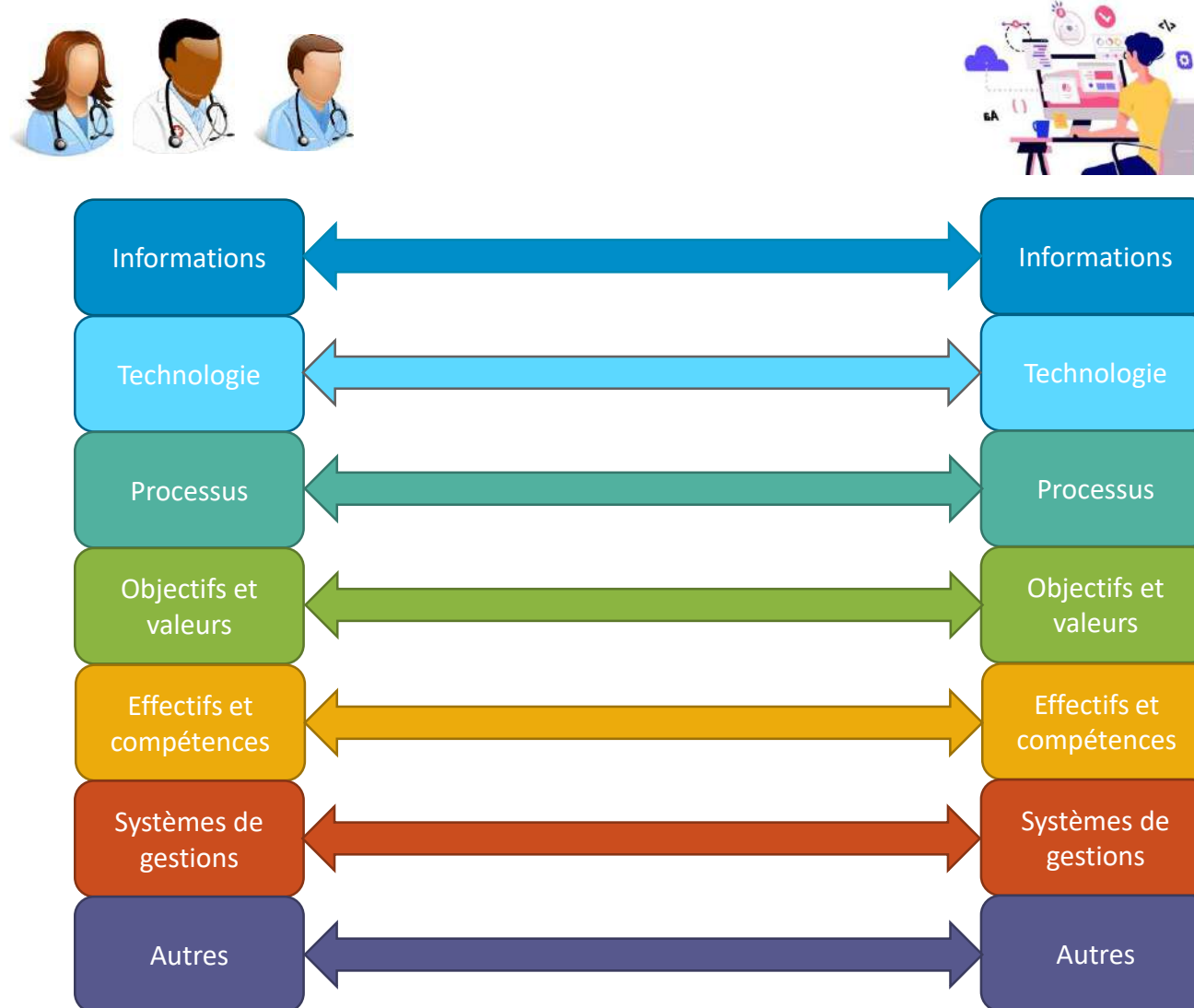
TRANSPARENCE ET EXPLICABILITÉ ^{1 2 3 4}



- Pourquoi ce résultat et pas un autre ?
- Quel degrés de confiance mettre dans l'outil et ses résultats ?
- Dans quel situation l'outil est bon et quand est-il mauvais ?
- Comment corriger une erreur de l'outil ?

1. [Gunning and Aha \(2019\) – DARPA's Explainable Artificial Intelligence \(XAI\) Program](#)
2. [Berredo-Arrieta *et al.* \(2020\) - Explainable Artificial Intelligence \(XAI\): Concepts, taxonomies, opportunities and challenges toward responsible AI](#)
3. [Mueller *et al.* \(2019\) - Explanation in Humain-AI Systems: A Literature Meta-Review, Synopsis of Key Ideas and Publications, and Bibliography for Explainable AI](#)
4. [Richard *et al.* \(2020\) – Transparency of Classification Systems for Clinical Decision Support](#)

RAISONS D'ÉCHECS: ÉCARTS CONCEPTION-RÉALITÉ ^{1 2}



1. [Heeks \(2006\) – Health Information Systems:: Failure, success and improvisation](#)
2. [Masiero \(2016\) – The Origins of Failure: Seeking the Causes of Design-Reality Gaps](#)

PRÉVENIR LES RISQUES

DANS LE DÉVELOPPEMENT DE SIH BASÉS SUR DU ML

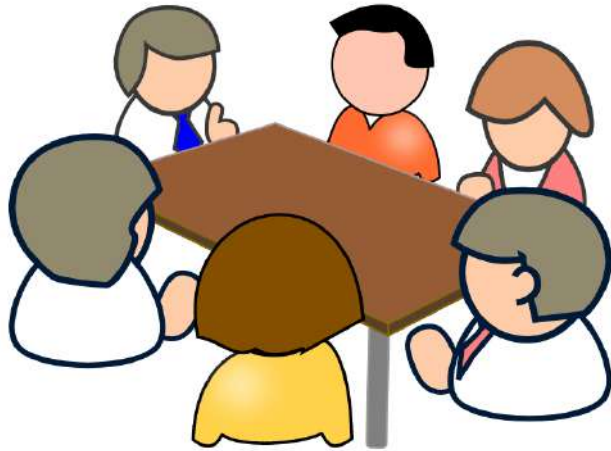
RETOURS TERRAINS



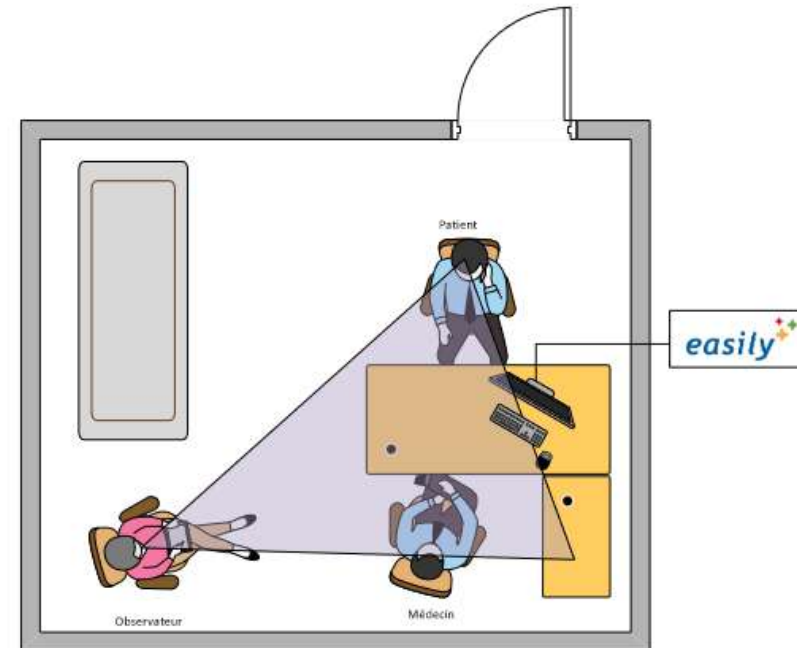
- Problématiques cliniques:
 - Améliorer un processus
 - Développer un test basé sur des données
 - Etc.
- Intégrer un outil déjà développé:
 - Par une entreprise/start-up
 - Par un médecin et/ou un interne
 - Par nous pour un autre service
- Curiosité/intérêt scientifique

IDENTIFIER LES PROBLÈMES ¹

Interviews

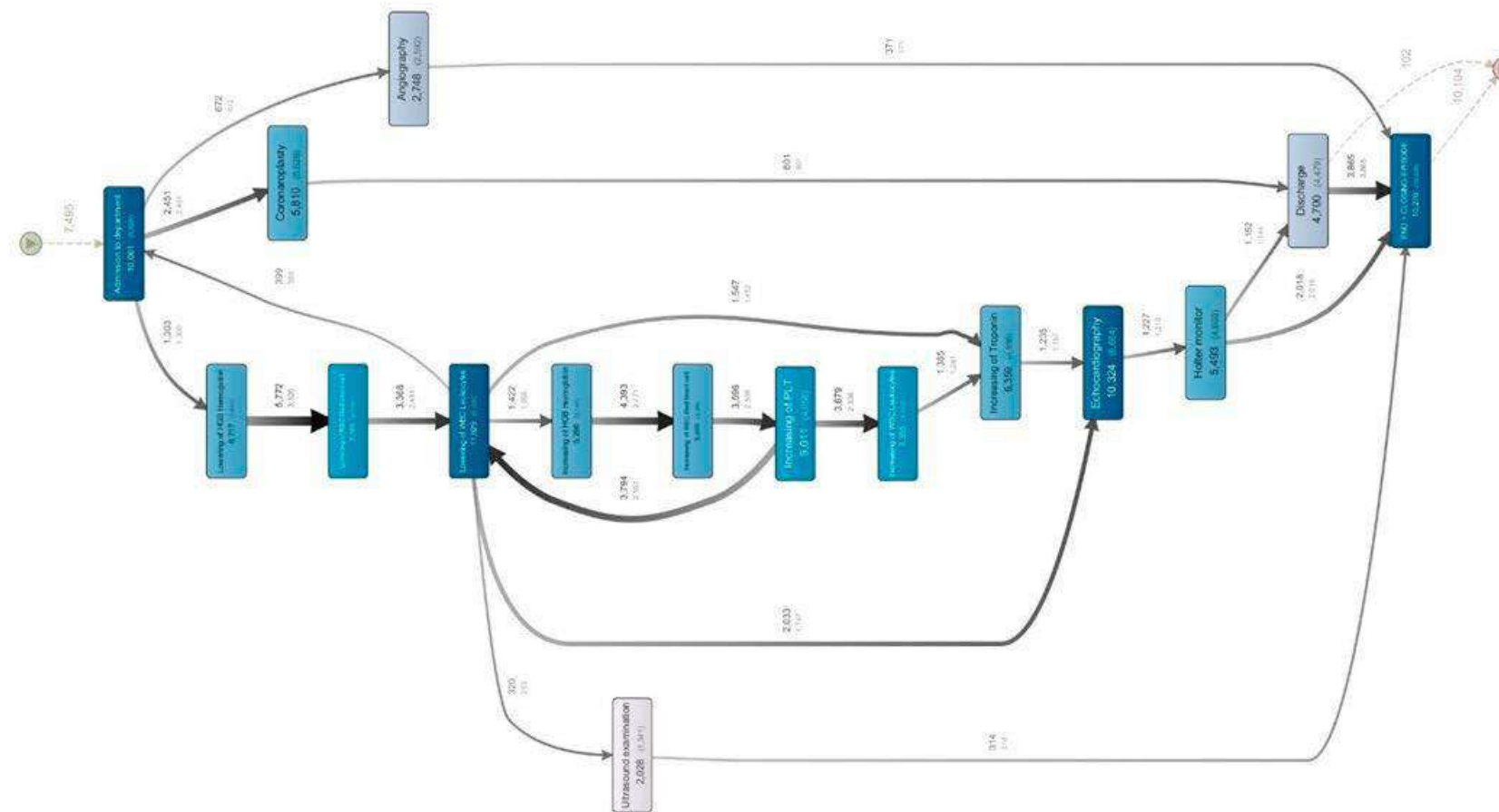


Analyses terrain ²



1. [Kim J.Y., Boag W., Gulamali F., et al. \(2023\) – « Organizational Governance of Emerging Technologies: AI Adoption in Healthcare »](#)
2. [Richard \(2021\) - Proposition d'un outil d'aide à la décision adapté aux contraintes et aux enjeux d'un soutien informatique aux consultations médicales coutumières](#)

ANALYSES NUMÉRIQUES^{1 2}



1. [Van der Aalst W., Adriansyah A., Alves de Medeiros A. K., et al. \(2012\) – « Process Mining Manifesto »](#)
2. [Rojas E., Munoz-Gama J., Sepúlveda M., and Capurro D. \(2016\) – « Process Mining in Healthcare: A Literature Review »](#)
3. [Metsker O., Yakovlev A., Bolgova E., et al. \(2018\) – « Identification of Pathophysiological Subclinical Variances During Complex Treatment Process of Cardiovascular Patients »](#)

DÉTERMINER L'APPROCHE LA PLUS ADAPTÉE POUR L'AIDE À LA DÉCISION ^{1 2}

Contrainte principale

Les décisions doivent se **conformer** à des directives non révocables venant de hautes autorités



Approche adaptée

Conformiste:

L'outil doit aider l'utilisateur à être **conforme** à ces directives



Objectiviste:

L'outil doit se baser sur ces faits et théories **objectivement** pertinents



Ajustive:

L'outil doit **s'ajuster** aux besoins des soignants et ne pas interférer avec leur processus de travail ou leur capacité d'initiative

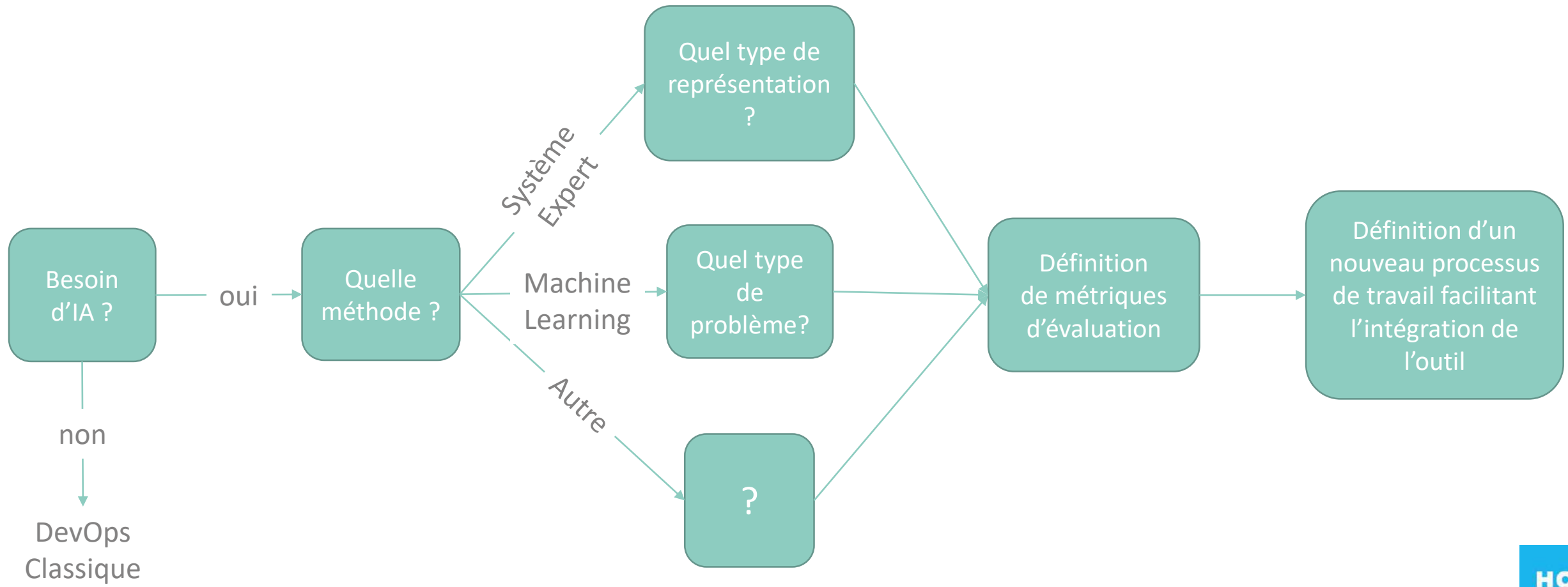
1. [Meinard and Tsoukias \(2019\) – On the rationality of decision aiding processes](#)

2. [Richard \(2021\) - Proposition d'un outil d'aide à la décision adapté aux contraintes et aux enjeux d'un soutien informatique aux consultations médicales coutumières](#)

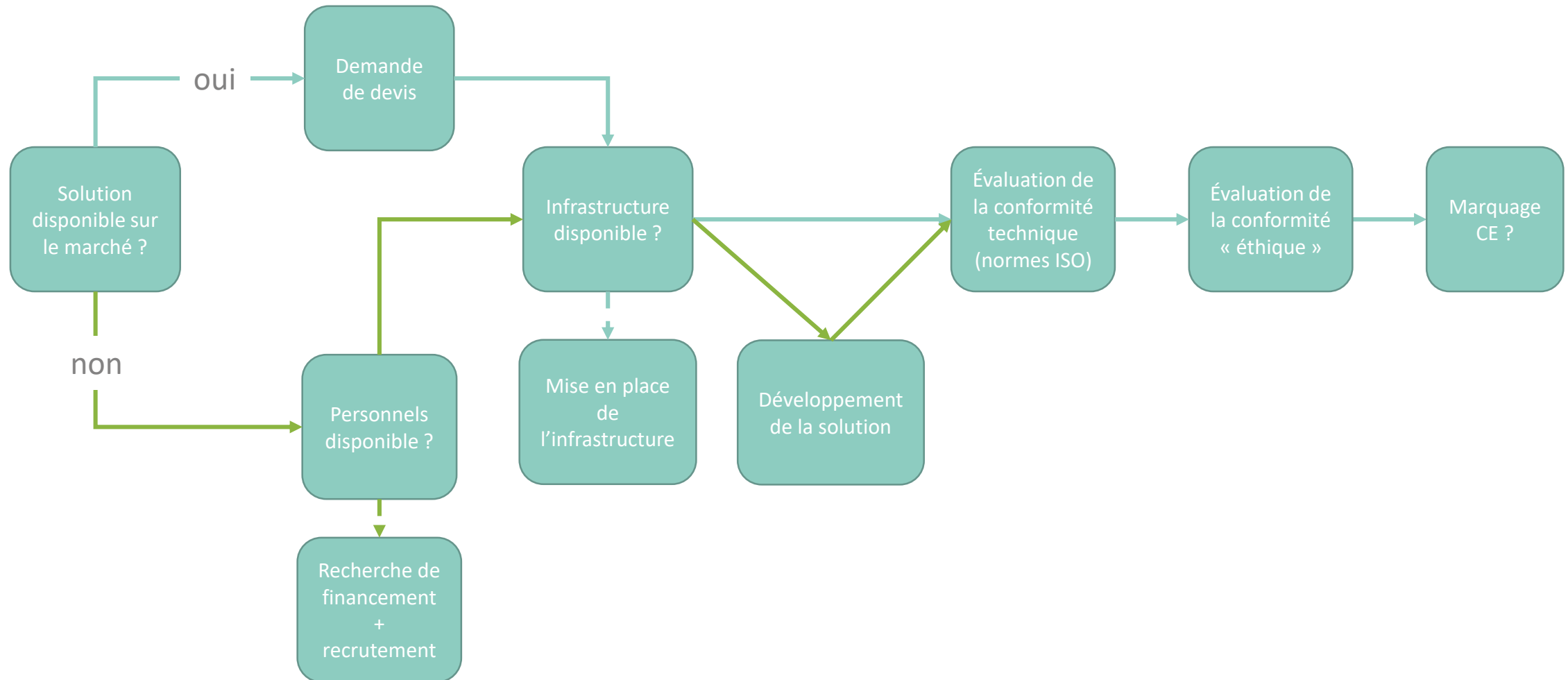
FORMALISER LE BESOIN

64

DÉTERMINER LE BESOIN EN IA/ML/DL ¹



IDENTIFIER LES RESSOURCES DISPONIBLES ¹



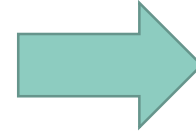
NORMES ISO



ISO 13485:

Quality management systems & Requirements for regulatory purposes

<https://www.iso.org/standard/59752.html>



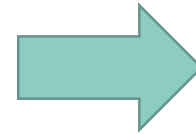
Applicable à l'IA ? ^{1 2 3}



ISO 62304:

Medical device software & Software life cycle processes

<https://www.iso.org/standard/38421.html>



Publiées:

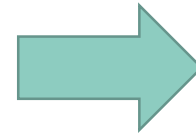
- [ISO 24029](#): Assessment of the robustness of neural networks



ISO 14971:

Application of risk management to medical devices

<https://www.iso.org/standard/72704.html>



En cours de développement:

- [ISO 18988](#): Application of AI technologies in health informatics
- [ISO 5259](#): Data quality for analytics and machine learning (ML)



ISO 62366:

Application of usability engineering to medical devices

<https://www.iso.org/standard/63179.html>

1. [O'Sullivan et al. \(2018\) – Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence \(AI\) and autonomous robotic surgery](#)
2. [Zhao \(2019\) – Improving Social Responsibility of Artificial Intelligence by Using ISO 2600](#)
3. [Natale \(2022\) – Extensions of ISO/IEC 25000 Quality Models to the Context of Artificial Intelligence](#)

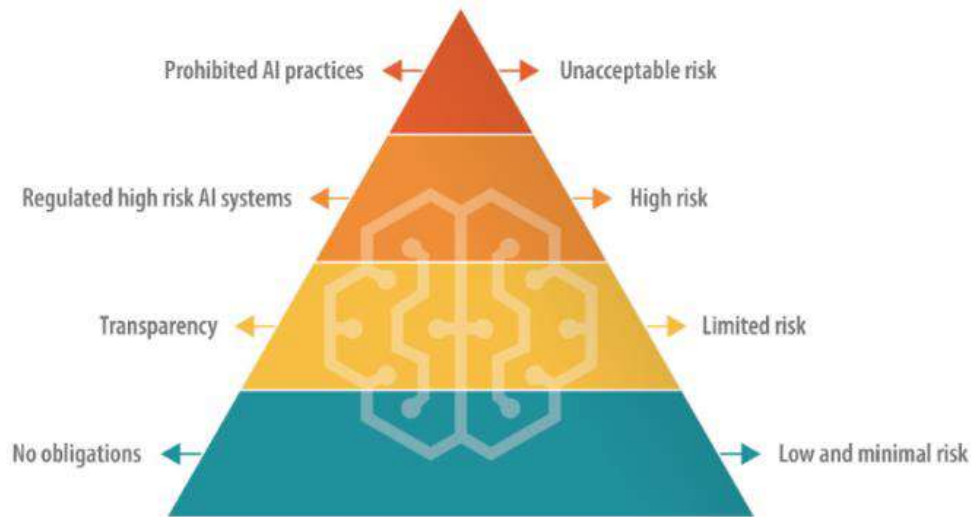
CONFORMITÉS « ÉTHIQUES »

67

PRINCIPES GÉNÉRAUX ^{1 2 3 4}

Le Serment
Holberton-turing ³

CNIL.



- Principe de Loyauté
- Principe de Vigilance/Réflexivité
- Principe d'Autonomie
- Principe de Justice
- Principe de Transparence

1. <https://www.cnil.fr/en/algorithms-and-artificial-intelligence-cnils-report-ethical-issues>

2. <https://www.cnil.fr/en/ai-systems-compliance-other-guides-tools-and-best-practices>

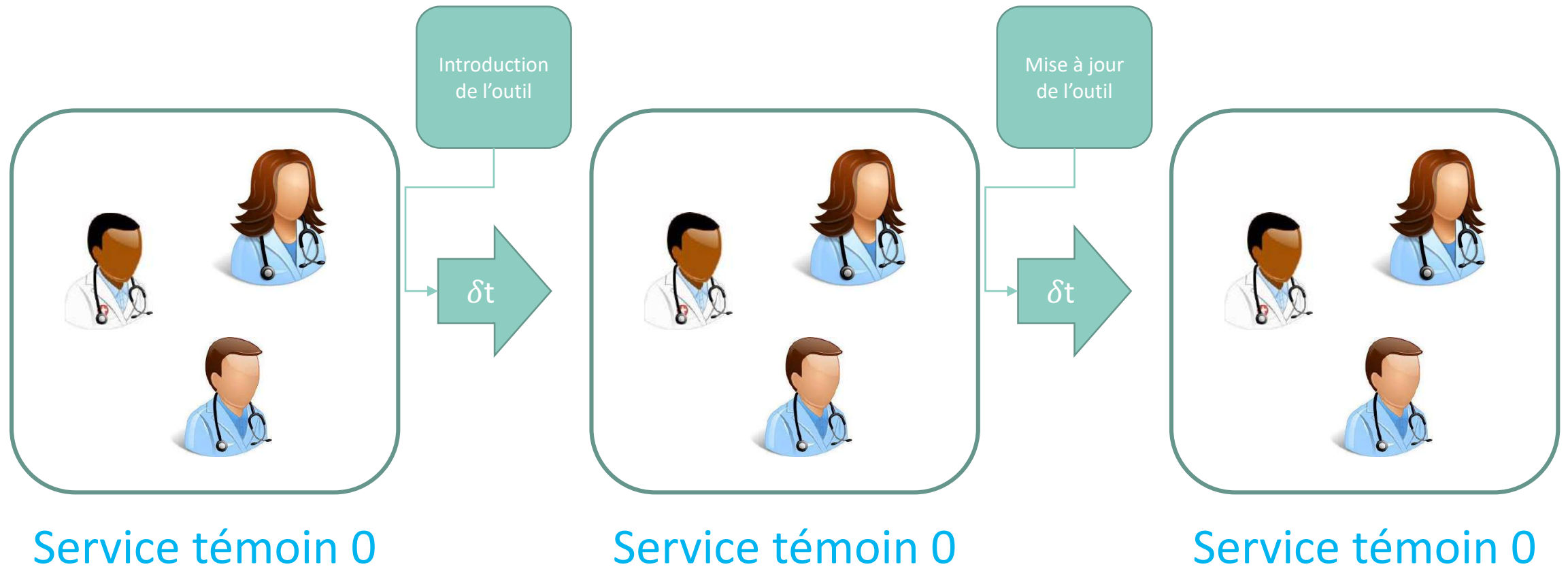
3. <https://www.holbertonturingoath.org/>

4. European Parliament (2021) – « Artificial Intelligence Act »

SUIVI À LONG TERMES

68

ÉTUDES LONGITUDINALES ^{1 2}

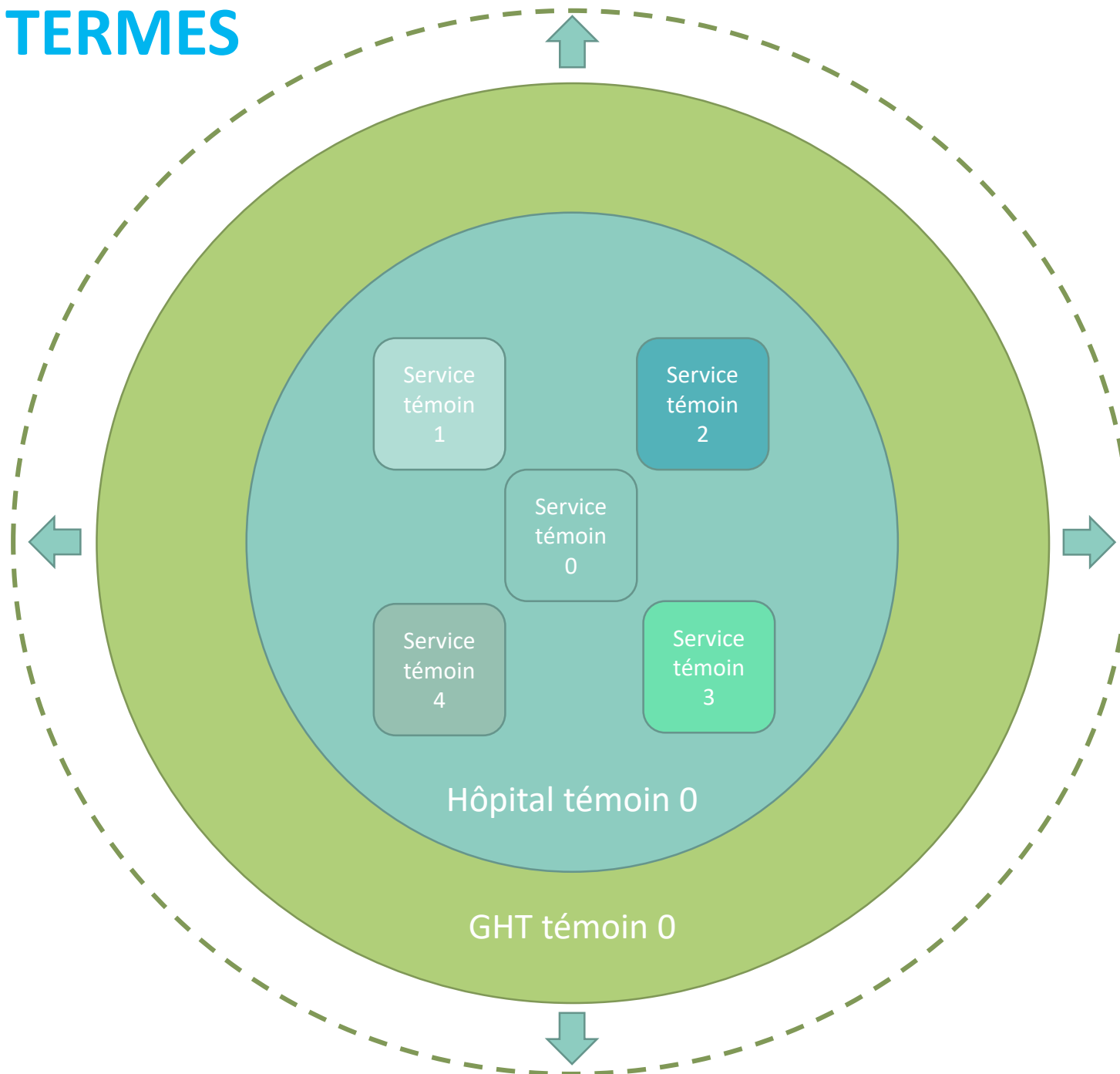


1. [Williamson G. R. \(2004\) – « The A-Z of Social Research: A Dictionary of Key Social Science Research Concepts »](#)

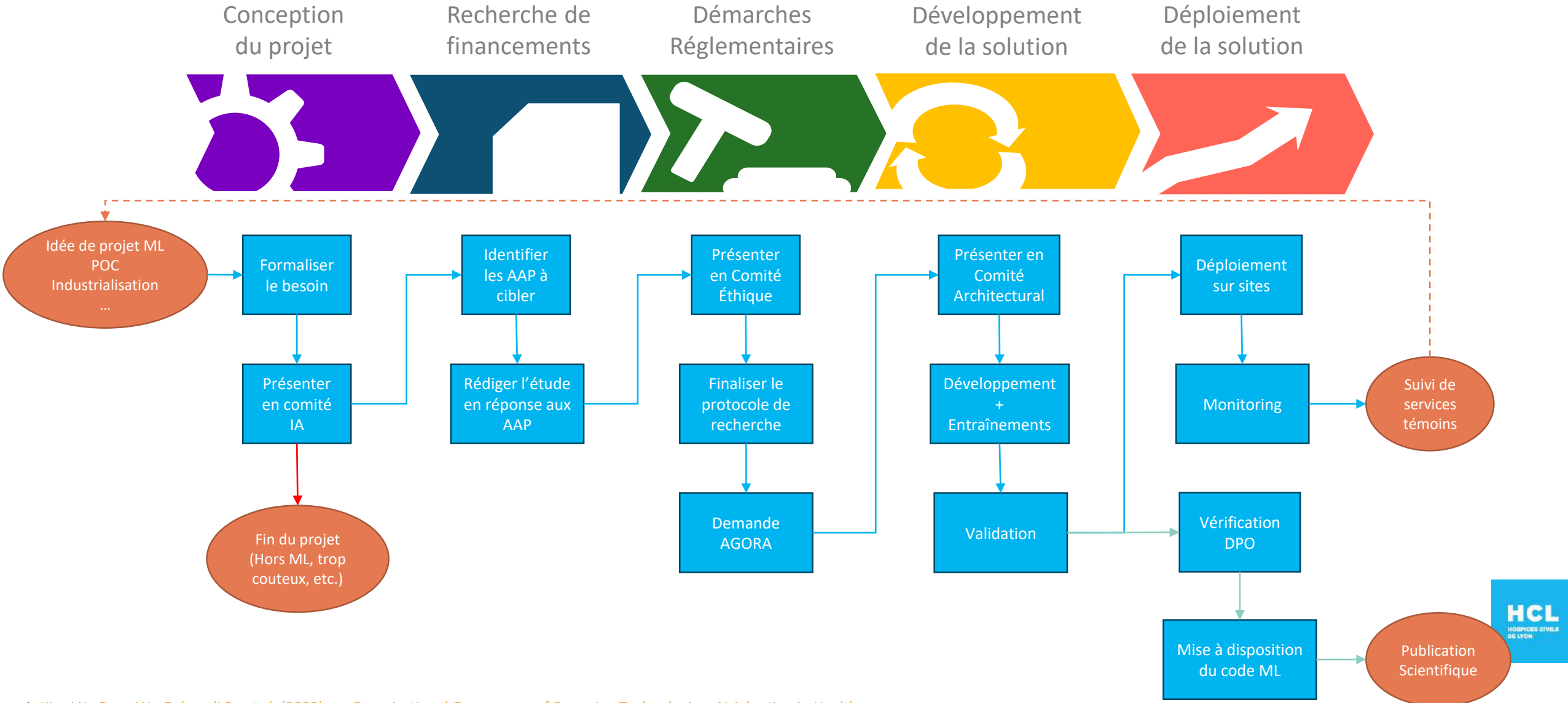
2. [Caruana E. J., Roman M., Hernández-Sánchez J., and Solli P. \(2015\) – « Longitudinal Studies »](#)

SUIVI À LONG TERMES

MONTÉE D'ÉCHELLE



INTÉGRATION DU ML DANS UN PROCESSUS ORGANISATIONNEL ¹



1. Kim J.Y., Boag W., Gulamali F., et al. (2023) – « Organizational Governance of Emerging Technologies: AI Adoption in Healthcare »

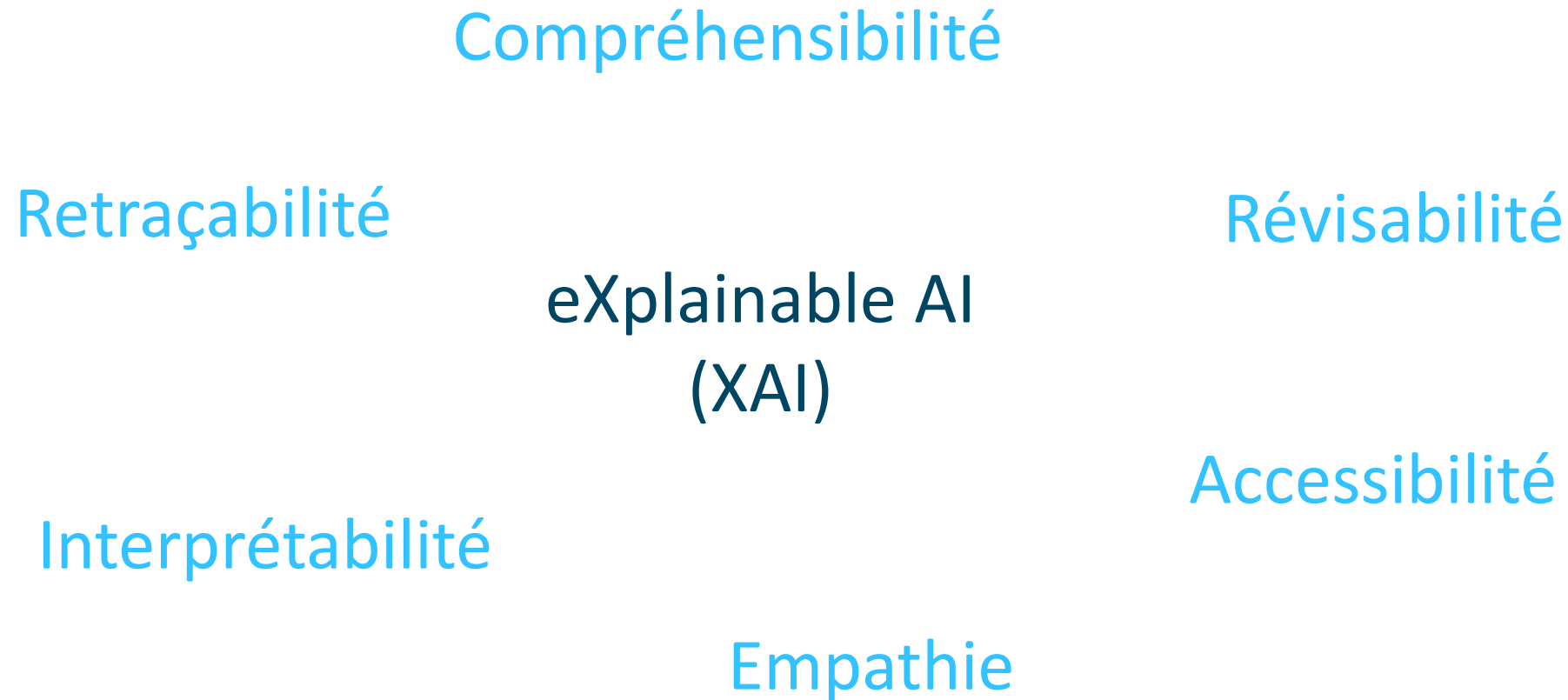
TRANSPARENCE ET EXPLICABILITÉ

PRINCIPES ET ALGORITHMES

TRANSPARENCE ET EXPLICABILITÉ

72

PLUSIEURS CONCEPTS CONNEXES ^{1 2 3 4}



1. [Ali S., Abuhmed T., El-Sappagh S., et al. \(2023\) – « Explainable Artificial Intelligence \(XAI\): What we know and what is left to attain Trustworthy Artificial Intelligence »](#)
2. [Berredo-Arrieta *et al.* \(2020\) - Explainable Artificial Intelligence \(XAI\): Concepts, taxonomies, opportunities and challenges toward responsible AI](#)
3. [Mueller *et al.* \(2019\) - Explanation in Humain-AI Systems: A Literature Meta-Review, Synopsis of Key Ideas and Publications, and Bibliography for Explainable AI](#)
4. [Richard *et al.* \(2020\) – Transparency of Classification Systems for Clinical Decision Support](#)

ÊTRE TRANSPARENT ENVERS QUI ?

73

IDENTIFIER LE PUBLIC CIBLE ^{1 2 3}



Questions clés:

- Pourquoi ce résultat et pas un autre ?
- Quel degrés de confiance mettre dans l'outil et ses résultats ?
- Dans quel situation l'outil est bon et quand est-il mauvais ?
- Comment corriger une erreur de l'outil ?

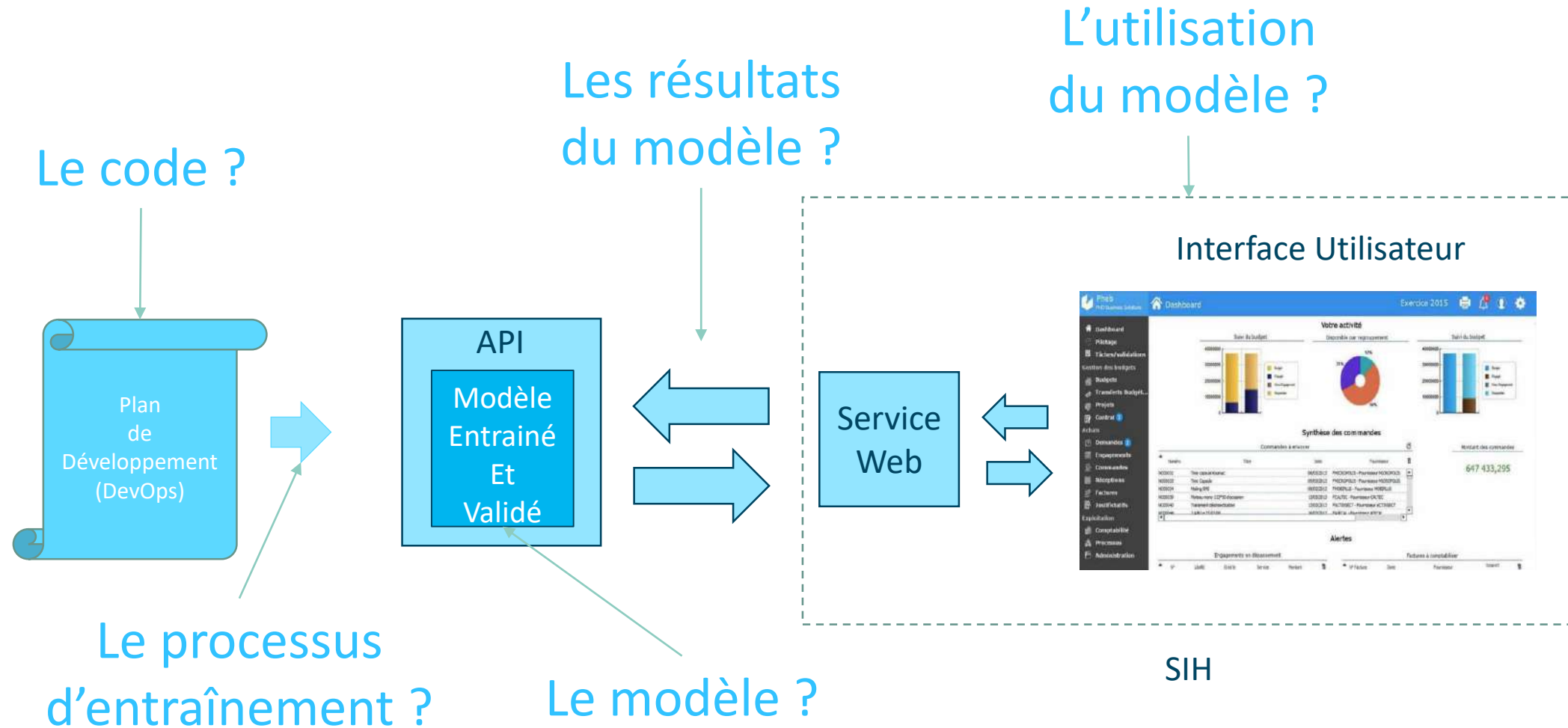


1. [Ali S., Abuhmed T., El-Sappagh S., et al. \(2023\) – « Explainable Artificial Intelligence \(XAI\): What we know and what is left to attain Trustworthy Artificial Intelligence »](#)
2. [Berredo-Arrieta et al. \(2020\) - Explainable Artificial Intelligence \(XAI\): Concepts, taxonomies, opportunities and challenges toward responsible AI](#)
3. [Mueller et al. \(2019\) - Explanation in Human-AI Systems: A Literature Meta-Review, Synopsis of Key Ideas and Publications, and Bibliography for Explainable AI](#)
4. [Molnar C. \(2023\) – « Interpretable Machine Learning: A Guide for Making Black Box Models Explainable »](#)

QU'EST-CE QUI DOIT ÊTRE TRANSPARENT ?

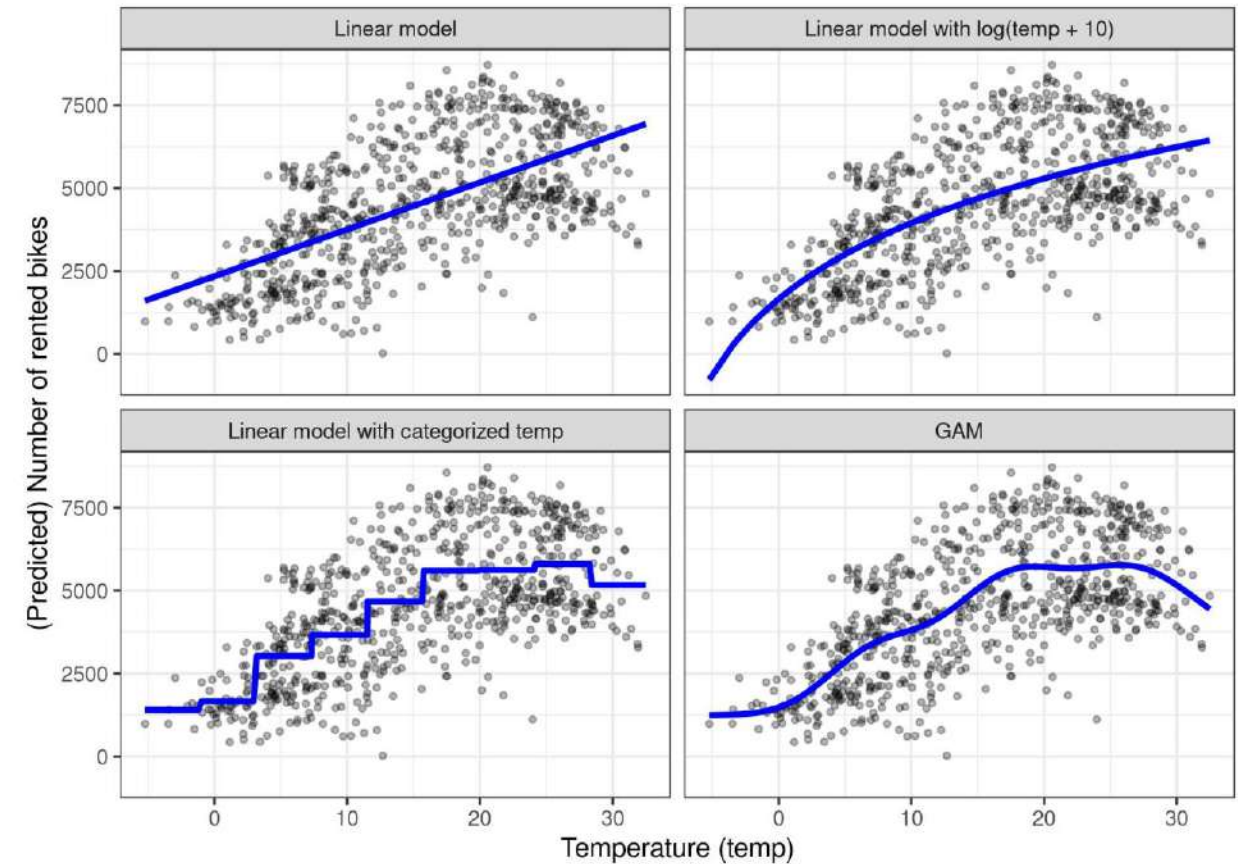
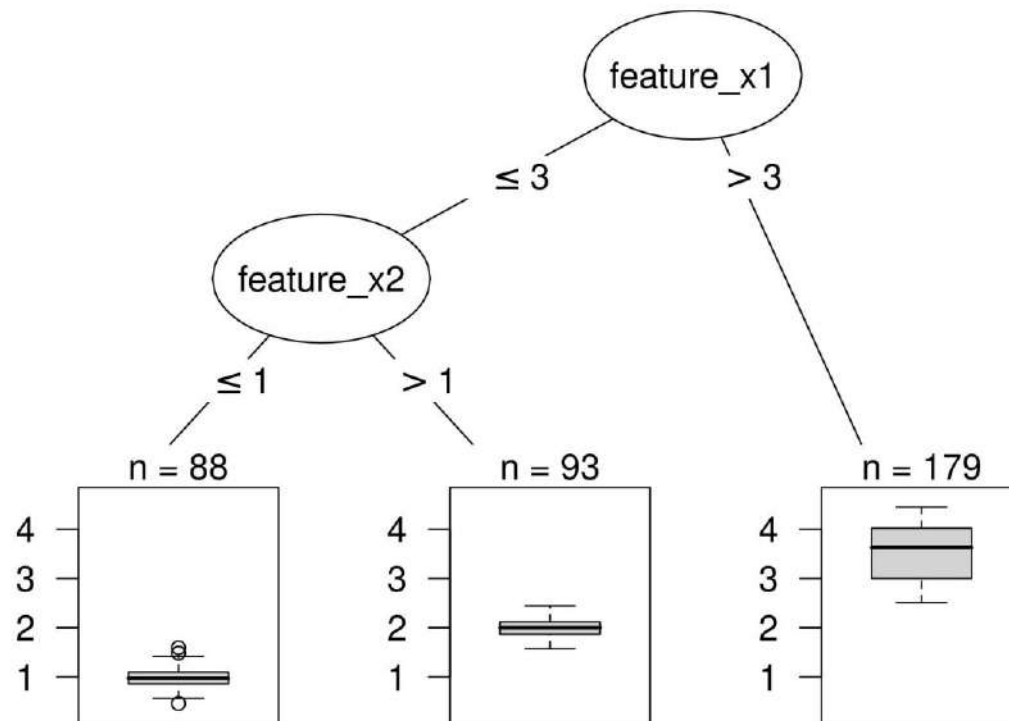
74

IDENTIFIER L'ÉLÉMENT À EXPLIQUER ^{1 2 3}



1. [Berredo-Arrieta et al. \(2020\)](#) - Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI
2. [Mueller et al. \(2019\)](#) - Explanation in Human-AI Systems: A Literature Meta-Review, Synopsis of Key Ideas and Publications, and Bibliography for Explainable AI
3. [Molnar C. \(2023\)](#) – « Interpretable Machine Learning: A Guide for Making Black Box Models Explainable »

MODÈLES INTERPRÉTABLES ¹

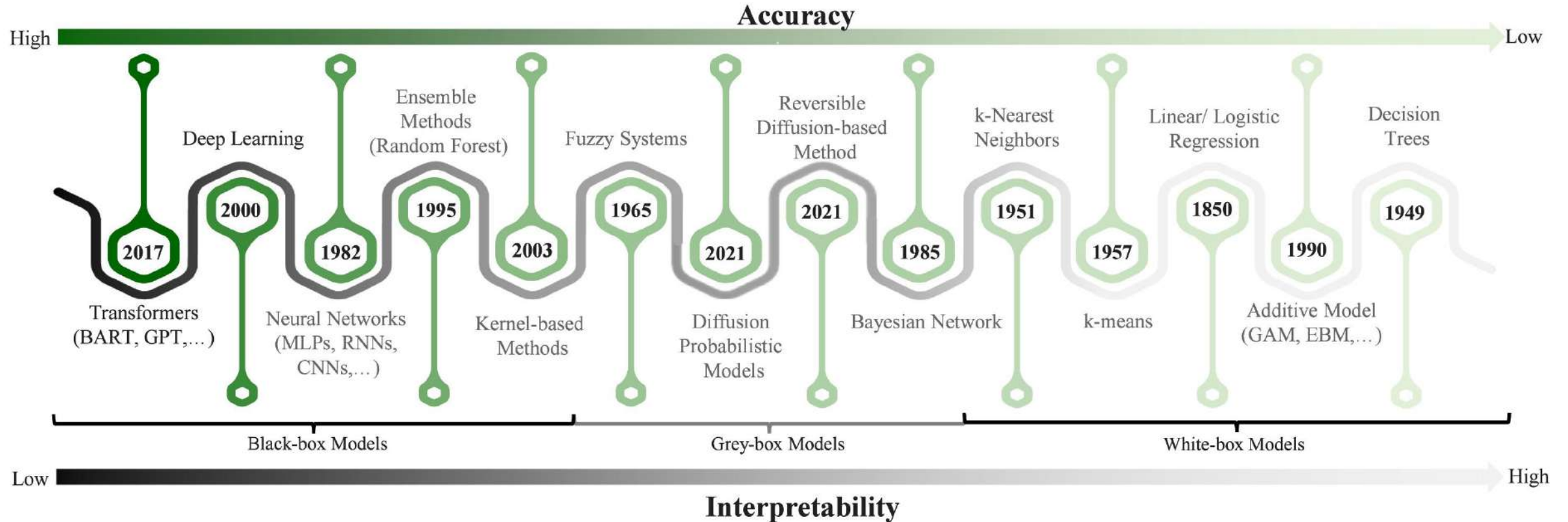


1. [Molnar C. \(2023\) – « Interpretable Machine Learning: A Guide for Making Black Box Models Explainable »](#)

INTERPRÉTABILITÉ DES MODÈLES

76

INTERPRÉTABILITÉ VS PERFORMANCES ^{1 2}



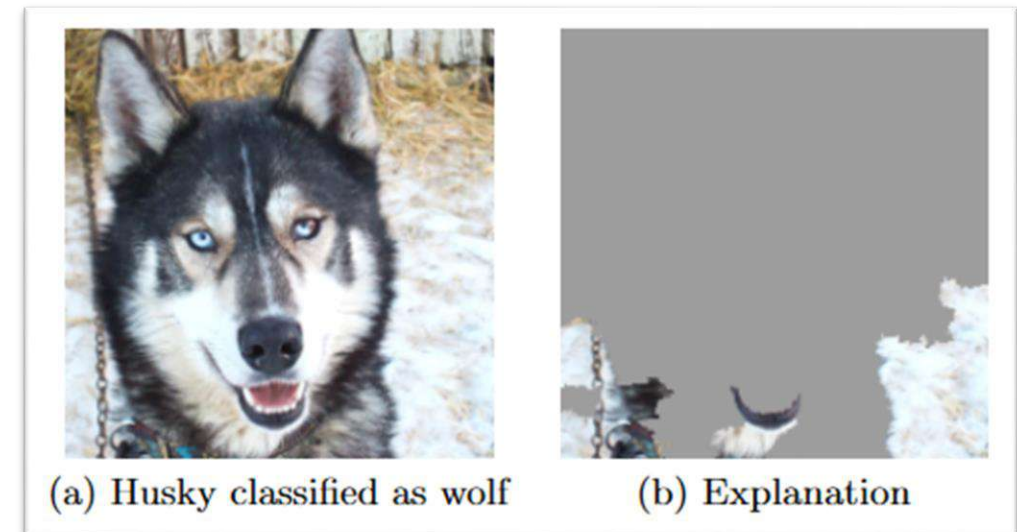
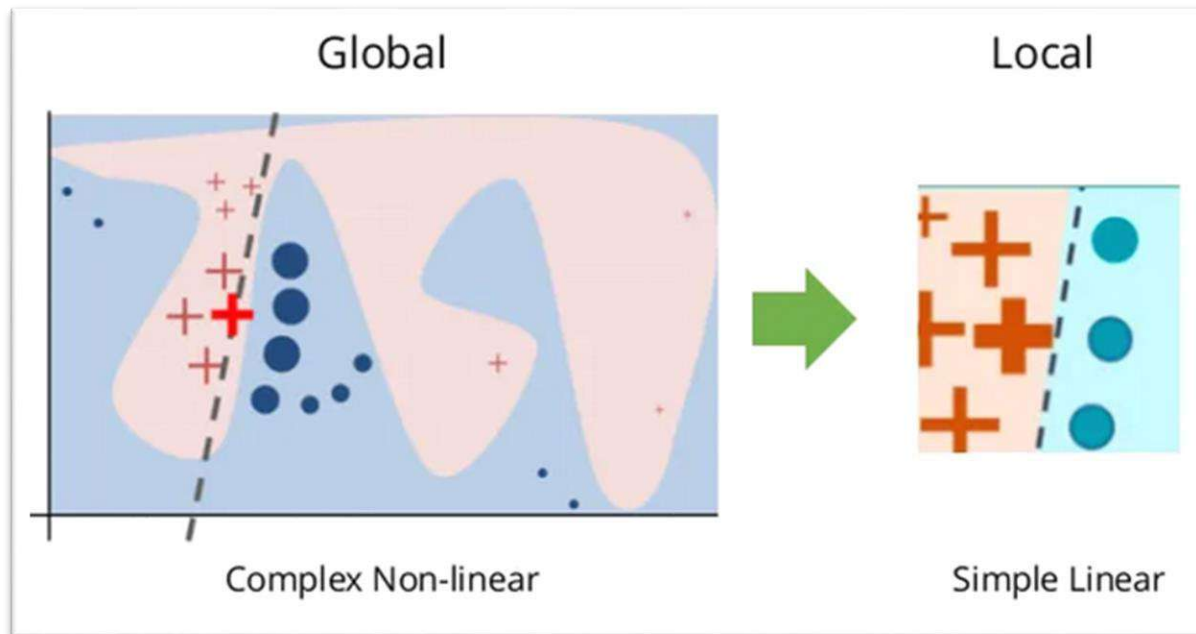
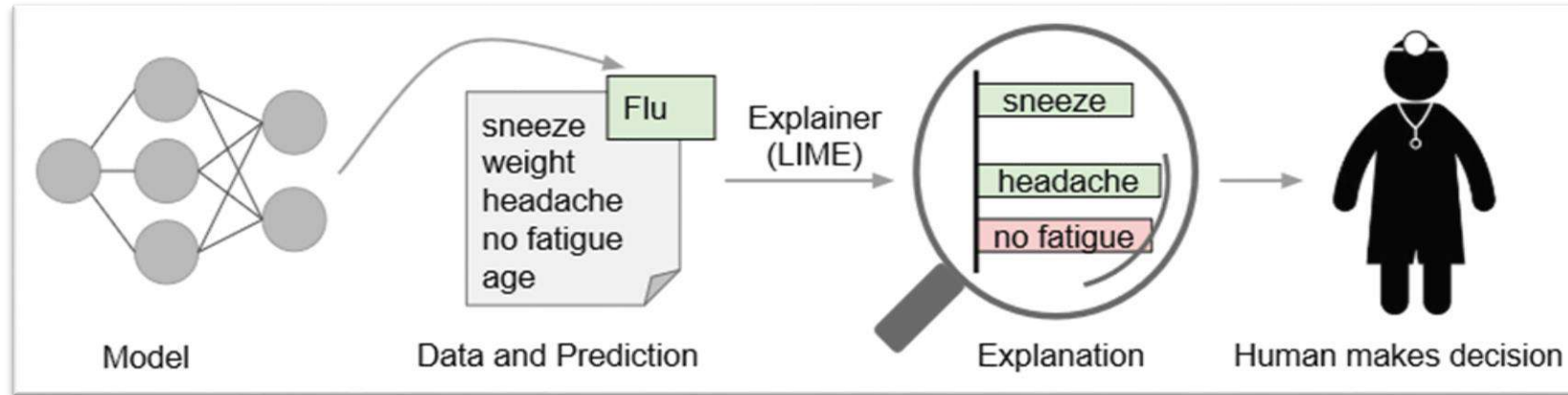
1. Ali S., Abuhmed T., El-Sappagh S., et al. (2023) – « Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence »

2. Richard A., Mayag B., Talbot F., et al. (2020) – « transparency of classification systems for clinical decision support »

ALGORITHMES « MODEL-AGNOSTIC »

77

LOCAL INTERPRETABLE MODEL-AGNOSTIC EXPLANATIONS (LIME) ^{1 2}



1. [Ribeiro M. T., Singh S., and Guestrin C. \(2016\) – « Why Should I Trust You? : Explaining the Predictions of Any Classifier »](#)
2. [Molnar C. \(2023\) – « Interpretable Machine Learning: A Guide for Making Black Box Models Explainable »](#)

SHAPLEY ADDITIVE EXPLANATIONS (SHAP)^{1 2}

Coalitions $\xrightarrow{h_x(z')}$ Feature values

Instance x

Age	Weight	Color
1	1	1

Instance with "absent" features

Age	Weight	Color
1	0	0

$x = \begin{matrix} \text{Age} & \text{Weight} & \text{Color} \\ 0.5 & 20 & \text{Blue} \end{matrix}$

$z = \begin{matrix} \text{Age} & \text{Weight} & \text{Color} \\ 0.5 & \cancel{20} & \cancel{\text{Blue}} \end{matrix}$

\downarrow 17 \downarrow Pink

Coalitions of super pixels $\xrightarrow{h_x(z')}$ Image

Instance x

sp1	sp2	sp3
1	1	1

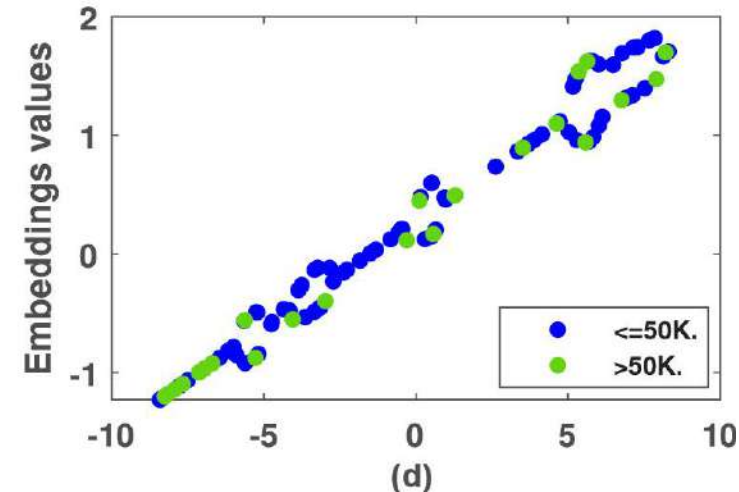
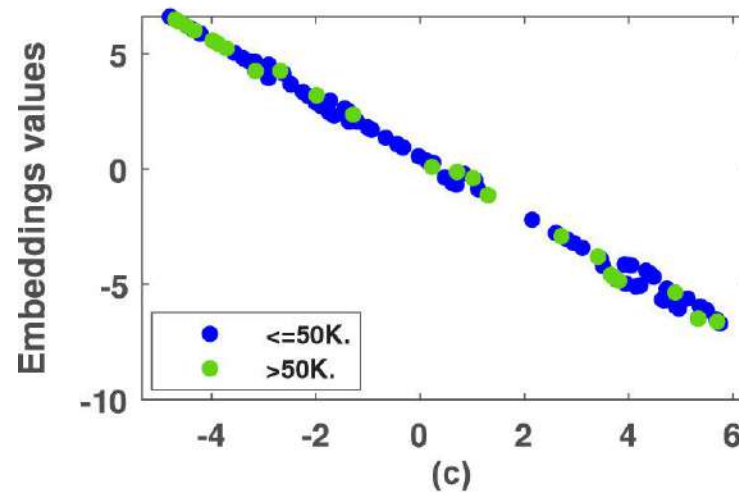
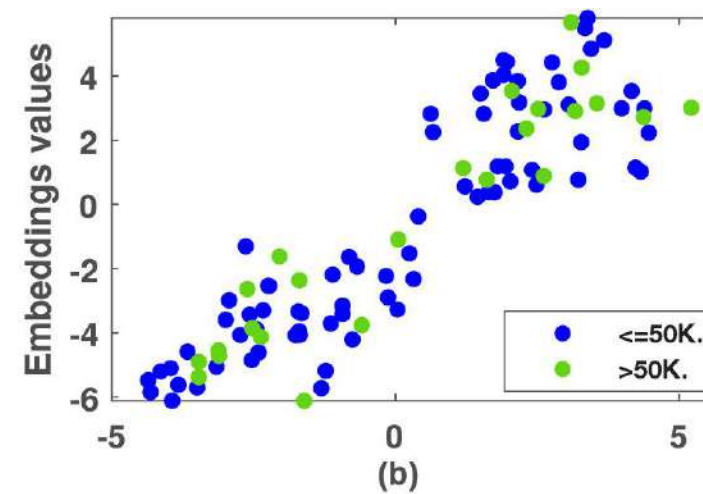
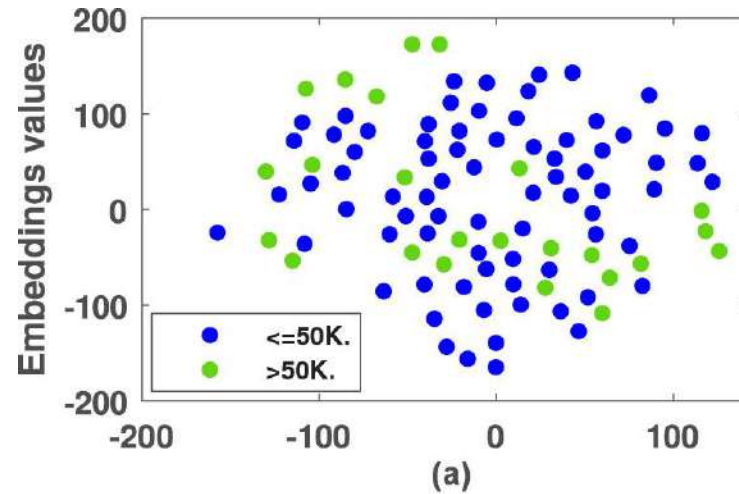
Instance x with absent features

sp1	sp2	sp3
1	1	0



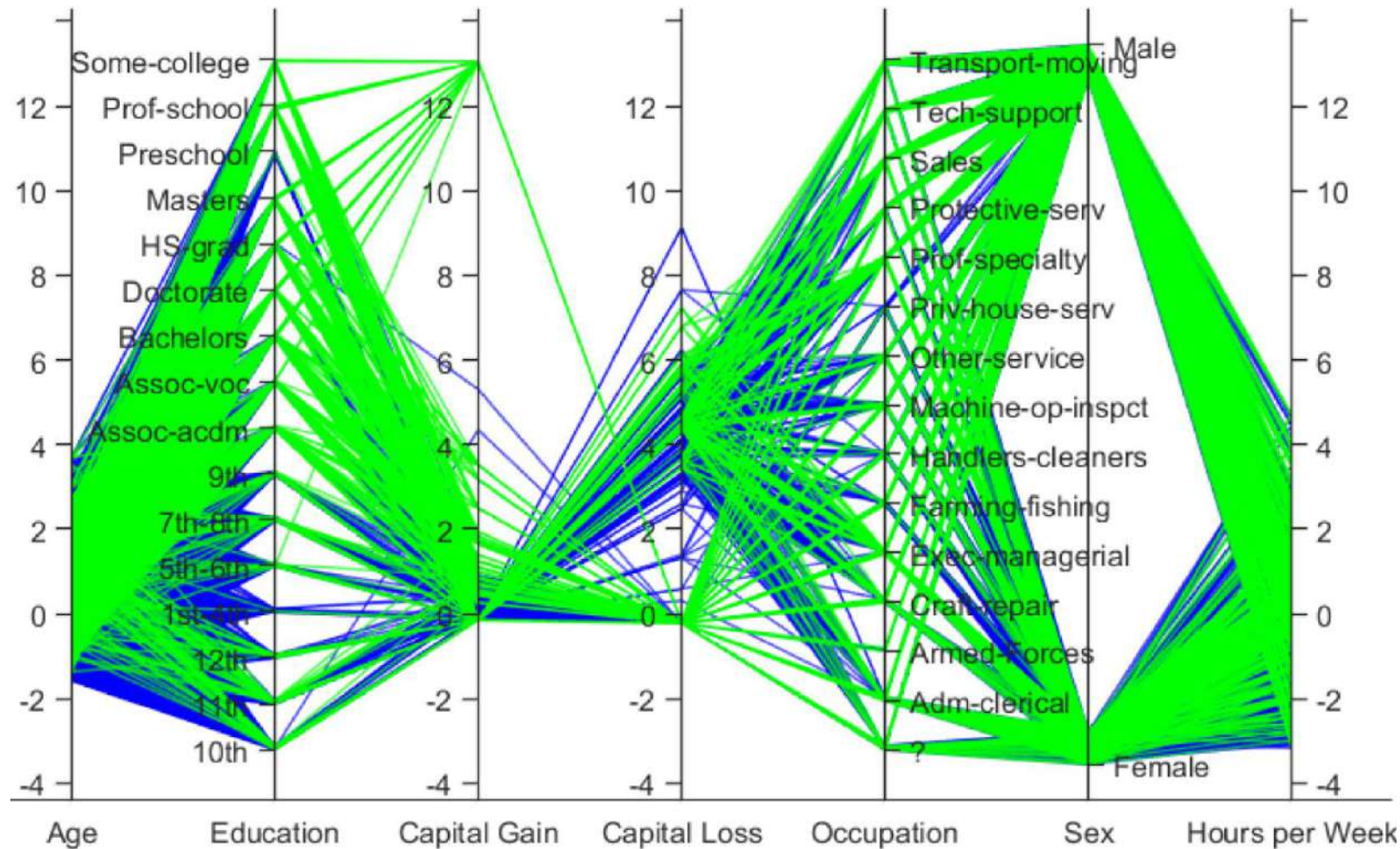
1. Lundberg S. and Lee S.I. (2017) – « A Unified Approach to Interpreting Model Predictions »
2. Molnar C. (2023) – « Interpretable Machine Learning: A Guide for Making Black Box Models Explainable »

T-DISTRIBUTED STOCHASTIC NEIGHBOR EMBEDDING (T-SNE) ^{1 2}



1. [Van der Maaten L. and Hinton G. \(2008\) – « Visualizing data using t-SNE »](#)
2. [Ali S., Abuhmed T., El-Sappagh S., et al. \(2023\) – « Explainable Artificial Intelligence \(XAI\): What we know and what is left to attain Trustworthy Artificial Intelligence »](#)

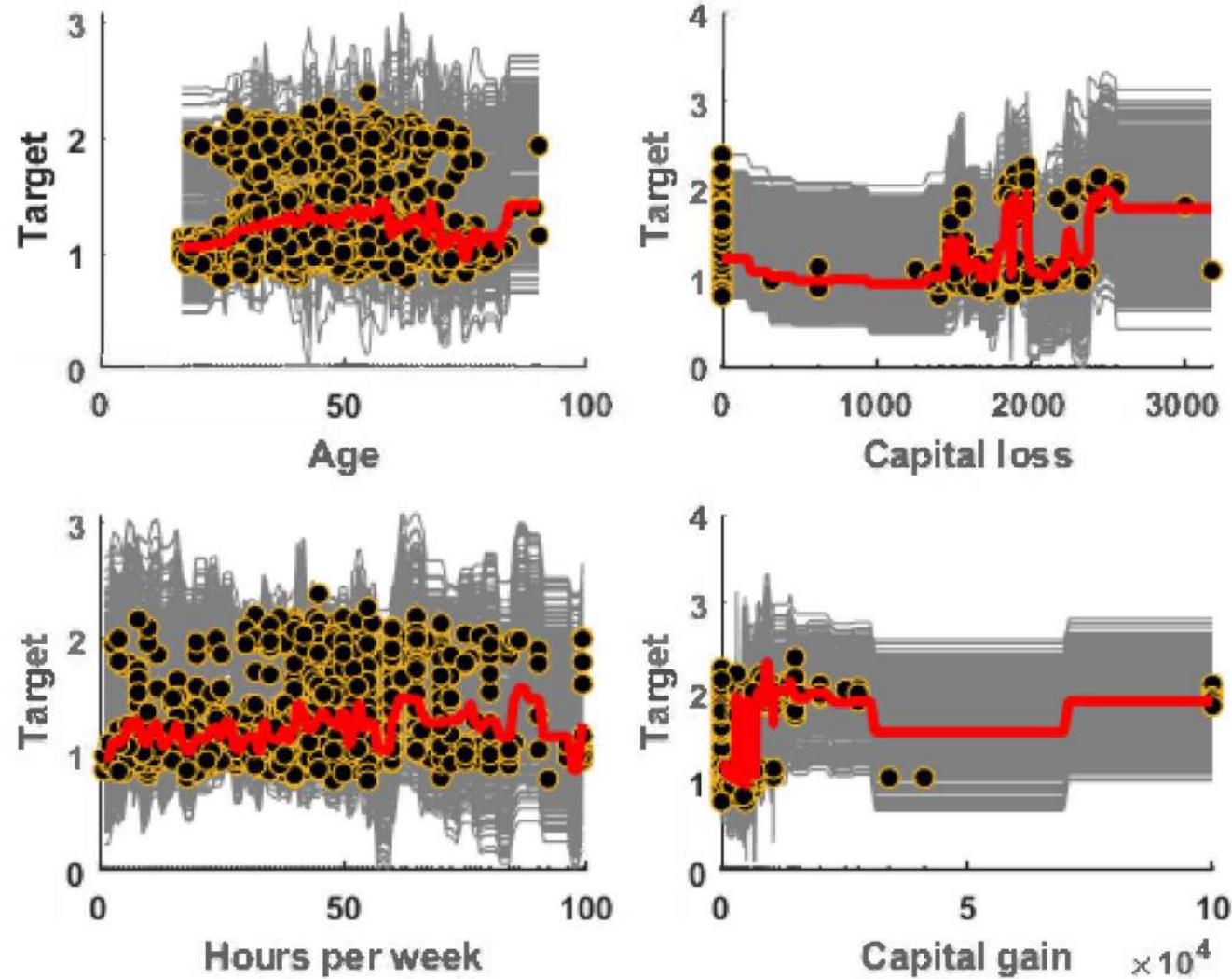
PARALLEL COORDINATE PLOTS (PCP) ^{1 2}



1. [Tilouche S., Nia V. P. and Basetto S. \(2021\) – « Parallel coordinate order for high-dimensional data »](#)

2. [Ali S., Abuhmed T., El-Sappagh S., et al. \(2023\) – « Explainable Artificial Intelligence \(XAI\): What we know and what is left to attain Trustworthy Artificial Intelligence »](#)

INDIVIDUAL CONDITIONAL EXPECTATIONS (ICE) ^{1 2}



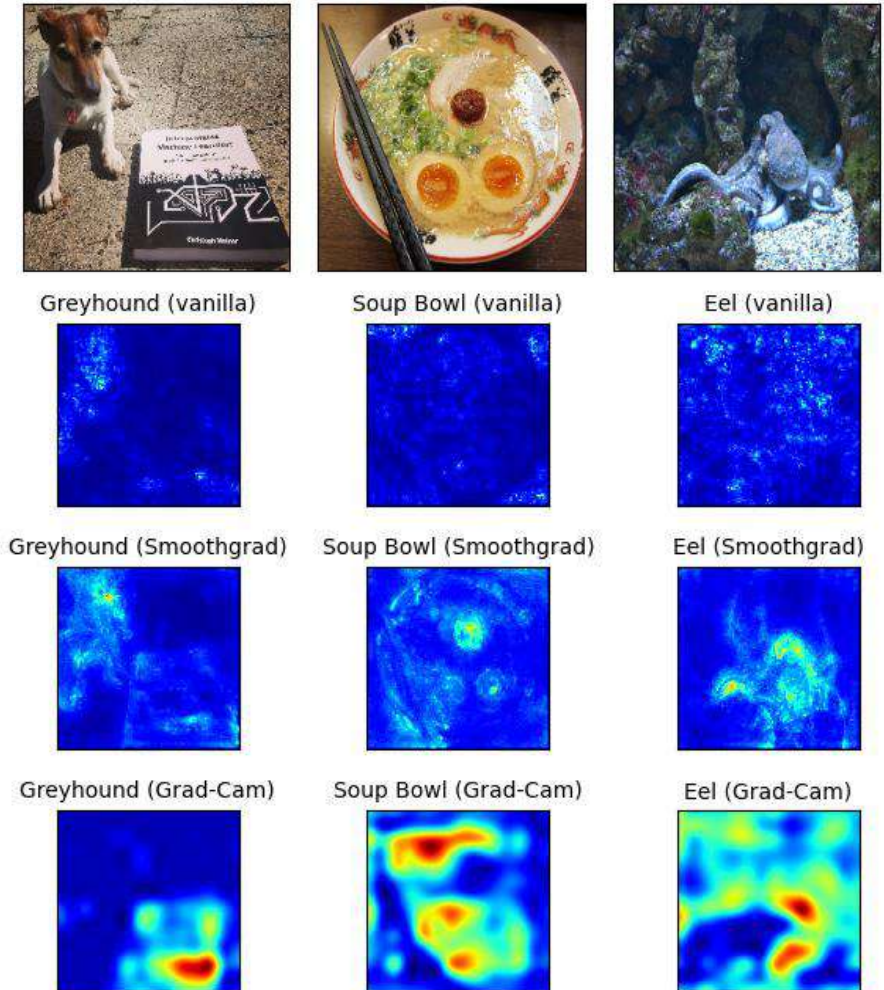
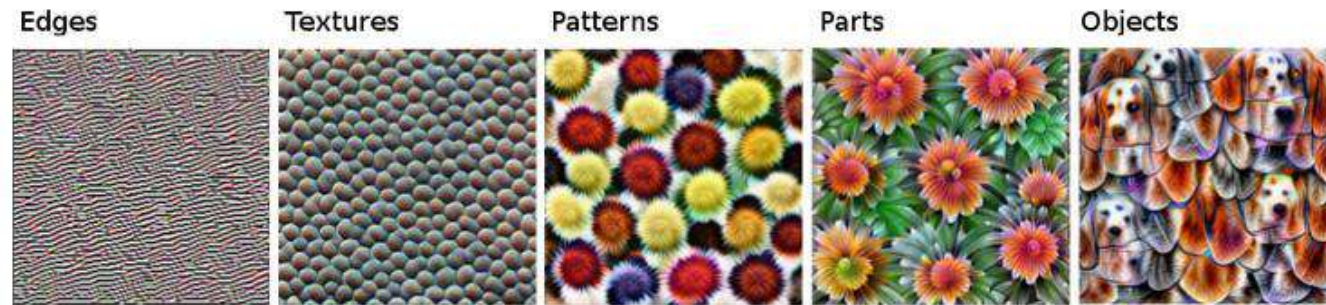
1. [Hyvärinen A. and Oja E. \(2000\) – « Independent component analysis: algorithms and applications »](#)

2. [Ali S., Abuhmed T., El-Sappagh S., et al. \(2023\) – « Explainable Artificial Intelligence \(XAI\): What we know and what is left to attain Trustworthy Artificial Intelligence »](#)

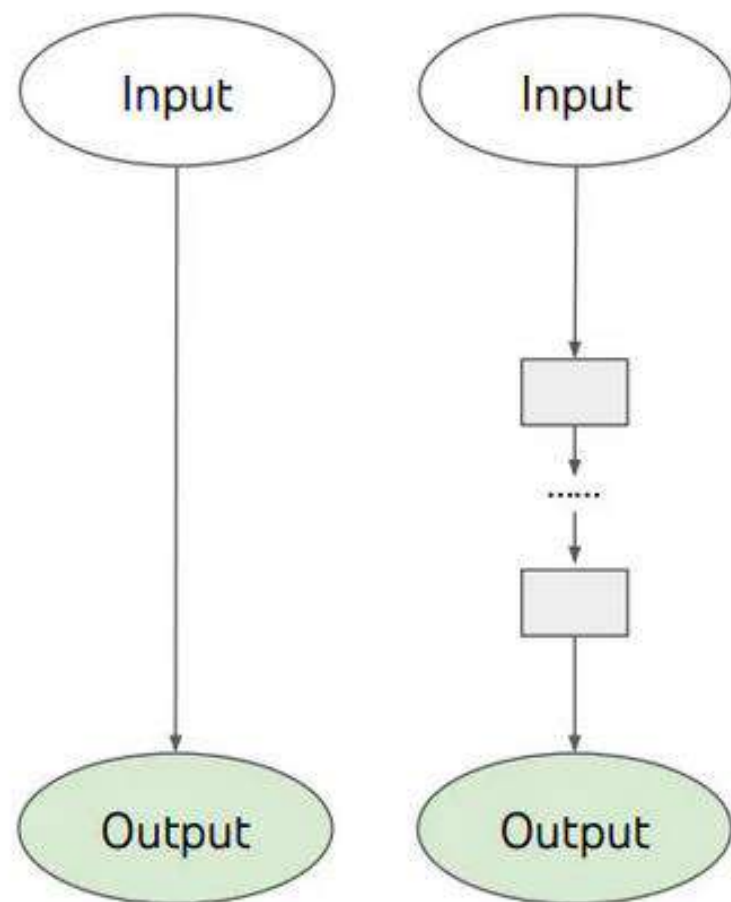
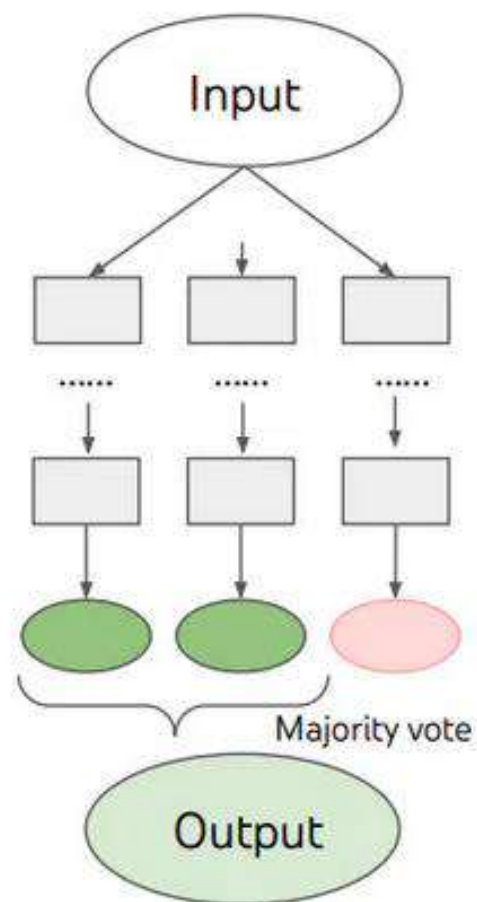
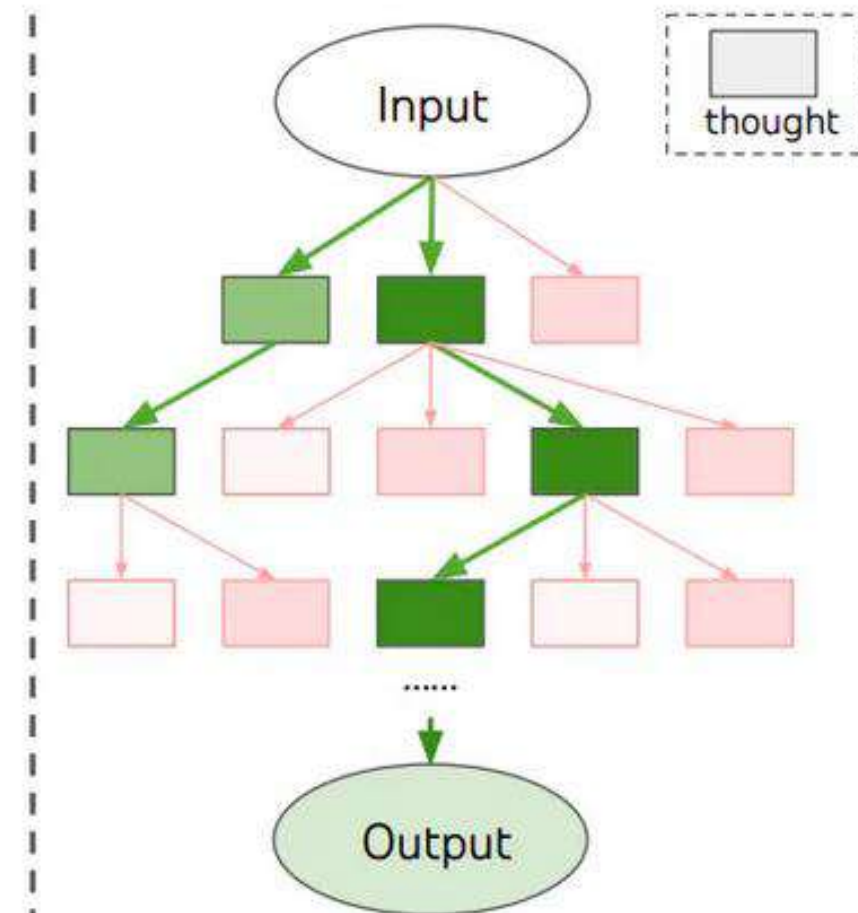
CLASSIFICATION D'IMAGE VIA RÉSEAUX DE NEURONES

82

LEARNED FEATURES AND PIXEL ATTRIBUTION ¹

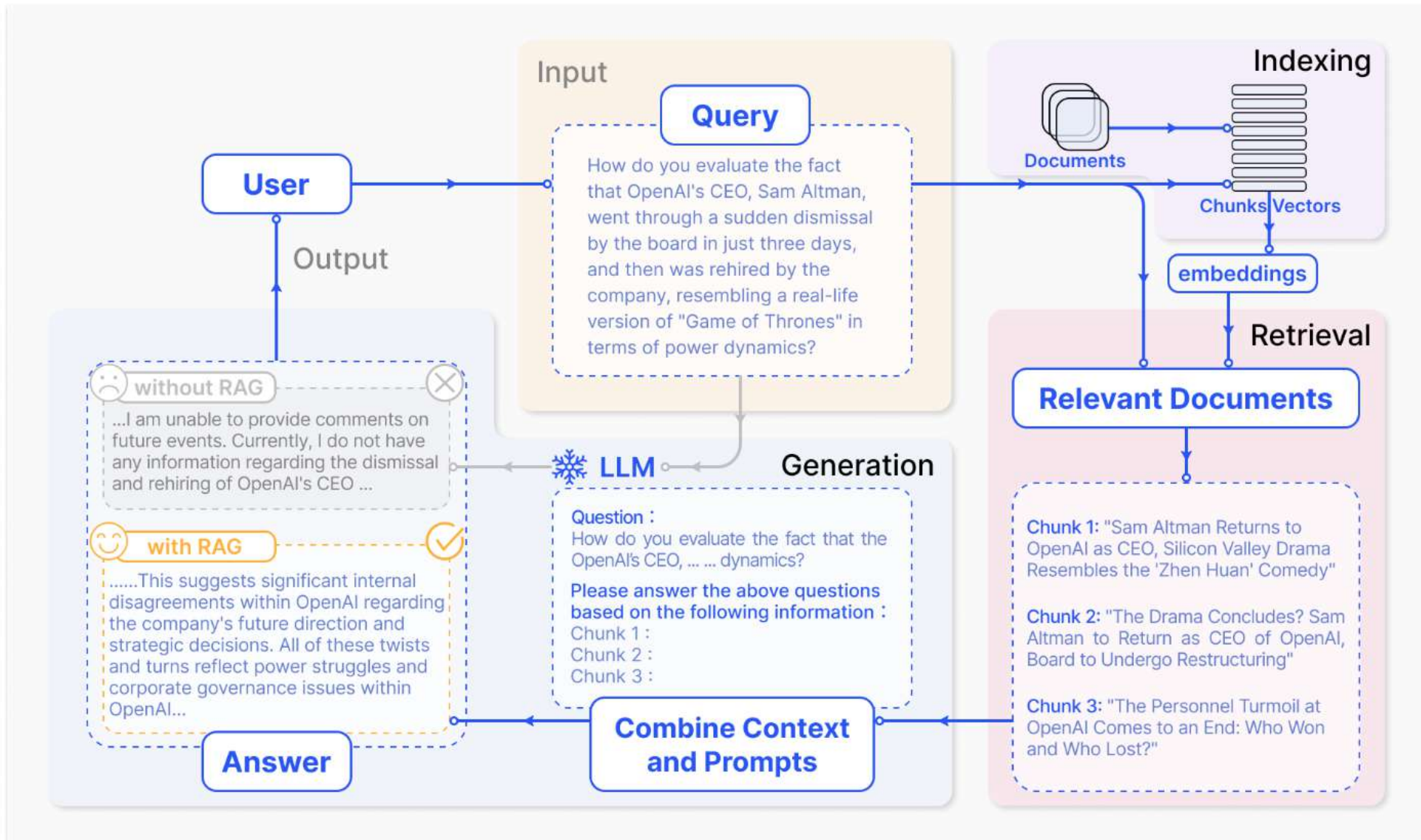


1. [Molnar C. \(2023\) – « Interpretable Machine Learning: A Guide for Making Black Box Models Explainable »](#)

PROMPT ENGINEERING ¹(a) Input-Output Prompting (IO)(c) Chain of Thought Prompting (CoT)(c) Self Consistency with CoT (CoT-SC)(d) Tree of Thoughts (ToT)

1. <https://www.promptingguide.ai/>

RETRIEVAL AUGMENTED GENERATION (RAG) ¹



1. Gao Y., Xiong Y., Gao X., et al. (2023) – « Retrieval-Augmented Generation for Large Language Models: A Survey »

UN CONCEPT VIABLE ?

Empathie « Cognitive » \neq Empathie « Affective »



Résultats de ChatGPT
au test LEAS ³

	French men's mean \pm SD	French women's mean \pm SD	ChatGPT score evaluation 1 (One-sample Z-tests)	ChatGPT score evaluation 2 (One-sample Z-tests)	Improvement between the ChatGPT evaluations
Total	56.21 \pm 9.70	58.94 \pm 9.16	ChatGPT score = 85 Men: Z = 2.96, p = 0.003 Women: Z = 2.84, p = 0.004	ChatGPT score = 98 Men: Z = 4.30, p < 0.001 Women: Z = 4.26, p < 0.001	Δ score = +13 Δ Men: Z = +1.34 Δ Women: Z = +1.42
MC	49.24 \pm 10.57	53.94 \pm 9.80	ChatGPT score = 72 Men: Z = 2.15, p = 0.031 Women: Z = 1.84, p = 0.065	ChatGPT score = 79 Men: Z = 2.81, p = 0.004 Women: Z = 2.55, p = 0.010	Δ score = +7 Δ Men: Z = +0.66 Δ Women: Z = +0.71
OC	46.03 \pm 10.20	48.73 \pm 10.40	ChatGPT score = 68 Men: Z = 2.15, p = 0.031 Women: Z = 1.85, p = 0.063	ChatGPT score = 78 Men: Z = 3.13, p = 0.001 Women: Z = 2.81, p = 0.004	Δ score = +10 Δ Men: Z = +0.98 Δ Women: Z = +0.96

MC, main character; OC, other character; Δ , the difference between the second and first evaluations. All statistically significant p -values remained significant after false discovery rate correction in the first, second and between examinations (q < 0.05, p < 0.041).



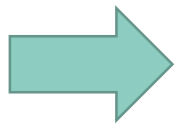
Déléguer les réponses aux patients et le support émotionnel aux LLM ? ⁴

1. [Sorin V., Brin D., Barash Y., et al. \(2023\) – « Large Language Models \(LLMs\) and Empathy – A Systematic Review »](#)
2. [Cuff B.M.P, Brown S. J., Taylor L., and Howat D. J. \(2014\) – « Empathy: A Review of the Concept »](#)
3. [Elyoseph Z., Hadar-Shoval D., Asraf K., and Lvovsky M. \(2023\) – « ChatGPT outperforms humans in emotional awareness evaluations »](#)
4. [Ayers J., Poliak A., Dredze M., et al. \(2023\) – « Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum »](#)
5. [Richard A. \(2022\) – « Can AI be conscious ? »](#)

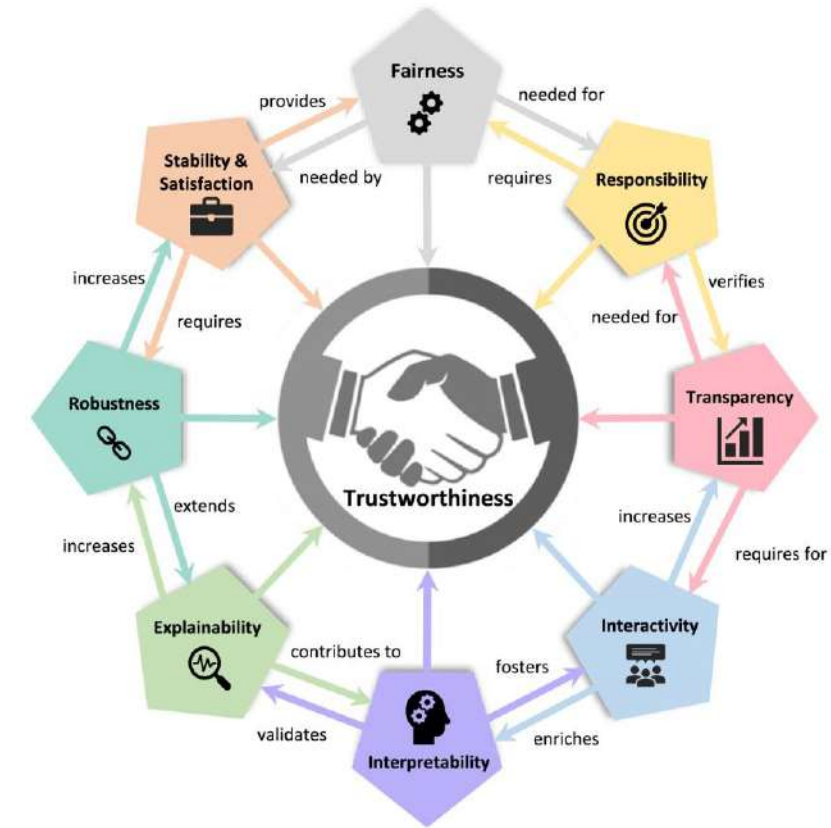
EXPLIQUER C'EST TRADUIRE



- La transparence tente d'approximer le fonctionnement d'un modèle pour l'expliquer
- Trop de transparence, tue la transparence
- La transparence est nécessaire, mais non suffisante, à instaurer une confiance envers un outil ¹



La clé pour l'adoption de l'IA par les médecins
c'est le travail d'équipe Humain-Machine ²



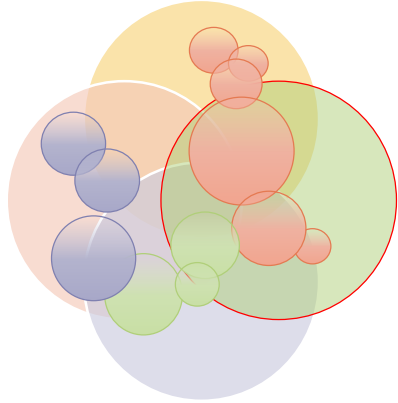
1. [Ali S., Abuhmed T., El-Sappagh S., et al. \(2023\) – « Explainable Artificial Intelligence \(XAI\): What we know and what is left to attain Trustworthy Artificial Intelligence »](#)
2. [Henry, K, Kornfield R., Sridharan A., et al. \(2022\) – « Human-machine teaming is key to AI adoption: clinicians' experiences with a deployed machine learning system »](#)

CONCLUSION

CONCLUSION

88

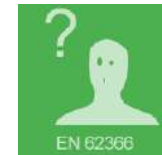
SYNTHÈSE ET PERSPECTIVES



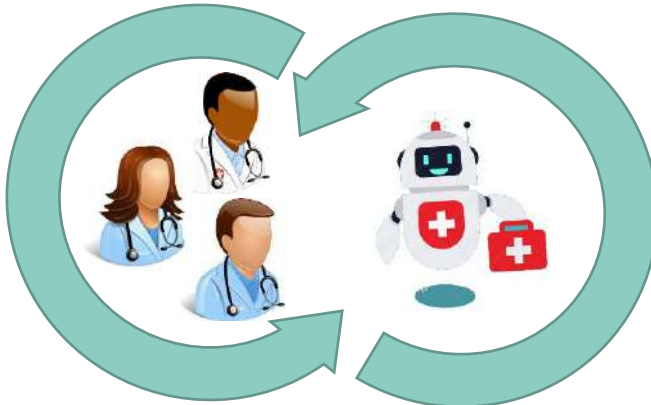
L'IA est un domaine de recherche très vaste



Les besoins des soignants et/ou des patients doivent être au centre du processus développement



L'IA en santé nécessite d'être encadré et réglementé



Une synergie doit être recherchée entre les professionnels de la santé et les SIH basé sur de l'IA pour apporter les meilleurs soins possibles

MERCI

www.chu-lyon.fr



HCL
HOSPICES CIVILS
DE LYON